IPIC 2021 Conference papers and posters Contributions

Draft Proceedings

Version 11th of June, 2021

Dear IPIC 2021 participant and contributor. This is the document that includes all papers and posters accepted out of IPIC 2021 call for contributions.

All contents and contributions will be available in the Programme IPIC Conference Webpage and for participants will be available in the Physical Internet Knowledge Platform (Link to IPIC 2021 contributions) together with other Conferences Contributions (see Resources in the Physical Internet Knowledge Platform)

IPIC 2021

IPIC2021 Sponsors and Partners

SPONSORS

PARTNERS & PROJECT COLLABORATIONS

CONFERENCE PAPERS

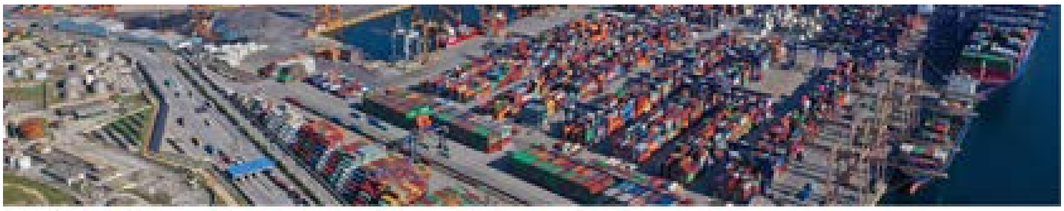
- **Submission 1** | *Framework Artifact for the road-based Physical Internet based on Internet Protocols*
- **Submission 2** | *Towards an open and neutral data sharing infrastructure – sharing state information*
- **Submission 3** | *Dynamic Containerized Consolidation in Physical Internet Enabled Parcel Logistics*
- **Submission 4** | *Assessing the potentialities of Physical Internet for Developing Countries Last Mile deliveries*
- **Submission 5** | *Introducing Services and Protocols for Inter-Hub Transportation in the Physical Internet*
- **Submission 6** | *Generating clusters for urban logistics in hyperconnected networks*
- **Submission 7** | *On Modelling and Solving Collaborative Tactical Transportation Planning*
- **Submission 8** | *Improving Demand Prediction and Reducing Out-of-Stock – Application of Advanced Data Analytics in Retail Supply Chains*
- **Submission 11** | *Port Digitalization through an Activities Scenario Model as a First Step for a Digital Twin of Port*
- **Submission 12** | *Towards PI Implementation: Interoperability and robustness of PI layers functionality illustrated through multi-context implementation*
- **Submission 13** | *Optimising Container Transports in Collaborative Roundtrip Scenarios*
- **Submission 15** | *Technical State-of-the-Art of Supply Chain Visibility*
- **Submission 16** |*Impact of High Capacity Vehicles on the future developments in the Logistics sector*
- **Submission 18** |*Smart IoT solutions for ports paving the way towards PI nodes*
- **Submission 19** |*GPICS a framework to create a digital twin for the Physical Internet*
- **Submission 22** |*Enhancing Logistics Demand Prediction Accuracy Through Client – Service Provider Hyperconnected Data Ensembles*
- **Submission 23** |*Potential Access Hub Network Design Optimization in Hyperconnected Urban Logistics*
- **Submission 24** |*Reduction of Food Wastage due to expiry at the Supermarket Level*
- **Submission 27** |*Optimising Product Swaps in Urban Retail Networks*

- **Submission 29** |*5G Enabled Video Analytics for Detecting Container Seals in Port Operations*
- **Submission 30** |*5G applications in ports, corridors and logistics hubs*
- **Submission 31** |*Framework and Research Roadmap for a Next-Generation Hyperconnected Logistics Hub*
- **Submission 33** |*Physical Internet inspired Atomic Modeling for Supply Chain Risk Management*
- **Submission 35** |*Towards a self-organizing logistics system: identifying the value of SOLs for different stakeholders and transition-phases*
- **Submission 38** |*Modularization of Delivery and Transportation*
- **Submission 39** |*Conceptual Framework for Hyperconnected Transport and Delivery Logistics Infrastructure*
- **Submission 40** |*Digital Twin Design Requirements for Durable Goods Distribution Physical Internet*
- **Submission 41** |*Design of a Simulation-based Experiment for Assessing the Relevance of the Physical Internet Concept for Humanitarian Supply Chains*
- **Submission 42** |*Resilient Hyperconnected Logistics Hub Network Design*
- **Submission 43** |*Modular and Mobile Design of Hyperconnected Parcel Logistics Hub*
- **Submission 48** |*Predictive demand modeling for new services in Hyperconnected Urban Parcel Logistics*

POSTER

- **Submission 9** | *Smart Track 4 Waterway – Regulatory Analysis*
- **Submission 10** | *Spatio-temporal prediction over hyperconnected logistics networks*
- **Submission 14** |*Dynamic Consolidation and Packing of Parcels in Modular Physical Internet Handling Containers*
- **Submission 34** | *Autonomous and interconnected technologies for the container supply chain: The MOSES Concept*
- **Submission 44** | *The Benefits and price of hyperconnectivity in last mile delivery*
- **Submission 46** | *Building on Synergies between Freight Logistics and People Mobility in Urban Areas*

- **Submission 49** | *Synchronizing public transport co-modality networks for city logistics*
- **Submission 51** | *Production-Intralogistics Synchronization in Physical Internet-enabled Manufacturing Systems*
- **Submission 53** | *A Dynamic Protocol with Payments for the Physical Internet A Simulation with learning agents*
- **Submission 54** | *Can Physical Internet support the integrated management of last-mile (reverse) logistics? An exploratory analysis*



Framework Artifact for the Road-Based Physical Internet based on Internet Protocols

Steffen Kaup[1], André Ludwig[2], Bogdan Franczyk[1, 3]

1. Leipzig University, Information Systems Institute, Leipzig, Germany
2. Kühne Logistics University, Computer Science in Logistics, Hamburg, Germany
3. Wrocław University of Economics, Wrocław, Poland

{kaup, franczyk}@wifa.uni-leipzig.de, andre.ludwig@the-klu.org*

Keywords: *Protocol Transformation, Road-Based Physical Internet*

Abstract

The Physical Internet (PI, π) raises high expectations for efficiency gains in transport and logistics. The PI represents the network of logistics networks for physical objects in analogy to the Data Internet (DI). Road based traffic represents one of these logistics networks. Here, many empty runs and underutilized trips still take place (International Transport Forum, 2019). Hence, there is a lot of potential in the road-based Physical Internet (RBPI), which will have an impact on transport and logistics strategies, but also on vehicle design. On the DI, logistics strategies are implemented in protocols. In order to transfer such concepts to the RBPI, relevant protocols of the DI had been analyzed and transferred to the world of physical objects. However, not all functionalities can be transferred one-to-one, e.g. a data packet in the DI can simply be re-generated by a hub in case of damage or loss. To compensate for the challenges, a framework artifact has been designed with appropriate transformation customizations based on design science principles (vom Brocke, 2007). From this, resulting requirements for future vehicles were derived. This paper makes a contribution to the implementation of the RBPI in order to fit road based vehicles to the future world of transport and logistics.

1. Introduction

The Physical Internet (PI) represents the network of logistics networks for physical objects in analogy to the Data Internet (DI). 72 percent of freight traffic in Germany is done by trucks on the road (Statistisches Bundesamt, 2019). This is why the road-based Physical Internet (RBPI) plays a central role within the PI. The capacity for the movement of goods on the road is restricted by physical constraints, such as the number and available volume of vehicles, as a part of the π-movers family, and limited road connections. Here, many empty runs and underutilized trips still take place (International Transport Forum, 2019). Through a road-based Physical Internet (RBPI), methods of a very established information network regarding resilience and efficiency, the Data Internet (DI), are transferred to the world of freight road transport. As a vision of the future, freight is being navigated through an existing transport network in a way analogous to how information packets are transferred from a host computer to another computer on the DI, as visualized in Figure 1.

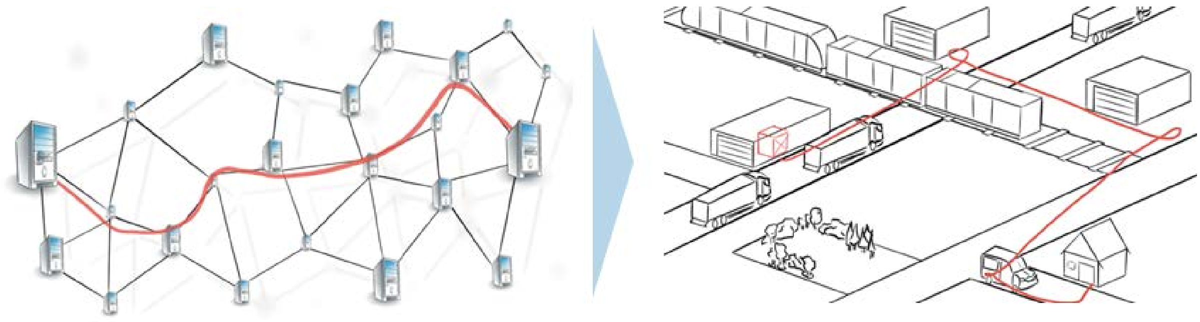


Figure 1: Transformation from DI to RBPI (own visualization)

On the DI, data transport methods are covered by Internet Protocols (Badach and Hoffmann, 2007). For the RBPI, an adequate concept is still missing. However, this concept as well, like the basic idea of the DI itself, can be adapted from the existing Internet Protocol family, as transferred in the following sections. After presenting related work on the topic in Section 2, research methodology is introduced in Section 3. Subsequently, Section 4 starts with the transformation of Internet Protocols and proceeds to develop a design artifact for the RBPI. In Section 5, conclusions are drawn and further steps on the road to implementation of the RBPI are suggested.

2. Related Work

Research related to the Physical Internet dates back to the year 2009, initiated by the idea of Benoit Montreuil to transform the Digital Internet to a Physical Internet. He was inspired by an article which bore this headline (Markillie, 2006). Together with Ballot and Meller he wrote a manifesto that describes a lot of facets of this transformation (Montreuil, Meller and Ballot, 2010). Since this initial publication, research has developed further through numerous contributions at conferences, such as the 'International Physical Internet Conference' (IPIC). Hardly any papers published there dealt with the research of protocols and if so, only on an abstract level (Zikria *et al.*, 2018). Fundamental work concerning network-based software architecture was carried out in a dissertation (Fielding, 2000). A conceptual approach to solve the problem of finding a feasible route with the lowest total cost and an appropriate time is designed with a graph based model (Dong and Franklin, 2020). A model transfer was carried out within the context of crowd logistics with regard to available cargo space capacities of current road traffic data (Kaup and Demircioglu, 2017). In recent years, two industrial standards for the Internet of Things (IoT) have developed, the 'Reference Architecture Model Industrie 4.0' (RAMI 4.0) and the 'Industrial Internet Reference Architecture' (IIRA). The models aim to capture production objects throughout the entire life cycle and to map them uniformly and consistently on the IT side. The differences between RAMI 4.0 and IIRA lie in the different emphasis in scope and depth. The layer of the RAMI model responsible for organising the IoT components is based on the Open Systems Interconnection (OSI) model (Shi-Wan *et al.*, 2017). Hence, our research focuses on the OSI model. Furthermore, a recent study recommends the detailed analysis of Internet Protocols in order to sharpen the analogy from DI to PI (van Luik *et al.*, 2020).

2.1 Open Systems Interconnection (OSI) model and existing transformations

IT network services have been structured into a seven layer service model, the Open Systems Interconnection (OSI) model. OSI is a reference model for network protocols as a layered architecture. It has been published as a standard by the International Organization for Standardization (ISO) since 1984. For each layer, functions and protocols are defined which have to fulfill certain tasks within the communication between two systems (Kaup and Neumayer, 2003). The layers 1-3 represent network-oriented functions, like transmission and switching. Layer 4 is called the transport layer and is intended to enable the different transport networks to be used for the connection of end-to-end systems. The protocols of layers 5-7 are application-specific. Montreuil et al. (Montreuil, 2012) introduced an Open Logistics Interconnection model (OLI) as an analogy to its digital equivalent OSI. The OLI model proposes the seven layers: physical (1), link (2), network (3), routing (4), shipping (5), encapsulation (6) and logistics web (7), as shown in Figure 2.

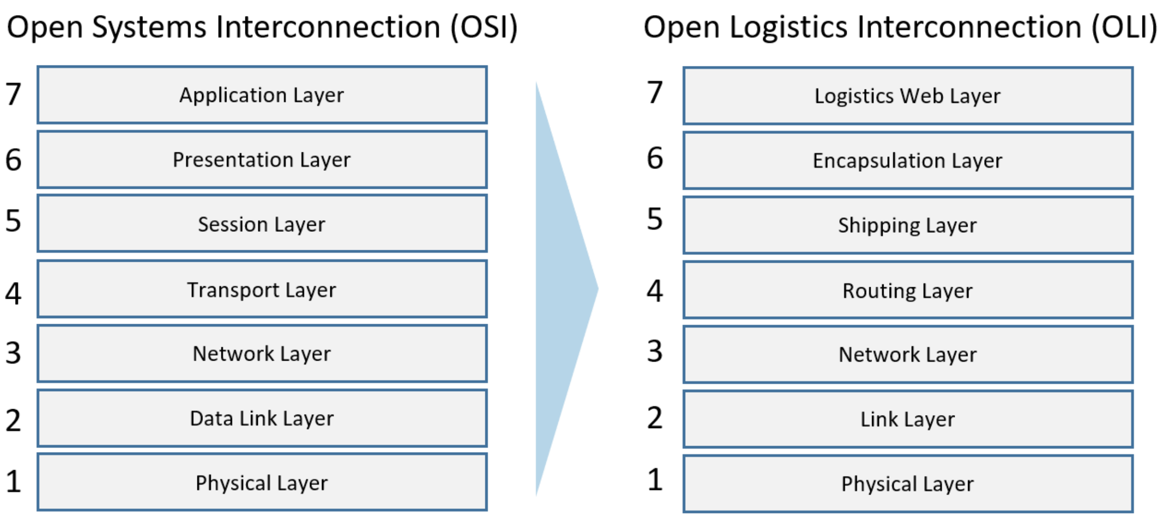


Figure 2: Open Logistics Interconnection model, adapted from (Montreuil, Meller and Ballot, 2010)

A working group of the European research group 'Alliance for Logistics Innovations through Collaboration in Europe' (ALICE) proposed to concentrate on a conceptual five layers model

(Liesa *et al.*, 2020), shown in Table 1, which goes along with the latest research (Dong and Franklin, 2020).

Table 1: Physical Internet Conceptual Five Layers Model (Liesa et al., 2020)

Protocol Layer	Description
"Application" Layer	This layer is where the actual goods are "defined" and human readable information about the goods is created. It is in this layer that this information and these goods are prepared for transmission/transport to their destination. As with the Internet, this packet of information and the associated physical goods (the shipment) is our "message."
"Transport" Layer	At the transport layer shipments are broken up into sizes that are transportable by standard sized containers or network defined standard transport mechanisms (for goods not amenable to standard containers). In addition, the transport layer provides services that ensure delivery of the shipment and manage flows between the sending location and destination. The standard loads that are shipped out from the transport layer are our "segments."
"Network" Layer	The network layer takes the "segments" constructed in the transport layer and manages all services required to deliver these "segments" to their destination. This layer defines how all nodes between source and destination should respond to handling and controlling the goods that are in the segments. The information concerning handling and control of the segments is attached to the segment and the combination of this information and the shipment segment forms our "datagram."
"Link" Layer	The link layer takes the "datagram" from the network layer and passes it from the current node to the next node in the network. The services that the link layer provides depends on the mode of transport between nodes. The encapsulated "datagram," which includes all information on how the particular transport mode is to handle the shipment, is called a "frame."
"Physical" Layer	The physical layer of the Physical Internet actually moves the "bits" of a shipment between the linked nodes. The services provided are both link and mode dependent and depend heavily on mode, carrier, regulatory bodies, etc.

Building upon this research, this paper conducts a deep dive into the five transport relevant layers, as developed by (Liesa *et al.*, 2020). From the results of the protocol analysis, functionalities and attributes are transformed to the RBPI, weaknesses of a one-to-one transfer are identified and possible fixes discussed and adopted as possible.

2.2 Knowledge gap and research questions

With the PI, methods of a very established information network regarding resilience and efficiency, the DI, are transferred to physical goods transport. Some concepts have been developed so far, but without sufficient depth in terms of protocols. These protocols implement the methods within the DI. This paper analyzes Internet Protocols in depth and dares the transformation of their functionalities and attributes to the RBPI. This transformation answers the leading Research Question (RQ): '*What are the corresponding functionalities and attributes within the road-based Physical Internet resulting from the analysis of Digital Internet protocols?*' that is solved with the following Sub-Questions (SQ):

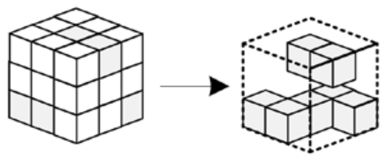
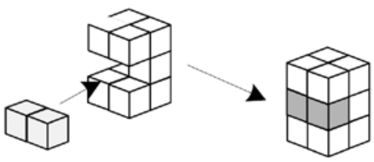
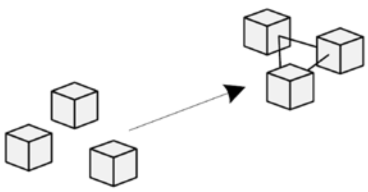
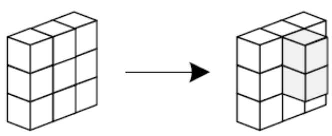
- SQ$_1$: Which functions and/or attributes can be transferred one-to-one from the DI protocols to the RBPI?
- SQ$_2$: Which of the non-transferable functions and/or attributes require replacement?
- SQ$_3$: What vehicle requirements can be derived from the protocol transformation?

The following Section describes relevant principles to solve these design questions.

3. Research Methodology

The purpose of this paper is to transfer the operating principles of the DI to the world of road-based freight transport. Hence, this is a design problem, even though design science research (DSR) has a methodological answer: design principles for reference modelling. Basically, five design principles are distinguished: analogy, specialization, aggregation, instantiation and configuration (vom Brocke, 2007). Each principle represents a special technique of reusing methods or content from the original model in order to build a target model.

Table 2: Principles for reuse in DSR reference modelling, adapted from (vom Brocke, 2007)

Principle	Visualization	Description
Configuration		The derived model is a selection of relevant properties and methods of the original model.
Instantiation		Instantiation offers the opportunity to construct models for which both the methods as well as the properties are reusable.
Aggregation		Aggregation offers the potential of combining model statements of original models in new contexts.
Specialization		Specialisation allows the taking over of general construction methods and/or properties and extending them to specific demands.
Analogy		The relation between the original model and the resulting model is based on a perceived similarity of both models regarding a certain aspect.

The similarity of the two models, original and target model, depends on the selection and arrangement of components, methods and properties within the systems (Fielding, 2000). The principles move further and further away from the original concept from top to bottom, i.e. starting with 'configuration' and ending with 'analogy'. While 'configuration' merely involves the configuration of existing components, and many methods and properties are reused on a one-to-one basis in 'instantiation', the principle of 'analogy' only leads to a roughly corresponding solution. The principle of transformation has to be chosen in such a way that the transferred method from the DI works in the targeted world, the RBPI. In the following Section, properties and methods of the OSI layers are analyzed in respect of reuse regarding the RBPI.

4. From DI to RBPI: Protocol Transformation

This Section builds on the research of the sense project (Liesa *et al.*, 2020), which identified five relevant layers for transformation. In contrast to the OLI model (Montreuil, Meller and Ballot, 2010), layers 5-8 are combined into one layer. This layer then also contains the protocols responsible for routing mechanisms, as Figure 3 shows.

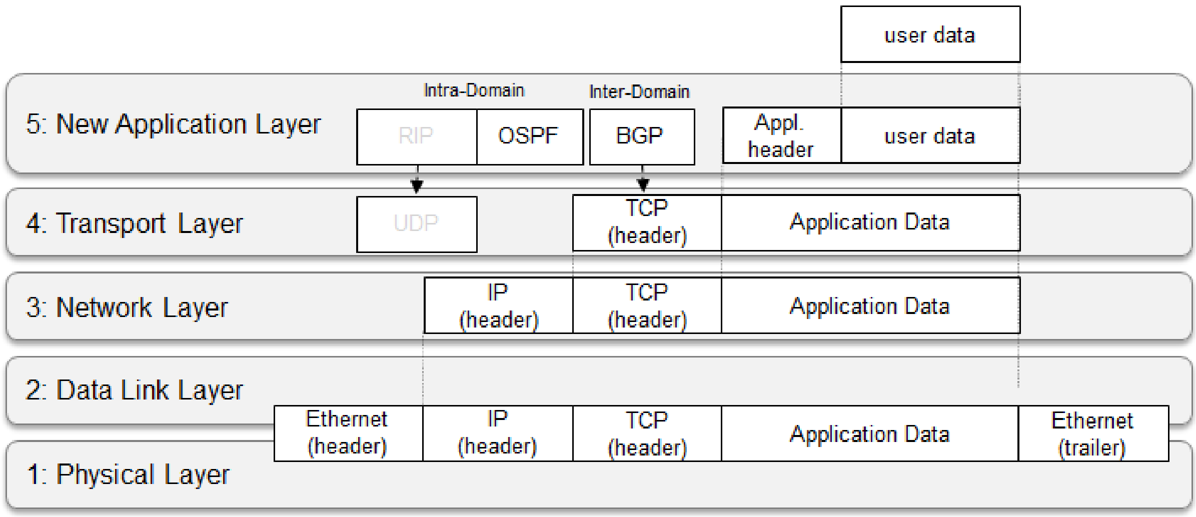


Figure 3: Internet Protocols based on the five layer model (own visualization)

In the following Subsections, the protocols and their functions and attributes are analyzed with respect to their reusability and transformation for the RBPI.

Transformation of Layer 5: The aggregated Application Layer (OSI Layers 5-7)

We start with the new aggregated application layer (Liesa *et al.*, 2020), where the transport journey begins. *'It is in this layer that this information and these goods are prepared for transmission/transport to their destination. (Liesa et al., 2020)'*. This information contains sender and destination addresses, treatment requirements, latest delivery date and the cost budget. This information aggregation in combination with the associated physical goods, or so-called shipment, is the counterpart of a message within the RBPI. On this layer, the protocols RIP (Routing Information Protocol), OSPF (Open Shortest Path Information First) and BGP (Border Gateway Protocol) implement different routing strategies, as visualized in Figure 6. The RIP protocol is a distance vector protocol. This means that decisions are made only on the basis of the number of hubs (or so-called π-nodes) on possible connection routes. The route with the lowest number of interconnected hubs wins, regardless of whether there is traffic with free capacity here at the desired time. It is based on the Used Datagram Protocol (UDP) and

works connectionless, i.e. for each sub-component this routing process takes place again instead of maintaining a connection for a number of transmissions. There are approaches to empower UDP to become connection-oriented in order to be nearly equivalent to a TCP connection, e.g. via QUIC (Kumar, 2020). The advantages lie in reduced latency for multiplex connections, which are, however, negligible for the RBPI.
The OSPF and BGP protocols work differently. They provide the network nodes with routing tables containing current network information. With OSPF, there are so-called autonomous systems that synchronize with each other. BGP, on the other hand, provides an overarching exchange, which makes it particularly suitable for cross-system exchange of routing information via routing tables as shown in Figure 4.

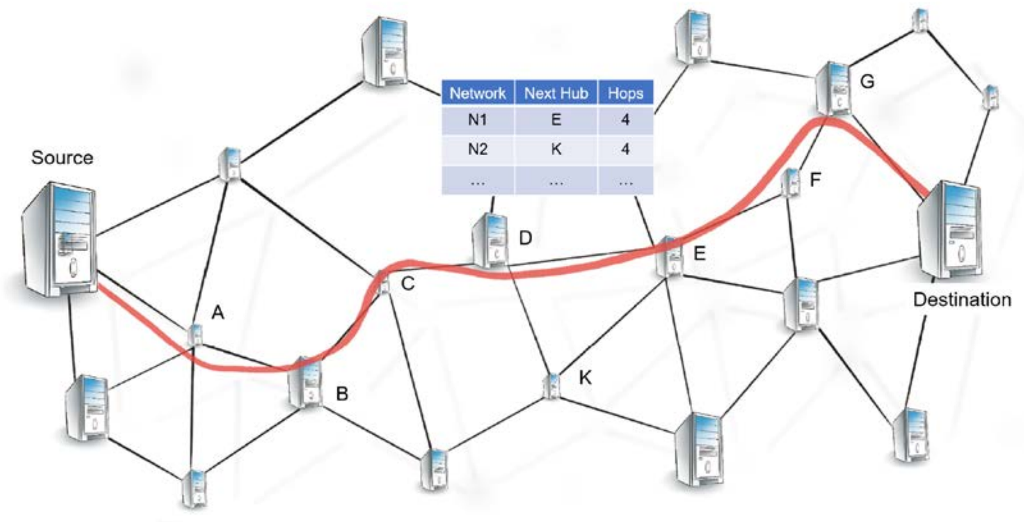


Figure 4: Operation of the BGP-Protocol, based on (Kaup, Ludwig and Franczyk, 2020)

The BGP-Protocol is based on the Transmission Control Protocol (TCP) on Layer 4 and the Internet Protocol (IP) on Layer 3. For this reason, only the TCP/IP protocol family was considered by the SENSE group (Liesa *et al.*, 2020), which is again confirmed by the evaluation of the routing protocols on Layer 5.

Transformation of Layer 4: From Transport Layer to Routing Layer

The Transport Layer ensures that all data is transmitted completely and arrives correctly at the receiver. '*At the transport layer shipments are broken up into sizes that are transportable by standard sized containers or network defined standard transport mechanisms' (Liesa et al., 2020)*. In addition, the transport layer provides services that ensure the faultless delivery of the shipment and manage the flows between the sending location and the corresponding destination. For this purpose, the segments are numbered sequentially to prevent double transmission of freight packets and to reconstruct their correct sequence at the destination side, as described and transferred in Table 3 with the principle of analogy.

Table 3: Transport Layer Transformation Table

Task within DI	Principle	Task within RBPI
Segmentation and reassembly		
A message is divided into transmittable segments. Each segment contains a sequence number. These sequence numbers enable the transport layer to reassemble the message correctly upon arriving at the destination and to identify and replace packets that were lost in transmission.	Analogy	An order can be divided into individual sub-components, e.g. bulky or heavy parts and small or low-weight parts. These components can take different routes through the RBPI to their destination. If freight is decomposed into sub-components, these sub-components and their order must be provided with instructions for reassembly by the responsible π-node.
Flow and error control		
This method ensures faultless operation, see also Table 3: Acknowledge and checksum attributes.	Analogy	It is not trivial to implement something like a checksum mechanism to automatically identify damaged packages. This can be done almost exclusively by visual inspection.

The tasks described in Table 3 are implemented by the TCP-Protocol. Its methods and properties are transferred in detail as follows:

Table 4: Transformation of TCP packet header to OLI TCP header

TCP attribute	# bits	Function within DI	Function within RBPI
Source/ Destination	16/16	Represents the port addresses of source and destination.	Ports could be used to identify a π-node's point of entry or exit.
Sequence number	32	The sequence number indicates the position of the data packet in the complete message.	The *sequence number* is required for the composition of the π-containers belonging to the same shipment.
Acknowledge-ment number	32	When the Acknowledge flag (ACK) is activated, the next sequence number is written in this field.	The *following sequence number* is of interest if some components are 3D-printed on the last π-node instead of being transmitted.

TCP attribute	# bits	Function within DI	Function within RBPI
Data offset	4	Indicates the size of optional information.	Indicates the size of all transport relevant information.
Reserved	3	For future use and should be set to zero.	Not relevant, can be used for additional transport information.
ECN/ NS/CWR	1/1/1	The Congestion Experienced Flag (ECN) serves as an indication of network congestion.	This flag could be used to indicate identified (or impending) congestions on specific routes.
URG	1	Indicates the significance of the URGent pointer field.	Indicates an *urgent* or *very valuable* freight item.
ACK	1	Indicates significance of the acknowledgement number.	*Acknowledgement information of delivery* is requested.
PSH	1	Trigger for forwarding buffered data to the receiving application.	negligible
RST	1	Resets the connection.	negligible
SYN/FIN	1/1	Indicates the first (SYN) or last (FIN) packet within a transmission.	Indicates the first/last π-container of a shipment with more than one π-container.
Window size	16	The number of bytes that the sender of this segment is currently ready to receive.	Size of the storage location at the π-node of the receiver.
Checksum	16	The 16-bit checksum field is used for error checking the header, the payload and a pseudo-header.	Damaged packages can be identified almost exclusively by visual inspection.
Urgent pointer	16	Points to the last urgent data byte.	No one-to-one transfer is obvious.
Options	32	Optional information	Optional information

Transformation of Layer 3: The Network Layer

The Network Layer describes procedures for exchanging files between addressable systems. The most important tasks of the network layer include the provision of cross-network addresses (IP), the routing or creation and updating of routing tables and the fragmentation of data packets (Badach, Hoffmann and Knauer, 1997). Hence, this layer *'defines how all nodes between source and destination should respond to handling and controlling freight contained in the segments' (Liesa et al., 2020)*, as detailed in the following table.

Table 5: Network Layer Transformation Table

Task within DI	Principle	Task within RBPI
Routing		
The router decides the next appropriate node based on the destination address and the use of routing tables.	Analogy	Determine the next π-node, depending on relevant information about the current traffic situation stored in the current π-node. Individual components of a delivery can take different routes to the destination.
Logical Addressing		
If a packet passes the network boundary, another addressing system is necessary to help distinguish the source and destination systems.	Speciali-zation	Different networks could be represented by different shipping companies with different address spaces. This would then be the task of the π-node to mediate.

Each involved element, such as servers and terminals in the DI, has a unique IP address. The first major version of IP, Internet Protocol Version 4 (IPv4), has now been exhausted by its address space. Hence, a successor has been in increasing deployment, Internet Protocol Version 6 (IPv6). For reasons of actuality, IPv6 is considered in this paper. The IP protocol frames the message into a datagram and ensures its consistent and correct handling. In the physical world, this datagram corresponds to a piece of freight and its packaging, a π-container. Table 6 shows the attributes of this datagram and transforms these into the RBPI. The routing table contains information about the next appropriate hub, as provided by the BGP-Protocol. For dynamic routing, this table has to be actualized with transport bandwidth information, represented by the free capacity of vehicles on the road that leads to a particular challenge.

Table 6: Transformation of IPv6 datagram from DI to RBPI

IP attribute	# bit	Function within DI	Function within RBPI
Version	4	Version of the IP-Protocol, distinguished between IPv4 and IPv6.	Distinguishable versions of the *handling of π-containers* [disposable or reusable].
Traffic Class	8	This field defines different priority levels in terms of quality of service.	Traffic class in the RBPI might signify requirements, like *special treatment* [temperature] and/or *express deliveries* [time].
Flow Label	20	Contains the randomly selected identification number of a virtual end-to-end connection.	*Label packages that require extraordinary treatment*, for example the transport of animals.
Payload Length	16	Specifies how many bytes follow the header as so-called payload.	Specifies how much space or weight is required in the π-container, like *payload, available transport volume* or *max. number of pallets.*
Next Header	8	Mode indication 'connectionless' or 'connection-oriented'.	Specification of routing method.
Hop Limit	8	This attribute specifies the maximum number of network nodes through which a packet may pass before it is dropped.	This attribute can be transferred 1:1 to physical logistics, this would correspond to the *maximum number of intermediate π-nodes within the transport chain*.
Source/ Destination	32/32	Determines the source and destination address.	

Transformation of Layer 2: From Data Link Layer (OSI) to Link Layer (OLI)

The Data Link Layer ensures '*a reliable and functioning connection between both the end device and the transmission medium*' (Badach and Hoffmann, 2007). It '*takes the datagram from the network layer and passes it from the current node to the next node in the network*' *(Liesa et al., 2020)*. In order to prevent transmission errors and data loss, this layer contains functions for error detection, error correction and data flow control. The physical addressing of data packets also takes place on this layer.

Table 7: Link Layer Transformation Table

Task within DI	Principle	Task within RBPI
Framing		
Divides bitstream messages into manageable data units (frames).	Speciali-zation	As every manageable unit must be transportable individually. It corresponds to the smallest unit, e.g. a pallet.
Physical addressing		
The link layer adds a header to a frame containing the address of sender and/or receiver.	Instanti-ation	Corresponds to the delivery address of the shipment.
Access control		
If two or more devices are using one link, protocols are necessary to determine which device has control over the link and when.	Analogy	Access to a bandwidth medium means access to appropriate π-movers with free capacity. Hence, remote access to the π-mover's cargo space must be ensured.
Flow control		
If data is sent faster than it can be received, a flow control mechanism is used to avoid overwhelming the receiver.	Speciali-zation	If a π-node is at its capacity or handling limit, transport vehicles on their way to the π-node could be informed in order to avoid congestion in the π-node.
Error correction		
When a damaged or lost message is detected, it is resent.	Analogy	Lost freight components cannot simply be sent again, but must be reordered. If damaged or lost freight components cannot be reproduced from the current π-node (e.g. 3D-Print), they must be reordered from the beginning.

On the DI, data packages that are handled are checked for damage. In the case that a package is either damaged or lost, it is simply generated and sent again from the current node. This leads to a problem with the transfer of this functionality into the world of physical objects. In some cases this can be compensated by generic part design at the hub location (3D print).

Transformation of Layer 1: The Physical Layer

The task of the Physical Layer in the DI is to '*ensure the transport of unformatted digital information units'* (Badach, Hoffmann and Knauer, 1997). Hence, this layer focuses on the electrical and mechanical properties of the transmission media. For this purpose, a transmission channel is provided via which information is exchanged, such as electrical cables, fiber optics or plug connections. Within the RBPI, the physical layer handles the movement of the physical objects in road-based vehicles. As discussed in (Kaup and Demircioglu, 2017), the correspondence to transmission media are the vehicles on the road. Hence, the bandwidth counterpart of transport media is the free capacity of vehicles moving on roads.

Table 8: Physical Layer Transformation Table

Task within DI	Method	Task within RBPI
Transport Channel Provision		
Bandwidth is given by the media, like electric wiring or fiber optics.	Analogy	Bandwidth definition correlates to free capacity of vehicles on the road.
Medium provisioning.	Analogy for π-movers	Load securing, so that freight is safely stowed in the vehicle.
Amplification		
In order to prevent a signal from becoming so weak over long distances that it can no longer be interpreted, it is amplified in the network.	Instantiation for π-movers	Refueling π-movers with gasoline, electricity, compressed or liquid hydrogen.
	Analogy for π-containers	Preserve cargo, for example, from breaking the cold chain. Maintenance of services for π-containers, e.g. the provision of temperature or power control.

Bandwidth is given in the DI by the media itself, like electric wiring or fiber optics. This corresponds in a one-to-one transformation to smaller streets (analogy to electric wiring) or highways (analogy to fiber optics). With respect to the RBPI this analogy is not best suitable as the term of bandwidth, because there the bandwidth correlates to free capacities of vehicles on the road. Nevertheless, the comparison to road traffic is dared or aspired here. This means that π-movers driving on a road with available capacities correspond to a transmission medium in the DI. In order to determine the transport bandwidth, information on the utilization rate of the driving π-movers is required (Kaup and Demircioglu, 2017). A second task of the physical layer is the amplification of signals (Badach, Hoffmann and Knauer, 1997) in order to prevent a signal from becoming more weak over long distances. An instantiation would lead to the refueling/-charging of vehicles or other kinds of π-movers, like starship robots. However, some freight requires special treatment, such as refrigeration. Hence, the maintenance of containers with temperature and/or power supply is the second analogy to the amplification of signals on the DI.

Artifact Framework Design for RBPI

All functionalities, methods and attributes, transferred from the DI to the RBPI in Section 3 lead to the design of a framework artifact, as shown in Figure 5. The corresponding functionalities of the DI in the RBPI were added to the five layer model from Section 2 that leads to the following artifact:

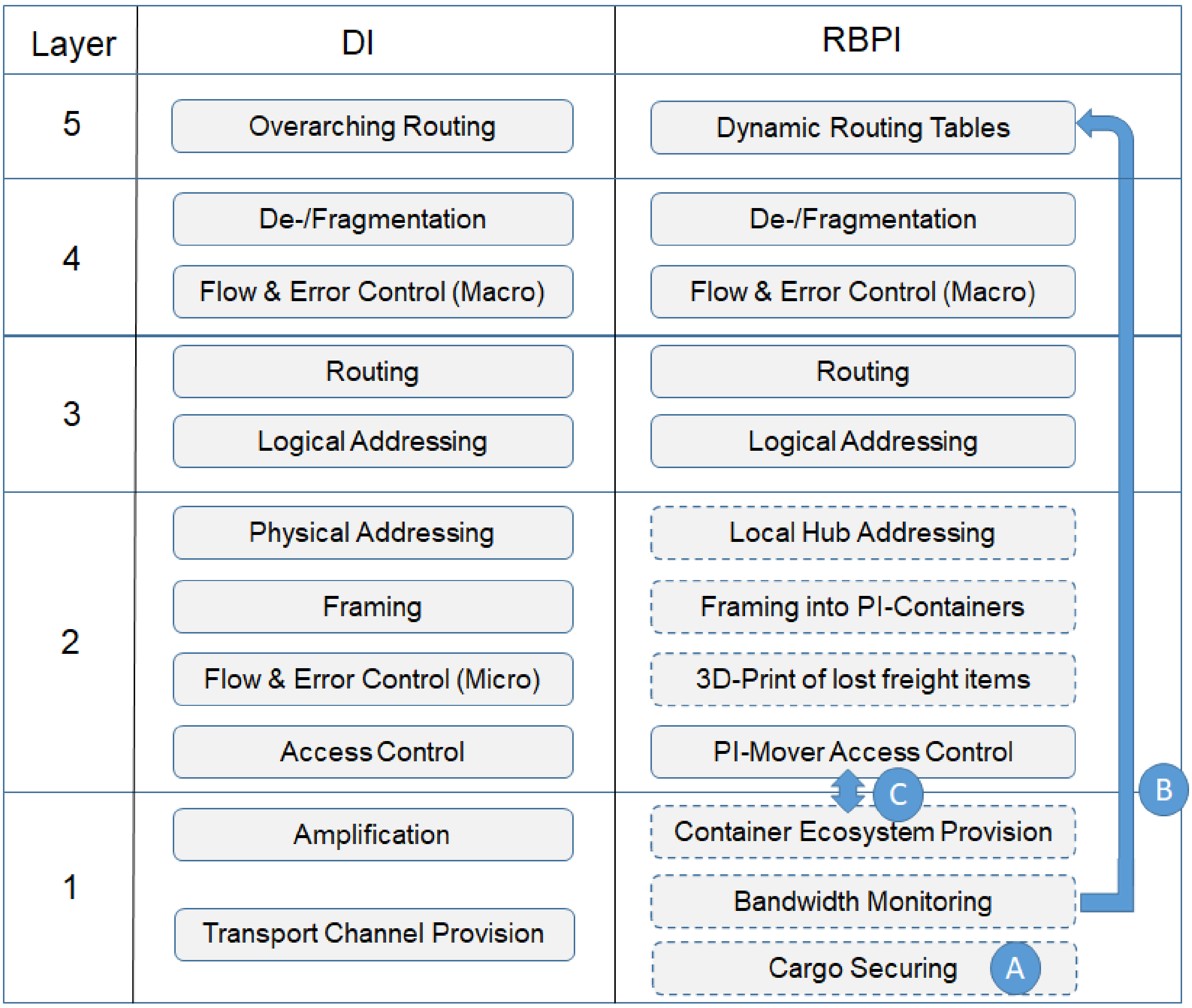


Figure 5: Artifact Framework for RBPI as a result of the Transformation

The rounded rectangles on the left show the functions in the DI and those on the right the corresponding functionalities in the RBPI. When a solid frame of a rectangle is shown in Figure 5 on the RBPI side, a closely coupled principle was performed, such as specialization or instantiation. In the case of a loosely coupled analogy, the frame is shown dotted. The components shown in blue represent identified vehicle requirements. The circle with the (A) indicates that the load is automatically secured against slipping in the cargo space. The circle in which the (B) is located represents the feedback of the freely available capacities of the vehicles to the dynamic routing tables. The freely available transport capacity feedbacks might be replicated to all hubs in the system. This feedback loop is represented by the blue arrow. The prerequisite for this is that the vehicles have knowledge of their load states or their free capacities. This knowledge can either be implemented via an in-vehicle tracking system or via a fleet management system. The circle with the letter (C) represents the optional power supply of containers that provide the goods with special treatment, such as cooling.

5. *Conclusions and further work*

In the future, the Physical Internet will gain importance in transportation and logistics, not least for road-based transport. This paper identifies protocols of the Data Internet that are relevant for transport and transforms their properties and methods to the RBPI, which answers SQ_1. The result of this work is a framework artifact, which represents transferred and adapted methods of the DI for the RBPI. This design artifact answers SQ_2. From this point, vehicle-relevant requirements are derived in order to answer SQ_3. In future, vehicles should have an automatic load securing system, as well as be able to provide remote access to the cargo space. In order to be able to use real-time information for routing freight within the RBPI, vehicles must have information about their loading status and communicate these to the RBPI's routing entity. It is supposed that future work might consider how π-containers can provide an ecosystem for including freight, e.g. a cooling mechanism. For this purpose, a power supply between vehicles and π-containers has to be designed. The automotive industry as well as logistics operators are recommended to take these requirements into account in order to make vehicles and π-containers fit to the RBPI.

References

Badach, A. and Hoffmann, E. (2007) *Technik der IP-Netze: TCP/IP incl. IPv6-Funktionsweise, Protokolle und Dienste*. Hanser Verlag.

Badach, A., Hoffmann, E. and Knauer, O. (1997) *High Speed Internetworking*. Addison-Wesley Longman.

vom Brocke, J. (2007) 'Design Principles for Reference Modeling: Reusing Information Models by Means of Aggregation, Specialisation, Instantiation, and Analogy', in *Reference Modeling for Business Systems Analysis*. IGI Global, pp. 47–76. doi: 10.4018/978-1-59904-054-7.ch003.

Dong, C. and Franklin, R. (2020) 'From the Digital Internet to the Physical Internet: A conceptual framework with a stylized network model', *Journal of business logistics*, (jbl.12253). doi: 10.1111/jbl.12253.

Fielding, R. T. (2000) *Architectural styles and the design of network-based software architectures*. University of California, Irvine Irvine. Available at: https://www.ics.uci.edu/~fielding/pubs/dissertation/fielding_dissertation.pdf.

International Transport Forum (2019) *ITF Transport Outlook 2019*. OECD Publishing. Available at: https://play.google.com/store/books/details?id=p8SZDwAAQBAJ.

Kaup, S. and Demircioglu, A. V. (2017) 'Von der Crowd-Logistik hin zu einem ganzheitlichen Ansatz hocheffizienten Warentransports', *Wirtschaftsinformatik & Management*, 9(3), pp. 18–27. doi: 10.1007/s35764-017-0052-z.

Kaup, S., Ludwig, A. and Franczyk, B. (2020) 'Design and Evaluation of Routing Artifacts as a Part of the Physical Internet Framework', *arXiv [cs.NI]*. Available at: http://arxiv.org/abs/2011.09972.

Kaup, S. and Neumayer, B. (2003) *Rechnernetze und Datensicherheit*. Shaker.

Kumar, P. (2020) 'QUIC (Quick UDP Internet Connections) -- A Quick Study', *arXiv [cs.NI]*. Available at: http://arxiv.org/abs/2010.03059.

Liesa, F. *et al.* (2020) *Physical Internet Roadmap*. D2.1. SENSE-Project.

van Luik, S. *et al.* (2020) 'On the value of the Digital Internet/Physical Internet analogy', *Journal of Supply Chain Management Science*, 1(3-4), pp. 87–103. Available at: https://journals.open.tudelft.nl/jscms/article/view/5300.

Markillie, P. (2006) 'The Physical Internet, A survey of Logistics', *The Economist*, 17 June.

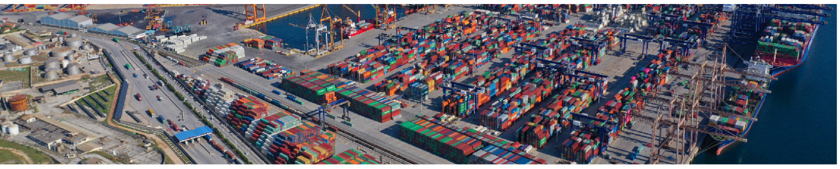
Montreuil, B. (2012) 'Physical Internet Manifesto, Version 1.11. 1', *CIRRELT Interuniversity Research Center on Enterprise Networks, Logistics and Transportation*, pp. 2–3. Available at: https://www.slideshare.net/physical_internet/physical-internet-manifesto-eng-version-1111-20121119-15252441.

Montreuil, B., Meller, R. D. and Ballot, E. (2010) 'Towards a Physical Internet: the Impact on Logistics Facilities and Material Handling Systems Design and Innovation'. Available at: https://digitalcommons.georgiasouthern.edu/pmhr_2010/40/ (Accessed: 16 June 2020).

Shi-Wan, L. *et al.* (2017) *Architecture Alignment and Interoperability*. Industrial Internet Consortium.

Statistisches Bundesamt (2019) 'Güterverkehr 2018'. Available at: https://www.destatis.de/DE/Themen/Branchen-Unternehmen/Transport-Verkehr/Gueterverkehr/_inhal t.html#sprg235064.

Zikria, Y. B. *et al.* (2018) 'Internet of Things (IoT): Operating System, Applications and Protocols Design, and Validation Techniques', *Future generations computer systems: FGCS*, 88, pp. 699–706. doi: 10.1016/j.future.2018.07.058.



Towards an open and neutral data sharing infrastructure – sharing state information

Wout Hofman[1], Mitchell Out[2] and Hasse Romer[3]

1. TNO, the Hague, the Netherlands
2. Dutch Customs Administration, Rotterdam, the Netherlands
3. Ericsson, Stockholm, Sweden

Corresponding author: wout.hofman@tno.nl

Abstract: *Data sharing is essential, not only for the realizing the Physical Internet, but also for efficiency – and effectiveness improvements of both business – and regulatory processes. Improved effectiveness of regulatory processes will contribute to seamless goods flows. This paper introduces a way of data sharing in supply and logistics with Linked Data. Data stays at the source and basically only links to that data are shared. Applying Linked Data requires additional functionality from a (cyber-)security perspective like Identification, Authentication, and Authorization (IAA) and nonrepudiation. This type of functionality is already applied in supplied and logistics, but only in a limited case within framework contract relations. Supporting data sharing in an open organizational network with resilient, agile, and compliant supply and logistics chains requires to implement these mechanisms. This contribution will also reflect on the way the required functionality is implemented by current platforms and solutions.*

Keywords: *data sharing, data sovereignty, infrastructure, federated network of platforms*

1 Introduction

Digitization and data sharing are a key for creation of a resilient and agile Physical Internet. Optimization of node operation and interconnection of logistics networks is based on the capability to share data amongst stakeholders, compliant with data sovereignty. There is a clear requirement to share state information for optimal capacity utilization of logistics networks and nodes, logistics customers would like to have predictability and increased visibility of their goods flows, and authorities require improved data quality for a seamless goods flows based on the data pipeline concept (van Stijn, et al. 2011). All stakeholders involved would like to know the past, present, and future state of their physical object(s) of interest, e.g. goods – and truck tracks.

Literature (Pan, et al. 2021) mentions that four aspects are relevant for digitization and transformation of the transport industry, namely digital platforms, blockchains with Application Programming Interfaces (APIs), Digital Twins and Cyber-physical systems, and data-driven solution design, planning, modelling, and control. Various platforms with different functionality can address data sharing (Kenney en Zysman 2016). Many platforms already exist in supply and logistics (Zomer, Hoesen en Zuidwijk 2021). These will have various consequences, like who owns the platform and thus who captures the value? Similar questions can be posed for blockchain based applications (van Engelenburg, et al. 2020), showing how rights can be distributed illustrated by the TradeLens example. APIs are a well-known technology that can be applied in many ways, not only for internal IT systems, but also by platforms and blockchain applications. These are part of a (technical) solution. New initiatives have been initiated by the founder of the Web, Tim Berners-Lee, for constructing a completely distributed solution for the Social Web (Solanki 2021) based on Linked Data (Bizer, Heath en Berners-Lee 2011).

Linked Data is on sharing links between data stored in heterogeneous IT systems. Several challenges have to be addressed, namely the data semantics and action set, required functionality, and the technology for sharing the data. Some take the approach for sharing links between heterogeneous semantic models, which implies that a model of a dominant player will be leading with respect to another (Choudry 1997). Less dominant players have to deal with all these different solutions; ideally, they could apply ontology alignment, which however requires additional effort (Mohammadi 2020). The second challenge addresses the various technologies that can be applied for sharing (links to) data. The reference architecture of the International Data Space Association is an approach (Dalmolen, et al. 2019), blockchain based solutions another one (Hofman, Dalmolen en Spek, Supply Chain Visibility Ledger 2019). In case of applying linked data and its technology, the semantic web architecture can be applied (Berners-Lee 2006). It is about specifying semantics with OWL (Ontology Web Language), rules with RIF (Rule Interchange Format), querying using SPARQL (SPARQL Protocol and RDF Query Language), and sharing with RDF (Resource Description Framework), all open standards developed by the World Wide Web Consortium. It implies that all stakeholders implement this stack in their IT environment, which is a challenge.

In this paper, we will present how Linked Data and supporting technology can be applied in the supply and logistics domain. First of all, a case illustrating the various requirements will be given and secondly the principles are introduced. Thereafter, the model and its implementation in an organizational network. Finally, conclusions are given.

2 An example illustrating the requirements

In order explain the requirements or necessities of sharing data globally in logistics, let us do a thought experiment of a pallet which suddenly stands at the doorstep of a warehouse operator.

The pallet and its boxes have several identifiers such as shipper's pallet and box numbers. However, they have very little meaning to the warehouse operator. Instead, what if we have a universal identifier that links to additional information of the pallet and its boxes, one that globally had meaning to everyone. It just so happens that the W3C (World Wide Web consortium) have defined just one like that, the URI (universal resource identifier), or more often the location reference of the URI expressed as an URL, i.e. www.mycompany.com/thelonelypallet.

The warehouse operator scans the URI to retrieve necessary information from the source and update related information such as goods receipt at warehouse or similar. Of course, the warehouse operator would like to know if the link on the pallet is not used for cyber security attacks. Thus, the link has to be authenticated. Furthermore, the data holder of pallet information wants to know if the warehouse operator is who he says he is and if he is authorised to access the pallet information. Identification, Authentication, and Authorisation (IAA) need to be arranged, supported by federated identity providers and agreed access rules, so that data holder and -user can safely share data via the link on the pallet.

The warehouse operator should also understand the information and instruction that it retrieves, for handling the pallet. The pallet, the boxes, or their content can be stored in the warehouse, where commercial conditions for storage are known ('who will pay the bill'). This is best solved by reference to a common dictionary and action set as part of the data that is shared, i.e. using semantic models or linked data to define the data sent. This would also enable the warehouse operator if it doesn't understand the data received to find a data transformation service to its preferred language for data.

There are many other cases that can be handled in (W. Hofman, A methodological approach for development and deployment of Data Sharing in Complex Organizational Supply and Logistics Networks with Blockchain Technology 2019) the same way, where data stays at the source and only links to data are shared. Think of a trip of a truck picking up and dropping of cargo, temperature-controlled transport of vaccines, and bridges opening and closing where two infrastructures cross each other.

The proposed solution properly works with a common semantic model with business process alignment, IAA, and applying URIs as identifiers for accessing data of a data holder.

3 High level requirements and operating framework

There is a number of drivers for data sharing. Of course, we can list the dominance of large platforms that take over the market, not only for supply, but also logistics. Another driver is the current COVID pandemium. It has learnt that supply chain resilience and agility are required. In many cases, resilience refers to creating the necessary safety stocks and agility to the means of making them available in time at the required locations. Resilience and agility prosper by visibility. A third driver for change is the so-called EU Green Deal (European Commission 2021). Sustainability of logistics operations can increase with optimization of capacity utilization. It requires visibility of available capacity and sharing capacity across by creating a system of logistics networks (Alliance for Logistics Innovation through Collaboration in Europe (ALICE) 2020).

To address these challenges, any solution should adhere to a set of (high level) requirements or principles and should operate accordingly. The latter is called the operating framework that has to be implemented (see for instance (FEDeRATED 2019). Various projects like FEDeRATED (federatedplatforms.eu) and FENIX (Fenix-network.eu) have established such a set of requirements that underpin the developments at European Union level (The Digital Transport and Logistics Forum (DTLF) 2017). The assumption is that there will always be different platforms and technical solutions. Thus, the following requirements and operating constraints can be formulated (see for instance (FEDeRATED 2019):

- Open, neutral, and inclusive: a solution has to be able to support seamless data sharing by all enterprises and authorities operating in trade and logistics, including Small and Medium sized Enterprises (SMEs). There needs to be a level playing field enabling all stakeholders to share data;
- Data sovereignty: each organization is in control of sharing data with any other organization, compliant with regulations;
- Mobility and freight: both mobility (passengers) and freight (cargo) can be combined in operations, using the same infrastructure and transport means;
- Enterprises and authorities: whereas enterprises share data for their processes (B2B – Business to Business), they also have to make data available to authorities for compliance to regulations (B2A – Business to Administration). At the same time, authorities can provide data to business (A2B) for optimization of goods flows;

- Distributed development and maintenance: it should be possible that a solution can be developed by several stakeholders, where each stakeholder adds his particular view to a total solution. The various parts forming a total solution should be consistent, thus enabling individual organizations to implement those parts of the solution that supports their business. This aspect refers to governance;
- Based on common agreements: each individual organization can join and share data. There should not be any prior (bilateral) agreement on which data to share and which regulations or data processing agreements to apply. A common rules set should be applicable, including organizational interoperability and a legal framework.
- Trust: a total solution should provide trust for its use by individual organizations. Trust is at various levels, like business -, data -, and technical level. Trust at business level is outside scope, but trust should be provided at data – and technical level, compliant with existing or new regulations like the Data Governance Act (European Commission 2020). Trust might require a new legal framework, additional to other regulations.
- Re-use of existing solutions and technology: where possible existing solutions like platforms and (innovative) technology should be re-used and existing implementations of open and defacto standards should be supported. This will allow a rapid implementation and enable migration; and
- Support innovation: future innovations in mobility and logistics should be easily supported by data sharing applications in an easy way. It means the solution is based on an open world assumption and innovative requirements should be deployed rapidly.

The aforementioned already indicates the necessity of a common legal framework for data sharing, of which the aforementioned Data Governance Act needs to be part, and a governance structure to ensure a complete and consistent solution developed and deployed in a distributed way. These two aspects are for further research.

These requirements and the operating framework result in some basic design choices, namely (The Digital Transport and Logistics Forum (DTLF) 2017):

- Technology Independent Service – there should be a set of IT services supporting B2B, B2G, and G2B data sharing. These services should support business processes of enterprises, include compliance with regulations, provide optimal, safe use of infrastructures and should have a shared semantics.
- Architecture – there should be an architecture specifying the basis functionality with visualization of support of different solutions and platforms.
- Plug and Play – organizations should be able to register with a solution of their choice and be able to share data with any other organization in a safe and trusted way, according the Technology Independent Services.

The following parts of this paper present the underlying model for the Technology Independent Services and the architecture.

4 Linked Data and Digital Twins for supply and logistics

Linked data is often referred to as a technology stack, i.e. the semantic web stack (Berners-Lee 2006). This stack enables sharing links to whatever concepts using RDF. These so-called triples, that can be represented by a Uniform Resource Identifier (URI), specify associations between different concepts, which can be anything ranging from a business document data set to a physical object like a container. A more formal and methodological approach can be found in (W. Hofman 2019).

First of all, we introduce the concept of 'Digital Twin', similar to that originally defined by (Glaessgen en Stargel 2012) as 'a composable digital representation of a cyber-physical system with its lifecycle'. The lifecycle is continuously updated from design and operational data, representing the state of a physical object in a timely manner, measured by for instance sensors. The lifecycle might include the predicted future state of a physical object, which can be predicted using data analytics with historic data. One of the properties of a Digital Twin is 'time', as second is 'geospatial'. The latter makes it possible to refer to so-called fixed objects, i.e. objects that stay at a location and perform a logistics function like a crane or a warehouse, and moving objects like for instance trucks, vessels, and containers. The lifecycle of these Digital Twins refers to their condition, like a fairway can have a change of water depth over time, the tires of a truck will wear out, and the temperature setting of a container may change over time. The state of a Digital Twin can also be its unavailability in a time period, like the opening time of a bridge in a railway infrastructure.

Links thus refer to a Digital Twin and its state over time. Since the location of a moving object will change over time, this can also be shared by linked data, where the link contains additional data:

- A physical object is at some location at a time, where that location provides some type of logistics function with constraints, e.g. a truck is on the road or a vessel uses a fairway. Constraints for road transport are for instance a speed limit and the safety of the truck; constraints for using a fairway are in the water depth and the passage of vessels using that fairway.
- The association between physical objects like a container transported by a truck. The truck will move from one location to another and based on the existence of the association of the cargo objects of the truck, those will also move. A set of rules can be specified to derive knowledge of the cargo from its transport means. There will also be constraints representing which associations may exist between physical objects, e.g. container vessels can only transport containers and not packages or pallets.
- The afore mentioned associations have to be represented 'timely', i.e. the past, present, and future, where the future can be predicted (e.g. an Estimated Time of Arrival) or prescribed (e.g. a required arrival time of a vessel in a port).

Secondly, we introduce the concept of '<u>business transaction</u>' as 'the structured set of interactions between a customer - and service provider role to perform a logistics activity according to its business service' (Hofman 2014). This definition contains various aspects like two roles: customer and service provider. Intuitively, organizations can have these roles. By representing 'person', which can be subtyped into 'natural –' or 'legal person', as Digital Twin, any two Digital Twins can have this role. This implies a natural person can have a business transaction with a legal entity, but moreover, a physical object like a container can be a customer of a transport means providing a transportation activity. Thus, a business service is a property of a Digital Twin expressing its capabilities and any relevant constraints in its relation to another Digital Twin. These capabilities can be expressed timely, e.g. a future port call of a vessel.

A structured set of interaction is further relevant in this context. It refers to a negotiation between a customer – and service provider role with respect to for instance conditions, prices, and time at which a logistics activity will be performed. Here the 'expected', 'planned' or ‘required’, ‘estimated’, and ‘actual’ time are introduced: the expected is the one provided by a customer, the others are provided by a service provider. ‘Time’ refers to an association between Digital Twins and a location or to an association between Digital Twins.

Thirdly, for liability and ownership, a service provider logs the Digital Twins and their state at the start and end of performing a logistic activity on behalf of a customer. The log may also contain the agreed (standard) conditions (and prices). This log is normally represented by a business document that functions as contract of carriage like a Bill of Lading (B/L. sea transport), a CMR (road transport), or Airway Bill (air transport). The B/L has also the function of ownership: the one that is able to present the original B/L is owner of the cargo transported by a vessel. The log may be extended with pictures of the cargo in case of damage during transport.

Finally, authorities may impose constraints on the state of a Digital Twin. We have already mentioned safety and security constraints of an infrastructure provided by an infrastructure manager, but other constraints relate to a safe, secure, fair, and sustainable society. Examples of the latter constraints are prevention of tax evasion and smuggling. Most traders want to be compliant, but their trade flows can be used for smuggling. These types of constraints are governed via risk assessment algorithms on trade data that represent goods flows. Authorities construct Digital Twins and their state and perform a (complex) state analysis for risk assessment. It may cause delays, meaning the state of (moving) Digital Twins is altered.

With respect to modelling Digital Twins, we already gave an example of reasoning. By applying semantic technology, other examples can be specified where additional knowledge is included in a model representing Digital Twins and their state. An example is ‘container track’ that is ‘the timed sequence of locations which a container has passed’ or an ‘itinerary’ representing the ‘timed sequence of locations where a Digital Twin representing a transport means will load and/or discharge other Digital Twins based on its constraints’. In the same way, all types of events can be formulated like ‘container loaded’, ‘gate-in’, and ‘ETA of a vessel’ referring to a location where a logistics activity will take place.

Concluding, links are shared to reflect the state of Digital Twins between a customer – and service provider role, and they are shared with authorities. The simplest link refers to a business document data set logged for liability and ownership; other links refer to the timely state of a Digital Twin like its movement and change of condition. This approach implements data sovereignty.

5 Functionality

Instead of applying linked data, data can also be made available by its data holder via a message to a data user. In such a case, a data user would like to know the identity of the sender of the data, i.e. the identity of the data holder. The data holder would also like to know that only the selected data user receives the data. In framework contracts between a data holder and -user in their roles of customer and service provider, this issue is addressed by applying end-to-end data security, either via peer-to-peer data sharing or applying an intermediate, trusted platform. End-to-end security can be achieved via asymmetric encryption where data is first encrypted with the private key of the data holder/-sender and this encrypted data is also encrypted with the public key of the data recipient/-user. Only the data recipient is able to decrypt the data using its private key and authenticate the data sender by decrypting the data with the public key of the sender.

The same mechanism of end-to-end security can be applied for sharing linked data. It requires what is called a Public Key Infrastructure (PKI). This is mostly applied at link level between any two systems that are interconnected. In case there are systems between that of a data holder and -user, individual links can apply PKI, but end-to-end security is not ensured. Applying standardized Application Programming Interfaces (APIs) supporting the semantic model and business interactions do not require the use of intermediate platforms; they can be applied end-to-end and thus end-to-end encryption and authentication can be applied. However, a number of stakeholders will use platforms. Two questions thus should be posed, namely (1) can a data user trust the link it receives and (2) can a data holder authenticate the data user that evaluates a link. The trust in receiving a link can be addressed by a chain of trust of the intermediate systems (platforms) and links, which answers the first question.

The Internet Engineering Task Force (IETF) has introduced the mechanism of an authorization token (Internet Engineering Task Force 2021), where a data user can get a token for accessing data of a data holder. Such a token cannot be traced to an employee of an organization. Each organization needs to have proper authorization mechanisms in place, based on its Identity Provider, where only authorized employees can get an authorization token. Furthermore, the authorization token can also have a validity, which can be in time (it is valid for a period of time) or per use (it is valid for a number of uses). A data holder will have the capability to authenticate the token and based on its access control mechanisms provide access to the data.

What is required is a federated identity where authorization tokens are issued. Multiple Identity Providers will have to register with one or more Identity Broker (European Commission 2021). According to this Regulations, European Union Member States have installed eIDAS brokers for certified Identity Providers. Similar mechanisms have been implemented by the private sector.

An alternative solution is where each natural person collects its identities and stores them in a so-called wallet. The mechanism is based on Decentralised Identifiers (DIDs, (Reed, et al. 2021)). The data model of DIDs and mechanisms to apply these DIDs are specified. However, DIDs in the context of evaluating linked data should also integrate with organizations providing and revocation of these DIDs to their employees. Certain blockchain based solutions for DIDs like Sovrin have developed and implemented these in- and revocation mechanisms (Lodder en Hardman 2021). These blockchain solutions with their wallet mechanism serve as globally applicable Identity Providers. It requires a data holder and – user to integrate with such a solution for invocation, revocation, and authentication of DIDs. There can be specific constraints of a global or EU trade community on the assertions to be applicable. For instance, EU member states or even particular authorities will only accept particular registrations like an EORI (Economic Operators Registration and Identification) for traders that can assign a DID to their employees. Applying DIDs makes an action traceable to a particular natural person, which is not the case when applying OAuth tokens. This needs further consideration in the context of liability.

Thus, IAA is still complex and requires further attention and agreement amongst various stakeholders. Additionally, other aspects will have to be addressed. For instance, the required business process action based on information retrieved from a data holder. It requires a standardization of interaction sequencing between a customer and service provider as part of a business transaction (see before). Other aspects are:

- Discoverability – finding the proper business services and linked data and processing the linked data. This will result in a Service Register, an Index, and a search mechanism. Search mechanism might be expressed as derived concepts in the semantic model, resulting in the evaluation of multiple links (see before).

- Logging and auditing – all data that is shared between a data holder and – user needs to be logged, including the Identity and Authentication mechanism applied for sharing and timestamps. This functionality supports non-repudiation.
- Data transformation – it has already been indicated that a data holder will provide transformation to the common semantic model, potentially by transforming derived concepts to the main concepts of the model.

Logs and audit trails don't only provide non-repudiation, they can also be a basis for payment related to data consumption. Potentially, a number of APIs are not for free, in such a case payment can be triggered from the log. The log and audit trail can also trigger the payment for business services. A particular log entry like sharing CMR data at delivery or registering a arrival notification can be a trigger for payment.

6 Implementation variants

We see three implementation variants for organizations, all of which can be implemented using different technology. We will map the various technology options to these implementation variants.

With respect to the functionality, each organization should implement or select identity provision and authentication. Each organization should also implement access control as part of their internal data store. In the proposed approach, we consider all other specific functionality bundled into a single solution. The functionality can of course be decomposed leading to hybrid solutions, e.g. a shared Service Register with a local Index.

The implementation variants are:

- IT adaptation – an organization implements all the functionality themselves. It implies that an organization will have to implement all the interfaces themselves.
- Single Organization Node implementation – an organization implements a (standard) node providing the particular functionality.
- Shared Organization Node implementation – two or more organizations share a node. A platform is a synonym for such a node.

Having a node, implies that the interfaces with all other organizations can validated once for that node and each organization can just link such a node to its internal IT systems.

A node or platform can also support part of the functionality, especially the business process interactions, and provide derived concepts to one or more users. For instance, a node can provide supply chain visibility for a particular modality, using the terminology for that modality. In road transport, an itinerary is for instance called 'trip'; in sea it is called 'voyage' and in air transport we talk about 'flight'. Such a node will also not support all concepts, sea will for instance focus on vessels and containers, road on trucks, trailers and goods.

The aforementioned implementation variants result in the following implementations, where a distinction between a data holder and – user is made (see next figure). Data holders and data users might implement shared functionality, especially authorities might combine functionality in their role of data users for monitoring transport flows from various regulatory perspectives.

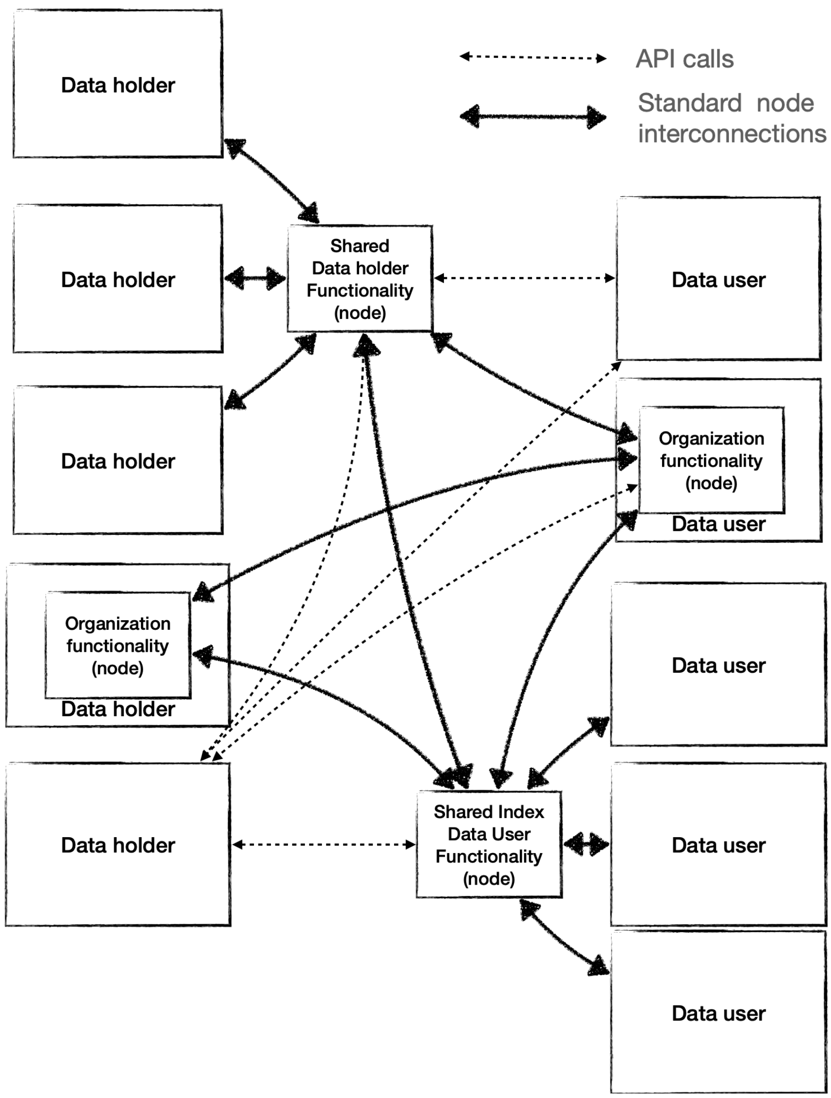


Figure 1: Implementation variants

The figure shows that standard interfaces can be provided by nodes, paving the way to a rapid deployment (bold arrows). Otherwise, each individual organization will have to implement the functionality itself and be able to get some type of certification (if required). Whether or not data is shared using a particular link is based on an organizational network. Such an organizational network is based on business transactions and compliance to regulations (see before).

To add complexity to this figure, there are various technology options to implement a variant. Each technology option will lead to implementation of functionality by an organization, since these technology options will not provide all functionality. Of course, a node can implement the complete semantic web stack, requiring an upload and download function with internal IT systems. Implementation of the semantic web stack also implies automatic support of derived concepts, that need to be programmed otherwise.

The following solutions can be applied for developing a single user or shared node:

- <u>Commercial or community platforms</u> – these basically provide all functionality that is required, including Identity Authentication and Authorization. However, they will all implement the functionality for their specific market or user community, integrating various implementation guides of open and defacto standards.

- International Data Space Association (IDSA) Reference Architecture – it provides a Service Register and a log and audit trail as infrastructural services, meaning they will be used by several organizations. Data sharing is peer-to-peer using so-called connectors. These connectors can deploy so-called apps, where these apps can provide particular functionality like an index.
- FENIX connector – FENIX is a CEF funded Action supporting the development and validation of all concepts. They have developed and validated a connector, where the connector can be implemented by an organization or a platform. A FENIX connector supports a Service Register, authentication, and access control.
- Blockchain technology – there is a variety of technology or blockchain protocols as the blockchain community calls it. Standard, blockchain technology provides nonrepudiation serving as an immutable log and audit trail. Additional functionality, also referred to as 'smart contracts', can specify the Service Register and an Index. There is also a choice whether or not to store data on-chain or only share linked data via a blockchain. Particular technologies will provide configurations for organizational networks, e.g. Hyperledger using channels, or need distribution mechanisms that can be configured by organizations, e.g. Corda where each organization can run a node that is connected to all other nodes, fully supporting data sovereignty.
  A community can define its functionality (van Engelenburg, et al. 2020). There are for instance various solutions sharing links to and supporting data integrity of business documents like B/L and eCMR.
- Broker technology – using broker technology like Kafka implies that all required functionality needs to be developed, with some type of integration with the semantic model.
- Managed cloud service – this is an alternative to the previous solution. Many cloud providers like Microsoft, Oracle, Amazon, and IBM can rapidly deploy the required functionality utilizing standard technical components like brokers, registers, and Identity Providers.

7 Conclusions

We have illustrated how the concept of linked data can be applied to data sharing in supply and logistics. This concept with its various implementation variants will meet all principles and operating rules identified. Using linked data, special attention needs to be given to Identity, Authentication, and Authorisation (IAA) and nonrepudiation to create trusted, safe, and secure data sharing. By only sharing linked data, two levels of security are implemented: a data holder shares only a link with particular data users and implements authentication and access control before sharing data.

What is still required is a governance structure for such an open and neutral data sharing environments based on technology independent services supported by a semantic model. It is all about interfaces, where these interfaces will evolve over time. Governance needs to deal with these aspects. Other important research topics are in the migration towards the proposed solution. Migration can be seen from various perspective like the evolvement of bilateral data sharing to platforms and finally completely distributed data sharing or from sharing links to business documents to event sharing between organizations and finally data sharing between Digital Twins. These all require decision making algorithms based on for instance data analytics.

A final aspect is that of adoption of the solution. One of the main challenges is to identify first movers that have a clear business case that can be solved by the solution. Customs authorities seem to be such a clear case (Hofman, et al. 2021), but developments can also be driven by the requirement to implement new regulations like the electronic Freight Transport Information regulation.

The authors are grateful for the contribution of all participants of the various working groups and teams of the development of the Basic Data sharing Infrastructure in the Netherlands, the FEDeRATED architectural team, and the architectural team of DTLF subgroup 2.

References

van Stijn, E., D. Hesketh, Y. Tan, B. Klieving, S. Overbeek, Heijmann. F., et T. Butterly. 2011. «The Data Pipeline.» *Global Trade Facilitation Conference 2011 Connecting International Trade: Single Windows and Supply Chains in teh Next Decade.* European Commission. 1-26.

Kenney, M., et J. Zysman. 2016. «The Platform Economy.» *Issues in Science and Technology*, April: 61-69.

Zomer, G., M. Hoesen, et R. Zuidwijk. 2021. «A wolf in sheep's clothing or a blessing in disguise? Booking Platforms as disruptors and enabler in the logistics industry.» repository.tno.nl.

Solanki, M.R. 2021. «SOLID: A Web System to Restore the Control of Users' Personal Data.» Dans *ICT Systems and Sustainability, Proceedings of ICT4SD*, de M Tuba, S. Akashe et A. Joshi, 257-267. Springer.

Bizer, C., T. Heath, et T. Berners-Lee. 2011. «Linked Data - the story so farq.» *International Journal on Semantic Web and Information Systems.*

Choudry, Vivek. 1997. «Strategic Choices in the Development of Interorganizational Information Systems.» *Information Systems Research* (pubsonline.informs.org) 8 (1).

Mohammadi, M. 2020. *Ontology Matching: Algorithm, Evaluation, and Application.* Delft: TUDelft.

Dalmolen, Simon, Harrie Bastiaansen, Erwin Somers, Somayeh Djafary, Maarten Kollenstart, et Matthijs Punter. 2019. «Maintaining control over sensitive data in the Physical Internet: towards an open, service oriented, network-model for infrastructural data sovereignthy.» *International Physical Internet Conference (IPIC2019).* Londen.

Hofman, Wout, Simon Dalmolen, et Jacco Spek. 2019. «Supply Chain Visibility Ledger.» *6th International Physical Internet Conference.* Londen.

Berners-Lee, T. 2006. «Artificial Intelligence and the Semantic Web: AAAI 2006 Keynote.» *World Wide Web consortium.* www.w3.org/2006/talks/0718-aaai-tbl/overview.html.

Pan, S., D. Trentesaux, D. McFarlane, B. Montreuil, E. Ballot, et G.Q. Huang. 2021. «Digital interoperability and transformation in logistics and supply chain management: editorialq.» *Computers in Industry 129 (2021) 103462.*

van Engelenburg, S., B. Rukanova, W. Hofman, J. Ubacht, et Yao-hua, Janssen, M. Tan. 2020. «Aligning stakeholder interests, governance requirements and blockchain design in information sharing.» *eGov2020.* online: Springer.

The Digital Transport and Logistics Forum (DTLF). 2017. «An outline for a generic concept for an innovative approach to interoperability in supply and logistics chains.» Discussion Paper, EC DG Move, Brussels.

FEDeRATED. 2019. *Vision Document - FEDeRATED Milestone 1.* FEDeRATED CEF Action.
European Commission. 2020. *Regulation of the European Parliament and of the Council on European Data Governance (Data Governance Act).* Brussels: European Commission Com(2020) 767 final.
Glaessgen, E., et D. Stargel. 2012. «The digital twin paradigm for future NASA and U.S. Air Force vehicles.» *Proc. 53rd AIAA/ASME/ASCE/AHS/ASC Struct. Dyn. Mater. Conference.*
Hofman, Wout. 2014. «Control Tower Architecture for Multi- and Synchromodal Logistics with Real Time Data.» *ILS2014.* Breda.
Alliance for Logistics Innovation through Collaboration in Europe (ALICE). 2020. «Roadmap to the Physical Internet - executive version.» etp-alice.eu.
European Commission. 2021. «The European Green Deal (COM(2019) 640 final).» *Communication from the Commission to the European Parliament, the European Council, the Council, the European Economic and Social Committee, and the Committee of the Regions.* 18 5. https://ec.europa.eu/info/strategy/priorities-2019-2024/european-green-deal_en.
Internet Engineering Task Force. 2021. «The OAuth 2.0 Athorization Framework.» *RFC6749.* 18 5. https://datatracker.ietf.org/doc/html/rfc6749.
Reed, D., M. Sporny, D. Longley, C. Allen, R. Grant, et M. Sabadello. 2021. *Decentralized identifiers (DIDs) v1.0 - Core architecture, data model and representations.* W3C Candidate Recommendation Draft 15 May2021.
European Commission. 2021. «On electronic Idenitfication and trust services for electronic transactions in the internal market and repealing Directive 1999/93/EC (eIDAS).» *Regulation (EU) no. 910/2014 of the European Parliament and of the Council.* 18 5. https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:32014R0910&from=EN.
Lodder, M., et D. Hardman. 2021. «Sovrin DID Method Specification.» 18 5. https://www.w3.org/TR/sovrin.
Hofman, W., J. Migeotte, M. Labare, B. Rukanova, et Y. Tan. 2021. «Using business data in customs risk management: data quality and data value perspective.» *Electronic Governence Conference (eGov2021).* Granada: Springer.
Hofman, W. 2019. «A methodological approach for development and deployment of Data Sharing in Complex Organizational Supply and Logistics Networks with Blockchain Technology.» *IFAC2019.* Delft.

Dynamic Containerized Consolidation in Physical Internet Enabled Parcel Logistics

Sara Kaboudvand and Benoit Montreuil
Physical Internet Center, Georgia Institute of Technology, Atlanta, USA
Corresponding author: sara.kaboudvand@gatech.edu

Abstract: *Many studies in the parcel logistics literature have accredited the hub-based network structures for better freight consolidation and economies of scale. A downside to this practice, however, is that sorting a large number of parcels at (intermediate) hubs requires a significant investment in real estate, human, and machine resources. Furthermore, the time spent in the hubs for waiting and processing increases the parcels' total in-transit time. Such re-sorting can be bypassed by smartly encapsulating parcels that share common service features and a subsequent destination. In this study, we rely on the capabilities of a Physical Internet enabled logistic network to introduce several decision protocols for dynamic containerized consolidation as pertinent to fast-paced parcel logistic environments. We use agent-based simulation to provide empirical results and elaborate on the benefits of containerized consolidation, specifically in saving handling efforts at hubs.*

Conference Topic(s): *Material handling, Modularization, PI Modelling and Simulation*

Keywords: *Physical Internet, Parcel Logistics, Containerized Consolidation, Agent-Based Simulation*

1 Introduction and Literature Review

In today's last-mile delivery logistics systems, transportation time is not the dominant part of the total in-transit time for the majority of parcels. Instead, the processing and waiting of parcels at hubs, constitutes the major portion of the total time that parcels spend in the system. Besides its impact on the parcels' total in-transit time, sorting parcels at every intermediate hub induces unnecessary costs by requiring significant real-estate, human, and machine resources. These costs are deemed unnecessary as they can be notably reduced through smartly encapsulating parcels that share common service features and a subsequent destination, which allows them to bypass the sorting process at several intermediate hubs. This is called containerized consolidation. Containerized consolidation offers several other benefits. It decreases the chances of in-transit damages to the parcels (by decreasing the number of parcel touches), frees up sorting capacity at critical hubs as containers bypass the sorting process, and simplifies handling, loading, and unloading operations.

Many studies in the parcel logistics literature have accredited the hub-based network structures for better freight consolidation and economies of scale. The freight consolidation problem (FCP) is primarily studied in the context of the Less-than-Truck-Load (LTL) industry. It seeks time/quantity-based aggregation of commodity flows into larger shipments at intermediate/intermodal hubs to save on the shipping cost (Min and Cooper, 1990; Tyan et al., 2003). A different but tightly close class of scientific studies is focused on the Freight Consolidation and Containerization problem (FCCP). The main difference between FCP and FCCP is that FCP deals with consolidating shipping flows to better fill the trucks, while FCCP addresses the problem of aggregating the shipping units into proper size containers of higher fill rates so that trucks can move with less empty space resulting in less transportation cost/footprint (Qin et al., 2014; Tiwari et al., 2021). FCCP is different from the problem we address in that FCCP only considers one-leg container consolidation, and as such, it is not concerned with the sorting process at intermediate hubs. In our problem, however, we specifically assess the impact of containerized consolidation on the hubs' material handling cost and effort. Another similar thread of studies relates to the Railroad Blocking Problem (RBP) in the context of train transportation which seeks the optimal assignment of train cars to blocks to minimize the handling cost at intermediate train stations (Newton et al.,1998; Yaghini et al.,2021). Despite its similarities, RBP is also differentiated from the context of our study in several technical and modeling aspects.

To the best of the authors' knowledge, only two published studies in parcel logistics literature consider the direct advantage of containerized consolidation on handling cost and effort. The first study was conducted in 2004 by Chayanupatkul et al. in a project for the USC METRANS transportation center (Chayanupatkul et al., 2004). They plan package routing and containerization for long-haul freight shipments to minimize transportation and sorting costs. In a recent study in 2021, Kaboudvand et al. (2021b) address this problem in the last-mile logistic delivery for indivisible shipments. Due to the operational dynamics in the last-mile logistics, they relax the assumption of fixed fleet schedules, which allows them to reduce the problem symmetry induced in the USC METRANS model and improve the solution performance. They also provide an extensive analysis over the impact of demand features and logistic network configuration on the savings achieved through containerized consolidation. Both of these studies address the containerized consolidation problem at a tactical level. The growth in customers' expectations for high-speed parcel delivery operations, however, narrows the planning horizon a logistic system can secure before finalizing any consolidation decision. In this study, we take advantage of the capabilities of a Physical Internet (PI) enabled logistic network to introduce decision protocols for the dynamic containerized consolidation, mostly targeting the megacities' fast-paced delivery environment.

The Physical Internet (PI) was first introduced by Montreuil (2011), with one of its main pillars relating to encapsulating parcels into smart containers of modular sizes. In a physical internet enabled logistic network, not only containers, but all system components, including hubs, movers, and workers, are equipped with a digital interconnection capability. This capability allows for smoothly and continuously monitoring components throughout the network and thus act timely if one falls out of the planned route, action, or schedule. Due to the computational burden of live processing of data at this large scale, we develop heuristic solutions for tackling the containerized consolidation problem. Furthermore, to evaluate the proposed algorithms, we use an agent-based discrete-event megacity logistics simulator developed by Georgia Tech's Physical Internet Center (Kaboudvand et al., 2021a).

The rest of the paper is organized as follows: Section 2 provides a formal description of the problem and underlying assumptions; In section 3, the solution methodologies together with the simulation model are explained; Section 4 summarizes the numerical results, and finally, the paper is concluded in section 5.

2 Problem Description

The framework introduced in this section, first and foremost, relies on the concept of Physical Internet introduced by Montreuil (2011). Our goal is to leverage the availability of live data in a PI enabled logistic network to introduce and assess dynamic containerized consolidation policies and study their impact on the hubs' overall handling cost and effort. We assume movers continuously send signals ahead of time to the hubs along their path to inform and update their expected time of arrival. We also assume that each parcel has a planned path in the network and a Target Time of Departure (TTD) from every hub along its path to assure a timely arrival at its delivery point.

At the hubs, individual parcels will be waiting to be called for the sorting process by the hub Information System (IS) at the proper time. The hub IS considers the information on arriving trucks and the parcels' service features reflected into their TTD to dynamically prioritize parcels for containerized consolidation. As such, the most priority parcels are released in waves toward the sorting stations to avoid delays or congestion. When the list of parcels of various priorities traveling with the next arriving truck is determined, it is fed into the consolidation and containerization modules. The former module decides on the parcels' next consolidation destination, and the latter determines parcels grouping into containers of potentially different sizes. Less priority parcels are used for filling the priority containers to reach higher fill rates.

Figure 1 provides a simple example supporting the explanation of the hub's IS operation. Assume that at time T_0, the hub's IS gets notified that truck m_1 is heading toward the hub and will be departing current hub at τ_{m_1}. Assume the IS already has estimates on other trucks' arrival and departure including truck m_2 which is heading toward the same next destination as m_1 , say d, and will be departing at τ_{m_2}. Consider three parcels p_1, p_2, and p_3 sitting at the hub with target time of departures TTD_{p_1}, TTD_{p_2}, and TTD_{p_3} , respectively, and all heading toward d as their next immediate destination. In this example, the IS decides that parcels p_1 and p_2 are prioritized for shipment with truck m_1, as their TTD is before the expected departure time for the second truck; while parcel p_3 can wait for truck m_2 and therefore is less priority to be sorted at this time.

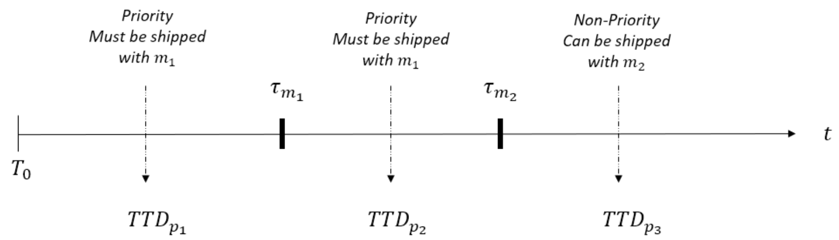


Figure 1: Parcels' Prioritizing for Consolidation at the Hub

Given the problem structure, we propose solution policies for the dynamic parcel consolidation and containerization in the following section.

3 Methodology

In this section, we first present the mathematical basis of the live data tracking in the target parcel logistic system. Next, relying on the availability of desired data, we introduce four heuristic policies for consolidating individual parcels into containers. Finally, we introduce the simulation testbed used for the assessment of proposed methodologies.

3.1 Parcel Tracking

As explained in section 2, in this study, we leverage the capabilities of a dynamic learning system for providing and continuously updating time-dependent estimates of parcels and vehicles' arrival time at every vertex along their planned path. Assume parcel p passes through a sequence of vertices $\{v_0, v_1, \dots, v_n\}$ from its pickup vertex v_0 to its delivery vertex v_n, where a vertex may refer to a hub, or a specific latitude and longitude associated with a customer's home or office. Let $t_{p,0}$ indicate the time at which parcel p is ready for pickup at v_0. Also, let

$\mu_{i,t}$ and $\sigma_{i,t}$ represent the average and standard deviation of shipping time from vertex v_{i-1} to vertex v_i given that the vehicle departs from v_{i-1} at time *t*. Similarly, $\mu'_{i,t}$ and $\sigma'_{i,t}$ show the average and standard deviation of processing time at vertex v_i for parcels arriving at this vertex at time *t*. This information is continuously tracked, learned, and updated through time based on the system's performance at different hours of the day. Assuming a normal distribution for processing and shipping times and considering ρ as the robustness factor, the Target Time of Arrival for parcel p at vertex v_i along its path ($TTA_{p,i}$) is equal to $ETD_{p,i-1} + \mu_{i,ETD_{p,i-1}} + \rho\sigma_{i,ETD_{p,i-1}}$ where $ETD_{p,i-1}$ refers to the Expected Time of Departure for parcel p from vertex v_{i-1} and is computed as $ETD_{p,i-1} = TTA_{p,i-1} + \mu'_{i-1,TTA_{p,i-1}} + \rho\sigma'_{i,TTA_{p,i-1}}$.

In order to secure a target service level, each parcel p is only offered those services that have a delivery time later than the $TTA_{p,n}$ for its fastest shipping path. Let D_p represent the promised delivery time for parcel p at its destination. Then the slack time S_p is computed as $S_p = D_p - TTA_{p,n}$. Let $TTD_{p,i}$ represent the target time before which parcel p should depart from vertex i to be on time at its final destination. To leverage the slack time when computing the parcels' TTD at each vertex v_i along its path, we proportionally distribute this extra time between different legs of the parcels' path using the formulation $TTD_{p,i} = ETD_{p,i} + S_p \times \left(\frac{TTA_{p,i}}{TTA_{p,n}}\right)$.

In the following subsection, we rely on the live availability of the above data to propose heuristic solution methodologies for dynamic consolidation and containerization of parcels.

3.2 Containerized Consolidation Policies

3.2.1 Next and Final Destination Consolidation

The first two consolidation policies lie at two extremes of complete- and no-sorting at intermediate hubs; they are considered as benchmarks for assessing the performance of the other more sophisticated policies. The first policy is the Next Destination (ND) policy, in which each parcel is consolidated toward its immediate next destination. As such, in this policy, parcels are shipped in containers but go through the sorting process at every hub along their path. Conversely, in a Final Consolidation (FD) policy, parcels are destined to the last hub along their path as their consolidation destination. After parcels are grouped based on their consolidation destinations, Algorithm 1 is used to pack each group into containers.

Algorithm 1: Next/ Final/ Top Destination Consolidation

Input: C: set of container types
P: set of priority parcels to get consolidated (sorted based on TTD)
P': set of less priority parcels to get consolidated (sorted based on TTD)
h_p: the consolidation destination for parcel *p*
Output: L: set of containers to be made and their associated parcels
Initialize: $D \leftarrow \cup \{h_p : p \in P\}$ // Set of Consolidation Destinations

for $d \in D$ do:
  $\bar{L} \leftarrow Containerization(\{p \in P : h_p = d\}, C)$;
  for $l \in \bar{L}$ do:
    for $p' \in P' : h_{p'} = d$ do
      if p' fits in l do
        $l \leftarrow l \cup \{p'\}$;
        $P' \leftarrow P' \backslash \{p'\}$;
      end if
    end for
    $L \leftarrow L \cup \{l\}$;
  end for

end for
return L

In order to create a proper set of container sizes for each consolidation destination in Algorithm 1, a containerization algorithm is developed (Algorithm 2), which is an extension of the first-fit decreasing bin-packing algorithm (Baker, 1985) to adapt to multiple container sizes. Due to its computational burden, we do not consider a 3D packing within the context of this study. Instead, we consider a maximum volume fill rate threshold for each container type to allow for extra space between loads with non-compatible dimensions. For each container type, a target minimum utilization threshold is also determined to prevent creating low-utilized containers. Furthermore, for each container created, we check if it can reach a higher utilization by loading less priority parcels that are heading toward the same consolidation destination.

Algorithm 2: Containerization (P,C)

Input: C: set of container types sorted from largest volume to smallest volume
- v_c = volume of container $c \in C$
- δ_c = minimum utilization threshold for container $c \in C$
- λ_c = maximum utilization threshold for container $c \in C$

P: set of parcels to be packed sorted from largest volume to smallest volume
- v'_p = volume of parcel $p \in P$

Output: L: set of containers to be made and their associated parcels

Initialize: $V' \leftarrow \sum_{p \in P} v'_p$

for $c \leftarrow 1$ to $|C|$ do:
  while $V' > \delta_c \times v_c$ or $c = |C|$ do:
    Create a new container φ of type c;
    $r_\varphi \leftarrow \lambda_c \times v_c$;
    for $p \leftarrow 1$ to $|P|$ do:
      if $v'_p \leq r_\varphi$ then:
        add parcel p to container φ;
        $P \leftarrow P \setminus \{p\}$;
        $r_\varphi \leftarrow r_\varphi - v'_p$;
        $V' \leftarrow V' - v'_p$;
      end if
    end for
    $L \leftarrow L \cup \{\varphi\}$;
  end while
end for
return L

The ND and FD policies have their own shortcomings. Parcels never bypass the sorting process in the ND policy, and the small volume of goods shipped between many ODs in the FD policy leads to creating smaller and less utilized containers. In the following subsection, we introduce a new consolidation policy intended to overcome these limitations by enabling parcels to bypass the sorting process at intermediate hubs while using larger and more utilized containers.

3.2.2 Top Destination Consolidation

In the Top Destination (TD) consolidation policy, for each hub h in the network, we identify a limited set of destination hubs that receive the largest flow of shipments passing through hub h. This information is logged per hour of the day and is dynamically updated by continuously tracking and learning the flow of parcels through the network. The TD policy leverages the top destination information for both parcel routing and parcel consolidation at the hubs. A logistic network operating under TD policy incorporates this information into the parcel routing decision so that instead of simply picking the fastest path, it proactively routes parcels through the paths that are expected to have better consolidation opportunities and less handling cost. Algorithm 3 is a scoring algorithm for finding the best consolidation path for a parcel.

Algorithm 3: Find Best Consolidation Path for Parcel p

Input: R: set of feasible paths for parcel p
τ_h: sorting penalty at hub h
H_r: sequence of hubs along path r where $h_{r,i}$ refers to the i'th hub along path r
$TTA_{r,i}$: target time of arrival for parcel p at $h_{r,i}$
Output: r^*: best scored path for parcel p
Initialize: $\rho_r \leftarrow 0$, for $r \in R$ // Penalty for path r

for $r \in R$ do:
    $i \leftarrow 1$ // Index of current hub
    while $i < |H_r|$ do:
        $i \leftarrow i + 1$;
        for $j \leftarrow |H_r|;\ j > i;\ j--$ do:
            if $h_{r,j}$ is top destination for $h_{r,i-1}$ at $TTA_{r,i-1}$ then
                $i \leftarrow j$;
                break;
            end if
        end for
        $\rho_r \leftarrow \rho_r + \tau_{h_{r,i}}$;
    end while
end for
$r^* \leftarrow \arg\min_{r \in R} \rho_r$;
return r^*;

When it comes to the dynamic parcel consolidation at a hub, say h, a parcel's consolidation destination is set to be the furthest hub along the parcel's path that is a top destination for hub h. If none of the hubs along the parcel's path is a top destination for hub h, the parcel is consolidated toward its immediate next destination. After the consolidation destinations for all parcels are determined, Algorithm 1 is used for grouping each batch into proper size containers.

Despite its fairly straightforward operation, a TD policy requires significant computational capacity and capability to enable dynamic tracking of the flow in the system and update the hubs' top destinations per hour. Furthermore, the number of top destinations is a subjective measure determined based on the experts' knowledge of the operating system; its true value can be quite different for each hub, and as such, difficult to identify. In the next section, we introduce a Backward Greedy consolidation heuristic built to capture the benefits of the TD policy while requiring significantly less computational effort.

3.2.3 Backward Greedy Consolidation

The Backward Greedy (BG) consolidation algorithm tends to create highly utilized containers toward parcels' furthest possible consolidation destinations depending on the available flow at the time of decision. Therefore, the consolidation decision is made independent of any specific target destination or time of the day. For each parcel, the BG algorithm first considers the last hub along its path as its target consolidation destination. Then, at each iteration, if enough volume does not exist for a consolidation destination to fill the smallest container size up to its minimum utilization threshold, for all parcels heading toward that destination, the consolidation destination is updated to the immediate hub (along the parcel's path) before the parcel's current consolidation destination. The BG algorithm keeps merging the flows in a backward direction through the parcels' shipping paths until either enough volume is aggregated for filling at least the smallest size container to its minimum utilization threshold or it is not feasible to go backward anymore. The pseudocode for the BG policy is summarized in Algorithm 4.

Algorithm 4: Backward Greedy Consolidation

Input: C: set of container types
P: set of priority parcels to get consolidated (sorted based on TTD)
P': set of less priority parcels to get consolidated (sorted based on TTD)
H_p: sequence of hubs along parcel p's path, where h_{p,i_p} , $i_p \in \{1, \dots, |H^p|\}$ is the consolidation destination for p

Output: L: set of containers to be made and their associated parcels

Initialize: $i_p \leftarrow |H^p|$, for $p \in P, p' \in P'$
$D \leftarrow \cup \left\{ h_{p,i_p} : p \in P, i_p \in \{1, \dots, |H^p|\} \right\}$ // Set of Consolidation Destinations

while $P \neq \emptyset$ do:
    for $d \in D$ do:
        $\bar{L} \leftarrow Containerization(\left\{p \in P : h_{p,i_p} = d\right\}, C)$;
        for $l \in \bar{L}$ do:
            if $D = D^{prev}$ or can fill l to its minimum threshold considering P' then
                for $p' \in P' : h_{p,i_p} = d$ do
                    if p' fits in l do
                        $l \leftarrow l \cup \{p'\}$;
                        $P' \leftarrow P' \setminus \{p'\}$;
                    end if
                end for
                $L \leftarrow L \cup \{l\}$;
                $P \leftarrow P \setminus \{p : p \in l\}$;
            else
                for $p \in l$ do
                    $i_p \leftarrow max(1, i_p - 1)$;
                    $D \leftarrow D \cup \left\{h_{p,i_p}\right\}$;
                    $D \leftarrow D \setminus \left\{d : \nexists p \in P \quad s.t. \quad h_{p,i_p} = d\right\}$;
                end for
            end if
        end for
    end for
    $D^{prev} \leftarrow D$;
end while
return L

3.3 Megacity Logistic Simulator as the Test Bed

In this study, we use the megacity logistic simulator developed by Georgia Tech's Physical Internet Center (GTPIC) to test the proposed containerized consolidation methodologies. This simulator is built based on the topological definitions of an urban territory underlying the concept of a multi-tier hyperconnected logistic web, introduced in Montreuil et al. (2018). The picture on the right side of Figure 2 shows a schematic view of an intracity multi-tier logistic web and a typical path for a parcel shipped through this web, moving up and down to different tiers (green arrows). In the multi-tier logistic web terminology, customer pickup and delivery locations lie within the unit zones and are served from the nearby access hub. As we get to the higher tiers of the network, we deal first with the local and then with the gateway hubs, which are relatively larger and more equipped hubs to facilitate sorting processes.

The GTPIC simulator is capable of modeling megacity logistic operations at single-parcel granularity. It is designed based on several key components, from proactive agents with complex decision-making capabilities (such as customers, demand managers, parcel and vehicle routers, and hub managers), to reactive agents who act according to instructions they receive (such as movers), and objects (including hubs, vehicles, and packages). The simulator is also equipped with a variety of optimization and heuristic algorithms taking over decisions relative to service offering as well as parcel, container, and vehicle routing across the network of hubs to deliver the required services. More details on the GTPIC simulator framework are provided in (Kaboudvand et al., 2021a).

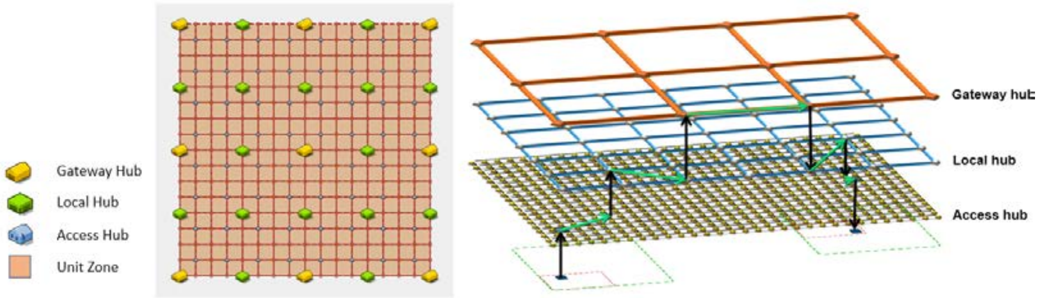


Figure 2: (Left) Simulation Network Under Study, (Right) and Hyperconnected Megacity Logistic Web (from Montreuil et al., 2018)

The left side of Figure 2 provides a screenshot from our target logistic network in this study embedded into the GTPIC simulator. The city has 256 unit zones, 16 local cells, and four urban areas, respectively served by 289 access hubs, 25 local hubs, and 9 gateway hubs. Leveraging the introduced simulator capabilities and functionalities, in this study, we specifically propose decision protocols for the hub managers to guide the containerized consolidation process and study their impact on the overall handling effort and cost throughout the network.

4 Numerical Experiments

For each heuristic policy, embedded into the simulator testbed, the results are reported for one week of the simulation run, leaving one day for the warmup period. We assume all other logistic and operational features, including network configuration, movers' routes, and the demand volume and distribution are the same across all scenarios.

The lower part of Figure 4 includes the assumptions on the container types, dimensions, and costs. These are average costs imposed at the local hubs. For the access and gateway hubs these costs are multiplied by 1.3 and 0.7, respectively, to reflect the larger hubs handling efficiency. The upper part of Figure 4 provides the results on the total handling cost for the default scenario with no consolidation versus the consolidation scenarios including policies ND, FD, BG, and TD for 10, 30, and 50 top destinations per hub. The results suggest that the ND policy incurs more costs even compared to the no-consolidation policy. This is because although savings are achieved in terms of unloading and loading cost using containers, the imposed handling cost associated with emptying and filling containers exceeds the overall savings. In general, the FD and BG have the least total handling costs as they enable more parcels to bypass sortation.

The light blue line in Figure 5 shows the percentage of parcels that are crossdocked in containers, without sorting, in each scenario. The results suggest a tradeoff between the use of larger containers and the percentage of parcels that can be sorted. Smaller containers are more expensive per item but allow for a larger volume of crossdocking and more savings in handling costs. Moreover, the savings in unloading, loading, and crossdocking costs associated with the use of larger containers in the BG policy exceeds the extra sorting cost compared to the FD policy, making BG the least expensive scenario overall (Figure 4). The TD policies lie somewhere between ND and FD both in terms of total handling cost and type of containers used; When we consider fewer top destinations per hub, the model tends to behave similar to the ND policy, while allowing more top destinations skews the results toward the FD policy.

Item	Dimensions (ft)	Cost ($)					
		Unloading	emptying	sorting	filling	xdocking	loading
Parcel	-	0.05	-	0.25	-	-	0.10
tote_1	2 × 1 × 1	0.10	0.05	-	0.10	0.05	0.20
tote_2	2 × 2 × 1	0.15	0.10	-	0.20	0.10	0.30
tote_3	2 × 2 × 2	0.25	0.15	-	0.30	0.15	0.50
tote_4	4 × 2 × 2	0.40	0.20	-	0.40	0.20	0.80

Figure 4: Handling Costs in Different Scenarios

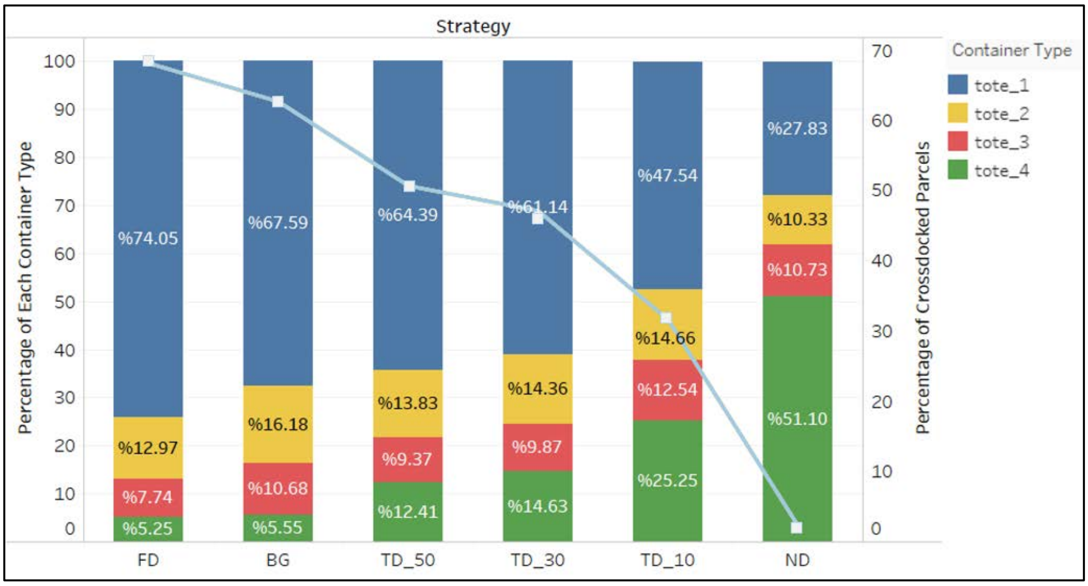


Figure 5: Percentage of Different Container Types Used in Different Scenarios

Although using smaller containers can benefit in terms of handling cost, it may have a negative impact on transport cost. Let $v^C_{m,a,t}$ and $v^P_{m,a,t}$ indicate the total volume of containers and parcels, respectively, transported by mover m, which starts traveling along arc a at time t. Also, let d_a show the distance along arc a. Then the ratio of excess transport capacity required in each consolidation scenario is computed as $\sum_{m,a,t}(v^C_{m,a,t} \times d_a) \,/\, \sum_{m,a,t}(v^P_{m,a,t} \times d_a)$.

Table 1 shows the results on the transport capacity requirement ratio in each scenario. The results suggest that the largest transport capacity is associated with the FD policy as it corresponds to the use of smaller and, as such, overall, less utilized containers. The smallest excess capacity corresponds to the ND policy as it moves larger and highly utilized containers.

Table 1: Shipping Capacity Ratio in Different Scenarios

	FD	BG	TD_50	TD_30	TD_10	ND
Container×km	803 485	765 636	770 340	764 730	722 963	670 072
Parcel×km	588 616	584 526	586 132	592 050	596 745	590 689
Transport Cap Ratio	**1.37**	**1.31**	**1.31**	**1.29**	**1.21**	**1.14**

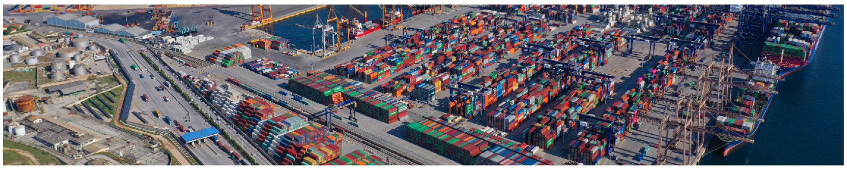
5 Conclusion

In this study, we relied on the Physical Internet capabilities to propose highly scalable solution strategies for the containerized consolidation in dynamic settings where neither time nor computation power is available for optimized decisions. Given the introduced framework, we tested the proposed decision protocols using a holistic simulator platform that models the last-mile parcel delivery operations in a high level of granularity. Simulation results suggest significant handling cost savings through containerized consolidation. They also demonstrate a clear tradeoff between the savings in handling cost through containerized consolidation and the required transportation capacity through the network.

To recognize the full potential of the containerized consolidation, in this study, we assume an unlimited capacity for the hubs. In practice, however, there are limitations associated with the capacity of each hub in the network, and therefore, future research is needed to help better understand the tradeoff between the network operational capacity and the potential for containerized consolidation. Moreover, depending on the size and complexity of the logistic network under study, hybrid optimization-heuristic algorithms can be developed to enable leveraging the full potential of containerization at minimum cost.

References

- Baker S. (1985): *A new proof for the first-fit decreasing bin-packing algorithm*. Journal of Algorithms, v6, no1, 49-70.
- Chayanupatkul A., R.W. Hall, D. Epstein (2004): Freight routing and containerization in a package network that accounts for sortation constraints and costs. Technical report, METRANS.
- Kaboudvand S., M. Campos, B. Montreuil (2021a): *Enabling scientific assessment of large scale hyperconnected urban parcel logistics: agent-based simulator design*. In Proceedings of 2021 IISE Annual Conference, Virtual, May 22-25.
- Kaboudvand S., B. Montreuil, M. Savelsbergh (2021b): *Hyperconnected megacity parcel logistics: joint parcel routing and containerized consolidation*. arXiv preprint (47 pages), arXiv:2103.02080
- Min H., M. Cooper (1990): *A comparative review of analytical studies on freight consolidation and backhauling*. Logistics and Transportation Review, v26, no2, 149-170.
- Montreuil B. (2011): *Toward a physical internet: meeting the global logistics sustainability grand challenge. Logistics Research*, v3, no2, 71-87.
- Montreuil B., S. Buckley, L. Faugere, R. Khir, S. Derhami (2018): *Urban parcel logistics hub and network design: the impact of modularity and hyperconnectivity*. In Proceedings of 15th International Material Handling Research Colloquium, Savannah, GA, USA, July 23-26, 2018.
- Newton H.N., C. Barnhart, P.H. Vance (1998): *Constructing railroad blocking plans to minimize handling costs*. Transportation Science, v32, no4, 330-345.
- Qin H., Z. Zhang, Z. Qi, A. Lim (2014): *The freight consolidation and containerization problem*. European Journal of Operational Research, v234, no1, 37-48.
- Tiwari S., H.M. We, Y. Zhou, L. Tjoeng (2021): *Freight consolidation and containerization strategy under business as usual scenario & carbon tax regulation*. Journal of Cleaner Production, v279, 123270.
- Tyan J.C., F.K. Wang, T.C. Du (2003): *An evaluation of freight consolidation policies in global third party logistics*. Omega, v31, no1, 55-62.
- Yaghini M., M. Mirghavami, A.A. Zare (2021): *A mathematical formulation and an LP-based neighborhood search metaheuristic solution method for the integrated train blocking and shipment path problem*. Networks, 1–20.



IPIC 2021 | 8th International Physical Internet Conference

Assessing the potentialities of Physical Internet for Developing Countries Last Mile deliveries

Sam BAN[1, 2], Andreea DAN[1], Félix GUINET[1], Julien PORTANGUEN[1], Sarot SRANG[2] and Matthieu LAURAS[1]

[1] Industrial Engineering Department, IMT Mines Albi, Toulouse University, Albi, France
[2] Industrial and Mechanical Engineering Department, Institute of Technology of Cambodia Tuol Kouk, Phnom Penh, Cambodia
Corresponding author: matthieu.lauras@mines-albi.fr

Abstract: *Developing countries are so particular as transports are often saturated, infrastructures are in poor condition and the demand is growing so fast. In such a context, the studied research question is: Does Physical Internet (PI) can improve the Last-Mile delivery performance in developing countries? And if so, how and how-much? To start answering this, the paper develops a quick literature analysis of PI solutions to identify which paradigms could bring added value to developing countries. Then, it presents a dedicated simulation-driven method able to assess the potential benefits and limits of applying PI to this specific situation. A Cambodian field-oriented case is finally presented as a very first ongoing experiment to validate the approach and a set of insights regarding further research able to deliver concrete recommendations for both practitioners and scholars is given.*

Conference Topic(s): *Last mile & City Logistics, PI Modelling and Simulation.*

Keywords: *Physical Internet, Supply Chain, Logistics, Last Mile Delivery, Simulation, Developing Countries, Cambodia.*

1 Introduction

Logistics represent a springboard and a key factor as an efficient supply chain allows for smoother and more competitive cities. This aspect is particularly impacting for developing countries which are mainly based on import-export economies (Tang and Bundhoo, 2017). Developing countries' main cities are designed differently from those of developed countries, in terms of organization and means of transport, which provides an opportunity for flexibility (Reda et al., 2020). Our study intends to apply the principles of the Physical Internet (PI) paradigm to developing countries' supply chains in order to assess its potentiality, notably regarding last-mile deliveries' issues. The research work is developed through an ongoing real application case made in the city of Phnom Penh, Cambodia.

Today, Cambodia's infrastructures, as other developing countries' infrastructures, are in poor condition with little maintenance. There is little to no collaboration between logistics actors as the market is rather competitive (Cervero, 2013). Moreover, there is no strategy or real-time monitoring that aims at consolidating the material flows. All of these factors lead to unoptimized roads where huge traffic jams occur on a regular basis, resulting in expensive logistics and prolonged lead-times (Pucher et al., 2005). There are many issues at stake, and they have a real economic interest because annually important financial loss is due to inefficient-logistics costs (Landschützer et al., 2015). In terms of economic development, in Cambodia, current solutions are the followings: investing in infrastructures, changing current import-export legal background or developing a Logistics Master Plan by the Royal

Government of Cambodia. These objectives are only very long-term solutions while the PI logistics approach of our study makes the bet that a better use of the existing current resources might deliver much more effective results. The PI solutions should have the consequence to reduce lead-times, which should increase customer satisfaction and the country's competitiveness thanks to optimized and fewer journeys (Neeraja et al., 2014). In addition, mitigating traffic leads should imply a reduction in environmental and noise pollution, an increase in the level of health and the quality of life of the inhabitants (Jardas et al., 2016). All of these progresses should support important benefits for developing countries such as the creation of jobs and economic activities. That is what our study intends to demonstrate.

The first section of this paper develops a quick literature analysis of both PI paradigms and developing countries logistics features in order to define and justify the research statement of the project. The second section is dedicated to the presentation of the flow simulation system that is currently developed to compare the current situation with some new PI-driven last-mile delivery options. The third section presents the ongoing application of the proposal to the case of an e-commerce company in Phnom Penh. Finally, the paper concludes on the positive benefits of using PI paradigm to logistics and last-mile delivery in developing countries' context that are beginning to be confirmed and open a set of perspectives for further research.

2 Background and problem statement

2.1 Developing countries' logistics features

The logistics performances in Cambodia, as many other developing countries, are difficult to measure due to the lack of data but approximations made by the World Bank indicate that these countries are systematically below-average of the world Logistics Performance Index (Kraay, 2018). Globally the Logistics Performance Index of a developing country represents something like 50% of a developed country… The lack of logistics data for developing countries also leads to the impossibility of drawing a complete picture with the goal of identifying the problems in the transport's networks. Nevertheless, in developing countries, transport seems to be the highest component in terms of costs followed by inventory carrying (Kraay, 2018). Some studies have shown that the most important factors affecting the logistics processes of some developing countries were in order: limited information centralization, either lack of or poor infrastructures, and some specific risks such as corruption (Yang and Wang, 2019). The World Bank also demonstrates a lack of standard contracts which implies the impossibility of implementing advanced partnerships for managing Supply Chains and Logistics' activities. However, most of the developing countries consider now this issue as a critical point and try to invest in the improvement of their capabilities. For instance, the Royal Government of Cambodia established recently a plan of Industrial Development Policy that is to be implemented between 2015 and 2025 which has the main goal of transforming and modernizing the national industrial structure, and particularly the Supply Chain Management capabilities of the country (Kraay, 2018). However, this kind of initiative is of great interest, it will take decades before achieving effective results. That is why, the aim of the current research work is to assess the potentialities of better using existing assets and means to improve the logistics performance.

On their side, (Reda et al., 2020) studied the features of developed, emerging and developing countries in matter of urban logistics. Based on an extensive literature analysis (more than 300 articles were studied), they demonstrated that the main logistics and supply chain issues for developing countries are urban and city logistics problems. Notably, they remind the key features of these countries, already mentioned in (Guimarães et al. 2020), regarding the last-mile delivery question: very high traffic congestion and high level of unsafety on roads; high growth of population (i.e., consumers but also drivers); few and poor-quality infrastructures (roads, rails, ports/airports…); expensive, polluting and unreliable vehicles. They also identify those developing countries are working now on some hot topics which are crowd-shipping, urban consolidation centers, e-commerce delivery, loading/unloading areas, alternative transportation modes, and parcel lockers (Reda et al., 2020). Regarding those challenges, Physical Internet (PI) appears as a potential lever able to support the developing countries' logistics performance, particularly in urban areas, and particularly regarding the last-mile delivery issue.

2.2 The PI opportunities

The PI is a recent and innovative concept of interconnected logistics networks capitalizing on the ability to share resources and information (Montreuil et al., 2012). The definition of the PI is a global logistics system built from the interconnection of logistics networks through a standardized set of collaboration protocols, modular containers, and intelligent interfaces for increased efficiency and sustainability (Ballot et al., 2014). As highlighted in this definition, PI proposes a rethinking of the fundamentals of logistics. The term "interconnection" thus refers to the close and intensive connection between actors and network components. The second key aspect of PI is the desire to open up logistics networks and share assets. Today, companies form private and relatively stable networks that own their own warehouses and vehicles in general. PI breaks with this logic and assumes that assets should be shared among all users of this new network and used as needed. Moreover, PI as a "network of networks" seeks to increase visibility and the possibility of new connections between actors.

This concept of PI was born in order to satisfy the increasing requirements in terms of environment and performance of services. Indeed, the current logistics system, particularly in developing countries, presents dysfunctions that are harmful to the environment and tend to compromise the sustainable development objectives (Ballot et al., 2014). Before it is too late, the PI was designed to offer a chance for logistics services to be more resilient, efficient, sustainable, and adaptable for its users by changing the way physical objects transit through the network (Montreuil et al., 2012). In terms of stakes, most of the theoretical advantages of PI usage have been studied and demonstrated in various application domains (Sallez et al., 2016). But almost all instantiations of PI concern developed countries and not really the rest of the world. This remains to be explored.

It also should be noted that such a paradigm shift in logistics requires significant transformations at various levels. In other words, the accessibility challenge is consequent. First of all, it is necessary to get adapted information systems using advanced technologies to enable the hyperconnection of actors, allowing increased and standardized information sharing and massive data storage (Montreuil et al., 2010). Then, it will be necessary to make the actors accept to evolve towards this new mode of operation and to engage efforts in this direction. The expected investments will be mainly financial and time related (Grest et al., 2019). In this respect, the ALICE (Alliance for Logistics Innovation through Collaboration in Europe) grouping has recently published a roadmap formulating the important steps and associated prerequisites for the implementation of the PI by 2050 (ALICE-ETP, 2020)… in developed countries. However, we can notice that digitalization of developing countries is quite good, and the key players seem ready to explore innovative solutions instead of duplicating old ones issued from developed countries. As a consequence, it seems reasonable to consider the PI accessibility issues feasible in the context of developing countries.

To sum up, the PI is slowly being implemented almost everywhere in the world, but it needs to be more widely known and especially convincing to be more widely adopted, especially in developing countries. This project therefore aims to show whether or not PI is of interest to the last-mile deliveries in developing countries' urban logistics context. Notably, according to (Ban et al., 2020), the following concrete objectives should be challenged and assessed: (i) reduce transportation lead times; (ii) improve on-time delivery ratio; (iii) avoid useless travelled distance; (iv) limit waste of goods due to bad transportation; (v) improve carriers' profitability; (vi) allow transportation multimodality; (vii) allow real time tracking of goods; (viii) optimize transportation costs; and (ix) reduce carbon footprint impact of transportation.

3 Proposal

Basically, the core contribution of this paper consists in designing both functional and technical architectures able to support the ambition of the whole research project described previously. Obviously, this is only a necessary but not sufficient step towards what will be the major contribution of this research work, i.e., a set of recommendations for significantly improving the logistics performance of developing countries by making better use of existing infrastructure and resources.

3.1 Functional architecture

To reach the goal of the current research, a nine-step methodology has been set up as a functional architecture, as shown in Figure 1. Inspired by the approach developed by (Vernadat, 2004) for supporting enterprise re-engineering and improvement processes, this methodology consists in four main phases which are: Conceptualize, Model, Experiment and Assess.

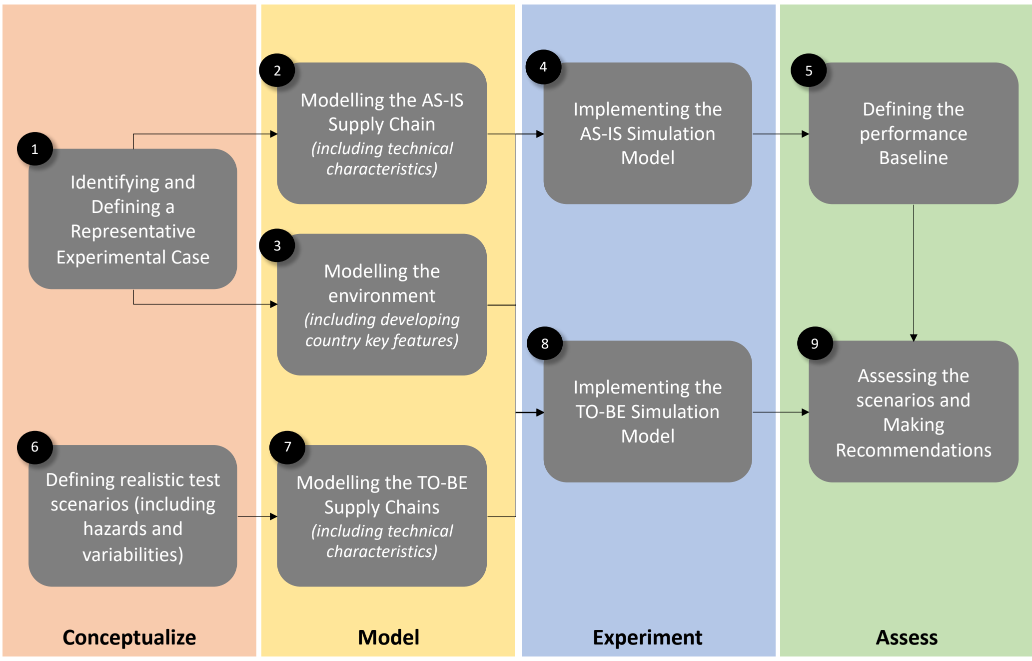


Figure 1: Functional architecture

These phases embedded the following steps:

- Step1: As our approach is field-oriented, this first step consists in selecting a concrete application case which is representative of the studied problem. In our context, the case should be a supply chain located in a developing country and having a high number of urban shipments to manage.

- Step 2: This step allows building a digital model of the current supply chain which includes the main features of the selected application case. This model would be able to simulate the AS-IS behaviors of the selected case.
- Step 3: This step consists in modelling the environment, and notably the urban characteristics, in which the select supply chain evolves. This model must be able to consider the specific developing countries' features (traffic congestion…).
- Step 4: At this stage, the two previous models are joint in a dynamic way to support simulations of the selected case behavior. Discrete event, multi-agent or systems dynamics models might be used to support this step.
- Step 5: This step consists in measuring objectively the performance of the selected supply chain in order to get a baseline. This baseline will be compared to other organizational alternatives, notably the ones which are based on PI.
- Step 6: From this step, the functional architecture moves to the definition, the modelling and the assessment of possible TO-BE organizations. Particularly, this step consists in designing a set of logistics scenarios likely to improve the overall supply chain performance.
- Step 7: As for step 2, this step supports the modelling part of the potential scenarios previously defined.
- Step 8: As for step 4, this step consists in simulating the scenarios in the context of the environment defined and modelled in step 3.
- Step 9: Finally, this step allows measuring the performance of each scenario defined in step 6 and comparing their results with the baseline calculated in step 5 to make conclusions and deliver some recommendations to practitioners.

3.2 Technical architecture

To the support the previous functional architecture, we have developed the technical architecture presented on Figure 2. First, this architecture supports the gathering of both qualitative and quantitative data from the teams and the legacy systems of the selected company. Additional open data bases are also considered to feed efficiently the step 3 of the functional architecture (particularly regarding the city mapping, the traffic mapping, etc.

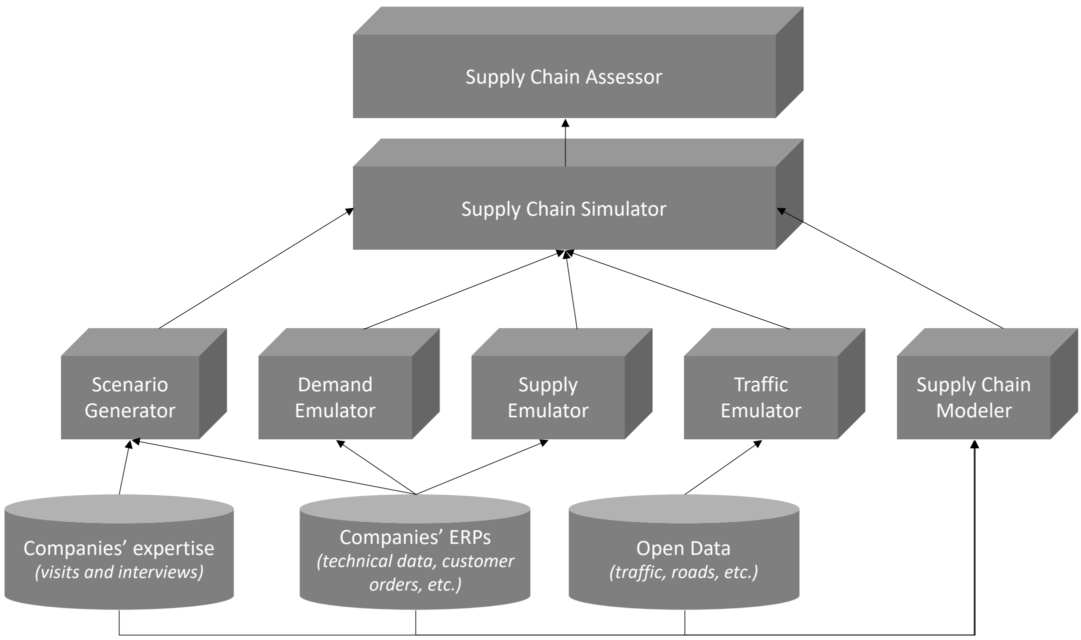


Figure 2: Technical architecture

From those data bases, several technical components have been built based on AnyLogic® simulation software, enriched with some Excel sheets and dedicated Python codes:

- A set of emulators have been set up to feed the simulation model according to the AS-IS and TO-BE scenarios. These emulators allow simulating the behaviors of respectively, the suppliers, the customers and the external traffic in the city.
- A module dedicated to the scenarios' definition has been developed to allow definition in AnyLogic® language of both AS-IS and TO-BE options.
- The core of this architecture is composed of a Supply Chain modeler and a Supply Chain simulator using a hybrid simulation model (multi-agent and discrete event simulation mainly) to support test-runs and assessments of the defined scenarios. This part of the technical architecture is developed on AnyLogic®.
- Last but not least, a dedicated dashboard including Key Performance Indicators such as logistics costs, On-Time delivery In Full, lead times, occupancy rates of used vehicles, carbon footprint, etc. has been set up to compare objectively all the scenarios and to support recommendations.

4 Ongoing application case

4.1 Step 1

The case study of this research is about an e-commerce company (mainly focused on fashion items), located in Phnom Penh city, typical of a developing country import business. Their consumers use a dedicated website to order and are supposed to be delivered within one day at a maximum if the product is in stock, and within few days or weeks if the product is out-stock. The turnover of the company is growing very fast (also typical from a developing country) and according to the CEO, the main challenge of such a company is to manage properly the logistics issues, and particularly the last-mile deliveries. He said that it is very hard to reach the target of logistics services such as same day delivery guarantees, hassle-free delivery and the right products at the right time at the right place and in the exact amount.

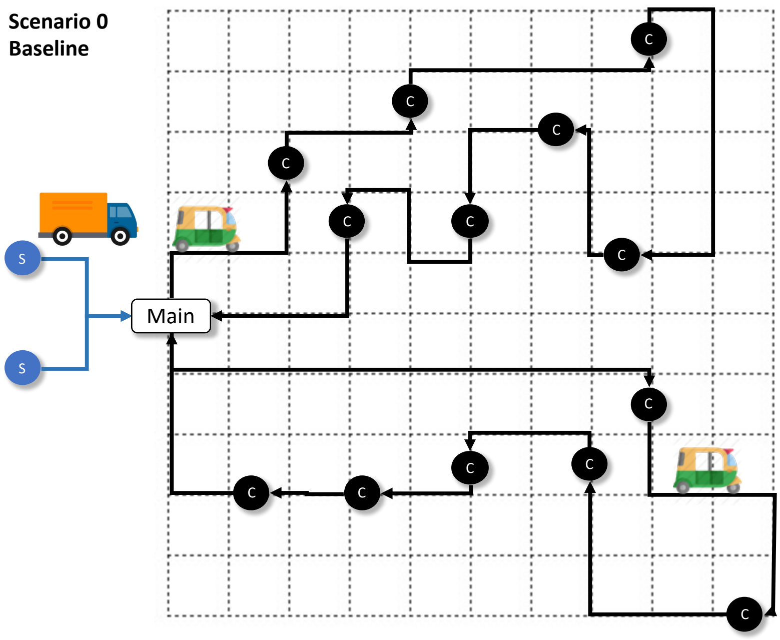


Figure 3: AS-IS Supply Chain scenario

4.2 Step 2

The company manages several thousands of items, both in-stock and out-stock, and its main suppliers are located in China, Thailand and Viet Nam. The company operates through one main warehouse and a small outlying one (not considered in this paper), both located in the city center and ownership of the company. The last-mile delivery is managed through a dedicated fleet (ownership of the company) of vehicles use through basic routings. Although, the company is thinking about a potential evolution of its network, its current supply chain is outlined on Figure 3. If the main activity concerns the management of its own products, the company also offers logistics services to some other companies (particularly regarding transportation). However, this last point is not in the scope of the study yet. The current order fulfillment process is described on Figure 4. We must notice that the IT systems use for managing the information and material flows is quite mature, and a bar-code tracking system is running at each step of the process.

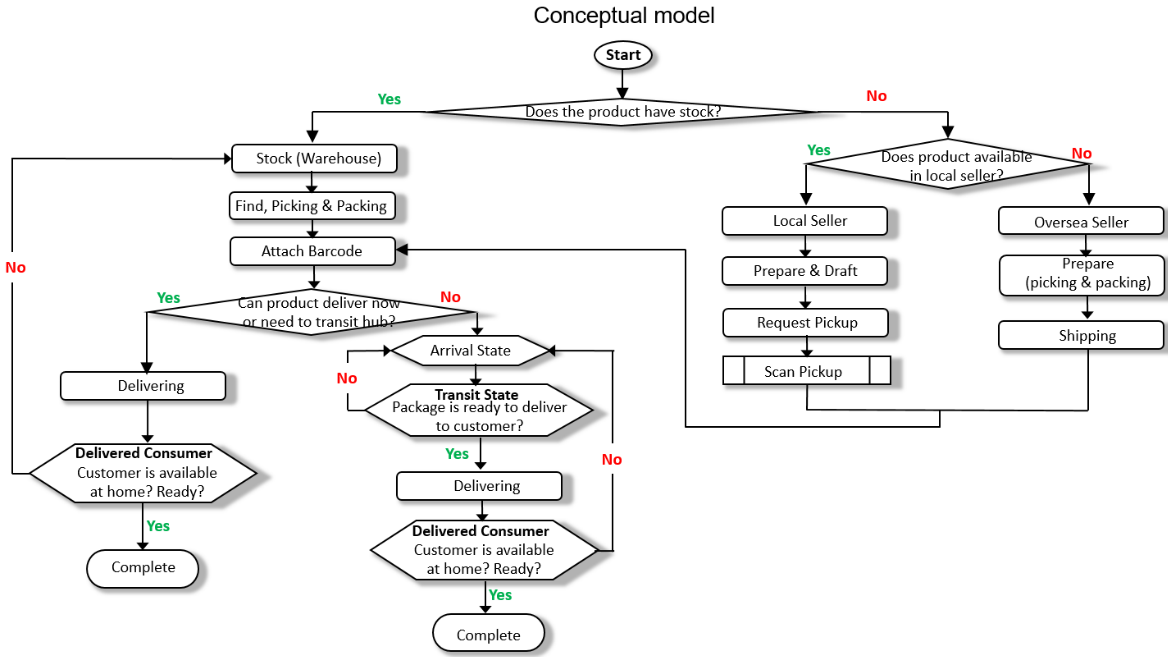


Figure 4: Conceptual model of the order fulfillment process.

4.3 Step 3

A case study by (Seiya Matsuoka, 2018) on traffic management on Phnom Penh clearly stated that traffic congestion and accidents are major issues in Phnom Penh. As a consequence, the speed of vehicles is hazardous and globally slow. Most of the time, companies (like ours) use their own fleet of vehicles for business operations. In addition, there is a huge number of individual transport businesses for passengers and/or goods (motorcycle taxi, tuk-tuk, Remork, Bajaj, etc.) These vehicles are moving randomly in the city to catch up passengers from one place to another. Most of the time, they are moving empty... (Seiya Matsuoka, 2018) also notes that the speed of vehicles is inversely proportional to their capacity, so motorbikes move faster than trucks or tuk-tuks. Basically, these features are the ones which have been implemented in our model by considering different speeds depending on the location, the time and the type of vehicles, by adding variabilities and hazards in the traffic, and by including external vehicles in the model that might be considered as potentialities for the business of the company. A last characteristic has been considered in the model. That’s the way of life of Cambodian people who lives outside most of the time during the day. The consequence is a very high number of no-show. This has been also included in the model.

4.4 Step 4

This work has been set up through a simulation model using AnyLogic® package. The model uses both discrete-event and agent-based modelling methodologies for simulating the environment and last-mile delivery process. Mainly, this scenario #0 (baseline) follows the process describes on Figure 4 and is illustrated on the left part of Figure 5.

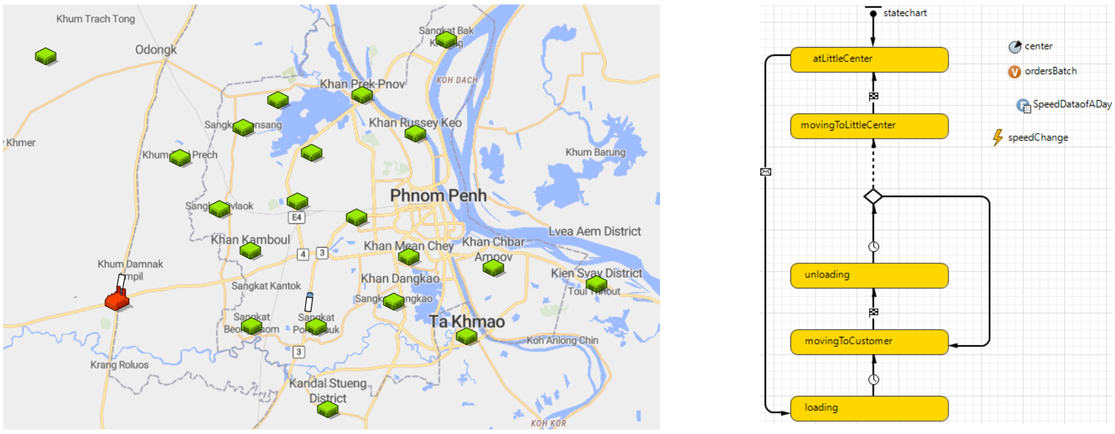


Figure 5: Illustration of simulations made on AnyLogic

In this first instance, in line with the field-reality, last-mile routes are ensured by truck, vans or tuk-tuk as shown on Figure 3. We made the assumption that all products are available in stock as we focused our analysis on the last-mile delivery issue. The transportation capacity is defined as 2 vehicles for 20 customers. All vehicles are attached to the main warehouse.

4.5 Step 5

Resulting from the previous model, we were able to establish a baseline on a set of Key Performance Indicators (KPIs). Mainly, we analyzed the situation regarding the lead-times, costs, no-shows, carbon-footprint and occupancy rate of vehicles. At this stage, as we did not have any reference, it was not possible to say if the KPIs were good enough or not.

4.6 Step 6

From this step, we defined a set of simple but PI-oriented scenarios in order to assess the potential improvements we can expect in such a context. The scenario #1 includes 4 PI-hubs in addition to the main warehouse (see Figure 6 on the left). Each PI-hub is located at the border of a part of the city and is used to manage all the customer orders of this area. Smaller vehicles (motorbikes) are now attached to these PI-hubs and trucks and tuk-tuks are only used to send massive group of orders from the main warehouse to the PI-hubs. Nothing is changed regarding the supply, the customers’ and traffic behaviors.

A scenario #2 (see Figure 6 on the right) is then defined which is almost the same than the scenario #1, except that here we do not consider an own fleet of vehicles for the company. Vehicles are not attached to the PI-hubs, they are chartered “on demand” depending on opportunities around the PI-hub when there is a need (like Uber).

A scenario #3 (not illustrated in this paper) consists in considering that two companies having two different main hubs but sharing the PI-hubs previously defined. Although these companies are on different businesses, have different products and different customers, they share facilities and process regarding the last-mile delivery.

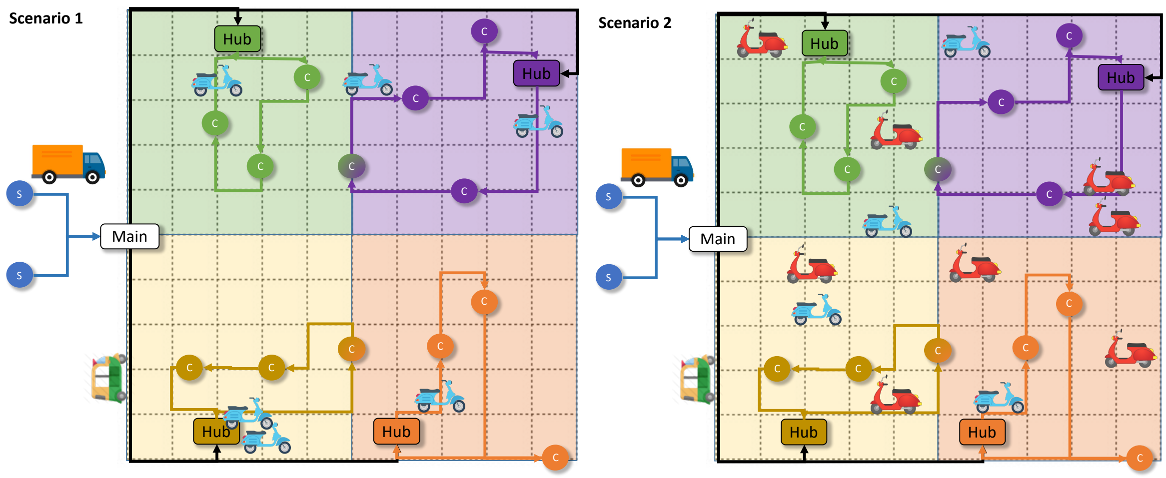


Figure 6: PI-oriented scenario #1 and #2

4.7 Step 7

In this step, the same modelling techniques than the ones discussed in Step 2 have been used.

4.8 Step 8

In this step, the same modelling techniques than the ones discussed in Step 4 have been used.

4.9 Step 9

Even if the experiments are still ongoing, first results start to appear thanks to the proposed methodology and simulation tools. Figure 7 for instance shows first results in matter of lead-time between scenarios #0 (baseline) and #1. It appears, that the introduction of PI-hubs in the context of the studied company, in Phnom Penh city speed up drastically the deliveries. Our experiments seem to indicate the same trend (i.e., better results with the PI scenario compared to the baseline) in matter of logistics costs, carbon footprint and vehicle occupancy rate. Obviously, additional studies remain necessary to analyze these very first results, particularly regarding the validity of the assumptions, the size of the model (coverage) and the sensitivity of the model.

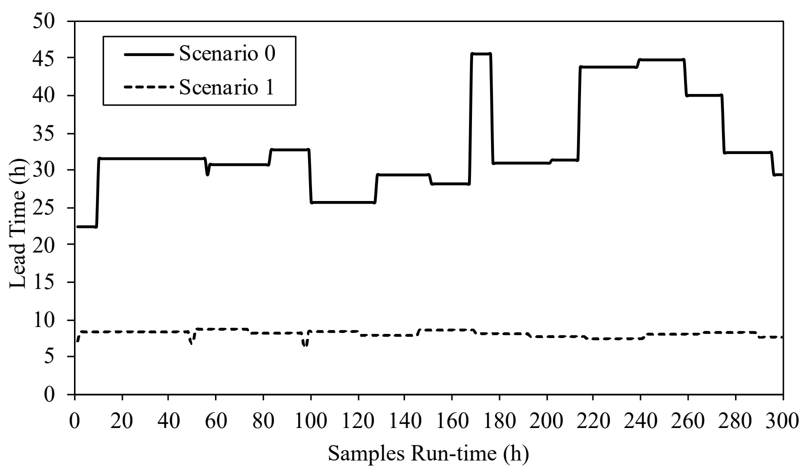


Figure 7: Lead-time comparison of scenarios #0 (baseline) and #1

5 Conclusions and further research

In this study, we addressed the question of the relevance of using the PI approach in a developing country context. In particular, we studied the question of how to improve the performance of last-mile delivery in such a context. Our ambition is to demonstrate that PI might be a great alternative for developing countries allowing them to have good logistics performance despite of poor infrastructures and means. For the moment, the proposal only consists in a concrete methodology composed of functional and technical architectures dedicated to this issue on one hand, and on a real ongoing illustrative case on the other hand. While the research work is in its infancy, first promising results appear and seem to demonstrate the relevance of using PI in developing countries logistics problems. Obviously, significative additional works remain to do to confirm these first conclusions and finally produce concrete recommendations for both scholars and practitioners. Rapidly expanding the proposed models to match much better the complexity of the field seems a necessity (more customers, more products, etc.). Also, including other dimensions of the studied supply chains such as supply is a room for improvement. Last but not least, enriching the environment simulation to better consider developing countries' features (as risks or opportunities) seem to be an essential perspective.

References

Cervero, R. (2013). Linking urban transport and land use in developing countries. *Journal of transport and land use*, *6*(1), 7-24.

Neeraja, B., Mehta, M., & Chandani, A. (2014). Supply chain and logistics for the present day business. *Procedia economics and finance*, *11*, 665-675.

Jardas, M., Dundović, Č., & Badurina-Tomić, P. (2016). Supply chain–a key factor of the sustainable development of city centers. Pomorstvo, 30(1), 45-50.

Montreuil, B., Rougès, J. F., Cimon, Y., & Poulin, D. (2012). The physical internet and business model innovation. *Technology Innovation Management Review*, *2*(6).

Montreuil, B., Meller, R. D., & Ballot, E. (2010). Towards a Physical Internet: the impact on logistics facilities and material handling systems design and innovation.

Sallez, Y., Pan, S., Montreuil, B., Berger, T., & Ballot, E. (2016). On the activeness of intelligent Physical Internet containers. *Computers in Industry*, *81*, 96-104.

Ballot, E., Montreuil, B., & Meller, R. (2014). *The physical internet*. La Documentation Française.

Grest, M., Lauras, M., & Montreuil, B. (2019, September). Toward Humanitarian Supply Chains Enhancement by using Physical Internet Principles. In *2019 International Conference on Industrial Engineering and Systems Management (IESM)* (pp. 1-6). IEEE.

Ban, S., Lauras, M., & Srang, S. (2020, November). Toward Physical Internet-Enabled Supply Chain and Logistics Networks in Developing Countries. In *Working Conference on Virtual Enterprises* (pp. 379-389). Springer, Cham.

Vernadat, F. B. (2004, June). Enterprise Modelling: Objectives, Constructs and Ontologies. In *CAiSE Workshops (3)* (p. 99).

Guimarães, Vanessa de Almeida, Gustav Carl Skroder, Glaydston Mattos Ribeiro, and Pedro Henrique González-Silva (2020). « Strategic planning of freight transportation to support smart cities design: The Brazilian soybean case ». Revista Facultad de Ingeniería Universidad de Antioquia.

Reda, A. K., Gebresenbet, G., Tavasszy, L., & Ljungberg, D. (2020). Identification of the regional and economic contexts of sustainable urban logistics policies. *Sustainability*, *12*(20), 8322.

Kraay, A. (2018). Methodology for a World Bank human capital index. *World Bank Policy Research Working Paper*, (8593).

Landschützer, C., Ehrentraut, F., & Jodin, D. (2015). Containers for the Physical Internet: requirements and engineering design related to FMCG logistics. *Logistics Research*, *8*(1), 1-22.

Tang, K. B., & Bundhoo, D. (2017). Foreign aid and economic growth in developing countries: Evidence from Sub-Saharan Africa. *Theoretical Economics Letters*, *7*(05), 1473.

Pucher, J., Korattyswaropam, N., Mittal, N., & Ittyerah, N. (2005). Urban transport crisis in India. *Transport Policy*, *12*(3), 185-198

Yang, N., & Wang, Y. (2019). Identify silent data corruption vulnerable instructions using SVM. *IEEE Access*, *7*, 40210-40219.

Introducing Services and Protocols for Inter-Hub Transportation in the Physical Internet

Sahrish Jaleel Shaikh[1,2,3], Benoit Montreuil [1,2,3] , Moussa Hodjat-Shamami[1,2,3], Ashish Gupta[1,2,3]

1. Physical Internet Center
2. Supply Chain and Logistics Institute
3. H.Milton Stewart School of Industrial & Systems Engineering,
Georgia Institute of Technology, Atlanta, U.S.A

Corresponding author: sahrish.shaikh@gatech.edu

Keywords: *Logistics Infrastructure, Mesh Network, Physical Internet, Routing Protocols, Network Architecture, Hyperconnected Service Network*

Abstract

The Physical Internet (PI) puts high emphasis on enabling logistics to reliably perform at the speed mandated by and promised to customers, and to do so efficiently and sustainably. To do so, goods to be moved are encapsulated in modular containers and these are flowed from hub to hub in relay mode. At each hub, PI enables fast and efficient dynamic consolidation of sets of containers to be shipped together to next hubs. Each consolidated set is assigned to an appropriate vehicle so to enact the targeted transport. In this paper, we address the case where transportation service providers are available to provide vehicles and trailers of distinct dimensions on demand according to openly agreed and/or contracted terms. We describe the essence of such terms, notably relative to expected frequency distribution of transport requests, and expectations about time between request and arrival at hub. In such a context, we introduce rigorous generic protocols that can be applied at each hub so as to dynamically generate consolidation sets of modular containers and requests for on-demand transportation services, in an efficient, resilient, and sustainable way ensuring reliable pickup and delivery within the promised time windows. We demonstrate the performance of such protocols using a simulation-based experiment for a national intercity express parcel logistic network. We finally provide conclusive remarks and promising avenues for field implementation and further research.

1. Introduction

The Physical Internet (PI, π) has been introduced with the specific purpose of increasing the efficiency and sustainability of logistics systems [1]. It uses the Digital Internet as a metaphor for inspiring, steering, and guiding transformations in designing, managing, and operating the supply, flow, and movement of physical objects. The Physical Internet is an open, global, hyperconnected and sustainable logistics system [2]. PI encapsulates goods in standard modular packaging, handling and transport containers. These PI containers move though distributed, multimodal transportation networks where containers from different origins are aggregated at logistic hubs depending on the next site they will travel to [3]. Open logistics facilities such as open semi-trailer transit centers and open crossdocking and consolidation hubs, as well as open warehousing, distribution, and fulfillment centers, are part of the interconnected networks, enabling a global Logistics Web [4,5]. The open logistics web is characterized by numerous logistics, handling, and buffering actors and infrastructure that are available on demand [6,7]. In this paper, we focus on hyperconnected transportation in the Physical Internet, moving modular containers from their source to through destination, through multi-segment inter-hub routes through the open mobility web.

Consider long-haul transportation as an example. Currently, long-haul drivers usually have generally a driving radius of 250 miles or more, with a maximum on the order of 11-hour trucking time per day, with variations from state to state. Going longer distances means long-haul truckers work throughout week and come home on weekends or be on the road for weeks or even months at a time. Transportation providers often have long-term contracts with shippers based on the routes their truckers will take and a fixed rate per mile. With a multi-segment network and the open mobility web such as that in the Physical Internet, the contractual terms between the shippers and transportation providers are to change drastically. On-demand trucking services have already hit the market with companies such as Convoy [8] and Uber Freight [9], offering transportation services coupled with tracking and traffic information. These are based on a pool of independent truck drivers who have registered themselves with such platforms and are able to look for shipments ready to be loaded near their vicinity. In this paper, we propose contracts between the shipper and the transportation service providers that complement such on-demand services within the Physical Internet.

Similar as in the Digital Internet, a central planning system will not be present in the Physical internet to determine how each package will move through the system. Rather, the Physical Internet can be established on a series of globally-informed and locally-applied protocols that will dynamically determine the next steps to take. The first aim of this paper is to define a set of algorithm-based transportation protocols for dynamically generating sets of modular containers to be shipped together and placing on-demand transportation service requests. These protocols are conceived for transportation in a network of open and interconnected networks where transportation service providers are available to provide vehicles of distinct dimensions on demand according to openly agreed and/or contracted terms. The second aim of the paper is to provide a simulation-based performance assessment of these protocols.

The rest of the paper is organized as follows. Section 2 focuses on the on-demand service contracts and the associated basic expectations. The design elements of the proposed protocols are described in Section 3 followed by the protocols in Section 4. The paper concludes in Section 5 by summarizing the design and motivation of the model and highlighting further research opportunities.

2. On-Demand Service Contracts

Freight rates are often broken down into two different categories: contractual rates and spot rates. Shippers usually sign contracts well in advance with a guaranteed volume and rates [10]. There are situations, such as inconsistent freight volumes, and seasonal or one-off shipments, in which shippers will opt for a spot rate instead [11]. However, spot rates are volatile and change with demand, and are generally significantly higher than contractual rates [10,11]. In our model, transportation service providers are available to provide vehicles and trailers of distinct dimensions, on demand according to a contracted expected frequency distribution of transport requests and order-to-arrival time of the trucks. When a package enters the system, its origin, destination, and delivery time promise are known. Based on the origin and destination of packages and modular containers, the total volume can be estimated for each route segment. The expected frequency distribution is then evaluated based on the container volume and the respective service levels expected on each segment. Important in an on-demand transportation contract is the time between the request of a vehicle and its arrival at the hub. This element is crucial as delay in the vehicle arrival can impact on-time package delivery performance. Although these times remain stochastic, the expected and latest time of arrival is included in the contract to ensure that the service level is not compromised. In the next sections, we discuss the protocol of information sharing which enables to reduce the uncertainty that revolves around expected time of arrival of vehicles.

3. Protocol Design Elements

In this section, we discuss the design elements involved in developing the algorithm-based protocols that dynamically generate consolidation sets of modular containers and requests for on-demand transportation services at the hubs. These elements are calculated at the beginning of a package journey and are subject to frequent update.

3.1. Assigned Dwell Time

The protocols consider three important location and time elements of a package; origin, destination and promised delivery time window. When a package enters the system, a route is assigned to it, which is a path from the origin to the destination, through a set of intermediary hubs. This route may be dynamically altered to deal with disruptions, yet important is maintaining a package-route duo. Based on its route, slack time is the surplus time a package has after accounting for the travel time from origin to destination and the processing time at the intermediary hubs. This excess time is distributed amongst the intermediary hubs based on the expected and relative capability and performance of the hub, and the time thus allocated is called the assigned dwell time of the package at the hub. At every intermediary node, we can calculate the latest departure time with respect to the promised delivery time and planned package route. The latest departure time of package p at the first hub then is:

$$T^{ld}{}_{p,1} = S^{a}{}_{p,1} + P^{p}{}_{1} + D_{p,1} \tag{1}$$

For every other hub in the path,

$$T^{ld}{}_{p,n} = S^{a}{}_{p,n} + P^{p}{}_{n} + D_{p,n} + \delta \tag{2}$$

$$\delta = (T^{sd}{}_{p,n-1} - T^{ld}{}_{p,n-1}) \times f_n \tag{3}$$

Where $T^{ld}{}_{p,n}$ is the latest departure time of package p from facility n, $S^{a}{}_{p,n}$ is the time of the arrival scan of package p at facility n, $D_{p,n}$ is the assigned dwell time of the package p at facility n, $T^{sd}{}_{p,n}$ is the time of the departure scan of package p at facility n, and f_n is the relative package flow of facility n. This set of formulas ensures that the latest departure time is adjusted based on the dwell time of the package in earlier hubs. If a package departed the facility earlier than its latest departure time, then it has some additional time that it can use at the next hubs. This can be also considered as an update to the assigned dwell time where we would add δ to the assigned dwell time of all remaining hubs. When calculating the assigned dwell time of a modular container, we consider all the packages in that container. The assigned dwell time of a container is the minimum of all assigned dwell times of the contents of the container. Let $\boldsymbol{P}_c$ be the set of packages in a container c at hub n, then:

$$D_{c,n} = \min_{p \in \boldsymbol{P}_c}(D_{p,n}) \tag{4}$$

3.2. Package/Container Signaling

We use ***package/container signaling*** where the dwell time of the packages/containers is used to request the on-demand trucks. When the assigned dwell time of a package is being approached, it is flagged as an urgent one. To ensure that there is enough time to react, the contractual terms with the service provider will define an expected order-to-arrival time of the vehicle. This input considers the time it takes a truck to arrive at a facility after it has been requested. Once the urgent packages/containers are flagged, we can check the estimated time of arrival (ETA) of the packages/containers that are enroute to that facility and have the same destination as the urgent packages/containers. If the arrival time and the sorting time of incoming packages/containers is less than the remaining dwell time of current urgent packages, it means that they can be sent in the same truck. Once the truck arrives, the loading of urgent packages/containers is prioritized, and to aim to fill the truck reasonably and not below a certain threshold, the truck may be filled with all other non-urgent packages/containers destined for the same segment.

3.3. Information Sharing

To be well positioned to make operational-level decisions at every hub, it is crucial that pertinent data flows seamlessly across the hubs. As soon as a package is assigned a path from its origin to its destination, all the hubs that the package is planned to visit are notified and provided with the estimated time of arrival based on the path, travel times, processing times and the assigned dwell time. This information is updated when a package arrives at a hub, if it is encapsulated into a modular container, when it departs from a hub, and when any significant disruption is registered. The updated assignment of dwell time notably accounts for any instance where the package/container may have left earlier or later than planned. The hubs are also notified of the containerization status of the package, that is, whether it is travelling independently or has been consolidated with other packages into a modular container. We propose to reduce the uncertainty involved in the order-to-arrival time of the vehicles by obtaining live data on the expected arrival time of vehicles after a request has been made to the contracted vendors for each segment.

3.4. Maximum Latency

It may be possible to consolidate a set A of modular containers at a hub with a set B of incoming containers, yet this may require containers in set A to stay longer at the hub than their assigned dwell time. Which may lead to package delivery lateness. In order to account for this flexibility while avoiding lateness, we introduce a flexibility margin for the assigned dwell time that a container c can spend at a facility n, called Maximum Latency ($L^M_{c,n}$):

$$L^M_{c,n} = {T^{ld}}_{c,n} + m\left(\sum_{i\in \boldsymbol{H_c}} D_{c,i}\right) \tag{5}$$

where ${T^{ld}}_{c,n}$ is the latest departure time of container c at facility n, based on the pre-calculated dwell time, $D_{c,i}$ is the dwell time assigned to a container c at each hub i element of $\boldsymbol{H_c}$, the set of hubs remaining in the assigned path of container c. Parameter m is an input between 0 and 1. Depending on the network and average number of hubs in a path, it can be altered. When $m = 0$, both the Maximum Latency and the latest departure time are same, and when $m = 1$, all dwell times are pulled to the first (current) hub. This means the package/container will have a lot of flexibility at the beginning but might be rushed at the end to meet the service level promise. As per the information sharing agreement, the hubs will receive information on the estimated time of arrival of all incoming packages; subsequently, the hubs will be able to identify any package arriving during this flexibility window.

4. Protocols

The protocols assume that trucking transportation service providers are available to provide vehicles and trailers of distinct dimensions on demand and that there is an established framework for sharing information. Hubs request the trucks based on the volume of containers in the hub and the incoming containers to the hub. Modular containers are flagged as urgent if they are approaching their assigned dwell time while present at a hub or if they are still en route to the hub but need to be moved as soon as they reach the hub. Hubs use these signals to request on-demand trucks, and to ensure that there is enough time to react we incorporate the contracted maximal order-to-arrival time of the vehicle into the signaling time. Therefore, if there are containers en route to a hub that require to be moved urgently, a vehicle order may be placed for them even before they arrive to the hub depending upon the order-to-arrival time of the vehicle. As the hub already has data on estimated time of arrival of the incoming containers as well as their assigned dwell time, it is able to re-evaluate the urgent containers and re-compute the queue of containers in terms of urgency. It is possible that some containers that are en route to the hub may be more urgent than the ones already present at the hub.

We hereafter discuss two protocols for creating container sets and placing vehicle requests. The first, called the Local Latency (LLT) Protocol with fixed dwell time, only considers the containers that are present at a hub and the containers that may arrive until the vehicle arrives. The second protocol, called the Maximum Latency (MLT) Protocol with flexible dwell time, proposes a flexibility margin to relax the assigned dwell time while respecting the promised delivery time window, so it also considers containers that arrive in that additional time.

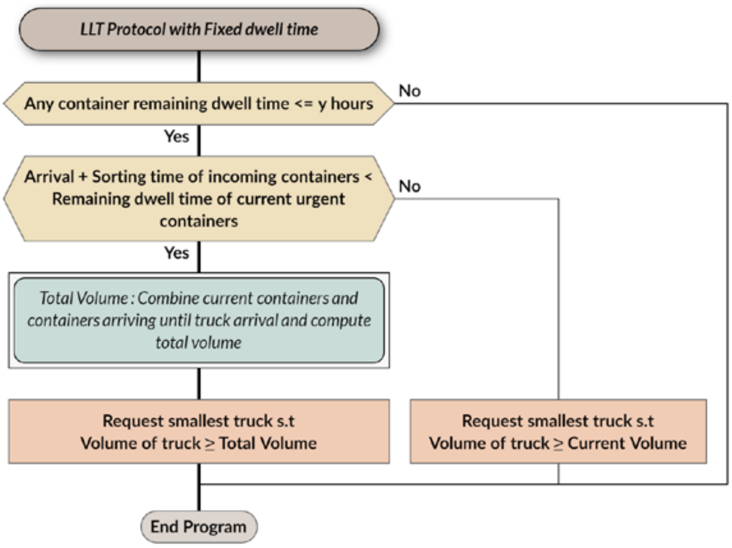


Figure 1:Algorithm Flow Chart for Local Latency Protocol with Fixed dwell time

4.1. Local Latency Protocol (LLT)

In this protocol, once the containers are flagged as urgent, the hubs verify the estimated time of arrival of incoming containers and their next destination. If there are containers that can arrive and be ready to ship within the remaining dwell time of current urgent packages, they are consolidated with the urgent packages. The facility will then place a request for the smallest number of on-demand truck(s) such that all packages can be accommodated. Once a truck arrives, it is loaded with the urgent containers, and to ensure a high fill-rate, the truck is filled with all other containers that are destined for the same segment but may not be urgent. In this paper we will use the terms trucks and trailers interchangeably. As the trailers can be of difference sizes, a truck can either be just a truck or a tractor pulling a trailer of a specific size. This algorithm ensures that the on-demand trucks are not overused, and we are able to send a lower number of larger trailers instead of higher number of small ones. The protocol is synthesized in Figure 1.

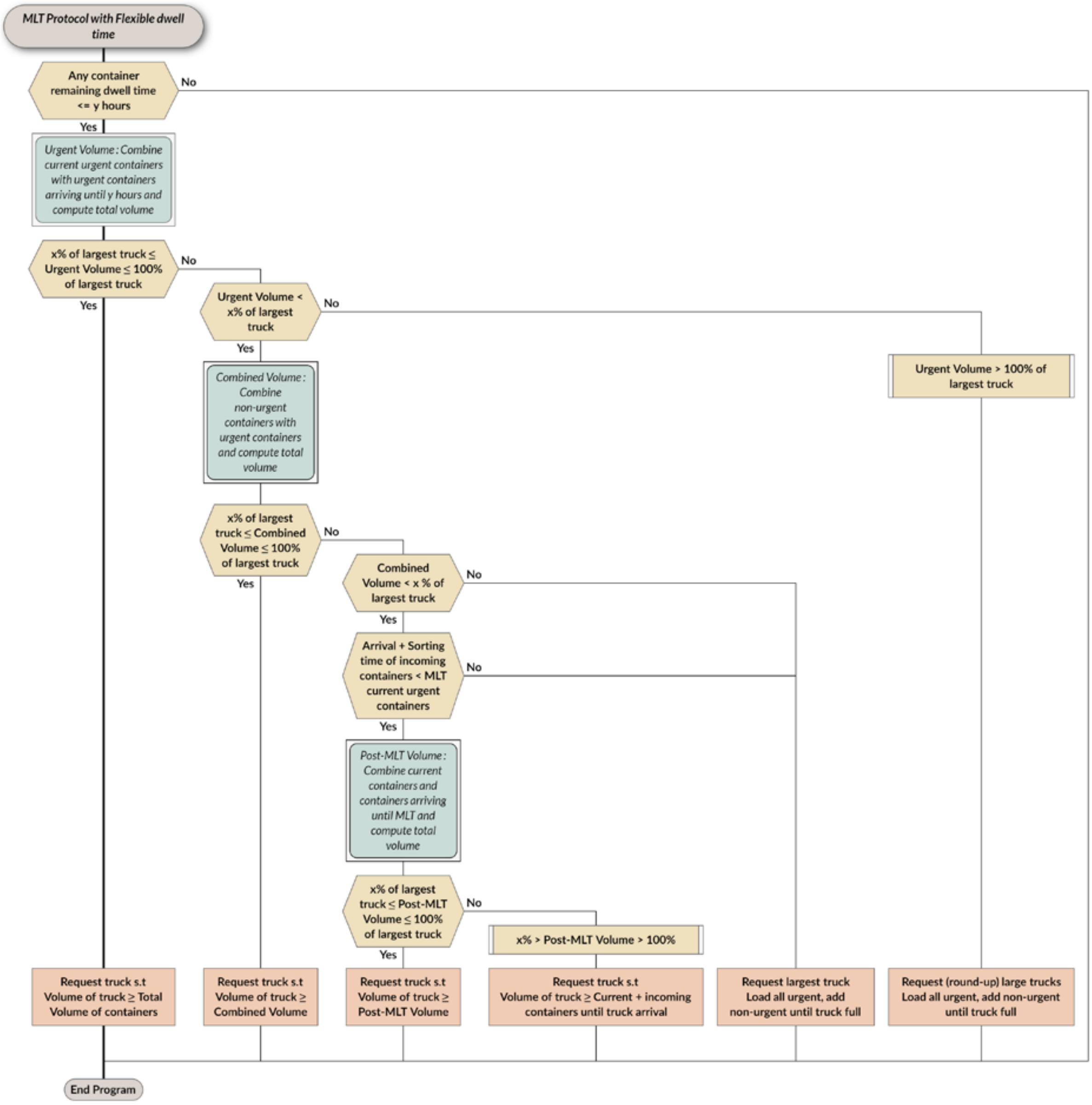


Figure 2: Algorithm Flow Chart for MLT Protocol with Flexible dwell time

4.2. Maximum Latency Protocol (MLT)

In this protocol, we introduce a flexibility margin using maximum latency. This allows the packages to wait longer at some hubs to encourage consolidation while still meeting the overall service level requirement. The hubs evaluate the volume of urgent containers currently present as well as the containers arriving within the order-to-arrival time of the vehicle. If there is not enough load to fill a large truck, the hubs check whether more containers will arrive during the Maximum Latency (MLT) of the current containers. The MLT uses a parameter *m* between 0 and 1, calibrated to robustly ensure that the packages in the containers are to be delivered within their promised time window.

In the MLT protocol synthesized in Figure 2, input parameters x and y respectively refer to the minimum acceptable fill rate of a vehicle and the order-to-arrival time of a truck. In the protocol, on-demand vehicle order request is triggered as soon as the latest departure time of any package present at the hub or en route to a hub approaches the contracted expected order-to-arrival time of the vehicle. Once this protocol's algorithm is triggered, we check the volume of urgent containers available at the facility and all that may arrive and be ready to be shipped within y hours. If the volume of these urgent containers is higher than x, a truck is requested, and the containers are shipped. In the case where the total volume is greater than 100%, we order the minimum number of trucks required to ensure that all urgent containers move. To improve the vehicle fill rate, we fill the remaining capacity of the truck with non-urgent containers such that the ones with lower remaining dwell time amongst others are moved first. The third case is where the truck fill rate with urgent containers is below the minimum threshold. In that case, the first step is to add all non-urgent containers to evaluate if they can fill the trucks. If, after combining the two types of containers, the volume has reached the minimum threshold, we request a truck and ship the containers.

In the case where after combining non-urgent containers, the volume is more than the truck capacity, then it is already known that the volume of urgent containers was less than $x\%$, so we do not load all non-urgent containers but rather just those which can help improve the fill rate and have lower remaining dwell times than other containers. The last case is where after combining all urgent and non-urgent containers, the total volume is less than the minimum threshold. This can occur in segments with low frequency.

To check for any consolidation opportunities in this case, we use the Maximum latency concept discussed above. We check if there are any containers that may arrive if we make the current containers wait a little longer. In case, even after waiting, the expected fill rate is still under the minimum threshold, we do not wait until the MLT and request the truck immediately. However, if by including containers arriving during this flexibility window can enhance the vehicle fill rate to be above the minimum threshold, we make the current containers wait and consolidate them with the containers that arrive. The last case then is when the fill rate exceeds 100% after including the containers arriving by the MLT cutoff. If that happens, we know that there are a lot of containers arriving and those containers will suffice to request and fill a truck with satisfactory fill rate. So, we do not make the current containers wait and rather request a truck immediately to ship them.

The aim of this algorithmic protocol is to have larger trailers moving less often rather than smaller ones. If the MLT does not affect consolidation, we do not make the current containers wait to avoid any congestion and compromise on service level promises. However, if during the MLT, more consolidation opportunities are foreseen, the containers wait, and the remaining dwell times are updated for all remaining hubs for those containers. There is always a trade-off between cost and service. This protocol's algorithm can be modified by changing some parameters in order to favor cost or service.

5. Simulation Model

5.1. Framework

In order to assess the performance of the proposed protocols, we leverage a hyperconnected logistics simulator developed in Georgia Tech's Physical Internet Center, built in the AnyLogic platform according to an agent-oriented discrete-event simulation modelling approach.

A similar simulation program is used in [12] where the process starts with generation of customer agent either stochastically based on demand model or as deterministic historic input to the simulation. As each parcel is generated, the Inter-city agent is notified, and it makes the decision of the choice of gateway hub at origin city that the parcel needs to arrive at. The Inter-city agent decides on how to route the package as well as consolidation of packages into modular containers based on the features such as the origin, destination, direction of travel and most importantly, the committed service level or the due date of the package to the final destination [13]. The consolidation of parcels before reaching the gateway hub is done with the aim of reducing the toll on the gateway hub and facilitating the sorting/cross docking operations. Subsequently, the agent assigns the ETA, dwell times, and maximum latency for next hub and all the future hubs except the last hub. Determination of these parameters occurs on the basis of protocols discussed in the earlier sections. Intercity Router further notifies all the Gateway Hub Routers of respective Gateway Hubs in the path of incoming containers. The Gateway Hub agents manage sorting/cross docking as well as vehicle scheduling, loading/unloading operations to execute the routing decisions with the aim of minimizing operational costs within the hubs.

5.2. Results of Proposed Protocols

We perform the computational studies for the two algorithms using the current network infrastructure of a large China-based urban parcel delivery service provider. We focus the results on the South China region served by a network of 658 hubs, using a demand of approximately 5,000,000 packages over a time horizon of 10 days with two sizes of vehicles available. To measure efficiency, we track the total number of vehicles used, the total capacity of trucks that was on the road, the average fill rate of the trucks and the total cost incurred. To track service measures, we evaluate the service level of the packages, i.e. whether the packages reach the destination by the committed time window, and the average delay/earliness of the packages.

Table 1: Simulation results from LLT and MLT Algorithmic Protocols

	Number of vehicles	Total Trailer Capacity	Small Trailers	Large Trailers	Reduction in Trailers	Vehicle Fill Rates	Fill Rate Small Trailers	Fill Rate Large Trailers
LLT Protocol	11,838	30,579,600	10,269	1,569	3%	86.9%	85.0%	99.0%
MLT Protocol	11,502	33,647,600	9,083	2,419		80.4%	79.5%	83.7%
	Total Cost ($)	Reduction in Total Cost	Service Level	Average Earliness (hours)	Sensitivity (Average earliness with tighter windows)			
					4 hours	6 hours	8 hours	10 hours
LLT Protocol	11,185,170	4%	100.0%	6.8	4.4	3.2	2.5	1.9
MLT Protocol	10,764,339		100.0%	10.0	7.2	5.8	3.9	2.1

Table 1 summarizes the results from both protocols. The service metrics look similar in both models as all packages reach the destination within the committed service time. We observe that although the total induced capacity from the trucks ordered increases in the MLT protocol, the total number of vehicles is reduced. This happens as MLT encourages more consolidation, so we are able to send the same packages in a lower number of larger vehicles than more frequent smaller vehicles. It is notable that both protocols perform well and all packages are delivered within the service window. That being said, we observe that the total cost is reduced by 4% in the MLT protocol. In this instance, packages move faster in the MLT protocol with average earliness of 10 hours compared to the Local Latency Protocol at 6.8 hours. We perform a sensitivity analysis on the performance of the protocols by reducing the service time windows of packages such that the promised window for delivery tightens while maintaining the feasibility of moving a package from the origin to the destination in that window. Although we make some packages wait longer in the MLT protocol, we are able to move a lot of non-urgent packages faster when consolidating. However, with very tight service levels, the earliness converges between the two protocols as the MLT protocol faces an even tighter window towards the end of the package journey once it uses the flexibility margin at the intermediate stops. We report the corresponding sensitivity results in the table above as well.

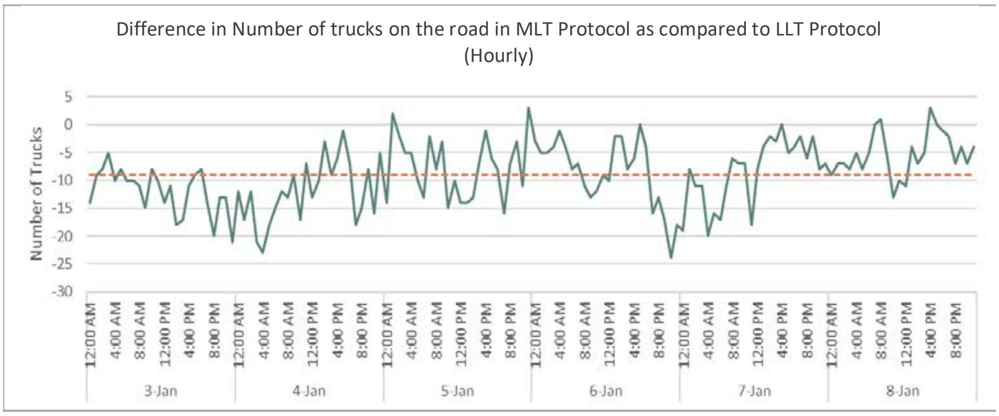


Figure 3: Difference in the active hourly number of trucks between the two protocols

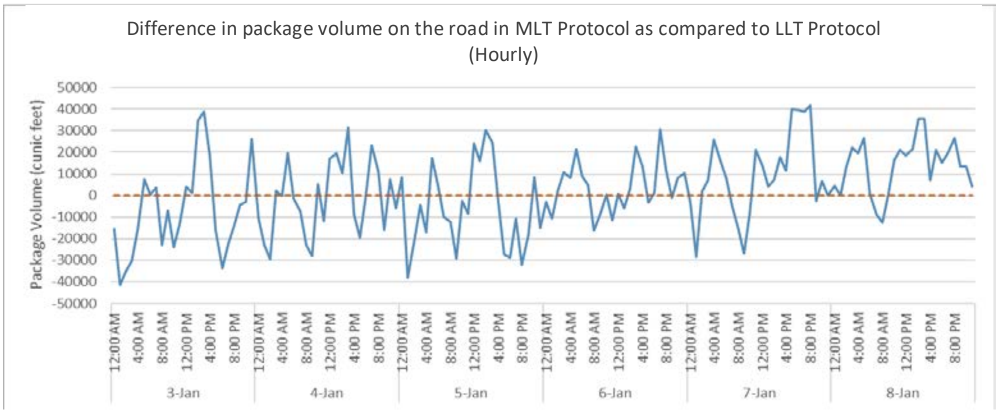


Figure 4: Difference in the active hourly package volume between the two protocols

Analyzing the two outputs further, we gain insights on the difference in the number of trucks, total trailer capacity and package volume on the road in the two algorithms as shown in Figure 3 and Figure 4. The total package volume in the MLT protocol is lower than LLT protocol before hiking up notably higher as the former one makes the packages wait at the hubs before they are sent to the next hub. We also observe that in this particular instance, although the number of vehicles in both protocols vary, on average we eliminate 9 trucks at any point in time using MLT protocol as compared to the LLT protocol.

5.3. Expected Frequency of On-Demand Vehicles

The simulation model stores and outputs the on-demand vehicle movements, identifying the origin, destination, the departure time and arrival time of each of the used vehicle along with the vehicle capacity and the capacity of the packages that travelled on that vehicle. These outputs can simply be translated into the expected frequency and sizes of vehicles required for each of the segments. In the current simulation results, we notice an overall trend for each of the segments. High flow lanes require a higher number of larger trucks while the low flow lanes have less frequent usage and mostly use smaller trucks. As there is a demand generator built within the simulation model and the package routes are an input to the simulation, these can be altered to obtain different expected frequencies in order to negotiate the on-demand contracts in a better way. The simulation model results for expected frequency will change if the demand distribution input is changed, therefore the simulation model provides an opportunity to perform scenario analysis with different distributions of demand, not just in terms on the quantity but also in terms of the origin and destination of the packages.

6. Current Contracts vs. Proposed On-Demand Models

In this section, we compare the proposed on-demand transportation model to the current model used by the national intercity express parcel logistic network. Using the historical data of the vehicle movements for the same demand distribution, we observe that the company is using dedicated fleet as well as outsourced contracts and occasionally obtains spot rates. The dedicated fleet usually travels long haul while the outsourced contracts may be longer segments but these can be one-way trips, and not necessarily a return journey. The spot rates obtained are for one-way trips as well. We compare the active trailer volume on the road in the current contracts and the proposed contracts in 5. In 6, we compare the number of trucks on the road at any hour during the considered time horizon in both models.

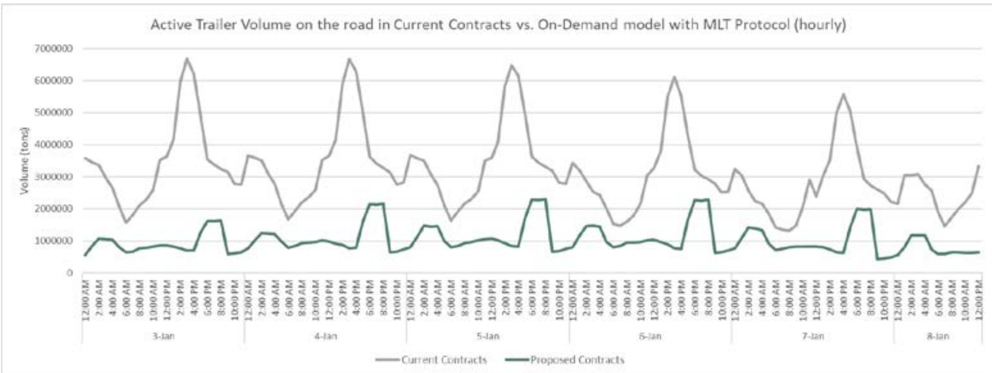


Figure 5 : Hourly Distribution of Active Trailer Volume

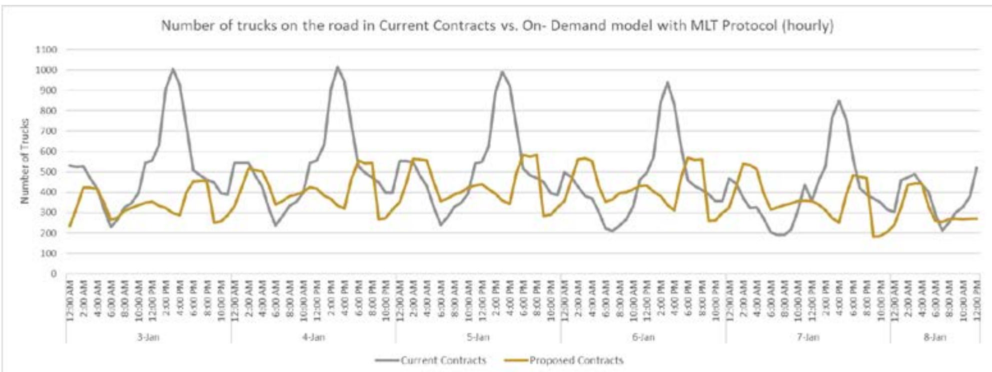


Figure 6: Hourly Distribution of Number of trucks on the road

We observe that the total number of vehicle movements has reduced in the on-demand contracts and the total trailer volume has also reduced. As the current contracts are made well in advance without information on the future behaviour of the system, there are many instances where a vehicle may only be 10% filled but still has to depart the hub to be able to reach the next hub within the scheduled time. The flexibility introduced by the on-demand contracts take advantage of the real-time data of packages such as their remaining dwell time to make an informed decision about the capacity and scheduling of the vehicle. With such a model, we are able to reduce the carbon footprint as the number of vehicle movements and total travelled miles required are reduced and are able to achieve a space-efficient solution in terms of higher vehicle fill rates while providing better work environment for drivers as well where they may not be too far from their home but rather shuttle back and forth between two locations.

7. Conclusion

The Physical Internet is an open hyperconnected logistics system that uses a set of protocols and interfaces to move physical goods contained within standard modular containers through an open mobility web. In this paper, we have introduced algorithmic protocols to create consolidated sets of modular containers and request the on-demand means to move such sets in the open mobility web, where each container encapsulates packages whose delivery is time sensitive.

We introduced two protocols that enable the scheduling of transportation requests to the contracted service providers, given the assigned dwell time and service promise for packages embedded in containers. We have provided preliminary performance assessment of the protocols using a simulation experiment on actual data of a large parcel delivery service provider. The parameters of the protocol algorithms can be altered according to the network topology as well as the flow and movement of the physical goods. As an example, some low-flow lanes can have a higher Maximum Latency Time ratio, i.e., the containers may wait longer for consolidation versus high flow lanes where the ratio can be set lower.

The Physical Internet requires an extensive set of robust protocols that can be applied to any setting. This paper has focused on one of the facets of such protocols. We believe that our introduction of the proposed protocols may open research avenues, including dynamically optimizing the algorithmic parameters , defining more setting-based parameters to improve the algorithm performance, and performing more extensive simulation-based assessments with scenarios varying notably in terms of demand mix, patterns, and uncertainty; vehicle mix; package size mix; transit time stochasticity; hub availability, capability, and capacity; and tightness of promised delivery times.

References

[1] Montreuil B. (2011). *Towards a Physical Internet: Meeting the Global Logistics Sustainability Grand Challenge*, Logistics Research, Vol. 3, No. 2-3, p. 71-87, https://doi.org/10.1007/s12159-011-0045-x.

[2] Sarraj R., E. Ballot, S. Pan & B. Montreuil (2012). *Analogies between Internet networks and logistics service networks: challenges involved in the interconnection*, Journal of Intelligent Manufacturing, www.springerlink.com/content/4402307w812271j4/

[3] Montreuil B., E. Ballot & F. Fontane (2012). *An Open Logistics Interconnection Model for the Physical Internet*, Proceedings of INCOM 2012 Symposium, Bucharest, Romania, 2012/05/23-25.

[4] Montreuil B., R.D. Meller & E. Ballot (2013). *Physical Internet Foundations*, in Service Orientation in Holonic and Multi Agent Manufacturing and Robotics, ed. T. Borangiu, A. Thomas and D. Trentesaux, Springer, p. 151-166.

[5] Montreuil B., E. Ballot, W. Tremblay (2016). *Modular Design of Physical Internet Transport, Handling and Packaging Containers*, Progress in Material Handling Research Vol. 13, Ed. J. Smith et al., MHI, Charlotte, NC, USA.

[6] Ballot, E., O. Gobet & B. Montreuil (2012). *Physical Internet enabled open hub network design for distributed networked operations*, in Service Orientation in Holonic and Multi-Agent Manufacturing Control, edited by Borangiu, T., Thomas, A., Trentesaux, D., Vol. 402, Springer, p 279-292.

[7] Crainic T. G. & B. Montreuil (2016). *Physical Internet Enabled Hyperconnected City Logistics*, Transportation Research Procedia – Tenth International Conference on City Logistics, v12, 383-398.

[8] Convoy, Future of Freight, convoy.com/

[9] Uber Freight, www.uber.com/us/en/freight

[10] Goldsby T. J., Iyengar, D., Rao S. (2014) The Definitive Guide to Transportation: Principles, Strategies, and Decisions for the Effective Flow of Goods and Services. United States, Pearson Education, 2014: pp 96-98.

[11] Roncoroni A. (2015) Handbook of Multi-Commodity Markets and Products: Structuring, Trading and Risk Management. Germany, Wiley: pp 355-360.

[12] Kaboudvand S. , Campos M. R. and Montreuil B. . (2021). Enabling Scientific Assessment of Large Scale Hyperconnected Urban Parcel Logistics: Agent-based Simulator Design. In IISE Annual Conference and Expo 2021 [974871] (IISE Annual Conference and Expo 2021). Institute of Industrial and Systems Engineers, IISE.

[13] Kaboudvand S. and Montreuil B. (2021). Data-driven Dynamic Containerized Consolidation in Hyperconnected Parcel Logistics. In IISE Annual Conference and Expo 2021 [974861] (IISE Annual Conference and Expo 2021). Institute of Industrial and Systems Engineers, IISE.

Generating clusters for urban logistics in hyperconnected networks

Cyrus Hettle[1], Louis Faugere[2], Simon Kwon[3], Swati Gupta[4], and Benoit Montreuil[5]
1, 3, 4, 5: Georgia Institute of Technology, Atlanta, USA
2: Amazon, Seattle, USA
Corresponding author: chettle@gatech.edu

Abstract: *In the hyperconnected logistics model, a city is represented as a continuous mesh of small regions called unit zones. The clustering problem is to partition the set of unit zones into larger local cells and urban areas, and is critical in defining network operations. We give a mixed integer programming-based method for solving the clustering problem, which combines aspects of graph partitioning and min-cost flow problems. Our model aims to minimize expected operating cost, accounting for s expenses throughout the network, while incentivizing clusters that are resilient, geographically compact, and have balanced demand. To generate meaningful warm-starts for our MIP and achieve computational speedups, we adapt a graph partitioning method called striping. Solutions for the clustering problem can be integrated with methods for other problems in hyperconnected network design, significantly improving their tractability. Our techniques work effectively in tandem with methods for choosing hub candidate locations and routing flow. We show the effectiveness of our methods in redesigning SF Express's hyperconnected network in Shenzhen.*

Conference Topic(s): *Interconnected freight transport, logistics and supply networks, Last mile, City logistics, PI Modelling and Simulation*
Keywords: *Logistics clustering, urban logistics, hyperconnected city logistics, last mile, GIS, Physical Internet, logistics space time network, mixed-integer linear programming*

1 Introduction

This paper proposes methods for designing a clustering of small atoms, called unit zones, into the larger pieces which serve as the backbone for hyperconnected intracity logistics, as proposed in the conceptual framework of (Montreuil et al., 2018). We work in a hyperconnected three-tier urban logistics model as depicted in Figure 1. A city is represented as a continuous mesh of unit zones (UZ), possibly of widely varying sizes, geography, population density, etc. Illustratively, an implementation in Shenzhen has unit zones covering in average about 5,000 inhabitants. These unit zones are served by access hubs (AH), where in the hyperconnected model each zone can be served by multiple nearby access hubs and each access hub can serve multiple unit zones, typically being positioned near their intersection. Unit zones are aggregated into disjoint larger spaces called local cells (LC), which in the USA could correspond to the size of a five-digit zip code. Similarly, each local cell is served by multiple larger processing centers called local hubs (LH), each of which can serve multiple local cells. Local cells are in turn aggregated into large spaces called urban areas (UA), served by gateway hubs (GH). We allow flow between any two hubs that are on the same level (horizontal flow) or one level apart (vertical flow).

In the overall network design problem, there are three high-level questions we seek to answer:

1. Where should hubs, of each type, be opened?
2. How should the set of unit zones be partitioned into local cells? How should the set of local cells be partitioned into urban areas?

3. Given the hubs and clusters, how should flow be directed?

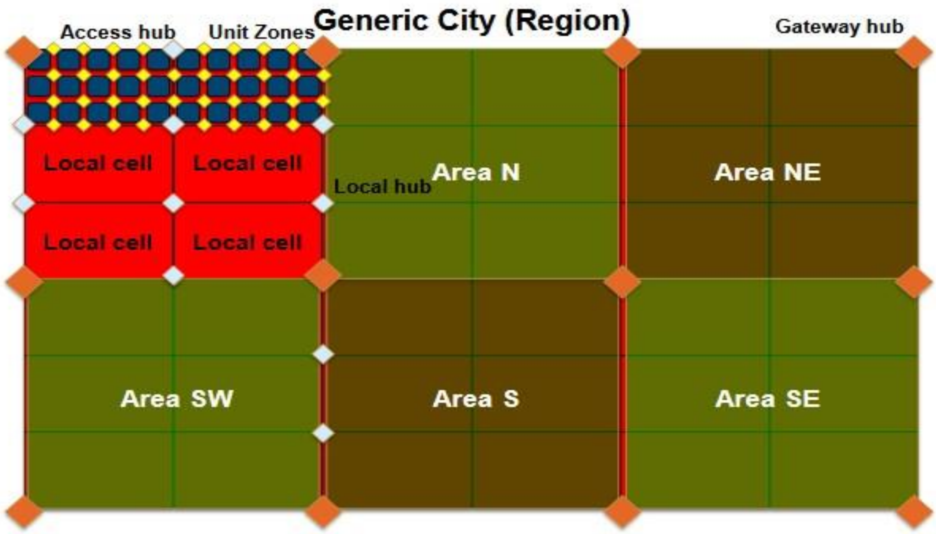


Figure 1: A generic city showing the various levels of region and hub (Montreuil et al., 2018) showing the smallest regions called unit zones, which are clustered into local cells and urban areas.

The clustering problem is to find a good solution to the second question. Such a solution is critical, as in the multi-tier web structure, it is very difficult to design operations without a shape structure. This structuring is analogous to the structuring of space in postal codes, which had been put in place to organize the postal flows yet have been frozen for decades notwithstanding the evolution of demographics and flow patterns, and are now mostly used for political, statistical, and location purposes. We generalize the structure in a wider logistics spectrum, and aim for the clusters to structure hyperconnected operations to evolve over time for optimized impact on logistics efficiency, capability, resiliency, and sustainability.

When performing the clustering, all hub locations are fixed and we approximate flow cost. We employ a robust mixed integer program method, combining aspects of graph partitioning and min-cost flow problems, to obtain a clustering which is expected to have low cost of operation and environmental impact, while enabling fast pickup and delivery and being resilient to network disruptions. We foresee such clusterings being used on two levels: an individual company may use them for its own networks, or a clustering can be used for space structuring for multiple stakeholders in an urban setting, requiring more complex criteria.

In Section 2, we discuss the assumptions and necessary data for our model. Next, in Section 3, we introduce our MIP approach for solving the clustering problem. We discuss its interaction with methods for other aspects of network redesign in Section 4, and give details of implementation in Section 5. Finally, in Section 6, we give an example of this clustering method in work with SF Express, applied to an urban area in northwest Shenzhen, China.

2 Modeling

In this section, we review the assumptions in modeling the city, its demand, and other input for the clustering model. We represent the city as a graph $G = (V, E)$, where the vertex set V is the set of UZ and the edges $e \in E$ represent pairs of adjacent UZ. Most of these adjacency pairs are determined by taking pairs of UZ whose boundaries nearly touch in a GIS map, but some additional edges may be added to connect more isolated parts of the city. We refer the interested reader to Muthukrishnan et al. (2021) on hub candidate selection for additional details of this process.

Unit zones are designed so each can be served by one or a small team of couriers. The size of unit zones can range dramatically to match the density of the urban landscape, from several blocks to a single skyscraper. Designing unit zones is outside the scope of the problems we consider here, but we want an efficient method to produce clusterings that can evolve through time as geography and demand changes require the unit zones to be redesigned. To estimate the operational costs of shipping, we have costs λ_{ij} for each pair of unit zones, and λ_{ih} for each pair of unit zone and hub. These costs may be taken to be simply a straight-line distance, but much more sophisticated measures that make use of city and geographical data are also possible. For instance, this cost may reflect elevation changes, traffic patterns, nearby building type and density, geographical variance of the demand profiles, etc. The other crucial element of the clustering instance is the demand profile. We abstract a demand profile to a single number ϕ_{ij} for each pair i, j of unit zones, representing the demand for commodities from unit zone i to unit zone j. The network structure design problem, which routes the flow more precisely than the clustering model, treats the multi-commodity aspect in more detail.

Note that in the method for determining clustering, described in the mixed integer program in Section 3, we capture the cost of vertical flow between unit zones and local hubs and/or gateway hubs, but do not capture the cost of flow between unit zones and access hubs. Routing this flow, which is carried by the couriers in the hyperconnected framework, is an important aspect of the overall network design problem. Our model can easily be augmented to handle this flow as well, but at significant computational cost given the large number of access hubs in the network. However, the feasibility of any specific UZ-AH flow is not impacted by clustering. Moreover, while the optimal pattern of flow to and from AH for a given UZ might be affected by changing the clustering, any such changes are on a local level, not large and consistent effects throughout the network. Therefore, we approximate the design at the lowest level and consider only higher levels in the web, and assume that UZ-AH assignments are made given the clustering, as in Muthukrishnan et al. (2021).

3 Mixed integer programming model

We propose using a large mixed integer program to solve the clustering problem. This MIP, detailed below, simultaneously produces clusters of unit zones into local cells and local cells into urban areas. Its objective is to minimize operational costs, which we approximate as a weighted sum of flow costs between zones and hubs; expressions for the compactness and balance of the clustering, which affect the efficiency and resiliency of low-level operations not captured in the higher-level flow; and costs for expanding the size of hubs. Horizontal flow is allowed between zones in the same local cell and local cells in the same urban area. Demand is treated on a global level where flow transit to and from a gateway hub are independent processes. Therefore, this model abstracts away some aspects of flow cost, such as the cost of opening different numbers of arcs for flow and the process of flow being aggregated as it moves up the tree towards a gateway hub. Much of this behavior is captured in the network structure design model, whose results can be reincorporated into the clustering model as detailed in Section 5.

3.1 Input and output

As input, the model takes a variety of data about the geography of the region, the unit zone network, and the demand, including a list of n unit zones indexed by i and a list of hubs, indexed by h, with base capacities κ_h and modules with capacities χ_{hm} and prices π_{hm} respectively. For each hub h, the neighborhood $N(h)$ is a set of unit zones within adequate distance to be served by h.

- Unit zone-to-unit zone demand ϕ_{ij}
- Unit zone-to-unit zone distances or costs λ_{ij}
- Unit zone-to-hub distances or costs λ_{ih}

Second, the model has a variety of parameters controlling aspects of the clustering and flow pattern, which can be set and tuned according to the problem instance:

- Maximum number of local cells K
- Maximum number of urban areas K'
- Parameters $hubs_{min}$ ($hubs_{min}{}'$) denoting the minimum number of LH needed to serve each LC, and GH needed per UA
- Parameters $cluster_{min}$, $cluster_{max}$ denoting the minimum and maximum number of local cells permitted in an urban area
- Parameters ρ^{LH}, ρ^{GH} that limit the proportion of flow between any pair of unit zones i and j that passes through a single local or gateway hub, respectively. Reducing these parameters creates a flow with increased resiliency to localized slowdowns.
- Within-local-cell cost parameter γ^L and within-urban-area cost parameter γ^A, representing the relative cost of horizontal movement compared to vertical movement
- Parameter δ controlling the proportion of the objective devoted to compactness
- Hub capacity thresholds R_l, with associated penalties ρ_l for exceeding them.

As output, the model gives a list of local cells, indexed by $1 \leq k \leq K$; a list of urban areas, indexed by $1 \leq k' \leq K$; and a high-level approximate zone-to-hub and zone-to-zone flows.

3.2 Decision variables and objective

Throughout, we index UZ by i and j, LC by k, and UA by k'. The key decision variables are binary assignment variables x_{ik}^C and $x_{ik}^A{}'$. Each x_{ik}^C is 1 if UZ i is in LC k and 0 otherwise. Likewise, $x_{ik'}^A$ is 1 if UZ i is in UA k' and 0 otherwise. Finally, c'_{kk} is 1 if LC k is contained in UA k' and 0 otherwise. Binary e_{ijk}^C are 1 if UZ i and j are both in LC k and 0 otherwise. Likewise $e_{ijk'}^A$ is 1 if UZ i and j are both in UA k' and 0 otherwise.

Flow variables d_{ijh}^{AH}, d_{ijh}^{LH}, d_{ijh}^{GH} are the (nonnegative) quantity of flow from UZ i to UZ j, passing through hub h, where h can be a local or gateway hub. Variables f_{ij}^C and f_{ij}^A are the (nonnegative) flow sent between UZ i and j at the LC level (through unspecified AH, not passing through any LH) and the UA level (through unspecified LH, not passing through any GH.) These capture vertical and horizontal flow, respectively, through the network.

3.3 Objective

The objective OBJ has five components: an estimate of the cost of vertical shipping operations, an estimate of the cost of horizontal shipping operations, a measure of the compactness of the local cells, a measure of the balance of the capacity used at each hub, and (optionally) the cost of opening modules for additional capacity at hubs. The other significant element of the cost of the network is opening and maintaining hubs, which is fixed in the input to the clustering model and which we discuss it in Section 4.

Compactness refers to the shape of LCs: i.e., an LC with a circular shape is more compact than an LC with a more elongated shape. In the hyperconnected framework, operations within

LCs are carried out by riders, who have several fixed routes between UZs in the LCs. In more compact LC, there is a greater choice of efficient routes for riders, creating better resiliency and more flexible operations. Determining these routes is itself a complex problem which this clustering model does not attempt to solve. We quantify compactness (and hence a measure of the corresponding operation cost) by the total pairwise distance between UZs in each LC.

Balance refers to the workload (i.e. capacity demands) on each hub. In operations, it is desirable to keep the workload of the hubs of each type (LH and GH) approximately equal, particularly for resiliency purposes. We quantify this by setting type-dependent thresholds B_l, with associated penalties β_l for exceeding them. If a hub exceeds B_l fraction of its capacity, cost β_l is incurred. Therefore, we express the objective as

$$OBJ = \sum_{i,j,h}(\lambda_{ih} + \lambda_{jh})d_{ijh}^{LH} + \sum_{i,j,h}(\lambda_{ih} + \lambda_{jh})d_{ijh}^{GH} + \sum_{i,j}(\gamma^C f_{ij}^C + \gamma^A f_{ij}^A) \qquad (1)$$
$$+ \delta \sum_{i,j,k} \lambda_{ij} e_{ijk}^C + \sum_{l,h} \beta_l \, b_{lh} + \sum_{h,m} \pi_{hm} z_{hm} ,$$

where the first line of the objective approximates the cost of vertical and horizontal flow, and the second includes a measure of compactness of the local cells, a measure of the balance of the hubs, and the cost of capacity modules added to the hubs.

3.4 Constraints:

We formulate the problem as a mixed integer program, minimizing the objective OBJ while satisfying the following constraints. Local cell and local hub constraints appear in left column, urban area and gateway hub constraints in the right column. Throughout, we index unit zones by i and j, local cells by k, urban areas by k', and local and gateway hubs by h.

Zone-to-cluster assignment constraints

$\sum_k x_{ik}^C = 1 \ \forall i$ (2) $\qquad \sum_{k'} x_{ik'}^A = 1 \ \forall i,$ (3)

$x_{ik}^C, x_{ik}^A \in 0,1$ (4)

Each UZ must be in precisely one LC and precisely one UA.

Tree flow constraints, ensuring contiguity of clusters

We use flow constraints, introduced in Shirabe 2009, to ensure the contiguity of each LC and UA. Each cluster has one of its UZ designated as the root of a tree in the subgraph of G induced by the vertices of the cluster tracked by the decision variables r_{ik}. The choice of root otherwise has no meaning. The remaining constraints force a flow t (abstract and unrelated to the flow of commodities in the network) on the tree. Alternative contiguity constraints may be used, such as the lengthier but stricter separator constraints introduced in Validi et al., 2020.

Multi-level clustering constraints, ensuring proper inclusion

$cluster_{\min} \leq \sum_k c'_{kk} \leq cluster_{max} \ \forall k',$ (5)

$\sum_{k'} c'_{kk} = 1 \ \ \forall k,$ (6)

$(x_{ik}^C + c'_{kk})/2 \geq x_{ik'}^A, \ \forall i, k, k',$ (7)

$c_{kk'} \in \{0,1\}.$ (8)

These constraints govern the relationship of LC to UA. The decision variables c'_{kk} track whether LC k is contained in UA k'. Each UA must contain between $cluster_{min}$ and $cluster_{max}$ LCs, and LCs may not overlap multiple UA.

Hub-to-cluster assignment constraints

$$a_{hk}^{C} \le \sum_{i\in N(h)} x_{ik} \quad \forall\, h,k, \tag{9}$$
$$a_{hk}^{\prime A} \le \sum_{i\in N(h)} x_{ik}^{\prime} \quad \forall\, h,k^{\prime}, \tag{11}$$
$$\sum_k a_{hk}^{C} \ge hubs_{\min} \forall h, \tag{10}$$
$$\sum_k a_{hk}^{\prime A} \ge hubs_{min}^{\prime}\ \forall\, h, \tag{12}$$
$$a_{hk}^{C}, a_{hk}^{\prime A} \in \{0,1\}. \tag{13}$$

Hubs (both LH and GH) may only serve a cluster (the corresponding decision variable $a_{hk(\prime)}$ is 1) if they are near to it. In particular, the cluster must intersect the hub's neighborhood $N(h)$. Furthermore, each hub must serve at least $hubs_{\min(\prime)}$LC or UA, respectively.

Hub-to-zone assignment constraints

$$y_{ikh}^{LH} \le (x_{ik} + a_{hk}^{C})/2 \quad \forall\, i,k,h, \tag{14}$$
$$y_{ik^{\prime}h}^{GH} \le (x_{ik^{\prime}} + a_{hk^{\prime}}^{A})/2\ \forall\, i,k^{\prime},h \tag{16}$$
$$y_{ih}^{LH} \le \sum_k y_{ikh}^{LH} \quad \forall i,h, \tag{15}$$
$$y_{ih}^{GH} \le \sum_{k^{\prime}} y^{A_{ikh}^{\prime}}\ \forall\, i,h, \tag{17}$$
$$y_{ih}^{AH}, y_{ikh}^{LH}, y_{ih}^{LH}, y_{ik^{\prime}h}^{GH}, y_{ih}^{GH} \in \{0,1\}. \tag{18}$$

These constraints combine the UZ-to-cluster decision variables x and the hub-to-cluster decision variables a to assign the values of variables y that track whether a UZ i can send flow to a hub h.

Total demand constraints

$$\phi_{ij} = f_{ij}^{C} + \sum_h d_{ijh}^{LH} = f_{ij}^{C} + f_{ij}^{A} + \sum_h d_{ijh}^{GH},\ \forall\, i,j \tag{19}$$

These constraints require that all demand ϕ_{ij} be met. All demand must either be met with horizontal flow within an LC or go to an LH, and likewise must either be met with horizontal flow within a UA or go to a GH.

Total capacity constraints

$$\kappa_h + \sum_m z_{hm}\chi_{hm} = c_h \quad \forall\, h, \tag{20}$$
$$\sum_{i,j} d_{ijh}^{LH} - b_{lh}\sum_{i,j}\phi_{ij} \le B_l c_h\ \forall\, l,h, \tag{21}$$
$$z_{hm}, b_{lh} \in \{0,1\},\ d_{ijh}^{LH}, d_{ijh}^{GH}, c_h \ge 0. \tag{22}$$

These constraints require that the capacity of all hubs be maintained. The base capacity of a hub h is κ_h, but this can be increased be adding modules (tracked with decision variables Z_{hm}) each with additional capacity χ_{hm}. The thresholds B_l, which determine the balance component of the objective, are checked and b_{lh} is forced to be 1 if the threshold is exceeded.

Horizontal flow constraints

$$e_{ijk}^{C} \le (x_{ik}^{C} + x_{jk}^{C})/2,\ \forall\, i,j,k, \tag{23}$$
$$e_{ijk^{\prime}}^{A} \le (x_{ik^{\prime}}^{A} + x_{ik^{\prime}}^{A})/2,\ \forall\, i,j,k^{\prime}, \tag{25}$$
$$f_{ij}^{C} \le \sum_k e_{ijk}^{C}\ \forall i,j, \tag{24}$$
$$f_{ij}^{A} \le \sum_k e_{ijk^{\prime}}^{A}\ \forall i,j, \tag{26}$$
$$e_{ijk}^{C}, e_{ijk^{\prime}}^{A}, \in \{0,1\},\ f_{ij}^{C}, f_{ij}^{A} \ge 0. \tag{27}$$

These constraints check the feasibility of horizontal flow. The variables e_{ijk}, which are 1 if both i and j are in cluster k and 0 otherwise, are set using the x_{ik}. Horizontal flow f_{ij} is permitted only if both unit zones are in the same cluster.

Vertical flow constraints

$$d_{ijh}^{LH} \le y_{ih}^{LH}\ \forall\, i,h, \tag{28}$$
$$d_{ijh}^{GH} \le y_{ih}^{GH}\ \forall\, i,h, \tag{30}$$

$d_{ijh}^{LH} \leq \rho^{LH} \cdot \phi_{ij} \ \forall\, i, h,$ (29) $d_{ijh}^{GH} \leq \rho^{GH} \cdot \phi_{ij} \ \forall\, i, h.$ (31)

The first two constraints enforce that flow from a UZ may only be sent to a hub that can serve that UZ (tracked by the decision variable y). We assume that all demand ϕ, and hence all flow d, is scaled to be bounded by 1 (if not, the RHS can be scaled appropriately.) The second two constraints enforce resiliency and robustness in the network. No more than ρ fraction of the flow for a given OD pair can be sent through a particular hub.

4 Integration and interaction with other aspects of network redesign

This clustering model is not a standalone method for network redesign. In this section, we examine how it integrates with methods for solving both hub candidate selection and network structure design problems. Each of these problems addresses one or more difficult computational aspects of the entire redesign problem, and approximates or does not consider the other aspects. The problems complement each other, so that repeatedly iterating between the problems provides a computationally tractable method that addresses all aspects of the high-level network redesign problem. Figure 2 (left) gives an overview of the information each of the three problems outputs and then gives as input to the other problems.

4.1 Hub candidate selection

The hub candidate selection process takes as input a very large set of points in the city (potentially tens of thousands of candidates.) It uses a combination of GIS methods, optimization techniques, and local expertise, incorporating zoning and other factors that may affect hub placement, to reduce the size of this set by roughly an order of magnitude. It both provides initial input to and interacts with the output of the clustering process. Further details, including the shape-based methods for determining candidate feasibility and the programs used to obtain optimal candidate sets, are given in Muthukrishnan et al. (2021). Figure 2 (right) depicts the steps in fine-tuning the candidate list, as well as the methods used.

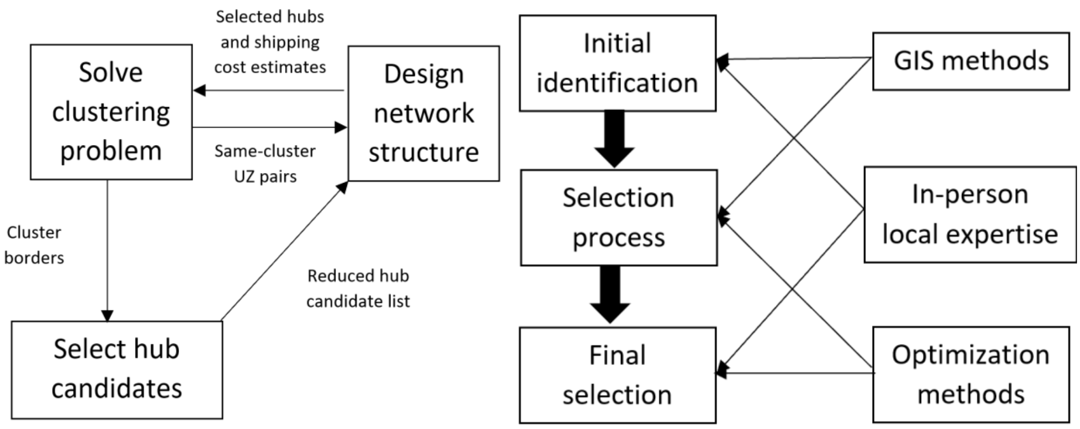


Figure 2: The iteration process and information flow between the methods of the network restructuring process; the steps and techniques used for hub candidate selection

4.2 Network structure design

The network structure design process takes as input a demand profile (modeled as commodities, each with a single origin and destination), a set of hub candidates (with associated costs), and a set of arcs between pairs of hubs, along which flow can be sent (with associated distances and costs.) The objective of structure design is to determine which hubs and which arcs to open and choose a corresponding feasible timed flow pattern, in such a way as to minimize cost. It also includes considerations for specific local requirements, such as varying traffic patterns and legal restrictions on freight transfers. This problem comes with considerable computational challenges in large instances.

While clustering is not a direct input to network structure design, it is implicitly captured in the set of arcs which are chosen out of all possible arcs between pairs of hubs. In the hyperconnected web, only pairs of AH in the same LC may have an arc between them. Therefore, choosing a clustering dramatically reduces the set of available arcs. Furthermore, in combination with the hub candidate selection methods as discussed above, hub candidates far from the boundary of clusters can be discarded. Both of these steps help reduce the size of the structure design instance and aid computational feasibility. In turn, the output of structure design provides a new potential set of hubs, as well as a detailed commodity flow on the graph, when used as input to another iteration of clustering. In particular, the shipping cost terms $d_{ijh}^{LH}\left(\lambda_{ih}+\lambda_{jh}\right)$ and $d_{ijh}^{GH}\left(\lambda_{ih}+\lambda_{jh}\right)$ in the objective of the clustering model are only approximations of the cost incurred in sending a commodity between i and j through hub h. Indeed, most parcels take a multi-hub voyage including AH, LH, and one or more GH, with cost depending on the number and type of vehicles employed on the arcs of that trip, rather than being fixed. Thus, after an iteration of the method for network structuring, which includes in its output the real costs for all arcs used by the plan, in the next iteration of clustering the objective term for all such arcs is replaced by the result from network structuring. This improves the approximation of the objective and allows the clustering to adapt and improve based on its performance.

5 Implementation

In this section, we discuss details of implementation and warm-starting for the MIP approach. This process assumes the following: a graphical model of the city; a demand profile on the city, as in Subsection 2.2; and an existing or starting list of LC and UA assignments for UZ.

5.1 Preprocessing

First, an auxiliary and artificial UZ is created for each GH. These UZ are used only to track intracity demand whose source or destination is a GH. Each such UZ i_g is assigned to its own auxiliary LC k_g, and the constraint $x_{ikg}=0,\ \forall\, i\neq i_g$ is added to the model for each such i_g, so that the LC consists only of the UZ. These LC exist only for the purpose of tracking flow in the MIP. Next, demand data is broken into intercity and intracity components. All intracity demand (UZ to UZ) becomes a corresponding term ϕ_{ij}. All intercity demand (UZ to GH or GH to UZ) is then assigned to a specific GH based on desired high-level flow patterns in the network. If the city contains multiple GH, this choice may depend on the facilities available at each GH or on the particular source or destination of the flow, so that parcels go to/from a GH that is close to their destination and source. Once this demand is assigned to a GH, it is expressed as a UZ-UZ demand, using the auxiliary UZ constructed for the GH.

5.2 Striping and warm-starting the MIP

To obtain a clustering that can be used to warm-start the MIP, we use the striping path-based algorithm introduced by Hettle et al. (2021). This algorithm takes as input a graph $G=(V,E)$ with vertex weights $w(v)$ for all $v\in V$, a desired number of parts k, a balance parameter ε, and a Hamiltonian path $(v_1,\ldots v_n)$ on G. We say that a partition is *consistent* with the Hamiltonian path used as input if each of its parts is a consecutive subpath of $(v_1,\ldots v_n)$. Using a dynamic programming framework, the algorithm returns a consistent partition of V into k parts, which correspond to local cells, and each of which is contiguous and is ε-balanced (has total weight within ε fraction of the average). Moreover, the total perimeter of the LCs in this partition is the smallest possible among all possible balanced ε-balanced consistent partitions. We apply this method to the graph G of unit zones, with the weight of each vertex given by its total demand. The choice of path is critical to producing a partition with compact parts, and we construct a Hamiltonian path by applying the uncrossing approximation algorithm for the traveling salesman problem on instances with a Euclidean metric (Van Leeuwen and Schoone, 1981). Even over multiple urban areas, both the path and the striping solution can be computed in under one minute and repeated with different paths and parameters to obtain multiple warm starts. Uncrossing requires at most $O(|V|^3)$ steps, and by using dynamic programming and exploiting the structure of the cut function used to calculate the perimeter objective for compactness, the striping algorithm runs in $O(|V|^2)$ time.

Given the initial warm-start clustering, as well as the auxiliary UZ and LC, we initialize variables x_{ik}, and thus the variables e_{ijk}, a_{hk}, y_{ikh}, and y_{ih}, which depend only on x_{ik}, are also set. Next, we determine the starting values for the variables used in the tree flow constraints for contiguity, for instance by using a breadth-first search tree on the subgraph of G induced by the vertices in each cluster k. Then we determine starting values for the vertical flow variables d_{ijh} and the horizontal flow variables f_{ij}. We assign flow so as to minimize the flow cost terms of the objective, while not yet considering the hub capacity balance term. For each pair (i,j) of unit zones, if the horizontal flow between them is cheaper than the vertical flow, we send all flow horizontally. Otherwise, we send flow vertically by greedily choosing the cheapest paths while respecting the resiliency constraints.

6 Example in southwestern Shenzhen

In this section, we describe an experiment in a large-scale setting, applying our model to a part of the SF Express network in Shenzhen. As input, we take a group of seven LCs in the southwest of Shenzhen, created using the striping method. The demand profile is based on customer behavior and the market share of SF Express, as well as on 1-day delivery times. Over 80% of demand associated with these cells is intercity, going to or from a GH. Therefore, auxiliary UZ are added as described above. For efficiency, some x_{ik} and other decision variables that depend on the x_{ik} were forced to be 0. For instance, for UZ in the easternmost part of the region, corresponding x_{ik} were set to 0 for $k = 1$ (the relatively remote dark blue LC in the southwest). All distances were computed using the length of a geodesic between the two points, with UZs represented by the centroid. Local hub neighborhoods were set to be those UZs whose centroids are within 1000m of the LH. This was sufficient in this instance, where density is still relatively great even in the less populated eastern UZs, but more sophisticated method where the distance changes throughout the city, increasing in less dense areas, are possible. Local hubs had starting capacity 3000 with up to ten optional modules of capacity 250, at cost of 5,000,000 units each.

6.1 Results

The results of applying our clustering model are shown in Figure 3. The model makes significant changes that resulted in more compact and roughly equal-demand clusters. Table 1 shows the breakdown of the various cost components of the objective function for both clusterings. The clustering changes are reflected in a significantly reduced compactness cost, modeling the expected reduction in rider operation costs. The costs of overall flow also decreased slightly, and flow was rerouted to require one fewer module to be added to local cells. Overall, the new clustering reduces the objective value by more than 6%.

Figure 3: The warm-start clustering of local cells (left) and a new proposed clustering (right).

Table 1: The costs of the clusterings in Figure 3.

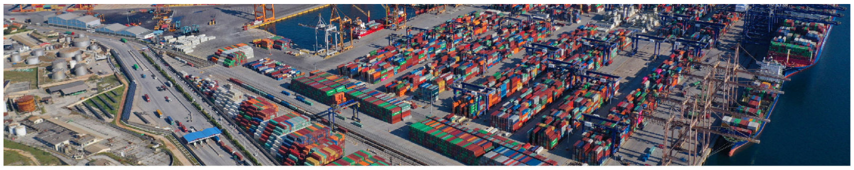
	Flow cost	Compactness	Balance	Modules	Total cost
Warm-start clustering	4.54×10^{8}	8.32×10^{7}	2×10^{7}	2×10^{7}	5.77×10^{8}
New clustering	4.32×10^{8}	7.62×10^{7}	2×10^{7}	1.5×10^{7}	5.43×10^{8}

7 Conclusion

The clustering problem in unit zone, local cell, and urban area assignment is critical in designing a logistics hyperconnected web network. We have shown that it can be effectively solved using a mixed integer program which incorporates considerations such as geographic compactness and contiguity, hub demand capacity, and resiliency. Furthermore, the MIP can be effectively warm-started using techniques from graph partitioning, such as striping, and has useful interactions with other problems in hyperconnected logistics network design. Given the large size and manifold considerations of the MIP, computational challenges do remain in very large instances, which further work may improve. Increased use of additional heuristics in concert with the MIP may be able to improve performance by streamlining some aspects of the model or exploiting geographic structure. Further development of the relationship between clustering, network design, and hub candidate selection is also a potentially rich area of study.

References

- Hettle C., S. Zhu, S. Gupta, Y. Xie (2021): Balanced Districting on Grid Graphs with Provable Compactness and Contiguity, https://arxiv.org/abs/2102.05028 (preprint)
- Montreuil B., S. Buckley, L. Faugere, R. Khir, S. Derhami (2018): Urban Parcel Logistics Hub and Network Design: The Impact of Modularity and Hyperconnectivity, Progress in Material Handling Research: 2018, Ed. A. Carrano et al., MHI, Charlotte, NC, U.S.A. https://digitalcommons.georgiasouthern.edu/pmhr_2018/19/
- Muthukrishnan, P., L. Faugere, B. Montreuil (2021): Filtering Potential Sites for Access Hub Deployment in Physical Internet Based Hyperconnected Urban Logistics Network, Proceedings of IPIC 2021 International Physical Internet Conference, 11p.
- Shirabe, T. (2009): Districting modeling with exact contiguity constraints. Environment and Planning B: Planning and Design, 36(6):1053–1066.
- Validi, H., A. Buchanan, E. Lykhovyd (2020): Imposing contiguity constraints in political districting models, http://www.optimization-online.org/DB_HTML/2020/01/7582.html (preprint)
- Van Leeuwen, J., A Schoone (1981). Untangling a traveling salesman tour in the plane. In *Proc. 7th Conf. Graph-theoretic Concepts in Comput. Sci., 1981*.



On Modelling and Solving Green Collaborative Tactical Transportation Planning

Lukas Gosch[1,2,*], Matthias Prandtstetter[1], Karl F. Doerner[2]
[1]AIT Austrian Institute of Technology, Center for Energy, Vienna, Austria
[2]University of Vienna, Austria
[*]Corresponding author: gosch.lukas@gmail.com

Abstract: *This paper presents a new mathematical model for tactical transportation planning in a horizontal collaboration defined by warehouse sharing and the joint organization of transport. The model features intermodal transport, handling and storage capacities, diverse products and realistic tariff structures with volume discounts. Furthermore, it allows for sustainable planning by associating estimated CO_2 equivalent (CO_2e) emissions to each logistic operation and then either optimize for transportation costs, emissions or both objectives. Subsequently, we derive a mixed-integer formulation for exact solution approaches and a hybrid heuristic to solve large-scale instances. The hybrid heuristic is composed of a slope scaling matheuristic we generalized to non-negative integer variables and a second local search based refinement step, which reroutes flow of multiple products at once along lowest-cost paths in the network. Results obtained from simulating collaboration in the Danube Region using the regional available transportation infrastructure including railway and shipping networks reveal significant saving potentials in both costs and CO_2e emissions. Cost minimizing solutions always lead to reductions of the carbon footprint. However, minimizing for emissions can significantly further this reduction but requires a minimum size of the collaboration to operate cost-efficiently.*

Conference Topic: *Interconnected freight transport, logistics and supply chain*

Keywords: *Freight Transportation Planning; Horizontal Collaboration; Intermodal; Emissions Reduction; Sustainability; Capacitated Network Design; Space Time Network; Hybrid Heuristic; Mixed-Integer Linear Programming; Local Search*

1 Introduction

The transport sector is responsible for over 14% of the total anthropogenic greenhouse gas emissions. Approximately 43% of these emissions can be attributed to freight transportation (Sims et al., 2014). Furthermore, the OECD predicts a tripling of the global goods traffic until 2050 (OECD, 2003; International Transport Forum, 2019) and when continuing with the status-quo, transport emissions are predicted to increase at the fastest rate compared to any other energy-related end-use sector (Sims et al., 2014). As a result, urgent measures must be taken to reduce the rising emissions in freight logistics. Central solution proposals for a more sustainable transport of goods are: modal shifts, increasing freight load factors for example through consolidation, and better optimized and integrated transport networks (Sims et al., 2014). A recent paradigm in logistics addressing all these points at once is collaboration between competitors taking the form of sharing warehouses and jointly organizing freight transportation. This so called horizontal collaboration (European Union, 2001) is a major stepping stone of the EU's strategy to achieve climate-neutrality by 2050 (ALICE, 2014, 2019) and an integral part of the Physical Internet vision (Montreuil, 2011).

Therefore, this paper introduces and solves a new mathematical model to optimally and sustainably plan and use logistic resources in a horizontal collaboration. Planning and decision making in logistics and supply chain management is structured into three different levels, strategic (long-term), tactical (medium-term) and operational (short-term) (Crainic, 2000). The model introduced here assumes that the relevant transportation infrastructure is already in place and is concerned with tactical planning which sets the general conditions for operational decisions.

1.1 Problem Description

Competing enterprises with diverse sets of products want to collaborate to reduce logistic costs and emissions. This collaboration is characterized by sharing transportation and opening up warehouses for mutual usage, both enabled by packing goods in modular and standardized load units. The companies together have a network of warehouses spread over different geographical regions with differing product demands, but not every enterprise on its own has a warehouse in every demand region. Furthermore, different companies can have product demands in similar geographical regions. Consequently, they could share certain transport routes. Now, to deliver goods into a demand region on time, the companies want to devise a shared transportation strategy which makes effective use of new consolidation and storage potentials and existing intermodal infrastructure whose usage is unlocked through higher product volumes. As these planning issues require some lead time, they are interested in developing a tactical freight plan. Hence, the model's main focus is in finding optimal paths freight should take through the existing network. This includes tariff, transport mode and storage choices making effective use of spatial and temporal consolidation potentials and of opportunities for economics of scale. The model is not concerned with concrete vehicle routing, packing problems or similar as these decisions are part of operational planning.

Key Model Aspects

Sustainable planning is achieved by associating estimated CO_2 equivalent (CO_2e) emissions to each logistic operation and optionally pricing them allowing for *internalization*. Internalization refers to estimating costs of wider effects of business activities on the community and ecosystem and integrating them into the companies budgets (McKinnon et al., 2015). As a result, the developed model allows to optimize for transportation costs, CO_2e emissions or both. The ***diversity of products*** is met with a holistic commodity-modelling approach. This includes if necessary, special transportation conditions such as cooling. Then, through adding a time dimension, the model allows for expiry-aware shared routing of perishable and non-perishable goods.

The model considers a network of facilities which can be warehouses, factories or transshipment points, connected by (capacitated) transportation relations of arbitrary types in space and time. As a result, ***intermodal transportation*** possibilities are integrated capturing their full implications on emission, cost and delivery time. Additionally, each node in the network can be endowed with ***handling capacity limits***. ***Storage*** possibilities are represented by transport relations in time between the same facility. Finally, the model incorporates ***realistic tariff structures*** with all-unit volume discounts often found in practice (Munson and Jackson, 2015). We design graph-structures to linearly model these tariffs. These allow us to formulate the problem as a ***capacitated network design problem*** (Magnanti and Wong, 1984).

1.2 Overview of the Paper

In Section 2 our model is formally introduced. Due to the linear mixed-integer programming formulation being NP-hard, we develop and combine two different types of heuristics to successfully solve large instances. In Section 3.1 we generalize the slope scaling heuristic (Kim and Pardalos, 1999; Crainic et al., 2004) to non-negative integer variables making it applicable to our model. Slope scaling is a matheuristic originally developed for network design problems with binary variables and constitutes an integral part of state-of-the-art hybrid heuristics for this problem type (Gendron et al., 2018; Akhavan Kazemzadeh et al., 2021). In Section 3.2 a local-search based approach - using the slope scaling heuristic as a fast and effective construction heuristic - is developed. Based on ideas from Harks et al. (2016), it jointly reroutes flow of multiple commodities using lowest-cost paths in the network. The aforementioned heuristics as well as a commercial MIP-solver are applied to generated problem instances based on the real intermodal transportation infrastructure in the Danube Region. The results of which can be found in Section 4. They show significant savings in costs and emissions. A conclusion can be found in Section 5.

2 Mathematical Model

This section presents the formal model developed and used in this work. Fundamentally, two stages of the underlying network are differentiated. First, a time-expanded network is given (see Section 2.1). This forms the basis on which the heuristics (see Section 3) operate. However, non-linear transportation tariffs between nodes result in a non-linear objective. Therefore, in a second stage simple arcs representing these transport relations can be replaced with more complex graph-gadgets to obtain a linear mixed-integer programming (MIP) formulation called the tariff-expanded network (see Appendix B). This allows to apply exact MIP-solvers to the problem[1]. However, only the heuristic solution approaches manage to effectively solve large instances. We call the developed network design problem the generalized tactical transportation problem (GTTP), as it can be shown to include the tactical transportation problem introduced by Harks et al. (2016) as a special case.

2.1 Time-Expanded Network

The time-expanded network $G_{\mathcal{T}}$ is constructed for a certain time horizon T (e.g. 7, 14 or 30 days) with the individual time periods summarized in the set $\mathcal{T} = \{1, ..., T\}$. $G_{\mathcal{T}}$ is constructed from a set of base nodes copied T-times resulting into a time-expanded node set $\mathcal{V}_{\mathcal{T}}$. Base nodes are either: physical facilities (e.g. warehouses, factories or transshipment points) with each facility being part of a demand region, demand nodes used to represent demand regions, or bin nodes to remove unused or expired goods form the system. There is only one bin node in the set of base nodes. Then, the arc set $\mathcal{A}_{\mathcal{T}}$ consists of transport relations between these nodes in different time periods. Transport relations can be of arbitrary types and we distinguish transportation modes $\mathcal{M} = \{L, R, S\}$ consisting of lorry (L), rail (R) and ship (S), storage arcs connecting the same facility at two consecutive time periods denoted by type C, and artificial arcs connecting each facility to its regional demand node as well as to the bin node. These use the type-symbol Ω.

Network Structure

The facilities are connected among themselves based on physical realities. If for example two facilities $i, j \in \mathcal{V}_{\mathcal{T}}$ are connected by a road network, then there exists an arc $a = (i, j, L) \in \mathcal{A}_{\mathcal{T}}$ with corresponding non-zero distance d_a and non-zero travel time τ_a. Storage arcs have a distance of zero but a travel-time of one. For each time-period, there is one bin node and each facility is connected to it with distance and travel time of zero. Facilities in a demand region are connected to the associated demand node with distance and travel times of zero using mode Ω. Hence, demand

[1]If one only optimizes for emissions, the time and tariff expanded networks are equivalent. This is due to the (very unfortunate) physical impossibility of volume discounts on emissions.

in the region the warehouse is located in can be satisfied directly out of stock. This is due to the fact that the model is only concerned with successful transportation into a region to fulfill its demand and the inner-regional distribution to customers should be addressed by an operational model. Furthermore, each facilities is connected to every other demand region except its own using mode L with non-zero distances and travel times. These connections represent direct deliveries by lorries into a demand region without first delivering into a regional facility or using any form of intermodal transport.

2.2 Commodities

All facilities are supplied with a certain stock of products, also called commodities, each time period acting as *source* nodes. Demand nodes have associated commodity demands for each time period acting as *sinks*. The set of commodities is denoted by $\mathcal{K}$. In general, products can be highly diverse in their volume and weight henceforth called the products *properties* (see also Harks et al. (2016)). As an example, it does make a non-negligible impact on the transportation costs and available resources if one transports a certain number of styrofoam sheets compared to the same number of steel beams. Therefore, each commodity $k \in \mathcal{K}$ has an associated extent p_{kl} for each property $l \in \mathcal{P}$. Additionally, products have different transportation and storage characteristics. As an example, some goods may require cooling to specific temperatures or can be perishable. To model different transportation requirements, we introduce a *product type* concept and associate to each commodity a specific type from a set of types Σ. Only goods of the same type can be transported together. In our instances, we separate two types of goods, those whose storage-containers require electricity and those whose storage-containers don't. Additionally, to address perishability commodities $k \in \mathcal{K}$ can have varying *lifetimes* Δt_k. As a result, a commodity is defined as perishable if its lifetime is smaller than the number of time-periods in the network. Consequently, commodities can be grouped into a set of perishable ones $\mathcal{K}_\Delta$ and a set of non-perishable ones $\mathcal{K}_\mathcal{T}$. In order to correctly handle perishability in the model, each perishable commodity has an associated *production timestamp* $t \in \mathcal{T}$ and is counted as expired in time-periods greater than $t + \Delta t_k$. With $x_{k_t} \in \mathbb{R}_0^+$ we denote the *flow* of perishable commodity k produced at time t and with $x_k \in \mathbb{R}_0^+$ the flow of non-perishable commodity k. These can be pooled together into a *flow vector* $x \in \mathbb{R}_0^{|\mathcal{K}_\mathcal{T}|+|\mathcal{K}_\Delta||\mathcal{T}|}$. The summed up extent of property $l \in \mathcal{P}$ over all commodities of type $\sigma \in \Sigma$ in x is given by the flow-sum function $P_l^\sigma(x) := \sum_{k \in \mathcal{K}_\mathcal{T}^\sigma} p_{kl} x_k + \sum_{t \in \mathcal{T}} \sum_{k \in \mathcal{K}_\Delta^\sigma} p_{kl} x_{k_t}$. The sets $\mathcal{K}_\mathcal{T}^\sigma$ and $\mathcal{K}_\Delta^\sigma$ comprise all non-perishable or perishable goods of type σ.

2.3 Tariffs

Each transport relation has one[2] associated tariff for each product type. Harks et al. (2016) give a good overview of possible tariffs and their resulting cost functions. In this work, we are especially concerned with *all-unit quantity discounts* as they are the most prevalent (Munson and Jackson, 2015). Corresponding tariffs usually have a set of cost levels $\mathcal{N} = \{1, \ldots, N\}$ defining different cost rates depending on the shipped volume or weight. Therefore, Harks et al. (2016) develop a cost function with different linear cost rates depending on the sent flow extent in some property $l \in \mathcal{P}$ as follows. Each cost level $n \in \mathcal{N}$ is applied starting from transport volume β_n and has the linear cost rate c_n. As a result, the cost function by Harks et al. (2016) for flow vector x and a specific

[2]Multiple available tariffs can be modeled by parallel arcs.

product type[3] σ is given as

$$C(x) = \min_{n \in \mathcal{N}} \{c_n \cdot \max\{P_l^\sigma(x), \beta_n\}\} \tag{1}$$

However, only linearly pricing transport volumes does not adequately capture the diversity of all-unit volume discount tariffs employed by freight transportation companies. Especially, it is not applicable to constant cost levels for different transport weight or quantity intervals or a combination of a fixed-cost rate with volume-dependent linear costs resulting in piecewise-linear cost levels (see Figure 1 in Appendix B). An overview of such tariffs with more complex cost structures is given in Kempkes and Koberstein (2010). Below, we formalize these tariffs used in our work.

2.3.1 Constant Cost Levels with All-Unit Discounts

Constant cost levels for different weight or quantity intervals are, among others, often encountered in rail cargo shipments. In them for each cost level $n \in \mathcal{N}$ a different fixed-cost rate f_n is assumed. In theory, these cost levels can have arbitrary interval lengths. We assume a fixed property- and type-dependent interval length of B_l^σ and interpret it as the capacity in direction $l \in \mathcal{P}$ of one installed *transport unit (TU)* on this transport relation. Exemplary, for rail or ship cargo shipments one transport unit could be one ISO container (see Figure 1 in Appendix B). Then, the number of installed TUs for product type σ is given by $\tilde{y}^\sigma = \max_{l \in \mathcal{P}} \{\lceil \frac{P_l^\sigma(x)}{B_l^\sigma} \rceil\}$. In the mixed-integer problem installed TUs are additional integer decision variables (see Appendix B.2). Now, β_n denotes the minimum number of TUs at which cost level n starts to be applicable. The starting cost of the n-th level is b_n and represents the base cost of shipping β_n containers. This yields the cost function (2).

$$C(x) = \min_{n \in \mathcal{N}} \{f_n \cdot \max\{\tilde{y}^\sigma - \beta_n, 0\} + b_n\} \tag{2}$$

Exemplary, b_n can represent the price of β_n TUs with the new cost rate f_n, i.e. $b_n = f_n \beta_n$. If $f_n \leq f_{n-1}$, this represents an all-unit discount structure. However, if b_n represents the costs of $\beta_n - 1$ TUs to the previous price rate f_{n-1} plus the costs of one TU to the current rate f_n, it would represent incremental discounts. Note that to formulate the cost function not in terms of installed TUs, but in term of single-property levels (as for example weight levels), replace $\tilde{y}^\sigma$ with $\tilde{y}_l^\sigma = \lceil \frac{P_l^\sigma(x)}{B_l^\sigma} \rceil$.

2.3.2 Piecewise-Linear Cost Levels with All-Unit Discounts

Tariffs with different piecewise linear cost levels are very similar to the constant cost levels case except that for each discount level n, we additionally assume variable costs c_n for the actual volume or weight of flow. The resulting cost function writes

$$C(x) = \min_{n \in \mathcal{N}} \{f_n \cdot \max\{\tilde{y}^\sigma - \beta_n, 0\} + c_n \cdot \max\{P_l^\sigma(x) - B_l^\sigma(\beta_n - 1), 0\} + b_n\} \tag{3}$$

Again, setting $b_n = f_n \beta_n + c_n B_l^\sigma(\beta_n - 1)$ with $f_n \leq f_{n-1}$ and $c_n \leq c_{n-1}$ results in all-unit discounts. Note that the linear costs only depend on one property. This cost function exemplary arises when modelling transportation costs with lorries. A fixed costs part arises for each commissioned truck due to various singular factors such as driver wages. Linear costs arise due to increasing fuel consumption based on its actual weight.

[3] Harks et al. (2016) do not have a product type concept and aggregate the property-extent over all products.

3 Heuristics

In Section 3.1 we generalize the slope scaling heuristic (SSC) to non-negative integer variables and develop an approximate linear program for our model. SSC is used as a fast construction heuristic for the local search presented in Section 3.2. We also experimented with iterative linear programming (Gendron et al., 2018) as theoretical properties such as finite convergence can be generalized to non-negative integer variables. However, we found it does not scale to larger instances.

3.1 Slope Scaling

Slope Scaling is an iterative solution process based on solving an approximate linear program $\mathcal{LP}$ exactly. The resulting solution $\tilde{x}$ is used to construction a feasible solution $(\tilde{x},\tilde{y})$ to the original problem $\mathcal{MIP}$. Then, $(\tilde{x},\tilde{y})$ is used to update the coefficients in the objective of $\mathcal{LP}$ yielding $\mathcal{LP}'$ in such a way that the objective function value $v(\mathcal{MIP}(\tilde{x},\tilde{y}))$ equals $v(\mathcal{LP}'(\tilde{x}))$. We denote the set of linearization coefficients in the linear program at iteration t as $\rho(t)$ and the corresponding linear program as $\mathcal{LP}(\rho(t))$. $\mathcal{LP}(\rho(t))$ takes the form of a network flow problem and the formal definition is presented in Appendix C. After efficiently solving $\mathcal{LP}(\rho(t))$, a solution $(\tilde{x},\tilde{y})$ to $\mathcal{MIP}$ is obtained by setting $\tilde{y}_a^\sigma = \max_{l\in\mathcal{P}}\{\lceil\frac{P_l^\sigma(\tilde{x}_a)}{B_l^\sigma}\rceil\}, \forall a \in \mathcal{A}_\mathcal{T}$. Denote with $\lambda_1 f_a^\sigma$ the weighted costs and with $\lambda_2\Delta_a^\sigma$ the weighted emissions of one TU (see $\mathcal{MIP}$-objective in Appendix B.3) and with $l(a)$ the property used for linear pricing on arc a. Now, using $(\tilde{x},\tilde{y})$, $\rho(t)$ can be updated as follows

$$\rho_a^\sigma(t+1) = \begin{cases} (\lambda_1 f_a^\sigma + \lambda_2\Delta_a^\sigma)\tilde{y}_a^\sigma / P_{l(a)}^\sigma(x_a) & \text{if } P_{l(a)}^\sigma(x_a) > 0 \\ \rho_a^\sigma(t) & \text{otherwise} \end{cases} \tag{4}$$

resulting in $v(\mathcal{MIP}(\tilde{x},\tilde{y})) = v(\mathcal{LP}(\rho(t+1)))$. Then, $\mathcal{LP}(\rho(t+1)))$ is solved and the above solution construction and update scheme repeated for $t+1$. Initial cost estimates are set to the linearized costs of one full TU on the respective connection, i.e. $\rho_a^\sigma(t=0) = (\lambda_1 f_a^\sigma + \lambda_2\Delta_a^\sigma)/B_{al(a)}^\sigma$. The slope scaling heuristic starts at $t=0$ and runs for as many iterations until an already seen solution is produced (resulting in a loop) or a specific time limit is reached. Additionally, it is possible to omit $\tilde{y}_a^\sigma$ from the update equation (4). Then, $v(\mathcal{MIP}(\tilde{x},\tilde{y})) \neq v(\mathcal{LP}(\rho(t+1)))$ but for $t \geq 1$ the update scheme becomes *monotonic* on each tariff level with respect to increasing flow extent. This can be interpreted as favoring higher transport volumes on an arc no matter the filling rate of the last TU, playing into the economics of scale. The results in Section 4 indicate that the monotonic update scheme outperforms the update scheme (4) matching objective-costs. This could indicate that the prime paradigm behind designing a slope scaling mechanism - matching objective-costs (Kim and Pardalos, 1999) - is not necessarily key to its success. Other properties such as monotonicity can be equally important. This insight can be used to devise novel problem specific update rules diverting from pure objective-costs matching.

3.2 Local Search

The presented local search is based upon the *flow decomposition theorem* (Ahuja et al., 1993) which states that every non-negative arc flow can be represented as a path and cycle flow. Thus, the basic idea of the local search is to construct a path-decomposition of flow on the acyclic time-expanded graph and then randomly dissolve paths and reroute the flow. The employed moves and path decomposition calculations are inspired by Harks et al. (2016). A path-decomposition $\mathcal{P}$ consists of a set of tuples (P, f_P). Each tuple consists of a path P in $\mathcal{G}_\mathcal{T}$ and the transported commodities $f_P \in \mathbb{R}_0^{|\mathcal{K}_\mathcal{T}|+|\mathcal{K}_\Delta||\mathcal{T}|}$. Paths carry only commodities of a specific type and due to the topology of the network, either end in a demand node or a bin node. Therefore, we call a path either a demand-path or a bin-path. Note that for a given flow $\mathcal{P}$ is not necessarily unique.

Now, we distinguish two types of neighbourhoods. The first consists of solutions constructed by removing one demand-path and all bin-paths and then, repairing the solution by rerouting the flow. The second neighbourhood differs by instead removing a group of demand-paths sharing the same transport relation. Therefore, the neighbourhoods are characterized by a given path-decomposition and the employed rerouting scheme R and are denoted $\mathcal{N}_1(\mathcal{P},R)$ and $\mathcal{N}_2(\mathcal{P},R)$, respectively. A move is defined by randomly choosing a solution from $\mathcal{N}_1(\mathcal{P},R)$ or $\mathcal{N}_2(\mathcal{P},R)$ and accepting it, if it improves upon the current solution. Furthermore, we distinguish two rerouting schemes: a *heaviest first* rerouting and a *cheapest relative cost* rerouting. The heaviest first rerouting searches a (source s, sink t)-pair allowing for the highest delivery weight and calculates a cheapest path from s to t and adds the found path to the path decomposition. It iteratively repeats these two steps until all demand has been satisfied or infeasibility is detected. The cheapest relative cost rerouting differs by choosing the (s,t)-pair which results in the cheapest path relative to the transported weight. The rerouting schemes are handling-capacity aware and if necessary, the flow of goods f_P is adapted to exactly fulfill handling-capacity limits. Arcs not allowing for additional flow are not anymore considered in finding cheapest paths. Lastly, two move-dependent ways to calculate the path decomposition are employed. For moves in $\mathcal{N}_1$ the path decomposition is calculated in a depth-first search manner. Starting at a source, for each incident arc, the maximum-weight flow vector that could be assigned to a path traversing this arc is calculated. The arc with the maximal maximum-weight-flow vector is chosen. For moves in $\mathcal{N}_2$ a bidirectional depth-first search is employed. Starting from the most heavily used arc with respect to unassigned flow, it chooses arcs in both directions that maximize the savings if one would reduce their flow by the assigned one.

To summarize, the local search works by in the beginning constructing a solution using slope scaling, followed by choosing an initial neighbourhood $\mathcal{N}_i(\mathcal{P},R)$ and then, repeating the following steps until convergence or a time limit:

1. Construct a path-decomposition for $\mathcal{N}_i(\mathcal{P},R)$.
2. Follow randomly chosen improving neighbouring solutions until the improvement over a certain number of iterations falls below a threshold.
3. Change R and go to step 1. If R has been changed in the last iteration, change i instead.

For the results in this paper, we chose $\mathcal{N}_1(\mathcal{P}, \textit{cheapest relative cost})$ as an initial neighbourhood.

4 Results

In this section we present the results of applying a state-of-the-art MIP solver and our heuristics to instances modelling horizontal cooperation in the Danube region.

4.1 Data

The instances are based on a dataset introduced by Wolfinger et al. (2019) including locations of major cities, train stations and Danube ports in Austria, Slovakia, Hungary, Romania, Serbia, and Bulgaria. In each city, we have added additional random locations and recalculated all road-distances using Ariadne (Prandtstetter et al., 2013). Distance and time-matrices of other transport modes were taken from Wolfinger et al. (2019). Additionally, each city is interpreted as a demand region and we randomly distribute commodity demands and stocks. Generated instances range in size from two collaborating warehouses in two different regions up to the exhaustion of our memory limits during the solution process. More concretely, input instances can be grouped into having two ($R2$), fife ($R5$) or ten regions ($R10$). $R2$-instances are generated with either one or four warehouses per region. $R5$-instances are generated with two or four warehouses per region. For $R10$-instances the five more populated regions have four warehouses and five lower populated ones have two. Warehouses are again grouped representing association to different companies. Some regions are equipped with a cross-dock, train station or port. All instances have a time period of one day and are generated in three different time horizon (T) versions of 7, 14 or 30 days. Lastly, each instance has a version with active handling constraints (tight t) and without (loose l). Therefore, instance groups are identified using $R<\#>\_T<\#>\_t/l$ and using a $*$ instead of $<\#>$ to indicate averages over all respective instance groups. Tariffs with three levels of volume discounts are employed. Piecewise-linear cost levels are used for lorry connections and constant cost levels for rail and ship connections. On these connections, TUs always correspond to 40-feet ISO containers. Handling costs are integrated into the arc costs. Prices are set based on information from an industry partner.

In total 60 instances have been generated leading to 180 optimization problems considering optimizing for costs, emissions or both. The input instances and the code to generate them together with an in-depth description of the data (including price and emission tables) are available on https://github.com/saper0/gttp-data.

4.2 Experiments

Solutions calculated on our test instances are evaluated with respect to a *direct delivery solution*. In a direct delivery solution demand for a product is satisfied by a lorry shipment without the possibility of cooperation. This means consolidation is only possible across warehouses associated to the same company. Direct delivery instances are generated by removing transshipment nodes and arcs connecting warehouses from different companies from the original input instances. Next, a MIP-solver is applied on the adapted instances for one hour. The thereby calculated direct delivery solutions are either optimal or proven to be within a 1% optimality gap. All experiments were written in C++ using CPLEX 12.10, had a time limit of one hour and were performed on an Intel(R) Xeon(R) CPU E5-2640 v3 @ 2.60GHz CPU with 32 GB RAM.

4.2.1 Algorithmic Insights

Table 1 shows that the local search based approach, using one iteration[4] of slope scaling (SSC) to generate a starting solution, outcompetes all other solution approaches by large margins. Additionally, SSC using the more monotonic linearization procedure slightly outperforms the objective cost-matching one. This highlights the sensibleness of problem specific slope scaling mechanisms parting from the paradigm of matching objective costs between the approximate and original problem. Lastly, the MIP solver - initialized with the direct delivers solution to aid its solution process - struggles with emission optimization and completely fails if transportation costs should be optimized. This is due to the fact that when transportation costs are included, the MIP operates on the tariff expanded network leading to a model blow up quickly exceeding the memory limits.

4.2.2 Managerial Insights

Table 1 shows the average percentual improvements of the objective functions in different instance groups obtained by applying the aforementioned solution approaches. The following discussion concerns the best obtained result on each instance (consistently by local search, see Section 4.2.1). Optimizing for emissions universally leads to significant CO_2e reduction possibilities averaging

[4] Applying SSC until convergence and then applying local search leads to slightly inferior solutions.

Table 1: Average percentual improvements (reductions) of the objective functions in different instance groups compared to the direct delivery solution. Hence, for emissions the improvement on CO_2e, for costs the improvement on the transportation costs and in both the improvement on the sum of emissions and costs by pricing one tonne of CO_2e with 100€ (Delft, 2019) (using the avoidance cost approach (McKinnon et al., 2015)) is shown. Due to its random components, the local search has been run with ten different seeds. Its reported results are averages over these runs including their sample standard deviation. For the MIP approach, the number of solvable instances versus all instances in the instance group are shown. The other solution approaches can solve all instances except two in $r10\_T*\_*$ due to exceeding memory limits.

Optimize for		Local Search [%]	SSC (monotonic) [%]	SSC [%]	MIP [%]
Emissions	$r2\_T*\_*$	30.7 ± 0.1	29.9	29.7	27.3 (24/24)
	$r5\_T*\_*$	31.9 ± 0.2	29.5	29.2	15.3 (22/24)
	$r10\_T*\_*$	30.9 ± 0.4	28.9	28.5	9.2 (10/12)
	$r*\_T*\_*$	31.2 ± 0.2	29.6	29.3	19.4 (56/60)
Costs	$r2\_T*\_*$	6.0 ± 0.2	0.9	-1.6	2.0 (22/24)
	$r5\_T*\_*$	20.1 ± 0.4	12.5	10.3	0.0 (7/24)
	$r10\_T*\_*$	23.6 ± 0.6	16.9	14.5	0.0 (1/12)
	$r*\_T*\_*$	14.9 ± 0.4	8.5	6.8	1.5 (30/60)
Both	$r2\_T*\_*$	8.0 ± 0.3	2.2	1.5	2.1 (22/24)
	$r5\_T*\_*$	21.5 ± 0.4	14.9	12.8	0.0 (9/24)
	$r10\_T*\_*$	24.3 ± 0.5	18.5	16.3	0.0 (1/12)
	$r*\_T*\_*$	16.4 ± 0.3	10.2	8.7	1.4 (32/60)

31.2%±10.2%[5] (13.8% - 53.0%) *independent* of the size of the collaboration. However, we find emission reduction potentials are strongly correlated with the planning horizon. A planning horizon of 7 days lead to on average 20.5%±3.7% CO_2e reduction, 14 days to 33.0%±5.2% and 30 days to 41.1%±7.6%. This is due to the fact that intermodal transport has longer lead times. Therefore, some changeover time from lorry transportation is needed to make effective use of its possibilities. Concerning transportation cost optimization, contrary to emission minimization we find a clear correlation with the size of the collaboration increasingly enabling cost-efficient consolidation (see Appendix A for more results on consolidation). Significant cost reductions are possible on all looked upon instances averaging 14.9%±9.2% (0.4% - 30.5%). Similar to optimizing for emissions, we additionally find a dependence on the planning horizon with less savings potential for a horizon of 7 days (12.0%±8.9%) than 14 days (15.9%±9.2%) than 30 days (17.0%±9.1%). This can be explained by increased consolidation possibilities (see Appendix A).

Table 2 in Appendix A shows that as expected optimizing for transportation costs simultaneously leads to emission reductions which are more pronounced the larger the size of the collaboration and the more time periods considered. However, emission reductions are not as large as when optimizing for them only. The picture is different when optimizing for emissions only. Here, solutions for small instances with collaborating companies between two regions results in slightly increased costs (7.0%±5.9%). Interestingly, the calculated small cost increases together with the average emission reductions of roughly 30% are consistent with a real life case study conducted in the EU in which four companies (shippers) from two regions collaborated (Jacobs et al., 2013). Hence, our results showcase the existence of a minimum necessary size of the collaboration such

[5] Different to Table 1, in the next two paragraphs the standard deviations are calculated based on the average results of the different instances in the referred to instance-group and not on the instance-group averages over multiple runs.

that decisions aiming at emission reduction also result in cost savings. Note that internalizing CO_2e emissions results in a good trade-off between optimizing for costs or emissions only.

5 Conclusion

In this paper we introduced a new mathematical model for tactical transport planning in a horizontal collaboration. The model incorporates realistic tariff structures, intermodal transport, handling capacities and storage possibilities. Furthermore, it can be applied to plan transportation with a diverse set of products among others supporting perishability or special transport needs such as cooling. The model allows for sustainable planning through internalizing CO_2e costs. Hence, it can be used to minimize transportation costs, emissions or both together. Subsequently, we developed a mixed-integer formulation of the model for exact solution approaches and a memory efficient hybrid heuristic. The hybrid heuristic is composed of two parts, a matheuristic which we generalized to non-negative integer variables and a local search. These solution techniques were successfully applied to generated problem instances based on the real transportation infrastructure - including railway and shipping - in the Danube Region. This revealed significant saving potentials in both costs and emissions. As expected, optimizing for transportation costs automatically leads to a reduced carbon footprint. However, optimizing for the carbon footprint only necessitates a minimum size of the collaboration for transportation costs to simultaneously decrease.

Our results support the assumption that horizontal collaboration in warehouse-sharing and unlocking intermodal freight transport opportunities positively impact emissions and climate-neutral actions. Therefore, further efforts need to be made to promote collaboration in real-world settings.

Acknowledgements

We want to thank Seragiotto Clovis for performing the computations with Ariadne and Georg Brandtstätter for his excellent compute cluster support. This work received funding by the Austrian Federal Ministry for Climate Action, Environment, Energy, Mobility, Innovation and Technology (BMK) in the research program “Mobilität der Zukunft” under grant number 877710 (PhysICAL).

References

Ravindra K Ahuja, Thomas L Magnanti, and James B Orlin. *Network Flows: Theory, Algorithms, and Applications*. Prentice hall, 1993.

Mohammad Rahim Akhavan Kazemzadeh, Tolga Bektaş, Teodor Gabriel Crainic, Antonio Frangioni, Bernard Gendron, and Enrico Gorgone. Node-based lagrangian relaxations for multicommodity capacitated fixed-charge network design. *Discrete Applied Mathematics*, 2021.

ALICE. *Global supply network coordination and collaboration: research and innovation roadmap*. 2014. URL `https://www.etp-logistics.eu/wp-content/uploads/2015/08/W46mayo-kopie.pdf`.

ALICE. *A framework and process for the development of a ROADMAP TOWARDS ZERO EMISSIONS LOGISTICS 2050*. 2019. URL `http://www.etp-logistics.eu/wp-content/uploads/2019/12/Alice-Zero-Emissions-Logistics-2050-Roadmap-WEB.pdf`.

Tolga Bektaş, Jan Fabian Ehmke, Harilaos N. Psaraftis, and Jakob Puchinger. The role of operational research in green freight transportation. *European Journal of Operational Research*, 274(3):807–823, 2019.

Dimitris Bertsimas and John Tsitsiklis. *Introduction to Linear Optimization*. Athena Scientific, 1st edition, 1997. ISBN 1886529191.

Teodor Gabriel Crainic. Service network design in freight transportation. *European Journal of Operational Research*, 122(2):272 – 288, 2000.

Teodor Gabriel Crainic, Bernard Gendron, and Geneviève Hernu. A slope scaling/lagrangean perturbation heuristic with long-term memory for multicommodity capacitated fixed-charge network design. *J. Heuristics*, 10:525–545, 09 2004.

CE Delft. *Handbook on the external costs of transport*. Publications Office of the European Union, Luxembourg, 2019.

European Union. Commission notice: guidelines on the applicability of article 81 of the ec treaty to horizontal cooperation agreements. *2001/C3/02*, 2001.

Bernard Gendron, Saïd Hanafi, and Raca Todosijević. Matheuristics based on iterative linear programming and slope scaling for multicommodity capacitated fixed charge network design. *European Journal of Operational Research*, 268(1):70 – 81, 2018.

Tobias Harks, Felix König, Jannik Matuschke, Alexander Richter, and Jens Schulz. An integrated approach to tactical transportation planning in logistics networks. *Transportation Science*, 50(2):439–460, 2016.

International Transport Forum. *ITF Transport Outlook 2019*. 2019. URL `https://www.oecd-ilibrary.org/content/publication/transp_outlook-en-2019-en`.

Kurt Jacobs, Steven Vercammen, and Sven Verstrepen. *CO^3 Test Project Report: Creation of an Orchestrated Intermodal Partnership Between Multiple Shippers*. 2013. URL `http://www.co3-project.eu/wo3/wp-content/uploads/2011/12/CO3-D4-2-exec-summary.pdf`.

Jens Kempkes and Achim Koberstein. A resource based mixed integer modelling approach for integrated operational logistics planning. *Lecture Notes in Business Information Processing*, 46:281–294, 01 2010.

Dukwon Kim and Panos M. Pardalos. A solution approach to the fixed charge network flow problem using a dynamic slope scaling procedure. *Operations Research Letters*, 24(4):195–203, 1999.

T. L. Magnanti and R. T. Wong. Network design and transportation planning: Models and algorithms. *Transportation Science*, 18(1):1–55, 1984.

Alan McKinnon, Michael Brown, Maja Piecyk, and Anthony Whiteing. *Green Logistics : improving the environmental sustainability of logistics*. Kogan Page Limited, 3th edition, 2015.

Benoit Montreuil. Toward a Physical Internet: meeting the global logistics sustainability grand challenge. *Logistics Research*, 3(2-3):71–87, May 2011.

Charles L. Munson and Jonathan Jackson. Quantity Discounts: An Overview and Practical Guide for Buyers and Sellers. *Foundations and Trends® in Technology, Information and Operations Management*, 8(1–2):1–130, June 2015. Publisher: Now Publishers, Inc.

OECD. *Delivering the Goods*. 2003. URL `https://www.oecd-ilibrary.org/content/publication/9789264102828-en`.

Matthias Prandtstetter, Markus Straub, and Jakob Puchinger. On the way to a multi-modal energy-efficient route. In *IECON 2013 - 39th Annual Conference of the IEEE Industrial Electronics Society*, pages 4779–4784, 2013.

Ralph Sims, Roberto Schaeffer, Felix Creutzig, Xochitl Cruz-Núñez, Márcio D'Agosto, Delia Dimitriu, M.J. Meza, Lewis Fulton, S. Kobayashi, Oliver Lah, Alan Mckinnon, P. Newman, Minggao Ouyang, J.J. Schauer, Daniel Sperling, and Geetam Tiwari. *Transport. In: Climate Change 2014: Mitigation of Climate Change. Contribution of Working Group III to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change*. 09 2014.

David Wolfinger, Fabien Tricoire, and Karl F. Doerner. A matheuristic for a multimodal long haul routing problem. *EURO Journal on Transportation and Logistics*, 8(4):397–433, 2019.

*Table 2: This tables gives more in-depth analytics on the solutions obtained by local search. The **first column** showcases percentual improvements of the respective quantity (costs or emissions) not optimized for compared to the direct delivery solution. Hence, for emission minimization the effects on the costs are shown. For cost optimization the effects on CO_2e emissions. When optimizing for both, Table 1 shows a weighted sum of costs and emissions. Hence, here both quantities are given separate. The **second column** gives a measure on the relative intermodal consolidation. It measures in percent how much less outgoing rail, ship and cross-dock to cross-dock transport units are ordered relative to the incoming truckloads (lorry-TUs). The **third column** shows the number of ordered ship and rail TUs relative to all ordered TUs. Again, reported results are averages over all runs including their sample standard deviation over the runs. The **last column** shows the percentual increase of the average filling rate calculated over all ordered TUs compared to the direct delivery solution.*

Optimize for	Results in [%]	Costs or Emissions		Consolidation	Rail & Ship	Filling Rate
Emissions	$r2\_T*\_*$	-7.0 ± 0.2		25.4± 0.2	6.3 ± 0.2	-5.6 ± 0.5
	$r5\_T*\_*$	13.4 ± 0.4		13.1± 1.2	7.2 ± 0.1	7.4 ± 0.8
	$r10\_T*\_*$	17.1 ± 0.6		20.2± 1.8	7.9 ± 0.1	14.1 ± 0.8
	$r*\_T*\_*$	5.6 ± 0.3		19.4± 0.9	6.9 ± 0.1	3.2 ± 0.7
Costs	$r2\_T*\_*$	8.5 ± 0.4		10.9±0.7	3.2 ± 0.2	2.5 ± 1.2
	$r5\_T*\_*$	23.5 ± 0.4		15.6±0.9	2.8 ± 0.1	14.6 ± 0.7
	$r10\_T*\_*$	24.3 ± 0.4		19.5±1.0	2.9 ±0.1	23.5 ± 0.9
	$r*\_T*\_*$	17.4 ± 0.4		14.3 ±0.8	3.0 ± 0.1	11.1 ± 1.0
Both (Costs | Emissions)	$r2\_T*\_*$	4.8 ± 0.3	20.0 ± 0.8	10.2±0.5	4.3 ± 0.2	1.5 ± 1.0
	$r5\_T*\_*$	19.8 ± 0.4	28.3 ± 0.4	13.9±0.9	4.0 ± 0.1	14.5 ± 0.8
	$r10\_T*\_*$	23.6 ± 0.6	27.8 ± 0.4	18.2±1.1	4.1 ± 0.1	22.7 ± 0.8
	$r*\_T*\_*$	14.3 ± 0.4	24.8 ± 0.5	13.1±0.7	4.1 ± 0.1	10.5 ± 0.9

A Further Results

Table 2 highlights intermodal consolidation and the use of greener transport modes such as ship and rail. It shows that when optimizing for costs or jointly for costs and emissions, consolidation effects are more pronounced when the size of the collaboration increases. Furthermore, internalizing emissions leads to higher usage of rail and ship connections. The small decrease in the measured consolidation effect when optimizing for both objectives compared to costs only can be explained due to the fact that greener intermodal opportunities start to be viable with less transport volume. This is due to the fact that emission reductions compensate high intermodal consolidation requirements which are otherwise necessary so that the chosen intermodal route is cost-efficient. When optimizing for emission the results show that there are a lot of consolidation and intermodal routing opportunities which could drastically cut emissions but are not yet viable from a cost perspective. An emission minimizing solution results in more than double the use of intermodal containers than optimizing for costs only.

Note that Table 2 only shows intermodal consolidation. Another type of consolidation involves shipments between warehouses. These can happen either to send out consolidated shipments or to bring in a large volume of products into a region to satisfy its demands over multiple time periods using one consolidated shipment. Such a large shipment of products then requires storage in a regional warehouse over one or multiple time periods. We find when optimizing for emissions, TUs on warehouse to warehouse connections make up on average 3.7%±4.2%[6](0% - 14.9%) of

[6]In this paragraph in-group standard deviations are reported.

the total number of ordered TUs. When optimizing for costs they make up 2.6%±2.5%(0% - 8.4%) and when optimizing for both they make up 2.4%±2.5%(0% - 8.5%). This small decrease when optimizing for both objectives can be explained by the increased intermodal usage. Without horizontal collaboration (in the direct delivery solutions) this quantity is effectively zero. Lastly, the average delivery time of a shipment on the last arc to a demand node drops on average to 0.3±0.2 days when optimizing for emissions compared to on average 1.2±0.2 days in the direct delivery solutions. Optimizing for costs results in 0.7±0.1 days and optimizing for both objectives results in 0.6±0.2 days. This indicates that products are very often shipped into a demand region by means different to direct delivery by lorry.

B Tariff-Expanded Network

Due to the complex tariff structures employed, if one would define a mixed-integer model on the time-expanded graph $G_{\mathcal{T}}$, the arcs representing transport relations would have highly non-linear cost functions. This would make it difficult to apply established and powerful solution techniques for integer linear programs. To address this problem, two specific concepts are used. First, we introduce a second set of decision variables y representing the number of installed *transport units (TUs)* on each arc. A TU represents a certain amount of capacity bought for product flow on this arc and is a non-negative integer variable. These TUs can be containers, trucks or similar. This concept naturally leads to a capacitated network design formulation but alone is not enough to result into a linear formulation. A similar concept is used by Harks et al. (2016) and more general *units of facility installed* (taken from Crainic (2000)) is common in the network design literature.

Secondly, each simple arc in $G_{\mathcal{T}}$ representing a transport relation is replaced by a more complex graph structure known as *graph gadgets*. These gadgets are again made up of arcs and nodes and - together with the above decision variables - allow to linearly model the non-linear cost structures induced by the employed tariffs. The resulting network $G = (\mathcal{V}, \mathcal{A})$ constructed from $G_{\mathcal{T}}$ is called *tariff-expanded*. Employing tariff-expansion allows to derive a linear mixed-integer model defined on G which is a variant of the fixed-charge network flow problem.

B.1 Graph Gadgets

Graph gadgets were first introduced by Harks et al. (2016) for their employed transportation tariffs. In this section, we extend their approach by deriving graph gadgets for tariffs with constant cost levels t_{const} and for tariffs with piecewise-linear cost levels t_{lin}. We assume the cost levels are discount structures, i.e. having monotonic cost levels $C_n(x,y) \leq C_{n-1}(x,y)$. We start by describing the graph gadget construction for the piecewise-linear case. First, the number of transport containers depending on the flow $\tilde{y}$ is replaced by the decision variable y leading to a cost function $C(x,y)$ depending on both x and y. Therefore, it is possible and done in practice to book more TUs than necessary to get a higher discount.

For the second step, assume one t_{lin} is active on a transport relation. Then, the graph gadget is constructed using the steps outlined in figure 2. At first, the transport relation is replaced by one arc for each cost level $n \in \mathcal{N}$ with cost function $C_n(x,y)$ (see Figures 2a and 2b). Each of these arcs is replaced by the graph structure shown in Figure 2c. Variable y_1 associated to e_1 is set to be binary, i.e. $y_1 \leq 1$. The TU-capacity of e_1 is set to infinity. The costs of e_1 are set to the starting cost of the tariff level this graph structure represents, i.e. $c(e_1) = b_n$. This means any flow over this graph structure has to pay at least the starting costs of this tariff. As these already include the costs of $\beta_n - 1$ fully filled TUs, $c(e_2) = 0$ with $y_2 \leq \beta_n - 1$. As b_n also includes the price of the

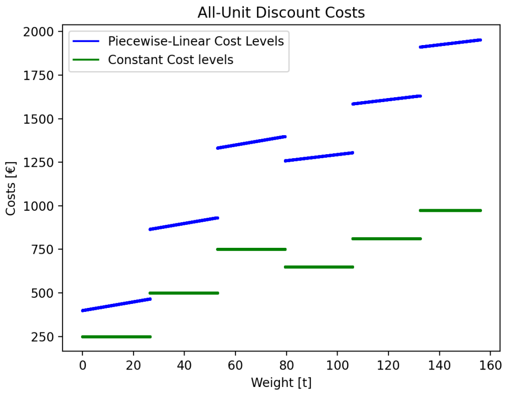
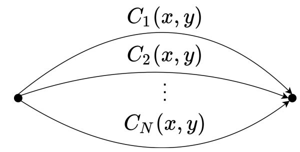
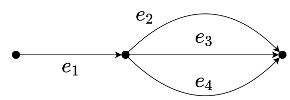


Figure 1: In this work, we use constant cost levels for rail and ship connections and piecewise-linear cost levels for lorry tariffs. The displayed tariffs have one cost level for the first three transport units (TUs) and a second discounted cost level starting with the fourth ordered TU. In this work, one TU always corresponds to one 40-feet ISO container.

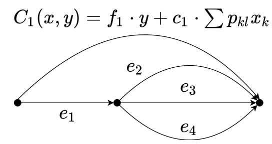
(a) Transport relation with non-linear cost function.



(b) Add parallel arcs for each cost level.



(c) Replace each parallel arc with four arcs and one gadget node.



(d) Simplify graph structure of first tariff level.

Figure 2: Graph gadget construction.

empty β_n-th TU, costs on e_3 are set to the linear costs only $c(e_3) = c_n \cdot P_l^\sigma(x_3)$ and e_3 is restricted to be chosen only once $y_3 \leq 1$. The capacity in property $l \in \mathcal{P}$ of each TU on e_2 to e_4 is set to the physical capacity B_l^σ of one real TU employed on this transport relation. Lastly, the costs of e_4 are set to the cost rate of the new tariff level $c(e_4) = f_n \cdot y_4 + c_n \cdot P_l^\sigma(x_4)$. An optimal solution will automatically first fill arc e_2 then e_3 and only then use e_4. If each tariff level is replaced by such a graph structure, again by an argument of optimality, only this graph structure will be chosen resulting into the minimum cost tariff level for a given flow x. However, the graph structure for the first tariff level can be simplified to only one arc with $C_1(x,y) = f_1 \cdot y + c_1 \cdot P_l^\sigma(x)$ as $b_1 = \beta_1 = 0$. If only a small number of tariff-levels exists, this significantly reduces model size. The graph gadget developed includes a graph gadget for constant cost tariff levels as special case when $c_n = 0$.

From a pure mathematical perspective, it is possible to model the employed tariffs using only two outer edges. However, this would not preserve the number of employed TUs for counting purposes necessary for handling capacity restrictions (see Section B.4).

B.2 Decision Variables

To formulate the objective and constraints, we formally define the decision variables on the network $G = (\mathcal{V}, \mathcal{A})$. Note that the node set $\mathcal{V}$ decomposes into a set of facilities $\mathcal{F}$, demand nodes $\mathcal{D}$, bin nodes $\mathcal{B}$ and gadget nodes $\mathcal{G}$ and that each nodes is uniquely identified by in index-time pair (i, τ). As a result, an arc $a \in \mathcal{A}$ connecting (i, τ) with (j, τ') is identified by a five tuple (i, j, τ, τ', m) with m representing the type of the transport relation. Arcs connecting to and from a gadget node inherit the transport mode m of the original transport relation.

The *flow of non-perishable goods* $f_a^k \in \mathbb{R}_0^+$ is defined on all arcs $a \in \mathcal{A}$ and for each non-perishable commodity $k \in \mathcal{K}_\mathcal{T}$. The *flow of perishable goods* $f_a^{k_t} \in \mathbb{R}_0^+$ is defined on all arcs $a = (i, j, \tau, \tau', m) \in \mathcal{A}$ with $\tau' \leq t + \Delta t_k$ and for each perishable commodity $k \in \mathcal{K}$.. The second decision variable $y_a^\sigma \in \mathbb{N}_0$ describing the *number of transport units* of type $\sigma \in \Sigma$ installed is defined for all $a \in \mathcal{A}$

and $\sigma \in \Sigma$. All flow-variables on an arc a can be collected into a flow vector x_a. These can be again stacked to one big flow vector x. Analogously a TU vector y is constructed.

B.3 Objective

The objective consists of a weighted sum of costs and emissions (for a review on green network design see McKinnon et al. (2015), for green planning techniques in general see Bektaş et al. (2019)). The (linear) cost term includes system-wide transportation, handling and storage costs as given by equation (5). Handling costs are incorporated into the transportation arcs and dependent on the connected facilities and mode of transportation.

$$C(x,y) = \sum_{a \in \mathcal{A}} \sum_{\sigma \in \Sigma} \left[f_a^{\sigma} y_a^{\sigma} + c_a^{\sigma} P_{l(a)}^{\sigma}(x_a) \right] \tag{5}$$

Fixed-costs per transport unit of type σ on arc a are f_a^{σ}. Flow extent in property $l \in \mathcal{P}$ of type σ on arc a is linearly priced by c_a^{σ}. The priced flow-extent property $l \in \mathcal{P}$ depends on the employed tariff. Therefore, it is dependent on the looked upon arc a and written as a function thereof $l(a)$. The emissions term (6) looks very similar, except it replaces fixed-costs with fixed CO_2e emissions Δ_a^{σ} per TU and variable CO_2e emissions δ_a^{σ} per flow extent.

$$\Delta(x,y) = \sum_{a \in \mathcal{A}} \sum_{\sigma \in \Sigma} \left[\Delta_a^{\sigma} y_a^{\sigma} + \delta_a^{\sigma} P_{l(a)}^{\sigma}(x_a) \right] \tag{6}$$

The combined objective reads

$$\min_{x,y} \lambda_1 C(x,y) + \lambda_2 \Delta(x,y) \tag{7}$$

Different choices of λ_1 and λ_2 allow different CO_2e pricing schemes. CO_2e pricing can be ignored by setting $\lambda_2 = 0$. Minimizing for emissions only is enabled by setting $\lambda_1 = 0$.

B.4 Constraints

The model is presented in a cut-set formulation and $\delta^-(v)$ and $\delta^+(v)$ have their usual meaning of all incoming or outgoing arcs, respectively, from node $v \in \mathcal{V}$. Additionally, mode-specific cut-sets are defined for each facility $v \in \mathcal{F}$ with $\delta_m^-(v)$ representing the set of all incoming arcs of type m. One could define $\delta_m^+(v)$ analogously. But to formulate handling capacity constraints in the model, we need to capture the correct number of outgoing transport units of a specific type using the arcs in $\delta_m^+(v)$. In its analogous definition, it will be distorted due to the graph-gadgets. To correct this distortion, in $\delta_m^+(v)$ we don't include the arcs connecting node v to graph-gadget nodes and in return add the outgoing arcs of graph-gadget nodes of type m.

The constraints read as follows:

$$\sum_{a\in\delta^+(v)} x_a^k - \sum_{a\in\delta^-(v)} x_a^k = b_v^k \qquad \forall k \in \mathcal{K}_{\mathcal{T}}, v \in \mathcal{V}\setminus\mathcal{B} \tag{8}$$

$$\sum_{a\in\delta^+(v)} x_a^{k_t} - \sum_{a\in\delta^-(v)} x_a^{k_t} = s_v^{k_t} \qquad \forall k \in \mathcal{K}_\Delta, t \in \mathcal{T}, v \in \mathcal{F}\cup\mathcal{G} \tag{9}$$

$$\sum_{t=0}^{\tau}\sum_{a\in\delta^-(v)} x_a^{k_t} = \omega_v^k \qquad \forall k \in \mathcal{K}_\Delta, v = (i,\tau) \in \mathcal{D} \tag{10}$$

$$\sum_{k\in\mathcal{K}_{\mathcal{T}}^{\sigma}} p_{kl} x_a^k + \sum_{k\in\mathcal{K}_{\Delta}^{\sigma}}\sum_{t=0}^{\tau} p_{kl} x_a^{k_t} \le B_{al}^{\sigma} y_a^{\sigma} \qquad \forall l \in \mathcal{P}, \sigma \in \Sigma, a \in \mathcal{A} \tag{11}$$

$$y_a^{\sigma} \le u_a^{\sigma} \qquad \forall \sigma \in \Sigma, a \in \mathcal{A} \tag{12}$$

$$\sum_{\sigma\in\Sigma}\sum_{a\in\delta_m^{\circ}(v)} y_a^{\sigma} \le h_v^{m\circ} \qquad \forall \circ \in \{+,-\}, m \in \mathcal{M}, v \in \mathcal{F} \tag{13}$$

$$\sum_{\sigma\in\Sigma}\sum_{a\in\delta_m^-(v)\cup\delta_m^+(v)} y_a^{\sigma} \le h_v^{m} \qquad \forall m \in \mathcal{M}, v \in \mathcal{F} \tag{14}$$

The constraints can be grouped into *flow-conservation constraint* (8)-(10) and *capacity constraints* (11)-(14). Constraint (8) defines flow-conservation for non-perishable goods with a positive b_v^k to represent a source for commodity k and negative to represent demand for commodity k. The following two constraints handle the more complex case of perishable goods. Equation (9) concerns the conservation of flow through facility and gadget nodes. For gadget nodes $s_v^{k_t} = 0$ as they have no supply of goods. If a facility acts as a source for commodity k in time period t, $s_v^{k_t} > 0$. The third flow-conservation constraint (10) regards demand satisfaction of perishable goods. Each region has a time-dependent demand ω_v^k for perishable commodity k which must be exactly fulfilled by its incoming flow. Demand for a commodity k can be satisfied by any commodity k_t independent of its production timestamp t.

Constraint (11) links the flow of goods with the necessary transport units. B_{al}^{σ} refers to the maximal extend of property l a TU on arc a for commodities of type σ can transport. For this study on arcs of mode $m \in \mathcal{M}$, B_{al}^{σ} always corresponds to one 40-ft ISO container except in the case when a graph-gadget requires adjusting the capacity. Constraint (12) establishes a maximum number of TUs for type σ installable on arc a. Together with (11) it capacitates the flow on a given arc. Storage is capacitated by setting $B_{al}^{\sigma} = 1$ and u_a^{σ} to the real world warehouse capacity for this property. Equation (13) capacitates incoming ($-$) and outgoing ($+$) handling operations separately. This captures the operational truth in cross-docks with separated incoming and outgoing docks or conceptually similar architectures like transshipment-points with mode-switches. Lastly, (14) capacitates the sum of containers handled both for incoming and outgoing operation. This is relevant for facilities like a classic warehouse with a specific number of docks for lorries not strictly split into incoming and outgoing docks. During handling commodity types are not distinguished (exemplary, a container crane is oblivious to the fact that a specific container has a cooling module or not).

Note that we have devised gadget-specific strengthening constraints for the mixed-integer formulation presented. However, they only result in small improvements on the solution performance not enough to solve most of the large instances. Therefore, we omit their presentation.

C Slope Scaling - Formal Model

The linear programming objective takes the form

$$v(\mathcal{LP}(\rho(t))) = \min_x \sum_{a\in\mathcal{A}} \sum_{\sigma\in\Sigma} \Big[\big(\lambda_1 c_a^\sigma + \lambda_2 \delta_a^\sigma + \rho_a^\sigma(t)\big) \cdot P_{l(a)}^\sigma(x_a)\Big] \tag{15}$$

with x defined as in the original problem (compare to the $\mathcal{MIP}$-objective in Appendix B.3). The linear program uses the same flow-conservation constraints (8) - (10). The other constraints are adapted as follows

$$\sum_{k\in\mathcal{K}_\mathcal{T}^\sigma} p_{kl} x_a^k + \sum_{k\in\mathcal{K}_\Delta^\sigma} \sum_{t=0}^{\tau} p_{kl} x_a^{k_t} \le B_{al}^\sigma u_a^\sigma \qquad \forall l\in\mathcal{P}, \sigma\in\Sigma, a\in\mathcal{A} \tag{16}$$

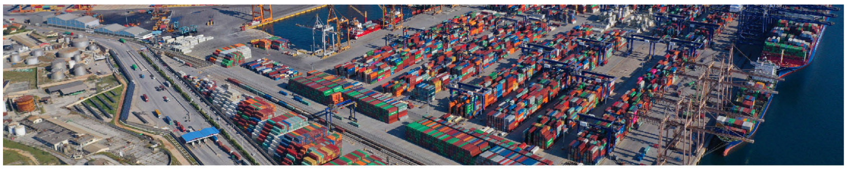
$$\sum_{\sigma\in\Sigma} \sum_{a\in\delta_m^\circ(v)} P_l^\sigma(x_a) \le B_{ml}^\sigma h_v^{m\circ} \qquad \forall l\in\mathcal{P}, \circ\in\{+,-\}, m\in\mathcal{M}, v\in\mathcal{F} \tag{17}$$

$$\sum_{\sigma\in\Sigma} \sum_{a\in\delta_m^-(v)\cup\delta_m^+(v)} P_l^\sigma(x_a) \le B_{ml}^\sigma h_v^m \qquad \forall l\in\mathcal{P}, m\in\mathcal{M}, v\in\mathcal{F} \tag{18}$$

Constraint (16) merges (11) and (12). The adapted handling constraints (17) and (18) use the mild assumption of same TU property-extents B_{ml}^σ on one mode m, valid for all our input instances. An exact solution $\tilde{x}$ of $\mathcal{LP}(\rho(t))$ can be calculated efficiently by linear programming algorithms. If an instance has no capacitating handling constraints, $\mathcal{LP}(\rho(t))$ takes the form of a multicommodity minimum-cost flow problem and we find for this case the *network simplex* algorithm (Bertsimas and Tsitsiklis, 1997) outperforms more general simplex algorithms.

As described in the main text, after solving $\mathcal{LP}(\rho(t))$, a solution $(\tilde{x},\tilde{y})$ to $\mathcal{MIP}$ is obtained by setting $\tilde{y}_a^\sigma = \max_{l\in\mathcal{P}}\{\lceil \frac{P_l^\sigma(\tilde{x}_a)}{B_{m(a)l}^\sigma} \rceil\}$ ($m(a)$ denotes the mode used by arc a). Implicitly estimating the number of TUs by summing up flow-extent over multiple arcs as done in (17) and (18) could lead to minor differences with the actual number of used TUs. If in instances with handling constraints some are slightly violated, we repair the solution by rerouting flow exceeding handling capacities using a variant of our local search moves (see Section 3.2).

Furthermore, we apply slope scaling on the much more memory-efficient time-expanded network by replacing $a\in\mathcal{A}$ with $a\in\mathcal{A}_\mathcal{T}$ and in each iteration adapt the cost factors c_a^σ and f_a^σ to the best available tariff level for the chosen $\tilde{y}_a^\sigma$. To the best of our knowledge, this results into the first application of a slope scaling procedure to a changing cost function.



Improving Demand Prediction and Reducing Out-of-Stock – Application of Advanced Data Analytics in Retail Supply Chains

Brandtner Patrick[1], Darbanian Farzaneh[1], Falatouri Taha[1], Udokwu Chibuzor[1]
1. University of Applied Sciences Upper Austria, Steyr, Austria
Corresponding author: patrick.brandtner@fh-steyr.at

Abstract: *Correct demand prediction is a key success factor of efficient and demand-driven supply chains. This is especially true for the retail sector, where out-of-stock products directly influence customer satisfaction. In a data-driven world, advanced analytics approaches offer huge potential for demand prediction. The paper applies two of the most acknowledged demand prediction approaches to a real-world retail case. Based on an extensive database of cashier as well as Supply Chain data, we apply ARIMA and SARIMA and evaluate their applicability to predict the demand of selected perishable products. In addition, the impact of adding SARIMA-based demand forecasting to out-of-stock detection is analysed. The results show high applicability and a good forecasting quality especially of SARIMA. The quality of out-of-stock detection can significantly be improved by adding advanced analytics to traditional approaches in this area. For reaching higher demand prediction quality, results indicate the need to add the effect of promotions and the implications of substitute products to the applied approach.*

Conference Topic(s): *Area 2 - Distributed Intelligence in Physical Internet: ML and AI tools for demand forecasting and inventory control; Area 2 – Systems and Technologies for interconnected logistics: Machine Learning, Big Data and Artificial Intelligence.*

Keywords: *Data Analytics, Demand Prediction, Supply Chain Management, Big Data, Supply Chain Optimization, Value Network, Product Availability, Retail Management*

1 Introduction

The retail sector is characterized by a high degree of demand uncertainty, volatile customer requirements and fierce competition by online retailers like Amazon. Having efficient and cost-optimized operational processes has become a key success factor for retailers, especially in the era of digitalization and the data-driven economy. Exact demand forecasting and optimal stock management are core abilities of every retail company and contribute significantly to efficient retail operations and subsequently to product availability and customer satisfaction (Pereira and Frazzon, 2020; Mandal 2020). Advanced data analytics based on Machine Learning (ML) and Artificial Intelligence (AI) offers huge potential to support and improve demand forecasting quality and correctness (Huber and Stuckenschmidt, 2020). However, for retail practitioners, the application of advanced data analytics is a challenging task and most approaches to demand prediction are done based on conventional statistical approaches and do not take into consideration additional factors on the level of the single stores and the specific conditions at this level (Lalou et al., 2020).

This research paper focusses on this issue and provides insights into the findings of a work package of a large research project conducted in the research field "Data Analytics and Foresight" at the department for logistics (Logistikum) at the University of Applied Sciences Upper Austria. The project partner is one of the main retailers in Austria with stores all across the country. The project's aim is to implement data analytics in retail Supply Chains. More precisely, the focused work package aims at improving the accuracy of demand prediction and out-of-stock detection for single stores by applying Machine Learning-based algorithms to optimize inventory control and demand forecasting for perishable items and products. The focal research question of the paper is hence defined as follows:

How can advanced analytics and Machine Learning-based approaches be applied in retail supply chains to improve demand forecasting and out-of-stock detection?

Based on an extensive analysis of cashier and Supply Chain data from the research partner company and for a selected region and a defined time span in Austria, the paper analyses two selected demand forecasting approaches and compares existing with improved out-of-stock detection methods. The remainder of the paper is structured as follows: section two provides an overview of forecasting methods and a discussion of innovative advanced analytics approaches and their benefits. Section three i) presents the data basis of the paper, ii) provides details regarding the approach used to fill missing data and iii) offers an overview of the selected products and their demand as the focus area of the paper. Section 4 discusses the results of applying two selected demand prediction approaches to the given retail case and shows the impact of this on out-of-stock detection. Finally, section 5 concludes the paper and provides and overview of its limitations and possible future research opportunities.

2 Demand Prediction and Out-of-Stock Detection in Retail SCM

Demand forecasting is one of the challenges that play an essential role in decision making in retail SCM (Da Fonseca Marques, 2020). However, it is not possible to calculate the real and actual customer demand in food retail stores because customers cannot place orders and can only buy products available at the point of sale. Therefore, product sales are considered as the real demand in the retail sector in general and in the current paper (Arunraj et al., 2016).

Inaccurate demand forecasting can cause understock or overstock problems. Understocking results in out-of-stock that ruins company's image and customers' trust. Overstocking does also bring negative effects on business, such as lack of storage space and waste of perishable products. Accurate demand forecasting helps businesses to increase their profit by managing their inventory and reducing out-of-stock. Data quality, data availability and forecast horizon are three factors that affect the accuracy of demand forecasting (Arunraj et al., 2016). Especially in the food retail sector, demand forecasting is crucial as perishable food products have short shelf lives and reducing both out-of-stock situations as well as outdating is important (Dellino et al., 2015).

Demand forecasting can be either short-term (between six to twelve months) or long-term (more than one year) and can be calculated with qualitative or quantitative methods. Quantitative methods are based on mathematical calculations. They aim to discover data patterns and provide accurate results. Machine learning techniques are among quantitative methods that can deal with big data with high processing speed and provide accuracy, adaptability and transparency (G and Prakash, 2020). Forecasting machine learning techniques can be classified into three groups. First, time series analysis such as autoregressive integrated moving average (ARIMA) and Holt Winters Exponential Smoothing (HW). Second, regression based methods such as Linear Regression (LR) and Support Vector Regression (SVR). And third, supervised and unsupervised methods like Support Vector Machine (SVM), Artificial Neural Network (ANN), Clustering and Hybrid modes (G and Prakash, 2020).

Time series analysis is the most popular approach for demand forecasting in retail industry due to its ability to capture trend and seasonality (Arunraj et al., 2016; Da Fonseca Marques, 2020). Shukla and Jharkharia (Shukla and Jharkharia, 2013) applied ARIMA models to predict daily sales of onion in a wholesale market is India. Based on their results, Monday seems to be a day with higher demand than other weekdays for a certain product. Evers et al. (Evers et al., 2018) stated that ARIMA models are able to capture seasonal effect of demand and perform better than simple linear regressions. They compared tree models with linear regression models and proved that the first group outperforms the second with more than 99 percent accuracy to predict daily sales of bread in an online supermarket in Netherlands. Besides ARIMA, seasonal ARIMA (SARIMA) is one of the most used demand forecasting approaches in the retail context (Arunraj et al., 2016; Roque et al., 2019). The current paper focusses on time-series analysis and more precisely on applying ARIMA and SARIMA to the food retail context to improve demand prediction as well as out-of-stock detection.

3 Data Basis

In this section, the data basis, the approach used to clean and prepare the data for the analysis as well as an overview of the selected data is presented.

3.1 Data Basis of the Paper

The data available for the analysis covers more than 17 million records related to cashier data of three perishable products for more than 90 stores in Austria from January 2017 to December 2019. Sample data for this study consists of aggregated daily sales data of stores (cashier data) that were consistently open within this time period (66 stores). Therefore, the aggregated daily sales dataset for these three products consists of more than 170,000 records. In addition, around 10,000 calculated missed data elements (about 5% of total records) have been added to the database following the formula in section 3.2. In addition we used operational Supply Chain data to fill missing elements in cashier data when needed. This SC data mainly included shipping records from distribution center to stores.

3.2 Filling missed data

Running a reliable statistical analysis requires complete and full data. However, missing values are an inevitable situation in big data series. To fill missing data elements, we used alternative data from operational Supply Chain process (shipping data from distribution center to store) if available. If not available, there are several methods to complete missing data in time series analysis such as likelihood methods and neural networks or total random methods. In this research, we have used Simple Arithmetic Average (SAA Simple arithmetic average) (Duan et al., 2013).

$$\hat{Y}_{m_k t} = \frac{1}{N}\left(Y_{m_k 1} + Y_{m_k 2} + \cdots + Y_{m_k N}\right) \tag{1}$$

In the formula (1), $\boldsymbol{\hat{Y}_{m_k t}}$ indicates calculated missed data, N represent the number of available homologous days and $\boldsymbol{Y_{m_k N}}$ are available data. If there is data missed in one special date in one year data, the homologous day for other years has been considered as the input of the formula to calculate estimation for the specific day.

In addition to missed data there are some noises in the data that could lead to forecasting results with lower accuracy. These noises may come either from external sources like promotions or holidays (Arunraj et al., 2016) or from logistic effects, bullwhip effects or temporal aggregation (Murray et al., 2018).

3.3 Overview of Data Selected for Demand Prediction

The data used in this paper focused on three selected perishable food items form the category of fruits and vegetables: i) cucumber, ii) salad and iii) avocado. The following figure provides an overview of the aggregated actual sales of the items in the selected 60 stores within the given time frame:

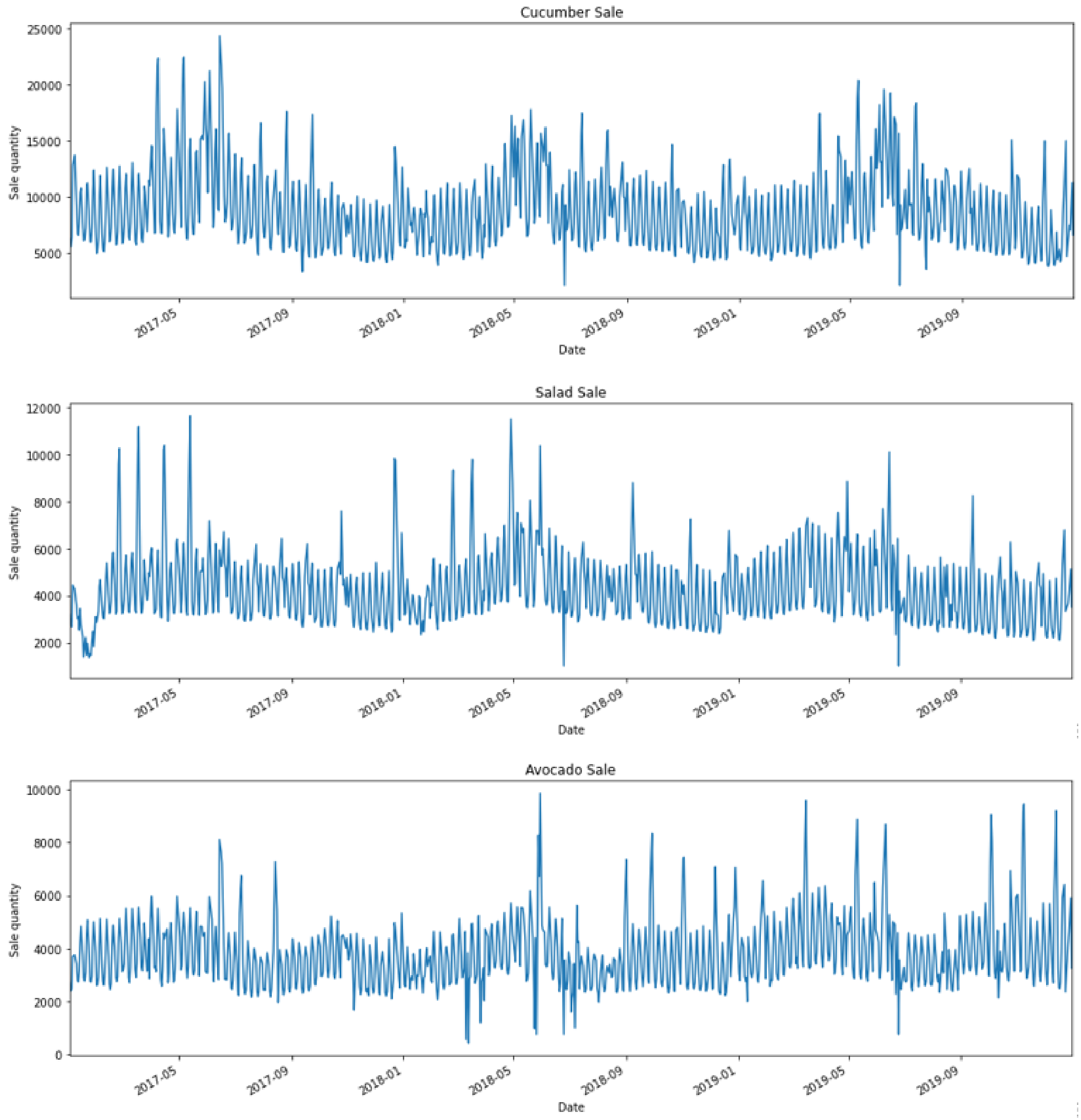


Figure 1: Aggregated actual sales of selected products

4 Results of Demand Prediction and Out-of-Stock Detection

Based on the selected products, ARIMA and SARIMA were analysed in terms of their applicability to predict demand. The results of the comparison of these two approaches is presented in section 4.1. Subsequently, section 4.2. discusses the effects of adapting out-of-stock detection based on the demand forecast by applying SARIMA.

4.1 Comparing ARIMA and SARIMA

To predict daily sales of selected stores in January 2020, the two most well-known time series algorithms ARIMA and SARIMA have been applied to historical sales data from January 2017 to December 2019. The two models have been evaluated with calculating the mean absolute percentage error (MAPE) (Lewis, 1982) based on the following formula:

$$\boldsymbol{MAPE} = \frac{\mathbf{1}}{\boldsymbol{n}}\sum_{\boldsymbol{t}=\mathbf{1}}^{\boldsymbol{n}} \frac{|\hat{y}_t - y_t|}{y_t} \tag{2}$$

In the formula (2) n represents the total number of records, $\boldsymbol{y_t}$ shows the observed value at time t and $\hat{y}_t$ indicates the predicted quantity. The difference between these two shows the deviation between reality and expectation. Dividing the summation of mentioned deviation shows the average deviation of expectation and reality. The less this number is the more accurate forecasting results are. The following table provides a framework for the interpretation of the calculated average (Lewis, 1982).

Table 1: Interpretation of MAPE.

MAPE	Interpretation
<10	Highly accurate forecasting
10-20	Good forecasting
20-50	Reasonable forecasting
>50	Inaccurate forecasting

In order to forecast the data we have used SPSS Modeler 18.2. with time series node. Table 2. illustrates the prediction results. The final model has been selected with Akaike (AIC) and Bayesian Information Criterion (BIC). As shown, SARIMA model performs better in all products with provided p, d and q parameters. The parameter p shows the number of time lag in auto regressive models (AR), d indicates the degree of differencing and q is the moving average order. For SARIMA, parameters P, D, Q represent the corresponding seasonal parameters to p, d and q. Although forecasting results are satisfying according to MAPE, there are some noises in the data which are not considered in our prediction model. This includes promotions and holiday effects that made irregular deviations from the fitted model. Hence, it is suggested to consider these parameters as external factors in future research.

Table 2: Comparing ARIMA and SARIMA results to predict sales from Jan. 2017 to Dec. 2019.

	AVOCADO ARIMA	AVOCADO SARIMA	CUCAMBER ARIMA	CUCAMBER SARIMA	SALAD ARIMA	SALAD SARIMA
p,d,q	p=0,d=0,q=7	p=3,d=0,q=3	p=0,d=0,q=7	p=0,d=0,q=4	p=0,d=0,q=10	p=0,d=0,q=2
P,D,Q	-	P=1,D=1,Q=1	-	P=1,D=1,Q=1	-	P=1,D=1,Q=1
RMSE	1,023.306	909	2,506.811	2,094.040	1,102.381	953
RMSPE	50	48	34	24	32	27
MAE	730	544	1,865.378	1,346.918	807	608
MAPE	**23**	**17**	**23**	**16**	**20**	**15**
MAXAE	5,451.908	5,363.323	15,856.889	16,736.227	5,416.038	5,520.990

MAXAPE	678	762	443	241	427	372
AIC	13,009.143	12,704.178	14,688.995	14,258.716	13,148.781	12,790.952
BIC	13,043.049	12,733.202	14,718.058	14,282.902	13,182.688	12,810.302

As shown in table 2, SARIMA provides a better approach to forecast the demand of the three selected products. The MAPE for SARIMA was found to be 15 for salad, 16 for cucumber and 17 for avocado. Following Lewis (1982) and considering the values provided in table 1, the forecast quality of SARIMA can hence be considered as "good forecasting". For ARIMA, MAPE is between 20 and 23, which represents a "reasonable forecasting" quality.

The following figure 2 compares the actual sales (blue line) with the predicted sales (orange line) based on the better performing SARIMA approach for cucumber, salad and avocados in January 2020. The high peaks for cucumber and salad (marked with red circles) represent the effect of promotions, which resulted in higher sales than forecasted:

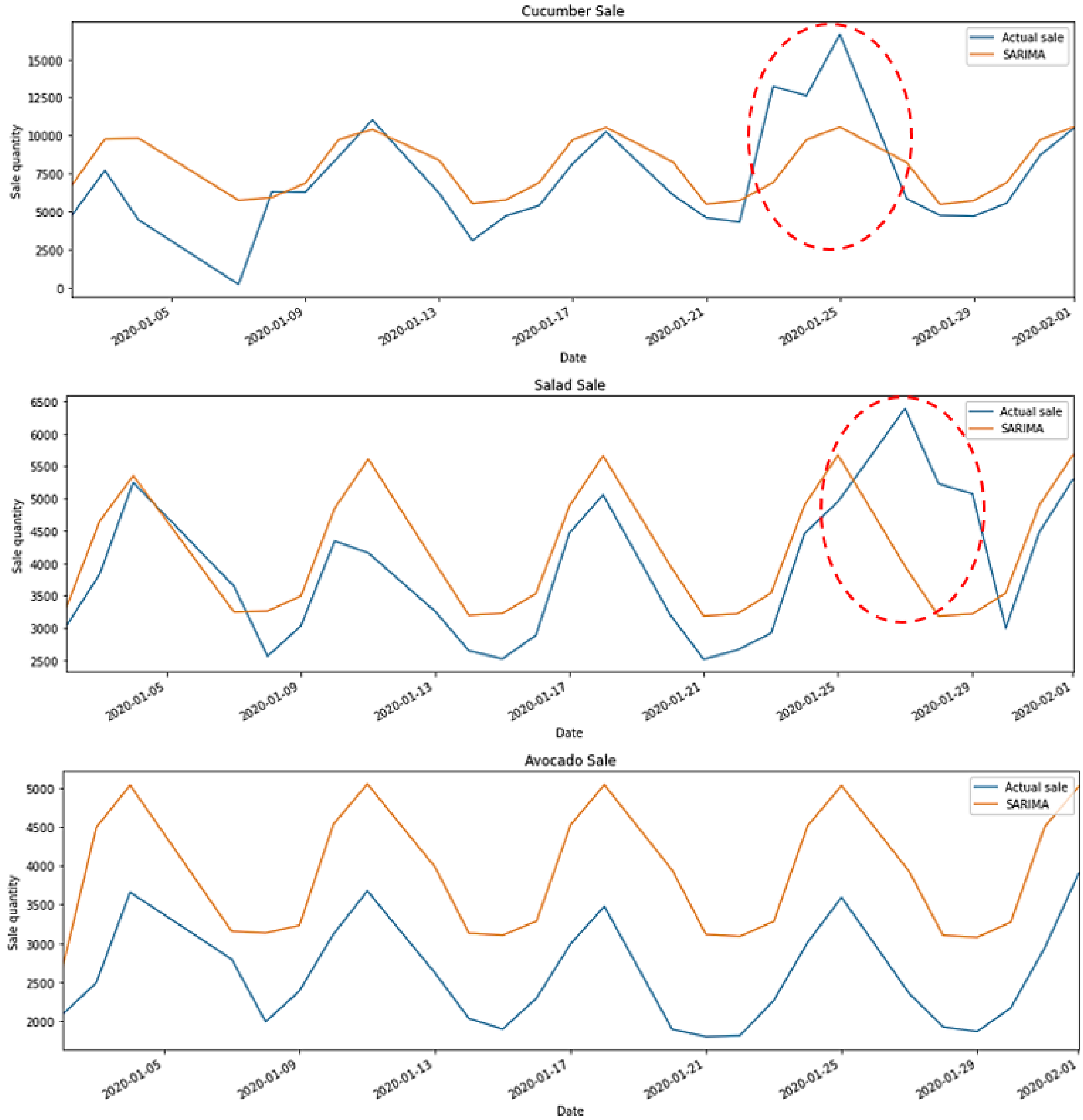


Figure 2: Predicted vs. actual sales of products in January 2020

Using the output of the forecasting model to test the data of the training records had satisfying results for cucumber and salad. However, for avocados (lowest visualization in figure 2) the results showed to be rather disappointing which raised the question of what external factors may have influenced the sale of this product in January. We paid particular attention to substitution effects between products as shown in table 3.

Table 3: Substitution effect of avocados.

Product	December 2019	January 2020
Avocado – sub1	20051	32569
Avocado – sub2	42082	46250
Avocado - target	103349	96141
Total	165482	174960

The table shows that while the total sale of avocados in January is higher than in the previous month, the trend of sale in the targeted product (i.e. the selected type of avocado) was declining. At the same time, alternative products (different pack-sized avocados as well as organic avocados) experienced a dramatic growth in sale. Hence, we suggest considering the replacement effect between products in future forecasting approaches.

4.2 Out-of-Stock Detection

Traditional methods for detecting out of stock situations were based on sale gaps in the last hours of a working day in a store. This approach has two main drawbacks: First, out-of-stock situations are considered to be the result of insufficient supply from distribution center to store only. Operational out-of-stock factors at the store level may not be detected, e.g. the fact that store personnel didn't manage to put products from storage to shelf. This is especially the case for product unavailability at shelves during the working hours. Second, the traditional out-of-stock detection method only works as a flag which indicates if an out-of-stock occurred or not. It does not allow for estimating how many product sales were potentially missed.

Based on the forecasted sales minus MAPE we identified out of stock situations as well as their amount. More precisely, the differences between expected sales and real sales (if higher than the average amount of MAPE) were used to determine out-of-stock and their respective amount. As shown in table 4, the proposed model identified more out-of-stock situations than the traditional approach. However, the main advantage of this adapted method is that it allows for estimating missed sale quantities.

Table 4. Out of Stock results for one selected store for January 2020.

Product	Out-Of-Stock situations Traditional method	Out-Of-Stock situations Forecasting method	Missed sale	Ave daily sale
Cucumber	1	2	221	157
Salad	1	9	168	79
Avocado	1	5	104	45

5 Conclusions and Outlook

The results of the paper indicate that there are big benefits of applying advanced analytics in retail demand forecasting and out-of-stock detection. Compared with the traditional out-of-stock detection approaches used in the given retail case, the applied improved model identified a far higher amount of critical out-of-stock situations at the point-of-sale. The improved out-of-stock detection approach does not only focus on the overall sold quantity and the sales gaps (i.e. times without sales) at the last hours of the day but also includes the forecasted daily sales of stores. Regarding the improvement of demand forecasting, time series forecasting and more precisely auto-regressive integrated moving averages (ARIMA) and seasonal auto-regressive integrated moving averages (SARIMA) was used to better predict the sales of the single stores on daily levels. In general, the results show good applicability of the applied approaches, especially SARIMA seems to be applicable and offers a good forecasting quality.

The practical implications of the paper are as follows: first, practitioners get an overview of how demand forecasting and out-of-stock detection is currently conducted in retail SCM practice. Second, innovative advanced analytics and Machine Learning-based approaches for improving these two tasks are discussed, providing practitioners with an overview of potential future improvements. Third, an actual practice use case from the Austrian retail sector is provided as a best practice example on how to improve demand forecasting and out-of-stock detection by means of Machine Learning-based approaches. This best practice example shows how advanced analytics can be implemented and adopted to a real-life use case in the retail sector. From a scientific point-of-view, the paper provides new insights in terms of presenting and discussing the applicability of time series forecasting with ARIMA and SARIMA in retail practice. For distributed intelligence in physical internet setting, the results of the paper provide the basis for establishing a more efficient demand planning across the network of retail stores, retail distribution centers, regional as well as trans-regional product suppliers and logistics serviced providers. By sharing the forecasted demand levels in an interconnected logistics network, out-of-stock situations, food wastage and unfulfilled customer demands can be reduced respectively avoided. In addition, transportation as well as warehouse infrastructure can be planned closer to real demand at an early stage, resulting in a better utilization of resources and the avoidance of express orders and shipping. Ultimately, these effects also contribute to lower emissions and lower product wastage due to expired items.

The limitation of the paper is the focus on one specific retail case in Austria. However, by having an extensive data base from actual sales at the point-of-sale as well as data along the retail Supply Chain provided a solid basis for data analysis and demand forecasting as well as out-of-stock detection improvement and evaluation. Future research should focus on adding and analysing the effect of substitute products as well as the effect of promotions on SARIMA-based prediction approaches.

Acknowledgements

This research was partly funded by the government of Upper Austria as part of the Excellence Network Logistics – Logistikum.RETAIL.

References

- Arunraj N. S., Ahrens D., Fernandes M. (2016): *Application of SARIMAX model to forecast daily sales in food retail industry*. International Journal of Operations Research and Information Systems (IJORIS), v7, no2, 1–21.
- Chen K.; Wang C. (2007): *A hybrid SARIMA and support vector machines in forecasting the production values of the machinery industry in Taiwan*. Expert systems with applications, v32, no1, 254–264.

- Da Fonseca M., Rúben A. (2020): *A Comparison on Statistical Methods and Long Short Term Memory Network Forecasting the Demand of Fresh Fish Products*. PhD-Thesis, Universidade do Porto.
- Dellino G., Laudadio T., Mari R., Mastronardi N., Meloni C. (2015): *Sales Forecasting Models in the Fresh Food Supply Chain*. Proceedings of ICORES 2015, 419–426.
- Duan W., Cao Q., Yu Y., Levy S. (2013): *Mining Online User-Generated Content: Using Sentiment Analysis Technique to Study Hotel Service Quality*. Proceedings of the 46th Hawaii International Conference on System Sciences: IEEE, 3119–3128.
- Evers J. M., Tavasszy L., van Duin R., Schott D., Gorte F. (2018): *Demand forecast models for online supermarkets*. Proceedings of Nectar Cluster 2: E-groceries, digitalization and sustainability: Which governance, planning and regulation mix do our cities need?, 12.
- G S.F., Prakash N. (2020): *Machine Learning in Demand Forecasting - A Review*. Proceedings of the 2nd International Conference on IoT, Social, Mobile, Analytics & Cloud in Computational Vision & Bio-Engineering (ISMAC-CVB 2020), 9.
- Huber J.; Stuckenschmidt H. (2020): *Daily retail demand forecasting using machine learning with emphasis on calendric special days*. International Journal of Forecasting, v36, no4, 1420-1438.
- Lalou P., Ponis S.T., Efthymiou O. K. (2020): D*emand Forecasting of Retail Sales Using Data Analytics and Statistical Programming*. Management & Marketing, v15, no2.
- Lewis C. D. (1982): *Industrial and business forecasting methods: A practical guide to exponential smoothing and curve fitting*. Butterworth-Heinemann.
- Mandal P. C. (2020): *Retailing Trends and Developments - Challenges and Opportunities*. International Journal of Business Strategy and Automation (IJBSA), v1, no2, 1–11.
- Murray P. W., Agard B., Barajas M. A. (2018): *ASACT-Data preparation for forecasting: A method to substitute transaction data for unavailable product consumption data*. International Journal of Production Economics, v203, 264–275.
- Pereira M. M., Frazzon E. M. (2020): *A data-driven approach to adaptive synchronization of demand and supply in omni-channel retail supply chains*. International Journal of Information Management, 102165.
- Roque L., Fernandes C. A. C., Silva T. (2019): *Optimal Combination Forecasts on Retail Multi-Dimensional Sales Data*. arXiv preprint arXiv:1903.09478.
- Ruiz-Aguilar J. J.. Urda D., Moscoso-López J. A., González-Enrique, J., Turias I. J. (2020): *Container Demand Forecasting at Border Posts of Ports: A Hybrid SARIMA-SOM-SVR Approach*. In : Bernabé Dorronsoro, Patricia Ruiz, de la Torre, Juan Carlos, Daniel Urda, El-Ghazali Talbi (Eds.): Optimization and Learning. Cham: Springer International Publishing, 69–81.
- Shukla M., Jharkharia S. (2013): A*pplicability of ARIMA models in wholesale vegetable market: An investigation*. International Journal of Information Systems and Supply Chain Management (IJISSCM), v6, no3, 105–119.
- Svetunkov I., Boylan J. E. (2020): S*tate-space ARIMA for supply-chain forecasting*. International journal of production research, v58, no3, 818–827.

Port Digitalization Through an Activities Scenario Model as a First Step for a Digital Twin of Port

Charles Garnier[1], Erwan Simon[1], Joao Pita Costa[2], Leonidas Pitsikas[3], Ignacio Lacalle[4], Carlos E. Palau[4]

1. Centre Aquitain Des Technologies De l'Information Et Electroniques (CATIE), France
2. Razvoj Programske Opreme In Svetovanjel (XLAB), Slovenia
3. PEOPLE, United Kingdom
4. Dpto. De Comunicaciones, Universitat Politècnica de València (UPV), Spain

Corresponding author: c.garnier@catie.fr

Keywords: *H2020 project, Digital Twin, Port of the Future, Port Management, Decision Support System, Simulation, Supply Chain, Simulation, Innovation Web Platform*

Conference Topic: *Systems and technologies for interconnected Logistics, PI Modelling and Simulation*

1 Introduction

Building on the momentum of the digitalisation of the maritime industry, the digital twin is arriving to the European ports and terminals to optimize operations and reduce costs. The technical term digital twin appears after 2010 [1] as a dynamical model which, given the current state of an observed system, is capable of a partial digital reconstruction of such a system. It got widely adopted throughout the years, especially in the context of IoT technologies and the industry 4.0 [2], as well as in healthcare to appropriately address the personalised medicine paradigm [3]. In the recent years there was a natural adoption of the approach by port authorities and container terminals [4], with the engagement of the main European ports in collaboration with the technology giants to explore several dimensions of it in this new context [5]. Though, the problems addressed by the port are much different than those in the digitalisation of a factory and, thus, need to be faced differently [6]. Data driven digital twins are still not in a general industrial practice due to the lack of AI know-how and possibly lack of relevant IoT data to reconstruct the underlying physical processes. The current marketplace for generic Digital Twin technology is not yet very mature, with key providers positioning between Bosh, IBM, Siemens, and General Electric. Most products are centred around internal businesses of the corresponding companies related to mostly IoT or manufacturing. The engineering paradigm of the Digital Twin arrives with: (i) online sensors becoming cheaper and ubiquitous; (ii) improved usefulness of Big Data analytics, processed for patterns and monitored for signals; (iii) the vast remote computing resources in the Cloud making them inexpensive and more accessible [7]. This allows businesses to reorganize organisational processes and workflows towards an improved cost-effectiveness deriving from the eminent digitalisation of the industry.

Modern ports face problems revolving around the issues of efficiency, environmental and financial sustainability. The difficulties they are facing are common with those of other nodes in the logistics chain and have to do (usually) with underutilization of resources: while there are empty warehouses and idle machinery at one given moment, at another moment there are demand peaks they cannot accommodate. One of the other challenges the ports are facing nowadays is that of the social integration. Major European cities have been developed from ancient times around ports as this eased the logistics of their time. However, in our era two issues are arising: i) high volumes of cargo entering and exiting the ports from the hinterland side, thus adding to traffic in cities, ii) usage of fossil fuels for energy production and / or machinery operations adding to atmospheric and sound pollution. An additional issue connecting the two previously mentioned problems is fluctuating employment in times of fluctuating supply and demand, which is not limited only to port workers, but also to activities related to the port industry. A great number of small and medium sized ports are not sufficiently equipped to utilize data already available or easily obtainable to face previously mentioned challenges. Examples of such data are i) vessel calls data which can assist in preparations and scheduling of energy peak demands, ii) usage of environmental sensors which can assist in normalizing emissions through better scheduling of activities, iii) usage of city traffic data which can assist in normalizing port traffic generation and iv) in the future exploitation of IoT enabled sensors in networked containers that can assist in better prioritization of port activities.

PIXEL platform allows to build a useful “what-if” scenario of the port activities which can be seen as an important contribution to the Digital Twin of the port paradigm. Thanks to an open-source tools called the Port Activity Scenario (PAS) based on vessel calls and use of handling equipment specifications and supply chain, PIXEL has allowed to establish an operational description of the port activities related to cargo handling. This description is composed of a set of data-model listing all the considered activities’ time series. These PAS outputs are then use as inputs for energy model or for the quantification of pollutants emissions. PAS model has been build considering needs and constraints of small and medium ports; thus becoming adaptable in terms of data availability, i.e.,working with a minimum set of data providing results of corresponding level of confidence.

2 PIXEL Platform a first step for a Digital Twin of Port

A Digital Twin is a digital representation of an existing physical element aimed at modelling and monitoring its behavior and status. This “digital copy” is unique and has direct communication with the actual element to remain updated, generally in a periodic fashion. The use of the Digital Twin concept offers a myriad of opportunities, focused on decision making, monitoring, predictive maintenance and others that bring clear benefits such as productivity optimization, efficiency and short-, medium- and long-term planning. When trying to characterize a Digital Twin, the main principles are: i) it is individualized, mapping 1 to 1 to a physical element, ii) it can simulate – if properly designed – the behavior of an element with high accuracy levels, but needs to be continuously maintained and updated, iii) the digital model can respond with a very low response time (latency) if and only if the proper technologies are used (IoT, 5G…) and iv) changes in one twin affect to the other, creating a closed-loop achieving physical-digital convergence.

The idea behind the concept is to allow workers (at all levels of the production chains) interact with digital copies as if they were acting over the physical elements, improving efficiency, optimizing costs and drastically reducing health risks at work. After performing a literature analysis, it is commonly accepted that the “Digital Twin” is not delivered by a single technology, nor even by a single technical domain, but rather it is an amalgam of disciplines that shall be put together on many ways and remain valid. Therefore, according to the previous, a “combination” of technologies is needed to deploy a realistic, successful Digital Twin of a physical element. In Figure 1, there is presented a summary of this multi-disciplinary view of the Digital Twin concept.

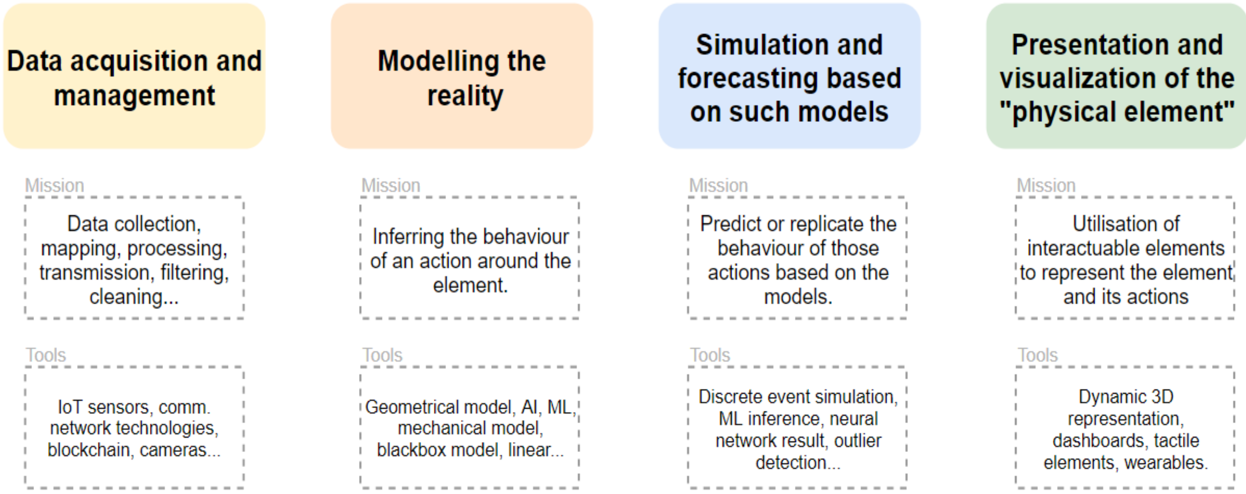


Figure 1: Digital Twin multi-disciplinary summary

Here, PIXEL comes into play. The on-going EC H2020-funded action has created the first modular platform combining strong methodologies and smart technologies for small and medium port ecosystems enabling optimization of operations through IoT while reducing environmental impact. It has been tested in four real ports (Bordeaux, Monfalcone, Thessaloniki and Piraeus) with great results in terms of digitalization and process modelling. Technologically (see Figure 2), it consists of a series of modules (services) framed within four “big blocks” (Data Acquisition Layer, Information Hub, Operational Tools and Dashboard) located at different layers (divided in three), wrapped under an ell-encompassing security framework. All the pieces of the architecture have been built using open-source technologies. The NGSI agents (in the Data Acquisition Layer) perform the actual data collection from the various sources, converting them into actionable information following a specific data format (based on the FIWARE KPI data model [8]) and sending it to an IoT Context Broker (FIWARE ORION [9]). In the case of the Port Activity Scenario (PAS) – of study in this paper –, the data collected is related to the vessel calls (announcing the ships to be operated by the port within the next few days) and the port parameters configuration. Then, these context data are stored inside the Elasticsearch database of the Information Hub for persistence. At this point, the data are ready for the simulation (using the PAS model), which is orchestrated via the Operational Tools (scheduler, controller, manager). This module manages to execute the simulation and defines the storage of the results on the proper place to be represented to the user through the Dashboard User Interface (UI).

As it has been exposed, the whole technological provision of PIXEL could play a role towards the “Digital Twin” of a port, however this article is centered in the Port Activity Scenario (PAS). In the structure of the architecture presented, PAS is conceived as a model that aims at simulating the terminal operations of a port after a series of inputs, priorities and contextual conditions. In particular the PAS might be considered as the “meeting point” between the Physical Internet theory (see Section 4.2) and the Digital Twin concept just outlined. While it aims at representing a “physical element” (that in this case would be a maritime port terminal), it remains a node within the logistic supply chain and could be leveraged towards the optimization and standardization goals of the Physical Internet realm.

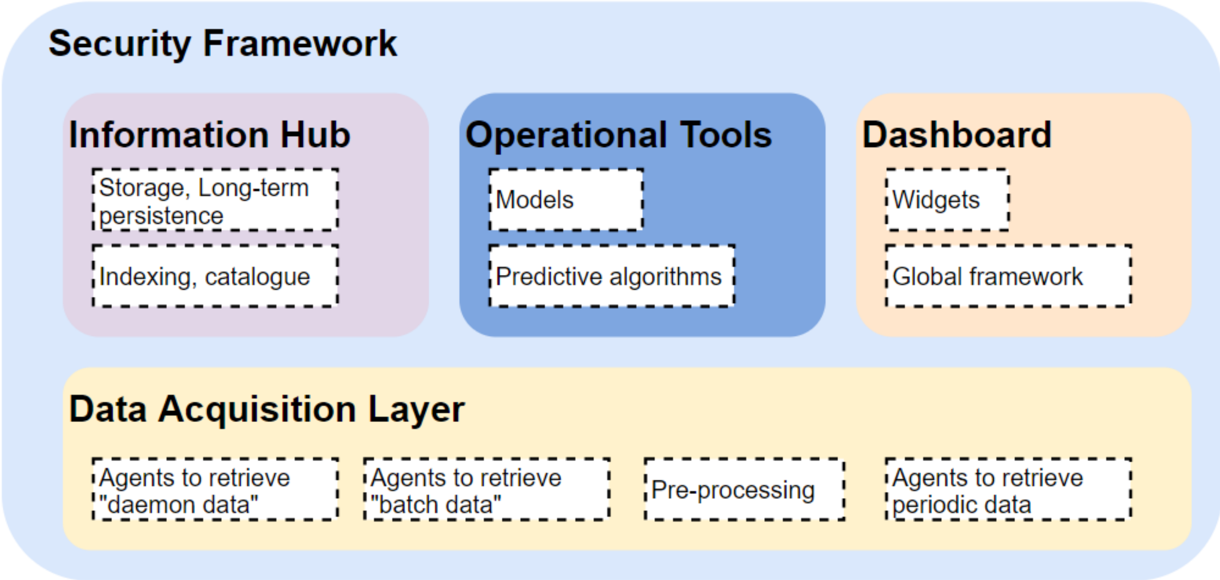


Figure 2: PIXEL modular architecture

3 The Port Activity Scenario model

3.1 Objectives of the model and context

Ports are complex ecosystems with a multitude of actors (port authority, terminal operators, carriers, citizens, legislators) producing and consuming each day huge quantities of information. These operational data are increasing, depend on different stakeholders and come from several sources: vessel operations, crane scheduling, resources tracking, container status, surface or berths available, air/water quality measurements, energy consumption and production. However, in many cases and especially in small and medium ports, the information is still exchange in a document-centric way. Today, European small and medium-sized ports are experiencing a lack of tools to calculate, estimate or predict impacts on energy consumption, transport networks and environmental pollution of port activities. To provide these tools we consider that the development and implementation of tools to model the impacts (in terms of energy, transport and pollution) of port activities is a needed step. Indeed, port activities undeniably have an impact on their environment, on the city and the citizens living nearby. To have a better understanding of these impacts, ports need tools allowing suitable modelling, simulation and data analysis. As the port activity is generated by incoming and outgoing cargoes into the port ecosystem, we have developed a model called the Port Activity Scenario, denoted as PAS, to consider cargoes' transitions and to be able to better understand the related energy consumption, pollutants emissions and organization of a port. PAS has been developed integrating constraints and need of the small and medium ports and with the objective to convert raw data into useful information about today and model port activities to predict future impact.

3.2 Definition of the model

The main purpose of the PIXEL modelling stack is to transform raw data (from sensors, various computing devices and fixed lists of data) into useful and actionable information. To do this, the approach developed makes a clear distinction between two steps, as illustrated in the figure below. The first step transforms the fed raw data into a Port's Activities Scenario (PAS) which is a list of all the upcoming port's operations with determined articulation across time. The second step calculates the outcome of the PAS on a certain period (as energies consumption, pollutants emission etc.).

Figure 3: From raw data to actionable information, the PIXEL modelling approach.

PAS focuses on cargo handling from one area to another through machines. For each cargo, there are several ways to arrange transition operations between areas. A hypothetical combination of those transition operations (for one or more cargo) is denoted thereafter as a scenario. Input data of the PAS model are mainly based on: i) vessel planning (arrival date, cargo type and tonnage, ii) activity data (details related to the transfer of cargo), iii) operational data (refer to the technical specifications of engines and equipment used) iv) emission source data (related to emissions factors of engine and equipment). The knowledge and modelling of the supply chain and port activities (machine type, duration of use, position in the port) enable the building of activity scenarios that are used to identify the energy sources and local emissions of pollutants but also to estimate the flow of cargo entering or leaving the port. Using this approach, the resulting modelling scenario might be used by the ports as a support for decision making.

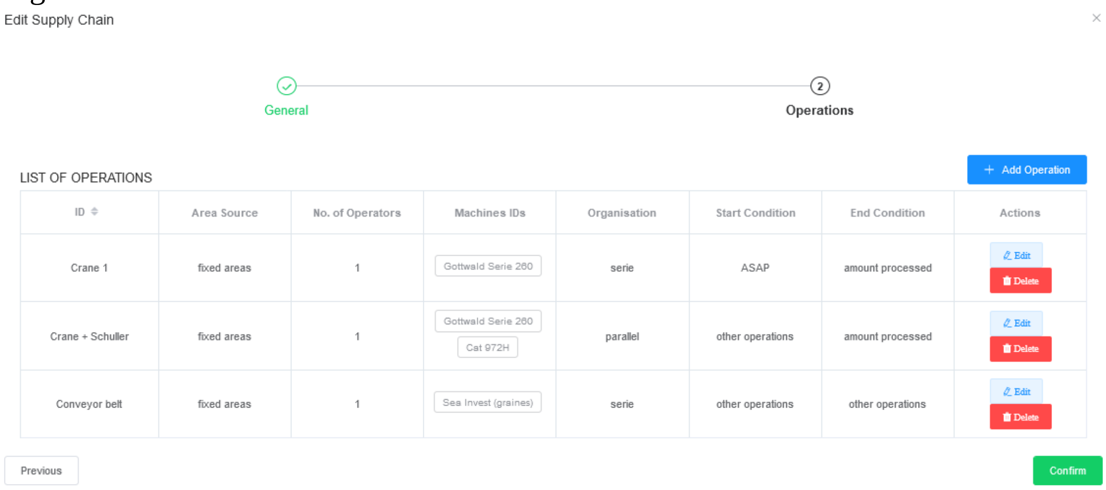


Figure 4: Supply chain definitions using the PIXEL GUI

A "PAS scenario" is a set of data describing all the activities and the equipment participating in those activities that are taking place within a timeframe for a specific amount of specified cargo. As such, it allows the calculation of energy consumption, emissions quantifications, and other externalities quantification. It consists of: a) a call for the handling of an amount of cargo starting at a specified time, b) a supply chain applied on the above-mentioned call. The supply chain contains machinery, areas, general port considerations or restrictions. PAS model can describe all operations related with a cargo transition that take place within a port. Its purpose is to establish an operational description of the port activities related to cargo handling. The produced PAS is a fulfilled data-model listing all the considered activities' time series. Because data is a key point to have useful models and because data collection is an intensive phase, PAS model can be used with different levels of details regarding input data. Depending on the purpose of modelling and the expected precision, inputs data could be less or more detailed. Thus, ports can use PIXEL models with comprehensive data (detailed model and low uncertainty), screening data (some local input and external data leading to a significant uncertainty) and scaled data (average and non-specific input data leading to great uncertainty and just giving an order of magnitude). If one port has installed enough infrastructure for collecting the data and has incorporated the needed specifications, a scheduled execution of the PAS can be set. A more detailed description of PAS model is available in [10].

Figure 5: PAS outputs and results visualize as a Gantt chart.

The sequence of actions from a vessel call to, as an example, a quantification of emissions is as follows:

1. The port agent, terminal operator or the port authority described the different supply chain with their associated machines and specifications. Figure 4 shows the GUI that have been developed to help ports fulfil these data. This work is just done once (and updated by the user as needed with the addition of new equipment procured or the deletion of obsolete machinery). Then the PAS model is configured to be run in an automatic way through the Operational Tools of the PIXEL platforms.
2. A ship approaching a port sends a vessel call and a FAL form (convention on Facilitation of International Maritime Traffic) to the port management information system. Thanks to the PIXEL platform this vessel call is automatically stored in the PIXEL Information Hub and made available as an input for the PAS model.
3. The PAS model considers the following item to create an ordered list of operations:
   a. The cargo.
   b. The estimated time of arrival.
   c. The port description.
   d. The rules and priorities set for the handling of similar cargo.

   This list of operation can then be visualize using for example a Gantt chart as shown in figure 5.
4. The energy consumption corresponding to the PAS predicted list of operations is calculated by the energy model.
5. With the use of emissions factors, the energy consumption calculated is translated to emissions quantifications by the PAS.
6. The emissions quantifications can be used by the Environmental Pollution Models and by the Port Environmental Index [11].

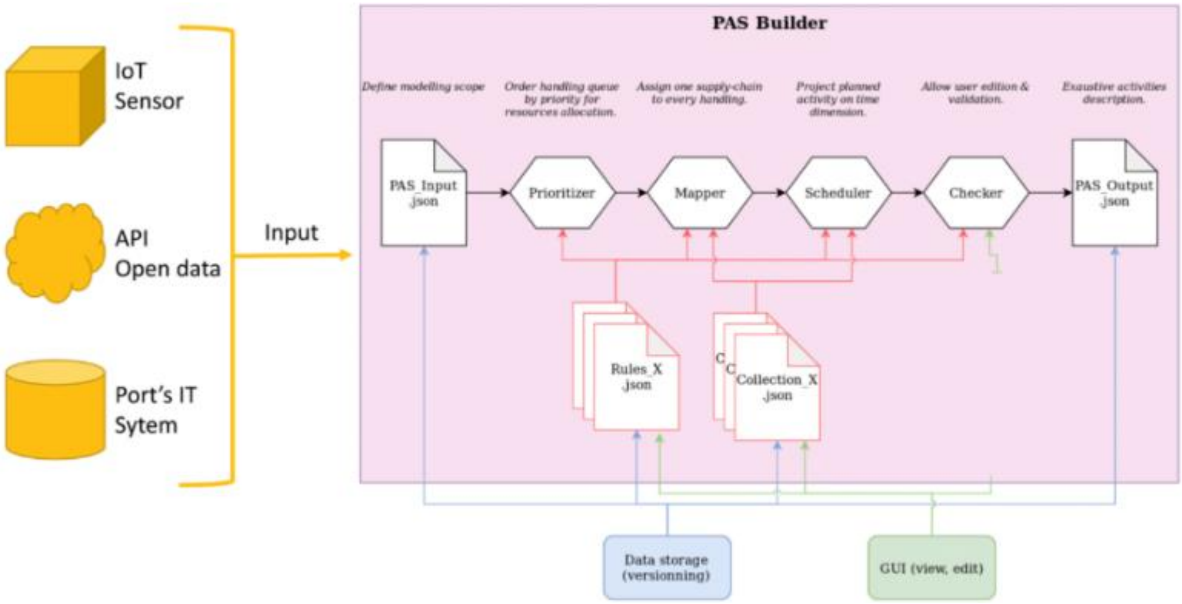


Figure 6: Schematic description of the PAS model.

4 Outcomes and results linked with the PAS model

4.1 PAS as an element for logistics innovation

The Port Activities Scenarios is used to link different models and guarantee interoperability as described in the figure below. Depending on the available data via the PIXEL Information Hub, a scenario can be completed with models related to energy, environmental pollution, area occupancy. In the following some examples are provided on how PAS and the modelling approach previously described can be used to improve logistics and the environmental port management.

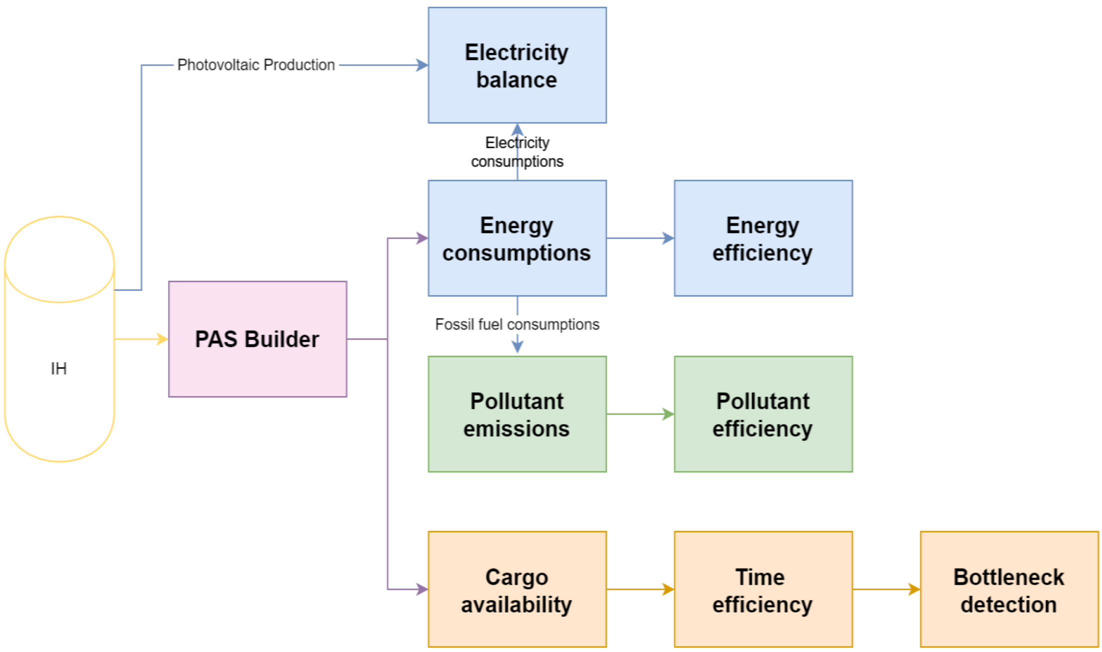


Figure 7: PAS and example of models' interoperability

PIXEL proposes a model that enables ports to calculate their energy consumption relating to their activities. For every considered cargo that is transiting inside the port, the corresponding sequence of operations across time is provided through the Port Activity Scenario described in the previous section. For each operation, the duration, the machine's energy type and its unit consumption values are available (from input and parameters combination). For each operation, the energy cost is calculated as the product of atomic operation's duration multiplied by its machine's unit consumption. Note that the energy consumption independent to the PAS (e.g., buildings with thermal regulation) could also be considered and add to the energy's consumption time-series. Then for every energy's type, all the consumptions occurring during the same timeframe can be summed to get the corresponding total consumption. The sequence of all those timeframes constitutes the energies consumption time-series. This functionality is directly included as a module of the PAS.

Figure 8: Example of results of the energy model (via the PIXEL platform GUI). Results can be filtered by energy type and be displayed as instant or cumulative consumption.

The Port Activities Scenario (PAS), combined with the energy model and emission factors, is a transferable and applicable tool for small and medium European ports that allows to model port supply chains. The output of this model can be used as an input for evaluating the energy consumption and with the use of emissions factors to provide an estimation of pollutants emissions. Following the previous, the usefulness of the PAS for the Port Environmental Index [11] can be understood two-fold: i) Providing information on the berthing time. Ideally, a port will provide the values of the time a vessel is being berthed, hoteling and manoeuvring within port maritime area. The PAS outputs can be used to obtain a (approximated) value of how much time is a vessel operated (loaded/unloaded), which can be considered a basis for the berthing/manoeuvring ratio. ii) Providing information of energy consumption of the terminal to calculate air emissions. Energy consumption is, usually, a complex metric to quantify, in contrast of what logical thinking would indicate. PAS output allows to know, with a simulation perspective, the total amount of energy consumed by the machinery needed to operate one vessel. Using that information, altogether with the type of fuel/power used, can be leveraged by the PEI to estimate the air emissions attributable to the terminal activity in a period of time. This functionality is directly included as a module of the PAS.

4.2 PAS as a node of the physical internet

The Physical Internet (PI, pi or π) is a concept that models logistics chains as an interconnected network of intermodal hubs that manage unified, standardised elements (pi-containers), encapsulating merchandise within them. The basic idea is to conceive the logistics network as a Digital Internet landscape, with a series of nodes (pi-nodes, representing logistic hubs) interacting among them to forward such elements behaving like a TCP/IP network. PI relies on collaborative protocols, and standardized, modular, and smart containers [12]. This change of paradigm relies on collaborative protocols, effective data sharing and the use of standardized, modular, traceable smart containers. PI, if successfully achieved, will improve overarching efficiency, reducing the impact of logistic chains on the environment and the time-to-distribution perspective. The PI requires the interconnection and intervention of a huge number of informational elements, which entails the co-operation of innovative technology applications for tracking pi-containers, managing the pi-protocols and sharing information among involved actors (from both horizontal and vertical sectors). In this sense, a major technology that comes to play here is Internet of Things. IoT can hugely benefit PI as it encompasses most requirements about containers tracking, end-to-end visibility of the objects, allowing operations over that digitalized data and facilitating ubiquitous information exchange [13].

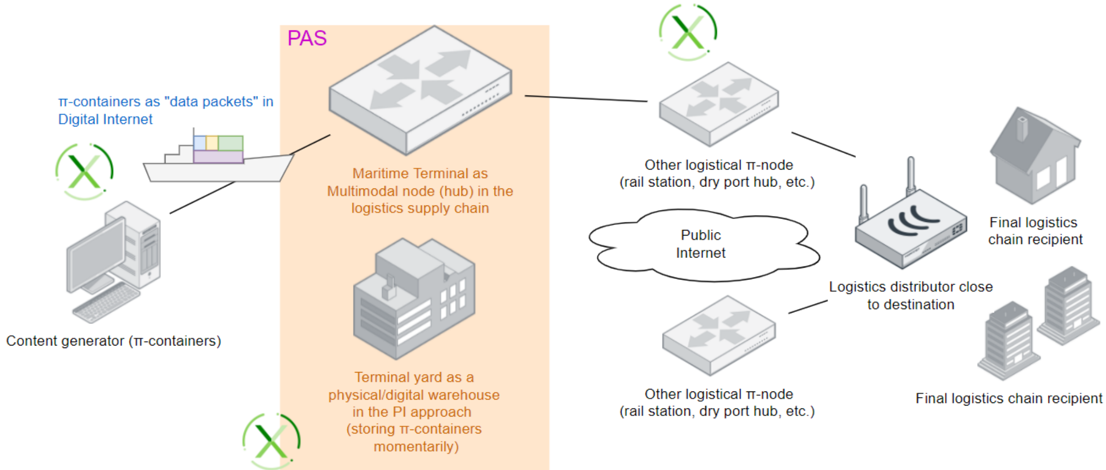


Figure 9: PIXEL and PAS as Digital Twin to represent a pi-node in the Physical Internet approach

Figure 9 aims at illustrating the fit of PIXEL (X logo) and PAS (orange rectangle) within the Physical Internet. On the one hand, PIXEL can be understood as the (aforementioned) IoT framework to achieve end-to-end communication and sharing of data/processes between diverse entities and actors of the logistics supply chain. PIXEL is prepared to bundle the components of the PI (pi-containers, pi-trucks, pi-trailers, pi-conveyors, etc.) as data pieces to be managed by the Context Brokers and Information Hubs (single-instance per each deployment in a node). In addition, PIXEL would allow their tracking (via GPS, RFID or others), allowing the logical operation of that data in each logistic node. Due to the multi-actor approach of the platform, they could share information (easy-to-connect APIs) in real time, letting various profiles of port managers to analyze and visualize the information.

5 Conclusion

In this whitepaper, we have described the Port Activity Scenario that has been developed during the PIXEL project. Based on an open-source IoT platform that allows to gather and store data coming from heterogeneous sources, the PAS model is able to simulate and predict the port activities and can be seen as a first step towards a Digital Twin of a port. The Port Activity Scenario (to be used as a model within PIXEL or outside the platform as a standalone tool) can be utilized to i) model and ii) simulate (two of the Digital Twin domains) a maritime port terminal as a pi-node in the logistic chain of a classic maritime transport picture. The fact of wrapping PAS as a pi-node provides various interesting traits towards optimization and efficiency: the PAS allows to i) “predict” the behavior of the pi-node (terminal) with regards to packet (pi-container) throughput, ii) forecasting how time it will take to operate each unit, and iii) how much energy will be used to do so and the internal operations required. The previous will allow port operators (and actually upcoming elements in the supply chain) to plan ahead and minimize round-trip-time of the pi-container (packet). Additionally, PAS is providing information that can be distributed in the internet regarding the expected timing of ports operations (such as loading and unloading of vessels) that can impact the logistics chain as well as their surrounding communities, through Port Community Systems. Thus, it can then be used to normalize a logistics chain, minimize storage times, warn about upcoming traffic, etc. The modularity of PIXEL allows PAS to connect and provide data to multiple models (calculation modules). Future research can focus on addition of models related to intramodality of transportations (selection of transport mode in relation with the expected cargo) and on integration of data available by pi containers in order to automate decision making regarding storage and shipment.

Acknowledgements

PIXEL project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement Nº769355.

References

[1] Elisa Negri (2017). "A review of the roles of Digital Twin in CPS-based production systems". Procedia Manufacturing. 11: 939–948.
[2] T. H. J. Uhlemann, C. Lehmann and R.Steinhilper (2017). The digital twin: Realizing the cyber-physical production system for industry 4.0. Procedia Cirp, 61, 335-340.
[3] C. Angulo Bahón, J.A. Ortega Ramírez, L. Gonzalez Abril (2019). Towards a healthcare digital twin. In Artificial Intelligence Research and Development vol. 319 (pp. 312-315). IOS Press
[4] W. Hofmann and F. Branding (2019). Implementation of an IoT-and Cloud-based Digital Twin for Real-Time Decision Support in Port Operations. IFAC-PapersOnLine, 52(13), 2104-2109.
[5] R. Boyles, "How the Port of Rotterdam is using IBM digital twin technology to transform itself from the biggest to the smartest," 29 8 2019. Online, https://www.ibm.com/blogs/internet-of-things/iot-digital-twin-rotterdam/. Accessed on: 9 Mar, 2021.
[6] E. Rademaker, "The Digital Twin of Port of Rotterdam," 17 10 2018. Online, https://youtu.be/nUbgHGJXj5c. Accessed on: 9 Mar, 2021.
[7] Maggie Mae Armstrong (2020) Cheat sheet: What is Digital Twin? https://www.ibm.com/blogs/internet-of-things/iot-cheat-sheet-digital-twin/
[8] Smart Data Models: Key Performance Indicator." https://github.com/smart-data-models/dataModel.KeyPerformanceIndicator (accessed Apr. 15, 2021)
[9] "FIWARE Orion Context Broker." https://fiware-orion.readthedocs.io/en/master/ (accessed Apr. 15, 2021), "pyngsi framework." https://pypi.org/project/pyngsi/ (accessed Apr. 15, 2021).
[10] E. Simon, C. Garnier, I. Lacalle, J. P. Costa and C. E. P. Salvador, "Small and medium ports' activities: introduction to the PIXEL approach.," WIT Transactions on the Built Environment , vol. 187, pp. 149-163, 2019.
[11] A novel approach for assessing the ports' environmental impacts in real time – The IoT based port environmental index, Matija Široka, StjepanPiličić, Teodora Milošević, Ignacio Lacalle, Luka Traven, in Ecological Indicators, Volume 120, January 2021, https://doi.org/10.1016/j.ecolind.2020.106949, (Available online 19 September 2020)
[12] E. Ballot, B. Montreuil, and R. Meller, The Physical Internet: The Network of Logistics Networks, L. D. Française, Ed. Paris, France: La Documentation Française, 2014
[13] H. Tran-Dang, N. Krommenacker, P. Charpentier and D. Kim, "Toward the Internet of Things for Physical Internet: Perspectives and Challenges," in *IEEE Internet of Things Journal*, vol. 7, no. 6, pp. 4711-4736, June 2020, doi: 10.1109/JIOT.2020.2971736

Towards PI Implementation: Interoperability and robustness of PI layers functionality illustrated through multi-context implementation

Konstantinos Zavitsas[1] and Makis Kouloumbis[2]

1. VLTN, Antwerp, Belgium
2. Inlecom, Athens, Greece

Corresponding author: k.zavitsas@vltn.be

Abstract: *This paper analyses the robust interconnectivity and functionality of the four core PI layers, as established through work undertaken for the recently completed EU funded project ICONET: 'New ICT Infrastructure & Reference Architecture to Support Operations in Future PI Logistics Networks' focusing on intra-European transport (TEN-T)[1]. The ICONET PI layers functionality and interoperability has been developed to fit multiple unique contexts, that represent significant components and functionalities of the PI. Utilising the Living Labs of the ICONET project, they account for:*

- *micro-level operations within PI hubs such as the Port of Antwerp,*
- *single-entity long-haul logistics PI functionality[5]*
- *multiple-entity e-commerce operations in urban areas,*
- *and Warehouse-as-a-Service offerings.*

This paper presents the functionality of each of the core PI layers that utilise the fundamental PI principles and illustrates their custom implementation into various T&L contexts. The performance of the layers' functionality is quantified using context specific KPIs developed in collaboration with Living Lab partners and tested in through a PI simulation environment. The findings illustrate benefits both in terms of utilisation of new PI capabilities, as well as towards improved fill rates and modal shift having an impact on transport cost, reduction of emissions, and operational efficiency.

Conference Topic(s): *From Logistics Networks to Physical Internet Network*

Keywords: *Workflow Algorithms, Decision Support Systems, Optimization, Physical Internet*

1 Introduction

Performance of freight transportation is one of the crucial elements for the sustainability of supply chains and logistics. Despite progress achieved, inefficiencies are evident by the high frequency of empty truck trips and relatively low utilisation of multimodal resources. According to Eurostat (2019), most trucks in Europe fell in the range between 15% and 30% of empty journeys. Moreover, freight transportation (in developed countries) is responsible for nearly 15% of greenhouse gas emissions. This ratio has been increasing despite ambitious

[1] https://www.iconetproject.eu/

reduction targets. Improved transportation efficiency is therefore an important objective for environmental and financial purposes.

The Physical Internet (PI) envisages to become a global system of freight transport across heterogenous networks and supply chains exploiting standard methods and protocols. It promises to revolutionise how transport and logistics is practiced, and to improve on critical variables such as cost, utilisation rates, and emissions through improved multi-modal integration and open accessibility to static and mobile infrastructure. Critically important for the implementation of the PI vision is the standardisation and interoperability of transport, logistics systems and processes (Montreuil, 2011). To integrate current Transport & Logistics practices into the PI concept, a classification based on the Open Logistics Interconnection (OLI) layers has been proposed (Montreuil et al., 2012; Colin et al.2016), with the four core layers being:

- Encapsulation: Standardises the packaging process of cargo and goods that are consolidated/deconsolidated into π-containers for transportation via the PI. It is also responsible for the consolidation/ deconsolidation of π-containers into π-movers.
- Shipping: Specifies what has to be transported as well as the transportation process conditions and constraints. It is responsible to make appropriate adjustments to the shipping instructions to ensure compliance.
- Networking: Networking defines the interconnected infrastructure of available processing, storage and transporting facilities (transport services, terminals, distribution centres, warehouses) through which the goods will be transported from their origins (manufacturing, distribution and other locations) towards their customer(s) locations.
- Routing: Routing is a process that creates a plan that describes the stage by stage detailed visiting and usage of networking nodes and links from origin to destination.

The above PI layers attempt to integrate the following four fundamental principles in the Transport and Logistics (T&L) infrastructure operation, in deviation from current practices:

- The ability to effectively handle **modular packaging** such as π-containers. Standardised containers have revolutionised the logistics industry through standardisation, however their size range offering is limited yielding operational inefficiencies. Increasing the modularity of containers, as well as well as the number of transhipments due to the compartmentalised routing, are addressed through the development of Operations Research (OR) based decision support tools for packaging, grouping and loading.
- The increasing **digitisation** of T&L infrastructure creates an information offering that remains largely unutilised. IoT sensors and Track and Trace capabilities are integrated into the PI workflow, enabling new functionality for cargo prioritisation, re-routing, identification and handling of damaged goods.
- The integration of robust **decision support tools** (DSS) that enable efficient cargo and fleet routing and distribution, under uncertain and adjustable terminal, warehouse and network conditions. Building on existing protocols (Sarraj et al, 2014), a toolbox comprising of established and novel OR and Machine Learning (ML) models is developed, adapted for the PI context and applied in a modular and robust way.
- The **open accessibility** to robust transport, terminal and warehousing services through the “as-a-Service” paradigm improves operational efficiency. The incorporate it in the

PI layers standardised collaboration protocols for the identification and formation of mutually beneficial agreements have been considered.

To integrate the aforementioned functionality and ripe the benefits arising from the core PI principles (REF), a substantial change to the resources and decision making process is required. This paper presents the outputs generated from the recently completed EU funded project ICONET: 'New ICT Infrastructure & Reference Architecture to Support Operations in Future PI Logistics Networks' focusing on intra-European transport (TEN-T). After an in-depth assessment of current T&L practices and their shortcomings, the project focused on development of an integrated workflow between the PI components, and its application into various contexts. This work is presented in the following three sections of this paper.

2 Redesigning the transport process

In principle, goods and products are transported from one location to another where they become more useful and valuable. There are several stages in the development of a product in a typical supply chain, starting from one or more raw materials, going through several processing and assembly stages, before it finally reaches a retail store, and is purchased by the final customer. Performance of freight transportation is one of the crucial elements for the sustainability of logistics and supply chain.

T&L involves the coordinating effort of several organisations, each of them focusing on a different part of the supply chain process. A supply chain includes not only the manufacturer and the suppliers, but also transporters, warehouses, retailers, and even customers themselves. Although this may include organisations that have only an indirect role such as for example banks and insurance companies, such organisations do not directly influence operational efficiency in the transport and logistics process and are therefore not considered further. Direct stakeholders in the transport and logistics processes can be due to them owning (initially or ultimately- i.e. as sellers and buyers), the goods that are transported, the equipment and other resources by which the goods will be processed and transported, or because they are coordinators of the different processes and activities involved.

The functioning of a supply chain involves three key flows between the stakeholders – information, products and funds. The goal when designing a supply chain is to structure the three flows in a way that meets customer needs in a cost-effective manner. Information flow is a crucial element for achieving such efficiency, as it allows for the better planning of resources in earlier stages of the supply chain. Currently, limitations to supply chain visibility (SCV) due to the lack of information sharing, are shown to be a cause for operational inefficiencies (Swift et al., 2018).

Operational information is typically located in organizational silos, and is analysed in a post-operational setting, to inform performance metrics and educate future collaboration and contractual agreements. Although this is in contrast to the PI open accessibility paradigm, multiple transitional stages are anticipated for bridging that gap[2].

Furthermore, supply chain stakeholders' perception of performance varies with the stakeholder role, operational context (e.g. urban or long-haul), and function in the supply chain (e.g. warehouse or transport). Therefore, the performance metrics each stakeholder utilises to measure operational efficiency do not always match, and in cases are contradicting. Through interactive discussions with stakeholders, several studies (Macharis et al, 2012) establish

[2] https://www.etp-logistics.eu

criteria and their associated weights per stakeholder. Due to this variability, collected information and decision processes vary greatly in each T&L stakeholder setting, hindering the motivation for standardization and integration of processes that PI promotes.

The definition of the OLI layers and their subsequent roles, based on Internet protocols, has motivated a top-down approach in the development of the PI, that does not yet clearly connect to ground operations for T&L stakeholders. The ICONET project has contributed in bridging this gap, through establishing an implementable data structure, built bottom up, from the projects Living Labs, that cover:

- micro-level operations within PI hubs such as the Port of Antwerp,
- single-entity long-haul logistics PI functionality
- multiple-entity e-commerce operations in urban areas,
- and Warehouse-as-a-Service offerings.

Based on the knowledge gain from the implementation of integration activities in the context of the four ICONET Living Labs, it has been possible to establish operational features for integrated T&L activities, based on the OLI functionality. The following sections cover those features in terms of integrated data structure, core operational flow of the transport process and its functionality, as well as functionality extensions to address specific contextual needs.

3 PI workflow and integrated functionality

3.1 Integrated data structure

The OLI Networking Layer is responsible for collecting and maintaining all crucial information of the operation of the PI network. Starting from identifying the relevant nodes and links of the network and representing them in terms of scale and aggregation level, as well as associating each physical infrastructure component with properties to be used for operational decisions.

3.1.1 Static and dynamic properties

Network information describe static infrastructure characteristics such as the length and bandwidth of a link, the modes that can accommodate the function of carrying cargo (e.g. truck, rail), or even more detailed information such as classification into motorway, or number of lanes. A similar concept is applied to the description of nodes, with capturing intermodal and processing capabilities being captured. A node may also represent a warehouse that has specific capacity for storage and docking capability. Static information enables the optimisation of routing decisions; however, such decision process is prone to bias due to dynamic parameters influencing the network, such as road traffic and conditions and processing/handling times and costs at nodes.

A networking service that is based solely on static information, will inevitably yield significant inefficiencies for the operation of the PI. For example, an incapacitated link due to roadwork or traffic, if not dynamically updated in the networking service, will be considered as a feasible route for cargo routing, and eventually cause delays. Similarly, for nodes, a broken-down refrigeration unit in a warehouse, if not considered by the networking service of the PI, will direct cargo to a location that cannot be serviced and eventually cause queues and operational inefficiency. It is therefore critical to accommodate network node and link operational status information as illustrated in Figure 1.

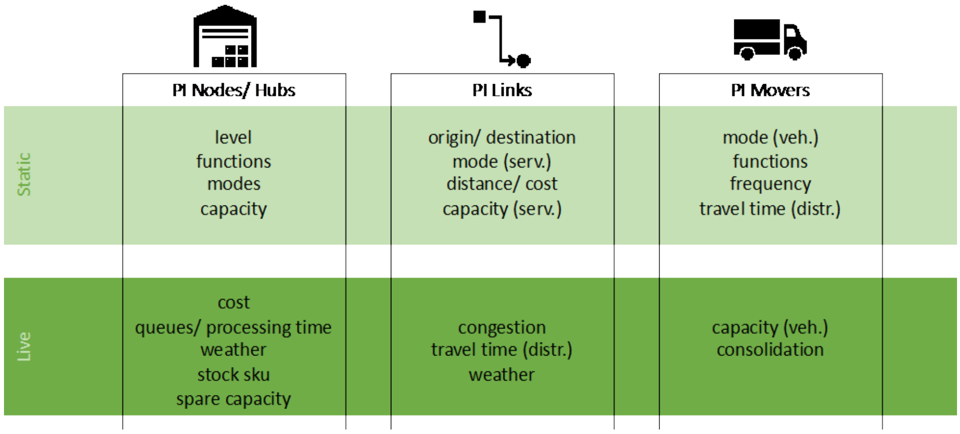


Figure 1: Integrated PI data structure

In the context of transport links and services, an additional data layer of the networking service is that of operational services and their availability. The aim of this layer is to account for the fact that roads and warehouses do not handle directly cargo transport, but rather indirectly. Truck, air, river/sea or rail services that operate on the infrastructure are responsible for undertaking the task of physically moving cargo from one location to another. With that in mind, it is essential that the networking service collects service schedules or track vehicles, that operate between specific locations. Such information is useful, as more enhanced freight routing and scheduling algorithms become available. A final layer of complexity is anticipated, that accounts for live information on the capacity of en-route service and warehouses. Only at that level of detail, an optimal with limited uncertainty allocation of freight to services to routes is achieved. The first two columns of Figure 1 capture the standardised static and live information required to inform decision making, while the third one looks into the data collected on the services of different modes that operate in the network.

3.1.2 Network representation

The integrated data structure presented in the previous section, allows for a detailed network representation that captures transhipment capabilities at nodes and complex PI Hubs such as the Port of Antwerp.

Transhipment shipment operations are captured through dedicated links that connect every arrival mode (source nodes), to every departure mode (sink nodes) within the PI Hubs, and each individual terminal, enabling the assignment of weights to PI Hub operations. This representation allows for transhipments to be associated to:

- Distances that require to be covered. This weight is more relevant to large PI hubs such as the PoA where transhipment legs are of considerable distance.
- Cost that can incorporate average handling cost
- Travel time that can incorporate average handling time and queues
- Capacity that represents the number of such transhipment PI hub infrastructure can undertake.

A similar network representation can be adopted for capturing Additional PI hub functions that extend beyond the direct T&L functionality, such as customs and infrastructure bottlenecks. A detailed representation of the PI Hub services contributes to accurately capturing the operational limitations that often drive stakeholder decisions.

3.2 PI transport process workflow

Due to the increased transshipment needs in the PI context, enhanced emphasis should be paid both for process coordination between unknown to each other parties, as well as on standardisation of the exchanged information. With a standardised data structure for representing the network and contractual operations, it is possible for the PI to make informed decisions on the utilisation of infrastructure, assets and decentralised capability. This also enabled the development of an integrated workflow of the transport process as illustrated in Figure 2. The proposed PI transport workflow aligns and builds on the OLI layers functionality and the process proposed by earlier PI studies such as Sarraj et al. (2013). The four core PI services are considered to have the following role and functionality:

- Shipping is responsible for initiating and high-level managing process where the transport requirements must be conveyed to the transport chain coordinator(s) to ensure feasible and efficient transport.
- Encapsulation is responsible for fitting the products to be shipped in modular units (π-boxes) that form π-containers by setting up and solving a bin-packing problem. These are consolidated and deconsolidated as appropriate during the transportation through PI to make efficient utilisation of the transport resources.
- Networking collects and standardizes information on the network conditions and the physical movement of goods through various sources (including IoT devices). These may include multiple transport modes and multiple 'changeover' points where cargo is bundled, unbundled, trans-shipped etc. It is also responsible to package relevant information to be consumed by the routing service or for other network exploration functions, in an efficient manner.
- Routing is an end-to-end framework for solving large scale routing problems. The input to the framework is PI network as property graph and the problem instance. The routing solution achieved is based on establish heuristics for given Travelling Salesman or Vehicle Routing Problem (VRP) instances and variants such as capacitated VRP and VRP with time-windows.

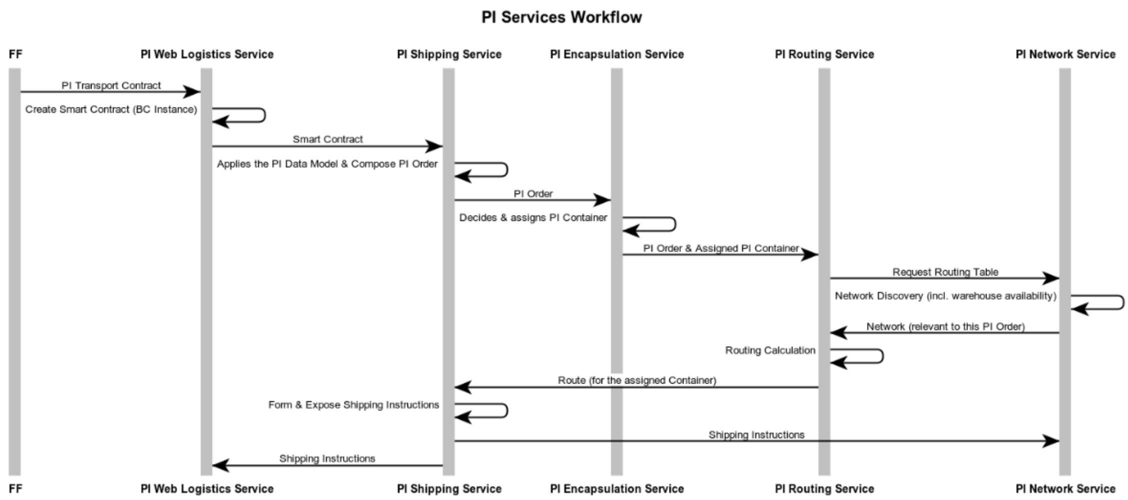


Figure 2: PI Services Transport Process Workflow

The Shipping Service initiates an order capturing it core contractual characteristics. At origin, the Encapsulation Service is responsible for fitting the cargo into optimal π-boxes that start to be tracked by the PI. The transport process is then initiated by the Shipping Service, that communicates the shipment info to the Routing Service. The Routing Service collects the relevant network status from the Networking Service, and identifies an optimal shipment route considering, travel time, consolidation, emissions and cost as per the shipment instructions. It returns the Shipping Instructions to the Shipping Service, requests from the Encapsulation Service to fit the π-containers to the allocated π-mover, and initiates the transport process.

The implementation and progress of each shipment is tracked through frequent requests initiated by the Shipping Service. The triggers for such requests can originate from geofencing that indicates arrival at the next node, or the expiry of an ETA. In the first case the Shipping Service requests updated Shipping Instructions from the Routing and Networking Services in cases of changes in the network status, and then calls the Encapsulation Service to fit the π-containers to the allocated π-mover, and initiates the transport process of the next leg. The process is repeated until the destination is reached. In the case of the expiry of an ETA, the Shipping Service, attempts to locate the π-container, and if needed propose better Shipping Instructions through the Routing and Networking Services.

This workflow allows for the integration of multiple transport process features, such as:

- The expedition of a π-container if it is running behind schedule. This is achieved, by increasing the weight associated to travel time when passing the instructions to the Routing Service.
- The prioritization of π-containers with higher travel time weight (express) at PI hubs. This is achieved by considering FIFO for regular shipments, but also considering a priority queue for express ones.
- The re-routing due to updated network conditions, when traffic or delays are anticipated in the Shipping Instructions route.
- The re-routing of damaged goods. The IoT connectivity of π-containers enables the identification of cargo that have been exposed to extreme conditions and require disposal. This can be handled on the go, rather than at the destination.

3.2.1 Handling Priority

The prioritization and expedition features of the PI workflow are accommodated through a variety of information maintained by the Networking Service on both infrastructure and services as discussed in Section 3.1.1. For example, the Networking Service provides information on distance, travel time, emissions estimates and handling/ transport cost estimates for every link and node (such as transshipment services) in the anticipated network. The Shipping Service maintains weights for each of those parameters depending on the progress status of the shipment. An on time regular shipment might has a higher weight for cost, while an express or a delayed regular shipment has a higher weight for time. This allows the Routing Layer to produce Shipping Instructions that better reflect the needs of each specific shipment increasing the changes of on-time arrival.

3.2.2 Handling Consolidation

The proposed workflow considers the challenge of increasing π-mover fill rates and achieving cargo consolidation. An additional metric is maintained by the Networking Service for each scheduled π-mover service, that captures the cost in consideration of how loaded the specific π-mover is. This metric attempts to capture that a half loaded π-mover, occurs 70% of the fully loaded transport costs and emissions. Therefore, the *consolidation price tag* for the first π-container that will be loaded to a π-mover that fits two π-containers, is 70% of its total running costs, while the *consolidation price tag* for the second is the remaining 30%. This approach encourages the further loading of already running (half-loaded) vehicles.

3.3 Extended functionality

Despite the versatile core PI workflow functionality, it is observed that without extended functionality, its applicability to ICONET's Living Labs is limited. Multiple special cases are found to arise for different T&L contexts, with the most predominant being urban logistics and the handling of eCommerce orders. The subcases arise from the unique nature of the transport process in the two contexts, that vary significantly from handling long-haul shipments. In the urban context, delivery is structured around consolidations centers, with dedicated van fleets responsible for undertaking delivery rounds to drop off locations. Delivery rounds cannot be scheduled or run on fixed routes, as the mix of deliveries varies highly from round to round. To address the challenge, a fleet routing problem was integrated in the PI workflow, utilizing information on van locations and loading, to design optimal delivery rounds. The process is handled by the Routing Service that solves variations of the VRP.

The PI workflow is also expanded to handle of eCommerce orders. The unique nature of this subcase arises, from the multiple stores each retailer has, that frequently hold similar stock-keeping-units (sku). Therefore, there are multiple candidate locations for the shipment origin. Furthermore, there are cases when an order requires to be fulfilled at multiple locations as there is no single store to hold all sku, and therefore order consolidation is required. To address the challenge, the Networking Service maintains stock level data as illustrated in the integrated data structure in Figure 1. The Shipping Service retrieves the information as well as a distance matrix. A assignment algorithm is implemented, to optimally allocate origins to all eCommerce orders prior to initiating the van fleet rounds design and the order preparation and fulfillment.

Figure 3 captures the core functionality of each PI Service as well as its extended functionality for addressing subcases that arose from ICONET's Living Labs.

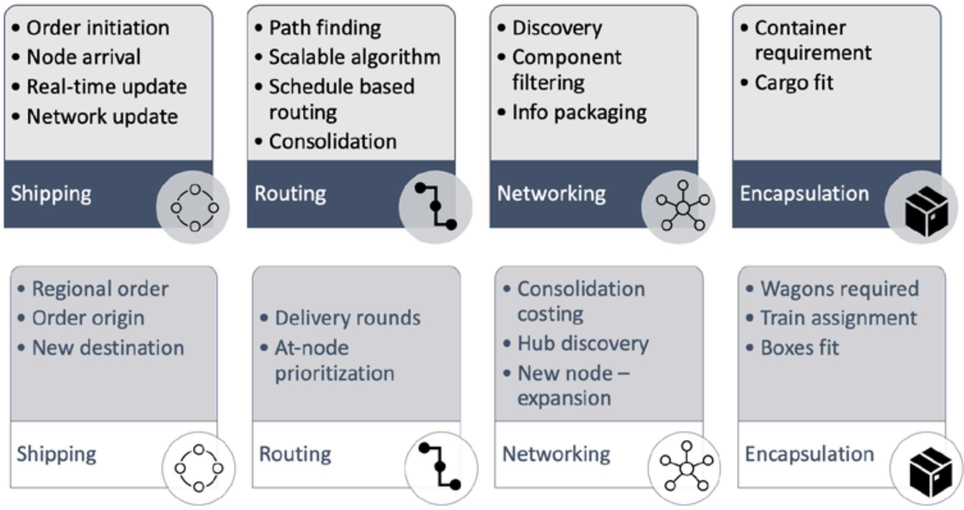


Figure 3: Core and extended PI Workflow functionality

4 Living lab implementation

To incorporate urban eCommerce delivery into the concept of the Physical Internet, a fulfilment store identification tool has been developed and incorporated in the Shipping Service. The tool simultaneously addresses the stockout challenge, by introducing a dynamic process for identifying which store will fulfil each order. The process attempts to improve on the current local (nearest store) fulfilment that causes stockouts. The aim of the tool is to associate each orders (specific delivery locations) with one of the candidate preparation stores optimally. This is handled through a mixed integer linear optimisation model that seeks to minimise the total distance for satisfying all orders. Assuming that the distance d_{ij} between every customer location j and every fulfilment store i is known, and that an additional integer variable s_{ik} captures the stock of products available at each fulfillment store i per sku k. And an additional integer variable o_{jk} captures the number of products k ordered in order j. A binary decision variable x_{ij} is equal to 1 if fulfillment store i is chosen to satisfy customer order j, and is 0 otherwise. Then, a cost minimization problem can be formulated as follows:

$$\min_{x_{ij}} \sum x_{ij} d_{ij} \tag{1}$$

$$\sum_i x_{ij} \geq 1 \quad \forall j, k \tag{2}$$

$$\sum_j x_{ij} * o_{jk} \leq s_{ik} \quad \forall i, k \tag{3}$$

The objective function is capture by equation (1). The first constraint (2) ensures that all orders are satisfied, while the second constraint (3) ensures that the store capacity for each SKU is not exceeded. The decision variables are binary.

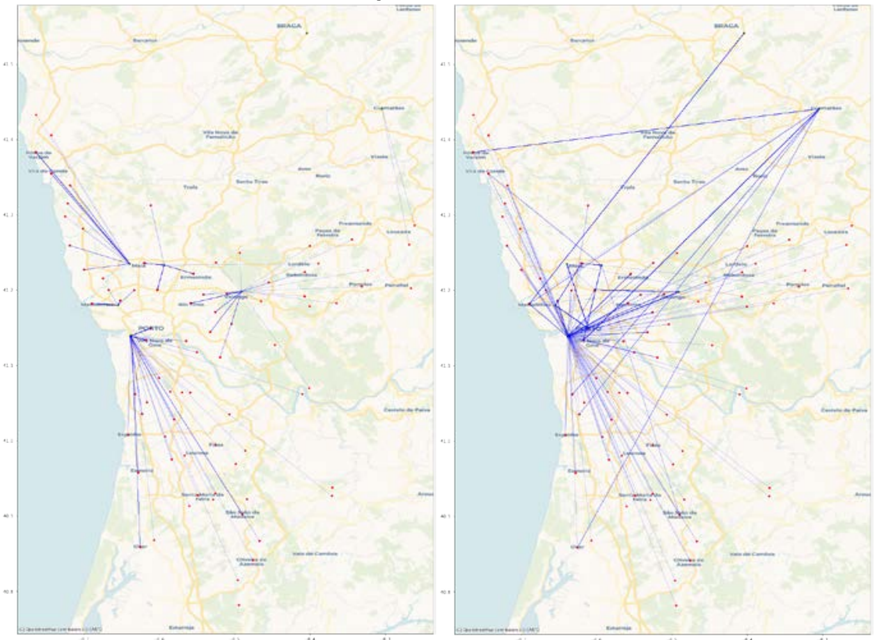


Figure 4: Store-order pairs with local fulfilment (left) and dynamic order fulfilment (right)

Table 1 illustrates the store-order pairs for 500 supermarket orders under BAU local fulfilment (left) and the PI enabled dynamic order fulfilment (right) in the region of Porto. In the business-as-usual case central Porto eCommerce orders are served predominantly by one store located in the west side of central Porto. The store eventually runs out of stock causing a significant amount of stockouts. In the PI enabled scenario, central Porto eCommerce orders are all served as the solution utilizes three additional nearby stores located north of the city center.

Table 1: Local and dynamic fulfilment performance metrics

Main KPIs	Scenario 1. Local fulfilment	Scenario 2. PI Network fulfilment	% Change
% Vans Fill Rate	63%	92%	+29%
Total Transport Cost (Eur)	2933	2961	+1%
Total Distance (km)	1047	1273	+21%
Distance per order (km)	2.52	2.32	-8%
Average Transport Cost per Order (Eur)	7.10	5.4	-24%
Average Delivery Lead Time per Order (min.)	59	77	+30%
Nº Orders Delivered	415	549	+25%
% Stockout	25%	0%	-25%

The order and stock data used for the analysis are based on real information provided by SONAE in early March 2020. Due to the COVID pandemic there was a significant increase in the number of eCommerce orders, compared to pre-pandemic levels, reaching approximately 500 per day. After pre-processing the dataset the analysis was initially undertaken with 549 orders in total. With the COVID pandemic putting a strain on logistics, the daily e-commerce orders in December 2020 in the Porto Region had more than doubled. To accommodate this significant increase the dynamic fulfilment model was revisited with a larger dataset of 1698 orders. The results were found to be consistent, even for a significantly increased order number. Under the local fulfillment scenario, it was found that 1148 out of the 1698 orders are delivered in full (no stockouts), with 32,4% of the orders having stockouts. Under the dynamic fulfilment model 1614 orders were fulfilled in full, with only 5% recording stockouts.

5 Conclusions

The paper presents an integrated data structure, and a robust PI workflow as well as extensions that cover ICONET project's Living Lab contexts. The proposed workflow and methods build on the OLI layers definitions and align with proposed PI transport process handling algorithms, aiming to expand the PI reach and functionality. The consideration of prioritization and consolidation features is illustrated, while extended functionality is shown to address urban logistics, and eCommerce challenges by incorporating delivery round design and eCommerce order handling.

The findings from the Case Study presented in Section 4, as well as results from other ICONET Living Labs (ref deliverables), indicate significant benefits for early adopters, if they can meaningfully integrate the limited PI functionality into their decision processes. At a PI level, it is illustrated that considerable functionality is currently implementable and can enhance individual and consolidated business operations, despite information not being openly shared. This is a significant finding, as several stages are still anticipated for comprehensive PI uptake of T&L processes.

The work is currently extended in PLANET EU funded project, that is focusing into international trade routes. Future work is anticipated to further bridge the gap between the conceptual PI and implementable solution, while also incorporating additional aspects in the PI functionality, such as social that are currently not considered.

References

- Montreuil B. (2011): Towards a Physical Internet: Meeting the Global Logistics Sustainability Grand Challenge, Logistics Research, v3, no2-3, 71-87.
- Montreuil B. (2013): Physical Internet Manifesto, v1.10, (Original v1.0, 2009), www.physicalinternetinitiative.org, 2014/02/20.
- Eurostat, 2019. Energy, transport and environment statistics. 2019 edition. Publications Office of the European Union, Luxembourg.
- Montreuil, B., Ballot, E., Fontane. F,. 2012. An Open Logistics Interconnection model for the Physical Internet. Proceedings of the 14th IFAC Symposium on Information Control Problems in Manufacturing, Bucharest, Romania.
- Colin, J-Y., Mathiew, H., Nackechbandi, M., 2016. A proposal fora n open logistics interconnection reference model for a Physical Internet. 3rd International Conference on Logistics Operations Management (GOL), Fez, Morocco.
- Sarraj, R., Ballot, E., Pan, S., Hakimi, D., Montreuil, B., 2014. Interconnected logistics networks and protocols: simulation-based efficiency assessment. International Journal of Production Research, 52:11:3185-3208.
- Swift, C., Guide Jr, V.D.R., Muthulingam, S., 2018. Does supply chain visibility affect operating performance? Evidence from conflict minerals disclosures. Journal of Operations Management 65 (5): 1–24.
- Macharis, C., Turcksin, L., Lebeau, K., 2012. Multi actor multi criteria analysis (MAMCA) as a tool to support sustainable decisions: State of use. Decision Support Systems, 54: 610–620.

Optimising Container Transports in Collaborative Roundtrip Scenarios

Matthias Prandtstetter, Georg Brandstätter and Lukas F. Krasel
AIT Austrian Institute of Technology, Center for Energy, Vienna, Austria
Corresponding author: matthias.prandtstetter@ait.ac.at

Abstract: *Within this paper, we discuss the importance of combinatorial optimisation problems arising in the context of the Physical Internet. We focus on one specific problem originating in one of the four demonstration pilots of the Austrian PhysICAL project which develops and demonstrates four best practice examples to path the way to the Physical Internet. It will be shown that the Collaborative Roundtrip Problem is similar to the well-known machine scheduling problem. Conclusions are drawn with respect to the computation solution approach as well as the importance with respect to the implementation of the Physical Internet.*

Conference Topic(s): *From Logistics Networks to Physical Internet Network; Interconnected freight transport, logistics and supply networks.*

Keywords: *Freight Transportation Planning; Horizontal Collaboration; Optimisation*

1 Introduction

In 2020, the Austrian model project PhysICAL (Physical Internet through Cooperative Austrian Logistics) has been started (PhysICAL, 2021). The main goal of the project is to showcase Physical Internet (PI) applications such that best practices examples are generated, and other companies and practitioners start to actively follow the PI idea. Hurdles and reservations shall be reduced by implementing four PI demonstration pilots. Ecological and economical assessment shall show the benefits on an individual (company) and general (societal) level.

The four PI demonstration pilots pursued within the PhysICAL project are:

- **Wood Logistics**: In Austria, the harvesting of wood is one significant economic branch as up to 40% of Austria's land is covered by woods (Statistik Austria, 2010). However, the transport of the harvested wood from forest to consumers is mostly done via trucks. While for the first-mile this is obvious (in most cases no other mode of transport is available in forests), the long-distance transports could be shifted towards rail (or even inland waterways). However, transhipment of wood from trucks to trains is rather cost-extensive as no intermodal transport unit is (currently) available. The goal is therefore to develop an appropriate hardware container which is capable of intermodal transports. In addition, digital service shall be provided to ease the ordering and implementation of these intermodal wood transports.
- **Intermodal Transport Platform**: The main goal of the intermodal transport platform is to provide a digital solution for ordering, processing and following intermodal transports. The idea is that the platform operates like a one-stop-shop where the shipper can order the whole supply chain like searching Google and clicking on the best result. That is, after specifying when to transport from where to where a list of feasible transport chains is presented. The shipper then only needs to click on the chosen one (e.g., the cheapest) which is then booked and (later) processed. All necessary intermediate steps (e.g., transhipments, coordination among the individual legs of the transport, etc.) are organised by the platform via direct (digital) links to the operators. In addition, the shipper can follow (in real-time) the processing of the transports. This platform is developed to be open to all participants following the given rules.

- **Supply Chain 3.0**: The main idea is to install a wholesale logistics supply chain for e-commerce. Currently, producers are either directly selling to end customers via their own online shops or via e-commerce platforms like Amazon or Alibaba. They have, however, to handle all logistics on their own. Even when selling on (large) online platforms, at least logistics up to the warehouses of the online platforms have to be organized by the producers themselves. In most cases, this is not the core competence of the producers and unnecessarily binds resources. Therefore, the idea is to install a wholesale logistics supply chain where the wholesaler takes over all (digital and physical) processes necessary starting at the loading ramp of the producer. This involves all physical transports but also digital registration of the products at platforms (including all eventually necessary accreditation processes). Obviously, this allows for horizontal collaboration as logistics processes (and therefore also transportation) is organized by one company. Furthermore, intermediate storages are shared among each other as well.
- **The New CEP Last-Mile**: In Courier, Express and Parcel (CEP) logistics the last-mile is known to be the most cost-extensive part of the transport chain with up to 53% share of the total costs (Honeywell, 2016). In addition, it is discussed that negative traffic-oriented impacts are existent (Accenture, 2021). Therefore, several projects exist focusing on collaborative white-label logistics for the CEP last-mile. However, we experienced in many discussions with CEP service providers that the white-label delivery is not desired as some questions are still open. Among others, transition of liability among CEP service providers is not yet fully regulated. Therefore, the main idea of this demonstration pilot is to come up with hardware and logistics process innovations enabling collaborative CEP last-mile services.

The whole concept is rounded up with the introduction of a Digital Twin (DT) of the (Austrian) transport system. Although the main focus is laid on a replication of those parts that are relevant for the four PI demonstration pilots. The main intention of the PhysICAL DT is on the one hand to objectively show the impacts of the collaborative logistics. On the other hand, the DT provides a decision support tool for decision makers (e.g., shippers) as they can easily see the impacts induced by their decisions. In addition, the DT presents historic data to decision makers. It is therefore easier for them to assess past decisions and their implication on the transports.

Within this paper, we focus, however, on just one specific optimisation problem arising in the context of the intermodal transport platform. As transport demand is automatically matched to possible transport offers (which are envisaged to be also booked automatically in the future), horizontal collaboration is fostered since decision makers are not aware of other goods to be transported on the same truck/train/vessel. That is possible as the platform itself is neutral (i.e., not operated by any company involved in the transport sector other than providing an open booking platform). Therefore, decision makers can decide on the (for them) best transport option without any resentments against other competitors.

The remainder of the paper is organised as follows: First, we give an introduction into the actual collaborative roundtrip optimisation problem. Then, we outline possible parallels with other well-known combinatorial optimisation problems and discuss possible solution approaches. Finally, we conclude on the importance of research in this sector for a successful implementation of the PI vision.

2 The Collaborative Roundtrip Problem

The problem as arising in our intermodal transport platform PI demonstration pilot is as follows: Assume a container train connection between two major cities, e.g., Hamburg, Germany, and Vienna, Austria. Further assume that this train connection is operated for economic reasons like a block train, i.e., no waggons are un/coupled at intermediate stops, cf. (Jänsch, 2016). The obvious advantage of this approach is that ordering a block train is in most cases cheaper (on average per container) than ordering individual container transportation (VCÖ, 2020). That is, in general, the train is ordered by a freight forwarder. Shippers are ordering individual container slots from the freight forwarder. Therewith the prices for shippers can be reduced (compared to direct ordering at the carrier).

Now, further assume that along the trip from Hamburg to Vienna an intermediate stop is scheduled at Enns, Austria. Although the train is operated like as block train for the carrier, un/loading of containers at the intermediate stop are possible. That means, however, that even though the train might be fully loaded with 40 40-feet containers when leaving Hamburg only a part of them might be scheduled for Vienna while the other part is unloaded in Enns. In an optimal scenario, additional containers have to be transported from Enns to Vienna such that the then empty waggons can be used for these additional containers. Of course, additional containers might have to be transported from Enns to Hamburg, others from Vienna to Enns and even further ones from Vienna to Hamburg. In an even more complex scenario, one can imagine that there more than three cities along the roundtrip.

In a more general setting, we define the (basic) collaborative roundtrip problem (BCRP) as follows. We are given:

- a roundtrip for a train with n stops in cities $(c^1, c^2, \dots, c^n, c^{n+1} = c^1)$;
- a train schedule for this roundtrip, i.e., actual departure times for each of the n cities;
- a capacity, i.e., number of containers, for each scheduled train;
- travel times between city c^i and c^{i+1};
- transport demands, i.e., a set of containers, from city c^i to c^j;
- for each container, an earliest departure date and a latest arrival time, specifying when the container is earliest available at its departure city and has to arrive latest at its destination city;
- with each container costs for lifting and transporting are associated;
- each container has a weight.

The goal is now to find a transport schedule such that each container arrives before its latest arrival time at its destination city. In addition, the number of needed roundtrips shall be minimised, i.e., the latest arriving container should arrive as early as possible. As a second, subordinated optimisation goal the number of container lifts should be minimised. Furthermore, the (total) costs should be minimised.

Figure 1: Example of a simple roundtrip A-B-C-A with three cities.

An extended version of this BCRP is the what we call subtour collaborative roundtrip problem (SCRP). For the SCRP, at least on pair of cities c^i and c^j exists, with $i \neq j$; $i, j \neq 1$; $i, j \neq n$, such that $c^i = c^j$. For example, a roundtrip of type A-B-C-B-A, cf. Fig. 1.

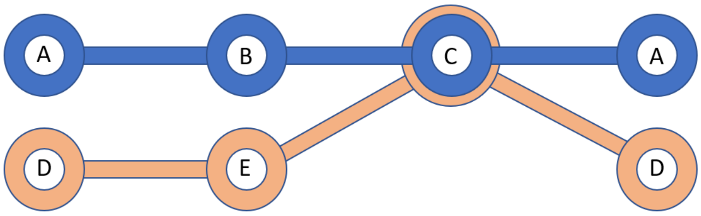


Figure 2: Example of a two overlapping roundtrip A-B-C-A and D-E-C-D.

A further extended version of the BCRP (and SCRP) is when multiple roundtrips exists with at least one common city. For example, roundtrip 1: A-B-C-A and roundtrip 2: D-E-C-D, cf. Fig. 2. Container transport from all cities to all other cities might occur such that transhipment from one roundtrip to another roundtrip at the common cities have to take place. We refer to this variation as multiple collaborative roundtrip problem (MCRP).

So far, we have no theoretical results on the computational complexity of these problems. However, some similarities with well-known combinatorial optimisation problems can be identified. We will discuss in more detail in the next section together with possible (algorithmic) solution approaches.

3 Computationally Solving the Collaborative Roundtrip Problem

3.1 Similarities with Well-Known Combinatorial Optimisation Problems

Rather obvious is a close relation to classical scheduling problems. Especially machine scheduling (Lenstra et al., 1977). The main goal of machine scheduling problems is to schedule a given set of jobs on a given set of machines. The goal is to find a schedule such that the make span, i.e., the finishing time of the latest job, is as early as possible. When considering now that containers are jobs and waggons are machines, the similarities are obvious. However, due to the site constraints, it is not so easy to directly link these two problems with each other. As extensively illustrated in (Lenstra et al., 1977), there are different version of machine scheduling which are computationally easy, i.e., are member of the class P of deterministic polynomial solvable problems. Other variants are, however, member of NP, i.e., there are no deterministic polynomial time solution approaches (unless $P = NP$).

As machine scheduling is closely related to knapsack problems (Lenstra et al., 1977) the similarities of the collaborative roundtrip problems to them is obvious.

3.2 Algorithmic Solution Approaches

Based on the observations in the previous subsection, i.e., the similarities with machine scheduling, it is obvious that (successful) computational approaches for machine scheduling should be followed. Therefore, we will examine them in more detail. Especially the basic variant of the CRP – or at least instances with, e.g., limited number of cities, are very likely to be polynomial solvable. Therefore, solution approaches based on integer linear programming (Nemhauser and Wolsey, 1999) or on dynamic programming are promising (Bellman, 1954).

However, for those variants which cannot be tackled via exact algorithms, local search-based approaches seem to be promising as local search operators can be naturally defined for this problem. Therefore, we propose to use metaheuristic approaches like variable neighbourhood search (VNS) and variable neighbourhood descent (VND), cf. (Mladenović, and Hansen, 1997). Operators which can be used for (efficiently) defining a (large) set of the necessary neighbourhoods will be based on swapping (exchanging container-to-waggons-assignments) and shifting (circular exchanging multiple container-to-waggons-assignments) operations.

4 Conclusions

Within this paper, we formulated the collaborative roundtrip problem (CRP) and variants of it. It arises in scenarios where containers need to be transported from one city (hub) to another one on an individual level, the carrying train, however, is operated on a roundtrip. The goal is to best possible utilise the train, i.e., to minimise the number of empty waggons such that the number of totally needed roundtrips/trains can be minimised. Due to its similarities to machine scheduling, it might be assumed that some variants of these problems can be time-efficiently solved via exact algorithmic approaches like inter linear programming or dynamic programming. However, more complex variants, e.g., those involving multiple roundtrips, are most likely too complex such that only (meta-)heuristic approaches can be meaningfully be applied as for larger real-world sized problem instances exact approaches will not come to an end.

The herein formulated CRP is not new as this problem already arises for (rail) carriers and/or freight forwarders on a daily basis. However, to the current status quo, the scheduling of containers on the train is on the one hand not very flexible and on the other hand only basically optimised. That is, a first-come, first-serve strategy is mainly applied when booking containers slots.

We, however, suggest that based on the complexity which can be easily handled via a (automated) digital transport mediation and booking platform efficiency of transportation can be increased. This increase in efficiency will, obviously, result in reduced transportation costs. At the same time the emitted CO_2e can be reduced as the increase in efficiency is mainly related to a higher utilisation rate of trains. Furthermore, the attractiveness of trains over trucks is increased and further shift towards trains can be expected. Finally, and not to neglect, is the positive impact on the society as a whole. Due to a modal shift towards train, it can be expected that negative by-products of (long-haul) road transports like decreased air quality and therefore deaths (WHO, 2021) are reduced.

We want to furthermore highlight that intermodal transport booking platforms ease the process of intermodal booking through their one-stop-shop mentality. Therefore, the hurdles of booking intermodal transports are significantly reduced. E.g., for a standard intermodal transport of first-mile, long-haul, last-mile only one instead of five bookings (first-mile, transhipment, long-haul, transhipment, last-mile) are necessary. Even though already nowadays the booking of intermodal transports could be outsourced to freight forwarders, there is still a person at the freight forwarder who has to organise all of the five bookings. That is, making intermodal as easy as unimodal transport one can expect a major impact.

In addition, it is essential that the booking platform is neutral. That is, that it is not operated, financed, or owned by one (or more) companies having a major interest in the transport sector. E.g., if a large freight forwarded or shipping company would operate such a platform it is most likely that other players on the market would not provide their transport capacities on this platform. This is since then the platform operator would have access to additional (and sensitive) data about capacities and transport operations of competitors.

The major goal of such a transport platform must be, however, that the booking processes are fully automated. That is, that it is not necessary for humans to interact with the platform on a daily basis. As soon as this fully automated booking process could be realised it will become easy to step forward from intermodality towards synchromodality. That is, since as soon as an interruption occurs on one of the booked transports the platform can automatically decide whether another transport option leads to better (e.g., more reliable) results. Furthermore, the platform can objectively decide whether a shift is beneficial for the system (and not only for the individual transport). It is therefore possible to optimise towards a system optimum (compared to a user equilibrium which is known to be inferior from a societal perspective), cf. (Wardrop, 1952).

However, in order to be able to automate the booking processes, it is essential that the needed computations in the background can be performed efficient (and fast). Therefore, it is crucial to investigate the underlying optimisation problems in more detail and to come up with fast, yet good solutions. We therefore suggest to further investigate the CRP (and other optimisation problems) arising in the context of the PI.

5 Acknowledgements

We would like to thank the partners in the PhysICAL project for the interesting discussions on the topic of the PI. Especially we would like to thank Nils-Olaf Klabunde from 4PL Intermodal for the intensive insights given in the transport booking platform laying the basis of this work.

This work received funding by the Austrian Federal Ministry for Climate Action, Environment, Energy, Mobility, Innovation and Technology (BMK) in the research program "Mobilität der Zukunft" under grant number 877710 (PhysICAL).

References

- Accenture (2021): The sustainable last mile. Faster. Cheaper. Greener. https://www.accenture.com/gb-en/insights/consulting/sustainable-last-mile-delivery (last visited 10.5.2021).
- Bellman, R. (1954): The theory of dynamic programming, Bulletin of the American Mathematical Society, 60 (6): 503–516.
- Honeywell (2016): Honeywell annual report, Washington, USA: Honeywell International Inc.
- Jänsch E. (2016): Handbuch: Das System Bahn, 2nd edition, DVV Media Group, 2016.
- Lenstra, J. K., Rinnooy Kan, A. H. G., & Brucker, P. (1977). Complexity of machine scheduling problems. Annals of Discrete Mathematics, 1, 343-362.
- Mladenović, N.; Hansen, P. (1997). Variable neighborhood search. Computers and Operations Research. 24 (11): 1097–1100.
- Nemhauser G., Wolsey L. (1999): Integer and Combinatorial Optimization. Wiley Interscience, New York 1999.
- PhysICAL (2021): The PhysICAL project website. https://physical-project.at/ (last visited 10.05.2021)
- Statistik Austria (2010): Agrarstrukturerhebung 2007.
- VCÖ (2020): *Güterverkehr auf Klimakurs bringen*, VCÖ-Schriftenreihe „Mobilität mit Zukunft", VCÖ, Austria, 2020.
- Wardrop, J. G. (1952): *Some Theoretical Aspects of Road Traffic Research*. Proceedings of the Institution of Civil Engineers. 1 (3): 325–362.
- WHO (World Health Organisation) (2021): Air pollution. https://www.who.int/health-topics/air-pollution#tab=tab_1 (last visited 10.5.2021)

Technical State-of-the-Art of Supply Chain Visibility

Andrea d'Auria[1], Wout Hofman[1], Erik de Graaf[1] and Jeroen Breteler[1]
1. TNO, Soesterberg, Netherlands

Abstract: *This article presents a literature study on the state-of-the-art of the technology adopted for Supply Chain Visibility (SCV) and identifies blockchain as the major player for innovation. The first section highlights the trade-off between immutability and confidentiality in distributed ledger technologies, and discusses the mismatch of interests between academic research and enterprise research and development. Then, three private blockchain platforms (Quorum, Hyperledger Fabric and Corda) are evaluated as solutions for confidentiality in shared databases. The third section examines the role of standards for logistics data. The paper concludes that blockchain is the number one technology contemplated for SCV both in academic discourse and in the private sector, however the former favors immutability and transparency while the latter favors confidentiality. By making this trade-off explicit and analyzing the different solutions offered by the main related technologies, this work contributes to a clearer identification of advantages and disadvantages of options currently available. Additionally, the role of standards is identified as a valuable topic for further research, and an open standard for SCV is deemed to be desirable to ensure interoperability between systems relying on different technology stacks.*

Conference Topic(s): *Systems and technologies for interconnected Logistics (blockchain, digital twins), New communication Networks enabling interconnected logistics*

Keywords: *supply chain visibility, blockchain, distributed ledgers, decentralization, digital twins, standards, data models, confidentiality*

1 Introduction

The future of interconnected logistics will rely on faithful representation of reality in data spaces. The movement of physical goods along supply-chains can no longer be considered as a separate self-standing instance from the data describing such movement or the digital twins describing the goods themselves. This interconnection creates new possibilities – like a higher degree of transparency for Supply Chain Visibility (SCV) – but creates new challenges as well – like the management of shared databases and the confidentiality of related data. This article presents the state of the art regarding the technological direction that SCV is taking both from an academical point of view and with an eye kept on the market. A selection of the most relevant scientific papers has been used to analyze what requirements are privileged and met. The analysis was further elaborated by integrating the technical knowledge on platforms that can support SCV.

In this work the concepts of transparency, visibility and traceability are used in the context of supply chain and data sharing, and they are closely related one to another: transparency, in supply chains, refers to the infrastructure layer where data is confined. Having this layer be transparent means that data in there is visible. Visibility refers to the ability of accessing and seeing a piece of data. Traceability is the capability of following "traces" of an object along a supply chain, for example, one could trace a cargo object being moved from one location to another. Such traces are data about events involving the object. The concept of traceability is inherently tied to the one of visibility: in order to track an object, one needs to have full visibility on the data that makes up the trace.

More specifically, Supply Chain Visibility can be defined as "awareness of and control over end-to-end supply chain information – including insight in sources of data and whereabouts of goods – enabling agile, resilient, sustainable as well as compliant and trusted supply chains" (Wieland & Wallenburg, 2013).

In section 2 of this document the current state of research is investigated, highlighting the difference between academic and enterprise environments when coming to preferred features and requirements, and a possible explanation for this mismatch is given. The trade-off between immutability and confidentiality is made explicit and formalized.

In section 3 a deeper insight about how the major private blockchain platforms (Quorum, Hyperledger Fabric and Corda) handle private data is given as well as a further elaboration about how their design choices relate to the immutability-confidentiality trade-off.

In section 4 the role of standards is introduced as well, and an overview of the major ones is given.

2 State of the Art

Scientific papers taken into account for the following work are mostly from IEEE, as it is the leading organization for technology advancement and a good compromise to narrow down the literature that could also be found elsewhere.

Through use of the IEEE search engine it emerged that about 41% of the results (20 out of 49) related to SCV in the last two years (2019-2020) also involved blockchain. This highlights a strong interest for distributed ledger technologies (DLT) for SCV.

Even among those not explicitly mentioning blockchain technology, some underline clear requirements for data transparency and smooth coordination (Yanamandra, 2019), which relates to DLT because transparency can be easily achieved with replicated databases, and the use of smart contracts can smoothen and automatize coordination between different stakeholders.

The first general remark that one can observe from a perspective of the functionalities required, is that most of the papers point their attention to visibility of data, tamper-proof characteristics, and trust among stakeholders. The main concerns are, therefore, driven by the data that is currently not available along the supply chain, the data being altered because of self-interest of one party against the others, and the lack of trust between such parties.

Consequently, within many academic papers DLT is identified as a good opportunity to tackle those issues.

However, there is a mismatch in expectations between the academical environment and the private sector, supply and logistics stakeholders. This discrepancy is especially evident when taking into consideration the transparency requirement. Transparency means that data will be available and auditable (visible) by all parties joining the network, and the academical works examined in this context very much rely on this assumption. However, transparency – which also relates to a higher degree of immutability – comes at the cost of confidentiality – which is an important requirement for companies. Confidentiality relates to commercial sensitivity of business relations that will be made transparent when blockchain technology is applied for SCV.

In order to better understand the nature of this trade-off, mentioned by Kannengiesser et al (2019), it is worthwhile to highlight what is the link between transparency and immutability. Blockchains are considered a subset of DLT, with the added characteristic that data is replicated at each node. This complete replication is fundamental to reach a consensus in the network and every node is responsible to keep and validate the whole set of transactions from the first genesis

block. Immutability can be guaranteed because the infrastructure is transparent and data is fully visible. As soon as a malicious agent operating a node tried to tamper with data previously shared with the network, such attempt would be immediately evident and could be prevented. The opposite situation is one where only one central non-transparent database exists, with users relying on its data. In this case a malicious agent controlling the central database could tamper with the data, and as long as other parties do not have visibility on historical data it is impossible for them to assess whether there have been alterations.
Most of the investigated papers did not work on a real-world use-case dealing with true needs of a company, and assumed that full transparency and data sharing among stakeholders was a fair assumption. Nevertheless, when the research deals with a real enterprise use-case, the need for confidentiality over data emerges clearly (d'Auria, 2020).
This ambiguity also showed up during the conference of the Blockchain Observatory of the Politecnico di Milano of January 2020: while their research highlighted that the future of the technology relies on public blockchains, almost all the companies invited on the stage were actually developing private solutions (Perego et al, 2020).
The demand for confidentiality over data is stronger among companies, and this is reflected in the implementation choices they make, as shown in Table 1.

Company	**Platform**	**Field**
ABI	Corda	Finance
Adledger	Hyperledger Fabric	Advertising
B3I	Corda	Insurance
Fnality	Private Ethereum	Finance
Food Trust	Hyperledger Fabric	Agrifood
HM Land Registry	Corda	Notarization
Komgo	Private Ethereum	Finance
LO3 Energy	Private Ethereum	Energy Utility
LVMH Aura	Quorum	Luxury
Marco Polo	Corda	Finance
Santander	Ethereum	Finance
SDX – Six Digital Exchange	Corda	Finance
Tradelens	Hyperledger Fabric	Logistics
uPort Zugo	Ethereum	SSI
Vakt	Quorum	Utility
Verified.Me	Hyperledger Fabric	SSI
Vinturas	Hyperledger Fabric	Logistics
Voltron	Corda	Finance
We.trade	Hyperledger Fabric	Finance

Table 1: Platforms chosen by companies for implementation of their services[4].

The data shown in Figure 1 clarifies that while scientific papers are focused mainly on the features of immutability and trustlessness, companies try to pull more towards the confidentiality side of the trade-off.

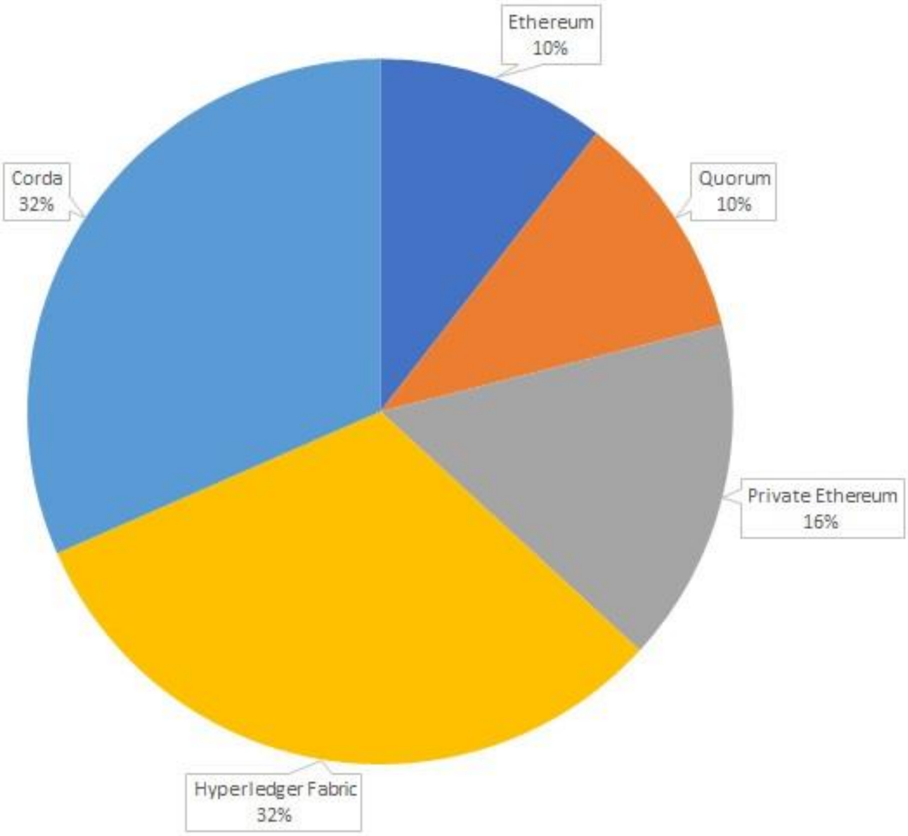


Figure 1: Platform distribution according to Table 1.

One approach (Hofman, 2020) goes deeper into formalizing the functionalities required and assessed that a platform for SCV should integrate:

- A subscription mechanism;
- Digital twins;
- Milestones and events.

These concepts are compatible with DLT.

3 Technologies along the spectrum of the trade-off

While the DLT landscape still maintains some characteristics of chaotic behavior – with new solutions and platforms being born and abandoned frequently – some tools are well established in the environment and regarded as solid and reliable, as the previous chart shows. In the following sections a deeper insight on how the main DLT platforms (namely: Quorum, Hyperledger Fabric and Corda) position themselves along the spectrum of the trade-off discussed in this article is given.

3.1 Quorum

In this article we refer to Quorum as the generic name for the project developed under ConsenSys umbrella which involves two pieces of software: GoQuorum and Hyperledger Besu. The choice of referring to it as one is motivated by the fact that they share the same design choices, and being under the same umbrella they will increase their interoperability over time.

Quorum is an Ethereum-like private blockchain which emulates Ethereum behavior and protocol in a confined environment – i.e. not necessarily open.
On top of the standard Ethereum-like functioning, Quorum provides private transactions (which translates into private smart contracts) allowing only specific nodes to receive private transactions.
When a transaction is submitted privately to the network (a call to a private smart contract) it originates a split in the state of receiving nodes:

- Everyone will receive a transaction with a hashed payload, which will not imply any update of the state per se;
- Only specific nodes will receive the plaintext transaction in a peer-to-peer communication, and will store it in a separate database, and link it to the hashed transaction in the main database, causing an actual update in the state;

So different nodes will see different states according to what they are allowed to see.
Quorum allows also the following "privacy enhancement" features:
Counter Party Protection (CPP): it prevents non-participant interaction on a private contract but allows state divergence (i.e. it will allow nodes to maintain different state through private transaction to "self" or "subset of nodes"). Nodes which are not part of a private smart contract, are still able to submit transactions to a private smart contract by default. They will not be able to record the update of the state, though. If CPP is enabled this behavior is prevented and it can substitute some strategies for access control.
Private State Validation: On top of all the verifications of CPP, it introduces further checks in order to prevent nodes from state deviations. Private transactions to "self" or "subset of nodes" will then fail, the full list of recipients is shared among all nodes and all transactions are validated against it. In standard private smart contracts – or with only CPP enabled – only the sending node knows – and it is free to choose – the list of recipients.

3.2 Hyperledger Fabric

Hyperledger Fabric is an open source private blockchain created by IBM, and it allows deep customization opportunities when coming to private data.
The offered options revolve around two main concepts:

- Channels
- Collections

Channels are wholly separate blockchains within the same networks, i.e. nodes and identities are kept within the network but different subset of nodes can run different blockchains, and there is neither communication nor access between different channels.
Collections are an implementation of private transactions within a channel. The core principle is the following: private transactions are hashed before submission to the ordering service – which appends them to the ledger together with public transactions – and those authorized to see such transactions have a "Private State Database" where plaintext is stored. Private data is disseminated encrypted peer-to-peer among nodes.
Important considerations:

- You do not necessarily have to be a member of a collection to write to a key in a collection, as long as the endorsement policy is satisfied. Endorsement policies can be defined at chaincode level, key level (using state-based endorsement), or collection level. In other words also nodes not allowed to read private transactions in a collection, are allowed to write a private transaction (visible to others) as long as they can make it consistent;
- Data hashes can be verified at chaincode level, meaning that one can prove that a transaction was already submitted to the ledger.

One should opt for channels when entire sets of transactions (and ledgers) must be kept confidential within the set of organizations that make up the channel. It is as effectively running different blockchains within the same network: those outside a channel do not have access to any of its information.
One should opt for collections when transactions (and ledgers) must be shared among a set of organizations, but when only a subset of those organizations should have access to some (or all) of the data within a transaction.

3.3 Corda

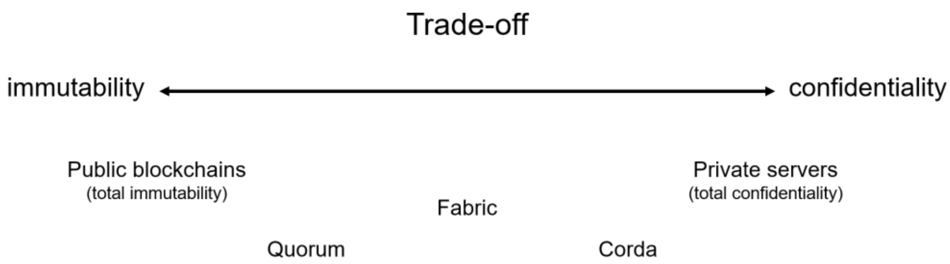
Corda is an open source platform backed by R3, which functions in a different way than standard blockchains. The platform is highly focused on confidentiality, yet borrowing concepts from blockchain to provide a certain degree of immutability.
Corda has two main entities participating to the network:

- Nodes
- Notaries

Notaries receive and store hashes of transactions and can validate them; nodes submit transactions on a need-to-know basis. Privacy is therefore granted by default because notaries, which have a broader overall view of the transactions happening in the network, only store fingerprints of them. Nodes have a full view of transactions they submit and receive, but cannot see transactions not related to them, resulting in an only partial view of the ledger.
Within Corda nodes can have a different view of what is there in the ledger, and consistency – and immutability – is guaranteed by notaries.
Communication develops on two levels: nodes disseminate data peer-to-peer, notaries disseminate transactions between them according to a pre-set consensus mechanism.

3.4 Considerations and Comparisons

One can roughly position the examined platforms as follows:



Hyperledger Fabric resulted in a very versatile tool thanks to the wide range of options that allow to use it both as very open and transparent classical blockchain and as a more closed environment with privacy measures in place. In this, it shares some similarities with Quorum for the way it implements the concept of collections: both are based on storing "publicly" (within the network) only a fingerprint of the actual data stored in a separated, confined, database. However, Fabric took this a step further with channels – wholly separated non-communicating blockchains within the same network of nodes – which made it get a step closer to confidentiality and one further from immutability.
Corda, as shown in the previous section, differs more significantly from both the other two and from blockchain platforms in general, as the concept of a replicated database owned by every node – crucial in blockchains – is not implemented in Corda. This makes it more

privacy-prone than the major counterparts, however it also results in a lower degree of immutability. Considering this design is the default in Corda, it pushes it towards the “confidentiality” side of the spectrum.

The borders presented in this paper are not fixed and rigid, they aim to give a rough overview on how different design choices can influence immutability and confidentiality instead: Quorum could enable strict policies and use only private smart contracts moving more to the centre, and Corda, likewise, could disseminate all the data openly and transparently with the whole network. This, in both cases, would imply not fully taking advantages of the strengths of the different design choices that were made for the various platforms.

Being at the “center” of the spectrum does not necessarily mean being in the best position. Saving the most of both equally means losing the most of both. The choice of a platform should rely on a more ad-hoc analysis of what is more important for a specific use-case and on the actual state of development of the platform and community support.

4 The Role of Standards

Going beyond technology, only a small portion of the examined papers took the requirement for a common data model into account, and only one paper considered it a crucial step for the achievement of effective software development (Grest, Lauras et al, 2019). As the industry is looking with interest at the subject, some standards are developed to support SCV. A common data sharing infrastructure – like a blockchain – can enforce a data model coded into smart contracts for everyone joining the network. However, such a system must be able to interoperate with external systems and thus share a common grammar. Agreeing on a common standard can help both to enhance adoption of a shared infrastructure, and to ease interoperability between different ones.

Besides these standards, there are also various platforms providing visibility, either with proprietary interfaces or based on (open or de facto) standards. Examples of these platforms are Tradelens (visibility of container transport via sea, developed as joint initiative by Maersk and IBM) and Transfollow (visibility of road transport).

We can find among others the following SCV-related standards:

- OpenTripModel (OTM)
- EPCIS Data Model (GS1 Standard)

4.1 OpenTripModel

OpenTripModel (OpenTripModel Documentation, 2021) (OTM) was funded by the Dutch government, developed initially by Simacan but now public and open for contribution. Its maintenance is organized by SUTC (Stichting Uniforme Transport Code), governed by the major road transport association TLN and the association of shippers with their own transport and forwarder (EVO/FENEDEX). Some important companies (DLG and PostNL Transport among others) have already experimented with it.

The main characteristics are the following:

- The model is independent of how transport within a supply chain is organized;
- The model is intended to be independent of any transport modality, although it is not applied as such and SUTC does not have participation of other transport modalities beside road;
- The data is human and machine-readable;
- The model is extensible;

The model is based on the concepts of Lifecycle, Entities and Events. The Lifecycle provides a context to the phase of an operation – "planned", "projected", "actual" or "realized".
Entities are the basic components of the supply chain – location, vehicle, actor, etc. Entities could be envisioned as Digital Twins. Events are also Entities, but they are tied to a Lifecycle. They model the relations between Entities.

4.2 EPCIS Data Model – GS1

The basic unit of data in EPCIS (Electronic Product Code Information System) is the EPCIS event (GS1 Official Website, 2021), which is a structure that describes the completion of one business step within an overall business process. A group of EPCIS events provides a detailed description of a business process.
Each EPCIS event is composed of the following:

- **What**
  The identifiers of the objects related to the event;
- **When**
  The date and time when the event took place;
- **Where**
  The identifier of the location at which the event occurred, and identifier of the location where the objects are expected to be following the event;
- **Why**
  Information about the business context.

This standard is currently adopted by for instance the Metro Group.

4.3 Considerations on Standards

Standards are a complex issue. Standards for interfaces, products, or services need to be produced in such a way that implementers can access them (publicly available, potentially against low costs) and can assess how the standard was developed. Standards can be categorized according to the nature of their development and maintenance processes:

- Open standards – those that are developed by recognized standardization bodies
- De facto standards – those that have not been developed and are not maintained by a recognized standardization body, but are used by most organizations.
- Proprietary standards – those that have been developed and are maintained by a limited group of organizations. The procedures are not transparent.

In this view, OTM can be considered as a proprietary standard. It is not yet a de facto one. TradeLens, for instance, develops a de facto standard for SCV for sea transport. It has been the basis for Vinturas, where IBM also participates in the development of the solution.

5 Conclusion

Blockchain is the number one technology contemplated for Supply Chain Visibility in research. However, there is a distinct view between academic research and research and development in the private sector: the former supports transparency and the latter requires confidentiality. Given this, research is required into access control and data management for SCV in blockchain technology. Different technologies have different solutions to this challenge. In the realm of private blockchains the major actors made different design choices in order to offer to the market different capabilities in terms of immutability and confidentiality. While the needs for immutability and confidentiality are the main drivers when choosing a platform for a given use-case, this trade-off often goes undiscussed in the literature. By making it explicit and analyzing it in this work, we have highlighted the possibilities and opportunities that DLT – blockchain technologies in particular – open for SCV, and at what cost they come. DLT can in fact

effectively enhance immutability and visibility of supply-chain data; at the same time a careful choice of the platform and a meticulous tweaking of its privacy features can effectively mitigate the loss of confidentiality.
Secondly, the role of standards is identified as a valuable topic for further research. In practice there are multiple efforts to develop a proprietary or de facto SCV standard. However, an SCV standard is rarely considered a relevant matter in the context of SCV despite the crucial role that has in the enhancement of interoperability. Solutions with a proprietary interface achieving market adoption are preferred by enterprises (the so-called network effect). Given this, it is useful to come to an open standard for SCV, i.e. a standard that is developed and maintained by a recognized standardization body. Once this is available, various SCV solutions can be developed. It would be of value also to map for instance OTM to the functionality described in Hofman (2020).

References

Consensys. Retrieved 2021, from GoQuorum Official Documentation: docs.goquorum.consensys.net/en/stable/

d'Auria, A. (2020). Porting blockchain smart contracts - A feasibility study.

Divey, S. J., Hekimoglu, M. H., & Ravichandran, T. (2019). Blockchains in supply chains: potential research directions. *2019 IEEE Technology and Engineering Management Conference (TEMSCON)*.

Grest, M., Lauras, M., Montarnal, A., Sarazin, A., & Bousseau, G. (2019). A meta model for a blockchain-based supply chain traceability. IEEE.

GS1 Official Website: www.gs1.org/standards/epcis - Retrieved 2021

Hofman, W. (2020). Semantics - Supply Chain Visibility (SCVL) - discussion paper.

Hyperledger Fabric Official Documentation: hyperledger-fabric.readthedocs.io/en/release-2.2/ - Retrieved 2021

Jangir, S., Muzumdar, A., Jaiswal, A., Modi, C. N., Chandel, S., & Vyjayanthi, C. (2019). A Novel Framework for Pharmaceutical Supply Chain Management using Distributed Ledger and Smart Contracts. *2019 10th International Conference on Computing, Communication and Networking Technologies (ICCCNT)*.

IEEE Search Engine: https://ieeexplore.ieee.org/search/searchresult.jsp?queryText=supply%20chain%20visibility&highlight=true&returnType=SEARCH&matchPubs=true&ranges=2019_2020_Year&returnFacets=ALL&searchWithin=blockchain - Retrieved 2021

Kannengiesser, N., & Lins, S. (2020). What Does Not Fit Can be Made to Fit! Trade-Offs in Distributed Ledger Technology Designs. *Hawaii International Conference on System Sciences (SSRN Electronic Journal)*.

Karuanchi, M. D., Sheeba, J., & Devaneyan, S. P. (2019). Cloud Based Supply Chain Management System Using Blockchain. *2019 4th International Conference on Electrical, Electronics, Communication, Computer Technologies and Optimization Techniques (ICEEECCOT)*.

Kshetri, N., & Loukoianova, E. (2019). Blockchain adoption in supply chain networks in Asia. *IT Professional*.

Li, Z., Wu, H., King, B., Miled, Z. B., Wassick, J., & Tazelaar, J. (2018). A Hybrid Blockchain Ledger for Supply Chain Visibility. *2018 17th International Symposium on Parallel and Distributed Computing (ISPDC)*.

OpenTripModel Documentation: www.opentripmodel.org/docs - Retrieved 2021

Perego, A. (2020). Annual Report.

Pundir, A. K., Jagannath, J. D., Chakraborty, M., & Ganpathy, L. (2019). Technology Integration for Improved Performance: A Case Study in Digitization of Supply Chain with Integration of Internet of Things and Blockchain Technology. *2019 IEEE 9th Annual Computing and Communication Workshop and Conference (CCWC)*.

R3. (n.d.). Retrieved 2021, from Corda Official Documentation: docs.corda.net/docs/corda-os/4.6/key-concepts.html

Raj, R., Rai, N., & Agarwal, S. (2019). Anticounterfeiting in pharmaceutical supply chain by establishing proof of ownership. *TENCON 2019 - IEEE Region 10 Conference*.

Sahoo, M., Singhar, S. S., Nayak, B., & Mohanta, B. K. (2019). A blockchain based framework secured by ECDSA to curb drug counterfeiting. *10th International Conference on Computing, Communication and Networking Technologies (ICCCNT)*.

Singi, K., Bose, J. C., Podder, S., & Burden, A. P. (2019). Trusted Software Supply Chain. *2019 34th IEEE/ACM International Conference on Automated Software Engineering (ASE)*.

Transfollow Official Website: transfollow.org - Retrieved 2021

Wieland, A., & Wallenburg, C. M. (2013). The influence of relational competencies on supply chain resilience: a relational view. *International Journal of Physical Distribution and Logistics Management*.

Yanamandra, R. (2019). A framework of supply chain strategies to achieve competitive advantage in digital era.

Impact of High Capacity Vehicles on the future developments in the Logistics sector

Andreas Lischke, Stephan Kirsten[1] Tim Bremeersch [2]
Christoph Jessberger [3]

1. DLR German Aerospace Center, Berlin, Germany
2. Transport & Mobility Leuven, Leuven, Belgium
3. MAN Truck & Bus SE, Munich, Germany

Corresponding author:
Andreas Lischke
DLR German Aerospace Center
Institute of Transport Research
Rudower Chaussee 7
12489 Berlin
andreas.lischke@dlr.

Abstract:

The reduction of the carbon dioxide emission in road freight transport in the next decades is a key issue. Focussing on this challenge, we analyse the impact of high capacity road transport with longer and heavier-trucks (European Modular System: EMS some examples see Table 1) on mode choice and CO_2e emissions at the EU 28 level. The presented results are part of research done in the EU project AEROFLEX (Aerodynamic and Flexible Trucks for Next Generation of Long Distance Road Transport) funded in the H2020 programme.

AEROFLEX project aimed to increase efficiency up to 33 % in long distance road transport and logistics. For assessing the impacts of these new vehicle types, this article describes the several approaches that are used to determine the expected impacts.

Finally, we address that AEROFLEX road transport innovations can take a role in the physical internet that is similar to that of broadband wireless connections in the digital internet: ultra flexible, capable of moving high volumes at high speeds, with the best possible coverage at much greater efficiency than past technologies.

Conference Topic(s): *Interconnected freight transport, logistics and supply networks and Vehicles and transshipment technologies*

Keywords: *Consolidation Center Commodities Emission Reduction H2020 Project Intermodal Load Factor Corridors, Hubs and Synchromodality*

1 Introduction

This article is structured into four chapters describing the impacts of high capacity vehicles of European Modular System: EMS 1 and 2 using an Advanced Energy Management Powertrain (AEMPT) with an e-dolly and aerodynamic-optimized tractors and semi-trailers (some examples, see Table 1). In the second chapter we quantify the impact from a use case perspective for selected commodities to show the improvements on logistics concepts or logistics pattern based on data collected through interviews with logistics service providers. In the third chapter, it is evaluated by freight transport modelling the impact of efficiency increases

- due to a higher average load factors (related to AEROFLEX innovations of Wabco CargoCam/Fraunhofer puzzle software). Further, it is considered already existing double-stack loading to increase cargo consolidation and cost reductions due to higher transport capacity per truck on road enabled by EMS 1 and 2 (related to tonne-kilometres due to higher weight and length limitations of EMS). The increase of efficiency on road transport is implemented in modelling parameters for these bot new vehicle configurations and shows the impact on freight transport at the EU 28 level based on several scenarios. In the last chapter, the impact on Physical Internet (PI) coming from innovations of AEROFLEX are elaborated.

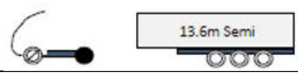
2 Impact of High Capacity Vehicles on logistics

Based on the network of the AEROFLEX project partners and sounding board members several online stakeholder surveys and in-depth expert interviews amongst logistic service providers (LSP) and shippers have been conducted. Out of this, 32 different use cases have been created. The regarding tours involved 19 countries, either as origin, destination or transit country (EU countries as well as Serbia and Turkey). Combined, 171 different combinations of tours, vehicle and load variants have been analysed. 24 Prime Candidates have been chosen by interviewees from a total number of 27 Prime Candidates as possible vehicle concepts to be used.

2.1 Combined results of interviews

Interviewees were also asked to select Prime Candidates per logistics segment and route type combination, which could be used in daily business providing biggest potential for economical and logistical benefits from their perspectives. The approach to use European Modular System (EMS) vehicles to improve efficiency is based on load consolidation as a crucial factor to realize the expected benefits. Thus, the impact of the use of the Prime Candidates is analysed with regard to the KPIs: €/tkm, €/tour and CO_2e [kg] emissions tank-to-wheel (ttw) and well-to-wheel (wtw). About 53 % of the interviewees vote for the following six most relevant Prime Candidates (in descending order of vote share): 6.1, 2.1, 3.1, 1.4, 2.2 and 4.7 (see Table 1). The shares ranged from 11.7 % to 6.2 %. An additional 10.1% was achieved by Prime Candidate 1.3, which is a standard 4x2 tractor unit with a 13,62 metres semi-trailer.

Table 1: Share of votes by interviewees of preferred Prime Candidates

No.	Prime Candidate	Share of votes
6.1		11.7 %
2.1		9.7 %
3.1		9.7 %
1.4		9.3 %
2.2		6.6 %
4.7		6,2 %
1.3		10.1 %

In addition, standard average loads by reference vehicles are compared to the maximum load for Prime Candidates to calculate average mean values and standard deviations of each KPI (see above). These mean savings potentials in percentage values for different KPIs for the overall sample are displayed in Table 2.

Table 2: Mean saving potential for overall sample in % for different KPI. Standard deviation in parenthesis. Negative values indicate advantages for the Prime Candidates.

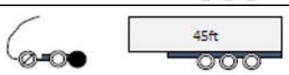
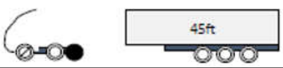
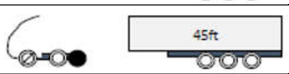
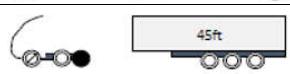
KPI	€/tkm	Cost/tour	CO2e TTW	Co2e WTW
Standard average load	18.7% (10.9)	19.0% (11.2)	28.8% (17.0)	20.9% (11.3)
Maximum load; average savings for all use cases	-28.2% (16.4)	-28.1% (16.5)	-16.9% (14.4)	-25.8% (33.7)

2.2 Results of two selected uses cases

To show the overall benefit in an exemplary way, the following two use cases are selected. Each use case shows the potential efficiency gain by shifting reference vehicles to EMS 1 or EMS 2 for a specific todays transport. The first use case reflects an intermodal logistics chain on road and water ways and involves multiple countries (Netherlands, Germany and Finland). Using Prime Candidate 6.1 (i.e. EMS 2) enables 74 tons instead of 40 tons Gross Combination Weight (GCW) and results in a CO_2e emission reduction potential of -129.6 kg or -25.8 % on one tour. The second use case distinguishes from the first use case and gives an explanation about EMS 1. In this case a single mode logistics chain (only road) is reflected by a tour between Germany and Austria using Prime Candidate 3.2 (i.e. EMS 1) with a maximum of 60 tons instead of 40 tons GCW permissible. Due to the lower transport distance between origin and destination an emission reduction potential of only -72.0 kg CO_2e could be achieved. Nevertheless, this is equivalent to a CO_2e potential of -32.4 % on one tour.

In relation to these two use cases table 3 shows the theoretical benefit of EMS 2 and EMS 1. Only 1 instead of 2 vehicles (EMS 2) and only 3 instead of 4 vehicles (EMS 1) would be needed to transport (nearly) the same load as the reference vehicles.

Table 3: Prime Candidates and re-allocations in selected use cases

No.	Reference vehicles (similar to 1st use case)	No.	Re-allocation w.r.t. EMS 2 (e.g. PC 6.1):
1.1		6.1	
1.1			 (saved)
No.	**Reference vehicles (similar to 2nd use case)**	**No.**	**Re-allocation w.r.t. EMS 1 (e.g. PC 4.3):**
1.1		4.3	
1.1		1.1	
2.3		2.3	
2.3			 (saved)  (saved)

But beside these very positive theoretical effects of EMS 1 and 2, there are much more complex decisions to be taken on fleet level, which is factored in by the results from the overall sample (cf. Table 2). Thus, on fleet level up to 30 % of tractors and drivers in suitable use cases could be saved by using EMS 1 and 2.

3 Impact of High Capacity Vehicles on freight transport of EU

3.1 Methodology approach

This chapter describes a freight modelling approach to calculate the impact of EMS 1 and 2 related to EU freight transport in 2040 by considered impacts like (i) the possible shift in mode choice, (ii) the change road transport mileage, and (iii) expectations on CO_2e emissions. The two topics of access policy and infrastructure requirements for EMS 1 and 2 are not addressed. The project AEROFLEX elaborated these two topics in a special approach and the impact assessment gives input to the argumentation. Thus, the described scenarios do not consider existing regulations for High Capacity Vehicles in several EU countries as well as current EU regulation on the maximum authorised dimensions in national and international traffic and the maximum authorised weights in international traffic (Directive 2015/719/EC, 2015) or other restrictions.

3.1.1 Freight modelling

For our projection we use the macroscopic freight model 'DEMO-GV' (Burgschweiger et al., 2017). It calculates the transported goods between c. 400 German an c. 200 other European traffic cells. The goods will be transported via three transport modes: 'rail', 'road' and 'inland waterways' and indicate the transport modal split. The goods transport on road can be realized by seven road-vehicle types (GCW: Gross Combination Weight):

Table 4: Table: truck types used in modelling approach

(I) Truck 3.5 ≤ 7.5 t GCW	(II) Truck 7.5 ≤ 12 t GCW	(III) Truck 12 ≤ 18 t GCW
(IV) Truck 18 ≤ 26 t GCW	(V) Truck 26 ≤ 40 t GCW	

and with European Modular System (EMS) vehicles:

(VI) Truck 40 ≤ 60 t GCW (EMS 1)	(VII) Truck 60 ≤ 74 t GCW (EMS 2)

The share between all truck types is the mean split. Modal split and mean split are calculated separately for every NST-2007 commodity class (NST, 2007) and the combined transport (CT). The model DEMO-GV imports the data of average load factors and average transport costs for every vehicle-type. Due to EMS 1 and 2 configurations: there are reduced costs per tonne-kilometres based on a higher available transport volume and GCW as well as a higher average load factor (by using AEROFLEX innovations like the Wabco CargoCam/Fraunhofer puzzle software as well as already existing double-stack loading to increase cargo consolidation, which help to optimize the ratio of transport volume and weight).

The input parameters for cost calculations (e.g. average fuel consumption) were fixed bases on exchange with AEROFLEX project partners, who give input from the results of their testing and simulations. The relevant transport costs distinguish time and distance related costs of different transport modes and vehicle types.

3.1.2 Upscaling the results for EU 28

The modal split and the means split on the road of 'DEMO-GV' have to be upscaled to the EU level. First, we calculate the freight transport in tonne-kilometres (transport performance) tp at German level, multiplying the transport volume tv by the distance d between the cells at German level. The unit is tonne-kilometre [tkm]:

$$tp = tv \cdot d_{Germnany} \tag{1}$$

The next step is an extension on the freight transport performance tp which exists at EU 28 level. For this reason, we assume:

$$\frac{tp_{German,c,i}}{total\ tp_{German}} = \frac{tp_{EU-28,c,i}}{total\ tp_{EU-28}} \tag{2}$$

- $tp_{German,c,i}$ = Freight transport performance at German level for commodity c with mode i [tkm]
- $total\ tp_{German}$ = Total freight transport performance at German level [tkm]
- $tp_{EU-28,c,i}$ = Freight transport performance at European level for commodity c with mode i [tkm]
- $total\ tp_{EU-28}$ = Total freight transport performance at European level [tkm]

The assumption (2) is the result of the same mode ratios in Germany and the EU 28 (EUREF projection, 2016). Based on equation (2) and the total projected freight transport performance in EU 28 of EUREF in 2016, a disaggregated freight transport performance in EU 28 in 2040 is derived. The freight transport performance is disaggregated by NST-2007-classification and the three modes. The maritime transport as well as short see shipping were not considered due to the assumption that this transport mode would not be influenced by use of EMS 1 and EMS 2 in land-based transport.

Based on the freight modelling output in tonne-kilometre by vehicle type, we derive the annual mileage using the average load factors for selected vehicle types in table 4: (V), (VI) and (VII). The CO_{2e} emissions are calculated based on emission factors of JEC (JEC, 2013, 2014).

3.1.3 Scenarios

The projection of EMS 1 and 2 is separated into 5 scenarios [short name in brackets]:

(1) baseline scenario 2040 (without EMS 1 and EMS 2) ['Baseline']
(2) implementation of EMS 1 without any restrictions 2040 ['EMS 1']
(3) implementation of EMS 1 and EMS 2 without any restrictions 2040 ['EMS 1+2']
(4) no EMS 1 and EMS 2 for 'heavy commodities': avoiding heavy cargo (e.g. bulk) will be shifted from rail to road ['EMS 1+2 + exclude commodities']
(5) consideration of external costs of transport e.g. study (Biehler, C., Sutter,D. 2019) from September 2019 ['EMS 1+2 + external costs']

Using these scenario projections, we are able to conclude the impact by comparison of the results of the different scenarios for EU 28.

3.1.4 Results

Corresponding to Figure 1, we observe the same increase of total transport tonne-kilometres from 2010 to 2040 in all scenarios and all modes will profit by increase of tonne-kilometres, that grows up from 2,556 billion tkm in 2010 to 3,801 billion tkm (49%) in 2040 for all modes.

The combined transport (CT) is disproportionately growing in the baseline scenario between 2010 and 2040 by 56% for inland water way (IWW) transport and for rail freight transport by 65%.

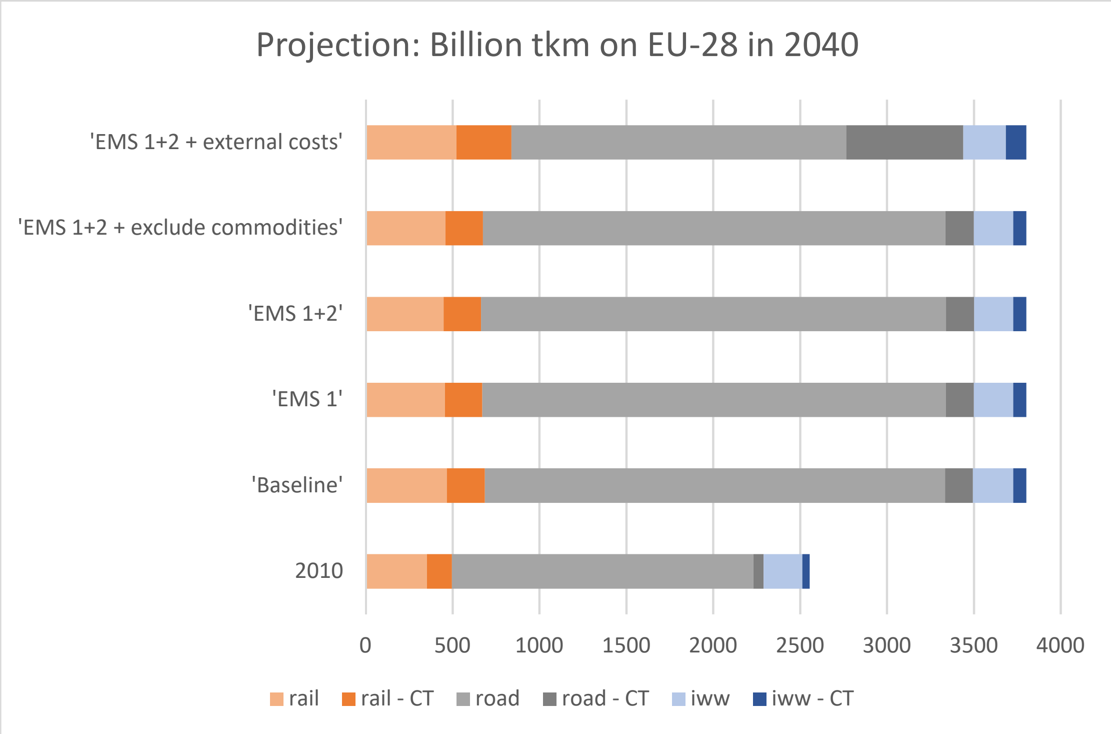


Figure 1: Projected transport performance for all scenarios

Related to the adjusted cost parameters, we see that the modal shift (in tkm) changes slightly:

- in in scenario ‘EMS 1’: There are an increase of 0.7 % on road and a reduction of 2 % on rail including CT and 1.7 % on IWW including CT.
- in scenario ‘EMS 1+2’: There are an increase of 1.1 % on road and a reduction of 3.2 % on rail including CT and 2.6 % on IWW including CT.
- in scenario ‘EMS 1+2 + exclude commodities’: There are an increase of 0.5 % on road and reduction of 1.6 % on rail including CT and 1.7 % on IWW including CT.

In scenario ‘EMS 1+2 +external cost’ the picture is different related to the other scenarios. There is a reduction of 7.4 % on road and rail including CT is growing by 22 % and IWW including CT by 18 %.

We could explain this modal shift by the higher efficiency of road freight transport in line with transport costs savings per tonne-kilometres (tkm) by using EMS. If this shift to road transport should be avoided, it is necessary to increase costs of road transport compensating the advantage by an increased efficiency due to use of EMS. Therefore, the external costs of all transport modes in scenario [‘EMS 1+2 + external costs’] is included and the figure shows a directly opposed impact by shifting transports to rail and IWW.

Figure 2 distinguishes the travelled road kilometres of the three heaviest vehicle types in all scenarios. The total travelled road kilometres grow up from 293.2 billion km (‘baseline’) to 298.5 billion km (‘EMS 1+2’). The implementation of external costs leads to 228.8 billion km, nearly 22 % less road kilometres as in baseline scenario. If there is assumed the exclusion of several commodities, we achieve the maximum: 301.7 billion kilometres. The strong increase of mileage in this scenario is determined by the shift of heavy commodities from EMS 1 and 2 back to the standard truck with up to 40 tonnes GCW.

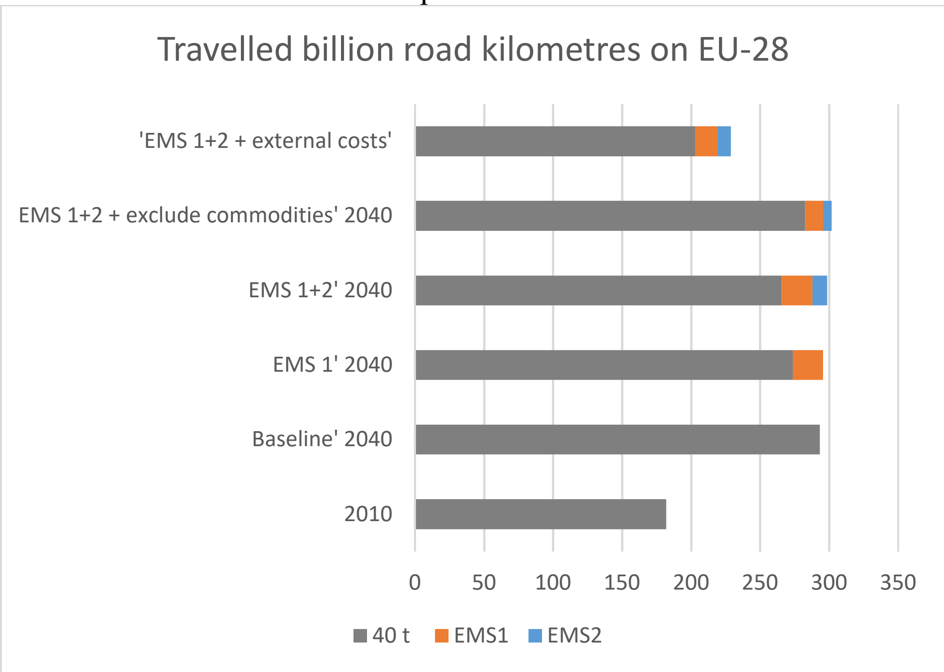


Figure 2: Travelled road kilometres of heavy trucks (40 t GCW , EMS 1, EMS 2) for all scenarios

The final step in the methodology is the calculation of CO_2e emissions of road freight transport in EU 28. The CO_2e emissions could be reduced by about 39 Mio. tonnes per year or about 18 % compared with the baseline in EU 28 (see figure 3) in the best case scenario 'EMS 1+2 + external costs'. This scenario does not assign the efficiency improvements of road freight transport to a reduction of transport costs (€/tkm) in comparison with the other modes.

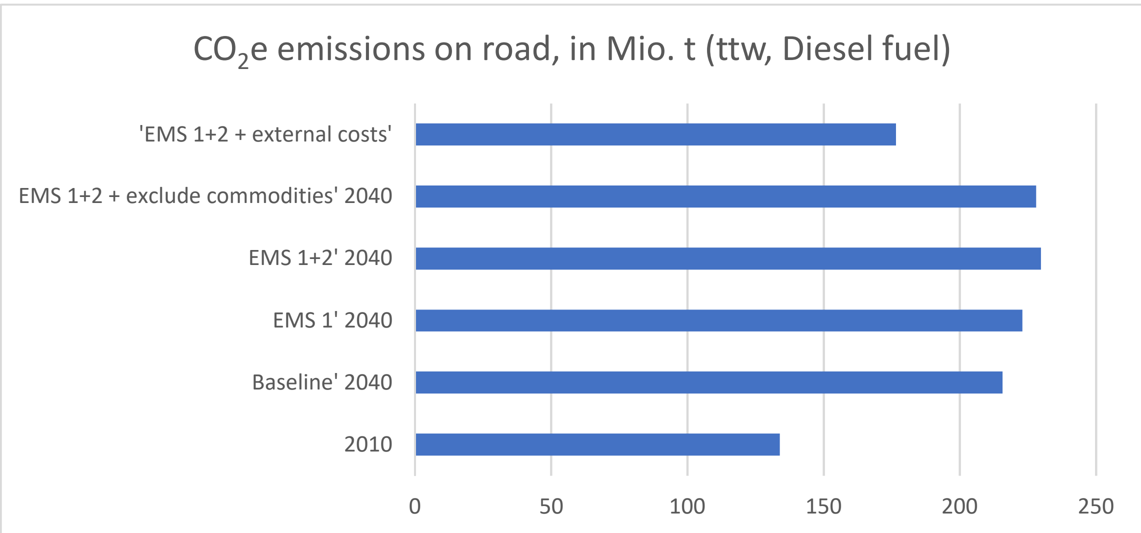


Figure 3: Impact on CO_2e emissions on road transport (ttw: tank-to-wheel)

In contrast, the modelling results of all other scenarios show that CO_2e emissions will increase due to mode shift from rail and inland water way to road. Based on these results, the policy regulation framework or the access policy for EMS should address on one hand side the realization of the possible improvements in road freight transport. On the other hand, the cost of road freight transport (in €/tkm) should be influenced by regulations in a way, that the cost advantages of the use of EMS would be compensated by measures to improve rail or inland water way (e.g. by including increased CO_2e emission costs in transport, internalisation of external cost, cost reduction on rail and IWW).

4 Application of AEROFLEX innovations in PI operations

The AEROFLEX project innovations can bring progress in the evolution towards the physical internet (PI) for three main areas of physical/digital/operational connectivity:

- encapsulation: standardized π containers: world-standard, smart, eco-friendly and modular
- flexible vehicles able to operate in diverse cycles, including (semi-)autonomously in logistic hubs using the electric drive train (of AEROFLEX trailer and/or dolly)
- high capacity transport for high volume major connections (e.g.: hub-to-hub transport with full truck load)

These aspects can be fitted in the physical internet roadmap. Each can be matched with innovative concepts developed in AEROFLEX:

- Advanced Energy Management Powertrain (AEMPT)
- Aerodynamic Features for the Complete Vehicle (AFCV)
- Smart loading units (SML)

Furthermore, the work regarding the regulatory framework should be of great help to facilitate the implementation of these concepts in homologation and standardisation processes that need to be set in motion if the concept is to find large scale adoption.

The rest of this section discusses some of these innovations in more detail.

4.1 Advanced Energy Management Powertrain (AEMPT)

The AEMPT is conceptually a distributed hybrid electric powertrain. In addition to its environmental savings potential (through a more optimal power management), the functional exponent of the AEMPT is an e-dolly. While a dolly is a vehicle component that is typically used to couple a rigid or a tractor and a semi-trailer, through electrification and built in communication equipment, the e-dolly can be operated remotely and without a towing vehicle. This allows the driver to split the whole EMS 1/2 in separate units for easy maneuvering and parking. The contributions of the AEMPT to the progress of the PI development are:

- Hybrid Electric, distributed powertrains can help the environmental performance (fuel consumption/climate change and local pollutants) of the vehicles in the first and last mile (maneuvering, high degree of start/stop driving).
- Physical Internet nodes are large or small logistics yards where autonomous maneuvering of loading units using the e-dolly can contribute greatly to the streamlined functioning of the yard.
- This also helps mitigate the issue of driver shortage and specialization. Drivers can focus on driving instead of loading and unloading, administration, etc. They can drop off their trailer at a gate and immediately pick up a new one to maximize their productivity.

4.2 Aerodynamic Features for the Complete Vehicle (AFCV)

The Physical Internet calls for high capacity vehicles (road, rail or ship, depending on the availability of infrastructure and the service requirements for the cargo) for the transport flows between the primary nodes of the network, in the most sustainable manner. In the case of road transport, this implies maximizing the energy efficiency of the largest vehicles travelling over motorways at high speeds in operational profiles that correspond to either “long haul” or “regional delivery” (as defined in the VECTO tool (European Commission DG CLIMA, 2021).

These cycles particularly lend themselves to the deployment of trucks that are aerodynamically optimized from front to back, and from top to bottom, so as to improve their fuel efficiency. The application is mainly in hub-to-hub transport, with high loads but essentially irrespective of distance. So long as there is an important part of driving under circumstances where the aerodynamic improvements developed in AEROFLEX achieve their maximal effectiveness (such as high speed driving on motorways), the deployment of AEROFLEX vehicles is useful.

4.3 Smart Loading Units (SML)

One of the most distinguishing features of the physical internet is the use of modular loading units that can be combined in an infinite amount of ways; from shoebox size to TEU container size. AEROFLEX works on ‘Smart Loading Units’ (SMLs), which cover the following features and functions:

- intelligent and safe,
- full access security,
- load optimization,
- fast interoperability,
- aerodynamic design,
- telematics-friendly,
- fit for intermodal.

Many of the design features of AEROFLEX SMLs translate seamlessly to the PI concept’s requirements.

In case, road transport is not the optimal choice, the standardized loading units studied in AEROFLEX are developed with the explicit objective to be suitable for intermodal transport. This is perfectly in line with the physical internet principle (and also with the synchromodality concept) to transport the cargo (or the loading unit to be exact) in the transport mode that maximizes efficiency while still meeting the customer’s requirements for delivery time. Another example of increased flexibility and load factor optimization called for by the PI concept is the use of double floor trailers and the CargoCam of project partner WABCO and the puzzle software of project partner Fraunhofer IML, a software tool based on the use of 3D sensors built into the trailer. Tests have shown, this can improve fill rate up to 38 %. These concepts are demonstrated in the project through practical use cases.

4.4 Other considerations

In addition to technological development, AEROFLEX work on the regulatory framework can be a stepping stone for further innovative legislative design to accommodate other PI-related advances.

5 Conclusion

AEROFLEX road transport innovations can take a role in the physical internet that is similar to broadband wireless connections in the digital internet: ultra-flexible, capable of moving high volumes at high speeds, with the best possible coverage at much greater efficiency than past technologies. While able to operate on its own, this new and improved characteristic of road freight transport is best supported by a strong wired network (rail, inland waterway and maritime transport) that is able to achieve even greater efficiency at higher volumes, between the main nodes, i.e. consolidation centers of the network. The process towards the uniform modularity that is required for all data/cargo transfers is advanced by the work on the smart, intermodal and fully modular loading units, which can be an inspiration for the Physical Internet containers and build upon initiatives of other EU projects such as MODULUSHCA and CLUSTERS 2.0.

Use cases show, that on a transport related level transport costs (per €/tkm) and CO_2e emissions per ton-kilometres could significantly be reduced. Macroscopic freight modelling compares different scenarios and shows that a positive impact on whole EU freight transport need an intelligent access policy to scaling up the existing benefits of use cases to the EU road transport level. More detailed information about AEROFLEX innovations and findings are available at the website www.AEROFLEX-project.eu.

References

- AEROFLEX (2018): Transport market and its drivers with respect to new vehicle concepts, Deliverable D1.1, https://aeroflex-project.eu/download/project_results/AEROFLEX-D1.1-Transport-Market-And-Its-Drivers-PU-FINAL-2018.04.30-PUBSUM.pdf.
- Biehler, C., Sutter,D. (2019): Externe Kosten des Verkehrs in Deutschland Straßen-, Schienen-, Luft- und Binnenschiffverkehr 2017, Schlussbericht, Zürich, 21. August 2019.
- Burgschweiger, S., Wolfermann, A., Liedtke, G. (2017): A macroscopic freight transport demand model to analyse effects of transport planning in Germany. In: Proceedings of the 3rd Interdisciplinary Conference on Production, Logistics and Traffic 2017, Darmstadt, Germany.
- EUREF (2016) projection: https://data.europa.eu/euodp/en/data/dataset/energy-modelling, (status: 2020).
- Directive (EU) 2015/719 of the European Parliament and of the Council (2015): amending Council Directive 96/53/EC laying down for certain road vehicles circulating within the Community the maximum authorised dimensions in national and international traffic and the maximum authorised weights in international traffic; Brussels 2015, https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX:32015L0719.
- European Commission DG CLIMA, (2021): Vehicle Energy Consumption calculation Tool – VECTO, https://ec.europa.eu/clima/policies/transport/vehicles/vecto_en (consulted on 21/04/2021).
- JEC (2013): Joint Research Center of the European Commission, Eucar and Concawe (JEC), Tank-to-Wheels Report. Version 4.0, Luxemburg 2013.
- JEC (2014): Joint Research Center of the European Commission, Eucar and Concawe (JEC), Well-to-Wheel Report. Version 4.a, Luxemburg 2014.
- NST (2007): Eurostat, satistics explained, Standard goods classification for transport statistics (NST) https://ec.europa.eu/eurostat/statistics-explained/index.php/Glossary:Standard_goods_classification_for_transport_statistics_(NST).

Smart IoT solutions for ports paving the way towards PI nodes

Achim Klukas[1], Maximiliane Lorenz[1], Franz-Josef-Stewing[2] and Christophe Joubert[3]

1. Fraunhofer Institute for Material Flow and Logistics, Dortmund, Germany
2. Materna Information & Communications SE, Dortmund, Germany
3. Prodevelop SL, Valencia, Spain

Corresponding author: achim.klukas@iml.fraunhofer.de

Keywords: *Emissions reduction, Internet of Things, Intermodal and Synchromodal Transport, Logistics Nodes, Multimodality, Port of the future*

Conference Topic: *From Logistics Nodes to PI Nodes, Ports and hubs in the PI, Interconnected freight transport, logistics and supply networks*

Abstract

Ports play an important role in global supply chains. The use of IoT is considered a key feature for port operators to improve the efficiency of port operations, to be able to better manage (container goods) traffic, to empower their workforces increasing throughput and to decrease carbon emissions while making traffic safer. In addition, smart IoT solutions support automation and intelligent transport control to realize the physical Internet.

The R&D project I²PANEMA helps ports to become (a network of) smart ports by exploring and demonstrating the applicability of IoT technologies. With this approach, I²PANEMA aims to make ports more efficient and sustainable by introducing IoT based measures such as Active Noise Control systems and keeping pollution (like noise, dust) under control while promoting multimodal transport as part of the physical internet. Case studies from the project show the possibilities for developing ports in Physical Internet Nodes.

1 Introduction

Following the globalization trend, inland and seaports become important logistics nodes for the economic sector. Nevertheless, the reaction to future growth and increasing requirements for digitization often leads to problems for the ports. Furthermore, as inland ports are often located in densely populated areas, conflicts with the surrounding environment arise. One solution for many of the problems that ports are currently facing is the implementation of IoT based measures such as sensors, platforms, and IT-Systems.

The European research project I²PANEMA addresses the above-mentioned problems and explores solutions in the field of the Internet of Things. Digitization is the key factor for ports to maintain their competitiveness and expand their importance for the entire economy. In addition, digitization measures allow processes to be more efficient, ports to better manage their traffic and implement sustainable strategies to reduce carbon as well as noise emissions. Hence, I²PANEMA addresses the current problems of ports and aims to develop solutions for them.

IoT based optimization of logistics processes builds the central idea for the digitalization of ports. Therefore, most processes in the port need to be digitized to improve the communication, standardization, and optimization of the whole port. Especially the introduction of standards for data transmissions, help the ports to optimize the complete process chain from transport to handling and warehousing. Hence, IoT solutions provide the foundations for the so-called Smart Ports concept. Smart Ports are defined by their use of automation and innovative technologies (Port Technology International Team, 2021). The *ITEA*3 project I^{2}PANEMA aims to implement IoT based solutions in ports and will be further described in this paper.

2 Project description and Physical Internet

The *ITEA*3 project I²PANEMA aims to implement IoT measures and the physical internet in ports by connecting processes, increasing the efficiency and focus on sustainability measures. To help the partner ports within the project to become smart ports, nine business scenarios are implemented. Common barriers like data security agony and IT systems are also kept in mind within the projects scope. Hence, I²PANEMA develops a reference architecture for ports which applies software components directly for the port environment. The project enables overall supply chain resilience and supports the growth of ports in Europe as important logistics nodes.

I²PANEMA focusses on developing different types of technological innovations. The reference architecture designed within the project, is specified for ports. It therefore includes a security and privacy layer by design and a flow based IoT interoperability layer. The reference architecture supports the implementation of IoT based implementations in the ports and reduces innovation barriers to help the ports to maintain their status as important logistic nodes (Stewing et al, 2019). The reference architecture is transferred to the business scenarios within the projects. Accordingly, I²PANEMA contributes to standardization bodies such as Industry 4.0 RAMI.

Physical Internet is a concept for an optimized, standardized global goods transport system based on the idea of the digital Internet. In contrast to today's approach, in which a single transport service provider transports goods over long distances, the idea of the Physical Internet relies on fragmented, provider-independent transports (logistikknowhow.com, 2021). The roadmap to the physical internet described the path from ports to nodes in the PI. In PI port the operations are standardized and the usage of a family of standard and interoperable modular load units from maritime containers to smaller boxes is extensive. Services in PI nodes are visible and digitally accessible and usable including planning, booking and execution operations. Also, at the end of the development a connection between networks are available the PI nodes are fully autonomous (ALICE-ETP, 2021).

I²PANEMA includes the main findings and recommendations of the roadmap to the physical internet provided by ALICE. The I²PANEMA business scenarios aim to optimize and standardize port processes and thereby connect the stakeholders along the ports supply chain.

3 Background

3.1 Supply Chain

Global supply chains often use ports as means for handling operations, leading to them being essential players within the supply chain. The supply chain process for ports, inland as well as seaports, goes from the delivery all the way to the payment of a product or service. Increases in global shipping transport lead to enlarged requirements for the port companies and associations (Geerlings et al, 2018). This trend can also be seen for inland waterway transports. The distribution networks within the countries are expanded leading to higher transport volumes on the inland waterways, resulting in augmenting requirements regarding efficiency and the performance of the ports (Stewing et al, 2019). These can be achieved by introducing organizational and IT-supported solutions and connecting stakeholders.

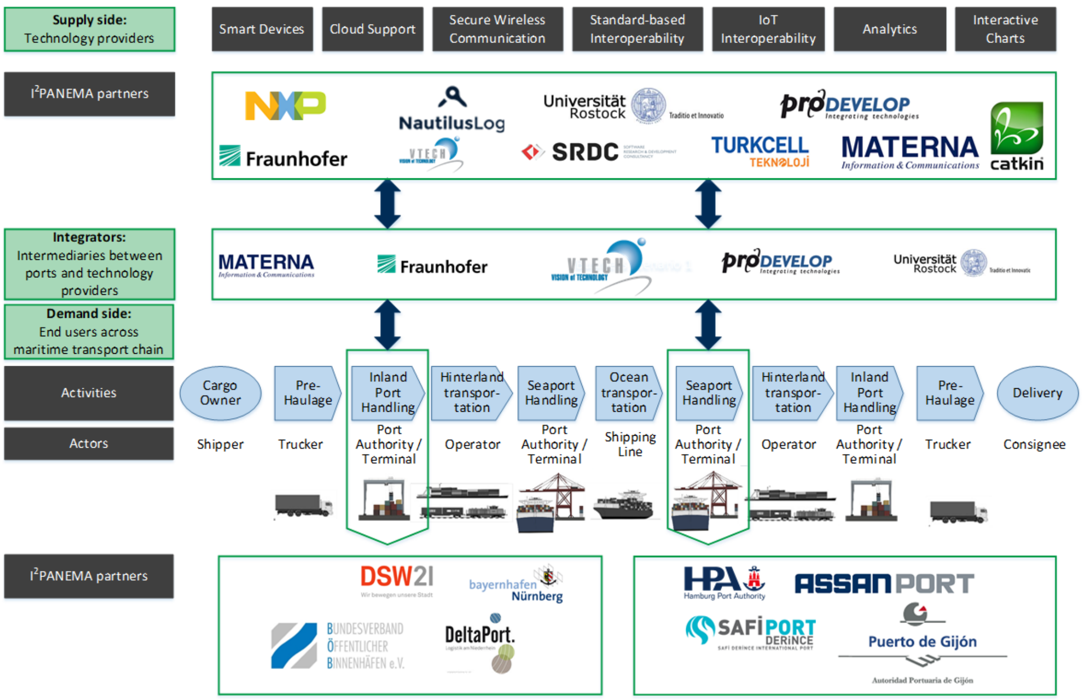


Figure 1: I²PANEMA business value chain (I²PANEMA project consortium, 2019)

The I²PANEMA consortium combines ports, technology providers and research organizations to accomplish the challenges mentioned above. Figure 1 shows the project structure of I²PANEMA. With the integration of the different partners, a network, which can share experiences and yields different perspectives, is created. I²PANEMA connects port stakeholders with IT companies and develops IoT based solutions for the problems the ports are facing. The project focusses on the port handlings, whilst keeping preliminary and follow-op processes in mind.

3.2 IDS and Gaia-X within I²PANEMA

The International Data Space is an initiative in which secure data spaces are created. These data spaces enable companies, independent from their size or operation sector. Digitalization is at the same time driver and enabler of innovative business models. In this context, data as an asset is becoming increasingly important (ten Hompel et al, 2018). For the successful implementation of smart services, innovative service offers and automated business processes, a regulated handling of data and a secure cross-company data exchange are necessary. I²PANEMA provides a data driven IoT solution to improve the performance of inland- and seaports in relation to the challenges of globalization. With respect to the heavy use of data and the concepts of Internet of Things (IoT), aspects such as data sovereignty, data space and economic platforms become more and more important. The target must be to build an agile, flexible and competitive infrastructure for the industry and port sector by exploring central key technologies.

Another, newer initiative is GAIA-X. GAIA-X is a European initiative which develops an open, transparent and digital ecosystem. The aim of GAIA-X is to enable secure, open and sovereign use of data by implementing a unique data ecosystem (BMWIa, 2020). Developers from many different sectors like economy, science and politics participate in it in order to develop a secure and connected data infrastructure which fulfils the highest standards of data sovereignty and data innovation. GAIA-X enables the user to make self-determined decision on data storage, processing and usage within its infrastructure. Portability, interoperability, and interconnectivity enable a secure data usage. (BMWIb, 2020)

3.3 ISO 4891 Smart Logbook

Open standards to speak a common language are the key to safer and more effective. Several problems occur now and hinder the automation of information flows. To mention are outdated and costly workflows (Paper, Excel, Mail, …), new regulations (e.g. environmental reporting) requires more data & exchange, and missing standards for digital data collection and exchange. From this follows that technical adaptions for new market needs are too slow in development (Hardware vs. Software). Also, the demand of new applications/collaboration between ship and shore and integration of new technologies (e. g. Sensors/IoT) takes place slowly.

For this reason, the topic of a Smart Logbook was brought to an ISO working group to develop ISO 4891 - Smart Logbook (ISO, 2021). The general goal of the ISO norm is to improve safety on board and shore through increased data availability and create a standard that helps to generate logbooks for smart devices. This should also allow new applications to fulfill new environmental requirements. A smart logging device as an add on to electronic record book (ISO 21745) for legal data. Also, support a smart logging device for non-legal data and faster development and reactions to new regulations and documentation requirements. The standard can support different solutions via structured parts and integrate more partners and applications.

This standard will allow flexible and easy integration of IOT and third-party systems to monitor and to fulfill contractual charter party requirements. Display and notify about vessel information, dangerous situations, and overall status of the systems on a smartphone in all areas of the vessel and at shore will be possible. Furthermore, the synchronization of ETA with port activities vessel calls rescheduling between crew and ports are possible. Additionally, monitoring of vessel condition with the integration of flag state procedures and integrated checklists could be done. Based on the standard legally required documentation from connected collected data (e.g. MRV reports) could be generated automatically.

4 I²PANEMA reference architecture

I²PANEMAs aim is to enable and advance the use of IoT solutions that lead to Smart Ports under consideration of the growing demands regarding efficiency, sustainability, and safety aspects. The project specifically intends to develop a reference architecture for ports and tries to implement it in the business scenarios within the project's duration. The I²PANEMA reference architecture is built with a functional block including different parts such as the data management layer, the interoperability layer, the M2M communication layer and the security and privacy layer. For the design of the different layers, an extensive research on the current state of the art has been conducted.

4.1 State of the art

During the project, a detailed state of the art report was prepared (Clemente et. al., 2020). It contains an analysis of the current state of the art and alternative technologies for the architecture proposed in the I²PANEMA project. The report highlights the main reference architectures for the industrial Internet of Things for Europe and globally, like RAMI4.0 and IIRA. Based on various studies, a deep analysis was made for each of the identified functional block consisting of IoT communication, IoT interoperability, IoT data management, IoT operational workflow, IoT visualization and IoT security & privacy.

For the data management layer, the analysis focused on data characteristics, namely data volume, data traffic, data criticality, and data variety, and technological solutions giving support to these characteristics (e.g. data broker Kafka, persistence with HDFS (Hadoop File System), MongoDB, Apache Cassandra). In terms of operational workflow layer, an overview was made of existing workflow management systems (e.g., Mendix, Nintex) with a particular emphasis on flow-based programming solutions. For the IoT visualization layer, several data visualization techniques (e.g., charts, tplots) and tools (e.g., Tableau, Plotly) were reviewed. Finally, for IoT security & privacy layer, a STRIDE Threat Analysis was recommended, as well as SotA security solutions (e.g., access control, asset identification, root of trust, secure execution platform). The STRIDE analysis was invented by Microsoft and contains exploration of threats on spoofing, tampering, repudiation, information disclosure, denial of service and elevation of privilege. Therefor all system components were independently evaluated, and potential security vulnerabilities were uncovered by giving each component a DREAD rate on for example the impact or the probability of an attack. This rating formed the basis for further decisions on how to handle the risk, e.g. develop countermeasures like remove the feature that causes the risk, pass the risk to an externality or migrate the risk, or for low risk rates accept the risk and its possible effects. In a context where the information is a major asset and the systems – especially IT based ones – change and evolve rapidly, it is important to find a proper balance between detailed technical specifications and enough flexibility for future changes, adapting to new conditions in the markets, the emergence of technologies or variations in the approach of the ports that could be faced in the near future. The results of the state-of-the-art report have also been incorporated into *ITEA*'s Living Roadmap (*ITEA*3, 2021), where they can be read in detail. Based on the findings from the analysis of the state of the art and the requirements from the business scenarios, the following reference architecture was developed.

4.2 Reference architecture

The goal of the I²PANEMA reference architecture is to provide an architecture, that is customized for ports. This includes data security aspects and interoperability. Furthermore, the reference architecture needs to be focused on industrial needs, as ports are in industrial environments and need to note their requirements. Figure 2 shows the layers of the I²PANEMA reference architecture. The “M2M Communication Layer” (MCL) includes the devices and the communication between the devices with the actual sensors and actors and their management. The central layer, the “IoT Interoperability Layer” (IIL), accesses the measurements of device by adapting and verifying the measurements and transforming them into data and events. Accordingly, it provides an abstraction of technologies from different vendors by converting measurements for further processing and translating actor intents into the MCL layer. The ILL transfers data and events to the uppermost layer, the “Data Management Layer” (DML). The DML stores, visualizes, and manages all given information by consolidating the actors’ representation and sensor measurement. Machine Learning components, workflow management and data storage are also part of the DML.

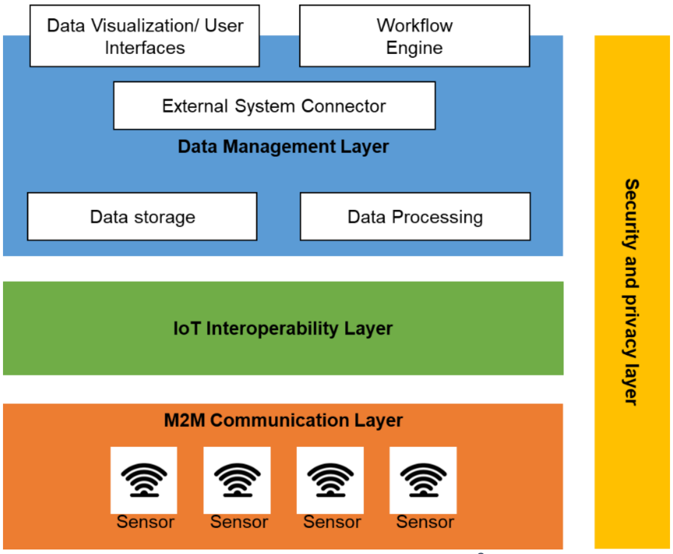


Figure 2: A schematic representation of the I²PANEMA reference architecture

5 Description of business scenarios

Within the I²PANEMA project, nine business scenarios in inland as well as seaports are implemented. The business scenarios aim to strengthen the ports positions in the global supply chain by introducing digitized and connected processes. The scope of the business scenarios reaches from noise reduction, to sustainability and process improvement measures. In the following paragraphs a short overview over six business scenarios in I²PANEMA and their technical innovations are given.

5.1 Intelligent parking management – Port of Dortmund

The shift of road freight transport to rail and waterways is being implemented via Combined Transport Terminals (CTT). By designing the infrastructure at these traffic junctions, traffic is concentrated and can lead to traffic jams and insufficient parking opportunities, especially at peak times. This problem arises at the port of Dortmund. The port is located right in the inner city of Dortmund, which leads to trucks having to drive over and cross highly frequented roads.

Within the Business Scenario in Dortmund, the goal was therefore set to reduce road congestion to achieve greater acceptance among other road users and residents. In addition, making the handling of goods more efficient to minimize time delays will contribute to an optimized utilization of the CTT.

For this purpose, the truck is tracked after it has been registered in an application. Based on IoT devices and camera recognition, the traffic, parking lot and factory gate situation is evaluated. The trucks are then called to the factory gate just in time.

5.2 Smart Port Ship – Cross Country

The Smart Port Ship business scenario is implemented in all the participating countries. The business scenario addresses the issue, that registration processes of ships in ports are not standardized yet. More precisely this means, that the data exchange for incoming ships varies from port to port and several different data packages. The European Maritime Single Window (e.g. National Single Window and the National Single Window Plus) already provide interfaces that enable ships to report their data to the port authorities. Nevertheless, the communication and data exchange is not automated and standardized. To share the data to all relevant actors, manual transmissions take place, leading to errors and delays.

The Smart Port Ship business scenario starts at this point and aims to optimize the registration of ships at the ports via a standardized interface. A digital logbook is installed on board the ship for this purpose. With the help of the digital logbook, data is collected and tracked. The logbook enables the connection of different stakeholders (i.e. manager, charterer, port authority, crew) and shares the relevant data with them. For the secure transmission of data, the business scenario will examine whether the International Data Spaces can be used as an interface, including control layers for the relevant stakeholders.

5.3 Environmental indicators – Port of Gijón

The port of Gijón is in Spain at the Cantabrian coast. The Port Authority of Gijón has implemented an environmental surveillance plan in to control the environmental quality of the surrounding environment. The surveillance of PM_{10} (Particulate Matter less than 10µm) stations is part of the I²PANEMA project. Within the project, an IoT platform working with real live data on the pollution situation is developed. The particle pollutions are interpreted and alarms and alerts according to emergency action protocols enter into force whenever the pollution created by the port processes exceeds given maximum values. This allows the port to control and reduce the environmental impact created by port operations and traffic around the port. Figure 3 shows the web application that monitors the pollution and provides alerts.

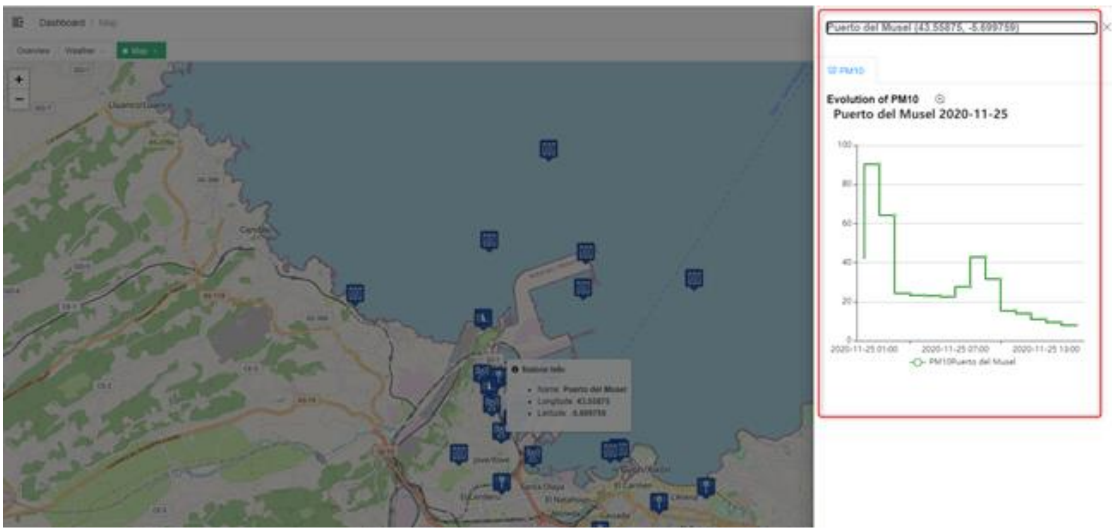


Figure 3: Evolution of PM10 levels over the last day from the I²PANEMA map view (own representation)

The application includes the PM_{10} emissions measurement, information on pollution peaks, current weather information, connections between berths and pollution levels and PM_{10} level forecasts.

5.4 RoRo localization – Assan / Safi Ports

The aim of the RoRo (Roll-on, Roll-off) business scenario, conducted in Turkey, is to increase the efficiency of RoRo operations in Assan / Safi Ports by using IoT devices. With the use of location sensors, an overview over the location of the drivers who are transporting the cars to their loading / unloading places within the port is given. The accuracy of the device is up to 50cm leading to exact position information for each vehicle at the RoRo terminal. The I²PANEMA partners use a NarrowBand IoT network for the location transmission. The use of the IoT sensors is expected to lead to a decrease in turnaround times by minimum 10%.

The IoT system developed within I²PANEMA will be implemented into the port management system of VTEK, an industrial partner within the project. The integration of a visual component in form of a dashboard enables port operators to see KPIs such as the number of cars in each parking lot, the occupancy rate of the parking lots and the average loading / unloading times per car.

5.5 Smart Ferry – Port of Hamburg

The business scenario in the port of Hamburg focusses on the ferry services for the local transport network. The port of Hamburg operates a fleet of 25 ferries. The aim of the business scenario is to implement a prototypical holistic IoT solution for the ferry fleet in Hamburg, as information on delays or the estimated time of arrival is currently not provided for the ferry transports. By forecasting passenger volumes, calculating ETAs (estimated time of arrival) and monitoring the waiting times of the passengers, the transparency shall be enhanced, and the processes shall be optimized. This should lead to an increased customer comfort and an optimized passenger flow.

For the technical implementation, the project team will install a local network on each ferry. These networks will provide real-time GPS data, automatic passenger counting and travel times.

5.6 Active noise control – TriCon Terminal Nuremberg

The TriCon Terminal Nuremberg is a combined transport terminal in Germany. The terminal is in an urban area, which results in handling restrictions, especially during night times, due to noise emissions. The location of the terminal leads to conflicts with the surrounding population as the handling processes such as the mounting and dismounting of containers and empty runs produce noise emissions.

Within the I²PANEMA project, a novel Active Noise System (ANS) will be applied at the CT terminal to reduce the noise emitted by the processes mentioned above. A test environment to test the feasibility of the ANS system is therefore developed within the project. The ANS systems aims to generate longer operation times for the terminal, especially during night times. The ANS works with microphones and speakers that analyze the emitted noise and produce counter-sound waves. These counter-sound waves will reduce the audible noise emissions directly at the terminal. Figure 4 shows the principle of the ANS system in a visualized way.

Figure 4: ANS system for the CT terminal

5.7 Conclusion on business scenarios

With the implementation of IoT measures, I²PANEMA helps ports to digitize their processes and transform from logistics nodes to PI nodes. The project uses hardware as well as software applications in order to connect data and stakeholders. I²PANEMA lays a special focus on security aspects regarding data exchange and data collection. Therefore, the potential of the Industrial Data Spaces and GAIA-X are tested within the business scenarios. This helps ports to transform to Smart Ports. With the development of a reference architecture, the project helps other ports to realize IoT projects. Furthermore, a standardization for data exchange in ports is developed within the project.

In summary, I²PANEMAs business scenarios support the development of the Physical Internet. The work on ISO supports the standardization of information processes, and standard services are defined, thus increasing overall efficiency. The reference architecture enables the uniform design of IoT services in ports. This is based on the RAMI 4.0 standard and thus enables the simple connection of services in ports to IT and logistics.

The business scenarios also create additional services in ports, which on the one hand improve efficiency, but also improve the flow of information, lay foundations for the automation of ports in various areas such as the maintenance of quay walls and the improvement of logistical processes such as RoRo loading, ANC and traffic management in port. In particular, the overarching SmartPortShip scenario aims at automated and standardized information flow.

6 Conclusion

IoT devices allow to digitalize processes in inland- and seaports by using new IT strategies. This leads to increased efficiency of processes in ports, but also allows the transformation to sustainable, green ports with low emissions and a cooperative coexistence with the population. Based on the Industrial Data Spaces and GAIA-X, secure exchange and linking of data in business ecosystems based on standards and collaborative governance models are enabled. Using IoT, GAIA-X and IDS ports can transform into interactive, digitized hubs, so called smart ports. To show the potentials, the project I²PANEMA has the goal to implement standards with IT-based solutions and the use of Internet of Things in ports, which will serve as a reference architecture to demonstrate the variety of possibilities opening up. Through the practical application of new technologies in broad business cases, the transferability of various IoT concepts to the port environment is tested. In the future, the development of the Port Data Space will help port-related supply chain actors to implement IoT applications, intelligent AI-powered data processing and data networking.

References

- ALICE-ETP (2020): Roadmap to the Physical Internet,
- BMWIa (2020). GAIA-X: A federated data infrastructure for Europe. https://www.data-infrastructure.eu/GAIAX/Navigation/EN/Home/home.html
- BMWIb (2020). GAIA-X: A Pitch Towards Europe – Status Report on User Ecosystems and Requirements.
- Clemente J. et al. (2020): I²PANEMA State of the Art. https://itea3.org/project/workpackage/document/download/6432/Deliverable%20D1.5%20-%20State-of-the-Art%20Overview%2FAnalysis.pdf.
- Geerlings, H., Kuipers, B., Zuidwijk, R. (2018): Ports and networks – Strategies, operations and perspectives. London, [England], New York, [New York]: Routledge. Chapter 2.
- I²PANEMA project consortium (2020)
- ISO (2021): ISO/WD 4891. Ships and marine technology — Navigation and ship operations — Smart logbooks for shipping.
- *ITEA*3 (2021): Living Roadmap challenges. https://itea3.org/living-roadmap-challenges.html. Last access: 12.02.2021.
- Port Technology International Team (2021) : What is a Smart Port ?, https://www.porttechnology.org/news/what-is-a-smart-port/, 2021/04/14.
- Stewing, FJ., Klukas, A., Lorenz, M. (2019): Full Project Proposal I²PANEMA intelligent, IoT-based Port Artefacts Communication, Administration and Maintenance, November 15, 2019.
- ten Hompel M., Otto B., Wrobel S. (2018): Industrial Data Space. In: Neugebauer R. (eds) Digitalisierung. Springer Vieweg, Berlin, Heidelberg. https://doi.org/10.1007/978-3-662-55890-4_8.
- https://logistikknowhow.com/materialfluss-und-transport/physical-internet/

GPICS a framework to create a digital twin for the Physical Internet

David Cipres, Alberto Capella, Lorena Polo, Jose L. Lopez Ramon[1]
1. ITAINNOVA (Instituto Tecnologico de Aragon), Zaragoza, Spain
Corresponding author: dcipres@itainnova.es

Abstract: *The GPICS (Generic Physical Internet Case Study) framework has been developed within the ICONET (New ICT infrastructure and reference architecture to support Operations in future PI Logistics NETworks) project. It is a methodology for designing and evaluating supply networks through PI. GPICS is a conceptual framework that is generated as an abstraction of different Living Lab cases and generic industrial scenarios. It is formed based on six interrelated dimensions. GPICS modelling is designed to allow the composition of a generic PI network through reusable modelling components and rules, via appropriate configuration, these components can represent different types of supply chain flows.*

The six interrelated dimensions range from the necessary modelling components and base configuration rules (Modelling Kit) to the scenario definition/parameterization capabilities (based on operational rules, business models and vertical and horizontal collaboration strategies between the different roles in the supply chain), including Master Data Sets, that concern and are relevant for a Geographical Area within the EU. As mentioned above, the GPICS also includes a set of benchmark KPIs for the assessment of different PI scenarios, based on a different combination of the configuration capabilities of these scenarios.

The network is represented by using a virtual simulation model. A multi-agent simulation model is used to create a digital twin of the supply network, utilizing GPICS main components and integrated with ICONET's PI services. The virtual model contains a general representation of the main nodes and their interconnections. It creates a representation of the main flows of freights in a PI network. It can include transports from different companies, with different restrictions. The simulation model developed is a tool that helps companies to visualize how the movement of products over a PI network can be, including flows from other companies.

The virtual models are used to quantify the impact of the different services. Economic (transport and handling costs), operational (reducing lead time) and environmental (CO2 emissions) indicators have been obtained in different living labs. The description of two use cases developed in the project is included.

Conference Topic(s): *PI Modelling and Simulation, PI Fundamentals and Constituents.*

Keywords: *Digital Twins, Multi-agent simulation, Simulation, Showcasing, Intermodal and Synchromodal Transport*

1 Introduction

This article describes a method for modelling and analyzing different Physical Internet (PI) configurations in different use cases. A part of the content has been inspired by the implementation of the ICONET project [1] where the main parameters of GPICS (Generic PI Case Study) have been defined. The modelling components of the framework are based on the

different living labs under a common PI framework. GPICS identifies the components, parameters, rules and indicators to create a PI network within the context of the Living Labs. The design of the GPICS framework serves as the basis for simulation-based analysis of the performance of PI networks, using defined and agreed KPIs. The ICONET project also has direct connections with the SENSE project [2], but with different goals. SENSE strategic goal is to accelerate the path towards the Physical Internet (PI), so advanced pilot implementations of the PI concept will be well functioning and extended in industry practice by 2030, and hence contributing to at least 30 % reduction in congestion, emissions and energy consumption.
The design of this framework begins with a review of the existing literature on the Physical Internet. The PI literature is constantly growing. The very first publication on the PI dates from 2006 while the concept of actual PI was initially introduced in 2010 by Montreuil et al. in [3], who laid its foundations and received the attention of academics and practitioners. Most PI publications are conceptual and try to supply practical solutions for certain PI components. Similarly, there are many studies and simulations aimed at providing real-life solutions for some of the PI components (e.g., simulations for the operations of PI-hub, PI-store and PI-sorter) but there are few case studies or experiments focused on the analysis of the potential impact and benefits at the level of the PI network and the integration of the models with the specific PI services.
A Digital Twin (DT) is a virtual environment that mirrors the real physical system (a physical twin) and its processes by updating its virtual real-time status from various sources of information about weather forecasts, congestion levels, positions of assets (barges, trains, trucks) and their ongoing working conditions. There are different definitions of the Digital Twin concept. Zheng et al. [4] define the DT as "an integrated system that can simulate, monitor calculate, regulate, and control the system status and process". The Digital Twin could simulate multiple virtual scenarios at once. Launching virtual scenarios allows key indicators to be measured in a virtual environment, unlike pilot projects which can be expensive or risky to implement.
In the ICONET project, a methodology has been developed to combine physical and digital elements through simulation models that can be integrated with PI services that provide real-time information on the position of containers and transports. Behind the idea of the mixed Digital/Physical simulation approach is the technical perspective, that the developed models should allow the inclusion of information from both digital elements as well as and physical sources. In certain simulation models, the information of some components (i.e. the position of a truck, or the storage capacity of one facility) could be obtained from a real physical entity, through APIs available in the cloud. The work on the development of modeling and simulation around Physical Internet is currently evolving, among other projects within the PLANET project [5].

2 GPICS ELEMENTS

The GPICS framework consists of six dimensions that are interrelated, in fact, these six dimensions that make up the GPICS are more than interrelated, they are interdependent, in the sense that each of them is the input to the next. The GPICS framework provides not only the components needed for a case study definition but also a process or cycle to drive it.

As shown in (Fig. 1) GPICS (Generic Physical Internet Case Study) is defined on the basis of six interrelated dimensions covering from the necessary modelling components and the rules of base configuration (Modelling Kit) up to the capabilities of scenarios' definition/parameterization (based on operational rules, business models and vertical and horizontal collaboration strategies among different roles in the supply chain) including Master Datasets, which concern and are relevant to a Geographic Area within the EU, which will allow the instantiation of the GPICS and the creation of the PI (Physical Internet) Hubs Plan. As mentioned above, the GPICS also includes a set of key performance benchmarks Baseline Key Performance Indicators for the evaluation of different PI scenarios, based in different combination of the configuration capabilities of those scenarios.

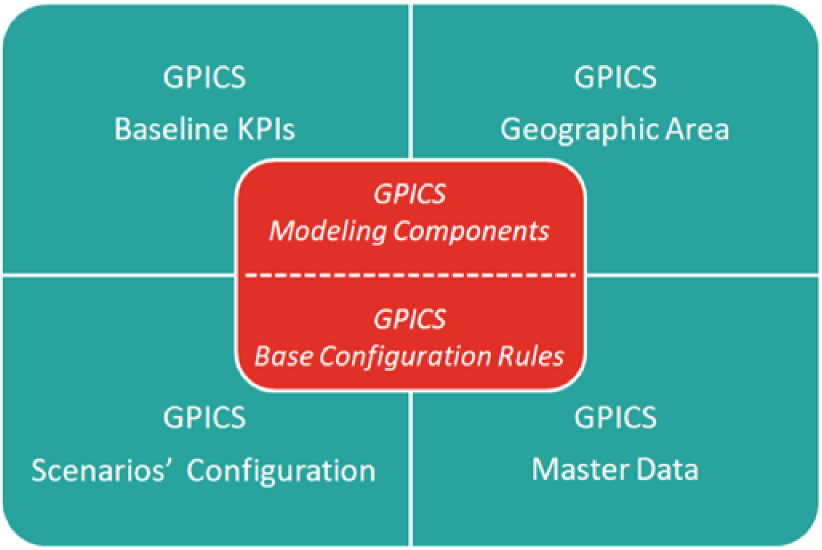


Figure 1: GPICS Framework/Dimensions

The modelling components and base configuration rules in the Modelling Kit meet the PI challenges posed by use cases as an abstraction level allow the integration of the four Key PI capabilities which correspond to each of them.

The GPICS framework is also the basis for simulation models. GPICS modelling elements and base configuration rules, which are included in the dimension "Modelling Kit", have a direct correspondence with simulation models. On the one hand, the modelling elements, such as hubs/nodes or corridors, have their representation in the simulation as objects, the so-called 'Atoms', and on the other hand these 'Atoms" have a behavior based on the basic configuration rules defined in the GPICS framework and instantiated in the GPICS definition.

The “Geographic Area” is the first dimension of the GPICS definition. The geographical area defines the regions covered by the case study and is the main GPICS parameter. The second dimension is the “Set of Master Data” associated with the geographic area selected in the previous dimension. This master data characterizes the current supply chains in the geographical area in terms of specific ports, multimodal hubs, TEN-T corridors, urban distribution centers, population coverage, cargo/freight load distribution, transport demand/flow, warehousing capacity, transport modes and frequencies, lead times, etc.
The core of the GPICS is the Modelling Kit, which consists of two dimensions. On the one hand, it includes the “Modelling Components” and, on the other hand, the “Base Configuration Rules”. The modelling components are a set of elements that represent physical elements in a PI network, such as: PI hubs/nodes, PI corridors, PI containers, etc., as well as a set of roles which interact and have an active participation in a supply chain in PI. Amongst these roles, we can highlight: PI sender, PI receiver, PI transport & logistics service provider or PI network

coordinator. The fifth dimension which is part of GPICS, is the “Scenarios' Configuration Capabilities”, based on What-if scenario analysis (WISA). WISA is a business planning and modelling technique used to yield various projections for some outcome based on selectively changing inputs parameters. A scenario, in this context, is a potential circumstance (i.e., parameter change) or combination of circumstances (i.e., combination of different parameters changes) that could have a significant impact -- either positive or negative -- in an organization. The last dimension of GPICs consists of a set of “Generic Key Performance Indicators”, which will allow a standard and common evaluation of the performance of the PI supply chains configured in the GPICS, between different scenarios. The GPICs Key Performance Indicators have the mission to provide a comprehensive vision of the impact of PI with respect to the current situation and to be an instrument capable of shedding light on the strengths and weaknesses of different PI scenarios.

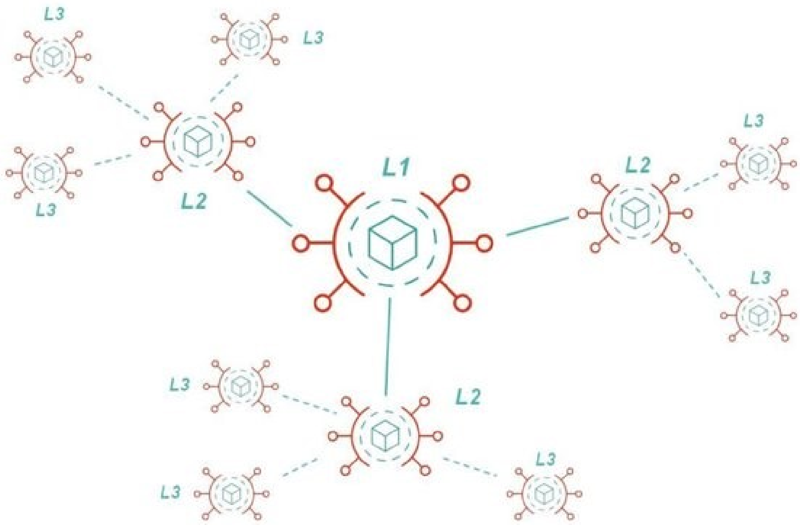


Figure 2: GPICS three-level structure of HUBS

3 SIMULATION COMPONENTS

The aim of Digital Twin model is to produce network simulation models in support of PI concepts prototyping efforts connected with the PI services designed. Simulations are used to estimate the network-wide performance associated with different routing policies and dynamic routing decisions, to find best networking for collaborating hubs, and to evaluate the efficiency of the PI-enabled services.

Regarding the design of the elements necessary to stand for a PI transport scenario, the elements defined by Montreuil, Meller and Ballot [3] have been used as a basis. They proposed three key types of physical elements as enablers of Physical Internet: the PI containers, the PI nodes and the PI movers. PI containers are described as the unit loads that are manipulated, stored, moved and routed through the systems and infrastructures of the Physical Internet. PI containers are generically moved around by PI movers. Moving in this context is used as a generic equivalent to different logistics and transport activities or processes such as transporting, conveying, handling, lifting and manipulating. PI nodes are defined as locations expressly designed to perform operations on PI containers, such as receiving, testing, moving, routing, sorting, handling, storing, picking, composing, decomposing and shipping PI containers.

Regarding modelling Sarraj and Montreuil [6] raise that, physically, a logistic service is carried out per a transport service based on a network consisting of nodes (including distribution centers, warehousing, plants, etc.), arcs to define the means of transfer of goods by freight services (road, rail, maritime service, etc.) and the final shippers/receivers (companies, organizations, or individuals). Applying the Internet analogy, a shipper sends his merchandise to a nearby node that manages it, stores it and sends it to its destination through one of the many accessible logistics plans. For this purpose, as in the case of Internet data, the merchandise is encapsulated in the form of standardized packets: PI containers.
Based on the current state of the art of the research of modelling PI physical elements and with the valuable insights from ICONET's forums and living labs, ALICE cluster and Advisory Board of the project, an innovative approach of modelling components has been defined. GPICS makes an abstraction of a real PI world system by creating a conceptual model and such a representation must be defined by four fundamental parts: lexical, structural, procedural, and semantic. In this regard, the GPICS modelling components cover and support two of these parts of the representation. On the one hand, the lexical part of the representation, which deals with the description of the symbols allowed in the vocabulary of representation, and on the other hand the semantic aspects of the representation that establish a way of associating meaning with the descriptions. This is one of the reasons why the GPICS modelling components are considered a fundamental part of the ICONET's GPICS framework. The following image shows the components that have been designed to represent the components of a supply chain in a PI environment.

GPIC structure	
GPICS Container	Unit load manipulated, stored, moved and routed through the systems and infrastructures of the Physical Internet.
GPICS Node/Hub	Location specifically designed to carry out logistics and transport processes and activities on PI containers.
GPICS Mover/Transport	Moving element used to carry PI containers through the PI nodes/hubs.
GPICS Corridor	Connection between two PI Nodes/Hubs directly connected.
GPICS Route	Set of GPICS corridors which connect a GPICS Node origin and a GPICS Node destination.
GPICS Network	Set of containers, nodes, movers/transport, corridors, and routes.
GPICS Roles	Actors/Agents involved in the operation of the PI Network.

Figure 3: GPICS modelling components.

Throughout the virtual model, different decisions must be made, for example, if it is necessary to separate the load from the receipt of a truck, how long a pallet should be stored, if it is necessary to group different containers, or what is the best transport to advance to the next destination. The following image (Fig. 4) supports the understanding of these decisions. In fact, this figure shows the schematic representation of a Generic Hub internal process. Starting from the left side, the flow begins with the reception of the products. Subsequently, there is a classification process, in which the decision is made as to whether the container should be separated or added, with other containers. There is also the possibility of temporary storage until a new order of movement for the products arrives. The lower part represents the state of the movers (i.e., trucks, trains). Initially, the movers perform the unloading of containers. Second, they can wait some time or, finally, they can load new products into the node, and continue to the next destination.

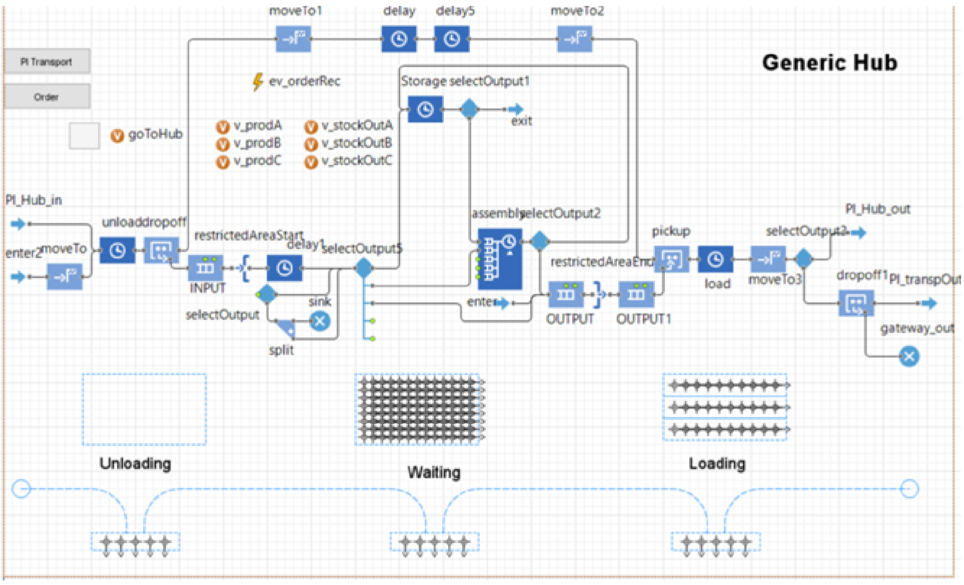


Figure 4: Generic PI Hub model

4 MODEL INTEGRATION

The designed simulation models allow to virtually represent the behavior of the different actors in the supply chain. A proposed added value is the integration with real services (e.g., networking, routing, shipping services) that supply information on the state of transports in the real world, and from this information it is possible to evaluate their evolution in different scenarios in the virtual model.

The concept of physical simulation in the PI framework refers to the simulation of models that can be performed with real data/events fed from the physical world. The physical simulation will materialize with sources of information that come from data sources available in the living laboratories.

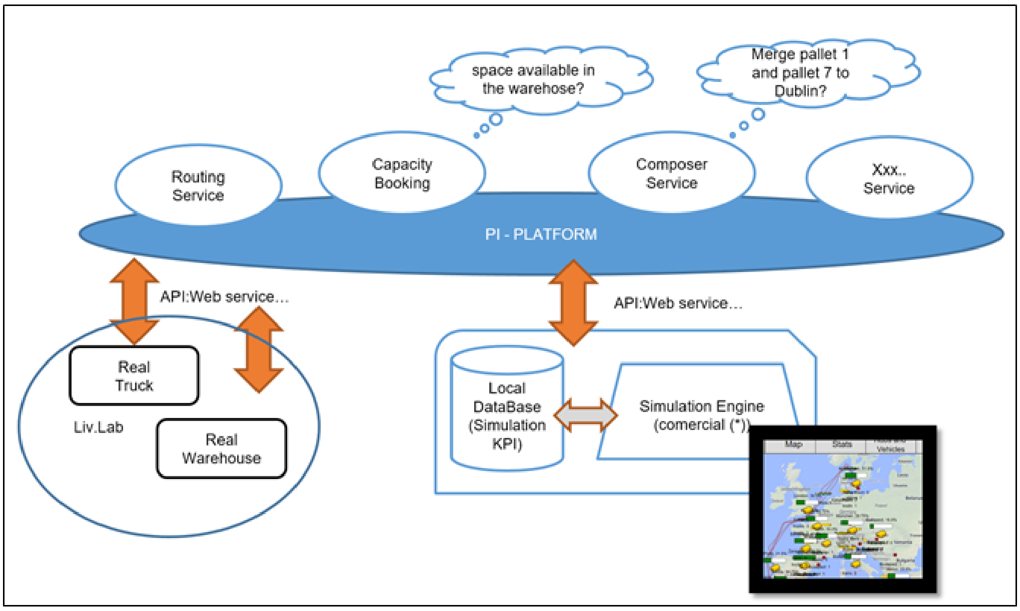


Figure 5: Simulation model integration.

The information of the physical world will be connected to the simulation models through data exchange interfaces. (i.e., from a common information management platform). In the information platform, the different physical elements, such as a truck or a container, can give information about their arrival at the destination, their departure from the port or their GPS position at a certain time through the IoT devices in the PI containers. The digital simulation can receive this information and, consequently, apply the management rules of physical internet management to evaluate the possible evolution of the network considering the real data that has just received. The following image describes the connection between the simulation (virtual model) and the different PI services developed in the ICONET project.

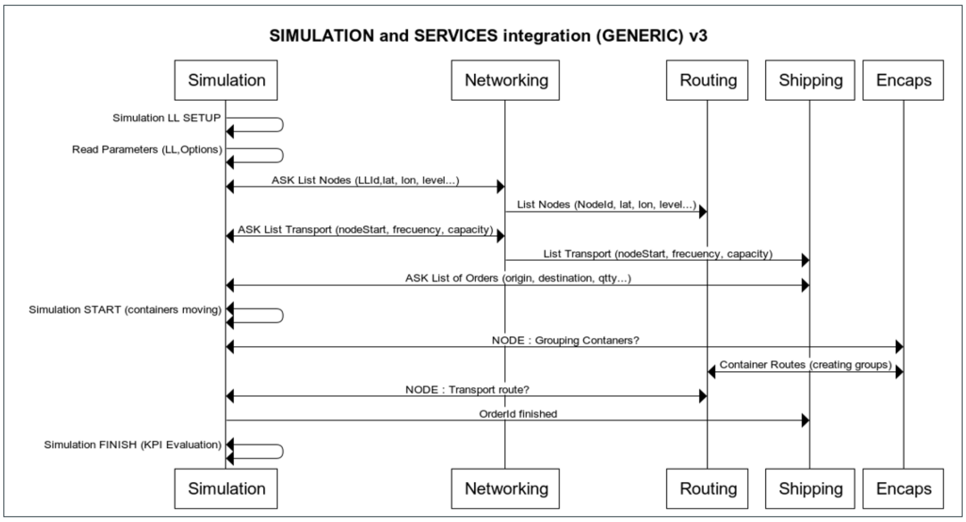


Figure 6: Simulation model and service integration.

The virtual simulation model is integrated with the different ICONET services to interact with the real elements of the physical network. Through these services the simulation models are dynamically configured, the number and position of the nodes are defined, the dynamic route for the containers is calculated, the necessary container groupings are created, and the best mode of transport is specified through the following services.

- *Networking*: Networking defines the interconnected infrastructure of available processing, storage and transporting facilities (transport services, terminals, distribution centres, warehouses) through which the goods will be transported from their origins (manufacturing, distribution and other locations) towards their customer(s) locations.
- *Routing*: Routing is a process that creates a plan that describes the stage by stage detailed visiting and usage of networking nodes from origin to destination.
- *Encapsulation*: How products to be shipped are encapsulated in modular packets that are then consolidated/deconsolidated into containers for transportation via the PI.
- *Shipping*: The shipping service specifies the cargo that is transported, the limitations and restrictions on its transport.

5 USE CASE EXAMPLES

The GPICS methodology and the described simulation models have been applied in various living labs to validate the PI modelling concepts and operating rules on real business scenarios. This article describes two use cases and how the virtual models have been applied.

The simulation of the first use case is focused on the evaluation of the performance of the IoT and the Cloud-based ICT Infrastructure for PoC Integration to foster the intermodal transportation. The goal of the Living Lab is the selection of the best route for containers in the Antwerp - Milano corridor, using intermodal transportation and considering the current position of the container and network status.

In this case, the virtual model is used to compare the impact of different transport options: direct road transport with intermodal transport, several steps are taken, in addition to the loading, road and unloading, there are terminal movements for the container and the rail transportation.

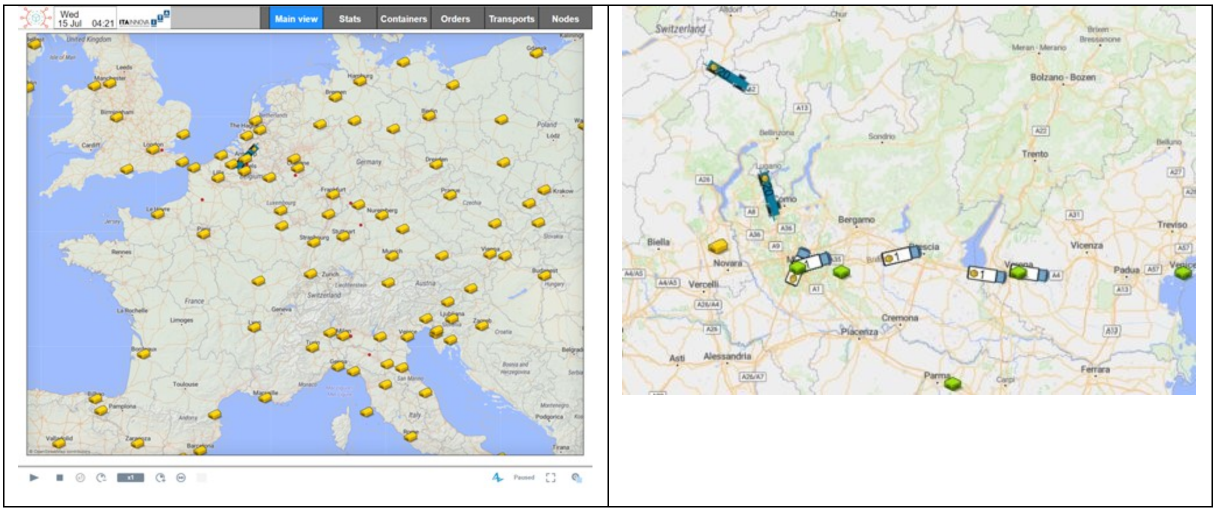


Figure 6: Virtual model representation in a European corridor.

The simulation of the second model is focused on the evaluation of the efficiency of the PI concepts in an eCommerce distribution network. It compares multiple locations (stores and warehouses) to assess best option to fulfil an order in terms of cost, lead time and stock out, considering different distribution networks, from different companies.

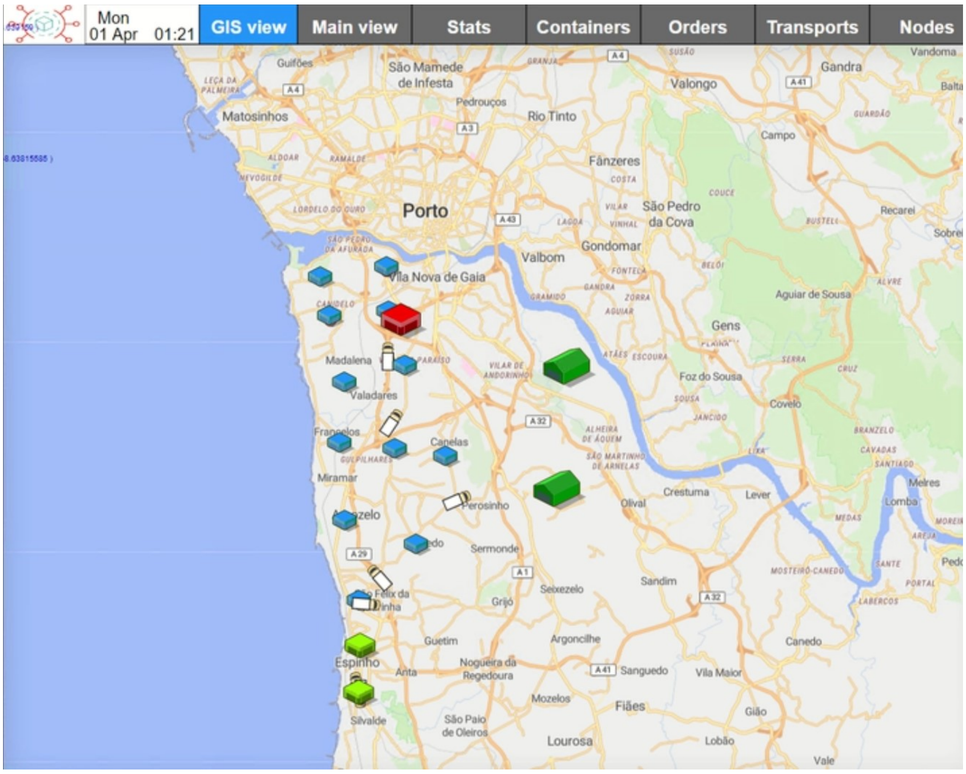


Figure 7: Virtual model representation in a urban environment

The simulation developed is a tool that helps companies to visualize how the movement of products over a PI network can be, including flows from other companies. The simulation helps to confirm the value generated by PI Services developed for the operation of Physical Internet. The virtual model helps with the quantitative assessment of the impact that PI services introduce on companies' supply chains. The visualization and sharing models between different users to evaluate the effect of PI on their own network considering economic, operational and environmental indicators.

6 CONCLUSIONS

This paper presents a methodology to create a digital twin of a transport network under the Physical Internet framework. It is based on the work developed in the ICONET project and will be continued in the PLANET project. The GPICS Framework enables the comprehensive representation of a real PI world system by creating a conceptual model that can be virtual simulated and connected with real PI services to generate accurate digital twins of the transport network. The GPICS Framework enables the comprehensive representation of a real PI world system by creating a conceptual model that can be virtually simulated. Companies need to have a better understanding of how they can take part in PI, how to benefit from PI, how to integrate in the network and what data they need to share. The simulation with virtual model is a powerful tool to visualize how the movement of products over a PI network can be, including the impact of the flows from other companies.

This report is the result of a dedicated effort in a relatively short period of time. The authors are very grateful to amongst others Makis Kouloumbis, Philippos Philippou, John Farren, Kieran Flynn and Kostas Zavitsas. Their input and feedback, direct or indirect, has been of great value to this report. Furthermore, we would like to thank all members of the ICONET consortium for having fruitful discussion and providing a productive atmosphere for applied research.

References

[1] url: https://www.iconetproject.eu/

[2] url: https://www.etp-logistics.eu/progress-of-sense-project-on-developing-physical-internet-pi/

[3] B. Montreuil, R. D. Meller, and E. Ballot, “Towards a Physical Internet : the impact on logistics facilities and material handling systems design and innovation,” Prog. Mater. Handl. Res., p. 23, 2010.

[4] Zheng, Y., S. Yang, and H. Cheng. 2019. “An application framework of digital twin and its case study”. Journal of Ambient Intelligence and Humanized Computing 10(3):1141–1153.

[5] url : https://www.planetproject.eu/

[6] R. Sarraj and B. Montreuil, “Analogies Between Internet Networks and Logistics Service Networks : Challenges Involved in the Interconnection Analogies Between Internet Networks and Logistics Service Networks :,” J. Intell. Manuf., vol. 25, p. 1207–1219., 2014.

Enhancing Logistics Demand Prediction Accuracy Through Client–Vendor Hyperconnected Data Ensembles

Xinyue Pan[1,2,4], Ashwin Pothen[1,2,4], Jana Boerger[1,2,4], He Wang[1,2,4] and Benoit Montreuil [1,2,3,4]

1. Physical Internet Center
2. Supply Chain and Logistics Institute
3. Coca-Cola Chair in Material Handling & Distribution
4. H. Milton Stewart School of Industrial & Systems Engineering

Georgia Institute of Technology, Atlanta, USA

Corresponding author: xpan67@gatech.edu

Abstract: *In this paper, we present a hyperconnected data ensemble framework under the Physical Internet (PI) paradigm. As the current world is more volatile, uncertain, complex and ambiguous than before, it is well known that today's logistics and supply chain management (LSCM) is facing greater difficulties and the idea of PI is introduced as a solution to the logistics sustainability grand challenge. PI aims to achieve significant logistics system efficiency and sustainability improvement through universal interconnectivity and smart open coordination. Meanwhile, PI will facilitate data sharing among supply chain parties. Instead of traditional data sharing by integrating datasets in common cases, we suppose that the hyperconnected data ensembles require as minimal data as possible. With the utilization of aggregated information, such as the overall activity forecast, the hyperconnected data ensembles can enhance the accuracy of the logistic demand prediction while preserving data privacy. A framework of logistics demand prediction with hyperconnected data ensembles is established and results of some experiments conducted based on the framework support our hypothesis that the demand prediction accuracy can be increased by integrating the forecast data that clients are willing to provide.*

Conference Topic(s): *Systems and Technologies for Interconnected Logistics*

Keywords: *Hyperconnected Logistics, Ensemble Methods, Demand Forecasting, Data Sharing, Physical Internet, Supply Network Coordination and Collaboration, Supply Chain Efficiency, Supply Chain Visibility*

1 Introduction

As the current world is more volatile, uncertain, complex and ambiguous than before (Bennett & Lemoine, 2014), it is well known that today's logistics and supply chain management (LSCM) is facing greater challenges, notably related to supply chain visibility and cost control. In order to minimize costs, LSCM should be more agile to quickly respond to the changing market. Meanwhile, supply chain visibility should also include visibility among partners such that collaborative decisions can be made more closely tailored to client needs. None of these could be done without the cooperation among supply chain stakeholders. Therefore, there is a need to transform the traditional independent and discontinuous LSCM to become more interdependent and continuously hyperconnected. This transformation can be done by adopting the Physical Internet (PI) paradigm.

Physical Internet was originally defined by (Montreuil, 2011) to address the issue of physical objects being transported, handled, stored, realized and used in an unsustainable way. PI aims to achieve order-of-magnitude improvements in logistics system capability, efficiency, resiliency, and sustainability through universal seamless interconnectivity and smart open coordination based on a metaphor of the Digital Internet. Previous researches have demonstrated potential economical, environmental, and societal benefits of applying PI in interurban and urban transportation and logistics (Montreuil et al., 2012; Ballot et al., 2014; Darvish et al., 2016; Sallez et al., 2016; Kim et al., 2020). Industry and territories are picking up the pace toward the Physical Internet, as illustrated by the recent publication by the industry-led European Alliance for Logistics Innovation through Collaboration in Europe of its Roadmap to the Physical Internet (ALICE-ETP, 2020).

PI facilitates data sharing among supply chain parties and vice versa. It makes the data sharing process among participants easier. As an illustration, PI promotes the cooperation between logistic providers and clients such that the necessary cloud-based IT systems for the data sharing process can be set up more smoothly. In addition, PI is a key trigger for digitalizing LSCM because it can serve as a practical paradigm for digital supply chain design (Pan et al., 2021) and the digitalization of LSCM makes data sharing simpler than before. Open collaboration under the PI paradigm, aiming at achieving efficient, effective, and sustainable logistics, is aided by data sharing. As demonstrated, data sharing supports seamless interconnection of logistic networks which requires accurate forecast of parcel arrivals. Data sharing also helps to promote dynamic decision-making, especially helpful for the modular containers of PI (Tretola et al., 2015).

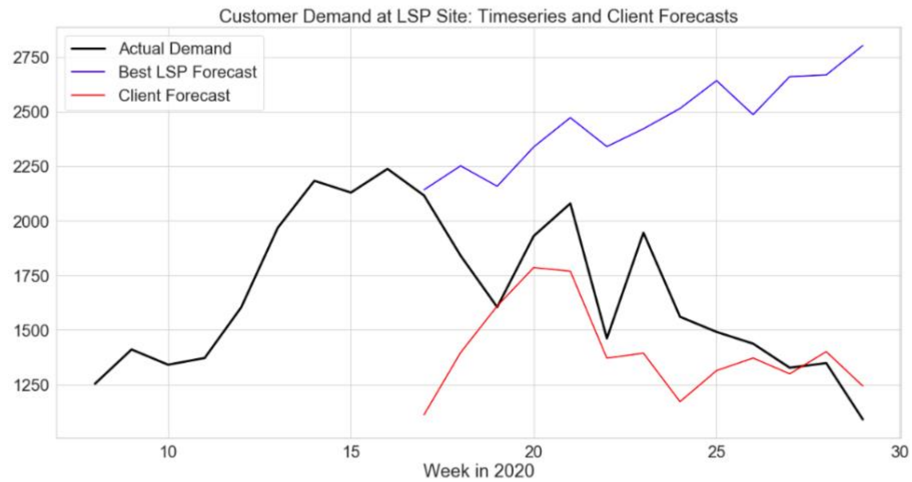


Figure 1: Comparing LSP forecasts and client-provided forecasts with real client demand at LSP site

Figure 1 illustrates the insights provided by client forecasts at a Logistics Service Provider (LSP) site, where traditional time series forecasting methods are influenced by the recent demand levels and trends while client forecasts capture the market information which cannot be deduced from data alone. However, despite the mutual benefits that can be gained from sharing data, most companies are hesitant to do so, often due to data privacy concerns. Companies are worried about losing their competitive edge in the market if sensitive data shared with their partners gets stolen and leaked to their competitors. In this paper, we show that the hyperconnected data ensemble method is different from traditional data integration methods that the contents of shared data is dependent on the choices of clients, such that the client can preserve the privacy of sensitive information.

In order to support our hypothesis that hyperconnected data ensembles can enhance the performance of LSCM without integrating entire clients' datasets, experiments of logistics demand prediction of the logistic service provider (LSP) with data ensembles from its upstream and/or downstream clients are conducted. The results show that the accuracy of the demand prediction increases by only integrating the aggregated forecast information shared from the client. An accurate forecast of warehouse activities can reduce the chance of service failures which include process failures such as an operational gridlock at warehouses, or outcome failures such as inability to receive incoming shipments or to pick and prepare outgoing shipments, and thus not meeting service agreement. Accurate forecasts can therefore increase the chance of the 3PL company to meet client service expectations. Enhanced information sharing can help supply chain players gain hyperconnectivity and alleviate the aforementioned service failures. It enables close cooperation and coordination among logistics service providers and upstream and downstream clients and hence enhances the performance of the entire supply chain system.

In this paper, we address the Predictive Layer of the analytics framework for logistics service providers introduced by Boerger & Montreuil (2020). In this layer, future logistics activities are forecasted, predicted and assessed. It builds on the Descriptive Layer in which logistics activities are monitored, analyzed and diagnosed and offers input to the Prescriptive Layer in which decisions are made. The rest of this paper is organized as follows. In section 2, we present a brief literature review of the data privacy issue. In section 3, we provide the framework of logistic demand prediction with hyperconnected data ensembles. In section 4, we describe the experiments to support our hypothesis that that the demand prediction accuracy can be increased by integrating the forecast data that clients are willing to provide. In section 5, the illustrative numerical results of the experiment will be shown. Finally, section 6 concludes the work and provides the avenues for future works.

2 Literature Review

Data sharing refers to the situation that collaborators' systems are allowed to access private data which attempts to provide timely and accurate information to the decision makers in the supply chain. A large part of the data sharing literatures focuses on illustrating the benefits of data sharing and the types of data that can be shared (Lee & Whang, 2000; Themistocleous et al., 2004). Few of these papers provide concrete examples showing how to utilize the shared data to improve the performance of the supply chain. In this paper, we ensemble the shared client forecasts into the existing forecasts from the logistics' provider and conduct experiments to support our proposed framework.

Data sharing is usually achieved through the integration of databases. As shown in (Tretola et al., 2015), the proposed method for data sharing is Enterprise Application Integration (EAI). EAI is popular for information integration on both intra- and inter organizational level (Linthicum, 2000). However, data security and privacy remain an issue on this kind of data sharing, where data leakage may cause severe consequence for both logistics' provider and clients. Existing literature mainly solves the problem by modifying the key contributions of the raw data, such as in (Sweeney, 2002; Dwork, 2010), but the shared data still faces the danger of algorithm or background knowledge-based attack. A popular way as shown in (Lu et al., 2020) to avoid the data leakage problem is federated learning. The main idea of federal learning (Konečný et al., 2016) is to collaboratively train a shared global model without storing the data centrally. In that case, clients store their own raw data locally and never share with their collaborators and the coordinating server. Instead, clients only share their locally updated model to the central server and the server combines the partially trained models to form a federated model. Therefore, federated learning eases the leakage of private data. However, there are some algorithmic and technical challenges or limitations in the real-world utilization. For instance, all collaborators must have some basic knowledge about how to train the global model. Moreover, all shared data points in the federated learning should follow the same distribution.

Either sharing the model updates using federal learning, or a revised version of raw data can nonetheless reveal sensitive information. We investigate the case of only utilizing insensitive data or the data that clients are willing to provide. In this paper, we present a hyperconnected data ensemble under the PI paradigm. Unlike data integration processes where clients are asked to provide all information, hyperconnected data sharing requires as minimal data as possible. The content of shared data is dependent on the choices of clients. For example, clients may only exchange aggregated data such as the overall activity forecast without providing the detailed information on product level or SKU level. Clients do not need to disclose confidential data like the key model, so the issues of data privacy can be avoided. Furthermore, results of our experiments support our assumption that the demand prediction accuracy can be increased by using the forecast data that clients are willing to provide.

3 Framework

In this section, we construct a basic framework of logistics demand prediction with hyperconnected data ensembles, shown in Figure 2. The black blocks represent the processes on the client side and the orange blocks belongs to the logistic service provider (LSP).

In the LSP side, historical data is to be processed using a variety of methods, including time-series and machine learning methods for model training, parameter tuning, and forecast generation. Each forecast method is to generate its corresponding forecasted results. Then, the LSP will also generate the ensembled forecasts without client forecasts, denoted as $F_e^c(1)$, using ensemble methods.

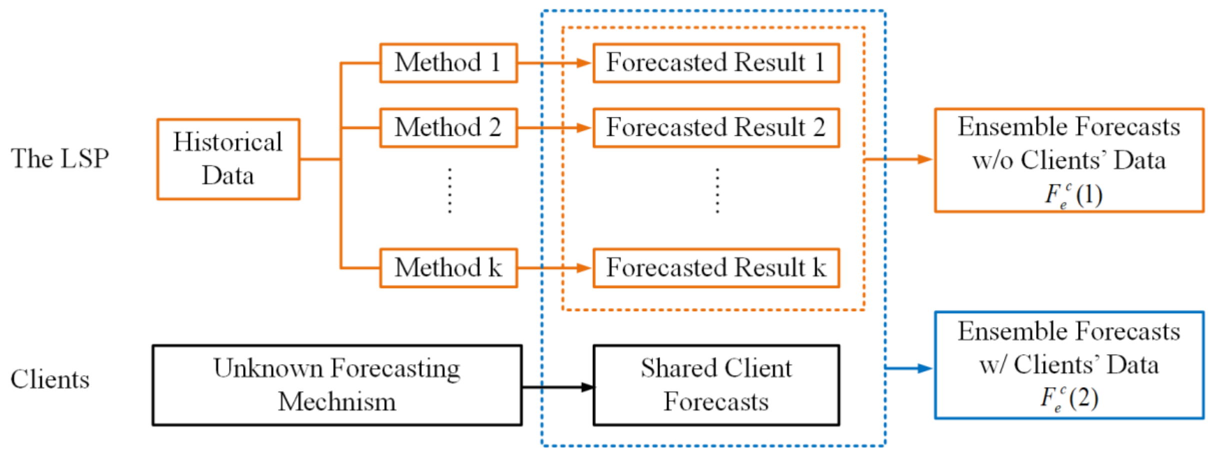


Figure 2: The Framework of logistics demand prediction with hyperconnected data ensembles

We assume that clients will generate forecasts using their own mechanism as well. The forecast mechanism and the data used for client's prediction are unknown to the LSP in order to preserve data privacy. Typically, the LSP is to receive the aggregated level of forecasts and the shared client forecasts can be ensembled with the forecasts generated by the LSP.

Multiple ensemble algorithms (e.g., stacking (Ting &Witten, 1997); max voting (Romera-Paredes et al., 2013); boosting (Mayr et al., 2014); etc.) can be utilized in the framework. In this paper, we use a weighted averaging method. Assume that the client is only willing to share its overall activity forecast and the LSP aims to exploit the shared information to generate a more accurate forecast for the demand induced by a specific client. In this case, the forecasted results for the LSP are for the specific client as well. The formulas of the ensemble method are shown in equations ((1)-(2)):

$$F_e^c(1) = \sum_{j=1}^{k} F_{LSP_j}^c \alpha_j, \ \sum_{j=1}^{k} \alpha_j = 1, \tag{1}$$

$$F_e^c(2) = \sum_{j=1}^{k} F_{LSP_j}^c \alpha_j + F_c^c \alpha_{k+1}, \ \sum_{j=1}^{k+1} \alpha_j = 1, \tag{2}$$

where the superscript c refers to the client level; $F_e^c(1)$ represents the ensembled forecast without shared client data and $F_e^c(2)$ represents the final forecast with client data ensembles; $F_{LSP_j}^c$ is the forecasted result generated by the j-th method from the LSP and F_c^c is the forecast shared by clients; α refers to the weighting parameter and the value of α falls between 0 and 1.

4 Experiments

We report on an experiment realized in the context of a collaborative research project between Georgia Tech's Physical Internet Center and major North American logistics service provider (LSP). The experiment focuses on demand forecasting for a major client of the LSP at two of its sites, this client having accepted to collaborate in the experiments. The demand forecasts are, expressed in terms of the number of pallets predicted to be shipped for the specific client in each site at a weekly level. The LSP has access to 2.5years of historical demand data for the client's activity in the 2 sites at the daily level. The historical data is subject to seasonal fluctuations, which has been validated through explanatory data analysis. As the demand data, (i.e., number of pallets) is a time-series data, it is intuitive to employ time-series methods. We apply four time-series methods to generate forecasts using the historical data obtained by the LSP: Holt-Winters (HW) (Winters, 1960), Double-Seasonal (DS) (Taylor, 2003), Multi-Seasonality (MS) (Gould et al., 2008) and Bouchard-Montreuil (BM) (Bouchard and Montreuil, 2009) method. Corresponding to the equations ((1)-(2)), we have:

$$F_{LSP_1}^c = HW, \tag{3}$$

$$F_{LSP_2}^c = DS, \tag{4}$$

$$F_{LSP_3}^c = MS, \tag{5}$$

$$F_{LSP_4}^c = BM, \tag{6}$$

$$F_c^c = \text{Clients' Forecasts}. \tag{7}$$

Historical data of the LSP is for the daily level, but the goal is to generate a forecast at the weekly level. In the experiment, daily forecasts are generated through each method, and the weekly forecasts are obtained by aggregating the daily forecasts into weekly forecasts.

The client provides weekly aggregated demand forecast with a 13-week horizon per site. In other words, at the start of every week, the client generates the weekly forecast for the next 13 weeks and share these forecasts with the LSP. With access to 50 weekly forecasts provided by the client, we generate the logistics demand forecasts with data ensembles.

The validation of the experiment is done by Monte-Carlo cross validation, where we perform a fixed number $N$of iterations (here set to 30). For each iteration $i \in \{1, 2, \ldots, N\}$, we randomly select 80% of dates from the total set of start-dates S (here $|S| = 50$). The training set consisted of forecasts generated for the selected start-dates $\hat{S} \in S$ and the forecasts by the unselected start-dates $S' = S \backslash \hat{S}$. The training error for i-th iteration, E_i^{TR}, is calculated as follow:

$$E_i^{TR} = \frac{1}{|\hat{S}|} \sum_{s \in \hat{S}} MAE_s^i = \frac{1}{|\hat{S}|} \sum_{s \in \hat{S}} \frac{1}{13} \sum_{h=1}^{13} \left| A_{s,h}^i - F_{s,h}^i \right| , \tag{8}$$

where the superscript/subscript i refers to the iteration number, $\hat{S}$ refers to the set of start-dates in the training set, and MAE_s^i refers to the Mean Absolute Error (MAE) of the ensemble forecasts for a given start-date $s \in \hat{S}$. $A_{s,h}^i$ represents the actual demand starting from date s for h period ahead and similarly, $F_{s,h}^i$ represents the ensembled forecast made from s for h period ahead using the equation ((1) or (2)).

After all the iterations, we calculate the cross-validation (CV) error E^{TR} by averaging the errors for all training sets using the following equation:

$$E^{TR} = \frac{1}{N} \sum_{i=1}^{N} E_i^{TR}. \tag{9}$$

Finally, the optimal weighting parameters are selected by minimizing E^{TR}. Then, the forecasts in the test set at each iteration are ensembled using the selected optimal weighting parameters and are used to calculate the test error E_i^{TE} and cross-validation test error E^{TE} similar to the procedure in the training set:

$$E_i^{TE} = \frac{1}{|S'|} \sum_{s \in S'} MAE_s^i = \frac{1}{|S'|} \sum_{s \in S'} \frac{1}{13} \sum_{h=1}^{13} \left| A_{s,h}^i - F_{s,h}^i \right|, \tag{10}$$

$$E^{TE} = \frac{1}{N} \sum_{i=1}^{N} E_i^{TE}. \tag{11}$$

5 Results

Based on the experimental setup, we compute the ensembled forecast using the equations ((1) – (2)) for two sites. For site 1, the optimal weighting parameter for the ensemble without the client forecast is shown in equation (12). Equation (13) shows the formula for calculating the ensemble forecast with the client forecast.

$$F_e^c(1) = 0.382 * F_{LSP_1}^c + 0.403 * F_{LSP_2}^c + 0.183 * F_{LSP_3}^c + 0.032 * F_{LSP_4}^c, \tag{12}$$
$$F_e^c(2) = 0.252 * F_{LSP_1}^c + 0.297 * F_{LSP_2}^c + 0.067 * F_{LSP_3}^c + 0.056 * F_{LSP_4}^c + 0.328 * F_c^c. \tag{13}$$

The CV errors $\boldsymbol{E}^{T:(TR\ or\ TE)}$for training and test sets of site 1 are shown in Table 1. Improvement I refers to the decrease in CV Error as described in equations (9) and (11), and is calculated for demand forecast method F versus the best LSP forecast F' as:

$$\mathrm{I} = \frac{E_{\hat{F}}^T - E_F^T}{E_{\hat{F}}^T}, \tag{14}$$

Among all the time-series methods, the DS method performs best with lowest training and test error. After ensembling all time-series models, there is a 1.66% and 1.53% increase of accuracy of the training and testing errors, respectively. The forecast provided by the client has lower accuracy than both the DS forecast and the ensembled LSP forecast. However, with the ensembles of client forecast, the accuracy of $F_e^c(2)$ is the highest among all forecasts with 4.95% and 8.15% decrease of errors in training and test errors, respectively.

Table 1: CV errors and improvement for the site 1

	Cross-Validation Error		Improvement	
Demand Forecast Method	**Train**	**Test**	**Train**	**Test**
Best LSP Forecast *DS*	467.27	456.71		
Ensemble of LSP Forecasts $F_e^c(1)$	459.50	449.71	1.66%	1.53%
Client Forecast $F_{customer}^c$	503.81	482.28	-7.82%	-5.60%
Ensemble with Client Forecast $F_e^c(2)$	444.15	419.50	4.95%	8.15%

Table 2: CV errors and improvement for the site 2

	Cross-Validation Error		Improvement	
Demand Forecast Method	**Train**	**Test**	**Train**	**Test**
Best LSP Forecast *DS*	450.09	485.31		
Ensemble of LSP Forecasts $F_e^c(1)$	450.67	485.75	-0.13%	-0.09%
Client Forecast $F_{customer}^c$	342.15	359.69	23.98%	25.88%
Ensemble with Client Forecast $F_e^c(2)$	324.42	338.09	27.92%	30.34%

Similarly, we did the same experiment in the site 2. The optimal weighting parameter for the ensemble without client forecast is shown on equation (13). The equation (14) represents the formula of calculating the ensemble forecast with the client forecast.

$$F_e^c(1) = 0.020 * F_{LSP_1}^c + 0.977 * F_{LSP_2}^c + 0.0032 * F_{LSP_3}^c, \tag{14}$$

$$F_e^c(2) = 0.0004 * F_{LSP_1}^c + 0.235 * F_{LSP_2}^c + 0.028 * F_{LSP_3}^c + 0.736 * F_c^c. \tag{15}$$

The CV errors for both training and testing sets of the site 2 is shown in Table 2. As in the site 1, the DS method outperforms all other time-series models in terms of training and testing CV errors. However, the ensemble of all time-series model does not improve the forecasting performance; but results in a 0.13% and a 0.09% decrease in accuracy of training and test sets, respectively. The client's forecasts are much more accurate, with a decrease in CV errors of 23.98% and 25.88%. Finally, with the ensemble of client forecasts, the accuracy of $F_e^c(2)$ is the highest among all forecasts, with 27.92% and 30.34% decrease in train and test errors, respectively. $F_e^c(2)$ in the site 2 is more precise than the site 1, owing to the quality of the client forecast.

6 Conclusion

In this paper, we propose a hyperconnected data ensemble framework based on data sharing under the Physical Internet (PI) paradigm. The difference between hyperconnected data sharing and traditional data integration is that hyperconnected data sharing requires as minimal data as possible and we suppose that hyperconnected data sharing still maintains the ability to enhance the performance of the logistics and supply chain management (LSCM). Under PI, the content of shared data is dependent on the choice of clients and the data privacy issues which commonly happen in the traditional data integration can be avoided. In order to support our hypothesis, we construct a framework and experiments on logistics demand prediction with hyperconnected data ensembles. The preliminary results of the experiments have proven that the forecast errors for individual client measured by the cross-validation error have decreased by 4.95% in the training error and 8.15% in the testing error in site 1, and by 27.92% in the training error and 30.24% in the test error in site 2 after integration with the clients' forecasts. This suggests that the ensemble of the LSP's data with the client's data improves overall forecasting accuracy which can set the basis for improved supply chain decision making.

There are numerous choices of ensemble methods that can be applied to the integration of client and vendor forecasts. In this paper, the conducted experiments utilize a simple weighted averaging method, paving the way for further research investigating more complex versions of ensemble methods. Here, we show the ensemble of the client forecast into LSP's forecasts. Further research may experiment the value of integrating different types of non-privacy-compromising information that clients are willing to provide into the demand prediction. Finally, further research is needed on how to protect data privacy in hyperconnected supply chains, while exploiting data sharing and new technologies, and enabling better predictions and mutual performance.

References

ALICE-ETP (2020). Roadmap to the Physical Internet, Executive Version, http://www.etp-logistics.eu/wp-content/uploads/2020/11/Roadmap-to-Physical-Intenet-Executive-Version_Final.pdf

Ballot, E., Montreuil, B., & Meller, R. (2014). *The Physical Internet*. https://hal-mines-paristech.archives-ouvertes.fr/hal-01113648

Bennett, N., & Lemoine, G. J. (2014). What a difference a word makes: Understanding threats to performance in a VUCA world. In *Business Horizons* (Vol. 57, Issue 3, pp. 311–317). Elsevier Ltd. https://doi.org/10.1016/j.bushor.2014.01.001

Boerger, J., & Montreuil, B. (2020). *Data-driven analytics-based capacity management for hyperconnected third-party logistics providers*.

Bouchard, J., and B. Montreuil (2009). Méthode de prévision journalière de séries temporelles saisonnières. *Congrès international de génie industriel–CIGI-2009*

Darvish, M., Larrain, H., & Coelho, L. C. (2016). A dynamic multi-plant lot-sizing and distribution problem. *International Journal of Production Research*, *54*(22), 6707–6717. https://doi.org/10.1080/00207543.2016.1154623

Dwork, C. (2010). Differential privacy in new settings. *Proceedings of the Annual ACM-SIAM Symposium on Discrete Algorithms*, 174–183. https://doi.org/10.1137/1.9781611973075.16

Gould, P. G., Koehler, A. B., Ord, J. K., Snyder, R. D., Hyndman, R. J., & Vahid-Araghi, F. (2008). Forecasting time series with multiple seasonal patterns. *European Journal of Operational Research*, *191*(1), 207–222. https://doi.org/10.1016/j.ejor.2007.08.024

Kim, N., B. Montreuil, W. Klibi, N. Kholgade (2020). Hyperconnected Urban Fulfillment and Delivery, *Transportation Research Part E Logistics and Transportation Review*, Vol. 145, 23p. https://doi.org/10.1016/j.tre.2020.102104

Konečný, J., McMahan, H. B., Yu, F. X., Richtárik, P., Suresh, A. T., & Bacon, D. (2016). *Federated Learning: Strategies for Improving Communication Efficiency*. http://arxiv.org/abs/1610.05492

Lee, H. L., & Whang, S. (2000). Information sharing in a supply chain. *International Journal of Manufacturing Technology and Management*, *1*(1), 79–93.

https://doi.org/10.1504/IJMTM.2000.001329

Linthicum, D. S. (2000). *Enterprise application integration*. Addison-Wesley Professional,. https://www.worldcat.org/title/enterprise-application-integration/oclc/42397664

Lu, Y., Huang, X., Dai, Y., Maharjan, S., & Zhang, Y. (2020). Blockchain and Federated Learning for Privacy-Preserved Data Sharing in Industrial IoT. *IEEE Transactions on Industrial Informatics*, *16*(6), 4177–4186. https://doi.org/10.1109/TII.2019.2942190

Mayr, A., Binder, H., Gefeller, O. and Schmid, M., (2014). The evolution of boosting algorithms-from machine learning to statistical modelling. *arXiv preprint arXiv:1403.1452*.

Montreuil, B., Ballot, E., & Fontane, F. (2012). An open logistics interconnection model for the physical internet. *IFAC Proceedings Volumes (IFAC-PapersOnline)*, *45*(6 PART 1), 327–332. https://doi.org/10.3182/20120523-3-RO-2023.00385

Montreuil, Benoit. (2011). Toward a Physical Internet: meeting the global logistics sustainability grand challenge. *Logistics Research*, *3*(2–3), 71–87. https://doi.org/10.1007/s12159-011-0045-x

Pan, S., Trentesaux, D., McFarlane, D., Montreuil, B., Ballot, E., & Huang, G. Q. (2021). Digital interoperability in logistics and supply chain management: state-of-the-art and research avenues towards Physical Internet. *Computers in Industry*, *128*, 103435. https://doi.org/10.1016/j.compind.2021.103435

Romera-Paredes, B., Aung, M.S. and Bianchi-Berthouze, N., 2013, April. A One-Vs-One Classifier Ensemble With Majority Voting for Activity Recognition. *In Esann.*

Sallez, Y., Pan, S., Montreuil, B., Berger, T., & Ballot, E. (2016). On the activeness of intelligent Physical Internet containers. *Computers in Industry*, *81*, 96–104. https://doi.org/10.1016/j.compind.2015.12.006

Sweeney, L. (2002). L. Sweeney. k-anonymity: a model for k-ANONYMITY: A MODEL FOR PROTECTING PRIVACY 1. In *International Journal on Uncertainty, Fuzziness and Knowledge-based Systems* (Vol. 10, Issue 5).

Taylor, J. W. (2003). Short-term electricity demand forecasting using double seasonal exponential smoothing. *Journal of the Operational Research Society*, *54*(8), 799–805. https://doi.org/10.1057/palgrave.jors.2601589

Themistocleous, M., Irani, Z., & Love, P. E. D. (2004). Evaluating the integration of supply chain information systems: A case study. *European Journal of Operational Research*, *159*(2 SPEC. ISS.), 393–405. https://doi.org/10.1016/j.ejor.2003.08.023

Ting, K.M. and Witten, I.H., 1997. Stacking bagged and dagged models.

Tretola, G., Biggi, D., & Verdino, V. (2015). *Enabling Logistics Data Sharing in the Physical Internet*. https://www.researchgate.net/publication/282607787_Enabling_Logistics_Data_Sharing_in_the_Physical_Internet

Winters, P. R. (1960). Forecasting Sales by Exponentially Weighted Moving Averages. *Management Science*, *6*(3), 324–342. https://doi.org/10.1287/mnsc.6.3.324

Potential Access Hub Network Design Optimization in Hyperconnected Urban Logistics

Praveen Muthukrishnan[1,2,4] , Louis Faugère [1,2,4] and Benoit Montreuil[1,2,3,4]

1. Physical Internet Center
2. Supply Chain and Logistics Institute
3. Coca-Cola Material Handling & Distribution Chair
4. H. Milton Stewart School of Industrial & Systems Engineering

Georgia Institute of Technology, Atlanta, USA

Corresponding author: pmuthukr3@gatech.edu

Abstract: *In the context of Hyperconnected Logistics Network, access hubs are first-tier consolidation and transshipment points which facilitate direct transfer of parcels between unit zones via couriers. These hubs can foster resiliency by being assigned to multiple unit zones when located at their intersections. In this paper, we propose an optimization based approach leveraging modularity to design a potential access hub network that need not be implemented as designed, but rather fed into large scale network design and clustering optimization models to improve its solvability when accounting for comprehensive inter-hub flow. Then, we study the efficiency and robustness of the proposed model through a set of experiments performed with an illustrative case representative of world's megacity inspired from a large parcel express carriers' operation. Results indicate that we should leverage on the capability of an access hub to serve more unit zones given its capacity restrictions. Finally, we identify promising future avenues for research and innovation enabling Logistics Service Providers (LSPs) in achieving their goals of serving efficiently and sustainably while offering fast and precise delivery services.*

Conference Topic(s): *Ports and hubs in Interconnected freight transport, logistics and supply networks*

Keywords: *Hyperconnected City Logistics, Physical Internet, Omnichannel Supply Chains, LSPs, Large scale hub-location problem*

1 Introduction

With the rise in e-commerce over the last decades, there has been a dramatic global shift in the demand for customer responsiveness. Buying items online has become a simple process which can be done within a matter of minutes at the click of a button. This growth is particularly pronounced during the COVID-19 pandemic, in which logistics firms have been inundated with online orders. For instance, between March and May 2020, when many stores were shuttered, online commerce increased 21% from March 2019 to March 2020, accompanied by a 45% rise in online purchases that used to be in-store purchases. In the view of many analysts, those behavioral changes may be permanent, Gurram et al. (2021). To cater to these changes and remain competitive by serving efficiently and sustainably, logistics service providers (LSP's) are challenged to break away from their conventional methods of operations such as hub-and-spoke networks and point-to-point networks.

In this direction, Montreuil et al. (2018) proposes a transformation towards a logistics topology web based on multi-tier hierarchical space structuring and an interconnected multi-plane logistics mesh networks, applying the concepts of hyperconnectivity underpinning the Physical Internet. This transformation is exacerbated in megacities which are the focus of this paper. At the first tier, megacities are represented as a mesh of unit zones constituting a demand area, such as a neighborhood, a city block, a university campus, an industrial park, a high-rise building, or a set of stories of high-rise building, clustering the pickup and delivery locations into courier work zones as shown in Figure 1, Montreuil et al. (2018). Each of the unit zones is a relatively small geographical area and can be defined as an intricate polytope except for high rise buildings, where we must specify height as well.

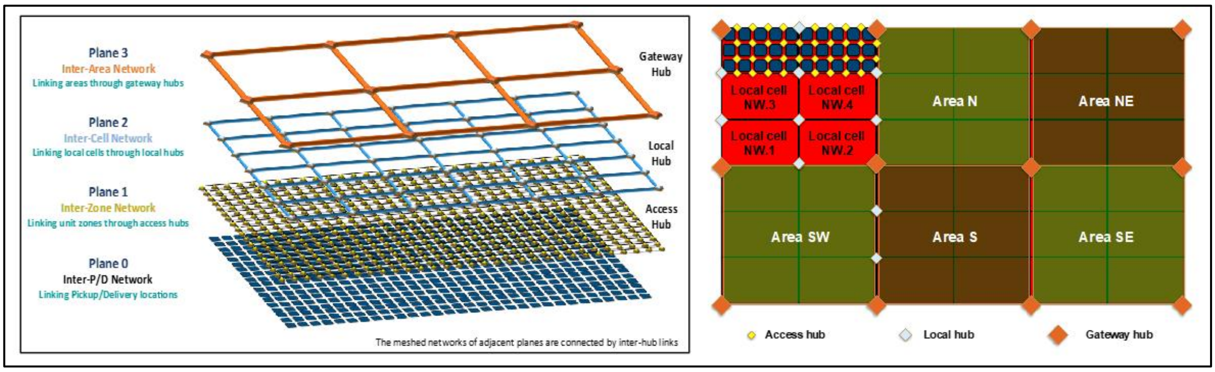


Figure 1: Urban Parcel Logistics Web Topology, Montreuil et al. (2018)

Access hubs are first-tier consolidation and transshipment points which facilitate direct transfer of parcels between unit zones via couriers and these hubs can foster resiliency by being assigned to multiple unit zones when located at their intersections. Due to urban infrastructure, a vast number of sites can be identified for the role of access hub in the neighborhood of each unit zone. Filtering out potential locations among this huge set considering the courier travel, real-estate availabilities and local government regulations is a key decision process for LSPs so that their optimization efforts taken towards overall urban logistic network design and urban space clustering remains tractable when dealing with large scale cities.

This paper proposes an optimization based approach for designing a potential access hub network in the presence of large set of candidate locations. Such potential access hub networks do not aim to be implemented as designed, but rather to be the input to a comprehensive urban logistic design process. In the case of a stationary access hub network, it provides candidate set of access hub locations that is of a scale amenable to optimization accounting for comprehensive inter-hub flow modelling. In the case of a mobile hub network, it provides the set of locations from which are to be selected the actual locations where mobile hubs will be dynamically located through the day (or week) depending on demand fluctuations, through the optimization framework proposed in Faugère et al. (2020). Potential access hub network design purposefully makes simplifying assumptions to enable optimization in a very large scale context. The proposed optimization modelling approach accounts approximately for pickup and delivery demand stochasticity, service level requirements and hub capacity. Given these considerations, the proposed model optimizes hub-zone covering given targets relative to the maximum number of unit zones that an access hub may serve, and the number of access hubs serving each zone.

Section 2 summarizes the advances in literature regarding the large-scale hub location allocation problems. Section 3 describes the problem studied and the proposed framework in detail. Section 4 presents the set of experiments performed with an illustrative case representative of a megacity inspired from a large parcel express carrier's operation. Section 5 underscores the key takeaways and guidelines for the researchers and managers and identifies promising avenues for future research and innovation.

2 Related Work

Hub location allocation problems have garnered a lot of attention from researchers over the past few decades. A comprehensive review of key developments related to this problem can be found in Campbell et al. (2012). It has applications in several areas but for the context of this paper, we study its evolution in addressing challenges faced by the freight transportation industry particularly parcel-delivery networks.

In most of the previous literature, hubs are referred to facilities that serve as switching, transshipment, and sorting points where most of the traffic are routed to tap into economies of scale (Alumur et al., 2008). Hub location allocation problems are concerned with deciding on the number, location, and size of these points along with its allocation to demand nodes. There are essentially two types of basic networks – single allocation and multiple allocation. In single allocation network, all traffic related to a demand node is routed through single hub while in multiple allocation, demand nodes can receive and send parcels through multiple hubs.

The first acknowledged mathematical formulation of hub allocation location problem started with the seminal work of (O'Kelly et al.,1986; O'Kelly et al., 1987). He introduces a quadratic programming model for uncapacitated *p*-hub location single allocation problem by studying airline passenger networks, where *p* uncapacitated hubs have to be located to exchange traffic among *n* nodes. Campbell (1994) presents an integer programming formulation for discrete hub location problems. Following this, Ernst et al. (1999) considers a capacitated single allocation hub location problem motivated by a postal delivery application and proposed simulating annealing to obtain good upper bounds that are used in a branch-and-bound algorithm. Bruns et al. (2000) discusses restructuring of Swiss Postal Services as a simple allocation location problem with more emphasis on determining the cost parameters of the model. Generally, in classical capacitated hub location problem, the flow that can be received by the hubs i.e.) capacity of hubs are limited. da Graca Costa et al. (2008) presents two alternative bi-criteria single allocation models that tries to minimize the time to process the flow entering the hubs rather than limiting the capacity of the hubs.

Several studies have been conducted in multiple allocation hub location problem as well. Aykin (1993) studied a capacitated hub-and-spoke design problem under a non-restrictive policy allowing two-stop services. Ebery et al. (2000) considers a capacitated multiple hub location problem and presents a mixed integer linear programming formulation along with an heuristic algorithm using shortest paths. Determining the locations of hubs and allocation of non-hub nodes simulataneously, complicates the development of solution techniques for these problems.

de Camargo et al. (2008) presents a Benders decomposition algorithm based on a well known formulation to tackle the uncapacitated multiple allocation hub location problem. They were successful in solving large scale instances in reasonable time which were considered to be out of reach of other contemporary exact methods. The emergence of global supply chains has led to the sustained demand for intermodal transportation. Design of transportation network for intermodal logistics is more complex compared to single mode logistics. Arnold et al. (2002) formulated a linear 0-1 program and applied it to a rail/road transportation system in the Iberian Peninsula. It shows that the modal shares of the goods which have their origin or their destination in Iberia is very sensitive to the variation of the relative cost of rail. Ishfaq et al. (2011) develops a mathematical model using multiple-allocation *p*-hub median approach that encompasses the dynamics of individual modes through transportation, modal connectivity, and fixed location cost structures.

These models are driven by the hub-and-spoke network topology, based on single-level view of hub as consolidation and sortation centers. Montreuil (2011) asserts that the way physical objects are transported, handled, stored through these networks are unsustainable economically, environmentally and socially. He introduces a new paradigm called The Physical Internet (PI,

π) as a response to what he termed the *Global Logistics Sustainability Grand Challenge,* where the movement of freight proceed in a consolidated way through a series of carrier services and relay facilities. Crainic et al. (2016) adds to this by exploring the necessary links and synergy between Physical Internet and City Logistics and introduces the notion of Hyperconnected Logistics along with nine fundamental concepts offering a rich framework for designing sustainbale urban logistics. Montreuil et al. (2018) used Physical Internet's hyperconnectivity and modularity conceptual pillars to develop a logistic topology web based on multi-tier hierarchical space structuring and an interconnected multi-plane logistics mesh networks enabling future generation of parcel logistics hubs to be capable of supporting X-hours (ultimately X-mins) delivery services. Access hubs are transshipment and consolidation points located at the neighborhood level and can be allocated to single or multiple demand centers termed as unit zones.

Transition towards Hyperconnected Urban Logistics requires restructuring of the overall urban logistic network design and urban space clustering. This paper fills the gap in literature by proposing an optimization modelling framework that allows multiple allocation of hubs to unit zones to generate the potential set of access hub locations which can be fed to large scale network design optimization models to make it solvable in appropriate time given its more in-depth modelling. Furthermore, with the advent of more sustainable solutions to urban logistics challenge such as mobile access hubs, it provides a solid base to filter the overall possible location set to be fed into more evolved planning models like the one proposed in Faugère et al. (2020).

3 Problem Description and Methodology

3.1 Problem Description

A large parcel express company provides pickup and delivery services to customers residing in an urban agglomeration such as a megacity. This territory is divided into a set of unit zones constituting a demand area whose borders are drawn by the transportation infrastructure (e.g., streets, boulevards) and geographic constraints (e.g., coast, hills, parks, and lakes). Each unit zone is assigned with a courier, typically one who is proficient about its geography and customer base. These couriers are responsible for transshipment of parcels between his respective unit zone and the access hub(s) at its neighborhood. Due to the urban infrastructure, a vast number of sites can be identified for the role of access hubs in the neighborhood of each unit zone. The objective of the company is to design a potential access hub network from a large set of candidate locations using gross approximations of the local hubs in a priori estimations of costs so that it can enable network design and clustering optimization models to remain tractable in time while accounting for comprehensive inter-hub flow which in turn can meet specific service level targets. In the context of this paper, the following assumptions are made:

- Each unit zone can be assigned to multiple access hubs and each access hub can be assigned to multiple unit zones.
- Operations are planned to serve the entire demand of the megacity.
- Capacity at access hubs is adjusted by adding modules of fixed capacity and there is a fixed cost associated with adding a module.
- Deployment of hub at a location has an associated geo-specific cost which can be interpreted as land reservation costs and is independent of number of modules.
- Percentage imbalances in demand flow assignments to hubs is known for each unit zone.
- Rider cycle times to each unit zone is given.

3.2 Methodology

3.2.1 Allocation of locations to unit zones

Before solving the optimization model, we must establish the relationship between the unit zones and access hub candidate locations i.e.) for each access hub candidate we must determine the set of unit zones it can serve. Owing to the considerations in access hub capacity along with the goal of reducing cost of couriers incurred in traversing to unit zones to meet service level requirements, an access hub is confined to serving their immediate neighboring unit zones only. A simple method would be to fix a threshold distance from the candidate and assign it to those unit zones which are covered under the imaginary circle drawn with threshold distance as radius. However, practically speaking, an urban agglomeration such as a megacity consists of a diversity of unit zone shapes and sizes depending on what it represents. Hence, it is not surprising to see regions with dense collection of unit zones smaller in area (e.g., cluster of high-rise buildings) and regions with sparse collection of larger unit zones as depicted in Figure 2. Having a fixed threshold for all candidate locations in a megacity fails to account for diversity in unit zone shapes, sizes, and demand profile. This calls for a geo-specific threshold that accounts for demand density of unit zones in the neighborhood of each candidate location. Then by using the corresponding geo-specific threshold for each candidate location, we can determine the set of unit zones it is capable of serving given the considerations on hub capacity and service level requirements. Further, for each zone, we can determine the set of candidate locations capable of serving it.

For example, as shown in Figure 2, if we assume same threshold radius of $'R'$ from location shown in sky blue in both denser and sparser regions, then a hub deployed in that location would be capable of serving 18 unit zones in denser region compared to 7 unit zones in sparser region which might not be beneficial in terms of meeting hub capacity and service level requirements. However, if we assume a threshold radius of $'r'$ for the dense region indicated by yellow circle in Figure 2(a), a hub deployed at that location would be capable of serving 7 unit zones only as expected.

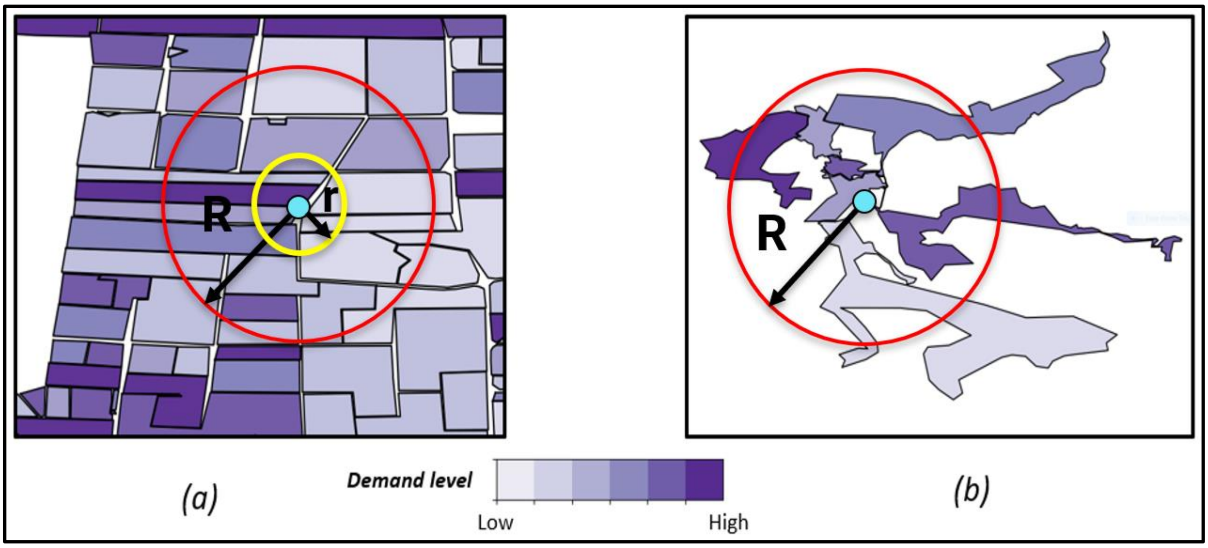


Figure 2: (a) Dense collection of unit zones smaller in area (b) Sparse collection of unit zones larger in area

3.2.2 Pickup and Delivery Routes

The starting point of the proposed modeling is the approximation of the vehicle routing problem via the route length estimation proposed by Daganzo (1984) and exploited in Faugère (2020). In case of couriers performing s stops in unit zone u from a hub, say h, the total distance travelled can be expressed as the combination of stem distance (from the originating hub to the area of service) and an in-tour distance (in the area of service) as follows:

$$D_u(s_u, h) = 2d_{hu}\frac{s_u}{Q_u^h} + s_u k(\delta_u)^{-\frac{1}{2}} \quad (1)$$

where d_{hu} is the average distance between the originating hub h and unit zone u, s_u is the number of stops to perform in unit zone u,Q_u^h is the length or routes of couriers operating in unit zone u from hub h and δ_u is the density of customer locations in unit zone u. k is a constant related to the distance metric used that can be computed by simulation, Daganzo (2005). Similarly, the total time required to perform courier routes can be expressed as follows:

$$T_u(s_u, h) = \frac{s_u}{Q_u^h}\left(t_{fixed}^{courier} + \frac{2d_{hu}}{v_o^{courier}}\right) + s_u\left(\frac{k(\delta_u)^{-\frac{1}{2}}}{v^{courier}} + t_{stop}^{courier}\right) + n_u v t_{handling}^{courier} \quad (2)$$

where $t_{fixed}^{courier}$, $t_{stop}^{courier}$ and $t_{handling}^{courier}$ are respectively the couriers' fixed time for each route, the stopping time at each customer location in the route and handling time per unit, $v_o^{courier}$ and $v^{courier}$ are respectively the couriers' stem and in-tour speed, and the average number of parcels handled per stop. Finally, the induced operations cost can be approximated as follows:

$$C_u(s_u, h) = \frac{s_u}{Q_u^h}\left(c_{fixed}^{courier} + 2d_{hu}c_o^{courier}\right) + s_u k(\delta_u)^{-\frac{1}{2}} c^{courier} + T_u(s_u, h)\, c_{wage}^{courier} \quad (3)$$

where $c_{fixed}^{courier}$, $c_o^{courier}$, $c^{courier}$ and $c_{wage}^{courier}$ are respectively the courier's fixed cost per route, variable cost on the stem part of a route, variable cost in-tour and variable wage per unit of time (e.g., $/hour).

3.2.3 Mathematical model

Indices:

h candidate location $h \in A$
u unit zone $u \in U$

Mathematical Sets:

U set of unit zones
A set of candidate locations
P set of {unit zone, candidate} pairs where $(u, h) \in$ P implies unit zone $u \in U$ can be served by access hub at location $h \in A$.

Parameters:

s_u total pickup and delivery demand of unit zone $u \in U$
m_u minimum number of hubs to be present at the assumed distance from unit zone $u \in U$
t_h maximum number of unit zones that can be served by hub at location $h \in A$
C_m fixed cost of adding a module
$C_u(s_u, h)$ operation cost estimates of courier (s) serving $u \in U$ from hub at $h \in A$
$C_d(h)$ fixed cost of deploying a hub at location $h \in A$
Q_u maximum of inbound or outbound parcels at unit zone $u \in U$ during rider cycle
α capacity of a module
g_u known % imbalance in demand assigned to hubs capable of serving unit zone $u \in U$,
$g_u = 0$ if $m_u = 1$

Decision Variables:

X_h 0-1 variable indicating if hub is deployed at location $h \in A$
Y_{uh} 0-1 variable indicating if hub at location $h \in A$ is chosen to serve $u \in U, \forall\, (u, h) \in P$
Z_h number of modules added at location $h \in A$

Model:

$$\min \ \left(\sum_{(u,h)\in P} C_u(s_u, h) * Y_{uh} + \sum_{h\in A} C_d * X_h + \sum_{h\in A} C_m * Z_h \right) \quad (4)$$

$$s.t. \quad \sum_{u\in U} Y_{uh} \leq t_h * X_h \qquad \forall \ h \in A \quad (5)$$

$$\sum_{h\in A} Y_{uh} \geq m_u \qquad \forall \ u \in U \quad (6)$$

$$\sum_{u:(u,h)\in P} \left(\frac{1}{m_u} + g_u\right) Q_u * Y_{uh} \leq \alpha * Z_h \quad \forall h \in A \quad (7)$$

$$Z_h \leq M X_h \qquad \forall h \in A \quad (8)$$

$0 \leq Y_{uh}, X_h \leq 1$ $\qquad X_h, Y_{uh}, Z_h$ *Integers* , M is a sufficiently large positive number

Objective function (4) aims at minimizing the sum of estimated costs induced by courier operations, deployment of hubs at candidate locations and deployment of modules at hubs. Constraint (5) enforces that a hub deployed at location $'h'$ can serve a maximum of t_h unit zones. Constraint (6) ensures that required minimum number of hubs are present at assumed distance from each unit zone $'u'$. Constraint (7) enforces that the number of modules added to hub at location $'h'$ is protected against demand flow imbalances and meet the respective capacity requirements during each rider cycle. Constraint (8) enforces that the modules are added only when a hub is deployed at location $'h'$.

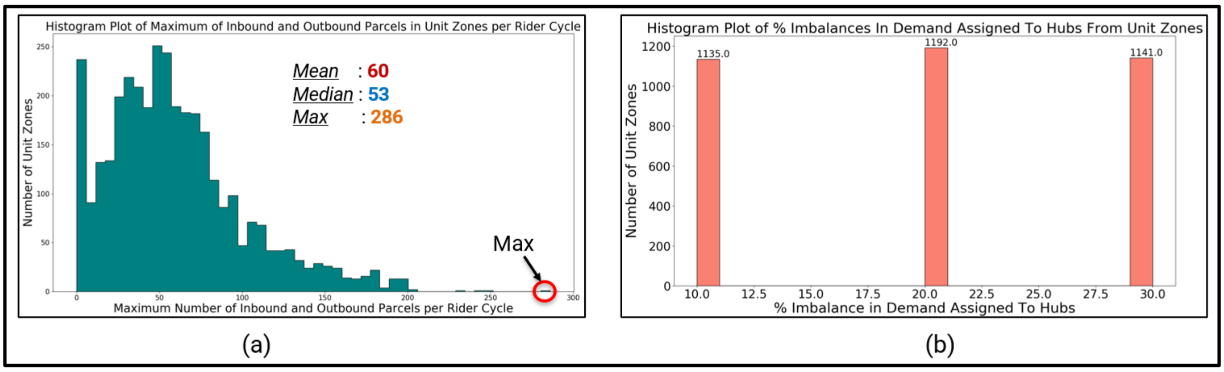


Figure 3: Histogram Plot of (a) Maximum of Inbound and Outbound Demand in Unit Zones per Rider Cycle (b) Percentage Imbalances in Unit Zone Demand Flow Assignments to Hubs for Illustrative Case

4 Results and discussions

In this section, we assess the potential of locations chosen by the model for access hub deployment through a set of experiments performed on an illustrative case inspired from a large parcel express carrier's operation. The experiments were implemented in Python 3.7 using Gurobi 9.0 as the solver and were computed on a laptop with an Intel(R) i7-9750H CPU @ 2.60GHz (Intel, Santa Clara, CA, USA).

The case is representative of a real-world Asian megacity with heterogenous demand profile. It consists of 3,468 unit zones and 52,148 candidate locations for access hub deployments. Figure 4 illustrates the considered megacity's geography and demand profile along with the set of initial candidate sites located either at the centroid or at the intersection of the unit zones. Demand averages between 0 to 4,704 parcels per unit zone with a total demand of 1.35 million parcels across the megacity. The rider cycle times are considered to be 2 hours for unit zones with maximum 2-hour inbound and outbound flow lesser than 200 parcels, 1 hour if maximum 2-hour demand is between 200 and 500 parcels and 30 minutes if it is above 500 parcels. Accordingly, maximum of inbound and outbound flow per rider cycle varies from 0 to 286 parcels per unit zone with an average of 60 parcels per unit zone per rider cycle as shown in Figure 3(a).

If the inbound (or) outbound flow of unit zone is split equally among assigned hubs, then it is a case of perfect balance. However, in real-world it is most likely that the flow is not balanced among assigned hubs to meet service level requirements, in which case hub has to be protected against this through additional capacity. Histogram plot of known percentage imbalances in unit zone demand assignments to hubs for the illustrative case is displayed in Figure 3(b), for example there are 1135 unit zones with 10% imbalance in demand assigned to hubs capable of serving them.

As discussed earlier, firstly we must determine the set of unit zones each candidate location is capable of serving. Setting geo-specific threshold distance from each location as min{50m, $\sqrt{Median\ Area\ of\ neighboring\ unit\ zones}$} found to be promising for our experiments. In future work, we can conduct sensitivity analyses to study the impact of different threshold distances on the total cost estimates. Capacity of access hubs are adjusted by adding modules of fixed capacity of 50 parcels (α) and cost of adding each module (C_m) is $50 irrespective of the hub location. Default values for parameters used in the model are listed in Table 1.

Table 1: Default Experiment parameters

Parameter	Value	Description	Parameter	Value	Description
$t_{fixed}^{courier}$	1 min	*Courier fixed time*	$c_o^{courier}$	$0.15/km	*Courier variable cost in stem*
$t_{stop}^{courier}$	0.5 min	*Courier stop time*	$c^{courier}$	$0.1/km	*Courier variable cost in tour*
$t_{handling}^{courier}$	0.05 min	*Courier handling time*	$c_{wage}^{courier}$	$10/h	*Courier hourly rate*
$c_{fixed}^{courier}$	$10/h	*Courier fixed cost*	$v_o^{courier}$	25 km/h	*Courier stem speed*
C_m	$50	*Cost per module*	α	50 parcels	*Capacity per module*
$\bar{Q}$	35	*Courier Capacity*	$v^{courier}$	8 km/h	*Courier in-tour speed*

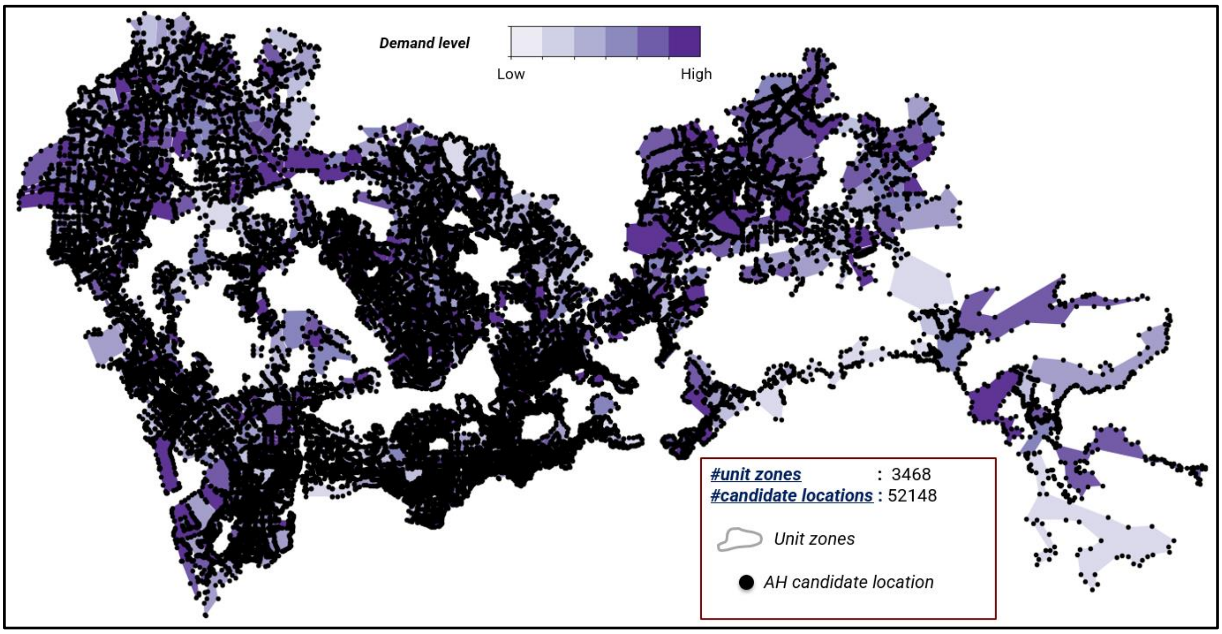


Figure 4: Initial set of candidate locations considered for access hub (AH) deployments

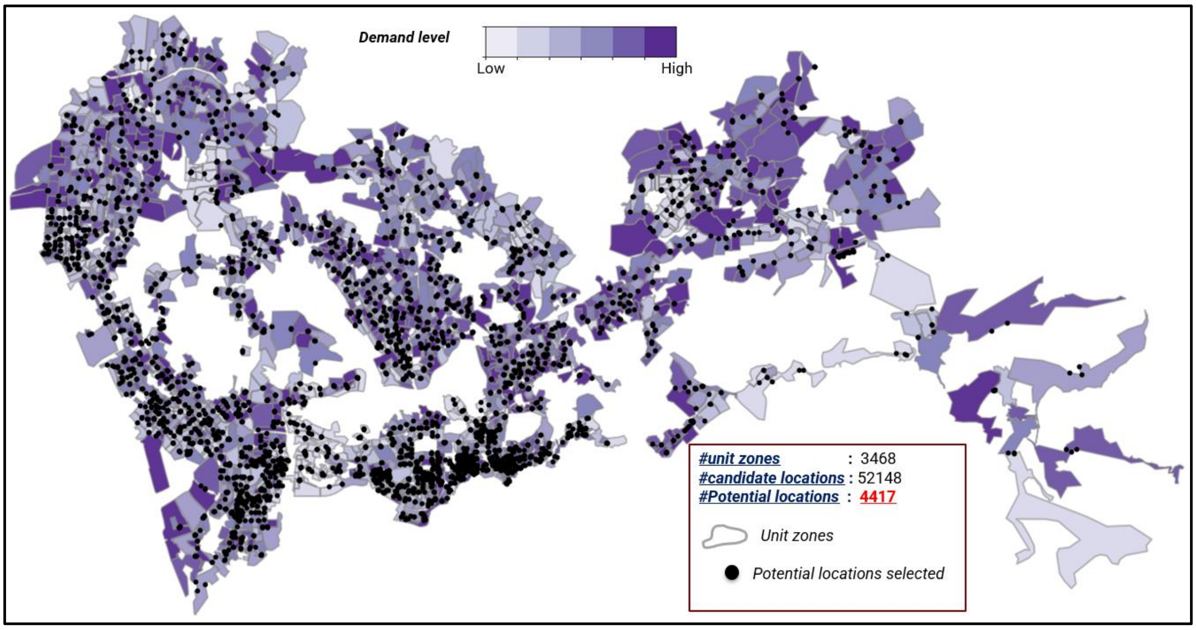


Figure 5: Potential locations selected by the model for the illustrative case with: $m_u = 3$, $t_h = 4$

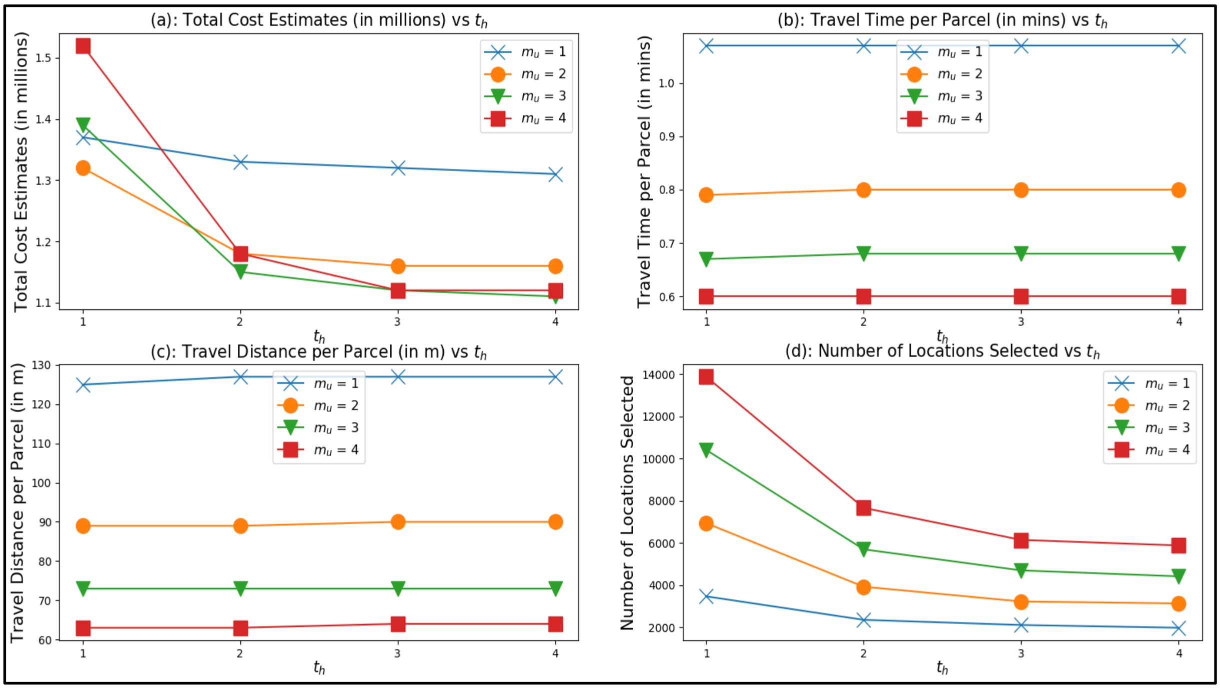


Figure 6: Several Performance Indicators vs t_h for different minimum number of hubs per unit zone (m_u) constraints

Output of the model provides a set of access hub locations to be activated, the set of zones expected to be served by each candidate location, the set of activated access hub locations from which a zone is expected to serve from, and the expected capacity to be deployed at each location to robustly be able to ensure adequate service.

Figure 6 provides the sensitivity analysis of several performance indicators against the maximum number of unit zones an access hub can serve (t_h) while enforcing different values for minimum number of hubs required near each unit zone (m_u) and solved to 3% optimality gap. First observation from Figure 6(a) is that, when access hubs are restricted to serve only one unit zone, as we increase the minimum number of hubs required per unit zone constraints from 1 to 2, total cost estimates reduces. This signifies that courier operational cost savings due to shared demand between 2 hubs for each unit zone is higher than the incurred extra deployment costs due to opening of new hubs and addition of modules to these hubs. However, as we increase m_u from 2 to 4, incurred deployment costs dominate over courier operation cost savings and the total cost estimates spikes up.

Secondly, when we allow access hubs to serve 2 unit zones, total cost estimates plummets for all m_u constraints. This can be ascribed to the fact that for most of the unit zones, the existing capacity in already deployed hub is used to fulfilling demand of one extra unit zone rather than deploying new hubs as depicted in Figure 6(d). For example, when we enforce at least 4 hubs near each unit zone, we see that the number of locations deployed drops from 13872 to 8000 as we allow access hubs to serve a maximum of 2 unit zones instead of 1.

Another general trend observed from Figure 6(a) is that as we make access hubs capable of serving more and more unit zones, total cost estimates drops significantly for all m_u constraints and the minimum total cost estimates is observed when we enforce atleast 3 hubs near each unit zone and each hub is capable of serving atmost 4 unit zones. 4417 locations were selected as potential locations in the case where minimal total cost estimates is obtained as depicted in Figure 5. This suggests that we should leverage on the capability of an access hub to serve more unit zones under the restrictions of its capacity to cut down on total cost estimates.

Figure 6(b) shows the travel distance per parcel against t_h and Figure 6(c) displays the distance travelled per parcel against t_h. For all t_h values, as we increase m_u, both travel time per parcel and distance travelled per parcel decreases because the inbound and outbound flow for each unit zone is split between the hubs, resulting in shorter courier routes as we leverage more hubs to serve unit zones.

5 Conclusion

This paper contributes to the literature by proposing an optimization model to design potential access hub networks, that need not be implemented as designed but can be fed into large scale network design and clustering models enabling them to be solved in appropriate time when accounting for more comprehensive inter-hub flow modelling. The framework suggests beginning by determining the set of unit zones that can be served by each candidate location. Then, we propose an optimization model using gross approximations of local hubs into priori estimates of cost and leveraging modular capacity deployments to find the set of potential locations that incurs minimum courier operations and deployment cost estimates while fulfilling the entire demand of all unit zones. These methods have been developed from realistic geometrical and operational approximations, making them readily applicable to different urban agglomerations.

It also assesses the potential of using such a solution in real-world context through an illustrative case inspired from a large parcel express carrier's operation. The analysis revealed that we should leverage on the capability of an access hub to serve more unit zones under its capacity restrictions to gain significant savings in operations and deployment cost estimates. Additionally, we can experiment with different values for threshold distance from hubs to study its impact on total cost estimates and other key indicators.

By specifying targets relative to the maximum number of unit zones that an access hub may serve, and the number of access hubs serving each zone, we can generate different sets of potential access hub locations. We can assess their computational solvability by feeding it to large scale design problem modelling comprehensive inter-hub flow and comparing its solving time and optimality gap with the baseline model starting with a large set of candidates as its input. In the context of mobile hubs, it can provide a set of potential locations from which actual locations can be selected for dynamic deployments of mobile hubs through framework proposed in Faugère et al. (2020).

References

- Alumur, S., & Kara, B. Y. (2008). Network hub location problems: The state of the art. *European journal of operational research*, 190(1), 1-21.
- Arnold, P., Peeters, D., & Thomas, I. (2004). Modelling a rail/road intermodal transportation system. *Transportation Research Part E: Logistics and Transportation Review*, *40*(3), 255-270.
- Aykin, T. (1994). Lagrangian relaxation based approaches to capacitated hub-and-spoke network design problem. *European Journal of Operational Research*, *79*(3), 501-523.
- Bruns, A., Klose, A., & Stähly, P. (2000). Restructuring of Swiss parcel delivery services. *OR-Spektrum*, *22*(2), 285-302.
- Campbell, J. F. (1992). Location and allocation for distribution systems with transshipments and transportion economies of scale. *Annals of operations research*, *40*(1), 77-99.
- Campbell, J. F. (1994). Integer programming formulations of discrete hub location problems. *European Journal of Operational Research*, *72*(2), 387-405.
- Campbell, J. F., & O'Kelly, M. E. (2012). Twenty-five years of hub location research. *Transportation Science*, *46*(2), 153-169.
- Crainic, T.G., & Montreuil, B. (2016). Physical internet enabled hyperconnected city logistics. *Transportation Research Procedia*, *12*, 383-398.
- da Graça Costa, M., Captivo, M. E., & Clímaco, J. (2008). Capacitated single allocation hub location problem—A bi-criteria approach. *Computers & Operations Research*, *35*(11), 3671-3695.
- Daganzo, C. F. (1984). The length of tours in zones of different shapes. *Transportation Research Part B: Methodological*, *18*(2), 135-145.
- Daganzo, C. F. (2005). *Logistics systems analysis*. Springer Science & Business Media.
- de Camargo, R. S., Miranda Jr, G., & Luna, H. P. (2008). Benders decomposition for the uncapacitated multiple allocation hub location problem. *Computers & Operations Research*, *35*(4), 1047-1064.
- Ebery, J., Krishnamoorthy, M., Ernst, A., & Boland, N. (2000). The capacitated multiple allocation hub location problem: Formulations and algorithms. *European journal of operational research*, *120*(3), 614-631.
- Ernst, A. T., & Krishnamoorthy, M. (1999). Solution algorithms for the capacitated single allocation hub location problem. *Annals of Operations Research*, *86*, 141-159.
- Faugère, L., White, C., & Montreuil, B. (2020). Mobile Access Hub Deployment for Urban Parcel Logistics. *Sustainability*, *12*(17), 7213..
- Gurram, G., and Fernandez, L.B. (2021 March 28) *Turning your retail space into distribution space*. SCMR, Available Online : https://www.scmr.com/article/turning_your_retail_space_into_distribution_space?platform=hootsuite&utm_campaign=JBLCampaign
- Ishfaq, R., & Sox, C. R. (2011). Hub location–allocation in intermodal logistic networks. *European Journal of Operational Research*, *210*(2), 213-230.
- O'kelly, M. E. (1986). The location of interacting hub facilities. *Transportation science*, *20*(2), 92-106.
- O'kelly, M.E. (1987). A quadratic integer program for the location of interacting hub facilities. *European journal of operational research*, *32*(3), 393-404.
- Montreuil, B. (2011). Toward a Physical Internet: meeting the global logistics sustainability grand challenge. *Logistics Research*, *3*(2), 71-87..
- Montreuil, B., Buckley, S., Faugere, L., Khir, R., & Derhami, S. (2018). Urban parcel logistics hub and network design: The impact of modularity and hyperconnectivity.

Reduction of Food wastage in Supermarkets due to Expiration

Keywords:food wastage, interconnected, supply chain, retail, sustainability, consolidation,distribution center optimization,dynamic pricing,FMCG,procurement

Apoorva Nuggehalli Srinivas*   Benoit Montreuil†

June 7, 2021

*Masters Student,ISYE,Georgia Institute of Technology, apoorvans@gatech.edu

†Professor and Coca-Cola Chair in Material handling and Distribution at Georgia Institute of Technology,benoit.montreuil@isye.gatech.edu

Contents

1 Abstract

Food disparity is one of the most widespread social problems, which can also be addressed as a supply chain problem. This paper focuses on methods to prevent wastage of food products due to expiry at the supermarket level, which contributes to almost 40% of the total food wastage in developed countries such as the US and Canada. With a literature review and a pertinent case study on the current methods in practice in supermarkets, the paper attempts to model an interconnected retail center which aggregates unsold products due for expiry in a region.The Interconnected retail center, termed **ReIntegrate**, helps in streamlining demand from the weaker economic strata and enables them to afford quality packaged and processed food.

In today's retail sector, where there is a constant hustle to maximize profitability, it is often ignored that food is beyond a replaceable retail commodity. Based on the increasing focus on the overlap and interaction between different supply chains, the model focuses on the interaction between the channels of distribution at the supermarkets. By efficient collaboration and coordination in terms of logistics and information sharing, different retail chains can coalesce their products due for expiry and find an ideal customer base.

Technology has been continually improving to allow for greater ease of inventory tracking, yet the demand volatility in the FMCG(Fast Moving Consumer Goods) sector renders the ideal algorithmic approach less efficient. By aggregating the products perceived as less desirable for consumption, and pricing based on perishability parameters, the proposed model accounts for the shift of customers towards delivery to home options, and preference to more affordable choices in the market.

We go on to discuss the different elements of the integrated retail distribution center's supply chain network, and its implications in terms of cost and operation. Modelled on the likes of a distribution center, the infrastructure lays a greater emphasis on categorizing and prioritization of picking to ensure minimum spoilage before reaching the customer . The paper also discusses other secondary methods such as efficient bar coding, on demand stocking for premium products and dynamic discounting guided by products' perishability and rate of degradation.

2 Introduction

2.1 Overview

This project was undertaken with the objective of identifying and recommending some methods of reducing food wastage, and making quality food affordable and accessible to all.

Food wastage can occur in many ways, beginning at the source due to conditions such as weather, improper storage conditions, and the changing demand patterns which require different variety of produce for sale/further processing.However, supermarkets play a major role in storage and distribution of a wide variety of food products, especially in developed and developing countries.

Apart from the commercial aspect , the sociological aspect is of considerable importance. With many sections of the society struggling to get affordable food, it is a pre-requisite to responsibly consume the food that is mass produced under strict quality control. Also with growing urbanization, the arable land area is shrinking, highlighting the importance of efficiency in production and storage.

This paper can be divided into three major sections. The first section discusses the problem of food wastage, the reasons for the same and its criticality.

Further,the second section highlights relevant literature review which highlights some relevant facts and figures and the current approaches to tackling the problem of reducing food wastage and preempting the same.

In the third section we will be highlighting potential ways to increase sales when the product is due for expiry and also diverting the products once the period of preempted sales have passed.The main focus solution of the paper is discussed in this section, the concept and organization of ReIntegrate Center and Food Transit Centers.

2.2 The Problem of Food Wastage

Statistics from [1] reveal that about 67 million tonnes of food is wasted everyday in India , which can be diverted to feed an entire state if used wisely. Similarly in the United States, 40 million tonnes is wasted everyday .Out of this, approximately 40% of the wastage is estimated to occur at supermarket level, and can be attributed to multiple factors, which can be efficiently preempted to eliminate food wastage.[2]. The following are the reasons why food wastage has a far more serious impact than other wastages:

- Loss of potential resource that has a societal impact and efficiently balancing the surplus with the growing demand can help improve distribution of affordable food.
- Production of food has a large carbon and water footprint due to the intensive processing.

2.2.1 Understanding the underlying reasons for Food Wastage

There are multiple reasons that contribute to food wastage, particularly at the supermarket level.The most predominant one is the demand uncertainity in the FMCG(fast moving consumer goods) industry.Customers also expect higher on shelf display quantities and high visual standards in terms of packaging which requires the supermarkets to carry higher inventory levels. Another common problem is the lack of differentiation between the various terminologies used to indicate expiry.The lack of standardisation between best before, date of expiry and other multitude of terminology used on products leads to disposal of products much in advance of its spoilage.

The advent of ecommerce has also benchmarked higher service levels for essentials and regularly consumed products and has made a larger array of products easily available. To compete with the same, supermarkets are pressured to match up service levels and options available for immediate purchase. The larger challenge is the expectation of lower price that has to be offered, which often pushes the stores to seek volume discounts and offer frequent promotions to lure customers.In the post pandemic world, the horizon has expanded with most supermarket chains having to offer deliveries to customer locations , complicating the logistic network and the associated costs.The customers are divided between having the luxury of shopping from the comfort of their homes, and the pleasure of making choices at the stores personally, and the supermarkets are under the pressure to satisfy both consumer bases.

2.2.2 Literature review

To gain a deeper understanding of the problem at hand and also be aware of the prevalent practices that have been formulated to maximize sales before expiry, an extensive literature review was conducted.

The main focus of search was to investigate the reasons for wastage and breakdown the problem to gain better understanding.The current best practices in industry to reduce supermarket level food wastage and other contemporary methods were used as a benchmark to help develop an innovative yet implementable solution.Major source of the established methods were research journals , supplemented by relevant news articles and management journals.

Among the majorly discussed causes,the misunderstanding of the terminology used in fitness of use of products has been attributed and discussed as a major cause of food wastage[6],[7].This highlights the need to standardize the terminology to be used to indicate the product's usability.

[8] , published by the FAO, summarizes the various causes of the food stages at the various stages of the food supply chain.Breaking down the problem into various stages helps efficiently tackle the problem and develop solutions for the same. Consider, the need for high on shelf storage, which can be efficiently tackled with an on demand stocking for premium products.

The large carbon and water footprint caused by food processing and the environmental impact is discussed in[15].This is of notable importance, as composting as a method of using expired foods to produce energy is less efficient in packaged foods due to the need to seperate the recyclable and non recyclable components, which makes it necessary to prevent wastage and also highlight importance of usage of sustainable packaging.Also, a greater emphasis can be placed on the foods that have a higher carbon foot print, like meat,dairy and fresh produce can diverted to discounted distribution as compared to the other less impactful products.

[17] provides a summary of the different practices that the retailers follow to deal with products due for expiry. These papers helped develop the base for a case study to investigate the retailer and consumer perspectives and analyze the different approaches and their associated challenges.

A noteworthy approach to minimizing domestic wastage of foods that have been purchased is discussed in [14] which alerts the user on the items that have to be used on priority to avoid wastage through a mobile application.

Many non profit organizations aim to redistribute "rescued food" among economically weaker sections of the society.One such major drives is the Meal Connect , which helps pledge donations directly as meals/grocery items, partnering with some mega food chains.[18]. The website helped us understand the current supply chain module followed by such non profit organizations to help reach out food to those in need.

2.2.3 Case Study

A detailed case study was conducted at Namdhari Fresh to gain a market perspective after understanding the academic and research work that was available.The brief summary of the findings are as follows:

Products that were most prone to wastage were dairy products, followed by the fresh produce.This was mostly due to the short shelf life of these products, though they had a much more stable demand as compared to the other products.It was noted that the store had a good shelving policy (First in First out), preempted the expiry of products atmost 15 days in advance of the indicated expiry dates. The ERP system used helped alert the store with all the products that were shortly due for expiry.Post alerting and initiating for disposal,return to Vendor was the most used channel of handling the products due for expiry, followed by Dynamic Discounting, explained in [13]. For farm produce, the products were composted on the company owned farms.

After gaining a preliminary understanding of the retailer's behaviours, customers were asked to complete a short survey to understand their treatment of products that are due for expiry.[3] discusses the behavior of different groups of customers to food that is closer to expiry, and the survey questionnaire was created on similar grounds. The respondents were well distributed among candidates of different age groups, genders and lifestyles. The results displayed a clear reluctance of candidates to purchase suboptimal food even at discounted prices and claimed that their decision would be governed by the category of product.

2.2.4 Current Practices and Shortcomings

After analysis of the results, some shortcomings were noted in the current models and practices, which can be noted as below:

- Lack of traceability of products post return to vendor and risk of reentry to market with repackaging which poses an ethical issue and also poses a potential health hazard.
- Inefficient activity based costing , for eg, though in the return to vendor policy, the store is benefitted by receiving new products,the cost of transportation is still borne by the manufacturer.
- Since food is an essential commodity with a high disparity of distribution, especially in developing countries, a greater weight has to be exerted on maximizing consumption before end of life than on profitability.
- Most dynamic discounting policies do not consider the impact of perishability and rate of spoilage, and are mostly driven by supplier promotion policies and margins.
- Diversion to the non profit organization is high non systemized, and this cannot be established as a demand channel, making return to vendor a safer option for the supermarkets.

3 Proposing Solutions

After a careful analysis of the current models and their shortcomings, we have proposed a major solution, the establishment and operation of a Food Web and other complementing practices to help consumption of product much in advance of its spoilage. A few things need to be established before implementation of complex networks:

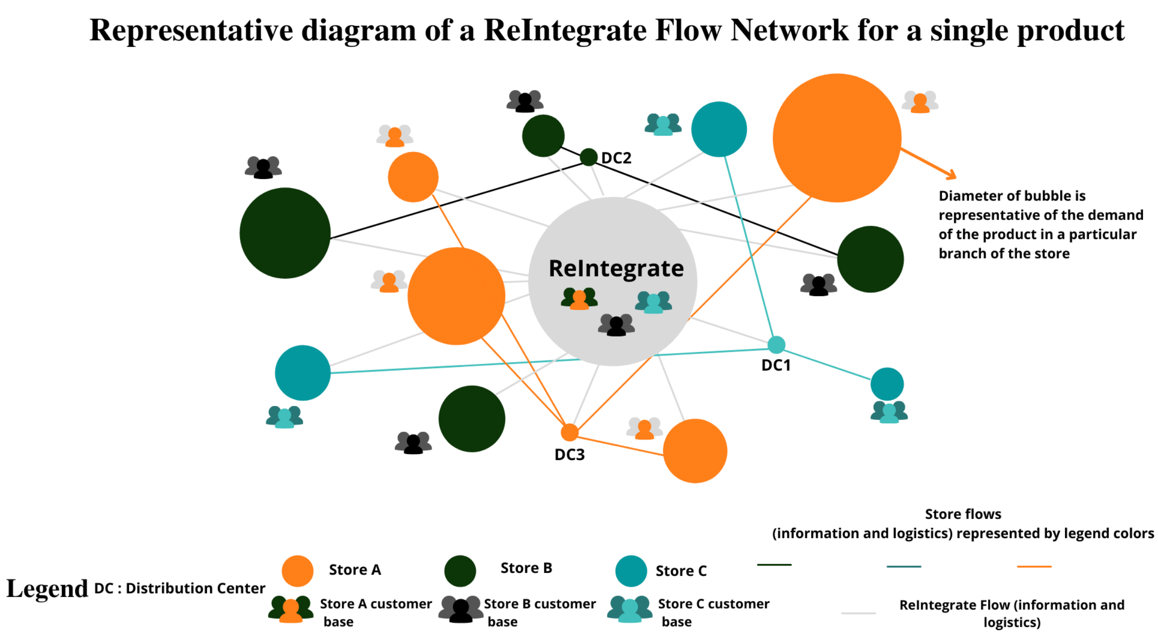
3.1 Prerequisites to Implementation of Food Web

A few important factors have to be set in place before the model is established.
The most important, perhaps, is the presence of a latent demand that exists and is damped by the factor of affordability, and the consumer capable of discerning that the discounting is not equivalent to a reduction in quality of product being sold.Following this , there is a need for clear packaging terminology. Given

the recommended storage conditions, when will the product be rendered unfit for consumption? 'Best Before' dates indicate the period in which the product is in the best condition, but does not indicate the degradation following the declared best before date.Another assumption of importance, as is validated by the survey conducted with consumers is that consumers are more conscious of brands and more loyal to brands than they are to supermarket chains.The last assumption, is perhaps, almost contradicting in nature. The costs involved in resale should be preceded with an analysis on the non quantifiable benefit of providing food to those in need.This should be the driving factor, and is ideally not limited in implementation by socioeconomic and legal policies in place in a particular region/sector. The next section clarifies and defines some terminology used in setting up the scenario.

3.2 Concept and Terminology of ReIntegrate and a Food Web

A **Food Web** can be defined in this context as a logistic network for material and information distribution of a food supply chain. The logistic network is augmented by the addition of a hub which acts as either a distribution center for sale or as a transit center. A brief diagrammatic representation can be denoted as below, where each supermarket chain is denoted by a different color and the size of demand of a particular product is represented by the size of the bubble.



Each set of supermarkets have their own set of internal networks and distribution centers through which goods are relayed based on the demand and minimum order and standard packing quantities.Alongside the distinct groups which have a preference to a particular store, there is among all these subsets with a latent demand for a particular product/category of products, which can be satisified by the ReIntegrate Center, which is defined as below.

The **ReIntegrate Center** is a term coined for a center to which the products which are due for expiry are redirected through aggregation to be redistributed through different distribution channels.

The major function of ReIntegrate can be characterised as below:

- **Demand Levelling**: With the aim of consolidating the latent/potential consumer base over an entire region, we are thereby levelling the demand, and utilising any trace demands that are unsatisfied due to economic factors, such as unaffordability.
- **Legal Stream of resale** : In reference to the earlier interviews conducted with the stores, there is a clear lack of transparency of the reverse flow of these products post / shortly before expiry. The possibility of repackaging and modifying the food before reselling is not an uncommon malpractice.
- **Streamlining supply to non profit organizations**: Many NGOs purchase the food that is not in good condition for retail sale.Streamlining and allocating the products to serve these channels is a major function of these centers.

3.3 Elements of a Food Web

- **Logistics Network**: The development of an optimal network comprising of facilities required to establish and distribute the items accumulated at the ReIntegrate center. This might include the vehicles,the drivers and establishing the network of transit centers and their associated routes.
- **Infrastructure**: Includes the internal infrastructure and facilities such as power, storage infrastructure and maintenance of a special enviroment to reduce the rate of food spoilage in the ReIntegrate Center.
- **Stores/Supermarket Chains** The participating stores are assumed to have an approximately normally distributed demand , with an average monthly sales, a variance component. In a more complex modelling scenario, seasonality and trend can also be accounted to accomodate for any demand volatility.

- **Customers** The general subset of a customers over a regional base of 2-5 mi remains constant and includes the customers who have a need for the category of the product, but make decisions governed by financial limitations.i.e, they will purchase the current product if available at a lower price, else opt for budget alternatives.

3.3.1 Cost Elements of a Food Web

- Inbound Costs : Includes all costs involved in inwarding products from the supermarkets,such as logistics expenses,quality checks and labour costs for storage.
- Outbound Costs : Includes all costs involved in outwarding the products, such as transportation expenses to customers,and the labour costs in picking and packing.
- Infrastructure costs - Fixed and Variable costs involved in establishment, maintenance and operation of the ReIntegrate Center.
- Administrative expenses and costs involved in information gathering and marketing expenses.

3.3.2 Comparing Costs to a Return to Vendor scenario

The purpose of Reintegrate is to divert costs from the return to vendor/reverse logistics to an extended redistribution chain. Typically , return to vendor consists of the following costs :[19]

- Time: The time elapsed from the intiation of pickup to collection and consolidation at retailer's warehouse can be effectively utilised in modeling a viable route for consolidation at Reintegrate center.
- Proximity and number of transits : The ReIntegrate center is in proximity of the supermarket chains,as oppposed to return to vendor, which undergoes several transits and handling, which might easily increase the rate of product spoilage.
- On the downside, return to vendor offers a higher return on products as compared to the discounted prices the products at ReIntegrate will sell at.

3.4 Channels of Distribution, Cost and Revenue Sharing

A complication of setting up the ReIntegrate center is identifying the entity in the Food Web to bear the costs of operation and share the revenue.Let us conceptually explore some models, which have varied stakeholders being accountable for different roles in the ReIntegrate Center.

The first channel, is to let the supermarkets in a given region invest in the establishment and operation of a Food Hub.Investments and revenues are established according to the share of products they agree to supply. For example, a supermarket might move 20% of it's milk cans when there are 5 days for delivery, and earns back the revenue based on its contribution to the total milk sale at the facility .It is ideal for a third party logistics vendor to be chosen to isolate the network from those of the supermarkets, and allow confidentiality in consolidation and information sharing.

The second channel, is to set the ReIntegrate center as an extended distribution center for ecommerce chains. Ecommerce companies have a larger share of the market in the current circumstances and an increased visibility and convenience factor , especially post the pandemic. Since the cost of operations of the chains is greatly reduced by the lack of a physical store, this can act as an incentive to obtain products at reduced rates.With greater visibility and emphasis on dynamic discounting, the latent demand can be reached out to with greater ease.

The final channel is involving NGO/non profit organizations. Most NGOs rely on stores / participating corporate partners to donate products based on their agreements, but might not satiate the required demand of consumers unable to afford quality food.By establishing a structure to estimate the demand and allocating the same among the products available at the Food Hub, the NGOs can immediately distribute the products for consumption well in time before the wastage.In this arrangement, the supermarkets can bear the operational and inwarding costs, as goodwill partners, and the outbound costs and administrative expenses can be managed by the NGOs.

3.5 Streamlining the NGO demand

Since a major portion of this paper and the model is aimed at distribution who those who could not afford the food otherwise, a great necessity is understanding the population and the demand figures and trends in this category. The population may include the underprivileged, the unemployed or those currently facing unfortunate circumstances. By understanding the demand of those who need these supplies and by filling it with the food that is unjustly wasted due to insufficient attention or corporate policies, we might be implementing the model with great efficiency.However, the greatest barriers in this case, is a reliable and continuously updated information stream . One of the possible methods to understand this demand is discussed in the following paragraphs.

One of the well tested methods,is to create an application, either mobile/web based for people to log in the kind of supplies and the estimated quantities that they would like over a given period of time.Considering some people might not have access to personal devices, most community service centers can use the application to gather information and this can be input to the ReIntegrate center database to divert quantities in the representative proportions to the NGOs/ regions.By doing so , each supermarket can make firm decisions on the product volumes for the particular time period that has to be forwarded to these demand hotspots.

3.6 Enhancing the operations

The effectiveness of these centers can be improved by using visual aids, and categorization.

- **Visual Aids**: Roughly based on the concept of Andon, items can be tagged to indicate the time left before expiry, and can play a major role in order pricking priorities.
- **Categorization**: The other major component in the fitness of use of an item is the category of item , which indicates perishability or rate of degradation of items. Wise use of these parameters can help tagging the products received from the supermarkets based on color coding / digitally readable encoding, which helps make the critical products a priority in order picking.

3.6.1 Challenges to Implementation

1. **Additional Investments required for Construction, Infrastructure and maintenance** : The most challenging part of implementing food webs is allocating investments for infrastructure, material handling and storage. For stores where return to vendor is a feasible option, the management may choose to not invest additionally in developing an external network.
2. **Logistics and Delivery from ReIntegrate Centers**: Post the phase of best usage of a product, the major costs incurred are due to logistics, and it is extremely crucial to optimize them by the use of good models and take care that the outbound costs do not have a major negative impact. It is a given that with the NGO distribution channel, this is a great possibility, in which case care should be taken to maximize outreach before spoilage.
3. **Lack of continuous flow of goods which justify employment/ facility utilisation**: Since this facility is mainly intended to route the flow of products due for expiry , the utilisation of the facility and labour tends to be highly erratic. By complementing the facility with another firm requiring stable labour, the uncertainity can be minimized.
4. **Pricing Issues** The focus must not be lost on the fact that what differentiates these hubs from the regular sales of products at the store is the pricing. The commitment to continually provide goods at a rate that is affordable to the lower economic strata of the society must not be deviated upon due to the obvious profit motive.

3.7 Alternative solution of focus: Use of ReIntegrate Centers as Transit Centers

In the solution discussed above, there are some decisions, such as the selection of appropriate stakeholders and involvement of third party logistics in transportation and last mile delivery. An alternative solution is to reduce the scale and scope of operation. With the ReIntegrate being maintained by the supermarket chains,it will act only as a temporary consolidation center for items due for expiry.This efficiently reduces infrastructure costs, and also eliminates labour costs, as packaging and sorting could be handled by the logistic partners.

3.8 Extended Barcoding / QR codes to Automate tracking

An important aspect of reducing wastage is to simplify the process of tracking the products due for expiry.[14] uses a barcode scanning to help consumers prioritize items in the fridge, and the same principle can be extended to stores. Manually tracking products due for expiry on a regular basis is highly labour intensive, especially in stores with higher stock.

Some advantages are that this method effectively groups together items that are due for expiry in a similar period of time and eases application of dynamic discounting when possible.It also prevents expired products from being sent back into circulation with tampered packaging.Bottleneck and slow moving products can be identified without manual intervention and shelving can be adjusted in accordance to sales and monitor success rate of the FIFO rule.
In the survey, it was pointed out that consumers often picked up newer products even if the older ones were available, which places an emphasis on reducing on shelf quantities and using QR codes for sorting to improve effectiveness.

3.9 Category Sensitive Discounting

In most merchandising stores, discounting is based on arbitrary values based on limits set by vendors. By categorizing the items that are to be discounted and are due for expiry, price sensitive consumers can be benefited, and wastage can be effectively reduced.Below is a table of some arbitrary factors contributing to wastage and some sample values and weights for the same.

Let us now try and inspect some parameter to understand the dependency and the relationship between these factors and the perishability of a product.

3.9.1 Parameters

t : days to expiry
q : Units on hand in stock, from a given batch due for expiry
s : Average shelf life
d : Percentage discounting
p: Profit margin factor (the desired percentage of profit, as a worst case scenario, as a fraction of the total

Table2.1:Arbitrary factors for average shelf life

Average Shelf Life	Weightage
1-2 days	10
2-10 days	8
10-20 days	6
> 20 days	4

Table2.2:Arbitrary factors for days to expiry

Days to Expiry	Weightage
1-2 days	10
2-10 days	8
10-20 days	4
> 20 days	2

Table2.3 :Arbitrary weightage factors for stock on hand from batch

Stock on Hand from Batch	Weightage
20-30 units	10
10-20 units	8
5-10 units	4
0-5 units	2

product price)

We can deduce from the above factors, that lesser the days to expiry, greater should be the emphasis placed on selling the product out. Similary, greater the remaining products in a particular batch, greater should be the discounting , as a larger volume of similar products have to sold out in a short period of time. The average shelf life indicates the rate of perishability of a product, and the emphasis can be placed on products that spoil faster, and need to be sold much in advance of their spoilage to be safe for consumption.

3.9.2 Advantages

- The arbitrary scale can be adjusted to suit consumers, products and supplier policies. Since the current dynamic discounting in the industry is chiefly governed by supplier profit margins and other lucrative schemes, placing an emphasis on sensitivity of environment should be given considerable importance while formulating discounting.
- In most regions with price sensitive consumers, the sale of products due for expiry can be maximised and food can be made affordable in exchange for placing condition on immediate consumption.

3.9.3 Challenges to application

- The company might find higher profits with return to vendor
- Excessive discounting may reduce perceived quality of products
- Dynamic discounting,at an increasing rate to maximize sale might not be acceptable to vendors
- Consumers may choose to purchase the products under the discounting scheme a day in advance of expiry, thereby drastically reducing profits for the business

3.10 On Demand stocking for Premium Products

Another feasible plan is to order premium products on a demand basis, so that stocking of products with high novelty, but low market demand can be avoided. As it can be noticed, in most stores, the highly priced products/ luxury food items often have a lower rate of consumption, but this does not give an opportunity to eliminate stocking due to the higher profit margins on these products. By creating a platform for reserving premium products and transporting from the DC on a case to case basis, the wastage can be reduced by levelling the demand.

4 Conclusion

This project briefly touches upon current best practices and recommended methods to reduce food wastage due to expiry at supermarkets. A greater transparency in disposal and product life cycle management is necessary to identify the bottlenecks in implementation of the ReIntegrate network.A conscious effort by the consumers and businesses can contribute greatly to a successful implementation of food salvaging models. A lot of papers are being published in this field, and one honorary mention is the proposal of biodegradeable packaging to help improve recyclability of items that are wasted inspite of best efforts. Hoping this paper will act as a starting point to continuing research and application models in this field and affordable &safe food is accessible to all sections of the society.

References

[1] Harshita Tyagi, How Food waste affects the economy *https://www.timesnownews.com/business-economy/industry/article/world-food-day-2020-how-food-waste-affects-the-economy/667896*

[2] Food Waste in America *https://www.rts.com/resources/guides/food-waste-america/*

[3] Consumer Behaviour towards price reduced Sub Optimal Foods *Jessica Aschemann-Witzel, Jacob Haagen Jensen, Mette Hyldetoft Jensen, Viktorija Kulikovskaja, Consumer behaviour towards price-reduced suboptimal foods in the supermarket and the relation to food waste in households, Appetite, Volume 116, 2017, Pages 246-258, ISSN 0195-6663, https://doi.org/10.1016/j.appet.2017.05.013. (http://www.sciencedirect.com/science/article/pii/S0195666316306687)*

[4] Food Perishability *https://steemit.com/food/@talha96/introduction-to-classification-of-food*
Geremy Farr-Wharton Marcus Foth Jaz Hee-Jeong Choi
https://onlinelibrary.wiley.com/doi/abs/10.1002/cb.1488

[5] *Montreuil, Benoit, "Omnichannel Business–to–Consumer Logistics and Supply Chains: Towards Hyperconnected Networks and Facilities" (2016). 14th IMHRC Proceedings (Karlsruhe, Germany – 2016). 19.*

[6] How food expiration dates contribute to food waste *https://www.foodindustry.com/articles/how-food-expiration-dates-contribute-food-waste/*

[7] Expiry Dates, Consumer Behavior, and Food Waste: How Would Italian Consumers React If There Were No Longer "Best Before" Labels?
https://www.mdpi.com/2071-1050/11/23/6821/pdf

[8] Causes and prevention of food losses and waste
http://www.fao.org/docrep/014/mb060e/mb060e03.pdf

[9] Survey conducted for Phase 1
Questionnaire Link for Survey:
https://docs.google.com/forms/d/18cDcirDw98xnrDVvnRERHW-AXEgUPP3lHLMxtRrSRk/prefill

[10] About Us *https://www.namdharifresh.com/about*

[11] Wikipedia Page : *https://en.wikipedia.org/wiki/Namdhari%27s_Fresh*

[12] *Consumer behaviour towards price-reduced suboptimal foods in the supermarket and the relation to food waste in households- Jessica Aschemann-Witzela,Jacob Haagen,Jensen,Mette HyldetoftJensen,Viktorija Kulikovskajaa*
https://www.sciencedirect.com/science/article/abs/pii/S0195666316306687

[13] *Liu, X., Tang, O. and Huang, P. (2008), "Dynamic pricing and ordering decision for the perishable food of the supermarket using RFID technology", Asia Pacific Journal of Marketing and Logistics, Vol. 20 No. 1, pp. 7-22. https://doi.org/10.1108/13555850810844841*

[14] *José Rouillard. The Pervasive Fridge. A smart computer system against uneaten food loss. Seventh International Conference on Systems (ICONS2012), Feb 2012, Saint-Gilles, Réunion. pp. 135-140. ffhal-00825886f ,https://hal.archives-ouvertes.fr/hal-00825886/document*

[15] *Environmental aspects of food wastage in trade – a case study Beata Bilska,Marzena Tomaszewska,Danuta Kołożyn-Krajewska,Krystian Szczepański,Robert Łaba,Sylwia ŁabaPublished Online 29 Jun 2020 DOI: https://doi.org/10.2478/oszn-2020-0009*

[16] *No time to waste: assessing the performance of food waste prevention actions Author links open overlay panel ValeriaDe LaurentiisaCarlaCaldeiraaSerenellaSalaa https://doi.org/10.1016/j.resconrec.2020.104946 https://www.sciencedirect.com/science/article/pii/S0921344920302640*

[17] *Food waste management: A review of retailers' business practices and their implications for sustainable value Iona YueluHuangaLouiseManning-bKaty L.JamescVasilisGrigoriadisdAnthonyMillingtonaVictoriaWoodaShaneWarde https://www.sciencedirect.com/science/article/pii/S095965262035530X*

[18] *https://mealconnect.org/*

[19] *https://qualitycorrections.com/2018/07/26/return-to-vendor/*

Optimising Product Swaps in Urban Retail Networks

Joyce Zhang and Russell G. Thompson
The University of Melbourne, Parkville, Australia
Corresponding author: rgthom@unimelb.edu.au

Abstract: *The Physical Internet (PI) and City Logistics is based on trying to achieve higher levels of consolidation on vehicles that typically requires exchanging loads at intermediate locations. A common problem in urban areas in swapping goods between retail stores where there is a small amount of goods moving between individual stores to satisfy customer requirements where there are stock shortages at some locations. This type of network is also common for deliveries between local post offices or B2B networks particularly with parcel lockers. Such networks are characterised by having multiple common origins and destinations requiring services operating from many to many locations. A number of performance measures can be considered for such networks including number of vehicles operated and distance travelled, service levels/reliability for customers, network efficiency, vehicle load factors and the number of times tasks/consignments are transferred between vehicles.*

This paper describes how networks for exchanging goods between stores can be designed using multi-objective optimisation modelling. A mathematical program has been formulated to include multiple objectives, namely vehicle operating costs, vehicle usage (number of vehicles used), labour costs (proportional to working time) and unreliability costs. Constraints considered include, vehicle capacity, unloading dock capacity and storage capacity. Decision variables are the vehicle routes with loads as well as waiting times at nodes. Coordination of transfers at customer considers vehicle-to-vehicle parcel transfer. Pareto optimal solutions for a small network are presented. A discussion of various solution procedures will be outlined.

Conference Topic(s): *PI network design*

Keywords: *Cross-Docking, Logistics Networks, Multi-objective optimisation, Transshipment*

1 Introduction

Low utilization of urban freight vehicles is contributing to rising costs of urban distribution and increasing levels of urban congestion that is leading to a deterioration in sustainability in many cities. E-commerce is leading to new distribution models with hybrid networks involving deliveries from stores as well as warehouses to homes becoming popular (Arslan et al 2020). However, transfers of goods between stores can provide cost savings and efficiency gains where there is excess stock at some stores and shortages at others. B2B parcel networks have discrete origins such as courier depots and destinations such as parcel lockers (Pan et al, 2021).

There is a growing need to transfer inventory between stores or outlets within retail chains in large metropolitan areas or regions due to growing range of specialty products becoming available, increasing pressure to maintain low stock levels at stores and retailers only a limited number of warehouses within metropolitan areas. Such networks have a discrete number of origins and destinations and a frequent demand for goods to be transferred between them. Similar networks exist for transporting books between libraries within local areas, medicine between hospitals within health care organisations and consignments between urban consolidation centres or hubs in large metropolitan areas.

City Logistics initiatives aim to reduce the economic, social and environmental costs associated with urban freight (Taniguchi and Thompson, 2015). Hyperconnected city logistics based on Physical Internet concepts provides a practical means of integrating urban freight transport systems (Crainic and Montreuil, 2016).

Increasing consolidation levels in vehicles was recognised as the key to achieving sustainable urban goods transport (OECD, 2003). Improved vehicle loading or tonnes moved by km driven has been identified is a key driver of emissions decrease in logistical field (ITF, 2018). This involves increasing the use of available capacity in vehicles and reducing the overall km driven by vehicles while delivering the same amount of goods.

2 Methodology

2.1 Problem description

The problem considers product swaps among retailor shops in an urban network. The aim is to determine the optimal vehicle routes for exchanging goods and identify the best locations for the goods transhipment for a given demand. The left plot of Figure 1 shows an example of a network with five retailor shops when product transhipment is not considered. To deliver 20 delivery tasks from shop $i \in \{1,2,3,4,5\}$ to shop $j \in \{1,2,3,4,5\} \backslash \{i\}$, five vehicles are employed, each from one shop, to deliver products to the other shops. This results in a long total travel distance as well as a lot of travel with empty/low load. On the other hand, if products can be transhipped and exchanged at some shops, fewer vehicles may be needed, and the total travel distance and load factor can be improved. The right plot of Figure 1 gives an example of the network when shop 5 is considered as an exchange point. Products to other nodes can be transhipped at shop 5.

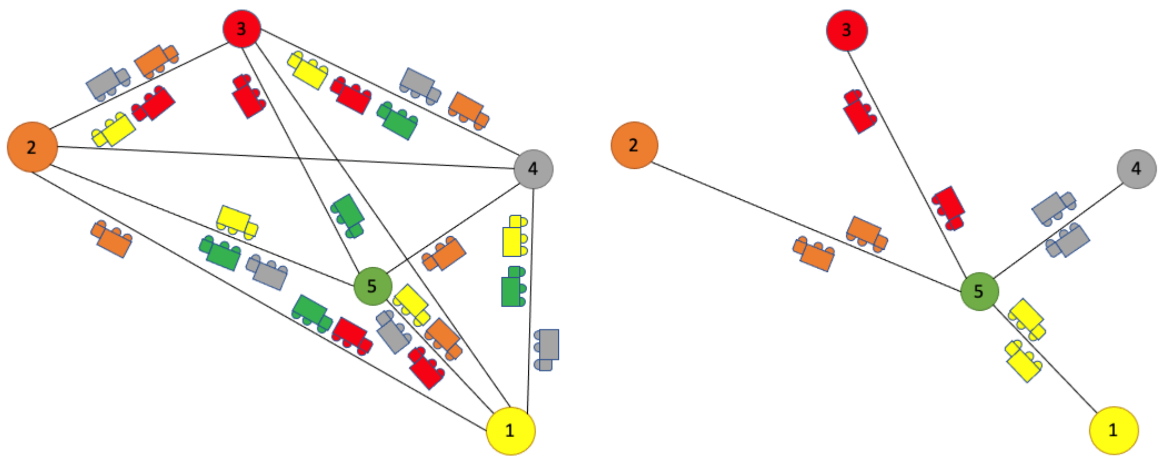


Figure 1: Illustrating product swap networks without (Left) and with (Right) product transhipment.

The objective is to minimise the total delivery cost as well as the reliability of the network, which is measured by the number of product transfers. In the next subsection, we develop a model to determine the vehicle routes and the location of exchange point(s). Key constraints such as vehicle capacity and vehicle coordination (a.k.a. vehicle synchronisation, which requires two vehicles present at the loading area of a shop at the same when transhipment occurs) are taken into account. In addition, vehicles need to occupy loading docks when transhipping goods, and so the dock availability constraint is considered.

2.2 Mathematical model

We consider a set of shops (i.e., nodes) $\mathcal{M}$ in the urban retail network with a set of delivery tasks $\mathcal{N}$; each task n is characterised by its origin and destination nodes (o_n, d_n) and size s_n. There is a set of vehicles $\mathcal{V}$ available to perform delivery; each vehicle is associated with a node, its depot (i.e., start node) f_v, and a capacity c_v. The model needs to decide when a vehicle v is to start from its depot, i.e., $x_v \geq 0$, which route the vehicle takes, which is represented by binary variable $y_{v,ij}$, and what products it picks up/drops off. Binary variable $z_{v,n,ij}$ indicates whether a task n is carried on by vehicle v on path $i \to j$, $i, j \in \mathcal{M}$. Vehicle waiting is allowed at nodes, which is controlled by decision variable $\omega_{v,i} \geq 0$. See Table 1 for the variable notation.

Table 1: Notation for variables

Decision variables	
x_v	Start time of vehicle $v \in \mathcal{V}$ from its depot, $x_v \geq 0, \forall v$
$y_{v,ij}$	$= 1$ if vehicle v traverses route ij; $= 0$ otherwise. $y_{v,ii} = 0$.
$z_{v,n,ij}$	$= 1$ if vehicle v transports task $n \in \mathcal{N}$ through ij; $= 0$ otherwise. $z_{v,n,ii} = 0$.
$\omega_{v,i}$	Additional waiting time of vehicle v at node $i \in \mathcal{M}$
Auxiliary variables	
$\mu_{v,i}$	$= 1$ if vehicle v unloads at node $i \in \mathcal{M}$; $= 0$ otherwise.
$\lambda_{v,i}$	$= 1$ if vehicle v loads at node $i \in \mathcal{M}$; $= 0$ otherwise.
$p\lambda_{v,n,i}$	$= 1$ if vehicle v picks up task $n \in \mathcal{N}$ at node $i \in \mathcal{M}$; $= 0$ otherwise.
$p\mu_{v,n,i}$	$= 1$ if vehicle v drops off task $n \in \mathcal{N}$ at node $i \in \mathcal{M}$; $= 0$ otherwise.
$D_{v,i}$	Departure time of vehicle v at node $i \in \mathcal{M}$
$A_{v,i}$	Arrival time of vehicle v at node $i \in \mathcal{M}$

$PD_{n,i}$	Departure time of task $n \in \mathcal{N}$ at node $i \in \mathcal{M}$
$PA_{n,i}$	Arrival/Available time of task $n \in \mathcal{N}$ at node $i \in \mathcal{M}$

The objective function consists of three parts: (i) vehicle operating cost (VOC), (ii) labour cost (LC), and (iii) unreliability cost. The VOC is composed of the fixed vehicle cost and the travel cost, which correspond to the two terms in Equation (1).

$$F_{voc} = VC_f \sum_{v\in\mathcal{V}} \sum_{i\in\mathcal{M}} y_{v,if_v} + VC_d \times V \sum_{v\in\mathcal{V}} \sum_{i\in\mathcal{M}} \sum_{j\in\mathcal{M}} y_{v,ij} T_{ij} \tag{1}$$

where VC_f is the fixed rate for employing a vehicle, VC_d is the cost per km, V is the average travel speed, and T_{ij} is the travel time from i to j.

The labour cost includes the salary of vehicle drivers and that of staff working at unloading docks for loading/unloading goods. It is assumed that the salary is proportional to working time. The working time for driving vehicle v is the difference in time between the arrival at and departure from its depot, i.e., $A_{v,f_v} - x_v$. Let binary variables $\mu_{v,i}$ and $\lambda_{v,i}$ indicate whether vehicle v unloads and loads any product at node i. Given the fixed unloading and loading times, U_i and L_i, the working time at node i is $\sum_v (\mu_{v,i} U_i + \lambda_{v,i} L_i)$. So,

$$F_{lc} = LC_d \sum_{v\in\mathcal{V}} (A_{v,f_v} - x_v) + LC_f \sum_{v\in\mathcal{V}} \sum_{i\in\mathcal{M}} (\mu_{v,i} U_i + \lambda_{v,i} L_i). \tag{2}$$

Here LC_d and LC_f are the salary rates for drivers and shop staff respectively.

Generally speaking, the system is more reliable when there are fewer product transfers. Hence the unreliability is measured by the number of transfers made. We let binary variables $p\mu_{v,n,i}$ and $p\lambda_{v,n,i}$ denote if task n is unloaded and loaded, respectively, by vehicle v at node i. Each task needs to be loaded and unloaded at least once. The right-hand side of (3) is the difference between the total number of loading and unloading and $2|\mathcal{N}|$ divided by 2, which gives the additional loading-unloading times, i.e., the number of transfers.

$$F_{ur} = \frac{1}{2} \sum_{i\in\mathcal{M}} \sum_{v\in\mathcal{V}} \sum_{n\in\mathcal{N}} (p\mu_{v,n,i} + p\lambda_{v,n,i}) - |\mathcal{N}| \tag{3}$$

The mathematical formulation of the problem is provided below.

$$\min_{\mathbf{x},\mathbf{y},\mathbf{z},\boldsymbol{\omega}} (F_{\$}, F_{ur}) \text{ s.t.} \tag{4}$$

$$\sum_{j\in\mathcal{M}} y_{v,ij} \le 1 \qquad \forall i \in \mathcal{M}, v \in \mathcal{V} \tag{5}$$

$$\sum_{j\in\mathcal{M}\setminus\{i\}} y_{v,ij} = \sum_{k\in\mathcal{M}\setminus\{i\}} y_{v,ki} \qquad \forall i \in \mathcal{M}, v \in \mathcal{V} \tag{6}$$

$$\sum_{v\in\mathcal{V}} \sum_{j\in\mathcal{M}} z_{v,n,o_n j} = 1 \qquad \forall n \in \mathcal{N} \tag{7}$$

$$\sum_{v\in\mathcal{V}} \sum_{i\in\mathcal{M}} z_{v,n,io_n} = 0 \qquad \forall n \in \mathcal{N} \tag{8}$$

$$\sum_{v\in\mathcal{V}} \sum_{j\in\mathcal{M}} z_{v,n,id_n} = 1 \qquad \forall n \in \mathcal{N} \tag{9}$$

$$\sum_{v\in\mathcal{V}} \sum_{i\in\mathcal{M}} z_{z_{v,n,d_n j}} = 0 \qquad \forall n \in \mathcal{N} \tag{10}$$

$$\sum_{v\in\mathcal{V}} \sum_{j\in\mathcal{M}} z_{v,n,ij} \le 1 \qquad \forall i \in \mathcal{M}\setminus\{o_n, d_n\}, n \in \mathcal{N} \tag{11}$$

$$\sum_{v\in\mathcal{V}} \sum_{j\in\mathcal{M}} z_{v,n,ij} = \sum_{v\in\mathcal{V}} \sum_{k\in\mathcal{M}} z_{v,n,ki} \qquad \forall i \in \mathcal{M}\setminus\{o_n, d_n\}, n \in \mathcal{N} \tag{12}$$

$$z_{v,n,ij} \le y_{v,ij} \qquad \forall i,j \in \mathcal{M}, n \in \mathcal{N}, v \in \mathcal{V} \tag{12}$$

$$p\mu_{v,n,i} \ge \sum_{k\in\mathcal{M}} z_{v,n,ki} - \sum_{j\in\mathcal{M}} z_{v,n,ij} \qquad \forall i \in \mathcal{M}, n \in \mathcal{N}, v \in \mathcal{V} \tag{13}$$

$$p\lambda_{v,n,i} \ge \sum_{j\in\mathcal{M}} z_{v,n,ij} - \sum_{k\in\mathcal{M}} z_{v,n,ki} \qquad \forall i \in \mathcal{M}, n \in \mathcal{N}, v \in \mathcal{V} \tag{14}$$

$$M\mu_{v,i} \geq \sum_{n\in\mathcal{N}} p\mu_{v,n,i} - 1 \quad \forall i \in \mathcal{M}, v \in \mathcal{V} \tag{15}$$
$$M\lambda_{v,i} \geq \sum_{n\in\mathcal{N}} p\lambda_{v,n,i} - 1 \quad \forall i \in \mathcal{M}, v \in \mathcal{V} \tag{16}$$
$$A_{v,i} = \sum_{j\in\mathcal{M}} y_{v,ji}\big(D_{v,j} + T_{ij}\big) \quad \forall i \in \mathcal{M}, v \in \mathcal{V} \tag{17}$$
$$D_{v,f_v} = x_v + \lambda_{v,f_v} L_{f_v} + \omega_{v,f_v} \quad \forall\, v \in \mathcal{V} \tag{18}$$
$$D_{v,i} = A_{v,i} + \mu_{v,i} U_i + \lambda_{v,i} L_i + \omega_{v,i} \quad \forall i \in \mathcal{M}\backslash\{f_v\}, v \in \mathcal{V} \tag{19}$$
$$A_{v,f_v} \leq T \quad \forall v \in \mathcal{V} \tag{20}$$
$$\omega_{v,i} \leq w \quad \forall i \in \mathcal{M}, v \in \mathcal{V} \tag{21}$$
$$PA_{n,i} = \sum_{v\in\mathcal{V}} \sum_{k\in\mathcal{M}} z_{v,n,ki}(A_{v,i} + U_i) \quad \forall i \in \mathcal{M}, n \in \mathcal{N} \tag{22}$$
$$PD_{n,i} = \sum_{v\in\mathcal{V}} \sum_{j\in\mathcal{M}} z_{v,n,ij}(D_{v,i} - \omega_{v,i} - L_i) \ \forall i \in \mathcal{M}, n \in \mathcal{N} \tag{23}$$
$$PA_{n,i} \leq PD_{n,i} \quad \forall i \in \mathcal{M}\backslash\{o_n, d_n\}, n \in \mathcal{N} \tag{24}$$
$$PD_{n,i} - D_{v,i} \leq \big(1 - p\mu_{v,n,i}\big)M \quad \forall i \in \mathcal{M}, n \in \mathcal{N}, v \in \mathcal{V} \tag{25}$$
$$PA_{n,i} - D_{v,i} \leq \big(1 - z_{v,n,ij}\big)M \quad \forall i,j \in \mathcal{M}, n \in \mathcal{N}, v \in \mathcal{V} \tag{26}$$
$$\sum_{n\in\mathcal{N}} s_n z_{v,n,ij} \leq c_v \quad \forall v \in \mathcal{V} \tag{27}$$
$$\max_{v\in\mathcal{V}} \sum_{u\in\mathcal{V}\backslash\{v\}} \mathbf{1}\big(A_{u,i} < D_{v,i} \& A_{v,i} < D_{u,i}\big) + 1 \leq l_i \quad \forall i \in \mathcal{M} \tag{28}$$
$$x_v, \omega_{v,i} \geq 0 \quad \forall i \in \mathcal{M}, v \in \mathcal{V} \tag{29}$$
$$y_{v,ij}, z_{v,n,ij} \in \{0,1\} \quad \forall i,j \in \mathcal{M}, n \in \mathcal{N}, v \in \mathcal{V} \tag{30}$$

The multi-objective function is given in (4), where F_1, F_2 are converted to the total financial cost $F_{\$}$. For vehicle routing, Constraint (5) ensures that each vehicle visits a node at most once, and (6) is the flow balance constraint. For performing delivery tasks, Constraints (7)-(10) guarantee that goods of each task are picked up from its origin and dropped off at its destination. For other nodes, Constraint (11) ensures that the goods can visit each of them at most once, and the product flow balance constraint is (12). Constraint (13) connects the decision variable for vehicle flow with that for product flow.

Equations (13) and (14) define the auxiliary variables for product unloading and loading respectively. Take (13) as an example. The term $\sum_{k\in\mathcal{M}} z_{v,n,ki} - \sum_{j\in\mathcal{M}} z_{v,n,ij}$ compares the incoming path to and the outgoing path from node i of v with task n. Task n must be unloaded at node i by v if it is positive. Likewise, Equations (15) and (16) define the variables for vehicle unloading and loading, respectively, through the corresponding variables for delivery tasks. Equations (17)-(19) define the vehicle arrival and departure times at nodes. The arrival time is defined through departure and travel times, while the departure time is the sum of arrival, loading/unloading and waiting times. Constraint (20) ensures all vehicles return to their depots by the end of delivery time period T. Constraint (21) imposes the waiting time limit.

The arrival and departure times for tasks are defined in (22) and (23). Product transfers are guaranteed by constraints (24)-(26). More precisely, (24) ensures that product's pick-up must be after its arrival, and furthermore, (25) requires that the corresponding vehicle's departure time cannot be earlier than that of the product. The vehicle v can transport product(s) of task n on route ij if the product's arrival time at i is not less than the vehicle departure time. The vehicle capacity constraint is respected by (27).

Loading dock capacity is respected by (28), where the left-hand side, including an indicator function $\mathbf{1}(\cdot)$ and a maximum function, calculates the maximum number of docks used at node i at any time and l_i is the number of available docks. Clearly, (28) is not a linear constraint. To linearise it, we define additional binary variables $AD^1_{v,u,i}$ ($DA^1_{v,u,i}$) to indicate whether vehicle v departs from node i before u arrives (v arrives at node i before vehicle u departs), and binary variable $l^1_{v,u,i}$ to indicate whether vehicles u, v are at node i at the same time. That is,

$$AD^1_{v,u,i}M \geq D_{v,i} - A_{u,i} \qquad \forall i \in \mathcal{M}, \forall u \neq v \in \mathcal{V} \qquad (31)$$
$$DA^1_{v,u,i}M \geq D_{u,i} - A_{v,i} \qquad \forall i \in \mathcal{M}, \forall u \neq v \in \mathcal{V} \qquad (32)$$
$$l^1_{v,u,i} + 1 \geq AD^1_{v,u,i} + DA^1_{v,u,i}, \qquad \forall i \in \mathcal{M}, \forall u \neq v \in \mathcal{V} \qquad (33)$$
$$l^1_{v,u,i} \leq AD^1_{v,u,i}, l^1_{v,u,i} \leq DA^1_{v,u,i}, \qquad \forall i \in \mathcal{M}, \forall u \neq v \in \mathcal{V} \qquad (34)$$
$$AD^1_{v,u,i}, DA^1_{v,u,i}, l^1_{v,u,i} \in \{0,1\} \qquad \forall i \in \mathcal{M}, \forall u \neq v \in \mathcal{V} \qquad (35)$$

Then (28) is replaced with

$$\sum_{u \in \mathcal{V} \setminus \{v\}} l^1_{v,u,i} + 1 \leq l_i \qquad \forall i \in \mathcal{M}, \forall v \in \mathcal{V} \qquad (36)$$

Finally, Equations (28) and (29) define the decision variables.

3 Model Application

3.1 Experiment settings

In this section, we perform a numerical experiment on a small network with five nodes. The locations of the nodes are selected to represent shops in southeast (node 1), western (node 2), northern, eastern (node 4) and CBD (node 5) areas of Melbourne. See Figure 1. The parameters used in the experiment is given in Table 2. There are two loading docks at each node for transhipment. All vehicles have the same capacity of 300 items, corresponding to a rigid truck, and the delivery distances and demand are provided in Table 3. We analyse the delivery efficiency in terms of total vehicle travel time (VKT), load factor (LF) as well as the delivery reliability in terms of product transfer times. The program is coded in Python and uses Gurobi 9.1.2 to find optimal solutions.

Table 2: Parameters

Labour cost		Vehicle cost		Travel speed	Max wait
LC_d =\$35/hr	LC_f =\$45/hr	VC_d =\$5/km	VC_f =\$10/veh	V =50km/hr	w =1hr

Table 3: Delivery distances ($dist_{ij}$ in km) and demand (s_{ij} in items)

From\To	Node 1		Node 2		Node 3		Node 4		Node 5	
	$dist_{ij}$	s_{ij}	$dist_{ij}$	s_{ij}	$dist_{ij}$	s_{ij}	$dist_{ij}$	s_{ij}	$dist_{ij}$	s_{ij}
Node 1	0	0	53	137	69	20	26	25	18.7	35
Node 2	53	154	0	0	30.2	35	53.4	43	38.8	39
Node 3	69	0	30.2	0	0	0	51	0	54	0
Node 4	26	0	53.4	0	51	0	0	0	18	0
Node 5	18.7	20	38.8	39	54	22	18	22	0	0

3.2 Impacts of product transhipment

We examine the results for three cases (i) dedicated vehicles are employed for shops (i.e., A vehicle only ships products originated from its own depot. See the left plot of Figure 1 for example), (ii) vehicles can transport products from any shop and product swap is forbidden, and (iii) a limit B is imposed on the number of product swaps. For Case (i), given there are three source nodes, 1, 2, 5, each dispatches one vehicle to transport products to other nodes. For example, the vehicle starting from node 1 carries goods for four tasks and drops off them at nodes 3, 4, 1, 5 in order. This results in a total financial cost of \$2902.6 (\$30 for vehicle cost, \$2470.5 for travel cost, \$402.1 for labour cost), VKT of 494.1km and LF of 0.326.

Figure 2 depicts the optimal routes for Case (ii) and Case (iii). Vehicle colours match their depot colours, and each rectangle with two coloured squares represents a delivery task for which the colours indicate its origin and destination. For Case (ii), two vehicles from nodes 2 and 5 are employed. The vehicle from node 5 is responsible for transporting products from node 5 to nodes 1 and 4 whilst the other transports the rest of the products. Relaxing the dedicated-vehicle constraint and keeping the no-product-swap condition, the VKT reduces to 227.4km and the total cost decreases to \$1390 (\$20 for vehicle cost, \$1137 for travel cost, \$240 for labour cost), saving more than 50% compared to Case (i). In Case (iii), we set the limit of swapping times to $B = 6$. The optimal solution uses two vehicles from nodes 1 and 2 and has four products swapped at intermediate node 5. Compared to Case (ii), the product swap helps improve the delivery efficiency; reducing the VKT by about 18% to 185.7km and the cost by 18% to \$1139.3 (\$20 for vehicle cost, \$928.5 for travel cost, \$192.8 for labour cost). Furthermore, Case (iii) raises the LF from 0.519 by Case (ii) to 0.754, which is a significant improvement compared with Case (i).

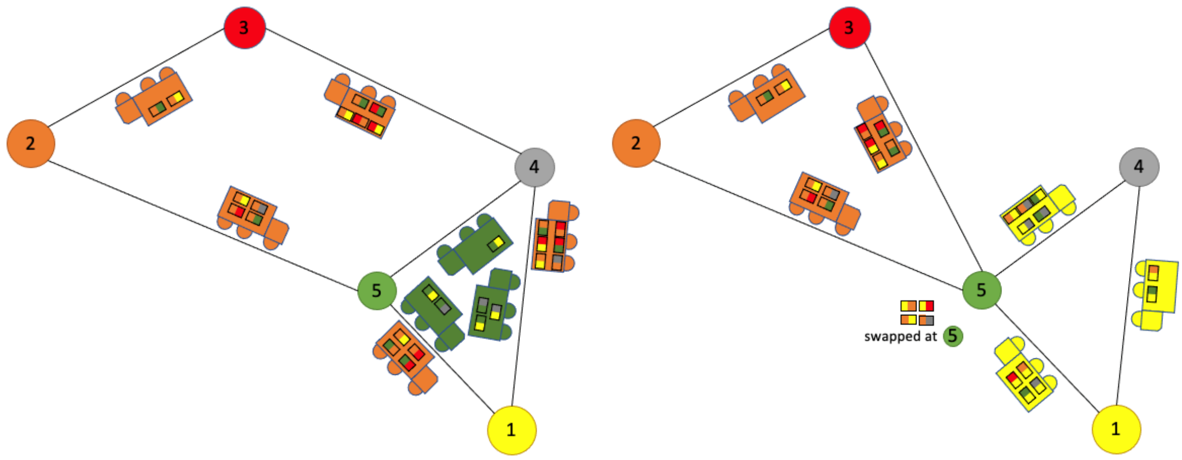


Figure 2: Vehicle and product routes without (Left) and with (Right) product transhipment.

We plot the space-time diagram for vehicles employed for Cases (ii) and (iii) in Figure 3, where x-axis is time and y-axis the index of node. A flat segment represents vehicle loading/unloading at a node. For both cases, the vehicle originated from node 2 starts at time 0, and it returns at time 233min in Case (ii) and 168min in Case (iii). That is to say, compared to Case (ii), Case (iii) has all product delivery complete 65min earlier. Case (i) has all vehicles return to their depots by time 223min. The second vehicle for both Cases (ii) and (iii) has a delayed start. In Case (iii), the vehicle originated from node 1 starts at time 14min so that it arrives at node 5 at the same time as the vehicle from node 2 for synchronisation and transhipment purposes. For Case (ii), since unnecessary delayed start is not penalised, the solution has the vehicle from node 5 start at 47min, which results in the same objective as the solution having it start at time 0. We remark that because the model allows flexible start times (through variable x_v), vehicle waiting at nodes is absent in both solutions for Cases (ii) and (iii). It is important to have flexible start times. Otherwise, the cost for vehicle waiting may jeopardize the benefit derived from product swapping. In that case, the transfer type of vehicle-to-storage-to-vehicle can be considered provided that the cost for temporary storage is reasonable.

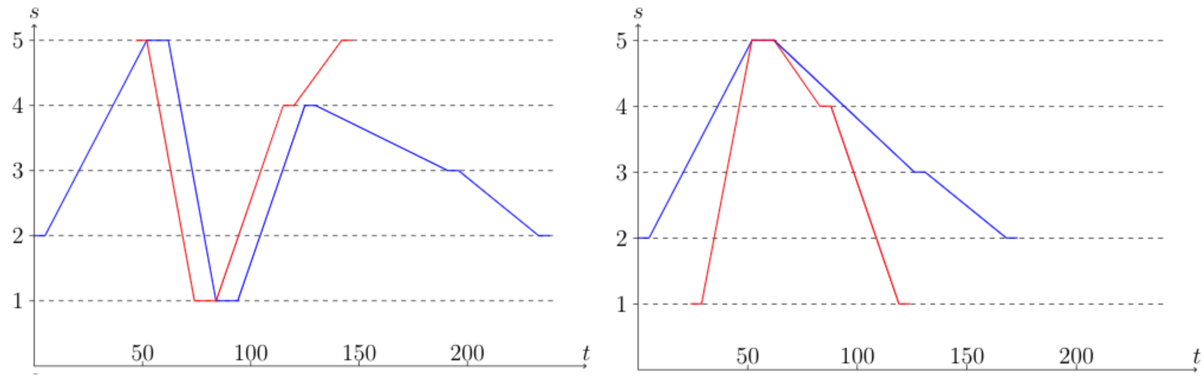


Figure 3: Time-space diagram for vehicle routing without (Left) and with (Right) product transhipment.

3.3 Pareto front

Through imposing a constraint on the number of product transfers, we can transform the multi-objective problem to a problem which minimises the financial cost. We plot the pareto front for this multi-objective problem and the resulting VKT and LF in Figure 4. We observe that the financial cost decreases with the number of transfers allowed. Allowing one task to be transhipped can result in a cost decrease of approximately 5%. For this small network, swapping four delivery tasks can achieve the best financial performance. In addition, product swaps benefit VKT as well as the LF; the VKT decreases with the number of transfers. The LF slightly drops when one swap is implemented, which is due to an empty-running leg. It increases to 0.745 when three swaps are allowed, very close to the largest LF 0.754 observed.

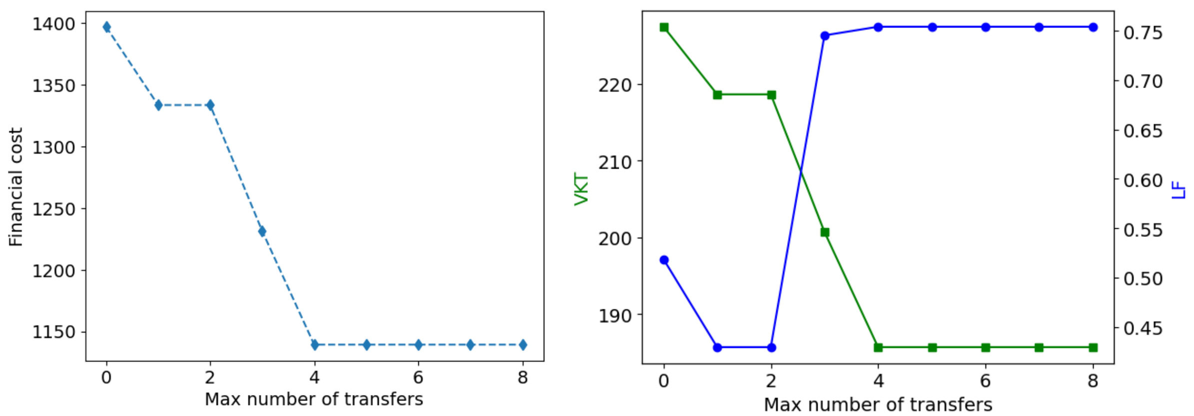


Figure 3: Left: Pareto front; Right: VKT and LF against the maximum number of transfers.

4 Conclusions

Designing more efficient goods transfer networks in urban areas requires consideration of origin and destination patterns as well exploring opportunities for transferring goods at intermediate shops. We analysed delivery efficiency in terms of total vehicle travel time (VKT), load factor (LF) as well as the delivery reliability in terms of number of product transfers. The model presented in this paper allows improved goods transfer networks to be designed and allows trade-offs between transport costs and reliability to be explored.

When product swaps at stores are allowed significant financial cost savings were estimated compared to networks where dedicated vehicles are employed for shops. Swap networks were found to have substantially increased load factors leading to savings in VKT that would lead to reduced emissions.

It is planned to enhance the model in several ways, including incorporating stochastic travel times between shops to allow reliability levels to be investigated in more detail, allowing short term storage of goods being transferred to increase opportunities for efficient transfers and developing heuristics based solution procedures to allow larger networks to be designed.

References

Arslan, A.N., W. Klibi and B. Montreuil (2020). Distribution network deployment for omnichannel retailing, European Journal of Operations Research (EJORS).

Crainic, T.G. and B. Montreuil (2016). Physical Internet Enabled Hyperconnected City Logistics, Transportation Research Procedia, 12, 383-398.

ITF (2018). *Towards Road Freight Decarbonisation – Trends, Measures and Policies*, International Transport Forum, Decarbonising Transport, Paris.

OECD, (2003). *Delivering the Goods, Challenges for the 21st Century,* Paris.

Pan, S., L. Zhang, R.G. Thompson and H. Ghaderi (2021). A Parcel Network Flow Approach for Joint Delivery Networks using Parcel Lockers, *International Journal of Production Research*, 59, 7, 2090-2115.

Taniguchi E. and R.G. Thompson, (2015). *City Logistics: Mapping the Future*, CRC Press, Taylor & Francis.

5G Enabled Video Analytics for Detecting Container Seals in Port Operations

Pavlos Basaras[1], Markos Antonopoulos[1], Konstantinos V. Katsaros[1], Giannis Kanellopoulos[1], Stavros Tsagalas[2], Angelos J. Amditis[1]

1. Institute of Communications and Computer Systems (ICCS), Athens, Greece
2. Piraeus Container Terminal S.A, Athens, Greece

{pavlos.basaras, markos.antonopoulos, konstantinos.katsaros, giannis.kanellopoulos, a.amditis}@iccs.gr, stavros.tsagalas@pct.com.gr

Keywords: 5G, NFV-MANO, Data Analytics, Container Logistics, Container terminal management

Conference Topics: *1) Systems and technologies for interconnected Logistics (3D printing, IoT, machine learning, augmented reality, blockchain, cloud computing, digital twins, collaborative decision making), 2) New communication Networks enabling interconnected logistics (5G), 3) Distributed Intelligence in Physical Internet*

***Abstract**: This article is focused on port control, logistics and remote automation, and aims at detecting the presence/absence of seals at cargo containers. The proposed use case is realized by developing a novel computer vision algorithm for the task at hand, enabled as far-edge computing services on board 5G connected Internet of Things (IoT) devices. The proposed service adopts key enabling technologies of the 5G ecosystem such Network Functions Virtualization (NFV) with MAnagement and Network Orchestration (MANO) support for (automated) lifecycle management of the various service components. The service and overall architecture can be deployed on commodity servers to facilitate interoperability across heterogeneous ports, at low cost. Our results, based on real datasets obtained from the Piraeus port operations illustrate that the subject computer vision approach can achieve up to about 94% accuracy for detecting the presence of absence of container seals.*

1 Introduction

European Ports and their day-to-day operations are key elements for the European economy and economic growth. About 74% of goods imported and exported into the European Union (EU) lands in at least one port before reaching its final destination. At the same time forecasted cargo volumes see an increase of about 57% by 2030 and double the current volumes by 2050. Such volumes introduce significant challenges for seaport operators especially when taking into consideration that almost all European ports along the TEN-T corridors currently operate almost at capacity level. To this end, it is of great importance to address operations at ports that are part of several work chains and thus can pose significant delays to subsequent/dependent operations, such as the loading/unloading phase of containers to/from vessels. Intelligent solutions based on the Physical Internet that require minimal investment and infrastructure requirements are key to the optimization of efficiency in container terminals.

Of particular importance, from the point of view of shipping companies, authorities and stakeholders is the seal-checking process of containers. Container seals are safely locking container doors and are installed right after container stuffing at the place of origin. Any attempt to access the contents of a container requires breaking the container seal. Container seal-presence check when entering/leaving container terminals verifies that container load remains intact during its stay at the terminal (and hence, is of paramount importance), proving to both shippers and customs authorities that the terminal has no liability in regards to container contents. This article focuses in the seal-checking use case in one of the busiest container terminals in Europe, the Piraeus Container Terminal (PCT) port in Greece. The Piraeus port is currently ranked 4th among the busiest European Ports of 2020 in terms of container throughput, and is presently moving about 5.5 million TEUs on an annual basis. In PCT, a mother vessel requires an average of 3000 stevedore moves for operation completion, e.g., for loading/unloading all containers that have either as final or intermediate destination the Port of Piraeus. The current seal-checking process requires human presence at the quay side and about 30 seconds per container to complete. Reducing this time by e.g., 3 seconds per container, results to 9000 seconds (or about 2.5 hours) reduction of vessel stay at the port and removes the need for human presence at an area with high safety risks. Attempts to introduce RFID technology in the past, either at seal or container level, has faced numerous barriers on behalf of container manufacturers, shipping lines and terminal operators mainly due to the infrastructure requirement and financial feasibility.

This paper will discuss 5G technologies enabling a computer vision analytics approach, to address this void, for automatic detection of the presence or absence of container seals as a far-edge computing service. In particular, the developed solution builds on the 5G network support for (automated) lifecycle management of the various service components, including both computing and communications/network resources. Leveraging but also further extending the operational scope of such capabilities, we develop a portable (far) edge computing service that on the one hand, integrates sensing, processing and communication functionalities, and on the other allows the remote and automated management and orchestration of the end-to-end computer vision analytics service; this includes a series of novel capabilities such as instantiating the service, (re-)configuring (computer vision) components, updating overall software, increasing bandwidth and/or resolution of inspection, etc.

Contributions

- We develop a holistic algorithmic methodology for the detection of container seals based on state-of-the-art computer vision techniques, and utilizing key enabling technologies of 5G networks.
- The proposed solution packs all necessary network and software components into Virtual Network Functions (VNFs) that can be orchestrated at scale and on demand to any compute node (or 5G-IoT device) based on opensouce could solutions and industry proven technologies.
- The proposed solution can be placed on commodity x86 servers and similar types of commercial off-the-shelf (COTS) hardware to facilitate interoperability across heterogeneous ports.
- We perform extensive analysis of our model from a real dataset obtained by PCT's daily port operations and find that our model can detect the presence/absence of container seals with about 94% accuracy.

The structure of the paper is summarized as follows: Section 2 references relevant computer vision techniques over a variety of applications and use cases. Section 3 details the selected computer vision methodology for the detection of container seals in port operations. Section 4 briefly discusses the 5G enablers used for the realization of the far-edge computing service, whereas the results are detailed in Section 5. Finally, Section 6 concludes the article.

2 Related Work

The video (or image) analytics task at hand is a classic computer vision task, which can be seen either as a straightforward classification task (i.e., images containing seals or not) or an equivalent object detection task (i.e., detect if the image contains a seal and where). There is a variety of approaches which have been successfully employed to tackle such problems, ranging from classic mathematical computer vision tools such as template matching [1], feature/keypoint descriptors and pertinent matching algorithms [2, 3, 4], to the state-of-the art approach of deep learning, particularly convolutional neural networks [5, 6]. References in the cited works point to an ample variety of successful real-world applications of these methods, with examples covering both image classification and object detection. Each of the aforementioned tools has its advantages and disadvantages; the successful application of each one of these general methods requires special effort and further customization to the specific task at hand and choosing one method over the others is usually a tradeoff. We provide further details of our approach in the next section.

3 Computer Vision Analytics

In Figures 1 to 4 we provide examples of non-sealed versus sealed containers. Seals are marked by red bounding boxes.

Figure 1. Unsealed Container

Figure 2. Unsealed Container

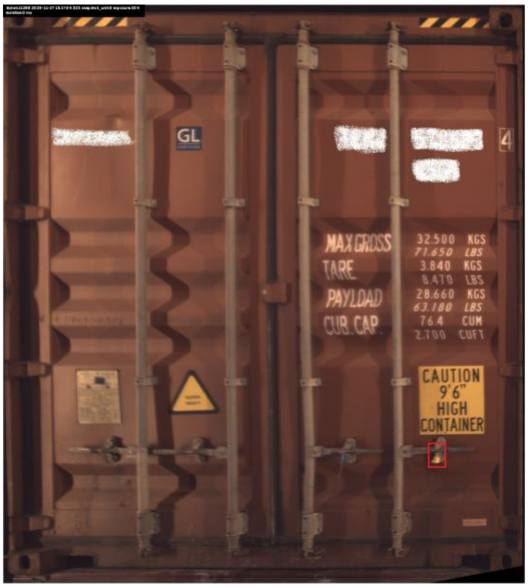


Figure 3. Sealed Container

Figure 4. Sealed Container

In general, seals are located in the lower-right quarter of the cotainer, placed in dedicated positions on the vertical bars or horizontal bar handles of the back side of a container. Given the fact that the images are taken from slightly varying distance and angle, and that bar and handle locations may also vary from one container to another, seals may be located virtually anywhere in the lower-right quarter of the image.

Rather than employing a deep learning based approach we have chosen to develop a custom algorithm utilizing mathematical computer vision tools, for the following reasons. As previously mentioned, the task at hand may be seen as either a classification task or an equivalent seal detection task. In the first case, a cropped part of the image (i.e. its lower right quarter) would be fed directly to a classification scheme (such as a convolutional network) and the output would be a binary variable corresponding to a “sealed” or “non-sealed” label. As can be seen from the Figures, such an approach essentially aims to classify a quite large image, rich in shape (e.g. letter characters), edge (e.g. shadows) and texture (e.g. rust, stickers, etc.) features by the existence or inexistence of a very small pattern in it. Despite their widely-celebrated merits, deep learning approaches have been shown to be heavily prone to unexplainable (and hence very difficult to remedy) errors in such occasions [7]. Attempting to build a seal detection scheme seems more promising; however, any pertinent deep learning approach would require large amounts of properly annotated images, something that is both tedious and extremely time consuming. Thus, we have opted for a more classical approach, whose details and advantages we discuss in this and the following sections.

By careful examination of sample images, we have chosen a collection of representative seal images (like the ones indicated by the red boxes in Figures 3 and 4) which we used as templates to be searched and matched within a given image. The developed algorithm essentially searches the image for templates matching one (or more) of the representative ones. If a match is found, the algorithm returns a bounding box around the matched segment of the image, and the container is labeled as sealed. If no matching is found, the image is labeled as unsealed. In detail, the algorithm works as follows.

1) The input image is cropped around its lower-right quarter. A noise reducing, edge-preserving bilateral filter is applied to the cropped image. Subsequently, we calculate the edges of the image, and an opening/closing smoothing filter is applied on the edges.
2) We repeat the steps of (1) for the seal image.

3) We perform a basic template matching on the edges by using simple matching measures including correlation coefficients and normed square differences [1].
4) If a matching is found, we perform an additional check using to reduce false positives; The histogram of oriented gradients (HOG) feature vectors of the seal image and the matched image area are computed and their Pearson correlation coefficient is computed. A significant correlation (>0.5) declares the image sealed and the algorithm is terminated by returning the seal location in terms of a bounding box. Otherwise, we repeat steps 2-4 for the next seal image.
5) If no matching is found, the image is declared non-sealed.

Note that the algorithm essentially performs seal detection; Based on the results of this detection the image is labeled as sealed or non-sealed; besides the class label, the output of the algorithm contains also the location of the seal in the image, if the image contains a seal. The algorithm was implemented by making extensive use of the OpenCV computer vision library. As detailed in the next sections, the algorithm was built by experimenting on a dataset of about 4,000 images. It has been tested on a separate, previously unseen and unused dataset of 2,000 images, resulting in an overall detection accuracy of 93.55%.

4 5G Enabled Video Analytics with NFV-MANO Support

This section describes 5G technologies enabling the discussed computer vision approach (Section 3) for automatic detection of the presence or absence of container seals as a far-edge computing service. In this framework, we present the overall architecture of our end-to-end solution, including both hardware and software components. The end-to-end edge computing service is composed of commercial off-the-shelf (COTS) hardware and open-source platforms, to expedite deployment and potentially facilitate interoperability across heterogeneous ports. The envisioned video analytics service will be deployed on novel 5G enabled Internet-of-Things (IoT) devices (c.f., 4.2) positioned at selected areas of interest within the Piraeus port, to automate the container seal detection process. In the sequence we present the architecture and respective devices that are used for our in lab testing at the ICCS 5G testbed.

4.1 MANO (ETSI MAnagement and Network Orchestration)

The adoption of NFV is considered as one of the enablers for a fully softwarized 5G architecture, that allows significantly higher flexibility for network service providers to instantiate and monitor services, configure and update them, commonly known as day0 to day2 operations of management and network orchestration [13].

Particularly, the MANO platform that is exploited for the container seal use case is based on Open Source MANO (OSM) release nine [8], an ETSI-hosted software stack aligned with ETSI NFV. The main platform is divided in three main components; the Virtualized Infrastructure Manager (VIM), which controls and manages the resources of an Network Function Virtualization Infrastructure (NFVI), i.e., the 5G-IoT devices that will host the VNFs and perform the video analytics tasks; the Virtual Network Function Manager (VNFM) taking care of the instantiation of VNFs, configuration, modification and termination of VNF instances, i.e., activating/deactivating/modifying the far-edge computing service; and the NFV Orchestrator (NFVO), which orchestrates the allocation of resources (compute, storage, networking, etc.) under the control of (potentially) different VIMs and manages the lifecycle of network services. This architecture will allow the instantiation of the container seal detection service, at scale, towards any 5G-IoT device under the control of the MANO platform, e.g., distributed at several sites of interest within the port premises.

The VIM exploited by the MANO platform for the container seal detection use case includes several components based on a subset of services offered by Openstack [9] (Victoria release). Other VIM solutions are also available e.g., OpenVim [10]. OpenStack is an opensource cloud operating system that controls large pools of compute, storage, and networking resources, i.e., the NFVIs, all managed and provisioned through APIs with common authentication mechanisms. The VIM orchestrator will be the interface towards the NFVI devices, i.e., the 5G-IoT nodes, that will host the VNFs (software applications that deliver network and computer vision service functions) and deliver the respective solution for facilitating the far-edge computing service of container seal detection. The VIM tool is controlled by the OSM to facilitate the MANO system, taking care also of the VNFM and VNFO services given the pool of NFVI nodes and the set of VNFs. For more details regarding the 5G-MANO stack please refer to [13] and references therein.

4.2 5G-IoT Device

The designed portable 5G-IoT edge device is composed of three main components: a generic compute node (that hosts the virtualized network and video analytics functionality) such as the NVIDIA Jetson Kit [11]; a high-resolution camera for data capturing (i.e., the input video feed for the analytics model of the container seal use case) and a 5G interface to establish communication with the backend system for visualization, database management, streaming of ultra-high-definition (UHD) videos etc. The prototype in-lab testing equipment of the 5G-IoT device as designed and tested at the ICCS 5G testbed is depicted in Figure 5.

Figure 5: 5G-IoT device components.

For the in-lab testing at ICCS we employ the 4G/5G stack of the OpenAirInterface platform following the Non-standalone (NSA) option of 5G technology [12]. USRP Software Defined Radio (SDR) devices are employed for the cellular connection, e.g., through B210 or N310 [14], and establish the high bandwidth connection necessary for the transmission of the UHD video streams. Upon migration to the port premises, the device will be placed at the quay side cranes in PCT, to enable the detection of presence/absence of container seals. The overall architecture is illustrated in Figure 6.

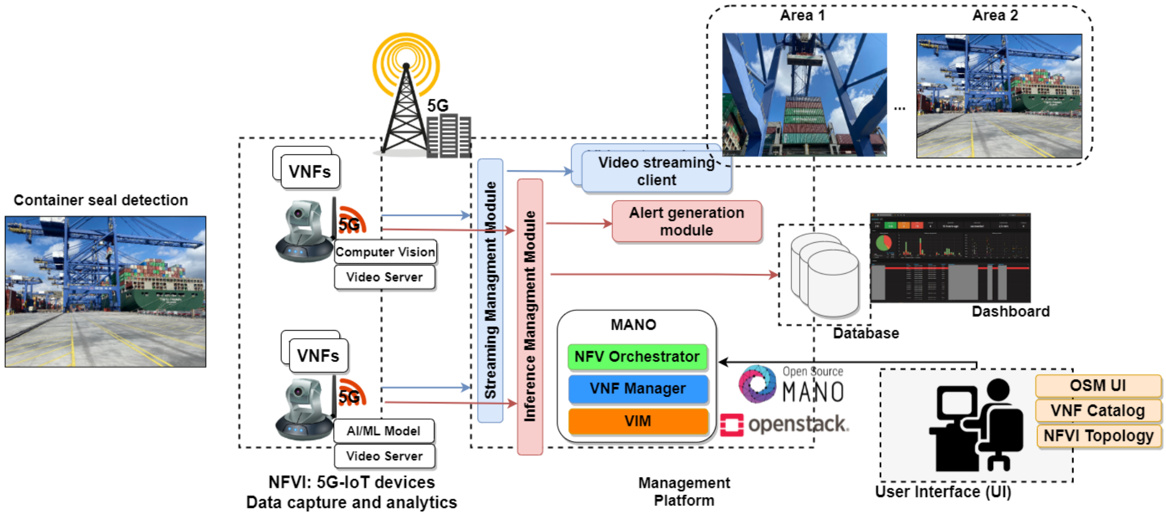


Figure 6: NFV-MANO enabled far edge computing architecture overview.

The instantiation process of a set of VNFs has the following workflow. Initially the user (administrator) interfaces a User Interface (UI), where the VNFs to be deployed are selected by a catalogue of supported services. The VNFs will bring all the necessary components for the analytics tasks to the 5G-IoT devices (software packages, libraries, etc.) as discussed in Section 3. When the user instantiates the service, the VNF descriptions are sent to the underlying VIM for preparing and configuring the physical infrastructure that will host them, i.e., the 5G-IoT nodes. These services will configure network interfaces, features regarding the virtualization technologies for the underlying physical resources (Virtual machines, Containers, Bare metal, etc.) as well as all software related services/components that need to be instantiated at the device for container seal detection analytics. The streaming management module (Figure 6) handles all video data transmitted by the 5G-IoT devices at the backend system enabling real-time monitoring of the operation, whereas the inference management module will receive the inference of analytics services from the IoT device, interface with the database (and respective dashboards) and alert generation module, e.g., container seal missing from the subject container.

For the overall deployment at PCT premises the required investment is minimal since it does not require special cameras and the cost of the compute node is less than the one of a standard user PC. At the same time, MANO capabilities lower the operational costs related to physical access to the device or downtime for maintenance.

5 Experimentation and Results

5.1 Dataset and Setup

A total amount of about 4,000 images, roughly 2,500 sealed and 1,500 non-sealed was used to select the representative seal patterns, and to build and tune the parameters and thresholds of the algorithm. The resulting algorithm was tested on a separate, previously unseen dataset consisting of 1,000 sealed and 1,000 non-sealed images.

5.2 Results

We visually inspected the output of the algorithm for each image of the test set, and manually measured correct versus incorrect seal detections. The confusion matrix of the resulting seal detection scheme is given in Figure 7.

	Actual Sealed	Actual Non-Sealed	**Accuracy**
Classified as Sealed	921	50	94.85%
Classified as Non-Sealed	79	950	92.35%
Recall	92.10%	95.00%	
	Overall Accuracy:	93.55%	

Figure 7: Confusion matrix of the proposed computer vision approach for the detection of container seals.

6 Conclusion

We have built a computer vision algorithm achieving high accuracy in detecting the existence of container seals in images of container back sides. For reasons mentioned above, we have opted to build the algorithm using mathematical computer vision tools, rather than deep learning. Contrary to deep learning approaches, the developed algorithm has the advantage that each of its steps is perfectly explainable, thereby facilitating further experimentations and improvements. Furthermore, by (optionally) tightening the algorithms' decision thresholds, the algorithm can be exploited as a reliable tool for annotating large amounts of images, thereby facilitating the use of deep learning-based object detection techniques. The overall service targets the far-edge computing vertical with NFV-MANO support for automation for ports: port control, logistics and remote automation. 5G enabling technologies have been considered based on open source solutions, COTS hardware and industry proven technologies to facilitate cross-knowledge sharing with heterogeneous ports or any other third party interested.

Acknowledgement

This paper is part of the 5G-LOGINNOV project, co-funded by the European Commission under the H2020 Research and Innovation Programme (grant agreement No 957400).

7 References

[1] R. Brunelli, "Template Matching Techniques in Computer Vision: Theory and Practice" Wiley;1st edition (April 20, 2009)
[2] E. Rublee, V. Rabaud, K. Konolige and G. Bradski, "ORB: An efficient alternative to SIFT or SURF," 2011 International Conference on Computer Vision, 2011, pp. 2564-2571.
[3] Hussain S., Triggs B. (2012) Visual Recognition Using Local Quantized Patterns. In: Fitzgibbon A., Lazebnik S., Perona P., Sato Y., Schmid C. (eds) Computer Vision – ECCV 2012. ECCV 2012. Lecture Notes in Computer Science, vol 7573. Springer, Berlin, Heidelberg. https://doi.org/10.1007/978-3-642-33709-3_51
[4] Joseph J. Lim, C. Lawrence Zitnick, and Piotr Dollar. 2013. Sketch Tokens: A Learned Mid-level Representation for Contour and Object Detection. In Proceedings of the 2013 IEEE Conference on Computer Vision and Pattern Recognition (CVPR '13). IEEE Computer Society, USA, 3158–3165. DOI:https://doi.org/10.1109/CVPR.2013.406

[5] Rawat W, Wang Z. Deep Convolutional Neural Networks for Image Classification: A Comprehensive Review. Neural Comput. 2017 Sep;29(9):2352-2449. doi: 10.1162/NECO_a_00990. Epub 2017 Jun 9. PMID: 28599112.
[6] Z. Zhao, P. Zheng, S. Xu and X. Wu, "Object Detection With Deep Learning: A Review," in IEEE Trans. on Neural Networks and Learning Systems, vol. 30, Nov. 2019.
[7] Heaven D. "Why deep-learning AIs are so easy to fool", https://www.nature.com/articles /d41586-019-03013-5. [8] O. ETSI, “Open Source Mano,” https:// osm.etsi.org/, 2020
[9] N. Makris, C. Zarafetas, A. Valantasis, and T. Korakis, “Service orchestration over wireless network slices: Testbed setup and integration,” IEEE Trans. Netw. Service
Manag., vol. 18, no. 1, Mar. 2021.
[10] T. Sechkova, M. Paolino, and D. Raho, “Virtualized infrastructure managers for edge computing: Openvim and Openstack comparison,” in 2018 IEEE International Symposium on Broadband Multimedia Systems and Broadcasting (BMSB). IEEE, 2018, pp. 1–6.
[11] NVIDIA, https://developer.nvidia.com/buy-jetson?product=jetson_agx_xavier&location=US, 2021.
[12] N. Nikaein, M. K. Marina, S. Manickam, A. Dawson, R. Knopp, and C. Bonnet, “OpenAirInterface: A flexible platform for 5G research,” ACM SIGCOMM Computer Communication Review, vol. 44, 2014.
[13] Girma M. Yilma, Zarrar F. Yousaf, Vincenzo Sciancalepore, Xavier Costa-Perez, “Benchmarking open source NFV MANO systems: OSM and ONAP”, Computer Communications, 2020.
[14] Florian Kaltenberger, Aloizio P. Silva, Abhimanyu Gosain, Luhan Wang, Tien-Thinh Nguyen, “OpenAirInterface: Democratizing innovation in the 5G Era”, Computer Networks, 2020.

5G applications in ports, corridors and logistics hubs

Margarita Kostovasili[1], Georgios Tsimiklis[1], Angelos Amditis[1]

1. Institute of Communication and Computer Systems, Athens, Greece

Corresponding author: margarita.kostovasili@iccs.gr

Abstract: *This research paper aims to present the main applications that have been performed in seaports based on 5G network communications, following a scoping review of available studies. The main applications that have been implemented in ports are presented, focusing mainly on improvement and automation of container terminal operations, vessel management and warehouse management. Such applications can support system multi-connectivity, sensor measurements, large-scale infrastructure monitoring and customizable network coverage, enabling the implementation of Physical Internet (PI) network. The outcomes of 5G applications implementation include increased efficiency of the port, providing flexibility and reliability to its daily operations, increased safety for the on-field employees, as well as control of the environmental impact and reduction of CO_2 emissions. Further research opportunities are presented, towards the adoption of a wider operational framework based on 5G technology, the full terminal automation, the automation of procedures among different terminals, the provision of services tailored to specific stakeholder groups and the promotion of PI approach.*

Conference Topic(s): *Interconnected freight transport, logistics and supply networks*

Keywords: *Internet of Things, 5G, port, container terminal management, automation, digitalization*

1 Introduction

As most cargo and goods are transported by sea, seaports have a key role serving as logistics hubs, where the cargo is imported/exported and transferred to/from other transportation means and services. Major seaports worldwide are called to serve an increasingly high demand of advanced cargo vessels, following the technological trends and adapting relevant innovations to their operational approach (Nikolopoulou, et al., 2019), taking also into account their impact to the port ecosystem, community and city, with respect to the environmental regulations, towards the minimization of the CO_2 emissions. A significant parameter for satisfying all the aforementioned needs is the direct, fast and real-time communication among all systems, components and stakeholders that are involved in the cargo management process (Zhong, et al., 2019). The digitalization of the ports can be supported by 5G networks that provide low latency, high capacity and increased bandwidth, allowing the collection and management of huge amounts of data from sensors and devices. 5G is considered as an enabler for the promotion of Physical Internet (PI) objective towards the development of sustainable, interconnected and highly collaborative logistics model (Nagendra, 2019). Currently, the 5G implementation is ongoing with many different applications in ports around the world, resulting to the need for this research paper, in order to further investigate the topic and provide relevant outcomes.

2 Methodology

In order to study the selected topic and compose an overview of it, the approach of scoping review was selected, allowing literature mapping and concept definition through the collection of relevant research studies and best practices that have already been tested and could be implemented in a broader context. Moreover, following the scoping review, possible research gaps can be identified and serve as the basis for further research opportunities. The first step of the selected methodology is to formulate the research question and possible sub-questions and define the corresponding search string. Following this, the sources of research are defined and the relevant databases are selected, along with the desired criteria for studies' inclusion and exclusion. The analysis of the scoping review steps is presented in the following sections.

2.1 Research question and keywords

The selected research question that was formulated was: "How 5G technology can contribute to the operational performance of major ports?". The following sub-questions were also selected for further analysis of the topic: (i) *"What aspects of the port operations can be improved by 5G applications?"*, (ii) *"What are the existing 5G-based techniques that have been implemented in major ports?"*, (iii) *"What are the potential 5G use cases that could be implemented and tested in major ports?"*. In order to identify relevant literature, the methodology followed was: (i) to determine the basic components to be used as keywords and (ii) to form the optimal query that could result the most relevant articles and papers. The selected keywords were "5G" and "port", in order to collect articles referring to 5G applications and innovations implemented in major ports and find possible use cases and gaps. The main issue with these keywords was that the term "port" was answered quite often in informatics literature and used mainly as hardware aperture or opening meaning (i.e. USB port). In order to encounter this, synonyms of port were used, such as seaport or harbor, while in parallel additional terms referring to maritime transport were added to the query (i.e. shipping, supply chain, container terminal etc.). The final query was the following: "5G" AND ("seaport" OR "container terminal" OR "inland port" OR "container port" OR "intermodal terminal" OR "cargo terminal" OR "port operation" OR "logistics hub").

2.2 Inclusion and exclusion criteria and refinement

Specific criteria were used for inclusion/exclusion of studies, setting that they should be peer-reviewed and published in scientific journals, while the publication date was selected to be between 2016 and 2020, as the desired topic is quite new and has not yet been broadly studied. The examined libraries were Google Scholar, Scopus and Taylor & Francis, while the queries were searched anywhere in the documents, in order to include relevant literature published under a broader topic. Also, studies that examine only 5G technology or solely port-related applications were excluded from the review. Only the papers that were written in English language were selected and the duplicates were removed, while grey literature was excluded from the analysis after double screening of the results. This process is presented in the following flow chart (Figure 1), while an overview of these studies is described in the following sections. The theoretical background is presented in the beginning, providing a broader perspective of the selected topic, while the 5G applications implemented in ports are grouped in three main categories: container terminal operations, vessel management and warehouse management. The results that occurred from the analysis and the answers regarding the defined research question are presented in the conclusions, followed by gap identification and further research opportunities.

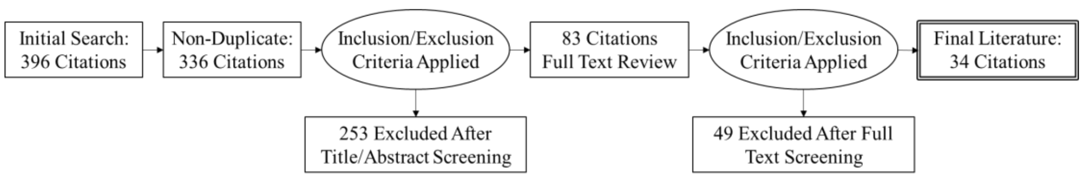


Figure 1: Scoping review process flow chart

3 Theoretical background

Many studies have been conducted in the last decade, describing the impact of technology and digitalization of ports and container terminals, but only few of them analyze the use of 5G technology in order to optimize the daily port operations. The study of Anwar et al. (2019) is such a case, where an extended study has been conducted and the digitalization in container terminal logistics is presented, analyzing the impact of many emerging technologies, but leaving aside 5G and other emerging technologies. This happens due to a twofold reason: as 5G technology is quite new, not all of its functionalities have been studied and thus the assessment of its impact cannot be complete; in parallel, 5G-based applications that have been tested in real-life conditions and are currently implemented and operative in a real port environment are quite limited. This leads to absence of significant number of use cases that can be studied and assessed, in order to draw valid conclusions. Furthermore, there are many studies dealing with the topic under a wider viewpoint, where digitalization and technology integration into the daily operations of a port are presented as part of a general "smart port" approach.

3.1 Smart port and IoT

There are many definitions of "Smart Port", describing all required technologies and functionalities that need to be implemented. According to Aslam et al. (2020), a port is considered "smart" when all its wireless devices, smart sensors, actuators, data centers, and other IoT-based port devices are fully connected with each other via the Internet and exchange information, enabling port authorities to offer more reliable information and various new services to their clients. Applying the IoT principles and technologies in the maritime sector, the interconnection of ships, crews, cargoes, on-board equipment, waterway environment, waterway facilities, shore-based facilities and other navigation elements enables these objects to collect and exchange data. A network of smart sensors and actuators, along with communications and computing, is the key infrastructure of the smart port, allowing the provision of services in a faster and more efficient manner (Zhong, et al., 2019). The efficiency of the ports can be improved by automating port operations, increasing the potential for economic growth (Muhammad, Kumar, Cianca , & Lindgren, 2018).

Competitiveness is a significant challenge for ports in emerging economies, leading to adoption of different technology-based solution approaches that can enhance the port's status (Moros-Daza, Amaya-Mier, & Paternina-Arboleda, 2020). Digitalization has been identified as one of the main game changers for ports (Acciaro, Renken, & El Khadiri, 2020), while the adoption of top ICT systems can increase informational connectivity of ports and strengthen their role in the global value chains and supply networks (Henriquez, Martinez, & Martinez, 2019). Another key aspect is the environmental excellence of port operations, which can be achieved and enhanced through new and emerging technologies (Lacalle Úbeda, Llorente, & Palau, 2019). In parallel, the current commercial ships are adopting new technologies based on ICT and automation increasingly to deal with global environmental regulations or energy-saving issues (Huh, 2020). The introduction of Internet of Things (IoT) and accelerated digitalization in port operations have globally improved productivity of ports and overall competitiveness (Jović, Tijan, Aksentijević, & Čišić, 2019). IoT technologies can also contribute to improve the safety of processes and the security of the port (Postolache, et al., 2019).

3.2 5G technology

The fifth-generation (5G) wireless systems have incorporated radical recent technological advancements for significantly enhancing the wireless capacity and adaptability (Ho, et al., 2019). The 5G is a new networking solution to overcome the challenges of future communication needs, where an enormous number of smart devices could communicate with each other anywhere and at any time (Taboada & Shee, 2020). The key element that differentiates 5G from the previous generations is that it offers the possibility to tailor mobile data services to the particular characteristics of specific (business) users (Anker, 2019). 5G mobile network enables applications and services that require better reliability, improved energy efficiency, massive connection density and lower latency, unlocking the full potential of industrial IoT for transport and logistics (Mazzarino, et al., 2019). Along with IoT and data analytics, 5G networks can support cargo ports to handle upcoming and future capacity, efficiency and environmental challenges (Nikolopoulou, et al., 2019). Some of the main features and services of 5G technology that can serve as enablers for advanced applications and innovations for the port of the future are presented below.

Network slicing: Network slicing is a key architectural feature of 5G, enabling networks to dynamically and flexibly adapt to the requirements of different applications (Patzold, 2018). The basic idea of network slicing is to "slice" the original network architecture in multiple logical and independent networks that are configured to effectively meet the various services requirements. Through it, the network is transformed from an infrastructure element to a service element, enabling numerous 5G smart services with diverse requirements (Khan, Yaqoob, Tran, Han, & Hong, 2020).

URLLC: Ultra-Reliable Low-Latency Communications service provides latencies of 1ms with high reliability (99.999 percent) for mission-critical applications. Low latency allows a network to be optimized for processing incredibly large amounts of data with minimal delay and to support instantaneous and intelligent systems. Through this service, mission-critical applications such as autonomous vehicles and machinery are enabled (Ho, et al., 2019).

mMTC: Massive Machine Type Communications service supports a very large number of connected devices (up to 1 million devices per km^2). It is aimed at supporting dense connections of various device types (e.g, mobile devices, IoT devices and sensors) in crowded areas (Ho, et al., 2019). Through mMTC, enormous number of devices can be connected in the same network with high coverage, supporting intelligent logistics and storage and smart cities applications (Taboada & Shee, 2020).

eMBB: Enhanced Mobile Broadband service provides large capacity for mobile broadband services (Zhong, et al., 2019). It defines a minimum level of data transfer rate, promising to deliver both vastly increased bandwidth and decreased latency compared to existing 4G services. It supports high data rate information exchange across a wide coverage area, utilized in high-resolution video streaming (4K, 8K) and truly immersive AR/VR applications.

4 5G applications in ports

As ports are significant nodes in larger transportation systems, high potentials are embedded towards digitalization and advanced connectivity, both for internal operations and hinterland interaction and collaboration (Inkinen, Helminen, & Saarikoski, 2019). The integration of 5G, new generation of IoT, Artificial Intelligence (AI) and other technologies can modernize the overall supply chain, providing increased service capabilities (Yi, 2019). More specifically, 5G technology can transform sea logistics and port operations, while its combination with IoT can make tracking deliveries possible in real time and provide just-in-time deliveries (Nagendra, 2019). This can be considered as part of a wider physical, digital and operational interconnectivity promoted by the PI approach. In a port environment, this interconnectivity can be achieved through solutions that increase data sharing and visibility between supply chain actors and enhance the hinterland connectivity of the port with the surrounding urban space (Nikolopoulou, et al., 2019).

Another common practice is the replacement of typical manually operated cranes that are used in container terminals with automated Rubber-Tyred Gantry cranes (AutoRTGs) supported by 5G network (Kokkoniemi-Tarkkanen, et al., 2019). In this study, the use of a wireless 5G ultra-reliable low latency communication (URLLC) link for the communication between the automated crane and the terminal operating system is examined and used for safety-critical communication, remote control and delivery of video streams between a remote-control center and AutoRTGs. Moreover, vehicular communication concepts and relevant network models can be set up, aiming to increase the occupational safety in the developing port environment. Such models involve communication equipment for on-port workers, pedestrians and machinery, along with the deployment of communication infrastructures, through the support of 5G technology, providing communication-based cooperative scheme for improving the safety of workers and optimizing the management of on-port vehicles (Bauk, Calvo, Schmeink, Azam, & Mathar, 2019).

4.1 Container terminal operations

Freight transportation industries are extremely interested in exploring the use of the emerging technologies, such as the fifth-generation cellular network technology (5G), high-resolution cameras and connected and automated guided vehicles (AGVs), for next-generation automated container hubs. The port of Rotterdam has one of the most automated terminals in the world that is powered by interconnected unmanned RTG cranes and AGVs, implementing the PI approach. With the support of private cellular network, the remote control of heavy machinery and automated port vehicles is possible in real-time, creating also a hugely more efficient and secure way of connecting and tracking all shipments and goods. It is a fully AI-powered autonomous port where IoT, big data and 5G are implemented and support the facilitation of advanced analytics and decision making for autonomous workflow management (Gurumurthy & Bharthur, 2019). In addition, several models and algorithms have been built to address the cooperation and synchronization among automated cranes and AGVs in an efficient, smooth and secure way, considering the inter-crane and inter-vehicle interferences, optimizing the scheduled-based and space-time vehicle route planning and performing real-life validation tests (Chen, et al., 2020).

In the Hamburg Port, a public 5G-based network is used to monitor vital infrastructure via enabling virtual reality (VR) applications. A network-slicing testbed has been conceived for large-scale 5G industrial networks in order to support multi-connectivity, proving that network slicing in a real, large-scale industrial environment is technically feasible (Ho, et al., 2019). Both the testbed and the use cases that have been implemented and tested are part of the European project 5G-MoNArch. The basic features of network slicing have been deployed, while multi-connectivity has been implemented, achieving high reliability throughout the testbed area (Rost, et al., 2018). Three services have been selected and demonstrated in the testbed: mMTC deployed for sensor measurements on mobile barges; eMBB applied for live remote site support; and URLLC tested for traffic light control (Taboada & Shee, 2020). Through the deployment of these services, the Hamburg Port Authority in collaboration with Deutsche Telecom and Nokia have tested the use of mobile communications to manage traffic lights within the port area, to collect and process environmental measurement data in real time and to monitor critical infrastructure through VR applications (Patzold, 2018).

A private non-standalone cellular-based network (LTE/5G) has been implemented in the Livorno port, in order to support main port operations. Through the use of remotely controlled quay cranes, the cargo loading/unloading process is smoother and safer, while the prediction accuracy regarding the maintenance for on-field machinery has been increased. In parallel, the automated remote control of Unmanned Ground Vehicles (UGVs) for loading and unloading operations in the port area has been demonstrated in the port of Livorno (Nagendra, 2019). The Livorno port has also served as a testbed for the research and innovation project COREALIS, which involves a 5G-based control module that can detect general cargo in a shorter time than usual human-driven communications, as well as enabling better management of the cargo (Cavalli & Lizzi, 2020). The required time for processes like vessels' berthing, pallet searching or forklift operation and the necessary yard movements can be reduced, leading to the increase of the capacity of the container terminal and the improvement of the overall performance of the port, through the deployment of cargo management optimization techniques with regards to loading/unloading, distribution into the storage area and handling during loading phases on the ship and selection of most appropriate forklift for handling (Nikolopoulou, et al., 2019).

Apart from the European ports and relevant use cases, Qingdao Port in China has one of the first fully automated terminals worldwide, supported by 5G network slicing communication, completing the verification of the world's first 5G smart terminal program at the beginning of 2019 (Jiang, 2019). One of the main container terminal operations that has been automated based on 5G connection is the remote wireless control of crane control operation, grabbing and transportation containers (Yang, Ding, Zhang, & Lin, 2019). In this case, the optical fiber communication has been replaced by 5G network, allowing real-time monitoring and fully automatic operation of the shore quay cranes and providing higher flexibility and reliability in the communication system. Another case of a real-time smart port implementation is in Xiamen Ocean Gate (Yang, et al., 2018), which is the first automated container terminal in China that follows the standards of a global automation terminal handling-system. Now, it has become one of the first terminal around the globe that uses 5G wireless networks in its operations.

4.2 Vessel management

In general, an autonomous ship is operated by a land-based surveillance and control center connected with a collection of digital technologies such as IoT, data analysis technology, and broadband communication, making the 5G communication system a major technological pillar for its implementation (Lambrou, Watanabe, & Iida, 2019). As the interconnection of vessels and shore facilities depends directly on the communication network and requires high reliability and performance, it can be based on 5G technology and further developed using wireless networking, wireless transmission, radio frequency and other 5G services. Following the Internet of Vessels (IoV) approach, the realization of intelligent navigation of ships is possible, supporting completed navigation, collision avoidance, information display, monitor, alarm, satellite communication and ship management and allowing the crew and shore-based commanders to observe, operate and optimize the information process collected from the equipment (Tian, Liu, Li, Malekian, & Xie, 2017).

In Port of Rotterdam, apart from the use of unmanned RTG cranes, a digitization project has been introduced along with IBM, including the implementation of sensors and the development of a centralized dashboard application, aiming to support the management of connected ships and collect all relevant information under an interconnected PI network. Through this system, multiple real-time data streams on tides, currents, temperature, wind speed/direction, water levels, berth availability and visibility can be collected and analyzed through an IoT platform, supporting decision making towards the reduction of waiting times, the determination of optimal times for ships to dock, load and unload and the prediction of arrival and offloading time (Gurumurthy & Bharthur, 2019). Another use case related to vessel management was tested as part of the Hamburg port 5G testbed, providing real-time location information of ships and relevant measurements from sensors deployed on barges (Rost, et al., 2018).

4.3 Warehouse management

Regarding warehouse management and the contribution of 5G technology to such a process, automation and remote control of cargo handling, monitoring and tracking systems is feasible through the implementation of available 5G services (URLLC, mMTC, network slicing, eMBB). In parallel, 5G technology can support the use of smart and autonomous vehicles for cargo handling and monitoring in port’s warehouses and storage areas. In both cases, the benefits enabled by 5G include lower times to find cargo, reduced accidents, less operational inefficiencies, fewer human mistakes, lower handling time per cargo unit, reduced economic costs and improved competitiveness (Cavalli & Lizzi, 2020). Moreover, the use of 5G and IoT devices can boost efficiency between warehouses and distributors, giving customers clearer visibility of their deliveries (Nagendra, 2019) and support decision making for intelligent storage, avoiding multiple and in many cases unnecessary handling (Postolache, et al., 2019).

In addition, the existing warehouse management systems facilitate operation management in their daily planning, organizing, staffing, directing and controlling the utilization of available resources, to move and store freight (Mazzarino, et al., 2019). Such systems that operate using 5G networks can enhance the performance of IoT-based ports, supporting services such as real-time cargo tracking, real-time communication with vessels, prediction of time of arrival and booking of docking slot, improving the delivery performance and optimizing the overall port and warehouse operational flow (Aslam, Michaelides, & Herodotou, 2020). Further to the operational perspective, the use of automated machinery and warehouse robots based on 5G technology to undertake tasks that are labor-intensive, repetitive and/or dangerous can reduce the human involvement in such processes, enhance the overall warehouse performance and increase the occupational safety (Gurumurthy & Bharthur, 2019).

4.4 Other applications

Apart from 5G-based applications that are implemented either in the shore (container terminal operations, warehouse management) or at the vessels, there are studies examining the use of 5G capabilities for underwater communication. 5G networks have a great impact on underwater wireless communication as they support the improvement of the data rate, connectivity, and energy efficiency (Ali, Jayakody, Chursin, Affes, & Dmitry , 2020). The sensors that are deployed under the sea surface are highly demanding and require a significant data rate for the transmission of their signals, making the use of 5G necessary in order to achieve continuous and real-time monitoring of sensors’ data and enhance the overall underwater communication. This can have significant effect in port infrastructure management, providing information related to the environmental conditions and supporting decision making for port operators.

On the other side, the topic is not limited only to studies presenting current uses cases and test beds, but also to studies that describe the constraints and limitations of such systems and try to provide supplementary solutions that will serve as backup and ensure the consistency of the port operations. Such a case is the study of Ali-Tolppa & Kajó (2020), where the capacity and coverage of the implemented network is analyzed and optimized. Using the Hamburg Smart Port 5G testbed and collecting data such as ship positioning, latencies and signal strength, analytics and machine learning techniques are used so as to predict possible network service degradations. As the system is aware of the ship position and the simultaneous network coverage, it can adjust the cell transmission power and optimize the coverage to avoid network failures and maintain the expected quality of service.

5 Conclusions

Based on the analysis of the examined studies, many ports have adopted a more digitalized and automated operational approach towards the PI concept implementation, with 5G networks playing a critical role. 5G applications implemented in ports can contribute to the improvement of its socio-economic and environmental impact, as they will allow the port to serve significant numbers of devices and data streams. Through data processing and performing optimization techniques, the port can improve its operational flow and thus manage more cargo, increase its profit and create new jobs, along with the reduction of the total CO_2 emissions. Also, the automation of port operations can reduce the human presence in the port yard and thus improve the overall safety. Answering the research question set in the beginning of this scoping review, the main port operations that can be improved by 5G applications were presented, along with already implemented 5G-based techniques, including but not limited to yard equipment monitoring, port traffic management, vessel navigation and berthing, cargo tracking, loading/unloading and handling. In all cases, the provision of high reliability and low-latency communication network resulted in operational flow improvement, efficiency increase and enhancement of port's role in the supply chain environment.

All applications that target to the automation of logistics processes and the development of innovative solutions to support port operations have a significant contribution to the implementation of PI towards a global, open, interconnected network, using a set of collaborative protocols and standardized smart interfaces. In parallel, it is clear that the implementation of such use cases is not limited only in port environments, as the proposed 5G port solutions have increased transferability to corridors and other logistics hubs. Innovations such as vehicle automation, smart traffic management, machinery and equipment remote control and warehouse operation management can be easily modified and adapted to other fields and sectors, incorporating all the aforementioned advantages and benefits. The collaboration between ports and other logistics hubs can promote the implementation of innovative concepts and solutions to interested stakeholder groups, highlight common goals and interests and increase the growth potential of the overall shipping industry and logistics sector.

6 Research opportunities

Through this scoping review on applications based on 5G technology and implemented in port environments, some gaps were identified that could be serve as further research opportunities, answering the last sub-question of the selected topic. A general conclusion is the lack of a wider operational framework based on 5G technology, as the majority of ports utilize 5G capabilities to single use cases. The expansion of its use and the adoption of an 5G communication network could highlight the full capabilities of 5G technology and enhance the overall port efficiency and competitiveness. This could be combined with a research on ports that provide fully automated terminals. Currently, there are only a few cases of full automated terminals and most of them utilize the 5G capabilities only in specific use cases and not as the main communication network under which all devices, equipment and systems are interconnected and exchange information, towards the improvement of the port's operations flow. Thus, 5G can serve as enabler for full terminal automation, increasing the port's performance and growth potential.

Another aspect that could be further investigated is the development of applications based on 5G technology that allow the real-time communication and collaboration among different terminals. The automation of operations among terminals can be supported by 5G networks and provide increased cargo shipping visibility, through data sharing and exchange. The most important potential of implementing 5G in ports that will be further investigated in future research studies is the capability of ports to provide services that are tailored to specific stakeholder needs and ensure their performance on demand, contributing to positive impact to the port ecosystem and the creation of new synergies. Such use cases can enhance the role of the ports and transform them into innovation hubs for the whole ecosystem they operate into.

Acknowledgements

The COREALIS project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 768994.

References

- Acciaro, M., Renken, K., El Khadiri, N. (2020): Technological Change and Logistics Development in European Ports. European Port Cities in Transition. Strategies for Sustainability. (pp. 73-88).
- Ali, M., Jayakody, D., Chursin, Y., Affes, S., Dmitry, S. (2020): Recent Advances and Future Directions on Underwater Wireless Communications. Archives of Computational Methods in Engineering, vol. 27, 1379–1412.
- Ali-Tolppa, J., Kajó, M. (2020): Mobility and QoS Prediction for Dynamic Coverage Optimization. NOMS 2020 - 2020 IEEE/IFIP Network Operations and Management Symposium, (pp. 1-2). Budapest, Hungary.
- Anker, P. (2019): The vision of 5G and the need for change in mobile spectrum access. Lecture Notes of the Institute for Computer Sciences, Social Informatics and Telecommunications Engineering (LNICST, volume 261).
- Anwar, M. (2019): Digitalization in Container Terminal Logistics: A Literature Review. 27th Annual Conference of the International Association of Maritime Economists (IAME), (pp. 1-25). Athens, Greece.
- Aslam, S., Michaelides, M. P., Herodotou, H. (2020): Internet of Ships: A Survey on Architectures, Emerging Applications, and Challenges. IEEE Internet of Things Journal, vol. 7, no. 10, 9714-9727.
- Bauk, S., Calvo, J., Schmeink, A., Azam, S., Mathar, R. (2019): On deploying vehicular communication at the developing seaport and related innovation success impediments. Transport, vol. 34, no. 1, 126-134.
- Cavalli, L., Lizzi, G. (2020): Port of the Future - Addressing Efficiency and Sustainability at the Port of Livorno with 5G. FEEM Working Paper No. 7.2020.
- Chen, X., He, S., Zhang, Y., Tong, L., Shang, P., Zhou, X. (2020): Yard crane and AGV scheduling in automated container terminal: A multi-robot task allocation framework. Transportation Research Part C: Emerging Technologies, vol. 114, 247-271.
- Gurumurthy, A., Bharthur, D. (2019): Impact of Digitalization in the Ports Sector. Artificial intelligence in India, vol. 5.
- Henriquez, R., Martinez, F., Martinez, J. (2019): IoT-Driven business model innovation. A case-study on the Port of Barcelona in the context of the belt and road initiative. Proceedings of the 16th International Conference on e-Business Engineering (ICEBE 2019), (pp. 302-314).
- Ho, T., Tran, T., Nguyen, T., Kazmi, S., Le, L., Hong, C., Hanzo, L. (2019): Next-generation Wireless Solutions for the Smart Factory, Smart Vehicles, the Smart Grid and Smart Cities. Signal Processing.
- Huh, J. (2020): Reefer container monitoring system using PLC-based communication technology for maritime edge computing. The Journal of Supercomputing, vol. 76, 5221–5243.
- Inkinen, T., Helminen, R., Saarikoski, J. (2019): Port Digitalization with Open Data: Challenges, Opportunities, and Integrations. Journal of Open Innovation: Technology, Market, and Complexity, vol. 5.

- Jiang, H. (2019): Research on Qingdao Port Logistics Development under the Background of “Belt and Road”. IOP Conference Series: Earth and Environmental Science, vol. 330, no. 5.
- Jović, M., Tijan, E., Aksentijević, S., Čišić, D. (2019): An Overview of Security Challenges of Seaport IoT Systems. 42nd International Convention on Information and Communication Technology, Electronics and Microelectronics (MIPRO), (pp. 1349-1354). Opatija, Croati.
- Khan, L., Yaqoob, I., Tran, N., Han, Z., Hong, C. (2020): Network Slicing: Recent Advances, Taxonomy, Requirements, and Open Research Challenges. IEEE Access, vol. 8, 36009-36028.
- Kokkoniemi-Tarkkanen, H., Horsmanheimo, S., Grudnitsky, A., Moisio, M., Li, Z., Uusitalo, M., Samardzija, D., Harkanen, T., Yli-Paunu, P. (2019): Enabling Safe Wireless Harbor Automation via 5G URLLC. IEEE 2nd 5G World Forum, (pp. 403-408). Dresden, Germany.
- Lacalle Úbeda, I., Llorente, M., Palau, C. (2019): Towards environmental impact reduction leveraging IoT infrastructures: the PIXEL approach. Internet and Distributed Computing Systems (pp. 33-45).
- Lambrou, M., Watanabe, D., Iida, J. (2019): Shipping digitalization management: conceptualization, typology and antecedents. Journal of Shipping and Trade, vol. 4.
- Mazzarino, M., Braidotti, L., Cociancich, M., Bottin, G., La Monaca, U., Bretagna, S., Marino, A., Bucci, V. (2019): On the Digitalisation Processes in the Adriatic Region. Proceedings of the 3rd International Conference of Nautical and Maritime Culture -CNM 2019, Naples, Italy.
- Moros-Daza, A., Amaya-Mier, R., Paternina-Arboleda, C. (2020): Port Community Systems: A structured literature review. Transportation Research Part A: Policy & Practice, vol.133, 27-46.
- Muhammad, B., Kumar, A., Cianca, E., Lindgren, P. (2018): Improving Port Operations through the Application of Robotics and Automation within the Framework of Shipping 4.0. 21st International Symposium on Wireless Personal Multimedia Communications (WPMC), (pp. 387-392). Chiang Rai, Thailand.
- Nagendra, A. (2019): 5G and its impact on Supply Chain. Management Vision, vol. 9, 77-91.
- Nikolopoulou, A., Amditis, A., Tsimiklis, G., Tsertou, A., Latsa, E., Krikigianni, E., Lu, M., Tardo, A., Cervera, C. P., Kanellopoulos, I., Hinkka, V., Slingenberg, A. (2019): Sustainable port development: towards the Physical Internet concept. 6th International Physical Internet Conference (IPIC 2019), (pp. 34-43). London, United Kingdom.
- Patzold, M. (2018): Countdown for the Full-Scale Development of 5G New Radio [Mobile Radio]. IEEE Vehicular Technology Magazine, vol. 13, no. 2, 7-13.
- Postolache, O., Yang, Y., Zhong, M., Yao, H., Yu, F., Fu, X., Mi, C. (2019): Modelling IoT technology for a cargo port. Journal of Physics: Conference Series, vol. 1379, no. 1.
- Rost, P., Breitbach, M., Roreger, H., Erman, B., Mannweiler, C., Miller, R., Viering, I. (2018): Customized Industrial Networks: Network Slicing Trial at Hamburg Seaport. IEEE Wireless Communications, vol. 25, no. 5, 48-55.
- Taboada, I., Shee, H. (2020): Understanding 5G technology for future supply chain management. International Journal of Logistics Research and Applications, 1-15.
- Tian, Z., Liu, F., Li, Z., Malekian, R., Xie, Y. (2017): The Development of Key Technologies in Applications of Vessels Connected to the Internet. Symmetry, vol. 9, 211-230.
- Yang, J., Ding, X., Zhang, L., Lin, J. (2019): Design of Intelligent Container Terminal Communication System Based on 5G Technology. Proceedings of the 2019 7th International Conference on Information Technology: IoT and Smart City, (pp. 279–282). Shanghai China.
- Yang, Y., Zhong, M., Yao, H., Yu, F., Fu, X., Postolache, O. (2018): Internet of things for smart ports: Technologies and challenges. IEEE Instrumentation & Measurement Magazine, vol. 21, no. 1, 34-43.
- Yi, Z. (2019): Study on Supply Chain Integration Strategic Model of Guangdong, Hong Kong and Macao Dawan District. 9th International Conference on Information and Social Science (ICISS 2019), (pp. 539-544). Manila, Philippines.
- Zhong, M., Yang, Y., Yao, H., Fu, X., Dobre, O., Postolache, O. (2019). 5G and IoT: Towards a new era of communications and measurements. IEEE Instrumentation & Measurement Magazine, vol. 22, no. 6, 18-26.

Framework and Research Roadmap for a Next-Generation Hyperconnected Logistics Hub

Benoit Montreuil,[1] Leon McGinnis[1,2], Shannon Buckley[1],
Sevda Babalou[1], Wencang Bao[1], and Ali Beranji[1]
[1]Physical Internet Center, Georgia Tech, Atlanta, GA USA
[2]Keck Virtual Factory Lab, Georgia Tech, Atlanta, GA USA
Corresponding author: leon.mcginnis@gatech.edu

Abstract: Today, parcel logistics hubs, where packages come in from many origins and are sorted to their many destinations, are both capital and labor intensive, with capacity that is largely determined by investments in conveyors. In this paper, in the context of Physical Internet growth, we propose a next-generation hyperconnected parcel hub concept that leverages parcel containerized consolidation, does not use conveyors, is robot-centric, with minimal requirement for human operators. Hub capacity can be readily adjusted to accommodate changing logistics patterns. The hub concept is described along with a demonstration case study, the fundamental hub design and operational decisions are identified, and a research roadmap is defined.

Conference Topic(s): Developing the System of Logistics Networks towards the Physical Internet.

Keywords: Physical Internet, Logistics Hubs, Hyperconnected Logistics, Parcel Logistics, Modular Containers, Consolidation, Robotics, Digital Twins

1 Introduction

Today, parcel logistics at both national and international scales is enabled by large-scale parcel logistics hubs, where parcels arrive in bulk from originating hubs, are sorted and consolidated to destination hubs and leave in bulk. The UPS Worldport (https://tinyurl.com/y3bq6tur) is the epitome of such hubs, with 155 miles (250 km) of conveyors occupying 5 million square feet (464,500 m^2), and 2.6 miles (4.2 km) of tilt tray sorters, providing capacity to sort over 416,000 packages per hour. Such hubs were a natural evolution as demand for rapid parcel delivery expanded from its early days. These kinds of parcel logistics hubs, however, present significant challenges. Basic physics limits the potential for expansion, begging the question of how to deal with ever-increasing parcel volumes. They require large amounts of part-time and temporary labor and involve working conditions and hours that are less than desirable. They also engage fixed-capacity transport, which can become inefficient when parcel volumes are not at their peak levels.

This paper addresses the following question: "Can the basic precepts of the Physical Internet (PI) [1,2] be used to develop an innovative parcel logistics hub concept that overcomes the limitations of contemporary logistics hub technologies?" In addressing the question, we assume the existence of some PI technologies that already have been investigated, specifically the use of modular containers [3] to consolidate small parcels for shipment from their originating hubs and delivery to their destination hubs. For example, we rely on two types of modular-sized handling containers, termed totes for those with sides up to near 2 feet, and boxes for those with larger sides of 2, 4 or 8 feet. We rely on modular mobile racks, whose external side dimensions are consistent with those of boxes, to store and transport totes. We further assume the existence of specific robotic technologies similar to those currently available.

The paper is organized as follows. Section 2 presents the basic concept for the innovative parcel logistics hubs using modular totes and boxes to consolidate parcels for handling, and modular mobile racks to consolidate totes for transport. Section 3 provides a system architecture definition for a particular realization of the robotic logistics hub concept and section 4 summarizes initial design of the boxes, totes and racks, consistent with standard truck and trailer dimensions. Section 5 describes in detail the "Shuffle Cell", a robotic cell in which totes are shuffled between racks to improve the racks' levels of consolidation. In section 6, a complete hub design is described, based on a case study from an existing parcel logistics hub. A detailed Anylogic™ simulation is described in section 7 along with some initial computational evaluation of the hub design. Section 8 compares the results from simulating the new hub design to the observed performance from the original benchmark hub. Finally, section 9 discusses a research roadmap for this novel approach to parcel logistics.

2 Basic Concepts

The idea of hyperconnected logistics is described in several publications, including [1, 4-6] and several previously explored PI concepts motivate this work. One is the consolidation of parcels into modular containers as discussed in [3, 7-9]. An originating hub will consolidate all parcels with the same destination into one or more modular totes or boxes. In the hub concept presented here totes are consolidated into racks for transport in standard trucks or semi-trailers.

A second motivating concept is a network of logistics hubs, as suggested in [1, 2, 5, 10]. The trucks from the originating hub connect to a logistics hub that is part of this network. At each logistics hub, an arriving rack gives up totes that are not going to the same next destination as the rack and acquires totes that are until the rack is fully consolidated and ready for transport to a next hub.

At an originating hub, parcels with same destination are accumulated in a tote. A tote in a rack from a particular originating hub may visit several logistics hubs and be transferred to other racks before it finally arrives at its destination hub where parcels are removed from totes. At each logistics hub, some totes may be removed from a rack, because their appropriate next destination is different from the rack's next destination, and some totes may be added. It is possible that a rack arriving to a logistics hub will be stripped of totes and stored temporarily because it is not needed at the moment to transport totes, and all the totes it contained can be accommodated by other racks.

A third essential concept is robotic transport of both racks and totes within the hub, thus dramatically reducing the numbers of humans involved in the logistics hub operations, in line with robotic mobile fulfillment systems [11]. Further, the resources for moving totes between hubs and for temporarily storing racks between these processes are organized according to a standard footprint, resulting in easily replicated cells.

This is a fundamentally new paradigm for parcel logistics. Rather than concentrating the capacity to sort parcels, it distributes that capacity across a hub network. Because the concept depends upon robotic technology, it is readily scalable; in fact, capacity at a given logistics hub can be adjusted as flow through the hub increases *or decreases*. Further, because capacity is based on robotic technology, the number, size and location of logistics hubs can be changed much more easily than in current parcel logistics systems.

We refer to this new type of logistics hub as a *hyperconnected logistics hub* or HLH.

3 HLH System Architecture

There are three main areas of operation in the HLH, unloading inbound racks, shuffling totes between racks to achieve consolidation, and loading outbound racks as shown in *Figure* 1. Note that hereafter, for conciseness purposes, we consider boxes to be racks that simply have to be crossdocked. Physically, unloading and loading may share the same docks but operationally they are different and may have different priorities. Both the loading and unloading centers have a staging area where racks may be located after unloading or before loading. The *ShuffleCenter* contains an area where racks may be stored temporarily (*BufferZone*) and a number of cells where totes are shuffled between racks (*ShuffleCell*s). The figure also shows the flows between the three areas of operation as well as within the *ShuffleCenter*.

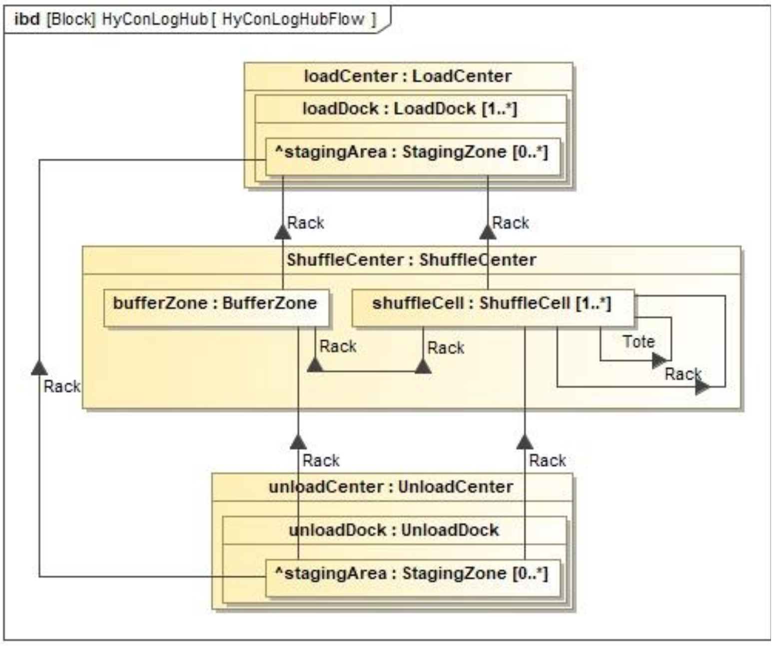


Figure 1 HLH Areas of Operation

In the HLH there are four distinct types of robots:

- *LoadBots* are capable of un/loading racks from/to trucks
- *MoveBots* move racks from/to staging areas, and within the *ShuffleCenter*
- *ShuffleBots* move individual totes between racks
- *ToteBots* move individual totes between *ShuffleCells*

As described in Section 4, the totes and racks are designed to maximize the utilization of standard semi-trailers and thus the racks are set flat on the floor of the trailer. Thus, *LoadBots* must be able to engage the racks, lift and move them. We assume the racks either have retractable "feet" or are placed on a specialized rack stand for movement within the HLH. Thus, *MoveBots* operate in a manner like conventional Kiva-style robots [11]. *ShuffleBots* are specialized for extracting/inserting totes from/to racks and moving totes between rack locations. Finally, *ToteBots* are conceptualized as small, quick robots that can transport individual totes between *ShuffleCells*. *ShuffleBots* can interface with *ToteBots* for their loading and unloading.

An agent-oriented control architecture for the HLH is shown in *Figure 2*. There are 11 distinct control domains. At the highest level of control is the *HLHAgent* that has visibility to the inbound trucks and the state of each of the control domains with which it directly interacts, namely the *UnloadCenter*, the *LoadCenter*, the *LoadBotPool*, and the *ShuffleCenter*.

The *LoadBotPool* has a control agent responsible for moving *LoadBots* between the un/loading operations. The *UnloadCenter* and *LoadCenter* are conceptually similar; each has a control agent, and each has an assigned pool of *LoadBots*. The center level agents are responsible for prioritizing the load/unload operations and the associated bot pool agents are responsible for managing the *LoadBots* to execute the prioritized loading and unloading.

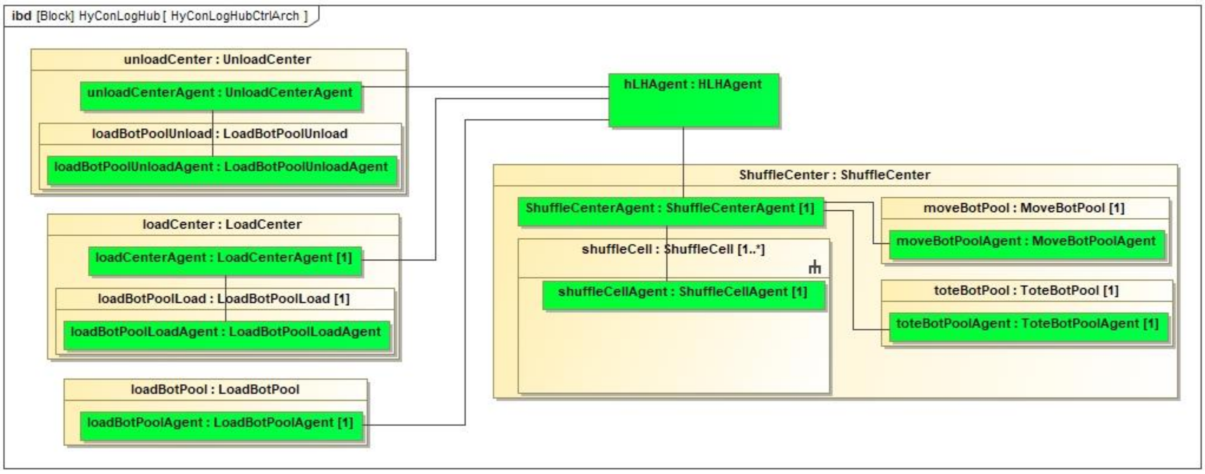


Figure 2 HLH Control Architecture

The *ShuffleCenter* has a control agent that directly interacts with the control agents for the *ShuffleCells*, *MoveBotPool*, and *ToteBotPool*. The *ShuffleCenterAgent* determines when racks should be moved and where to move them, and the *MoveBotPoolAgent* manages the execution of these moves. Similarly, the *ShuffleCenterAgent* determines when totes should be moved between *ShuffleCells* and the *ToteBotPoolAgent* manages the execution of the tote moves. Not shown in Figure 2 are the individual *MoveBot* and *ToteBot* agents that manage the execution of assigned rack and tote moves within the *ShuffleCenter*.

Within each *ShuffleCell*, the *ShuffleCellAgent* is focused on operations within the cell and determines which tote moves should be made. Possible tote moves include rack-to-rack, rack-to/from-buffer, rack-to/from-*ToteBot*, and buffer-to/from-*ToteBot*. The *ShuffleCellAgent* determines the sequence of moves and the *ShuffleBotAgent* manages the *ShuffleBot* execution of moves.

The execution of rack and tote moves are relatively straightforward and managed by the corresponding bot agents. More interesting are the operational decisions made by the agents for the HLH, the *ShuffleCenter*, the individual *ShuffleCells* and the various bot pool agents. The control architecture in *Figure 2* allows these decisions to be made using methods from simple heuristics to very sophisticated optimizations.

4 Initial Design of Totes and Racks

To make the HLH concept concrete, the totes and racks must be given physical configurations. Here we assume a single size with nominal dimensions of 2x2x2 feet, although other configurations are possible. We assume racks to have capacity for eight such totes, four high and two wide. Suppose the racks are to be transported in standard semi-trailers, with inside dimensions of 47’3”x99”x108-1/2”and rear door dimensions of 8’3"x8’9”. If the racks are placed in the trailer side by side, as shown in Figure 3, and we wish to use as much as possible of the trailer volume, then the external dimensions of the racks will be 1’10-1/2”x4’x8’4-3/4”.

Alternative designs have been considered, based on placing the racks in the semi-trailer aligned with the long dimension of the trailer, i.e., back-to-back rather than side-to-side. The two designs achieve slightly different volume utilizations within the trailer, and result in different configurations of the load/unload staging areas because of the nature of the connection between the *LoadBot* and the rack.

Assuming internal structural members with dimensions 4”, and allowing for ½” clearance around the totes, the resulting dimensions for a single standard tote would be 1’9”x1’9”x1’9”.

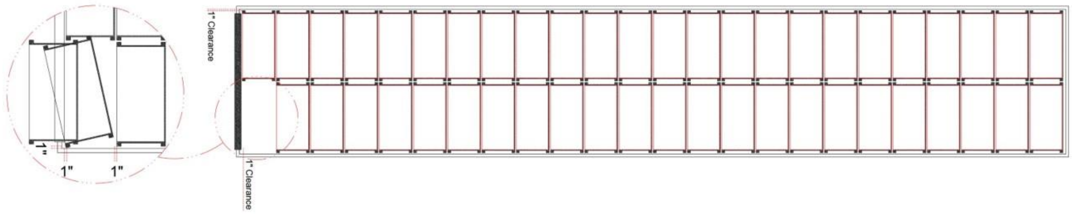


Figure 3 Racks Arranged in Semi-Trailer

5 Shuffle Cell Concept and Design

Conceptually, the ShuffleCell comprises a set of racks, ToteBot interfaces and a ShuffleBot that moves totes. The ShuffleCell operates on racks and totes but does not determine which racks or totes are assigned to it. Given a set of racks, constraints and priorities, the ShuffleCellAgent determines the specific tote moves to be made and their sequencing. From an operational perspective, avoiding deadlock conditions ([12, 13]) can be accomplished if there is always at least one empty location where totes can be placed to create opportunities for consolidation. Within this concept, the following must be determined: the number and arrangement of the racks, the ToteBot interfaces, the ShuffleBot specifications, and the operational control of the cell.

To make the concept concrete, consider a ShuffleCell configuration as shown in Figure 4. Eight racks are arranged in two rows, with space for a ShuffleBot to move between the two rows. There are locations for ToteBots that are accessible by the ShuffleBot. For this kind of configuration, the number of racks could be different from that shown. More racks would increase the chances for improving tote moves, but also might increase the dwell time of racks in the ShuffleCell.

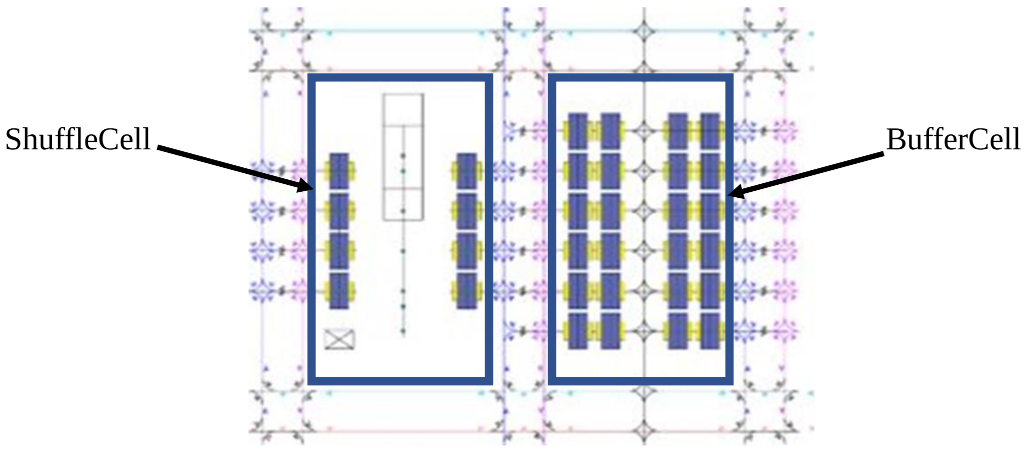


Figure 4 Example of a ShuffleCell (on the left) and BufferCell (on the right)

Conceptually, the ShuffleCell requires no “hard” physical infrastructure. The ShuffleBot is not confined to a rail and the racks do not require fixtures in the floor. Given the physical configuration illustrated in Figure 4 there are many different possible operational strategies, and the best strategy will likely depend on the nature of the arriving racks, i.e., how much consolidation already has been accomplished before the racks reach the HLH.

6 HLH Modular Design and Layout

We have designed modular components for each functional area to facilitate the overall facility design process. Figure 5 is a snapshot from a particular HLH simulation, showing the physical arrangement of the HLH. Docks are arranged around the perimeter of the building, and there is a staging area associated with each dock. The grid in the center corresponds to the *ShuffleCenter* and is configured by aisles along which the *MoveBot*s and *ToteBot*s can travel. Each space between aisles has the same footprint and can be used for either a *ShuffleCell* or a *BufferCell*. In Figure 4, the fifteen spaces on the left edge of the *ShuffleCenter* and the fifteen spaces on the right edge are used to store racks either before they first move to a *ShuffleCell* or after they have completed the consolidation process but are not yet ready to be moved to an outbound staging area.

There are thirty-five cells in the middle of the *ShuffleCenter*, of which thirty are allocated for use as *ShuffleCell*s and five for use as *BufferCell*s. *ShuffleCell*s are only active when there is a need for them, and the figure indicates that five of the potential cells are not in use when the snapshot was taken. Similarly, only two of the *BufferCell*s contained racks when the snapshot was taken.

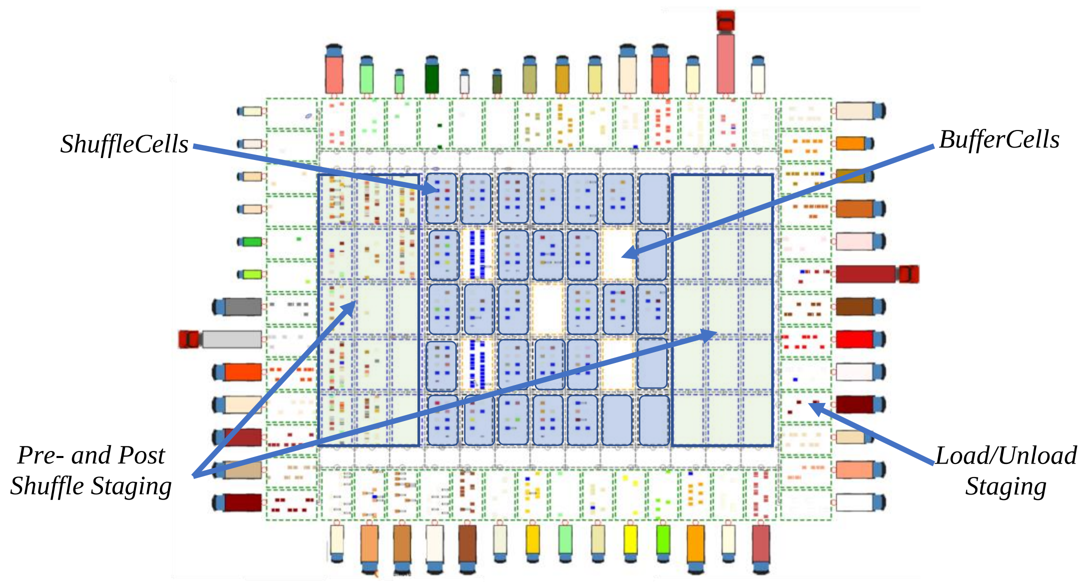


Figure 5 Simulation Snapshot for a Particular HLH Configuration

In fact, the cellular designs of the *ShuffleCell* and *BufferCell* admit a tremendous range of options with regard to the physical flow within the HLH. In Figure 5, an operational decision determines which cell location is activated “next” during the HLH operations, constrained by the predefined function (pre-shuffle buffer, in-process buffer or *ShuffleCell*). Clearly, this is not the only or even necessarily a good way to operate the HLH. For example, individual cell locations within the *ShuffleCenter* could have a predefined function, as in Figure 5, or the function of a cell location could be determined as needed. The numbers of cells of each type could be varied, as could the strategy for determining which locations to activate over the course of a day. Much more detail about the modular design is given in complementary paper [12].

7 HLH Simulation

Assessing the potential of this innovative parcel logistics concept requires a high-fidelity simulation model. Essential requirements for the simulation are the accurate representation of the operations within the *ShuffleCell*, the movements of totes and racks within the HLH, and the arrival and departure processes of trucks containing racks.

In designing the HLH simulation, three considerations were of utmost importance. First, it must be easy to modify the HLH configuration, changing the numbers of *ShuffleCell*s and *BufferCell*s and their physical arrangement. Second, it must be easy to modify the operational decision-making to accommodate experimentation with different strategies, policies and priorities. Third, it must be possible to display a visual representation of the HLH operations.

To meet the first requirement, we developed a layout specification in Excel that is processed through a Python™ script to create input to AnyLogic™ for the layout of the HLH. An example of the Excel specification is shown in Figure 6, corresponding to the simulated layout shown in Figure 5. Each cell in the spreadsheet corresponds to a standard square of defined dimension in the HLH. Cell entries indicate either a boundary of a cell or a segment of bot flow path.

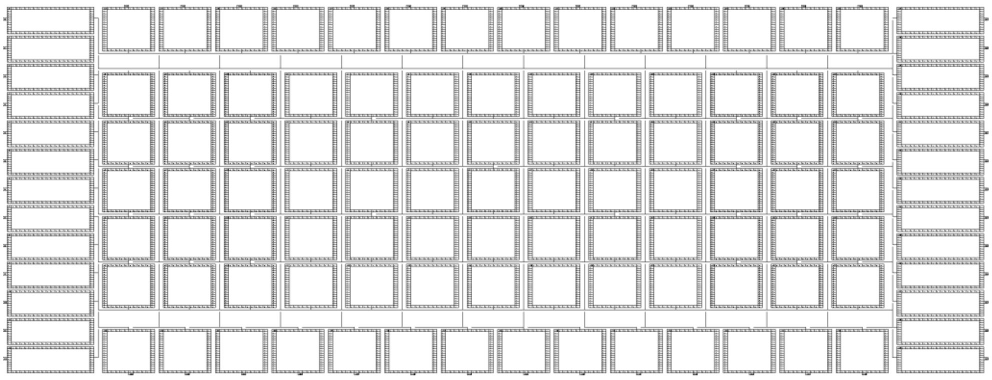


Figure 6 Example of layout specification in Excel

Our HLH simulation carefully separates modeling of the physical operations from modeling of the operational control decision making as shown in Figure 7. As discussed in much greater depth in [13], the key concept is to use state charts to model the control logic (i.e., deciding which totes or racks to move and the location to which they should be moved) and to use queues

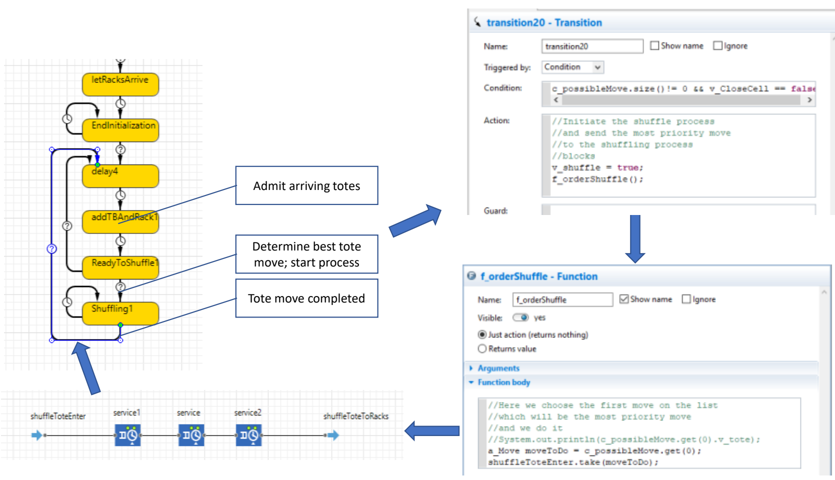


Figure 7 Plant-control separation in AnyLogic

and servers to model the execution of physical processes, such as moving totes or transporting totes or racks between cells.

In AnyLogic™, Java-defined state chart conditions can control the transitions between states; and Java-defined actions can be executed upon entering or leaving a state. As an example, in Figure 7, the transition between states *ReadyToshuffle* and *Shuffling* occurs when the condition on the transition is satisfied by finding the best tote to move, which is done using a Java-defined function. When there is a tote to move, the condition also sends a corresponding simulation agent through the simulated "shuffle process" where the delay associated with the tote move is realized. When this simulation agent departs the service2 block, the condition is satisfied on the transition from the state *Shuffling* to the state *delay4* and the control loop repeats.

The third requirement, focused on visualization, is accomplished in two ways. First, animation shows the movement of totes in shuffle cells and the movement of racks within the *ShuffleCenter*. Figure 5 is a snapshot of the hub-level animation. Second, there are several dashboards that show real-time results from the simulated operations. Figure 8 shows part of the full dashboard, tracking the number of open *ShuffleCell*s (top) and the number of racks in pre- and post-shuffle staging and in the in-process buffers.

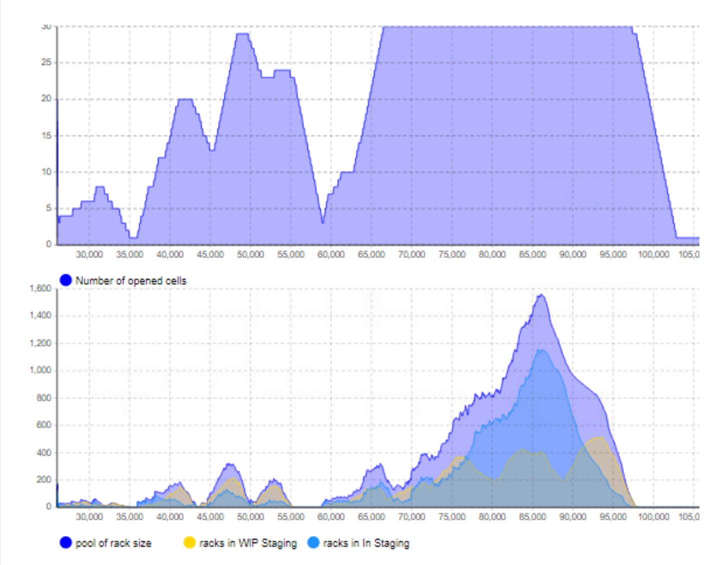


Figure 8 Dashboard Elements Showing ShuffleCells and Rack Pools

8 HLH Comparison to Conventional Hub

Our assessment starts with detailed data from an existing parcel hub, describing the arrival and departure of trucks and their parcel contents. This parcel hub handles on the order of 400,000 parcels per day. We "containerized" the parcels to convert the conventional parcel loads into corresponding loads of racks containing totes containing parcels. This resulted in approximately 72,000 totes passing through the simulated HLH per day. The baseline parcel arrival-and-departure information constitutes the performance baseline against which the simulated HLH is to be compared. In particular, the fundamental question is "Can the HLH satisfy the cutoff times for arriving parcels to be leaving the hub in a departing truck and do so efficiently?"

Our initial implementation of the *ShuffleCell* identifies racks in the cell as either "strip"—meaning the rack will only give up totes—and "stack"—meaning the rack will only receive totes. A strip rack may become empty or if not, it may return to a WIP *BufferCell*. It also may be redesignated as a stack rack if all remaining totes have the same next destination. Simulation parameters allow varying the number of strip and stack racks in the *ShuffleCell*s.

Our initial experimentation is based on the HLH configuration shown in Figure 5 and reasonable parameters for the various bot operations, as well as the parcel throughput from the baseline use case. Clearly a great deal of development and experimentation remains to be done. Given that caveat, we can report that the HLH with 25 *ShuffleBot*s successfully meets the parcel departure cutoffs from our benchmark parcel hub. Future papers will provide much more in-depth results from systematic experimentation of HLH performance.

9 Research Roadmap

Based on our very preliminary results, the HLH concept holds significant promise for revolutionizing parcel logistics. There are many research challenges arising from the initial proof of concept analysis and simulations. Clearly, a major category of research and development challenge is robotic technology *per se*. Given the robotic advances of the recent past, we are confident those challenges can be identified and resolved by the robotics community. Our focus in this section is on the logistics-related design and operational challenges.

Detailed design of shuffle cells. The proof-of-concept design allowed eight racks to be in a shuffle cell and assumed a specific type of shuffle robot. Overall, the best shuffle cell layout remains an open question. Moreover, the *ShuffleCell* concept as we have presented it here is only one of several, or perhaps many possible ways to employ robotics and automation to manage the destination-based consolidation of containers in HLH.

Overall consolidation strategy. In the proof-of-concept simulation, simple heuristics were used to choose racks to move into and out of shuffle cells. It seems likely that a more intelligent approach will yield better results in terms of the average time to consolidate a rack, and thus perhaps reducing the number of *ShuffleCell*s required. For example, should the shuffle cells be arranged logically into levels ranging from initial consolidation to final consolidation, where racks become more consolidated as they move through the levels? Are there other strategies that would significantly improve the time to consolidate?

Staging area design. In the proof-of-concept design, the staging area was dedicated to dock doors. It seems clear that a shared staging area approach could significantly reduce the overall staging capacity requirement, although it might impact the required number of *MoveBot*s. How to specify these capacities and how to operate the *LoadBot*s are interesting areas for investigation.

Layout. Given a shuffle cell specification, a fundamental issue is the physical arrangement of resources in the HLH, i.e., the locations of *ShuffleCells*, *StagingZones*, *BufferZones*, and dock doors. We showed only one possible configuration in the simulation results but clearly many alternatives remain unexplored. Also, we used very simple heuristics for assigning inbound/outbound trucks to docks, a decision that interacts with *ShuffleCenter* layout and operation in a very significant way to impact the total *MoveBot* travel.

Parcel de-containerization. In the proof-of-concept, boxes and totes arriving into the HLH do not have to be opened to reassign parcels to other totes or boxes, assuming adequate parcel-to-container assignment. In practice, specially at the lower-tier logistics mesh networks, there is significant probability that such sorting of some parcels into other containers may be pertinent for improving their consolidation in the next part of their journey. This requires integrating a *SortingCenter* into the HLH, potentially with *SortingCells* similar to order picking cells in goods-to-person fulfillment centers. Integrating such a *Sorting Center* has impact on the layout as well as the control of the HLH, and is a rich avenue for further research.

In summary, the HLH concept clearly is worthy of further investigation, and presents a broad range of research opportunities, both for the robotics community and for researchers focused on the strategic, tactical, and operational issues entailed in the realization of the hyperconnected logistics system and its hyperconnected logistics hubs.

References

1. Montreuil B. (2011): Towards a Physical Internet: Meeting the Global Logistics Sustainability Grand Challenge, Logistics Research, v3, no2-3, 71-87.
2. Ballot, É, B. Montreuil, R.D. Meller (2014), The Physical Internet: The Network of Logistics Networks, La Documentation Française, Paris, France, 205p.
3. Montreuil B., E. Ballot, W. Tremblay (2016). Modular Design of Physical Internet Transport, Handling and Packaging Containers, Progress in Material Handling Research Vol. 13, Ed. J. Smith et al., MHI, Charlotte, NC, U.S.A. https://digitalcommons.georgiasouthern.edu/pmhr_2014/1/
4. Crainic T. G. & B. Montreuil (2016). Physical Internet Enabled Hyperconnected City Logistics, Transportation Research Procedia – Tenth International Conference on City Logistics, v12, 383-398. https://doi.org/10.1016/j.trpro.2016.02.074
5. Montreuil B., S. Buckley, L. Faugere, R. Khir & S. Derhami (2018). Urban Parcel Logistics Hub and Network Design: The Impact of Modularity and Hyperconnectivity, Progress in Material Handling Research: 2018, Ed. A. Carrano et al., MHI, Charlotte, NC, U.S.A. https://digitalcommons.georgiasouthern.edu/pmhr_2018/19/
6. Kim, N., B. Montreuil, W. Klibi, N. Kholgade (2020). Hyperconnected Urban Fulfillment and Delivery, Transportation Research Part E Logistics and Transportation Review, Vol. 145, 23p. https://doi.org/10.1016/j.tre.2020.102104
7. Gazzard N. & B. Montreuil (2015). A Functional Design for Physical Internet Modular Handling Containers, 2nd International Physical Internet Conference, Paris, France, 2015/07/06-08, 19 p.
8. Landschützer C., F. Ehrentraut & D. Jodin (2015). Containers for the Physical Internet: requirements and engineering design related to FMCG logistics. Logistics Research, v8.
9. Sallez Y., S. Pan, B. Montreuil, T. Berger & E. Ballot (2016). On the Activeness of Intelligent Physical Internet Containers, Computers in Industry, v81, 96-104.
10. Montreuil B., R.D. Meller & E. Ballot (2013). Physical Internet Foundations, in Service Orientation in Holonic and Multi Agent Manufacturing and Robotics, ed. T. Borangiu et al., Springer, p. 151-166. https://link.springer.com/content/pdf/10.1007%2F978-3-642-35852-4_10.pdf
11. Renan de Costas Barros, I. & T. Pereira do Nascimento (2021). Robotic Mobile Fulfillment Systems: A survey on recent developments and research opportunities, Robotics and Autonomous Systems, 103729.
12. Yoon, H J. (2010). Scheduling for Deadlock Avoidance Operation in Robotic Manufacturing Cells. Proceedings of the Institution of Mechanical Engineers. Part B, Journal of Engineering Manufacture, vol. 224, no. 2, SAGE Publications, pp. 329–40, doi:10.1243/09544054JEM1422.
13. Reveliotis, S A, and Roszkowska, E. (2011). Conflict Resolution in Free-Ranging Multivehicle Systems: A Resource Allocation Paradigm. IEEE Transactions on Robotics, vol. 27, no. 2, IEEE, pp. 283–96, doi:10.1109/TRO.2010.2098270.
14. Babalou S., W. Bao, A. Barenji, B. Montreuil, L. McGinnis, S. Buckley (2021). Modular and Mobile Design of Hyperconnected Parcel Logistics Hub, Proceeding of IPIC 2021 International Physical Internet Conference, 10 pp.
15. McGinnis L., S. Buckley, A. Barenji (2021, to appear). Designing and Implementing Operational Controllers for Robotic Tote Consolidation Cell Simulation, Proceedings of the 2021 Winter Simulation Conference, S. Kim, B. Feng, K. Smith, S. Masoud, Z. Zheng, C. Szabo and M. Loper, eds.

Physical Internet inspired Atomic Modeling for Supply Chain Risk Management

Thibaut Cerabona[1], Frederick Benaben[1,3], Jean-Philippe Gitto[2], Matthieu Lauras[1,3] and Benoit Montreuil[3]

1. Centre Génie Industriel, IMT Mines Albi, Campus Jarlard, 81013 Albi CT Cedex 09, France
2. Scalian, 17 Avenue Didier Daurat, Batiment Pythagore, 31700 Blagnac, France
3. ISyE, H Milton Steward School of Industrial and Systems Engineering, 755 Ferst Drive, Atlanta, GA 30332, USA

Corresponding author: thibaut.cerabona@mines-albi.fr

Abstract: *At a time when instability is the norm, as the global health situation confirms, managers have to deal with increasingly complex situations. Managers expect to have decision support tools that allow them to manage this instability in order to suffer as little as possible. Simulation is one of the main tools to meet this demand. This paper presents the work in progress for the development of a modular supply chain simulation model inspired by the Physical Internet (PI). Its modules are developed from the processes defined in the Supply Chain Operations Reference (SCOR) tool. This simulation model will generate the inputs to apply Physics Of Decision (POD) approach, an innovative approach to risk management approach that draws on analogies with physical forces. This approach is dedicated to steering the performance trajectory of systems evolving in an unstable environment.*

Conference Topics: *Business models for open & interconnected logistics, modularization and PI Modelling and Simulation.*

Keywords: *supply chain management, risk management, performance measurement, SCOR, system modeling and physical internet.*

1 Introduction

The observation that instability is the norm made in Benaben et al. (2021) makes sense today in view of the global health situation. Supply chains are not exempt from this instability, they are even all particularly sensitive to it. Because of its network structure, connecting independent entities that do not always have the same objectives at the same time, supply chains evolve in uncertain, dynamic and risky environment. Thus, supply chain is a complex system. Many factors, as mentioned in Serdarasan (2013), increase this complexity: globalization, sustainability, customization, outsourcing, innovation and flexibility. Supply chain managers have to deal with increasingly complex situations. To manage these situations and instability, mangers need information about the system, its environment and the supply chain network organization. But they also need tools to identify the causes and assess the impacts of this instability on supply chain activities (Cope et al., 2007). Among the many tools and methods developed to help managers dealing with instability, the solution studied in this paper is simulation. However, modeling complex systems like supply chains requires a lot of time, resources and energy. Often, much of this effort for one model cannot be reused when creating another model. Therefore, this is where the use of standard reference models like SCOR becomes important. By providing standard processes and performance indicators that can be adapted to any type of supply chain, SCOR should build simulation models more efficiently and quickly.

This article aims to present the first developments of an "atomic" simulation model based on SCOR and inspired by the concepts of the Physical Internet. This simulation model will be used to generate the data needed to feed the POD approach, a physics-based approach dedicated to risk management and decision support. The fundamentals of the POD approach will be presented and illustrated by a case study of a fictitious aeronautical supply chain impacted by the Covid-19 crisis.

This article is structured as follows: section 2 provides a quick overview of existing tools and methods in the field of risk management. Section 3 presents the POD approach. Section 4 describes the application of this approach through the study of a fictitious aeronautical supply chain. Section 5 presents the first work done on a modular simulation model of a supply chain. Section 6 proposes an extension of this atomic vision to help address some of the challenges of the Physical Internet. Finally, section 7 concludes with a roadmap to turn that atomic modelling workable.

2 Background: Risk Management

During the last decades, researchers have provided numerous results and contributions in the field of risk management. One of them is the risk management process, which is typically divided into three or four steps as presented in White (1995) and Tummala and Schoenherr (2011):

- Risk Identification: determines the potential risks that may impact the system by studying the system and its environment,
- Risk Assessment: estimates the probability of occurrence and the impact of the risk in order to estimate its significance,
- Risk Response Strategies: identifies the actions and strategies to be implemented to limit the impact of risks,
- Risk Monitoring: monitors the evolution of the status of identified and treated risks.

The ISO 31000 standard, identified more than thirty tools focused on the risk assessment phase (de Oliveira et al., 2017). White (1995) in her comprehensive literature review gives an overview of the most known and used techniques: Failure Mode and Effects Analysis (Clifton, 1990), Fault Tree Analysis (Bell, 1989), Hazard and Operation Study (Hambly and Hambly, 1994), Cost Benefit Analysis (Lave 1986), Sensitivity Analysis (Covello 1987), Hertz-type simulation (Ho and Pike, 1992) and Monte Carlo simulation (Merkhofer, 1986).

According to White (1995) and Benaben et al. (2021), the study of these methods allows us to draw the following conclusions: risk response strategies and risk monitoring phases do not seem to be treated satisfactorily. Thus, half of the risk management process (two phases out of four) is weakly covered. The first two phases, risk identification and risk assessment, gather the necessary material to enable decision makers to decide. The last two phases, risk response strategies and monitoring, rely on the analysis and skills of the decision maker.

3 Physics of Decision Approach

This article presents the Physics of Decision (POD) risk and opportunity management approach, introduced in Benaben et al. 2021. This approach is founded on analogies with physics in order to help supply chain managers deal with instability and to support them in their decisions. Several pieces of information are essential for making decisions in an unstable environment, such as an understanding of the system being managed and its environment, the potential consequences generated by instability and change, and the mechanisms for selecting the best possible alternatives (Benaben et al., 2021).

In order to provide managers with all this information in a single tool, the POD approach is based on two modeling spaces: description and performance spaces. The **description space** allows to describe the studied supply chain and its environment, by positioning the supply chain in the framework of its attributes. Figure 1 illustrates this description space whose dimensions are its attributes. As part of the study of a supply chain, the attributes retained can be for example: the lead times, the number of suppliers, the price of raw materials, etc. The value of the attributes varies over time and according to decisions made by managers. The control space (blue shape) allows to consider these changes, by defining freedom constraints on each attribute. This sub-space is therefore in constant evolution. In this subspace, the studied supply chain can move freely. In certain areas of this description space, modeled by the context characteristics (orange shape), the system is more sensitive to the following potentials:

- Environment: includes all the potentials generated by the system's environment (e.g. new taxes related to the Brexit, etc.),
- Charges: represent the mandatory costs for the proper functioning of the system (e.g. wages, energy expenses, etc.),
- Innovations: represent all the actions and measures put in place to enhance the system (e.g. purchase of new equipment, etc.),
- Interactions: reflect all the potentials created by the interactions between the network actors (e.g. customer demand, lead times, etc.).

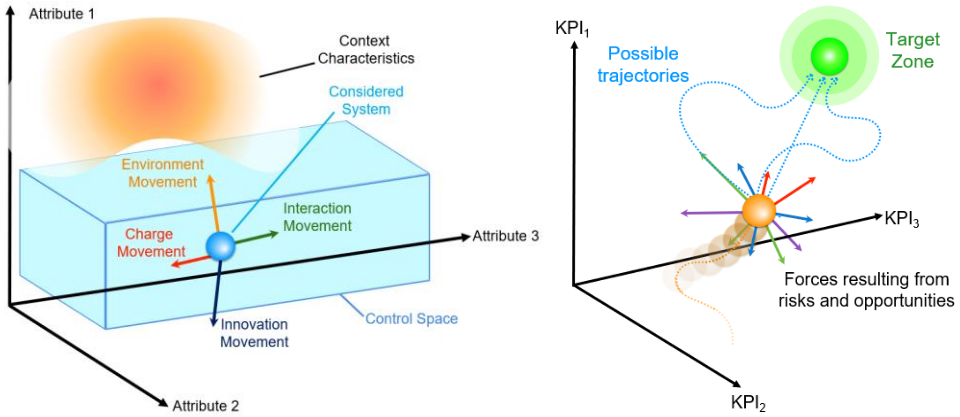


Figure 1: Description (left) and Performance (right) spaces (from Cerabona et al., 2020)

The **performance space** (illustrated in Figure 1) depicts the evolution of the supply chain performance. It allows to position the supply chain in relation to its current performance. The interest of this space is to monitor the evolution of the performance over time by obtaining the performance trajectory (Benaben et al., 2021). The supply chain performance evolves following the occurrence of disruptions, whose impact on the performance is designed by forces (color vectors). Schematically, this impact can be observed by comparing the performance trajectories of the perturbed system with the inertial trajectory. It reflects the initial performance trajectory of the supply chain subjected to no perturbations and serves as the basis for this study. In this approach, this framework supports managers in their decisions (Cerabona et al., 2020), in particular in the choice of the best possible alternatives to join the target zone at the lowest “costs” (for instance money, time, etc.). The target zone is an area of this performance space representing the performance to be achieved in order to be in line with the performance objectives set. In Figure 1, the target zone is represented by a green sphere but it could be a surface, a subspace, etc.

4 Application to Supply Chain Risk Management

4.1 Studied Supply Chain

Let's consider a fictitious aeronautical supply chain (called Air-POD) to illustrate this approach. This supply chain is strongly inspired by one of the supply chains of a famous European aircraft manufacturer. For this example, only the last phase of manufacturing an aircraft is considered, i.e. the assembly of the constitution assemblies. Figure 2 describes the organization of this supply chain network, composed of eight partners. A and B are the two assembly sites. B also supplies some fuselage sections to A. C, D, E, F and G are the suppliers of A and B, they supply respectively the vertical stabilizer, the cockpit, the horizontal stabilizer, the wings and the engines. H is the customer of A and B. The use-case presents a very simplified version of an aeronautical supply chain (for example, only one customer is considered in this model because the processes allowing to customize an aircraft are not considered currently), but includes a large potential of evolution and complexification. The associated model and processes have been developed on the Anylogic© simulation software.

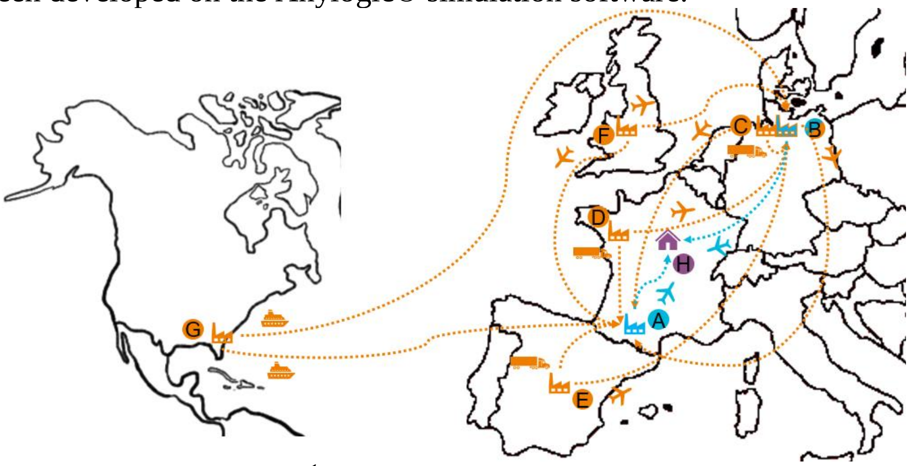


Figure 2: Air-POD network[1]

4.2 Experiments and Results

In this application example, the experiments carried out focus on the study of the impact of predefined risks on the performance of this supply chain. In order to facilitate the visualization of the results, the description and performance spaces will be limited to three dimensions. The description space will be built from the following attributes: the sales price of the airplanes, the monthly production and the number of employees assigned to the assembly of the airplanes on the two production sites (A and B). The performance space will have the following dimensions: the number of planes sold, average cycle time and capital asset. The description space and the inertial performance trajectory of Air-POD supply chain are illustrated in Figure 3.

[1] The maps used in this figure are taken from the websites: https://www.pinterest.fr/pin/497295983852497660/ for the European map and https://www.waouo.com/carte-du-monde-vierge/ for the map of America.

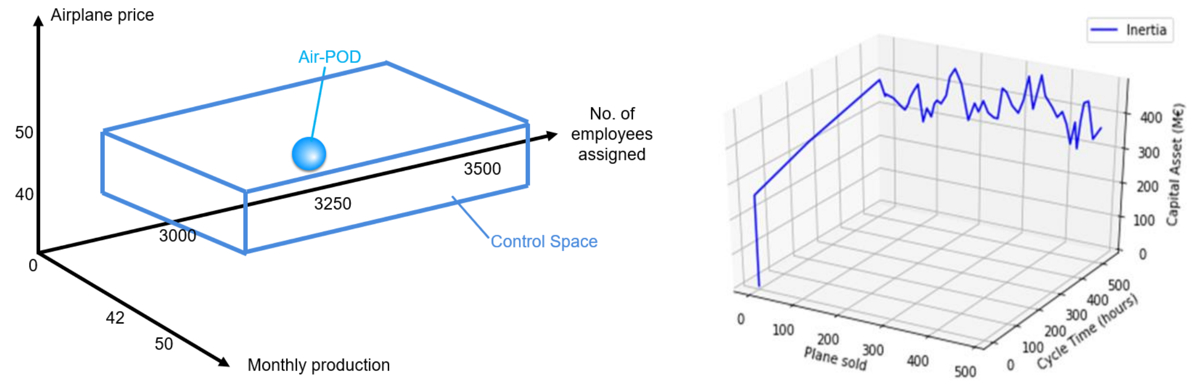


Figure 3: Air-POD description (left) and performance space (right)

Two disruption scenarios (DS) are studied, these two scenarios are used to evaluate the impact of the Covid-19 crisis on the performance of this supply chain. The results associated with the studied scenarios focus mainly on the visualization of the Air-POD performance trajectories and the study of the deviations they generate (these results are correlated with the realism of the model). DS1 is relative to a demand that is brutally divided by two. Logically, the number of aircraft sold has decreased. This disruption reduces the downtime costs by 27%, the average cycle time is reduced by about 13%. This reduction in cycle time is normal, in this scenario the available resources are not correlated to demand. DS2 represents a decrease of 10% in the number of employees available due to Covid-19 (e.g. due to illness, childcare, etc.). For this scenario, all of these lost human resources are not replaced. The performances are identical over the first six months simulated. This disruption slightly increases the average cycle time by seven hours and the average capital asset by about 2%. The differences in the number of aircraft sold can be explained by all the stochastic factors present in this model. The performance trajectories of the perturbed system are shown in Figure 4.

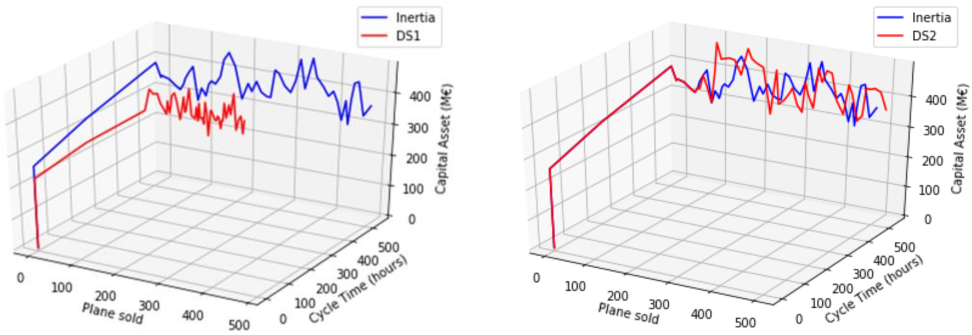


Figure 4: Inertial trajectory versus disruption scenarios trajectories

5 Atomic Supply Chain Model based on SCOR

The POD approach, presented in the section 3, makes it possible to monitor the evolution of performance following the realization of risks and opportunities, especially by measuring their impacts and benefits. However, in its current state, this approach does not allow decision-makers to measure and understand the micro-consequences created by these macro disturbances impacting the studied supply chain, its actors and processes. The study of these micro-consequences is a key component to effectively analyze a system as complex as a supply chain network. One way to solve this problem and bring this information to managers is to create a simulation model of the elementary processes, atoms, of a supply chain. This model is developed from the SCOR model in order to benefit from all the performance indicators, standard processes adaptable to any type of supply chain and their hierarchical organization, that this model provides. Moreover, building on SCOR also allows us to develop a model recognized and approved by academics and professionals working in the field of supply chain management. According to Fayez et al. 2005, SCOR is the only model that this community widely accepts.

5.1 SCOR – A hierarchical process model

The SCOR model provides an inter-process (horizontal) and hierarchical (vertical) breakdown of processes (Huang et al., 2005). This process structuring aims to improve the understanding of all supply chain processes in order to improve its management and performance. Thus, the defined processes are divided into four hierarchical levels (Persson, 2011). Figure 5 illustrates this process decomposition. At level 1, six macro processes are defined:

- Plan: coordinates the actions of the other five processes to meet business objectives,
- Source: includes the processes that enable the supply of products and services to satisfy expected or current demand (Huang et al., 2005),
- Make: is the set of activities and processes that transform raw materials into finished products,
- Deliver: includes all the processes that enable the delivery of finished products or services corresponding to the customers' needs. (Huang et al., 2005),
- Return: defines the actions to manage the flow of flawed products (Huang et al., 2005),
- Enable: supports the other processes, defining all actions associated with supply chain management (Council, 2017).

At Level 2, these macro processes are divided into three process categories based on the business-products of the supply chain studied: make to stock (MTS), make to order (MTO) and engineer to order (ETO). For instance, SCOR version 12 offers a toolbox of 32 process categories, allowing to model any type of supply chain. At Level 3, each of these 32 process categories is broken down into standard process elements representing and describing one step or action of each of these process categories. Thus, this inter-process breakdown by process brick allows, once assembled according to a predefined scheme, to plan all the activities of a supply chain, to source raw materials, to make products, to deliver goods or services and to manage the flow of non-compliant products. At level 4, each process element is detailed according to the tasks specific to each organization (this level is therefore not considered in our study).

5.2 Positioning of atoms in the SCOR framework

This hierarchical decomposition of the processes performed in SCOR defines the level from which the elementary processes, the atoms, will be built. As illustrated in Figure 5, the atoms will be built from the processes defined in SCOR level 2. For this first version, this model will essentially consist of six atoms, modeling the *Source*, *Make* and *Deliver* processes in the case of MTS and MTO business. In order to ensure the main function associated with them, each atom must perform a sequence of predefined and configurable actions. These actions, called atom process, will be modeled from the process elements defined in SCOR level 3 (as illustrated in Figure 5).

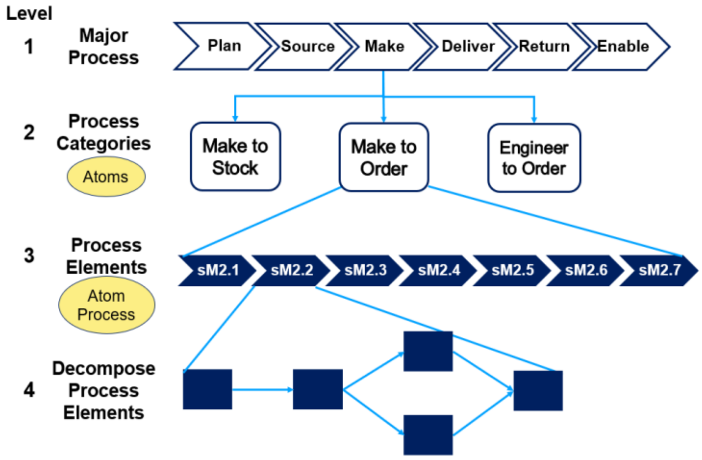


Figure 5: SCOR hierarchical process structuring and Positioning of atoms in the framework

5.3 Links between atoms, simulation model and POD approach

One of the strong expectations of this model is to offer reconfigurable modules that can be adapted to any type of supply chain, in order to reduce the effort required to develop new models. To achieve this, each atom will be modeled so as to be the most parameterizable possible. As stated in the preamble of this section, this model is developed in order to generate the data necessary for the application of the POD approach. Thus, this parameterizable character of the atoms, will also make it possible to cover a broad zone of the possible of the description space which will be associated to it. In particular, by adapting the values of the parameters of the model to the variations of values of the attributes following the realization of disturbance or the decisions taken by the managers. The description spaces at the atomic scale will be subsets of the description space describing the supply chain studied in its entirety. Each atom will be associated with a performance space and its own key performance indicators. In order to keep the root cause analysis of the performance proposed in SCOR thanks to the hierarchical decomposition of the metrics, aggregation functions will be defined and will allow to calculate the global performance of the supply chain from the metrics assessed by the atomic models. In SCOR, these global indicators are called performance attributes and defined as the strategic performance characteristics on which to align supply chain performance (Council, 2017). Figure 6 illustrates the links between atoms, simulation models and the POD approach.

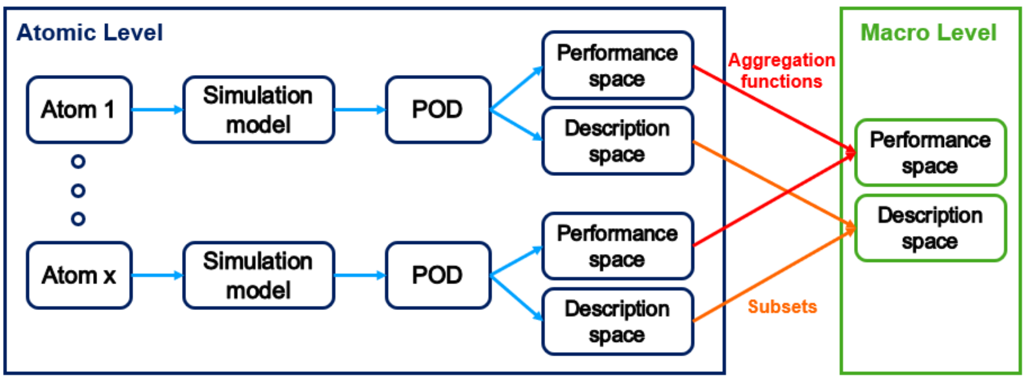


Figure 6: Links between atoms, simulation models and POD approach

6 Physical Internet

6.1 Research areas on the Physical Internet

The Physical Internet (PI) was introduced to address the Global Logistics Sustainability Grand Challenge (Montreuil et al., 2012a). In Montreuil et al. (2012a), PI is defined as “an open global logistics system founded on physical, digital and operational interconnectivity through encapsulation interfaces and protocols”. PI aims to develop a universal logistic network interconnecting all logistic services through the encapsulation of goods in smart modular containers. Thus, PI completely reshapes the current logistic network models, which are fragmented and most often dedicated to a specific company or market (Montreuil et al., 2012b). As mentioned in Pan et al. (2017), the current contributions on the Physical Internet can be divided into four axes: *conceptual research*, *assessment research*, *solutions design research* and *validation research*. *Conceptual research* focuses primarily on describing and expanding the fundamental concepts on which the PI is based (Pan et al., 2017). *Assessment research* focuses on the rigorous dimensioning of opportunities for improving the economic, environmental, and societal efficiency and sustainability induced by the PI, through experiments based on optimization and/or simulation modeling (Pan et al., 2017). *Solution design research* questions what is blocking the industry from committing to and implementing such a process and network. To address these shortcomings, it focuses on designing methodologies and technologies that enable the exploitation of the PI in the management of logistics flows across the globe. Solution design research is articulated around two main axes: designing the PI key components (for instance smart modular π-containers) and building methodologies and models to address key planning and operational decisions driven by the PI in logistic networks (Pan et al., 2017). *Validation research* aims to conduct studies to understand and determine the efforts needed to embed the PI in the industry and define the roadmap for implementing this approach (Pan et al., 2017).

6.2 Towards π-atoms

This atomic approach is inspired by some concepts developed in the PI approach. In the same way that the PI divides the logistics network into interconnected, adaptable and smart nodes, called π-nodes (Montreuil et al., 2010), the idea is to divide the supply chain processes into interconnected, adaptable and smart sub-processes, called atoms. Thus, this atomic vision of the supply chain operationalizes the multi-level PI architecture from the network level down to the enterprise level, as illustrated in Figure 7. The entire supply chain network will be reconstructed by linking the atoms together (Figure 7). The bonds between the atoms will therefore be a major issue in the realization and functionality of this model. Indeed, a lot of work on the connectivity between the atoms will have to be done. This work will be inspired by the way PI manages the interconnectivity developed between all the actors of the network.

Moreover, the goal of this approach is to contribute to the *assessment research* and *solution design research* axes, by creating π-atoms. These π-atoms would allow to evaluate the impact generated by a PI transformation on the performance of a supply chain, and thus determine the forces generated. But also, to be able to compare the robustness and resilience of a classical logistical network versus a network using PI principles in front of certain disturbances. The π-atoms will be created at first by modifying the actions of the delivery atoms. In particular by replacing some of the current actions by actions respecting PI protocols, for example by using π-containers for the transport of goods, by modifying the shipping routes so that the products transit from π-node to π-node, etc. These new atoms will be called π-Deliver, as in Figure 7. In a second step, π -atoms will be created from the protocols and processes defined in the literature, in order to develop atoms allowing to manage π -nodes, π -hubs, etc.

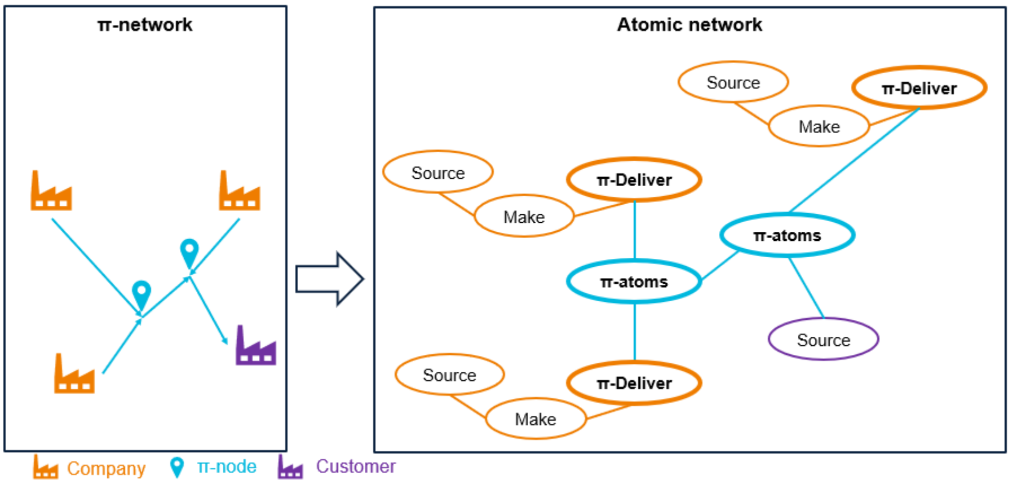


Figure 7: Link between Physical Internet and atomic modeling vision

7 Conclusion and Perspectives

All the works presented in this article depict some preliminary applications of that physics-based approach to supply chain risk management and future works in progress, such as that "atomic" vision of the supply chain modeling. The following points and perspectives define the roadmap for making this approach workable:

- Develop and design the atomic model equivalent to the supply chain presented in section 4. The aim is to use the "classical" model in order to validate the "atomic" vision, by comparing the results (the performance trajectories) of these two models subjected to the same type of disturbances.
- Check and study if it is possible to sum the atomic perturbations.
- Carry out simulation campaigns to cover the most widely accessible areas of the description space. They will generate enough data and materials to conduct analyses on the sensitivity of the system under study to some predefined disturbances and thus formalize the associated forces. In a second phase, all this generated material will be used to feed and train neural networks, which will in the future replace models of atoms and offer much more freedom and flexibility in predicting and studying the sensitivity of a supply chain to many different risks. These neural networks will avoid having to find a mathematical equation for each force.

References

- Bell, T (1989): Managing risk in large complex systems – special report, Spectrum.
- Benaben F., L. Faugere, B. Montreuil, M. Lauras, N. Moradkhani, T. Cerabona, J. Gou, W. Mu (2021): Instability is the norm! A physics-based theory to navigate among risks and opportunities, Enterprise Information Systems, 1-28. https://doi.org/10.1080/17517575.2021.1878391.
- Cerabona T., M. Lauras, L. Faugère, J.P. Gitto, B. Montreuil, F. Benaben (2020): A Physics-Based Approach for Managing Supply Chain Risks and Opportunities Within Its Performance Framework, In: Camarinha-Matos, L.M. et al. (eds.) Boosting Collaborative Networks 4.0, 418-427, Springer International Publishing, Cham. https://doi.org/10.1007/978-3-030-62412-5_34.
- Clifton, J.J. (1990): Hazard prediction. In Disaster Prevention, Planning and Limitation, edited by A.Z. Keller and H.C. Wilson, University of Bradford and The British Library.

- Cope D., M. Fayez, M. Mollaghasemi, A. Kaylani (2007): Supply chain simulation modeling made easy: An innovative approach, In: 2007 Winter Simulation Conference, 1887-1896, IEEE, Washington, DC, USA. https://doi.org/10.1109/WSC.2007.4419816.
- Council S.-C. (2017): Supply-Chain Operations Reference-Model, Scor Version 12.0, Supply-Chain Council.
- Covello, V.T. (1987): Decision analysis and risk management decision making: issues and methods, Risk Analysis, v7, no2, 131-139.
- Fayez M., L. Rabelo, M. Mollaghasemi (2005): Ontologies for Supply Chain Simulation Modeling, In: Proceedings of the Winter Simulation Conference, 2005, 2364-2370 IEEE, Orlando, FL. USA. https://doi.org/10.1109/WSC.2005.1574527.
- Hambly, E.C., E.A. Hambly (1994): Risk evaluation and realism, proceedings of the Institution of Civil Engineers, Civil Engineering, v102, no2, 64-71.
- Ho, S.S.M., R.H. Pike (1992): The use of risk analysis techniques in capital investment appraisal, In Risk: Analysis, Assessment and Management, edited by J. Ansell and F. Wharton, John Wiley & Sons, Chichester.
- Huang S.H., S.K. Sheoran, H. Keskar (2005): Computer-assisted supply chain configuration based on supply chain operations reference (SCOR) model. Computers & Industrial Engineering, v48, no 2, 377-394. https://doi.org/10.1016/j.cie.2005.01.001.
- Lave L.B. (1986): Approaches to risk management: a critique, In Risk Evaluation and Management, New York, Plenum Press.
- Merkhofer M.W. (1986): Comparative analysis of formal decision-making approaches, In Risk Evaluation and Management, edited by V.T. Covello, J. Menkes and J. Mumpower, New York, Plenum Press.
- Montreuil B., R.D. Meller, E. Ballot (2010): Towards a Physical Internet: the Impact on Logistics Facilities and Material Handling Systems Design and Innovation, 11th IMHRC Proceedings, Milwaukee, U.S.A, 40, https://digitalcommons.georgiasouthern.edu/pmhr_2010/40
- Montreuil, B., R.D. Meller, E. Ballot (2012a): Physical Internet Foundations, IFAC Proceedings, v45, no6, 26-30. https://doi.org/10.3182/20120523-3-RO-2023.00444.
- Montreuil B., E. Ballot, F. Fontane (2012b): An Open Logistics Interconnection model for the Physical Internet, IFAC Proceedings, v45, no6, 327-332. https://doi.org/10.3182/20120523-3-RO-2023.00385.
- de Oliveira, U.R., F.A.S. Marins, H.M. Rocha, V.A.P. Salomon (2017): The ISO 31000 standard in supply chain risk management, Journal of Cleaner Production, v151, 616-633.
- Pan S., E. Ballot, G.Q. Huang, B. Montreuil (2017): Physical Internet and interconnected logistics services: research and applications, International Journal of Production Research, v55, no9, 2603-2609. https://doi.org/10.1080/00207543.2017.1302620.
- Persson F. (2011): SCOR template—A simulation based dynamic supply chain analysis tool. 7.
- Serdarasan S. (2013): A review of supply chain complexity drivers. Computers & Industrial Engineering, v66, no3, 533-540. https://doi.org/10.1016/j.cie.2012.12.008.
- Tummala, R., T. Schoenherr (2011): Assessing and managing risks using the Supply Chain Risk Management Process (SCRMP), Supply Chain Management, v16, 474-483.
- White, D. (1995): Application of Systems Thinking to Risk Management: A Review of the Literature, Management Decision, v33, no10, 35.

Towards a self-organizing logistics system: identifying the value of SOLs for different stakeholders and transition-phases

Hans Quak[1,2], Elisah van Kempen[1,3], Janneke de Vries[1]

1. TNO, The Hague, The Netherlands
2. Breda University of Applied Sciences (BUas), Breda, The Netherlands
3. RSM Erasmus university, Rotterdam, The Netherlands

Corresponding author: hans.quak@tno.nl

Abstract: *The way the logistics system is organized has to change in order to be sustainable in the long term on an economic, environmental and social level. A more self-organizing logistics system (SOLs) could deal with external factors such as decarbonization requirements, developments in automation and digitization, connectivity and more demanding end receivers. We combine the insights of two practical studies on self-organizing logistics (SOLiD and SOLport). Next, by using a value proposition canvas we identify the opportunities and challenges on the transition-path towards a more SOL system. To move in the direction of such a system of connected networks, and eventually towards a Physical Internet (PI), it is important to examine the barriers for players in the current logistics system. Although a SOL system might relieve pains of, or create gains for other stakeholders, the question remains what stakeholder from what role needs to take initiative towards the transition to a more SOL system. The concise analysis in this contribution suggests that, without such a stakeholder, it might be very likely that we end up in (a number of competitive) closed networks that are dominated by stakeholders that have a strong position in the existing logistics system.*

Conference Topic: *Business models for open & interconnected logistics; Last mile & City logistics; Interconnected freight transport, logistics and supply networks*

Keywords: *Self-organizing Logistics (SOL), Business model, City logistics, Container logistics, Decarbonization, Delivery networks, Parcel Distribution, Transition, Value proposition,*

1 Introduction: drivers towards SOL?

The way the logistics system is organized has to change in order to be sustainable in the long term on an economic, environmental and social level (Quak et al., 2019; Montreuil, 2011). A transition is necessary and has already begun in some areas. We identify several (external) factors that are driving the existing system towards changes, i.e.:

- Decarbonization requirements. There is a high pressure to improve efficiency and sustainability in transport in order to reduce greenhouse gas emissions (such as CO_2). The Paris climate agreement (2015) has resulted in CO_2 reduction targets set by the European Commission: 15% in 2025 and 30% in 2030.
- Developments in automation and robotization ensure higher productivity. At the same time, handling costs decrease in various parts of the supply chain (for example in warehouses). It will be possible to have operations run 24 hours a day. This can result in new services that meet the wishes of the customer and end recipient even better at lower costs – also in areas that are not automated yet (e.g. transportation).

- Full connectivity in the physical world. As the Internet of Things (IoT) becomes more and more a reality, this creates opportunities for digitally connecting means of transport to the (logistics) infrastructure. New logistics services may become possible, ideas are developed in for example the concept of hyper-connected city logistics systems (Kim, Montreuil, & Kholgade, 2021).
- A more demanding end receiver. Due to technical developments, receivers (customers) are becoming more and more demanding. E-tailers can distinguish themselves by offering services like same day (or even within an hour) deliveries (or other personalized services) and as a result the receivers, are getting used to these possibilities in home deliveries (including parcels, groceries and meals). As it is possible to offer these services, end receivers are making more and more demands on this and want to be able to be supplied at any time, at any location. Other logistics services are essential to achieve "customer intimacy" at an acceptable price (Quak et al., 2019).

These different external factors are changing the logistics system already and require the system to change even more. A more self-organizing logistics system (SOLs) could deal with these external factors. These factors, either opportunities that arise through automation, robotization and connectivity, or threads that will change the existing way of organizing logistics operations, lead to a different logistics system in the future, but whether this system will be more self-organizing or exactly the opposite (in which a few large organizations centrally control most activities, like the tech-industry) remains to be seen.

This contribution combines the insights of two practical studies on self-organizing logistics, i.e. SOLiD (Quak et al., 2019) and SOLport (Hopman & Van Meijeren, 2020). Based on these insights, we identify the opportunities and challenges on the transition-path towards a more self-organizing logistics system. Hereby we use the view of the Physical Internet Roadmap (Ballot, et al., 2020).
Starting with two examples of practical experiments in scenario 3.1 (in the projects SOLiD and SOLport, where we cooperated with industry partners to set-up and do experiments with SOL elements) we show that it was difficult to actually set-up SOL experiments that go beyond the own logistics network and that require cooperation between different logistics networks. To really move in the direction of a SOL system that extends beyond organizational boundaries, and eventually towards a PI, it is important to examine the barriers for players in the current logistics system to move (more) in the direction of SOL. Next, we examine in two scenarios (3.2 and 3.3) what a more self-organizing logistics system would imply for especially the LSPs and their value proposition towards their customers, and based on that their likely attitude towards a more SOL system.

2 SOL and stakeholder perspectives in the literature

2.1 Self-Organizing Logistics and the transition to the Physical Internet

Pan et al. (2017) consider the Physical Internet as an application of a Self-organizing Logistics System, in which physical assets, information systems and organization models are modularized and standardized to enable connectivity. In this view a Self-Organizing Logistics system can be characterized by openness, intelligence and decentralized control. This idea of system or network openness implies that a Self-Organizing Logistics System can be extended beyond the boundaries of a single organization or logistics network as actors and assets can easily enter or leave the system. This is also recognized in the ALICE Roadmap to the Physical Internet (Ballot, et al., 2020). In the vision of ALICE, the Physical Internet will be realized in a gradual process (ALICE, 2015). This process was first conceptualized as evolving through (1) fully owned supply chains to (2) horizontal collaboration and vertical coordination and ultimately

(3) the Physical Internet (ALICE, 2015). In the most recent views, development of the PI through five generations is proposed (Ballot, et al., 2020). In the current status we are in the first generation of ‘silos within silos’ (separated subnetworks) whereby freight transport is highly fragmented and goods are not able to flow seamlessly through the system. As a subsequent step, network to network connectivity would allow organizations to connect with one another and new partners. Ultimately, in the fifth generation the Physical Internet is conceptualized as a system of logistics networks.
As we already argued in the introduction, we observe several (external) factors that are driving the existing logistics system towards changes Quak et al. (2018) point out that these external trends and drivers such as digitization and automation enable a further integration of the physical and digital world. These trends however do not necessarily have to lead to a more decentralized self-organizing system, but can also be direct towards a more central coordinated system or a hybrid form in between these options. In the remainder of this paper we will analyze how a possible transition towards more self-organizing scenarios could be influenced by the role and the changing value proposition and position of LSPs in the existing logistics system and SOL system.

2.2 Stakeholder perspectives and business modelling

Literature on SOL is relatively scare, and as – following Pan et al. (2017) PI can be seen as an application of SOL, we use the growing knowledge-base on PI to position our insights from the two SOL cases (see section 3.1), in this paper; the PI literature can be categorized in several themes of which business models are a prominent theme (Treiblmaier et al., 2020). A business model can be described as the design or archiecture of how value is created, captured and delivered by an organization (Teece, 2010). Most of the PI literature that is earmarked as the business model theme by Treblmaier et al. (2020) is concerned with the conceptualization of the main PI components such as interconnected logistics networks, PI hubs, containers and vehicles. In the emergence of the PI literature it was already recognized that the PI could have a potential impact on current business models: “The Physical Internet provides numerous opportunities for enhancing existing business models and designing novel business models. It can transform unprofitable or unreachable markets and ideas into attractive business opportunities” (Montreuil et al., 2012; pp.35).
Recently, Plasch et al. (2021) investigated the motives and success factors for collaborating in a Physical Internet network. They show that physical resources, digital/intanginble resources and relational resources are important for organizations entering a PI network as these expect competitive advantage from these resources. We link our research to the future recommendations of Plasch et al. (2021) of further developing knowledge on the business and operating models that are relevant in a PI context. This will make the vision and the way collaborations can take place more tangible.
In order to make this more tangible, we analyze how a possible transition towards more self-organizing logistics could look like. We make use of a value proposition canvas to uncover the various objectives of relevant stakeholders (Osterwalder, Pigneur, Bernarda, & Smith, 2014). The value proposition canvas helps to map the needs of the customer segment and the solution (value-proposition) that an enterprise can offer. Our focus here is on mapping the possible pains and gains of customers of the logistics system (in different SOL phases that we examine). Customers are defined broad here, and includes shippers, receivers and the society.
In this way we aim to reveal the organizational motives and barriers. In our practical application of SOL, in the projects SOLiD and SOLport, we cooperated with industry partners to set-up and do experiments with SOL elements. During the definition of these experiments we learned that it was difficult to actually set-up SOL experiments that go beyond the own logistics network and that require cooperation between different logistics networks (think of the future

envision PI generations as presented in 2.1 by ALICE). To really move in the direction of a SOL system, and eventually towards a PI, it is important to examine also the barriers for players in the current logistics system to move (more) in the direction of SOL.

3 Transition towards a SOL system: exploring stakeholder value

There are various drivers that give a possible push for the logistics system to move more towards SOL. But we learned in two applied-research trajectories, i.e. SOLiD and SOLport (see section 3.1, for short introduction) that such a transition does not take place automatically, and that it is quite complicated for companies with their role within the current logistics system to – even if they see the benefits in the longer term – drive change into the SOL direction of network cooperation.

In parallel to the PI generations (see section 2.1) we developed three SOL-collaboration configurations, that we call scenarios, in which we identified the main barriers and opportunities based on their current position in the logistics system. For these three configurations we examine the LSPs' value proposition in the current logistics system, and examine in which way a more SOL system would relieve their (customers') pains or create gains in the (near) future, based on the 'value proposition canvas' (Osterwalder et al., 2014). For the purpose of clarity we limit ourselves explicitly to three situations as we believe these are the most distinctive. Note that we explicitly call these scenarios, because a chronological order is not inherent to these steps according to our vision.

1. Optimize the own logistics network: what possibilities does SOL offer in the current situation, given the existing policies and external development within the boundaries of control of a company. Building on the developments in SOLiD and SOLport.
2. Exchange with other networks: the second scenario examines the pains and gains of cooperation between a limited number of networks from the customer and the service-provider's perspective.
3. Self-organizing logistics system: the third scenario contains the case of a self-organizing logistics system with cooperation taking place between different networks. So there is one connected system

For these three scenarios we examine the potential (future) value and role for the different stakeholders in a SOL system based on a value proposition canvas (Osterwalder et al., 2014). We identify the (future) value proposition - based on the products and services and the (future) gain creators and pain relievers for the customer segment - and the expected pains and gains of the customers. As such, we examine the potential value proposition to see what future drivers or barriers arise from a business perspective based on the position of existing players in the existing logistics system. By identifying these factors, we are able to see how and in which way existing or new stakeholders are to be incentivized to make a transition towards a more SOL system.

3.1 SOLiD and SOLport

We used our insights from two applied-research projects on SOL to examine the position and value propositions in the different SOL scenarios, i.e. SOLiD and SOLport. In this contribution we do not go into details of these research projects, nor do we discuss the project results (for that we refer to the project deliverables and other papers), but we only use the insights and reactions of the cooperating companies to examine how their position, role and value proposition (i.e. relation to their clients) changes in different stages of a development towards a SOL system. The possibilities LSPs have within the boundaries of their own organization to go towards more SOL was examined and experimented with in SOLiD and SOLport. The other scenarios were developed more theoretically, as this was outside the boundaries of the companies cooperating in the projects. (As a result, scenario 1 is described in more details).

Within the SOLiD project, practical experiments were conducted which were focused around optimizing the network within the daily logistics operations of parcel distribution given decentralized control, intelligence and openness of local real-time data in different assets (Quak et al., 2019). These experiments provided an outlook with respect to the feasibility of a more self-organizing logistics system and the extent to which it possibly adds value for the parcel delivery- and logistics industry. These experiments were conducted in the organizations of DPD and subsequently Fietskoeriers.nl. As these organizations are regarded as subsystems of a larger distribution network (Van Ommeren, et al., In press), steps towards SOL were only made in subsystems rather than in the larger system of different networks.
Within the SOLport project, a practical experiment was conducted regarding the transition of hinterland transport towards SOL by decentral form of control. This project provides an outlook with respect to the feasibility of SOL within a logistic network of hinterland transport. In the SOLport case the supply of bulk goods from a production facility to three locations is fully executed by barges (Hopman & Fransen, 2020). Comparable to the SOLiD project, in SOLport only the supply chain of one logistics actor was considered and therefore steps could only be made for making the single organization network move towards more decentralized control and self-organization.

3.2 Scenario 1. optimize the own logistics network

Given the characteristics of a SOL system, what possibilities does SOL offer in the current (and future) situation for a logistics company within the existing logistics system, given the existing policies and external developments within the boundaries of control of a company? We examine two different logistics systems: in SOLiD the main focus is on last-mile parcel deliveries, whereas in SOLport the main focus include hinterland container transportation and dry bulk inland navigation. The experiments in SOLiD and SOLport are a way to learn more on what self-organization could mean in daily logistics operations of a single logistics partner. In both projects experiments are conducted that introduce more dynamic and decentralized coordination compared to the pre-existing more central coordinated logistics activities.

3.2.1 Optimizing parcel distribution through dynamic planning and local intelligence

In the value propositions towards SOL in the first scenario, we identify the pains and gains of the different stakeholders in the existing system, as well as in the transition towards a more self-organizing logistics system. In the SOLiD case the logistical service suppliers DPD and Fietskoeriers.nl create value by delivering parcels as a service in a city environment as efficient as possible by optimizing their routes and limiting downtime during a single stop. However, due to these efficient route planning for the customer only one daily timeslot is available for the consumer to receive the goods. Also this time slot is fixed and communicated only shortly in advance. The existing system is efficient, but hardly allows for extra services (like specific slots, late changes in time or address, etc.). Table 1 shows the main gains for the different stakeholders in the current system, as well as the changes due to more SOL within the optimization of the LSP's logistics network (which mainly provide receivers with the option to make some last minute change, although limited as a parcel is already in van or cargo bike).
This scenario of optimizing the own (LSP) network shows almost no big changes, expect for the possibilities to customize services at a relative low extra costs. For society however the optimization of a single logistics network is not always beneficial, since fast delivery due to more demanding end-receivers requires multiple vehicles in the street with a higher chance on congestion, air pollution and higher vehicle emissions. A first step towards SOL is the optimization of routes using local intelligence and autonomy of the planning. In this case time is saved by automation of route planning while ensuring route efficiency (Van Ommeren et al., In press). However, major societal challenges such as congestion, air pollution and vehicle

emissions still need to be mitigated since optimization in a single subsystem in the logistics system of last-mile parcel deliveries has limited benefits for these major societal threats.

Table 1 Scenario 1 – Optimization of a single logistic network (parcel delivery): role and potential benefits from SOL in LSP's network

Stakeholder group	***Role/ Cares about***	✓ ***Potential gains***	***Potential pains***
Shipper	Selling products. Agreeing service contracts with transporters. Good service to final receiver.	✓ Goods delivered ✓ Being able to create customer intimacy by offering customized delivery services via LSP	- Offering specific delivery options for reasonable costs
Receiver	Receives ordered goods at preferred location. Wants to feel serviced and treated well (customer intimacy)	✓ Goods delivered at time, location of choice and receiver does not necessary have to go to the store	- Only one daily timeslot is available for the consumer to receive the goods. - No/ limited influence on timeslot.
Municipality, Community, Society	Cares about livable environment; clean air, congested-free roads.		- Number of delivery vans/bike messengers remains the same, increased customized services might increase freight traffic - Optimization single logistics network does not affect congestion, air pollution and vehicle emissions
	Value proposition	***Gain creator***	***- Pain reliever***
Logistics Service Provider/ Transporter	Fulfilling shippers request by servicing final receivers at low costs. Timely delivery of goods while optimizing vehicle movements.	✓ Optimizing load factor of delivery vans/bike messengers ✓ Automation of planning eliminates central planner; time is saved as decisions on routes are made based on local information ✓ Flexibility in offering extra services	- Feeling of lack of control in case there is no direct receiver-contact; - Lack of trust with collaborating partner.

3.2.2 Transitioning hinterland transport towards SOL by decentral form of control with local intelligence

Self-Organizing Logistics (SOL) comes into scope, where operators and carriers make local decisions to organize their logistic operations and no longer require central control (gains for carrier and operators). The requirements for this system include security of stock supply for the shipper and efficient use of assets for carriers and operators. This requires the carrier to have full access to local data; ship data, location data, data of stock at the consumption locations in order to steer logistic operations. The availability and interoperability of these data between assets is not yet evident.

The gains in the practical experiment of this project are comparable to the case in parcel distribution; mainly the shippers/logistical service provider experiences benefits from a automation of the operations; i.e. saving time for planner in optimizing routes, efficient use of assets in operational planning, possibility for the shipper to include personal preferences in the planning for receiver locations. However, major societal challenges such as congestion, air pollution and emissions still need to be mitigated since optimization in a single subsystem in the logistics system of last-mile parcel deliveries or container hinterland transport has limited benefits for these major societal threats.

3.3 Scenario 2. exchange with other networks

In the previous section we described what might happen when the various partners optimize their own processes. For the logistics operators (such as DPD, Fietskoeriers.nl in SOLiD) it is beneficial to focus on optimizing their own operation. For receivers, municipality and society however it is debatable whether it brings the highest value when logistics operators focus on optimizing their own processes solely.

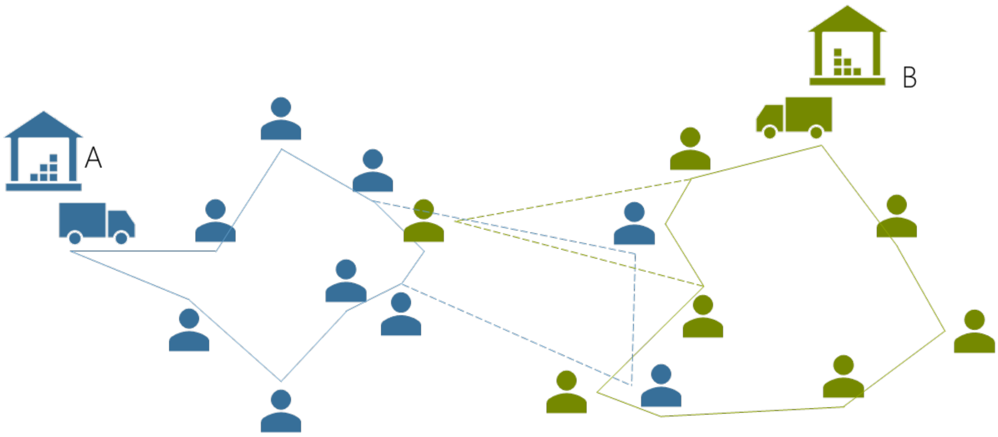


Figure 1 Horizontal collaboration between network A and network B (dotted line reflecting no collaboration)

In addition to the previous scenario, another scenario can be formulated whereby individual transport partners seek (horizontal) collaboration and exchange with other transport networks. Horizontal collaboration is defined as cooperation between competing firms that are at the same level in the supply chain, for example exchanging or pooling transport capacity to improve load factor (Cruijssen et al., 2007; Ferrell et al., 2020). The driving force behind collaboration is that each partner experiences a net positive value (Cruijssen et al., 2007). Although theoretically there are high benefits of collaborating with other networks – such as resource sharing, cost sharing, increasing flexibility, there exist also quite some barriers for logistics companies to engage in such collaborations (Simmer et al., 2017). Crucial factors for successful collaboration are trust between the actors, setting of precise conditions and shareable IT structures (Simmer, Pfoser, Grabner, Schauer, & Putz, 2017).

Given the above, if we consider an example for two parcel delivery companies A and B (see Figure 1) they could (use the external developments mentioned in section 1 to):

1. (Further) optimize their own network (scenario 1): network A would include all 'blue receivers' and include the dotted line receivers as well and vice versa for Network B.
2. Exchange with other networks (scenario 2): network A would service the 'green receiver' which is already close to its 'own blue receivers'; and network B would service the two 'blue receivers' which are already close to its 'own receivers.'

The exchange-scenario was not tested in practice in SOLiD, but mainly theoretically with the industry partners. As a result, we claim the exchange with other networks has several implications for the different stakeholders (see Table 2). First of all, if two networks start cooperating, the LSPs are able to (further) optimize the load factor of their vans on neighborhood- or even street-level, while at the same time minimizing travel distance (gain). However, this will come with possible disadvantages such as a possible lack of control when a collaborating partner serves the final receiver (pain). Collaborating will come with challenges of deciding on the rules that allow decentral assigning parcels to one partner or the other. Both the receiver and the municipality might gain from the more bundled deliveries. Notice that the

LSPs are mentioned twice in Table 2; as they are affected by the (SOL)cooperation, but it could also contribute to their value proposition.

Table 2 Scenario 2 - exchange with other networks (parcel delivery): role and potential benefits from SOL in LSP's network.

Stakeholder group	***Role/ Cares about***	✓ ***Potential gains***	***Potential pains***
Shipper	Selling products. Agreeing service contracts with transporters. Good service to final receiver.		- Lack of influence on which transporter will service final receiver
Receiver	Receives ordered goods at preferred location. Wants to feel serviced and treated well (customer intimacy)	✓ Bundled deliveries	- No influence on delivery time
Municipality, Community, Society	Cares about liveable environment; clean air, congested-free roads.	✓ Bundled deliveries might reduce vehicle movements in neighborhood	- Double parking by transport companies in neighborhood
Logistics Service Providers/ Transporter	Fulfilling shippers request by servicing final receivers. Timely delivery of goods while optimizing vehicle movements.	✓ Optimizing load factor (of delivery vans);	- Feeling of lack of control in case there is no direct consumer-contact; - Lack of trust with collaborating partner. - Less visible at receiver
	Value proposition	***Gain creator***	***- Pain reliever***
Logistics Service Providers/ Transporter	Together with partner provide higher service at lower costs, reduce number of vans in neighborhoods	✓ Higher drop density provides possibilities for service customization ✓ Higher service to receiver	- Less vans in neighborhood and / or streets (not necessary fewer kilometers though)

3.4 Scenario 3. self-organizing logistics system

Ultimately we would go even one step further, towards a SOL system. The SOL system is a form where the containers or parcels find their own way in the open, intelligent en decentralized controlled system (see Pan et al. 2017). Starting with examining the main concerns in such a system for the different stakeholders we see that the value proposition for the LSP is really different than in the other two scenarios. No longer does the LSP provide the value proposition to the customers, but the LSP is more an asset-provider in the SOL system; and the SOL system provides – based on the conditions set by all stakeholders – value for the final customer. In this scenario the LSP does no longer provide direct value to a shipper or receiver, as this comes from the SOL system. This situation also shows the barrier to reach this full SOL system, considering the current stakeholders, as the LSPs currently aim at providing value for their customers. The question is if, and if so how these LSPs will change from providing service to customers to providing services (or even only assets) to the open SOL system. In the first scenario, the LSP used SOL ideas to provide a better customized service, whereas in the second scenario the LSP provides these services at either lower costs, lower environmental and societal costs, or at a higher frequency due to cooperation between networks. The third scenario shows a complete other role for the LSP, that is not in line with the existing and developing value propositions in the earlier steps towards a SOL system. The outcomes in this scenario show that from a LSP perspective it might be better to develop scenario 2 further, so that the collaborating LSPs can provide services in a closed collaborative system to their customers, who then might not feel attracted to the potential gains and pain-relievers the SOL system offers (in scenario 3, see Table 3).

Table 3 Scenario 3 – SOL system: role and potential benefits from SOL in LSP's network.

Stakeholder group	***Role/ Cares about***	✓ ***Potential gains***	***Potential pains***
Shipper	Selling products. Agreeing rules with receiver on conditions of transported goods.	✓No contracts with specific LSPs needed ✓Customized service based on receivers' demands	- Hard to differentiate in offering (delivery) services as receiver can set conditions
Receiver	Receives sets conditions for delivery: ordered goods at preferred time and location. Wants to feel serviced and treated well (customer intimacy).	✓Bundled deliveries ✓Possibilities to steer on deliveries and set conditions	
Municipality, Community, Society	Cares about liveable environment; clean air, congested-free roads; sets conditions on space utilization, emissions and …	✓Bundled deliveries might reduce vehicle movements in neighborhood	- Reduction externalities - System does not automatically develop in desirable direction (risk for race to the bottom in delivery-employment)
Logistics Service Providers/ Transporter	Sets conditions on availability of assets		- Lack of control in case there is no direct consumer-contact; - Lack of trust with collaborating partner. - Less visible at receiver and shipper
	Value proposition	***Gain creator***	***- Pain reliever***
SOL system	Provide open, intelligent and decentralized controlled network to fulfil requirements within conditions	✓Possibilities for highly customized services	- Reduction externalities

4 Concluding remarks

By our analysis of the value propositions of SOL in the different scenarios for the different stakeholders, we identify the incentives and barriers the different stakeholders have and face in the transition towards a more self-organizing logistics system. These results help us to understand how to steer the developments, that follow from the external factors, in the changing logistics system in a more sustainable way, as well as to identify the main actors in the different transition-phases that have an interest and are able to cause a change the system.

As the scenarios show, given the stakeholders' positions in the current system and the external development, we can argue that SOL (applications) might improve the operations within logistics networks or even in (closed) collaboration between networks. Although in the later situations incentives for LSPs and shippers are becoming more challenging as these stakeholders also face potential pains. The last scenario, the SOL systems shows that this requires – especially for LSPs – a new way of thinking: rather than servicing the shippers by delivering to the end receivers, LSPs have to service the system. On a system level, this SOL system might lead to the highest benefits – improved resilience, adaptability and sustainability – for receivers and society if agreed on the right boundary conditions. However, it remains uncertain whether this is enough for making a transition towards SOL happen by itself, since there are quite some barriers for LSPs.

The question is what stakeholder from what role needs to take initiative towards the transition to a more SOL system. The concise analysis in this contribution suggests that, without such a stakeholder, it might be very likely that we do not arrive at a SOL system, but might end up in (a number of competitive closed cooperative) networks that are dominated by stakeholders that have a strong position in the existing logistics system; as in scenario 2.
The ALICE Physical Internet Roadmap (Ballot, et al., 2020) has included – we believe justifiably – steps to be taken with respect to the development and adoption of new business models. Based on our analysis we would like to stress that analyzing the incentives for stakeholders to adopt new practices is key if we want to move forward to a logistics system that is resilient, sustainable and futureproof.

Acknowledgements

The authors like to thank the Dutch Topsector Logistics (TKI Dinalog and NWO) for the support to Project SOLiD (NWO project number 439.17.551).

References

- ALICE. (2015). *Global Supply Network Coordination and Collaboration.* Retrieved from http://www.etp-logistics.eu/wp-content/uploads/2015/08/W46mayo-kopie.pdf
- Ballot, E., Barbarino, S., Van Bree, B., Liesa, F., Franklin, J., 't Hooft, D.,….., Tavasszy, L. (2020). *Roadmap to the Physical Internet.* Brussels: ALICE. Retrieved from http://www.etp-logistics.eu/wp-content/uploads/2020/11/Roadmap-to-Physical-Intenet-Executive-Version_Final.pdf
- Cruijssen, F., Cools, M., & Dullaert, W. (2007). Horizontal cooperation in logistics: opportunities and impediments. *Transportation Research Part E: Logistics and Transportation Review, 43*(2), 129-142.
- Ferrell, W., Ellis, K., Kaminsky, P., & C., R. (2020). Horizontal collaboration: opportunities for improved logistics planning. . *International Journal of Production Research, 58*(14), 4267-4284.
- Guo, C. (2020). *The physical internet for city logistics. PhD thesis.*
- Hopman, M., & Fransen, R. (2020). *SOLport praktijkcasus – aansturing logistieke aansturingsvormen.* In press.
- Hopman, W., & Van Meijeren, J. (2020). *Logistieke aansturingsvormen van centraal tot decentraal. TNO report 2020 P11269.* The Hague: TNO.
- Kim, N., Montreuil, B. K., & Kholgade, N. (2021). Hyperconnected urban fulfillment and delivery. *Transportation Research Part E, 145.*
- Montreuil, B. (2011). Towards a Physical Internet: Meeting the Global Logistics Sustainability Grand Challenge. *Logistics Research, 3*(2-3), 71-87.
- Montreuil, B., Rougès, J. F., Cimon, Y., & Poulin, D. (2012). The physical internet and business model innovation. *Technology Innovation Management Review, 2*(6).
- Osterwalder, A., Pigneur, Y., Bernarda, G., & Smith, A. (2014). *Value proposition design: How to create products and services customers want.* John Wiley & Sons.
- Pan, S., Trenesaux, D., & Sallez, Y. (2017). Specifying Self-Organizing Logistics System: Openness, Intelligence, and Decentralized Control. In Borangiu (Ed.), *Service Orientation in Holonic and Multi-Agent Manufacturing, Studies in Computational Intelligence,* (pp. 93-102).
- Plasch, M., Pfoser, S., Gerschberger, M., Gattringer, R., & Schauer, O. (2021). Why collaborate in a physical internet network? *Journal of Business Logistics, 42*(1), 120-143.
- Quak, H., Van Kempen, E., & Hopman, M. (2018). Moving towards practical implementation of self-organizing logistics-making small steps in realizing the PI vision by raising awareness. *5th International Physical Internet Conference IPIC.* Groningen.
- Quak, H., Van Kempen, E., Van Dijk, B., & Philipson, F. (2019). Self-organization in parcel distribution: SOLiD's first results. *IPIC 2019. 6th International Physical Internet Conference.* London: IPIC.

- Simmer, L., Pfoser, S., Grabner, M., Schauer, O., & Putz, L. (2017). From horizontal collaboration to the Physical Internet - a case study from Austria. *International Journal of Transport Development and Integration, 1*(2), 129-136.
- Teece, D. (2010). Business models, business strategy and innovation. *Long range planning*, *43*(2-3), 172-194.
- Treiblmaier, H., Mirkovski, K., Lowry, P. B., & Zacharia, Z. G. (2020). The physical internet as a new supply chain paradigm: a systematic literature review and a comprehensive framework. *The International Journal of Logistics Management*, *31*(2), 239-287.
- Van Ommeren, C., Quak, H., Van Kempen, E., Philipson, F., & De Vries, J. (In press). *Feasibility of Self-Organisation and Field Trials for Parcel Delivery Optimization by Bike Messengers*. The Hague: TNO.

Modularization of Delivery and Transportation

Nayeon Kim[1,4], Benoit Montreuil[1,2,3,4] and Walid Klibi[1,5]
1. Physical Internet Center
2. Supply Chain and Logistics Institute
3. Coca-Cola Chair in Material Handling and Distribution
4. H. Milton Stewart School of Industrial & Systems Engineering
Georgia Institute of Technology, Atlanta, United States
5. The Centre of Excellence for Supply Chain Innovation & Transportation (CESIT),
Kedge Business School, Bordeaux, France
Corresponding author: nkim97@gatech.edu

Abstract: *Traditional static and optimization based transportation planning and operation system has limitations on supporting highly dynamic, hyperconnected, multi-player logistics system in Physical Internet. Modularizing delivery and transportation can enable flexible transportation operations to adapt to such highly dynamic environment with potentially reduced planning effort. The modularization schemes are categorized into regional, hierarchical, and functional modularization. Regional modularization replaces a long route with many shorter routes, or modules, interconnected via PI hubs. Hierarchical modularization implements a multi-tier transportation system where each tier becomes a module connected via PI hubs. Functional modularization is to ensure each route serves a single functionality, each of which requires different resources. All three modularizations can increase flexibility of planning and operation while increasing the consolidation level and efficiency. Experimental results applying hierarchical and functional modularization for last mile delivery shows up to 27% of operational cost savings.*

Keywords: *Modularized transportation; Regional modularization; Hierarchical modularization; Functional modularization; Transportation design; Physical Internet*

1 Introduction

Efficient and cheap, yet fast and punctual, delivery and transportation are one of the main goal of logistics. Not surprisingly, there has been plethora of research and practice to improve transportation efficiency which ultimately can reduce cost and transportation-induced impacts on environment and society. The main challenge is to ensure the service capability, which is often measured as responsiveness and punctuality, while reducing costs. The growth of e-commerce, home delivery and globalization imposes even more challenges as origin and destination pairs become more diversified, each shipment size becomes smaller, and delivery lead time becomes tighter. Typically, such transportation problems are tackled by solving them with more advanced optimization modeling, such as stochastic optimization or online/dynamic routing, or solution methodologies, such as branch-and-price, metaheuristics or even machine learning (Gutierrez et al., 2018; Nazari et al., 2018; Koc et al., 2016). In fact, the transportation and routing optimization problems are one of the famous examples of mixed integer programs, which is well known to be a NP-hard problem (Toth & Vigo, 2014; Laporte, 2009). That is, in practice, companies spend substantial time and capital for finding an optimal transportation plan, which tends to be modified at the time of operations. However, for the Physical Internet based hyperconnected delivery and transportation operations, which is highly dynamic and where multiple players' operations are synchronized together, the traditional approach may not be suitable and can even be inefficient.

The Physical Internet (PI) is recently rising logistic paradigm shift (Montreuil, 2011). Here, we focus on the delivery and transportation in PI system (Kaboudvand et al., 2021; Kim et al., 2021). The traditional transportation and delivery planning methodologies are not easy to be applied in highly dynamic PI environment where multiple players are interacting. This paper aims to propose a flexible transportation planning scheme to facilitate PI transportation through modularization of delivery and transportation, as an extension of Kim et al. (2021) further developing transportation and delivery scheme. This is in line with the fundamental idea of the PI, a digital internet analogy applied for physical assets. The analogy clearly implies that modularization serves a key role under PI. Van Luik et al. (2020) provides good analogy and link between digital internet and PI. In this paper, we divide the transportation modularization into three types: regional, hierarchical, and functional modularization.

This paper contributes to literature by conceptualizing the modularized delivery and transportation and categorizing the modularization scheme into regional, hierarchical, and functional modularization. Then, the potential impact of modularized delivery is demonstrated via simulation-based experimental results in urban delivery. Through the conceptualization and simulation, we aim to develop a novel framework for delivery and transportation system to better facilitate PI operation. The structure of the paper is as follows. In Section 2, the modularized delivery and transportation is defined and conceptualized. Three types of modularization are described and linked to literature. Section 3 and 4 present the design and results of simulation-based experiments. Section 6 summarizes the paper and concludes with future research avenues.

2 Modularization of transportation and delivery

This section define modularization of delivery and transportation. We first review modularity in logistics and the hyperconnected delivery and transportation in PI context, from which the transportation modularization is motivated. Then, different types of modularization are defined, namely, regional, hierarchical, and functional modularization.

2.1 Modularity

Modularity is a general concept that has been used in various contexts such as science, software and programming and system design. Baldwin et al. (2000) defines a module as "*a unit whose structural elements are powerfully connected among themselves and relatively weakly connected to elements in other units*" and the core of modularization is to divide a complex system into smaller subsystems, which encapsulate their complexity in them and interact in the original system through interfaces, emphasizing the key concepts are *abstraction, information hiding,* and *interface*. Holtta and Salonen (2003) defines a module as "*a structurally independent building block of a larger system with well-defined interfaces*". In the context of supply chain, Bask et al. (2010) has provided the definition "*a modular system is a system built of components, where the structure ["architecture"] of the system, functions of components ["elements", "modules"], and relations ["interfaces"] of the components can be described so that the system is replicable, the components are replaceable, and the system is manageable*" through extensive literature review. Pekkarinen and Ulkuniemi (2008) identify the modularity into three categories: modularity in service, process and organization and emphasize the needs of good platform that links modules. All modularization definitions emphasize the independent components, interfaces, and system architecture where modules are connected to other modules. Another key feature we emphasis is, as stressed by Baldwin et al. (2000), each module should contain strongly-related elements that can be independently treated to other modules while inter-module dependency is minimized. Building upon the definition of module and modularization, we aim to apply the core logic of modularization to transportation and delivery system. Also, the paper aim to address how modularization of delivery and transportation can improve flexibility, reduce planning and operational complexity, and save costs, which are pointed out as a common benefits of modularization by Rajahonka (2013). Bask et al. (2010) categorized modularity in supply chain into modularity of product, production/process, supply chain/organization, and services. The transportation and delivery modularization discussed in the paper belongs to supply chain and service modularization.

2.2 Modularization for hyperconnected delivery and transportation

We define hyperconnected delivery and transportation as a *multi-player routing/shipping modularized by region, layer, and/or functionality enabling dynamic and broad range of flow consolidation*. The key notion in this definition of hyperconnected delivery and transportation is modularization. In this paper, a module is defined as a smallest planning and/or operation unit for transportation. For example, often the transportation to the last hub and the last mile delivery to customers are planned and operated separately, in which case, they can be considered as two separate transportation modules. Each module may be planned independently, served by different vehicles, or operated by different players. The major benefits of modular transportation stems from this characteristic. Meanwhile, in hyperconnected logistics, the modules are seamlessly interconnected with each other via standardized physical and digital interface. This is closely related to PI enablers of modular transportation, such as PI hubs (Ballot et al., 2012A; Meller et al., 2012) and containerization (Kaboudvand et al., 2021; Sternberg & Denizel, 2020; Gontara et al., 2018; Sarraj et al., 2014; Montreuil et al., 2010).

2.3 Regional modularization

Regional modularization is the most straightforward modularization scheme. It is to divide a long delivery route into multiple smaller and synchronized delivery routes by region. It lets each route covers smaller area. Montreuil (2011) describes it with an example of truck-based transportation from Quebec, Canada to LA, USA, where it is converted from long-haul to distributed multi-segment travel. Montreuil (2011) also emphasizes the flexibility of decision making achieved via such modularization, which can allow the use of other transportation modes such as rail for some segments. Ballot et al. (2012A) designed and applied open hub network to food distribution in France and showed remarkable results achieving regional modularization. The proposed network reduces the maximum length of trips by more than 71%, to 400 km. Sarraj et al. (2014) and Gontara et al. (2018) also modeled and experiment container routing problem with a network of PI hubs. Fazili et al (2017) further develop container routing model building on Sarraj et al. (2014). Figure 1 describes regional modularization where long routes can be modularized into shorter segments (or modules). Each module can be served in a regular work hour, and more consolidation can be achieved.

The regional modularization can facilitate flexible routing and dynamic consolidation by having more decision making and consolidation points. Also, with well-established network of PI hubs, each route module can be short enough to be served in regular working hour. This has positive impact of driver's work environment, being able to 'sleep-at-home'. It is one of the crucial factor to determine drivers' quality of life (Ferrell, 2016). Improving work environment for drivers can also contribute to lower the operation cost by reducing driver turnover (Min & Lambert, 2002). Also, typically long haul routes are served by a driver team, not a single driver, which increase labor cost per mile. When the shorter route modules replace long hauls, it can reduce labor hour per travel mile as short route modules can be served by a single driver.

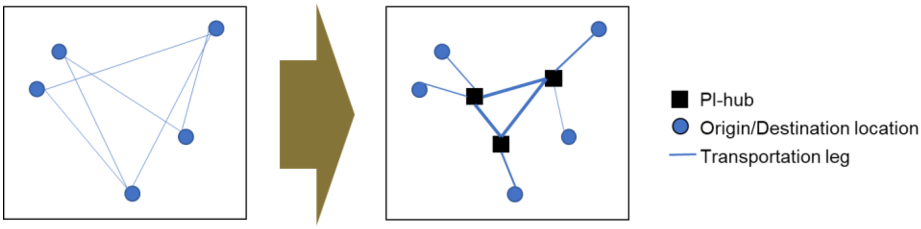


Figure 1:Illustrating regional modularization

2.4 Hierarchical Modularization

Hierarchical modularization is to divide delivery operations into multiple tiers which can be served in different time by different types of vehicle. It can be easily seen in the multi-tier or multi-echelon systems (Crainic & Montreuil, 2016; Winkenbach et al., 2016; Crainic et al, 2009; Smilowitz & Daganzo, 2007; Crainic et al, 2004). Montreuil et al. (2018) further developed the multi-tier system to hyperconnected delivery urban setting. Figure 2 describes 2-tier hierarchical modularization. The 2-tier system achieves better consolidation for tier 1, the delivery from fulfillment center to PI-hubs (which were urban consolidation centers (UCC) or satellite hubs in traditional settings). The tier 1 deliveries can be done during non-congested hours, overnight or early morning. The tier 2 delivery route is a lot shorter than tier 1 routes and carries smaller amount. Therefore, smaller vehicle or alternative transportation modes such as bicycle can be used depending on the scale. This is particularly beneficial when serving an area with special transportation restriction such as historical city center or when the road are very narrow inside a city as seen in many old cities. Generally speaking, higher tiers can utilize higher consolidation level and larger flow, which increases truck fillrate and shipping frequency in general. Lower tiers can adjust truck sizes to compensate relatively lower flow while maintaining shipping frequency.

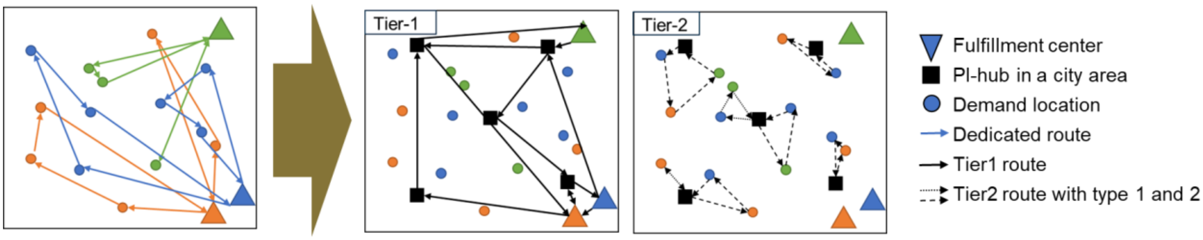


Figure 2: Illustrating hierarchical modularization

2.5 Functional modularization

Functional modularization is to divide multi-functionality routes into multiple routes of single functionality. This is less common than the other two types of modularization. Note that the in functional modularization, each functionality must require different resources which have different cost or require different time commitment. For example, pickup-and-delivery routes are common and proven to be more efficient than having pickup routes and delivery routes separately. Parragh et al. (2008) provides good survey of pickup and delivery literature and it is still actively studied in different context such as crowdsourcing (Arslan et al. 2019) or e-commerce distribution (Bergmann et al., 2020). The pickup and delivery are not the candidate of functional modularization as both requires same skillset of personal and same types of vehicles. Another multi-functionality route example is a route performing both physical item delivery and service at the same time. This route serves two functionality at once and it is a candidate of functional modularization. The two functionality require different time, equipment, and/or personnel. Note that Bask et al. (2010) pointed out that installation or recycling services can be bundled to the delivery as an additional value-added service module under e-commerce operations, as an example of *customer service* modularization which should be differentiated with the functional modularization presented here which is a *logistics service* modularization.

A more concrete example is a last mile delivery of furniture and large appliances, which often includes both delivery and installment (white-glove services) as described in Figure 3, which require different resources. Delivery requires large truck to fit the large items and a team of two personnel. Installment, on the other hand, can be done in small vehicle, but requires more skilled personnel and takes more time in general. When separated, installment routes can be served with small vehicle and delivery routes can be served without skilled personnel for installment. Also, installers can serve maintenance stops along with installation stops. Such functional modularization can increase efficiency by reducing the inefficient usage of more expensive resources, large trucks and skilled personnel. That is, each module can be more specialized.

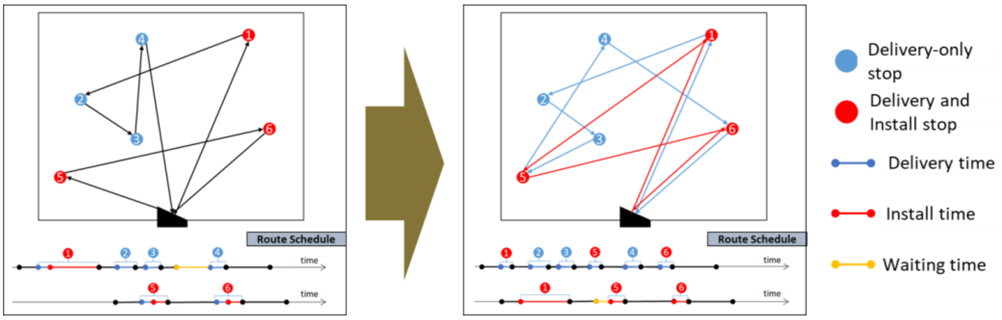


Figure 3: Illustrating functional modularization

3 Simulation-based experiment design

We designed a simulation-based experiment to measure the impact of hierarchical and functional modularization. The simulation experiment is built on the last-mile delivery of large items in a city. Here, we only briefly describe the scope of simulation and focus on the experiment design relative to the transportation modularization. See Kim et al. (2021) for details of base simulation platform structure.

3.1 Scope of simulation

The simulation experiment is built for the context of last-mile delivery of large items in urban area using the simulation platform presented by Kim et al. (2021) with customization. Urban delivery is suitable to implement the hierarchical modularization and large item delivery which often requires both delivery and installment can be a great example to implement the functional modularization.

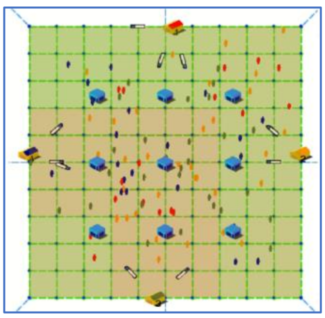


Figure 4 Simulation snapshot: four peri-urban fulfillment centers and nine PI hubs on a grid map

A grid-based model city is used where there are four retailers, each owns a single dedicated peri-urban fulfillment center (FC) and makes about 25 deliveries per day on average. When PI hubs are used, a network of nine PI hubs are assumed to be located in the city area. The snapshot of simulation with the 4 peri-urban FCs and 9 PI hubs are shown in Figure 4.

All stops involve delivery but only 80% of them requires installation. Each product has different delivery and installation times associate with it. Customers require a specific delivery time window between 9AM and 8PM, duration of which ranges from 2 to 4 hours. As pre-confirmation is needed for delivery time window, customers to be served on any day can be assumed to be known in advance without loss of generality.

3.2 Modularization scenarios

Two modularization schemes are used: hierarchical and functional modularization. For hierarchical modularization, current direct delivery (single-tier, ST) to customers from FCs and multi-tier delivery (MT) via PI hubs in the city are compared. In the latter, tier-1 delivery refers shipping from FCs to PI hubs and tier-2 from PI hubs to customers. For functional modularization, current multi-function delivery (MF) where delivery-only stops and delivery-and-install stops are mixed in a single route is compared to single-function delivery (SF) where each route performs either delivery or installation only. Four scenarios are constructed as shown in Figure 5: ST-MF, MT-MF, ST-SF, MT-SF.

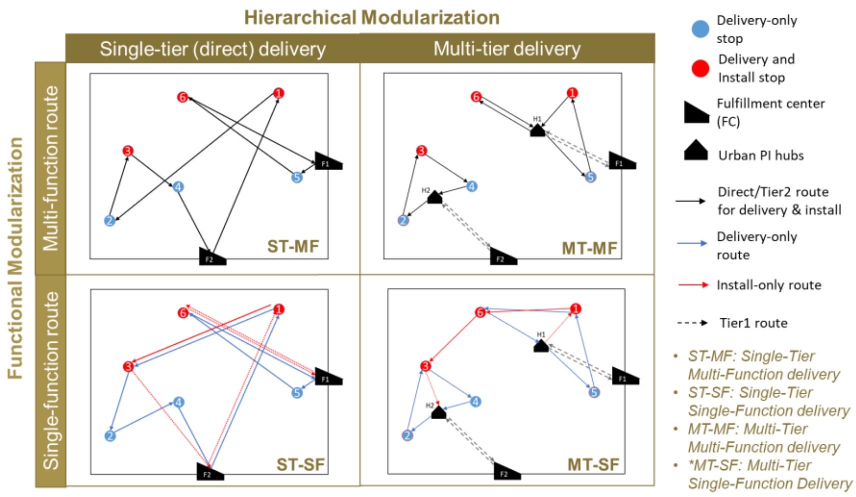


Figure 5: Four delivery scnenarios by hierarchical and functional modularization

Routes are served by different size of vehicles and have different requirement depending on tier and functionality. Table 1 summarizes the type of vehicles, labor requirements, and scheduling constraints for routes by tier and functionality.

Table 1: Transportation configuration by types of route with respect to hierarchical and funcational modularization

Functionality	Tier	Vehicle	Labor	Schedule
Delivery & install route	Direct delivery	17' truck	Driver & Installer	Time window
	Tier-2	10' truck	Driver & Installer	Time window
Delivery route	Tier-1	26' truck	Driver	Overnight
	Direct delivery	17' truck	2 Drivers	Time window
	Tier-2	10' truck	2 Drivers	Time window
Install route	Tier-2	9' cargo van	Installer	Time window

Delivery routing problem in this case belongs to routing with time window and capacity constraints problem. Here, the routing heuristics in Kim et al. (2021) is used. However, they did not present the routing with functional modularization. The most critical constraint in this case is that installation cannot start before delivery is completed. Therefore, when delivery and installation routes are separated, delivery routes are first constructed with same heuristics to the mixed functionality routes where the expected time at each customer stop only include delivery time. Then, the expected finishing time of delivery at each customer, which then becomes expected earliest installation start time, is fed as an input for installation routing. When an installer may arrive before delivery team due to stochastic travel and service times in simulation, the installer waits until the delivery is completed. Note that, for installation routes, any size of cars can be used as they do not carry large products and it does not need to start from hubs or FCs unless specific tools need to be picked up at FCs. This means that an installer can park the installation vehicle at his/her place or anywhere preferred. To reflect the flexibility, it is assumed that installation routes are modeled as path, starting from nearest hub or FC and end at the last customer location. The associated costs for installation routes are also charged for the path only.

4 Experimental Results

This section presents the experimental results from simulation-based scenario analysis. The four scenarios (ST-MF, ST-SF, MT-MF, MT-MF) are evaluated on the urban delivery simulation, built on 8.5.1. University. The most important key performance indicator (KPI) is total operational costs. It consists of labor cost, vehicle rental cost, fuel cost and hub usage cost. As shown in Figure 6, the total operational cost is reduced with modularization. When soley applied, functional and hierarchical modularization reduces the average cost by 16% and 21% respectively. Together, up to 27% of operational cost savings is achieved.

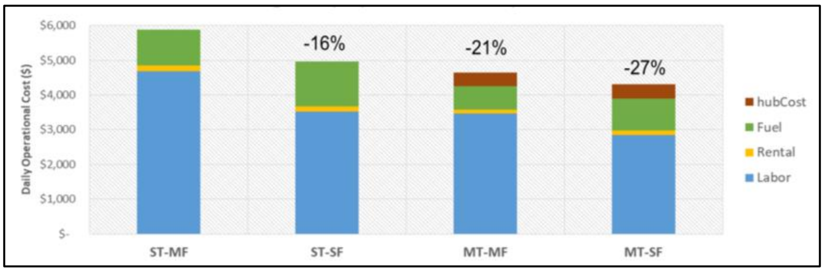


Figure 6: Average daily operational cost by scenario with percentage reduction

The average daily travel miles are shown in Figure 7. In fact, the functional modularization, increased the total travel miles due to the additional travel for installation. However, smaller, cheaper and more environmentally friendly vehicles used for installation routes make the cost increase not proportional to the travel mile. Similar impact is made with tier-2 routes. This also has indirect environmental and social impacts which is not measured explicitly in this paper.

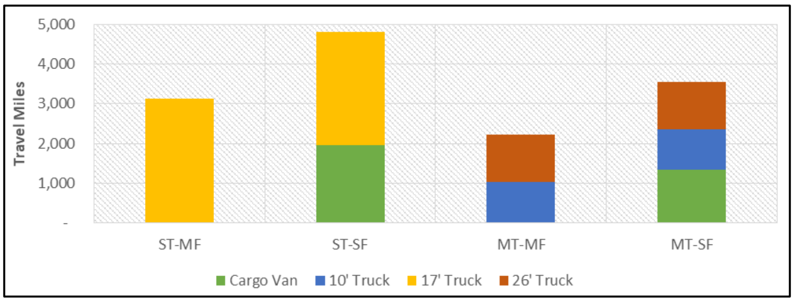


Figure 7: Average daily travel miles by vehicle type by scenario

Functional modularization, especially, changes labor requirement significantly. Figure 8 shows average daily labor hour by scenario, distinguished by types of personnel: driver and installer. Functional modularization reduces the labor hour of installers, who tend to have higher pay rate than drivers as they have more skillset. The total labor hours can also be saved with functional modularization by removing the waste of labor hour for 2-personnel team. Hierarchical modularization also has similar impact as tier-1 routes are served by a single driver with more consolidation.

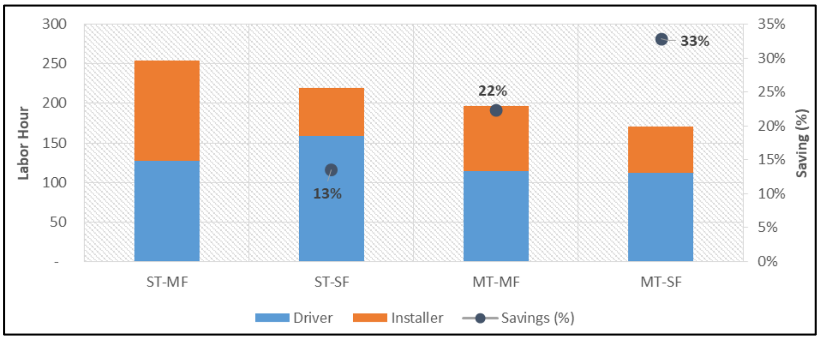


Figure 8: Average daily labor hour by type of personnel by scenario

5 Conclusion

This paper propose a concept of modularization of transportation and delivery in a hyperconnected logistics system. Three modularization schemes are described: regional, hierarchical and functional. All three types of transportation modularization have potential to facilitate dynamic consolidation, routing, and operation. The flexibility is especially beneficial in context of Physical Internet (PI) and multi-player operations. Modularized delivery and transportation better supports PI inspired hyperconnected logistics system while PI is also an enabler of transportation modularization. For example, PI containers can greatly reduce handling cost and complexity to justify the increased number of stops under modularized transportation. On the other hand, modularized transportation maximize the benefit of container operations. Then, the impact of hierarchical and functional modularization is demonstrated for the last mile delivery of furniture and large items. Simulation results shows that operational cost can be saved by 21% with hierarchical modularization and by 16% with functional modularization. Together, up to 27% savings can be achieved.

The paper opens various future research avenues by proposing an alternative transportation planning and operation scheme for more dynamic logistics system. Several limitations of the paper and future research topics are listed here. Firstly, rigorous mathematical routing and scheduling models and practical algorithms for each or any combination of the modularization scheme need to be developed. There are a few literature on regional or hierarchical modularization (Fazili et al., 2017; Ballot et al., 2012A; Crainic et al., 2009), but less literature exists on functional modularization or on combination. This is not only limited to routing problem. Integrated solution approach spanning a hub operations modeling or network design as well can generate more unbiased intuition. Secondly, the platform or system architecture which integrates the transportation and delivery modules needs to be designed thoroughly. A distributed platform structure can fit better than a single platform model especially under PI context which results in dynamic on-demand multi-player operations. Thirdly, digital and physical interface designs are crucial for modularized transportation. Digital interface needs to ensure the access to critical information such as destination information or arrival times for seamless operations while protecting confidentiality. Physical interface for the transportation modularization is PI hubs. Efficient facility design and use of PI containers can facilitate a smooth connection between modules. Forthly, the impact of modularization on multi-modal transportation system can also be a promising future research avenue. In such case, designing crossdocking or handling operation and capacity requirement measuring at multi-modal hubs as in Ballot et al. (2012B) can be a critical part of the study. Lastly, an extensive application to a real case study or actual implementation can better demonstrate the impact of modularized transportation.

References

- Arslan, A. M., Agatz, N., Kroon, L., & Zuidwijk, R. (2019). Crowdsourced delivery—A dynamic pickup and delivery problem with ad hoc drivers. Transportation Science, 53(1), 222-235.
- Baldwin, C. Y., Clark, K. B., & Clark, K. B. (2000). Design rules: The power of modularity (Vol. 1). MIT press.
- Ballot, E., Gobet, O., & Montreuil, B. (2012A). Physical internet enabled open hub network design for distributed networked operations. In Service orientation in holonic and multi-agent manufacturing control (pp. 279-292). Springer, Berlin, Heidelberg.
- Ballot, E., Montreuil, B., & Thivierge, C. (2012B). Functional design of physical internet facilities: a road-rail hub.
- Bask, A., Lipponen, M., Rajahonka, M., & Tinnilä, M. (2010). The concept of modularity: diffusion from manufacturing to service production. Journal of Manufacturing Technology Management.
- Bergmann, F. M., Wagner, S. M., & Winkenbach, M. (2020). Integrating first-mile pickup and last-mile delivery on shared vehicle routes for efficient urban e-commerce distribution. Transportation Research Part B: Methodological, 131, 26-62.
- Campbell, A. M., & Savelsbergh, M. (2004). Efficient insertion heuristics for vehicle routing and scheduling problems. Transportation science, 38(3), 369-378.
- Clarke, G., & Wright, JW (1964). Scheduling of vehicles from a central depot to a number of delivery points. Operations Research, 12 (4), 568-581.
- Crainic, T. G., & Montreuil, B. (2016). Physical internet enabled Hyperconnected City logistics. Transportation Research Procedia, 12, 383-398.
- Crainic, T. G., Ricciardi, N., & Storchi, G. (2004). Advanced freight transportation systems for congested urban areas. Transportation Research Part C: Emerging Technologies, 12(2), 119-137.
- Crainic, T. G., Ricciardi, N., & Storchi, G. (2009). Models for evaluating and planning city logistics systems. Transportation science, 43(4), 432-454.
- Fazili, M., Venkatadri, U., Cyrus, P., & Tajbakhsh, M. (2017). Physical Internet, conventional and hybrid logistic systems: a routing optimisation-based comparison using the Eastern Canada road network case study. International Journal of Production Research, 55(9), 2703-2730.
- Ferrell, C. L. (2016). High driver turnover among large long-haul motor carriers: causes and consequences (Doctoral dissertation, UNIVERSITY OF NORTH TEXAS).
- Gontara, S., Boufaied, A., & Korbaa, O. (2018, October). Routing the PI-Containers in the Physical Internet using the PI-BGP Protocol. In 2018 IEEE/ACS 15th International Conference on Computer Systems and Applications (AICCSA) (pp. 1-8). IEEE.
- Gutierrez, A., Dieulle, L., Labadie, N., & Velasco, N. (2018). A hybrid metaheuristic algorithm for the vehicle routing problem with stochastic demands. Computers & Operations Research, 99, 135-147.
- Hakimi, D., Montreuil, B., & Ballot, E. (2012). Simulating a Physical Internet Enabled Logistics Web: the Case of Mass Distribution in France, ISERC 2012 Proceedings.

- Holtta, K. M., & Salonen, M. P. (2003, January). Comparing three different modularity methods. In ASME 2003 international design engineering technical conferences and computers and information in engineering conference (pp. 533-541). American Society of Mechanical Engineers Digital Collection.
- Kaboudvand, S., Montreuil, B., & Savelsbergh, M. (2021). Hyperconnected Megacity Parcel Logistic: Joint Parcel Routing and Containerized Consolidation. arXiv preprint arXiv:2103.02080.
- Kim, N., Montreuil, B., Klibi, W., & Kholgade, N. (2021). Hyperconnected urban fulfillment and delivery. Transportation Research Part E: Logistics and Transportation Review, 145, 102104.
- Koç, Ç., Bektaş, T., Jabali, O., & Laporte, G. (2016). Thirty years of heterogeneous vehicle routing. European Journal of Operational Research, 249(1), 1-21
- Laporte, G. (2009). Fifty years of vehicle routing. Transportation science, 43(4), 408-416.
- Meller, R. D., Montreuil, B., Thivierge, C., & Montreuil, Z. (2012). Functional design of physical internet facilities: a road-based transit center.
- Min, H., & Lambert, T. (2002). Truck driver shortage revisited. Transportation journal, 5-16.
- Montreuil B. (2011): Towards a Physical Internet: Meeting the Global Logistics Sustainability Grand Challenge, Logistics Research, v3, no2-3, 71-87.
- Montreuil, B., Buckley, S., Faugere, L., Khir, R., & Derhami, S. (2018). Urban parcel logistics hub and network design: The impact of modularity and hyperconnectivity.
- Montreuil, B., Meller, R. D., & Ballot, E. (2010). Towards a Physical Internet: the impact on logistics facilities and material handling systems design and innovation.
- Nazari, M., Oroojlooy, A., Snyder, L., & Takác, M. (2018). Reinforcement learning for solving the vehicle routing problem. In Advances in Neural Information Processing Systems (pp. 9839-9849).
- Parragh, S. N., Doerner, K. F., & Hartl, R. F. (2008). A survey on pickup and delivery problems. Journal für Betriebswirtschaft, 58(2), 81-117.
- Pekkarinen, S., & Ulkuniemi, P. (2008). Modularity in developing business services by platform approach. The International Journal of Logistics Management.
- Rajahonka, M. (2013). Views of logistics service providers on modularity in logistics services. International Journal of Logistics Research and Applications, 16(1), 34-50.
- Sarraj, R., Ballot, E., Pan, S., Hakimi, D., & Montreuil, B. (2014). Interconnected logistic networks and protocols: simulation-based efficiency assessment. International Journal of Production Research, 52(11), 3185-3208.
- Smilowitz, K. R., & Daganzo, C. F. (2007). Continuum approximation techniques for the design of integrated package distribution systems. Networks: An International Journal, 50(3), 183-196.
- Solomon, M. M. (1987). Algorithms for the vehicle routing and scheduling problems with time window constraints. Operations research, 35(2), 254-265.
- Sternberg, H. S., & Denizel, M. (2020). Toward the Physical Internet—Logistics Service Modularity and Design Implications. Journal of Business Logistics.
- Toth, P., & Vigo, D. (Eds.). (2014). Vehicle routing: problems, methods, and applications. Society for Industrial and Applied Mathematics.
- van Luik, S., Fiebig, T., Fahim, P., de Waard, P., & Tavasszy, L. (2020). On the value of the Digital Internet/Physical Internet analogy. Journal of Supply Chain Management Science, 1(3-4), 87-103.
- Winkenbach, M., Kleindorfer, P. R., & Spinler, S. (2016). Enabling urban logistics services at La Poste through multi-echelon location-routing. Transportation Science, 50(2), 520-540.

Conceptual Framework for Hyperconnected Package Transport Logistics Infrastructure

Sahrish Jaleel Shaikh[1,2,3] , Nayeon Kim[1,2,3], Benoit Montreuil [1,2,3,4], Priidik Vilumaa[1,5]

1. Physical Internet Center
2. Supply Chain and Logistics Institute
3. H. Milton Stewart School of Industrial & Systems Engineering, Georgia Institute of Technology, Atlanta, United States
4. Coca-Cola Material Handling & Distribution Chair
5. VOPT Consulting

Corresponding author: sahrish.shaikh@gatech.edu

Abstract: *In the context of Physical Internet, where dynamic consolidation and flexible multi-player operations are required, the current large-scale integer programming based package transportation network operations may not only become inflexible and computationally heavy, but can often create significant ad-hoc plan changes. This paper presents a conceptual framework of an alternative way of designing, planning, and operating a package transportation network, Hyperconnected Package Transport Logistics Infrastructure (HyPTLI), inspired by the Physical Internet Initiative. The HyPTLI framework leverages five distinctive features of Physical Internet: multi-tier interlaced mesh networks, service-based package routing schemes, dynamic containerized consolidation, service-driven on-demand inter-hub shuttling, and open protocols and digital platform. The paper describes the design components of HyPTLI and positions them into a comprehensive design framework.*

Keywords: *Package Transportation, Logistics Infrastructure, Physical Internet, Hyperconnectivity, Mesh Network, Routing Protocols, Network Architecture, Multi-tier Network, Service Network*

1. Introduction

In recent years, the role of freight transportation and package delivery has increased, supporting the growth of e-commerce and home delivery. However, in today's urban package delivery and e-commerce logistics, plans are often changed, and ad-hoc decisions are taken to ensure service levels. The industry requires a more dynamic and flexible model similar to the one used by internet-based communication networks, where packets travel across a mesh network, a concept that has been explored in the Physical Internet Initiative. We propose a conceptual framework for the functional design of Hyperconnected Package Transport Logistics Infrastructure (HyPTLI) inspired by the Physical Internet Initiative. HyPTLI is a sub-element comprising Hyperconnected Logistics Infrastructure (HLI). HLI is a logistics infrastructure that supports the operations of the five logistics webs: mobility, distribution, realization, supply, and service webs (Montreuil et al., 2012). HLI includes physical resources ranging from factories and distribution centers to hubs, as well as open operational protocols and interfaces for each party. HyPTLI is a part of HLI that supports mobility.

The HyPTLI framework leverages distinctive features: (1) multi-tier interlaced mesh networks, (2) package routing schemes based on service requirements, (3) dynamic containerized consolidation, (4) service-driven on-demand inter-hub shuttling, and (5) open protocols and digital platforms.

First, the logistics networks are structured in multiple interconnected tiers where each tier induces a distinct degree of consolidation and covers different territorial sizes. Such networks are composed of customer interfaces, logistics hubs, and deployment centers (including warehouses, distribution and fulfillment centers). An early example of tiered networks can be found in two-tier urban delivery networks (Crainic et al., 2009) and a more advanced architecture for urban multi-tier logistics networks is proposed by Montreuil et al. (2018).

The second feature is its emphasis on dynamic package routing, where package is used here as a generic term referring to the basic freight unit ordered, shipped, transported, stored, and delivered. Traditional approaches assign a pre-determined path to a package based on its origin, destination, and service level, and then focus on vehicle routing. The HyPTLI supports semi-dynamic or fully dynamic package routing, where the next destination and loading timing are determined to maximize carrier utilization while respecting service level and maximum hub dwell time constraints.

The third feature is dynamic containerized consolidation. This is enabled by containerized operations across logistics networks, according to the Physical Internet in modular standardized packaging, handling, and transport containers such as introduced by Montreuil et al. (2015). This consolidation groups together up to a specific hub packages that share this hub as an intermediate or final hub destination, and whose service level engagement is expected to be respected by joining this group based on its target arrival time to the hub. Using consolidated containers reduces drastically the number of handled items, and the dimensional modularity of container sizes helps maximizing space usage while enabling appropriate shipment frequency.

The fourth feature is service-driven on-demand inter-hub shuttling, beyond relying on detailed vehicle routes and schedules often optimized way ahead of their implementation. Service-driven shuttling utilizes simple routes (single leg or a few legs) with adequate frequency to satisfy service requirements. On-demand transportation, which can be exemplified by businesses such as Convoy, Quay and Uber Freight, adds transportation capacity flexibility which is especially beneficial when demand is highly uncertain. Combining inter-hub shuttling and on-demand transportation into service-driven on-demand inter-hub shuttling minimizes planning determinism and complexity while maximizing agility of operations.

The fifth feature is the reliance on open protocols and digital platforms easing the interconnection of network design components and operations via seamless and smart information flow among players in the system, leveraging sensor networks and cyberphysical systems.

The rest of the paper is organized as follows. Section 2 reviews relevant literature and positions the paper. The design elements of the proposed model are described in Section 3. The paper concludes at Section 4 by summarizing the design and motivation of the model and listing further research opportunities.

2. Literature Review

2.1 Physical Internet

The Physical Internet Initiative is applying the principles of the Internet to inspire positive transformation of supply chains, logistics, and transportation, leading to order-of-magnitude improvements in the capability, efficiency, and sustainability of fulfilling society’s worldwide demand for physical objects. According to Montreuil (2015), the Physical Internet (PI) may be defined a hyperconnected global logistics system enabling seamless open asset sharing and flow consolidation, through standardized encapsulation, modularization, protocols and interfaces. It is said to be hyperconnected as its components and actors are intensely interconnected on multiple layers, ultimately anytime and anywhere. These interconnectivity layers include digital, physical, operational, business, legal, and interpersonal layers.

It relies on open interconnected networks, using a set of protocols and interfaces, in order to send and receive physical goods contained in standard modular containers – synonymous to packets of information in the wireless Internet technology (Crainic & Montreuil (2016)). Current logistics is dominated by a combination of point-to-point transport and hub-and-spoke transport. Even though these two ways are feasible in PI, distributed multi-segment intermodal transportation are more popular (Gakis & Pardalos, 2017).

Currently, long-haul transportation is either done by a solo driver or a 2-driver team assigned to the multi-day trip. Team drivers can take turns driving to cope with labor hour regulations, while a solo driver has to stop the truck and rest. Once reaching the destination and having completed the delivery, the drivers will be dispatched to pick up a new shipment, perhaps returning towards the original point of departure, in order to avoid empty travel. Due to poor planning or lack of deliveries, a return with an empty truck might also occur. As the drivers are usually away from home for long durations, this poor lifestyle has caused driver retention to become one of the biggest problems being faced by the trucking industry (Mittal et al., 2018). Montreuil et al. (2010) proposed to shift from point-to-point travel to a distributed multi-segment travel of π-containers (Sallez et al., 2015) through PI networks. If a package needs to move from Hub A to Hub B, it may dynamically move through a series of hubs in its best possible way using cooperative routing algorithms and protocols. The driver relay concept in the Physical Internet works in a way that once a driver transports a trailer a few hours away and would then pick up another delivery returning to the point of departure – so the driver teams usually work on certain segments going back and forth, but most often loaded with a trailer. Similarly, all segments are serviced this way meaning that a second driver-truck duo would soon afterward pick up the delivery and move it another segment forward . The process would be repeated until all the containers of the delivery have reached the final destination (Montreuil, 2011).

2.2 Multi-tier delivery

Multi-tier systems can be found in urban logistics, mainly as two tiers. It has been actively studied for more than a decade (Crainic et al., 2004; Crainic et al., 2009; Bektas et al. 2015; Cuda et al., 2015; Crainic et al., 2016; Winkenbach et al., 2016;). In general, the two-tier urban logistics system involves satellite facility which serves as intermediary transshipment and consolidation point between urban consolidation center (UCC) or peri-urban distribution center (DC) and point of delivery or pickup. The first tier covers more consolidated flow from UCCs to satellites or from peri-urban DCs to satellites. The second tier covers last mile in smaller area from satellites to point of delivery or pickup. This enables the use of smaller vehicles or other types of transportation mode such as bicycle for the second tier operations. While most of city logistics studies involve only inbound flow into the urban area, Crainic et al. (2012) and Crainic et al. (2016) modeled the 2-tier operation including inbound, outbound and intra-city flows.

More elaborate multi-tier systems have been introduced as a major component of hyperconnected logistics, notably in urban environments (e.g. Montreuil, 2011; Crainic and Montreuil, 2016). In this paper, as detailed subsequently, we build upon and extend the hyperconnected multi-tier urban package logistics systems introduced by Montreuil et al. (2018).

2.3 Logistics Infrastructure

The design problem of Hyperconnected Transport and Delivery Logistics Infrastructure is closely related to a network design problem, and more specifically a transportation service network design (SND) problem (Magnanti and Wong, 1984; Balakrishnan et al. 1989; Crainic, 2000; Wang et al, 2019; Perez et al., 2020). Although majority of SND research focuses on general or large-scale transportation networks, it can be focused on urban delivery network as in Scherr, et al. (2018). The SND problem generally tackles tactical planning issues such as service design, scheduling, routing, and empty resource repositioning (Crainic, 2000) via integer programming on time-space network (Wong and Qi, 2019; Ng and Lo, 2016; Crainic, 2000; Magnanti and Wong, 1984). The critical dimensions of the SND problem are capacity, commodities, service, and modes. Capacity is the ability of network to handle flows or volume which often includes time dimension as well. Depending on the situation, it can be modelled as constraints or can be considered a decision variable. Commodities or demand determine the flow to be served in the network, defined through origin, destination, service, and volume. Deterministic demand was the dominant assumption is earlier literature, yet some recent models tend to be built on stochastic demand (e.g. Ng and Lo, 2016). The types of commodities can vary ranging from packages to people (Wang and Qi, 2019). Service refers the maximum time to complete transportation of a commodity from origin to destinations which gives different priorities to commodities. Service is also closely related to the design of routes and choice of transportation modes. For example, most express or one-day deliveries currently need to be done in air in large area such as USA or Europe (Perez et al. 2020).

The proposed hyperconnected logistics system also targets to make tactical level decisions to efficiently satisfy uncertain demand as typical SND models. Distinctively to the core of the SND literature which models it as a large-scale complex integer program, and aims to solve this program optimally or heuristically for a given shipper or service provider, it is rather proposed to recognize from the ground up the huge multi-stakeholder size of a hyperconnected logistics infrastructure, and to emphasize its design for resilience as recommended in Montreuil (2011). It must therefore be aimed to support seamlessly dynamic adjustment of planning, scheduling or routing decisions by stakeholders based on demand realization. It must be designed for agile avoidance, smooth absorption, and fast rebound when faced with disruptions induced by demand changes, natural disasters, pandemics, and aggressions.

2.4 Mesh networks

In computer architecture, a mesh network as a network where each node is connected to many other nodes, configured to allow connections to be rerouted around broken paths, with the signal hopping from node to node until it reaches its destination (Navda et al., 2005; Akyildiz, 2009) define . A mesh network can either be fully or partly connected . In a full mesh topology, each network node is connected directly to each of the others with total connections n(n-1)/2. In a partial mesh topology, only some nodes have multiple connection partners. The decision of which nodes to mesh depends on the overall network traffic as well as the risk of failure of a connection or node.

In a mesh network, the nodes not only act as hosts but rather as routers, forwarding packets on behalf of other nodes to help them reach their destination. A wireless mesh network (WMN) with such mesh routers and clients is dynamic and self-organized as the data packets may re-route while the nodes maintain mesh connectivity amongst themselves. Such operations essentially create an ad hoc but resilient network leading to reduced upfront costs and easy network maintenance (Akyildiz (2009)). With the multi-hop and alternate routing architecture, there is a high fault tolerance in WMNs. Zhao and Raychaudhuri (2007) have evaluated the performance of multi-tier wireless mesh networks where the mesh routers are classified into different hierarchical tiers. They propose a network architecture that is based on three tiers of wireless devices: low-power nodes with limited functionality called the mobile nodes, higher-power radio forwarding nodes that route packets between radio links, and access points that route packets between radio links and the wired infrastructure. Zhao and Raychaudhuri (2007) classify all traffic as inter-cluster or intra-cluster; intra-cluster traffic is if the origin and destination hubs belong to the same forwarding node while otherwise, it is called intra-cluster traffic. They assume that the intra-cluster traffic is relayed by the forwarding nodes as there is no direct connection between mobile nodes. The inter-cluster traffic is forwarded from the origin to the destination through multiple hops that may involve several Forwarding Nodes and Access Points depending on the network and flows. In line with Physical Internet, our proposed approach is inspired by digital mesh networks, yet adapted to the logistics and transportation reality, where the hubs become nodes and packages are forwarded like data packets through the intermediary nodes of multi-tier mesh networks.

3 Proposed HyPTLI Conceptual Framework

In this section, we present the conceptual framework for the functional design of HyPTLI. We discuss the key concepts of our proposed framework: multi-tier interlaced mesh networks, package routing schemes based on service requirements, containerized consolidation, as well as service driven shuttling and on-demand transportation models.

3.1 Multi-tier Interlaced Mesh Networks

Multi-tier interlaced mesh network is where the one or more tiers of mesh networks are interlaced in hierarchical layers. Mesh networking introduces an innovative shift from the topology of traditional networking models. It is an interlaced structure where the nodes allow point-to-point while eliminating the need for a central entity. The hubs in the network then act as routers and forward the traffic to adjacent hubs, e.g. within 4-hour distance, enabled by seamless consolidation and crossdocking operations. Without a centralized entity managing the network, the network is not sensitive to a single point of failure. Even if a certain hub is compromised, the network is still capable of operating by finding alternative paths to the desired destination. Consequently, this means more resilience in the network. Such structure also enables more flexible routing. A mesh network would benefit from node density to form a greater number of alternate paths while subsequently, a sparse network of hubs may not form an effective mesh network. That being stated, there are different business models that can be exploited by a group of sparse networks to become a mesh network with shared resources. These models can comprise of hubs can have distinct ownership yet be tied together by a business model whereby the infrastructure is shared by all businesses. When there is more than one tier, hubs are assigned into different tiers where each tier has different level of geographical coverage and consolidation level. Montreuil et al. (2018) propose to leverage a multi-level structuring of space such as unit zones, local cells, (urban) areas, regions, blocks, and the world. Each tier of meshed networks covers a level of space structuring, as generically depicted in Figure 1. The multi-tier interlaced mesh network can potentially create simpler logistics flow than traditional network topologies such as Point-to-point (P2P) and Hub-and-spoke (H&S). Figures 2 and 3 contrast logistic flows on a traditional network and multi-tier mesh networks.

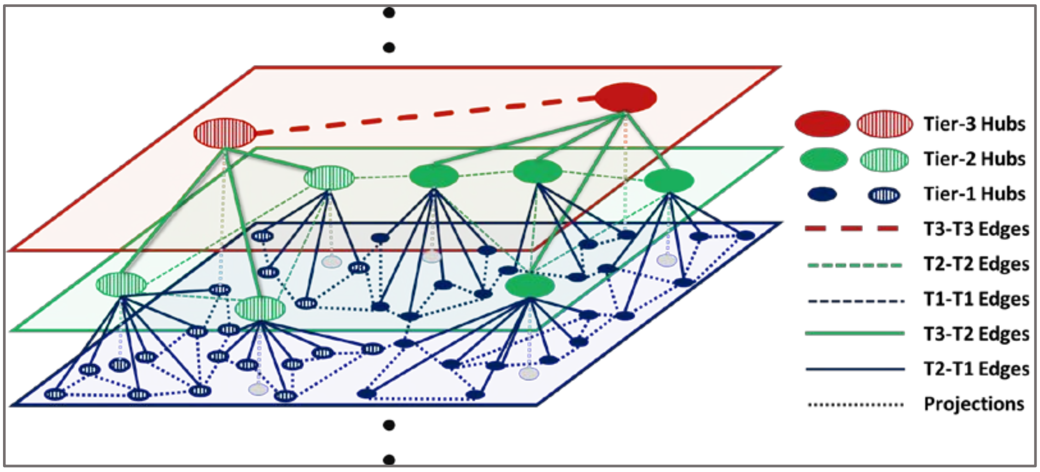


Figure 1: Conceptual diagram of interlaced mesh networks showing three tiers

Figure 2: Typical logistics flow in a traditional logistic network (P2P and H&S)

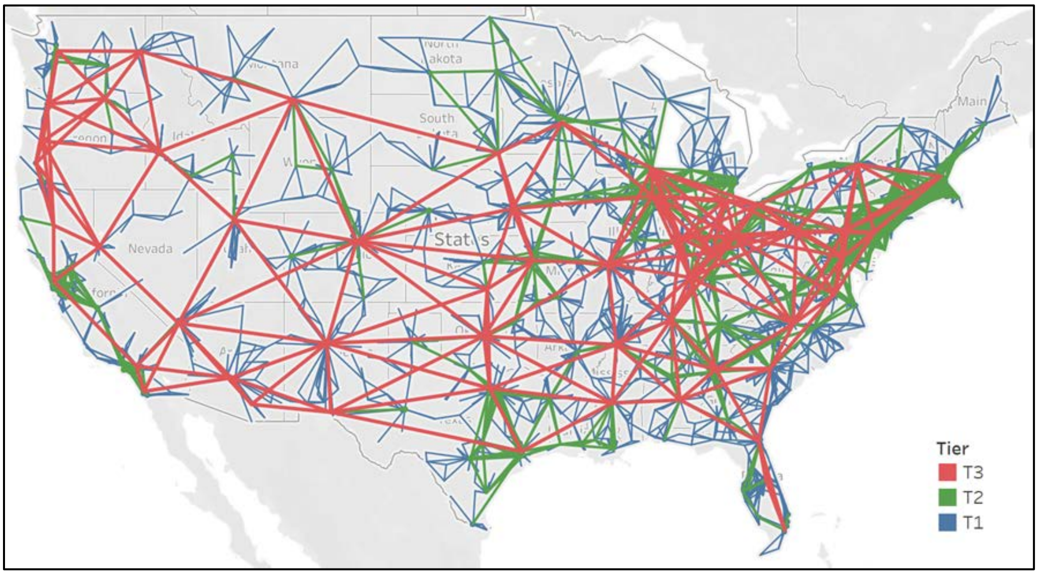


Figure 3: Hyperconnected Package Transport Logistics Infrastructure based on Multi-Tier Mesh Networks

3.2 Package Routing

The hierarchical mesh structure of the proposed interlaced mesh networks allows us to reduce the size of the pathfinding problem. Instead of relying on heavy optimization and planning of long multi-hub routes, packages are routed flexibly at each hub based on current flow and transportation schedule. In general, sending packages to a higher tier increases the chance of consolidation and carrier fill rates in the higher-level hub networks. High-tier hub networks consist of crossdocking hubs, railyards, airports, and their unimodal and multimodal connections. As suggested in Figure 4 , an alternative way is to not escalate to the next tier but rather continue towards the destination on the current tier. For certain origin-destination pairs with short distance, shipments may not need to be forwarded to a higher tier.

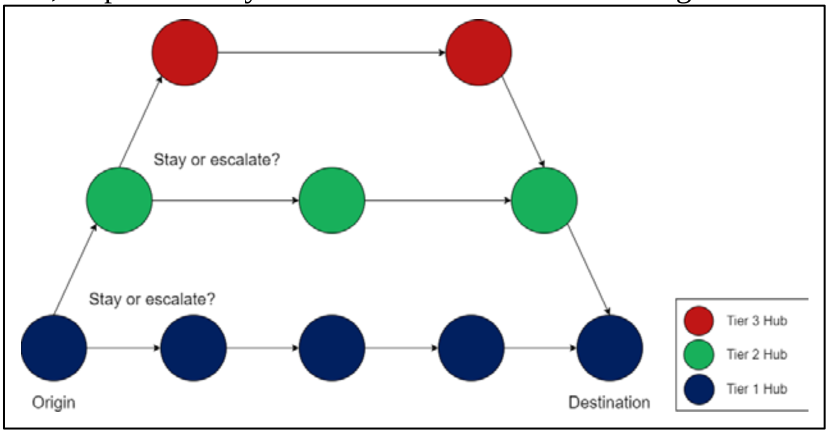


Figure 4. Hierarchical Mesh Routing

When a package enters the system, three elements already known: origin (O), destination (D), and service level (S). Using the service level representing maximum lead time (i.e. 2 days, 3 days etc.) and an initial route, we can calculate the slack time of the packages where the ***slack time*** is defined as the excess time a package has, after accounting for the travel time from origin to destination and the sum of processing and waiting times at the hubs. The sum of processing and waiting times at a hub is defined as ***dwell time***. Package paths with travel and processing times are illustrated in Figure 5. Once we know the slack time, this can be distributed amongst the intermediary hubs, we call this the ***maximum dwell time*** of the package at the hub. The slack time and the maximum dwell time of each package are dynamically updated based on the operations of the network. When containers are used, the slack time and maximum dwell time of a container are respectively determined by the most urgent package's slack and maximum dwell times in the container.

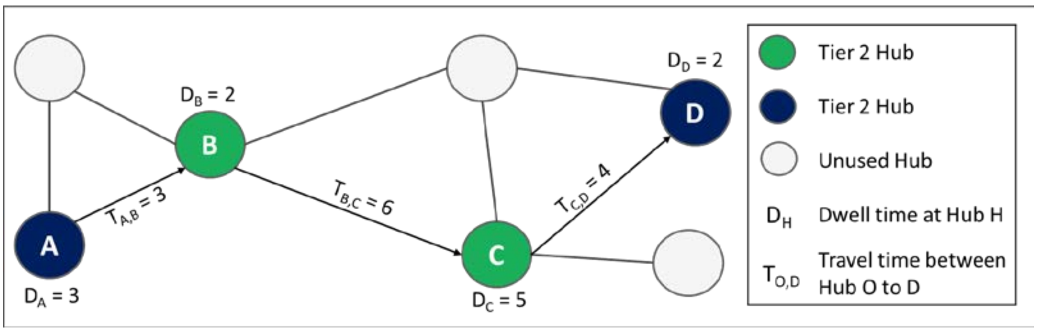


Figure 5. Package Path with Travel and Dwell Times

The routes of packages are re-evaluated and possibly reassigned if there are more chances of consolidation and better truck fill rate on the alternate routes. Since the dwell times are updated dynamically, it is important to evaluate the feasibility of all alternative routes as well. Moreover, in the case of more advanced information sharing, potential package paths can be evaluated at every step of the route based on transportation availability, consolidation opportunity, urgency of the packages as well as their destination. Under such a dynamic model, packages with same origin, destination, and service level can take different paths to fully utilize logistics resources. In addition to higher cost efficiency, packages could arrive to the destination faster, as better decisions are being captured during the shipping process.

3.3 Dynamic Containerized Consolidation

Consolidation in the mesh network is executed using π-containers of modular sizes. Products or packages can be consolidated in a container and reconsolidated dynamically based on their path and service requirements. For example, products sharing same destination, next hub, or next-K hubs, with similar slack time, can be consolidated in a single container. Containerization decisions balance tradeoff between maximizing consolidation and container fill-rate, minimizing re-containerization effort and maximizing routing flexibility. Containerization enables automation at hubs as π-containers can be loaded, unloaded and cross-docked easily and may not require manual labor through standardization. It offers a smart, more reliable and efficient tracking with the use of connected sensors. With the use of modular containers, the operations at the hub change from a dominancy of package sorting to container handling, which means that containerization has the potential of increasing hub throughput within a given hub footprint. For example, if a container is comprised of 30 packages, then they will be scanned just once rather than 30 separate package scans, and they will require a single handling rather than at least thirty.

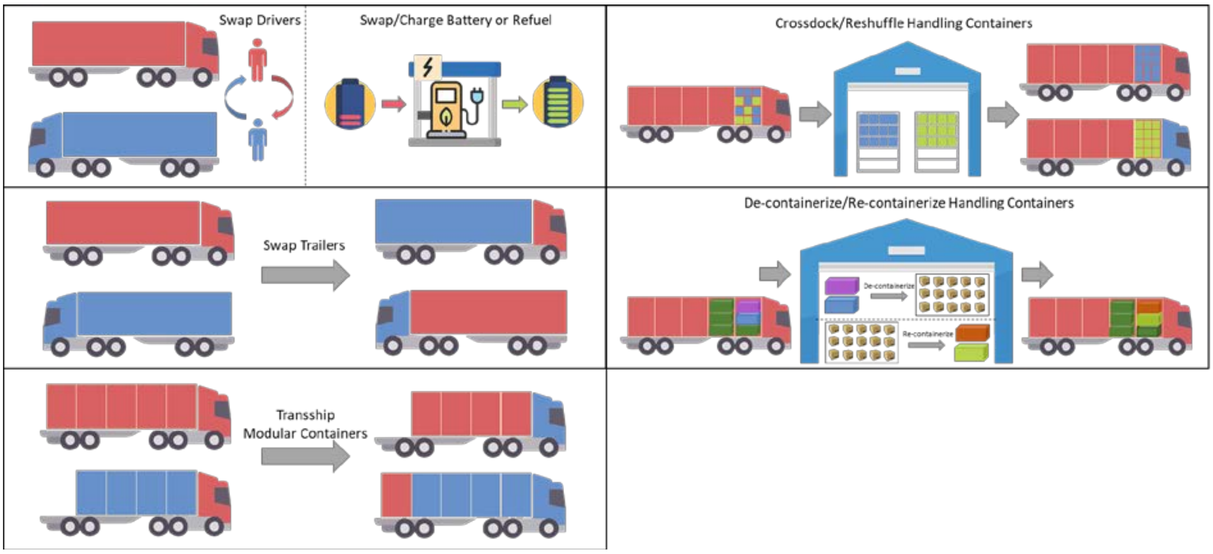


Figure 6: Various types of hub-based operations

Hub operations are visualized in Figure 6. As we propose relay-style driving, the drivers may swap the entire vehicle, or they may stop at a charging station to replace batteries in case of an electric vehicle or re-fuel the truck. The drivers may also just unhook the trailer and swap the trailer while keeping their own flat-bed truck. As we are using modular containers, it is easy to transship selected modular containers onto another truck. There are handling containers within these modular containers which can be reshuffled at a hub. This is done using racks which are placed at the staging station and as the modular containers are opened, the handling containers are placed into their respective racks. In Figure 6, each rack color represents a different destination, and the containers are placed accordingly. Later, these racks can be put into a trailer directly for the next destination. We minimize the intra-facility travel of the containers by using such re-shuffling of containers using racks. Lastly, the modular containers may need to be de-containerized and then re-consolidated in a hub. These operations can be expanded to a more granular level such as the package level.

3.4 Service-driven shuttling and on-demand transportation models

Here, we introduce the concept of ***service-driven shuttling***, coupled with ***on-demand transportation.*** In brief, on-demand transportation encompasses the transportation options that can be used flexibly bypassing planning. Although it has more flexibility, it is typically more expensive than planned transportation and imposes its own risk as the availability of transport resources is not guaranteed, especially when the market is not mature. The early business model can be exemplified with Uber Freight or Convoy. Service-driven shuttling refers to pre-contracted transportation guaranteeing each leg to be served with agreed frequency and responsiveness. The on-demand trucking and service-driven shuttling supplement each other, setting balance between service guarantee and operational flexibility. In other words, it ensures the transportation planning to remain simple and flexible while not compromising service capability. Figure 7 illustrates and contrasts the traditional transportation scheme with the one proposed here.

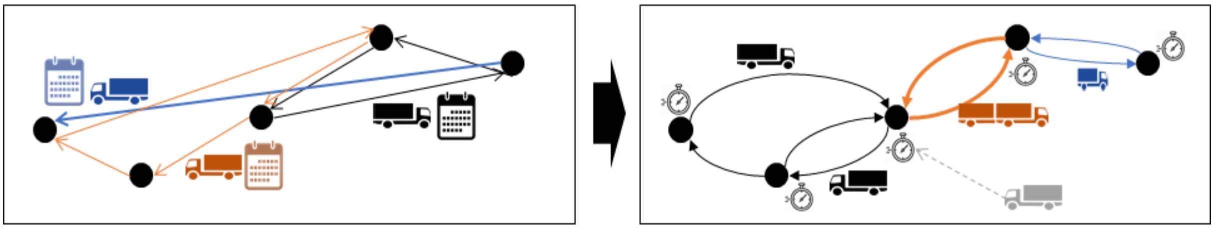


Figure 7: Illustration of traditional planned transportation (left) and service-driven shuttling with on-demand transportation (right) over network of 5 logistic hubs

Service-driven shuttling can be seen as frequency-based contracted transportation. It is called shuttling as each route typically serves closed routes consisting of only one or few legs repeatedly. The frequency is determined to meet service requirements, and it can be a constant or a range. The frequency is based on the forecasted volume of packages at each leg and the tier that segment belongs to. As meeting service level engagements is essential in package delivery, this must be considered as well, and it may impose a minimum frequency to each leg. For example, for the legs with very low volume, carriers of smaller size can be used instead of reducing frequency too much. This is to ensure that the packages do not keep waiting too long to be picked up at the hub. As the system determines the path of the packages as soon as they enter the system, it is possible to track the estimated time of arrival of packages at each hub in advance. Using this predicted arrival data, the frequency of the trucks is further refined to encourage consolidation and improve fill rate of trailers.

Note that when there is no or very low volume to be shipped, and none of the packages is urgent, there is no need for shuttling. In the same line, when such service-driven shuttling service is contracted with a third-party logistics provider, desirable frequency is determined as a 'minimum' frequency or range to have flexibility to adjust to actual flow. In other words, the system should be able to send confirmation to incoming trucks x-hours in advance so that shuttles will be used when there is enough volume or urgency that justifies shuttling. This is also closely related to information integrity as not only demand forecast but also incoming flows from other hubs must be projected so as to estimate the volume in each leg in next x-hours.

On-demand transportation supplements service-based shuttling, which is operated based on a plan or contract. This is mostly when the actual flow significantly exceeds the forecast over which the shuttles are planned. The reason why not all transportation is done by on-demand transportation service providers is due to the higher cost and uncertainty in supply. However, in an ideal case where on-demand transportation costs the same or not significantly more than planned transportation and there is sufficient and reliable market, it can replace planned shuttling completely. We propose that on-demand carriers be called based on the basis of containerized package urgency to respect maximum dwell times. When there is no departing carrier within the range of maximum dwell time of a package, it is flagged as an urgent package. If the existing transportation cannot be adjusted to ship the urgent package on time, on-demand transportation is triggered.

3.5 Open protocols and Digital Platform

In our proposed model, information flows seamlessly throughout the network. For example, once a truck departs from a hub, the next hub is updated with the estimated time of arrival of the truck and the contents inside the truck including both containers and packages. It also transmits data on the path of each package, while protecting privacy of private confidential information on packages. This enables the next hub to evaluate which containers can be cross-docked and need not be opened. This decision is taken based on the consensus within the packages in a container. If all packages share the same next hub, then there is consensus (100%) and the container can be easily cross-docked. If the packages do not have a consensus, then alternate feasible routes may be evaluated to keep the consolidation intact for as long as possible.

To enable the open protocols to work seamlessly, there will be pieces of information that will be stored and updated continuously. This information includes the package's logistic details such as origin, destination, and the committed service level. Once the package enters the system, the slack time and dwell times are calculated and updated continuously until the package reaches its destination. For every carrier, there is information stored about the packages and containers it carries. The departure of a carrier from a hub and the estimated arrival time to the next hub are also continuously updated to enable the entities to always have the most recent data.

In the proposed model, the entities have access to relevant information that is required to make the decisions. For example, when the hubs have to order on-demand trucks, the decision is to be based on the containers and their encapsulated packages that will be loaded onto the truck. The hubs will also have information on the capacity of the trucks available to order as well as the total volume of the packages to be shipped so the best fitting vehicle can be ordered for each trip. Moreover, the hubs will have information on the slack and dwell time of the packages as the on-demand trucks will be ordered based on the package urgency and the estimated order-to-arrival time of the truck.

There are many advantages of having better tracking system facilitated by the smart IoT technology embedded in π-container. One advantage is the ability to track urgency of each container based on maximum dwell time of packages in the container. That is, when loading trailers, more urgent containers can be loaded first. This is logical order and does not determine the physical order, where the physical order is determined by capacity utilization problem or unloading orders. The logical order ensures more urgent packages to be shipped first to prevent unnecessary emergency shipping in the future.

4 Conclusion

The package logistics industry heavily relies on hub-and-spoke and point-to-point networks, and large-scale planning models followed by significant operational adjustment efforts to adjust misfitting plans due to uncertainties and disruptions. Moreover, the trucking industry faces a serious and chronic problem of high driver turnover rate - typically more than 100% - with staggering associated costs. We have introduced a functional design for a Hyperconnected Package Transportation Logistics Infrastructure that facilitates flexible operation with potentially lower cost and higher resilience and sustainability, while providing better work environment for drivers as well where they can ‘sleep at home'.

In our proposed logistic infrastructure model, we structure the logistics network in multiple interconnected tiers where each tier induces a distinct degree of consolidation and covers a distinctive territorial size. Unlike the traditional approaches that assign a pre-determined path to a package, the HyPTLI supports semi-dynamic or fully dynamic package routing, where the next destination and loading timing are determined to maximize carrier utilization while respecting service level and maximum hub dwell time targets. We leverage dynamic consolidation that is enabled by containerized operations across logistics networks, according to the Physical Internet in modular standardized packaging, handling, and transport containers. The proposed model uses service-driven on-demand inter-hub shuttling, beyond relying on detailed vehicle routes and schedules often optimized way ahead of their implementation. Service-driven shuttling utilizes simple single-leg routes with adequate frequency to satisfy service requirements while the on-demand transportation adds transportation capacity flexibility which is especially beneficial when demand is highly uncertain. Combining inter-hub shuttling and on-demand transportation into service-driven on-demand inter-hub shuttling minimizes planning determinism and complexity while maximizing agility of operations.

HyPTLI relies on open protocols and digital platforms to ease the interconnection of network design components and operations via seamless information flow among players in the system, leveraging sensor networks and cyberphysical systems. Through this paper, we provide the foundation of a logistics framework that enables the interactions between the tiers of the logistics network, such as the synchronization and coordination of the fleet and hub operations.

Subsequently, this framework can be a motivation for further research. The proposed model involves optimization and modeling challenges that are related not only to the operations of vehicles and facilities in each layer, but also to the interactions between the tiers of the logistics network, such as the synchronization and coordination of the fleet and hub operations. Moreover, the heavy reliance of the model on open protocols suggests that robust protocols will need to be developed and tested rigorously to ensure that the service levels are met in a cost-effective manner.

References:

[1] Akyildiz, Ian F., and Xudong Wang. Wireless mesh networks. Vol. 3. John Wiley & Sons, 2009

[2] Balakrishnan, A., Magnanti, T. L., & Wong, R. T. (1989). A dual-ascent procedure for large-scale uncapacitated network design. Operations Research, 37(5), 716-740.

[3] Bektas, T., Crainic, T. G., & Van Woensel, T. (2015). From managing urban freight to smart city logistics networks.

[4] Crainic, T. G. (2000). Service network design in freight transportation. European journal of operational research, 122(2), 272-288.

[5] Crainic, T. G., Ricciardi, N., & Storchi, G. (2004). Advanced freight transportation systems for congested urban areas. Transportation Research Part C: Emerging Technologies, 12(2), 119-137

[6] Crainic, T. G., Ricciardi, N., & Storchi, G. (2009). Models for evaluating and planning city logistics systems. Transportation science, 43(4), 432-454.

[7] Crainic, T. G., Errico, F., Rei, W., & Ricciardi, N. (2012). Integrating c2e and c2c traffic into city logistics planning. Procedia-social and behavioral sciences, 39, 47-60.

[8] Crainic, T. G., & Montreuil, B. (2016). Physical Internet Enabled Hyperconnected City Logistics. Transportation Research Procedia, 12, 383-398. doi:10.1016/j.trpro.2016.02.074

[9] Crainic, T. G., Errico, F., Rei, W., & Ricciardi, N. (2016). Modeling demand uncertainty in two-tier city logistics tactical planning. Transportation Science, 50(2), 559-578.

[10] Crainic, T. G., Nguyen, P. K., & Toulouse, M. (2016). Synchronized multi-trip multi-traffic pickup & delivery in city logistics. Transportation Research Procedia, 12, 26-39.

[11] Cuda, R., Guastaroba, G., & Speranza, M. G. (2015). A survey on two-echelon routing problems. Computers & Operations Research, 55, 185-199.

[12] Gakis, K., & Pardalos, P. M. (2017). Network design and optimization for smart cities. New Jersey, NJ: World Scientific.

[13] Mittal, N., Udayakumar, P. D., Raghuram, G., & Bajaj, N. (2018). The endemic issue of truck driver shortage - A comparative study between India and the United States. Research in Transportation Economics, 71, 76-84. doi:10.1016/j.retrec.2018.06.005

[14] Magnanti, T. L., & Wong, R. T. (1984). Network design and transportation planning: Models and algorithms. Transportation science, 18(1), 1-55.

[15] Montreuil, B. (2011). Toward a Physical Internet: meeting the global logistics sustainability grand challenge. Logist. Res. 3, 71–87.

[16] Montreuil, B. (2015), The Physical Internet: A Conceptual Journey, Keynote Presentation, International Physical Internet Conference, Paris, France.

[17] Montreuil, B., Ballot, E., & Tremblay, W. (2015). Modular design of physical internet transport, handling and packaging containers.

[18] Montreuil, B., Buckley, S., Faugere, L., Khir, R., & Derhami, S. (2018). Urban package logistics hub and network design: The impact of modularity and hyperconnectivity.

[19] Montreuil, B., Meller, R. D., & Ballot, E. (2010). Towards a Physical Internet: the impact on logistics facilities and material handling systems design and innovation.

[20] Montreuil, B., Meller, R. D., & Ballot, E. (2012). Physical internet foundations. IFAC Proceedings Volumes, 45(6), 26-30.

[21] Ng, M., & Lo, H. K. (2016). Robust models for transportation service network design. Transportation Research Part B: Methodological, 94, 378-386.

[22] Pérez, J. M. Q., Tancrez, J. S., & Lange, J. C. (2020). Express shipment service network design with complex routes. Computers & Operations Research, 114, 104810.

[23] Sallez, Y., Montreuil, B., & Ballot, E. (2015). On the activeness of physical internet containers. In Service orientation in holonic and multi-agent manufacturing (pp. 259-269). Springer, Cham.

[24] Scherr, Y. O., Neumann-Saavedra, B. A., Hewitt, M., & Mattfeld, D. C. (2018). Service network design for same day delivery with mixed autonomous fleets. Transportation research procedia, 30, 23-32.

[25] Sobral, J., Rodrigues, J., Rabelo, R., Saleem, K., & Furtado, V. (2019). LOADng-IoT: An Enhanced Routing Protocol for Internet of Things Applications over Low Power Networks. Sensors (Basel, Switzerland), 19(1), 150. https://doi.org/10.3390/s19010150

[26] Summary of Hours of Service Regulations. (n.d.). Retrieved from https://www.fmcsa.dot.gov/regulations/hours-service/summary-hours-service-regulations

[27] Zhao, S. and D. Raychaudhuri, "Multi-Tier Ad Hoc Mesh Networks with Radio Forwarding Nodes," IEEE GLOBECOM 2007 - IEEE Global Telecommunications Conference, Washington, DC, 2007, pp. 1360-1364, doi: 10.1109/GLOCOM.2007.261.

[28] Navda, V., A. Kashyap and S. R. Das, "Design and evaluation of iMesh: an infrastructure-mode wireless mesh network," Sixth IEEE International Symposium on a World of Wireless Mobile and Multimedia Networks, Taormina-Giardini Naxos, 2005, pp. 164-170.

[29] Wang, Z., Qi, M., Cheng, C., & Zhang, C. (2019). A hybrid algorithm for large-scale service network design considering a heterogeneous fleet. European Journal of Operational Research, 276(2), 483-494.

[30] Wang, Z., & Qi, M. (2019). Service network design considering multiple types of services. Transportation Research Part E: Logistics and Transportation Review, 126, 1-14.

[31] Winkenbach, M., Kleindorfer, P. R., & Spinler, S. (2016). Enabling urban logistics services at La Poste through multi-echelon location-routing. Transportation Science, 50(2), 520-540.

Digital Twin Design Requirements for Durable Goods Distribution in Physical Internet

Miguel Campos[1,2], Shahab Derhami[1,3], Leon McGinnis[1,2],
Benoit Montreuil[1,] Ali V Barenji[1,2]

1. Georgia Tech Physical Internet Center, GA, USA
2. School of Industrial & Systems Engineering, Georgia Institute of Technology, USA
3. School of Management, Binghamton University, NY, USA

Corresponding author: mcampos@gatech.edu

Abstract: *Today the practice for distributing large products manufactured at few original equipment manufacturers (OEMs) consists of a dedicated Point-to-Point (PtP) logistics system, typically requiring long haul transport from the factory to the wholesale destination. A growing problem is the shortage of commercial drivers willing to be away from home for several days to move products cross-country. Hub relay network logistics systems are an alternative solution to P2P logistics systems that allow reducing drivers' away-from-home times. Operating a relay-based logistics system requires accounting for multiple interrelated operational decisions that become more complicated as the system becomes larger and encompasses more players. To deal with such complexity we propose utilizing a digital twin of the distribution and logistics system as a decision-making support tool to manage the system and make operational decisions efficiently. This paper explores the design and assessment of a hyperconnected relay network of transport hubs supporting the movement of durable goods from factory to wholesale destinations. It describes requirements and challenges in developing and implementing a digital twin for such systems.*

Conference Topic(s): *From Logistics Networks to Physical Internet Network, Developing the System of Logistics Networks towards the Physical Internet, Interconnected freight transport, logistics and supply networks, Systems and technologies for interconnected Logistics, PI Fundamentals and Constituents, PI Modelling and Simulation, Distributed Intelligence in Physical Internet*

Keywords: *Digital Twins, Physical Internet, Logistics System Design, Hub Relay Network*

1 Introduction

While parcel logistics has gained wide visibility and importance during the 2020-2021 pandemic, there is another logistics system of great importance that remains largely invisible to the public. Durable goods–appliances, computers, televisions, automobiles, and similar large products–are produced at a few original equipment manufacturers (OEMs) and distributed to many retail outlets, sometimes directly, sometimes through a local or regional warehouse. Today the practice is for each OEM to have a dedicated Point-to-Point (P2P) logistics system, typically requiring long haul transport from the factory to the wholesale destination or a set of nearby such destinations. In addition to long-haul transports, carriers often must break the backhaul routes to multiple legs to reduce deadheads, imposing an extended away-from-home period to drivers. A growing problem is the shortage of commercial drivers willing to be away from home for days or weeks to move products cross-country. Multiple studies, such as Hu et al. (2019), show the long driving distance causes several mental and physical issues for drivers, and is the primary culprit for high driver turnover in the trucking industry.

Hub relay network logistics systems are an alternative solution to P2P logistics systems that allow reducing drivers' away-from-home times. A hub relay network consists of multiple relay hubs where the truckloads are relayed. This includes the change of drivers, tractors, trailers, and/or loads. Smart placement of relay hubs can significantly reduce the drivers' away-from-home times while maintaining satisfactory service levels and minimizing costs and environmental impact. Hakimi et al. (2015) showed the conceptual feasibility of relay-based transportation in which drivers return home every day or after two days. Designing and implementing such a logistics system requires making several complex and interrelated design and operational decisions (Campos et al., 2021). From the design standpoint, a primary decision is designing the network of hubs. This includes finding the optimal number of hubs, their locations and capacities, while considering potential travel distances and flow between hubs, and available fleets of domiciled drivers and trucks in each region (Vergara and Root, 2013; Kewcharoenwong and Üster, 2017; Hu et al., 2019).

For illustration purposes, let us analyze a logistics system with some OEMs that need to deliver products to various retailers. If we think about a national or regional network, the retailers may be hundreds or even thousands of miles away from the OEM's, requiring long haul to distribute the goods. The point-to-point (P2P) logistics system for this example is depicted in simplified form in the left side of Figure 1.

Let us now assume that in order to address the shortage of commercial drivers we aim to implement a Physical Internet enabled hyperconnected relay logistics network which is illustrated in the right side of Figure 1. With this network we intend to allow the drivers to return rapidly to their domicile, avoid the high turnover of drivers in the trucking industry, and reduce the adverse effects of long haul on truckers' health. This network has additional hubs between the OEMs and the retailers, which we will call transit hubs, for enabling consolidation, transshipment, and crossdocking of goods. This network can be used by a single company or can be open to many companies which produce similar products, as the example shown in which the last OEM corresponds to a facility of a second company. Furthermore, the OEMs and retailers can be used as transit hubs as well depending on the context.

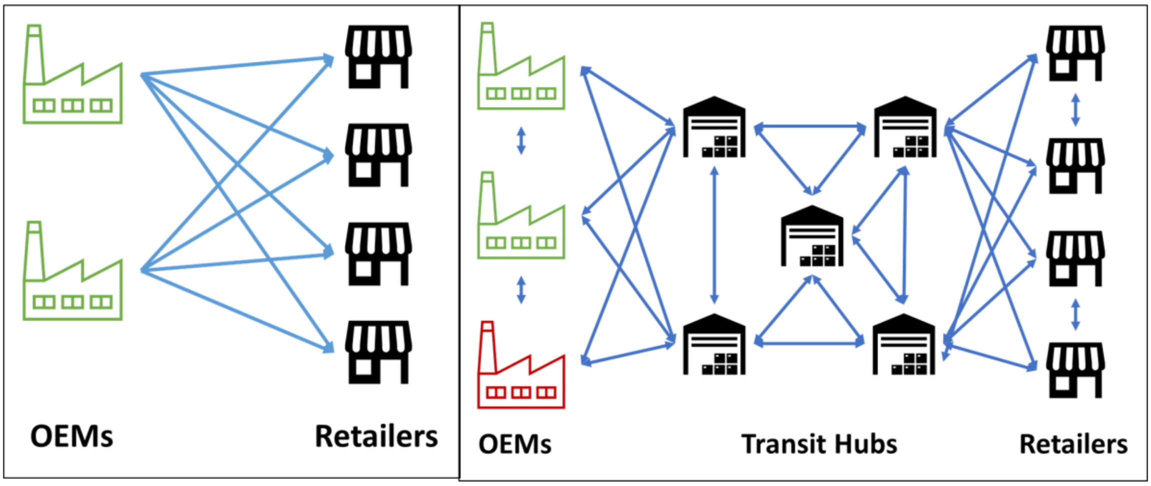


Figure 1. Long haul network vs. PI enabled hyperconnected relay network

While the hub relay network design focuses mainly on static optimization models, the operations of this system require accounting for multiple interrelated operational decisions in a system whose state changes dynamically by endogenous and exogenous factors. The decisions on how to route products in the network and when to dispatch vehicles need to be made on a daily or even hourly. This decisions are based on the status of the system and the availability of the resources, which dynamically change by exogenous factors such as traffic, weather, and breakdowns. Hence, the operating system should be equipped with proactive decision capabilities and contingency plans to react appropriately to any supply disruption and prevent propagating it to the entire system.

Moreover, relaying drivers, tractors, or trailers requires a high level of synchronization between planning and operational decisions for such resources. The system operation becomes more complicated when the logistics system is open to multiple carriers and OEMs, which offers more opportunities for flow consolidation and deadhead trip reduction but requires accounting for potential conflicts of interest between different players. To deal with such complexity, we propose utilizing a digital twin of the distribution and logistics system as a decision-making support tool to manage the system and make operational decisions efficiently.

Adapted from the works of Glaessgen and Stargel (2012), and Marmolejo et al. (2020a), a digital twin can be defined as a computational representation of a physical system that has real-time interaction with the latest state of the physical system and has analytics and simulation capabilities aimed to provide visibility, feedback and insights to be used in the decision-making process of the system, forming an improvement cycle. In logistics systems applications, the digital twin enables monitoring and scenario assessment and planning. As a real-time decision support tool, the digital twin should run faster than the real system; thus, the efficiency of the embedded algorithms is crucial.

Designing such a digital system is a complex challenge which requires considering various aspects and parameters, as shown in section 5. It also requires efficient use of available technologies such as Internet-of-Things (IoT) and sensors for efficient physical-digital communication. There have been several studies on the implementation of digital twins for logistics systems, particularly in manufacturing, warehousing, and inventory management, such as Agalianos et al., 2020; Kritzinger et al., 2018; Marmolejo, 2020b.

Nevertheless, there has not been an adequate study on the proper implementation of a digital twin of logistics systems that is connected to real-time data and used for real-time decision making. This is beyond using an offline simulation model for scenario testing. Furthermore, there are companies offering commercial digital twin software and implementations, such as Microsoft Azure, O9 Solutions, GE, Moicon, Siemens and Honeywell. These efforts together will boost the spread of digital twins in systems' assessment and improvement.

Building a digital twin for a system requires an objective definition of the system’s elements and interactions. For this purpose, the domain, and conceptual models, which are often used interchangeably in system modeling, should be defined, and implemented as tool-agnostic models, meaning they should enable the complete modeling and assessment of the system without targeting a specific methodology or tool (Thiers and McGinnis, 2011). In this paper, we will make a distinction between the two. The models should be clear enough to enable building analytical, optimization, and simulation models alike with the use of any software.

This paper explores the requirements for a digital twin supporting operational control of a hyperconnected physical internet-enabled relay network of transport hubs supporting the movement of durable goods from factory to wholesale destinations. Stakeholders in this new approach include the OEMs whose product must be distributed, the retailers who want product available to sell, the carriers who require profitability through sufficient utilization of their transport resources, and the drivers who want to be fairly compensated and desire fewer days away from home. We propose developing the digital twin of such a system for efficient management of its components. We will describe the requirements and challenges in developing and implementing digital twins of logistics systems.

2 Design of physical internet-enabled hyperconnected relay logistics systems

The Physical Internet (PI) concept was first described by Montreuil (2011) as an innovative logistics system meant to tackle the global logistics grand challenge toward improvement in efficiency and sustainability. It has been defined as "an open global logistics system founded on physical, digital and operational interconnectivity through encapsulation, interfaces and protocols" (Ballot et al., 2013).

Designing a digital twin for a Physical Internet-inspired hub relay network requires inclusion of various operational decision-making tools. Campos et al. (2021) propose a toolkit for configuring and assessing physical internet enabled logistics systems. Some of the most important tools for enabling the digital twin are: Demand scenario generation, logistics zone clustering, hub network design, service network design, logistics hubs configuration, flow routing, containerization, consolidation, system description, and simulation. These tools are interrelated, forming a feedback optimization loop towards generating better overall performance.

Regarding the simulation, Kaboudvand et al. (2021) presents a simulator of a large scale hyperconnected urban parcel delivery logistics system, which could potentially turn into a digital twin due to the level of detail modeled. Similarly, Kim et al. (2021) presents the decisions and system architecture for hyperconnected urban logistics in the context of large items. Both these models are offline discrete-event simulation models of physical internet enabled hyperconnected logistics systems. These models are run beforehand for assessing different designs and operational decision making towards improving systems' performance.

Nevertheless, digital twins of physical internet enabled hyperconnected logistics systems are often difficult to implement due to the lack of capabilities of organizations to use real time decision making. Furthermore, offline policies and algorithms might not perform well under all circumstances and should be implemented based on the current state of the system, reducing with this the medium and long term of uncertainty in the system modeled.

There are three options to operate the system. The first options is that the drivers are swapped at the transit hubs, meaning they just exchange trucks going in opposite directions at each hub. The second option is that each driver keeps the tractor for the backhaul, but the semi-trailers are swapped at the hubs. The last option is that each driver keeps his/her own tractor and semi-trailer, and the products are transferred between trailers at the hub. Given the size of the products, the last option might be inefficient as the cargo can take a long time to be loaded and unloaded. Nevertheless, whenever vehicles and facilities are PI enabled, a fast unload/load process between trucks becomes feasible.

The advantage of the first option is that the swapping time is the fastest, requiring minimum time to swap trucks and reducing the handling cost. The downside is that if the drivers own the trucks, the truck swap will be hard to implement. The second option has the advantage of allowing each driver keeping its own tractor, but imposes additional handling for enabling the trailer swap, furthermore, the compatibility truck-trailer can also be a limiting factor. The last option allow the drivers to stay with both their tractor and trailer. Nevertheless, it imposes additional loading and unloading time and cost, while also risking the integrity of the products in the handling process.

If there are hard constraints on the ownership of tractors and trailers, the third option might seem the best fit, but it will require handling equipment in each hub, and additional time and resources, making the implementation more expensive than the other two options. Depending on the product, this option might be a good fit if there is a way to do the product reshuffling at the hubs fast, safe, and reliable, which could be the case for PI enabled vehicles and facilities. If the drivers own the trucks but not the trailers, the second option might seem the best. Similarly, if there is feasibility to implement the truck swapping, option one might be the best fit.

In any case, to operate options one and two there are hard problems to be solved, notably the planning and scheduling of vehicles and drivers. This scheduling requires synchronization to avoid nonvalue added times at the hubs, meaning drivers and trucks going in opposite directions should arrive to a hub during a short time window. For option three, facilities would need bigger space for storage and handling, meaning having both warehousing and crossdocking roles, needing additional resources which also need to be scheduled.

As mentioned, the operational decisions regarding this system are very complex. The main decisions to be made to design and operate such system are the location, size and capabilities of the hubs, the truck routing, the product routing, and the driver scheduling. This paper will not discuss methodologies for making these decisions. Instead, we will analyze how to assess this type of system through discrete event simulation, and how to use the simulation as a digital twin for implementing the solution in a real context.

Discrete event and agent-based simulation models can be used as an initial step towards achieving a digital twin of a system. If modelled with enough detail, the offline simulation model can be used for assessing the systems' design and operations, as well as become a digital twin when connected to live data. The use of a simulation-based digital twins enables combining the assessment of systems' design with an implementable operational real-time decision-making tool. In any case, the first step towards analyzing a system is to properly describe it, which will be addressed in the following section.

3 Domain and conceptual models

The models to be used for accurately describing the system are the domain and conceptual models. In this paper we differentiate the two, each having a particular structure and objective. These models are meant to allow stakeholders to intervene in the improvement process. As this field progresses, a standardized formal language for describing logistics systems should be implemented. Such a language should allow representing specific domains, providing a rich set of fundamental abstractions, and allowing easy computations (Thiers and McGinnis, 2011). Although using a unified language is not necessary, it is encouraged for enabling tool agnostic standardized system modelling methodologies.

3.1.1 Domain model

A good domain model, in essence, creates a *language* for discussing instances in the domain. If the domain is plant-to-customer delivery of durable goods, the domain model defines the semantics and syntax for describing any instance of such a delivery system, involving any number of plants, customers, and carriers. To do so, it must incorporate definitions for those aspects of the system that are of interest to the three main stakeholders: the OEMs (shippers), the carriers, and the system itself. Both *structure* and *behavior* must be adequately defined.

Structure refers to the observable elements of the system, their properties, and inter-relationships. What is shipped is a key element of *structure*. From the shipper's perspective a *shipment* is a set of *unitsOfHandling* (each with specific properties), with properties that include: *shipmentID*, *customer*, *tenderedDate*, *pickupDate*, *serviceLevel*, and *deliveryDate*. From the carrier's perspective, this *shipment* may be associated with one or more *transportTask*s, whose properties include the *shipment* properties but add properties such as *driverID*, *tractorID*, *trailerID*, etc. Note that *driverID* identifies a particular *driver* with properties like *domicile*, *schedule*, etc. Each resource type will have an associated set of properties. A hub relay network will require definitions for *hub* and *route*, or the ability to travel directly between two hubs. The domain model must describe these and all the other relevant components of a hub relay network.

Behavior has two key elements. The observable actions of resources like drivers and rigs are examples of resource behavior. Often this behavior can be described in a domain model using activity networks. The largely not-observable decision-making processes also are behaviors that must be captured in the domain model. State machines have proven to be one useful approach to capturing decision-making behavior. The domain model must incorporate all the generic behaviors relevant to the decisions to be supported.

Figure 2 illustrates an example of a refrigerator manufacturer that distributes its products to retailers through a relay network. Such figures would be part of the domain model defining the overall movement of flow through facilities and the system elements.

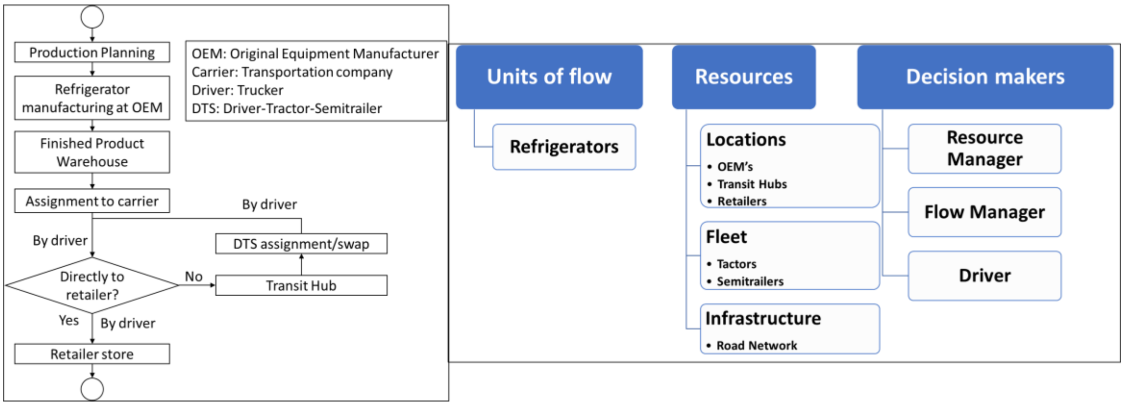


Figure 2. Domain model illustration of a refrigerator manufacturing relay network distribution

3.1.2 Conceptual model

Design and operational decision making for a hub relay network may be supported by a variety of analysis methods, from simple queuing analyses to simulation, and spreadsheets to math programming. For large scale and complicated systems like hub relay networks, it is critical that all these analyses reflect the same understanding of the system. One way to approach this challenge is through a conceptual, analysis-agnostic model, using the semantics and syntax defined in the domain model.

Each resource type defined in the domain model has a set of instances in the conceptual model. For example, the definition of *hub* from the domain model is used to identify the set of potential hubs in a proposed hub relay network, each with its own property values. The definition of *shipment* is used to create a database of shipments, both those completed, those in process and those not yet tendered. To be fully useful, the conceptual model must provide all the information needed to populate any specific analysis model that will be used to support decision-making.

The conceptual model will offer a more specific definition of the system elements and interactions, building on the domain model to describe the movement of flow in the system, the decision-making architecture, the interaction between stakeholders, resources and units of flow, and all other necessary concepts that will enable reproducing the systems' operation in a digital environment. Particularly, for the digital twins this model should also include the concepts regarding the interaction between the digital and physical systems. Figure 3 shows the flow manager decision making logic. Notice the connection between the digital and physical models is explicitly included. These types of figures would be part of the conceptual model.

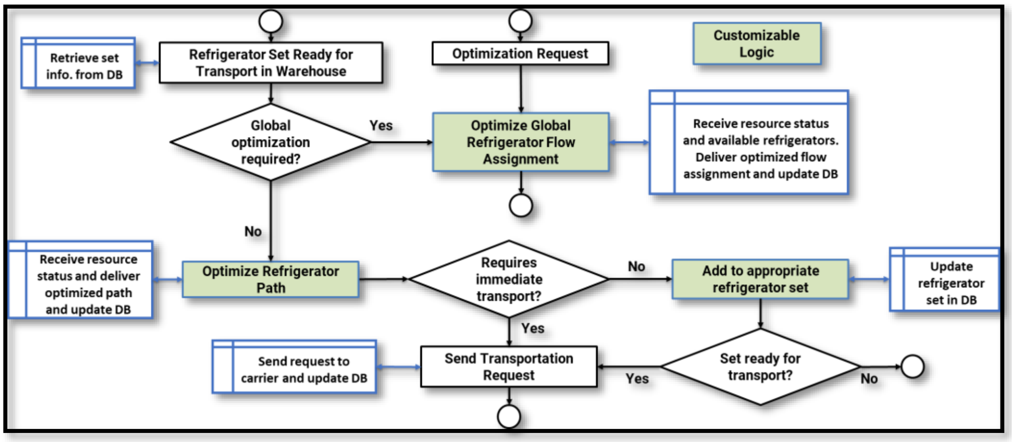


Figure 3. Flow manager decision making process.

Having these two types of models is of great help to communicate with stakeholders and guide the modeling that follows. However, in practice many times modelers skip this step, jumping directly to model the system. This practice is not recommended, as makes it hard to collaborate and allow other modelers to continue building on the existing model, especially when the original modeler is absent. Also, without domain and conceptual models it is hard for stakeholders to interact in the design process and make sure the digital model accurately represents the physical model.

4 Cyber-physical systems (CPS)

Cyber physical systems are multidimensional and multifaceted systems that integrate the virtual world and the physical world. Through the integration and collaboration of computing, communication, and networking, CPS deliver real-time sensing, feedback, control, and other services (Vatankah Barenji et al., 2020a). With intensive connection and feedback loops, physical and computing processes are highly resilient. In this way, cyber–physical integration and real-time interaction are achieved to monitor and control physical world in a reliable, safe, collaborative, robust, and efficient way. Digital twin is a paradigm for realizing the interaction and integration between the physical world and the virtual world, which has attracted full attention from the relevant academic circles and enterprises (Vatankah Barenji et al., 2020b).

Furthermore, digital twins are related to CPSs, creating a high-fidelity virtual model of the physical world, simulating objects behaviors in the real world, and providing feedback (Vatankah Barenji et al., 2020b). A digital twin is a cyber-physical system, but not all cyber-physical systems are digital twins. The concept of CPS considers the digital-physical interaction for implementing any process, which can be an isolated part of a system. The difference with digital twins lies in the fact that digital twins are digital representations of complete and well-defined systems. As the literature on CPS is also extensive, these two concepts should be analyzed together towards having a more comprehensive understanding of the implementation of digital twins.

5 Digital twins of logistics systems

One of the first challenges towards implementing a digital twin is defining the level of abstraction at which the digital system is going to represent the physical system, which is not always obvious (Singh et al., 2019). A proper system description (conceptual model) and collaboration of modelers with stakeholders will help tackle this challenge. The objective is to build the simplest version of a conceptual model which accurately represents all relevant aspects of the physical system.

After this, there are various operational decisions to be made before simulating this type of system. Some of the main models required are, the demand generation, the logistics zones clustering, hub network design, service network design, configuring logistics hubs, product routing, consolidation, and containerization. These decisions are complex in nature, and there are different exact and heuristics models to solve them. For more information about such algorithms and their interactions, see Campos et al., 2021.

For the physical internet-enabled hyperconnected relay logistics systems, there are additional challenges as mentioned in section 2. Notably, the synchronization required for both drivers and trucks imposes additional complexity to the service network design, consolidation, and product routing. This problem can be tackled using global optimization models which use heuristic algorithms to find good solutions daily. Other proposal is to find shipping protocols such as implementing chain type shipments per origin and destination with constant takt time. Other idea is trying to implement a live control tower where all vehicles are tracked, for increasing or reducing the vehicle velocity for improving the synchronization. On the simulation side, the use of global optimization models might make the run time big as the instance gets bigger, thus, dynamic protocol or algorithmic approaches are recommended if the instance gets big.

After defining the scope of the digital twin, the next big challenge is to implement the use of IoT, sensors and automation. The information from these sources needs to be compiled in a database, which needs to be updated live and must be accessible by the digital twin. The implementation of such technology can be costly and complex. There is a broad implementation of ERP systems for managing integrated processes inside companies. Some examples of areas that can be managed by such systems are the inventory, manufacturing, supply chain, dispatches, finance, etc. Nevertheless, problems have been reported with the use of ERP systems regarding real time data and identification of disruptions (Marmolejo, 2020b).

Nowadays, many organizations still have data silos managed in spreadsheets with no real time information or sharing of any kind. The information between areas is not connected, and this makes impossible the implementation of a digital twin. Furthermore, even when companies use ERP systems, such systems may not be properly integrated with partner companies such as suppliers and clients (Marmolejo et al., 2020a), due to lack of information technologies, data security and trust. This type of implementation results in myopic decision making in the supply chain management.

Novel technologies such as blockchain and smart contracts might help sorting some of the challenges with data security and trust. Nevertheless, they should be carefully implemented to make sure they do not impose excessive time for transferring the information between the physical and digital model. The next step would be to define the set of key performance indicator (KPI's) to be used in the decision-making process. Is key that stakeholders agree on the set of KPI's, the definition of how decisions will be made based on this information and the level of automation of the decision-making process. When implementing a digital twin, all these challenges need to be addressed before getting into the modeling, so that such model can be built knowing the available information, the format, and the accessibility features.

Regarding the large-scale simulation, the mix between discrete-event and agent-based simulation is the most appropriate combination of paradigms to assess the PI enabled hyperconnected relay network. On the discrete-event side, as it allows representing simple behavior efficiently and to make animations, and on the agent-based case as it allows modelling more complex proactive agent behavior and the scalability of the model through agent instance generation (Majid et al., 2016). For this type of system, being able to place agents into a GIS map is desirable for enabling realistic distance, movement, and animations. Therefore, among the existing commercial simulation software, Anylogic would probably be the best choice as it allows using together discrete-event and agent-based paradigms, and the use of GIS maps for agent's movement.

Other challenge of the simulation model comes with the size of the instance. For big instances, running such a simulation model with enough detail might take more than 80 gigabytes of RAM memory. Thus, a server with good computational power is required for running these models. For the instance generation, is recommended to create the network from input files, for liberating space in the model. In the case of digital twins, such files should be allocated in databases connected live with the model. Similarly, if the number of agents is too big, the memory required for this might be too high. Therefore, a useful modeling technique is to model flow objects (products) as data packages instead of complete agents, for reducing the model size and improving run speed. Another good practice is to turn off the automatic logs of the software and generating the output data in text files for analysis in exogenous data tools. Note in the case of a digital twin, the output generated should go directly into a database connected to the physical system, thus, storing, and computing KPI's inside of the model might be necessary.

6 Conclusion

Physical internet enabled hyperconnected relay networks will help reduce the driver's shortage and high turnover. Testing and properly managing such complex systems is difficult, so digital twins will be of great help in designing, assessing, and implementing improved operations. There are big challenges towards the implementation of digital twins, but companies will have to shift towards data driven operations to remain competitive. Academia should partner with industry regarding digital twin implementations to provide meaningful research avenues, results, and the development of the field. Simulation models with actual connections to databases for enabling digital twins are still incipient, so the implementation of such models is yet to be explored in the literature. There is need for more application cases of digital twins for proving the power of the tool; hopefully, this research will help motivate industry and academia to implement such tools in the years to come.

References

- Agalianos, K., Ponis, S., Aretoulaki, E., Plakas, G. & Efthymiou, O. (2020). Discrete Event Simulation and Digital Twins: Review and Challenges for Logistics, *Procedia Manufacturing*, v51, 1636-1641.

- Ballot, E., Montreuil, B. & Thivierge, C. (2013). Functional Design of Physical Internet Facilities: A Road-Rail Hub, *CIRRELT*.
- Campos, M., Montreuil, B., McGinnis, L., Kaboudvand, S., Kwon, S., Bahrami, Z., Faugere, L. & Buckley, S. (2021). Enabling Scientific Assessment of Large Scale Hyperconnected Urban Parcel Logistics: System Configuration and Assessment, *IISE 2021 Annual Conference,* Virtual, 2021/05/22-25.
- Glaessgen, E. & Stargel, D. (2012). The digital twin paradigm for future NASA and U.S. air force vehicles, *53rd Structures, Structural Dynamics, and Materials Conference: Special Session on the Digital Twin.*
- Hakimi, D., Montreuil, B. & Hajji, A. (2015). Simulating Physical Internet Enabled Hyperconnected Semi-Trailer Transportation Systems, *Proceedings of 2nd International Physical Internet Conference*, Paris, France, 2015/07/06-08, 11p.
- Hu, Z., Askin, R. G., & Hu, G. (2019). Hub relay network design for daily driver routes. *International Journal of Production Research*, *57*(19), 6130-6145.
- Kaboudvand, S., Campos, M. & Montreuil, B. (2021). Enabling Scientific Assessment of Large Scale Hyperconnected Urban Parcel Logistics: Agent-based Simulator Design, *IISE 2021 Annual Conference*, Virtual, 2021/05/22-25.
- Kim, N., Montreuil, B., Klibi, W. & Kholgade, N. (2021). Hyperconnected urban fulfillment and delivery, *Transportation Research Part E: Logistics and Transportation Review, 145.*
- Kewcharoenwong, P., & Üster, H. (2017). Relay Network Design with Capacity and Link-Imbalance Considerations: A Lagrangean Decomposition Algorithm and Analysis. *Transportation Science*, v51, no4, 1177-1195.
- Kritzinger, W., Karner, M., Traar, G., Henjes, J. & Sihn, W. (2018). Digital Twin in manufacturing: A categorical literature review and classification, *IFAC-PapersOnLine*, v51, no1, 1016-1022.
- Majid, M., Fakhreldin, M. & Zuhairi, K. (2016). Comparing Discrete Event and Agent Based Simulation in Modelling Human Behaviour at Airport Check-in Counter. *Human-Computer Interaction. Theory, Design, Development and Practice,* 510-522
- Marmolejo, J. & Hurtado, M. & Suarez, R. (2020a). Digital Twins in Supply Chain Management: A Brief Literature Review, *Intelligent Computing and Optimization,* 653-660.
- Marmolejo, J. (2020b). Design and Development of Digital Twins: A Case Study in Supply Chains, *Mobile Networks and Applications*, v25, 2141–2160
- Montreuil, B. (2011). Toward a Physical Internet: meeting the global logistics sustainability grand challenge, *Logistics Research*, v3, 71-87
- Singh, J., Settanni, E., Tsolakis, N. & Kaur, P. (2019). Supply Chain Digital Twins: Opportunities and Challenges Beyond the Hype, *23rd Cambridge International Manufacturing Symposium*, University of Cambridge, 26 – 27 September 2019
- Thiers, G. & McGinnis, L. (2011). Logistics systems modeling and simulation, *Proceedings of the 2011 Winter Simulation Conference (WSC)*, Phoenix, AZ, 1531-1541.
- Vatankhah Barenji, A., Li, Z., Wang, W., Huang, G. & Guerra-Zubiaga, D. (2020a). Blockchain-based ubiquitous manufacturing: a secure and reliable cyber-physical system. *International Journal of Production Research,* v58, no7, 2200-2221.
- Vatankhah Barenji, A., Liu, X., Guo, H., & Li, Z. (2020b). A digital twin-driven approach towards smart manufacturing: reduced energy consumption for a robotic cellular. *International Journal of Computer Integrated Manufacturing*, 1-16.
- Vergara, H. A., & Root, S. (2013). Mixed fleet dispatching in truckload relay network design optimization. *Transportation Research Part E: Logistics and Transportation Review*, v54, 32-49.
- Villarreal, C. & Kathiresh, K. (2020). Literature review on Cyber Physical Systems Design, *Procedia Manufacturing*, v45, 295-300.

Design of a Simulation-Based Experiment for Assessing the Relevance of the Physical Internet Concept for Humanitarian Supply Chains

Manon Grest[1,2], Mahmut Metin Inan[2,4], Yaarit M. Cohen[2,3,4], Ali V. Barenji[2,4], Mathieu Dahan[2,3,4], Matthieu Lauras[1,2,4], Benoit Montreuil[2,4,5]

1. Center of Industrial Engineering, University of Toulouse, IMT Mines Albi, France
2. Physical Internet Center, Supply Chain & Logistics Institute
3. Center of Health and Humanitarian Systems
4. School of Industrial and Systems Engineering, Georgia Institute of Technology, Atlanta, USA
5. Coca-Cola Chair in Material Handling and Distribution

Corresponding author: manon.grest@mines-albi.fr

Abstract: *The challenges faced in delivering relief items to victims of natural disasters and the growing external pressures urge humanitarian supply chain organizations to initiate some change. In this regard, the physical internet concept can offer a paradigm shift in relief organization and resource mobilization. To convince humanitarian actors to embrace this path, we propose a rigorous methodology leveraging a prototypical agent-oriented discrete-events simulator built within the AnyLogic platform, to conduct scientific experiments enabling to investigate the suitability and relevance of PI concepts for HSCs by systematically quantifying their benefits and drawbacks on HSC performance, sustainability, and resilience. We provide preliminary experimental results contrasting the baseline shaped by the current HSC structures, behaviors and practices, notably relative to sourcing, transporting, and warehousing, with those of hyperconnected HSCs in line with the Physical Internet at distinct degrees of maturity. In the experiment, we study past disaster scenarios that occurred in Indonesia and response efforts under different behaviors simulated with this platform. Initial results show that PI concepts are smoothly fitted to HSCs and the performance of hyperconnected HSCs is better than the current baseline.*

Conference Topic(s): *PI Modelling and Simulation*

Keywords: *Physical Internet, Humanitarian Supply Chains, Disaster Relief Operations, Hyperconnected Supply Chains*

1 Introduction

Natural disasters affect millions of people and cause damage to communities all around the world. Humanitarian response efforts take place immediately after a disaster and gather many different actors that plan for pre- and post-disaster efforts. Governments, humanitarian organizations, private sector partners, as well as volunteers aim to be prepared for potential disasters and respond to them successfully to help people and communities overcoming the consequences of disasters. As the intensity and frequency of natural disasters are expected to increase with the impact of global warming and climate change, the importance of Humanitarian Supply Chains (HSCs) operations, which concentrates most of the efforts (Van Wassenhove, 2006), is ever increasing. Besides, since human lives can be at stake after a disaster, it is crucial for HSCs to be effective while trying to help as many people as possible within the shortest possible time. However, achieving these goals is challenging for HSCs that concentrates most of the relief efforts (Van Wassenhove, 2006). Indeed, the efforts toward better coordinating actors, minimizing costs and environmental impacts, managing HSC operations and resources with a holistic view, and accounting for long-term impacts, are hampered by multiple factors. Among these are urgency contexts forcing decision-makers and other actors to work under pressure; damaged post-disaster infrastructures (unusable roads, lack of electricity and clean water); lack of collaboration culture; competition for limited funding and donations; HSC actors often having their agendas, making it hard to rally behind a shared goal; donors and media putting pressure on actors to improve monitoring and provide evidence for the quality of their relief operations; and missing concrete modelling of coordinated operations and the underlying structure. (e.g. (Tomasini and Wassenhove, 2009); (İlhan, 2011); Ergun, 2013). As a result, current practices and operations of HSCs are being highly criticized for their performance (Haavisto and Goentzel, 2015), and emerged the need for reorganizing HSCs and improving their operations.

In the literature, there is a recognized need for the effective measurement of HSC performance (Beamon and Balcik, 2008). Indeed, the current approaches, indicators, and dimensions used to measure performance are deemed ineffective (Abidi et al., 2014). Notably, there are growing concerns regarding integrating environmental considerations in operating humanitarian supply chains. Furthermore, the highly volatile environment in the humanitarian context requires more attention in measuring and improving the resilience of HSCs so they can adapt to unexpected changes. In the last decades, several approaches have been introduced to improve the efficiency and effectiveness of HSCs. For example, as reported by (Jahre et al., 2016), there have been significant innovations in preparedness toward improving the efficiency and effectiveness during response efforts: prepositioning inventory, advanced coordination between stakeholders (public and private), as well as enhanced education and training. Relative to the environment, there has been a drive toward sustainable HSC operations and reverse logistics (Peretti et al., 2015). Green efforts include the collection of relief item waste that can be reused, repaired or recycled appropriately (Farahani and Rezapour, 2011). The innovations above are insufficient for the HSCs to meet performance expectations, as on one side they do not alter the flawed core of their current schemes, and on the other side, disasters are becoming more frequent yet highly uncertain in terms of their occurrence and their impact.

To overcome all these challenges and improve the HSCs' performance, Physical Internet concepts can be applied as their induced capabilities have already been proven to be attractive for concurrently improving the efficiency, resilience, and sustainability of commercial supply chains. However, PI requires changes of long-lasting habits and paradigms, adaptation to new cutting-edge approaches, and some investment. So, the potential value and suitability for HSCs' operations still need to be investigated, especially given the inherent differences between commercial and humanitarian supply chains. Abdoulkadre et al. have been the first to investigate the conceptual applicability of PI concepts on HSCs, and to propose concrete PI-induced transformative avenues for HSCs (Abdoulkadre et al., 2014). This has revealed the need for further research on rigorous design and modelling of hyperconnected HSC scenarios and implementation efforts, and on assessing the potential impacts of such efforts on HSC performance.

2 Simulation-Based Research Methodology and Simulator

Due to the complexity of humanitarian systems, the use of simulation has appeared to be a suitable solution to answer our research question (Sheard and Mostashari, 2011). However, to ensure the reliability of the results and to avoid a series of pitfalls when the time comes to modeling and coding the simulation (Law, 2014), a functional framework for creating and developing the simulation has been followed and adapted to our experiment context. Adapted from the simulation study process of Law, the methodology consists of eight steps; 1) Research problem and solution design, 2) Grasp data and define a model, 3) Construct a computer program and verify, 4) Make pilot runs and observe results and validate, 5) Perform the experiment design, 6) Make production runs and verify, 7) Analyze output data and validate and 8) Document and present result. Such methodology is not necessarily a sequential process and may require going back over some steps as new elements are added and the vision becomes clearer (see Figure 1).

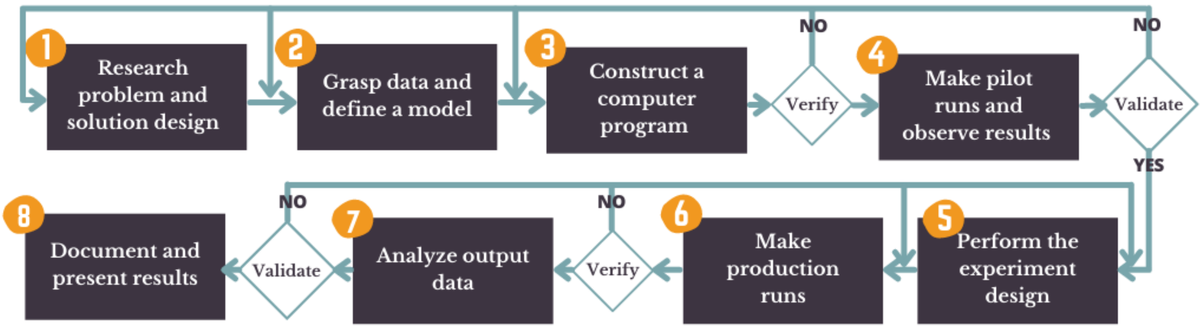


Figure 1 Adapted simulation study process (Law, 2014)

In parallel, a prototypical agent-oriented discrete events simulator was designed as a technological framework to support adequate simulation creation and development (Grest et al., 2021) (see Figure 2). The simulator architecture includes three key systems for properly experimenting: the scenario system, the virtual humanitarian ecosystem, and the performance system. The scenario system, as a configuration interface for the tester, provides to the ecosystem context elements and supply chain parameters inputs to integrate in order to form a particular scenario as a unique virtual humanitarian ecosystem to observe. Different settings at the scenario system level and runs of simulation will offer a set of scenarios to compare thanks to the performance system that keeps track of the evolution of performance indicators over time.

Figure 2 Prototypical agent-oriented discrete-events simulator embedding three different interrelated systems (Grest et al., 2021)

3 Case-Based Simulation Assessment of Hyperconnected HSCs

We illustrate end-to-end the functional framework by describing our approach in assessing the perspective of the PI for humanitarian operations through an Indonesian case study and a simple set of scenarios.

Step 1. Research problem and design. The hypothesis under study is the following: a reorganization of HSCs toward a hyperconnected version will positively impact the performance results of humanitarian players in assisting affected people. From there, an experiment needs to be undertaken to assess the impacts of such an original association understudied to date. To do so, a comparison through performance results will be conducted between a humanitarian supply chain baseline scenario representing current practices and some hyperconnected alternatives to test. Such a set of scenarios are instances at the factor definition level (see step 5) and are derived from a common model - detailed in the following step - named the theoretical model. For building the model, Indonesia, one of the most disaster-prone countries, will be used as a case study. Since HSCs are complex systems, using simulation for experimenting is necessary. Consequently, by recreating an environment prone to disasters and by programming the behaviors to be tested the evaluation will be performed through performance indicators. Furthermore, the study perimeter is limited to the improvement of HSCs' performance in delivering relief items to victims of natural disasters. Finally, from this design, different working areas have been identified and are detailed in the following sections.

Step 2. Grasp data and define a model. Gasping data for building the baseline scenario has started with a field survey within the Indonesian Red Cross (IRC) to enable the identification of the main objectives, strategies, and available resources within the country. Major findings were that IRC's logistics network is organized in a hierarchical structure and heavily depends on the population positioning. Indeed, the country is administratively divided into provinces which are composed of regencies. Therefore, IRC has positioned district warehouses in nearly every regency nearby major cities. Those are covered by a dedicated province warehouse which itself is supplied by a regional warehouse. The survey also enabled us to get a process-oriented vision regarding how IRC prepares and responds to natural disasters. Specifically, once a disaster occurs, small teams are sent to make a people and material damage assessment. Then, a global relief item need is calculated and sent to the warehouse in charge of the affected zone. In parallel, Points of Delivery (POD) for relief item distribution are deployed nearby population-dense areas and wait for supplies arrival. Next, the warehouse in charge responds to the demand with inventory on hand, finds an appropriate fleet, organizes the delivery to the PODs, and manages shortages. Finally, regarding the grasping data phase, historical records about past natural disasters have also been analyzed to generate statistics related to their impact and occurrence frequency in Indonesia. In addition, we have learnt about the need for the disaster type as well as the affected territory and population consideration into the model (Cosgrave and Herson, 2008).

Based on the knowledge gained, the theoretical model integrates four essential components in a relief context: i) the territory emulator that represents a territory and its key social and geographical characteristics having an influence on the relief organization, ii) the disaster generator that generates natural disasters events and impacts on people and materials, iii) the demand estimator that estimates the demand in relief items based on victims number and damages and iv) the humanitarian response simulator representing the humanitarian supply chain and logistics operations. In this paper, we base the model on an Indonesian case study and limit the study perimeter to the Aceh province and its 21 district parts of the mainland, as it is the site of major natural disasters in the past and displays disparity in terms of population distribution. Currently, only earthquakes, as a disaster type, are considered due to their frequent occurrence in this region. Regarding damages, only districts hit from a strong (with light damage) to extreme (very heavy damage) value according to the MMI shaking scale are regarded as requiring humanitarian assistance. From there, the focus is on affected people as Internally Displaced Persons (IDP) and are regarded as beneficiaries or a household in demand for five typical Red Cross items: blankets, jerrycans, hygiene kits, kitchen sets and tarpaulins. Rapidly, PODs are deployed in affected districts and wait for supplies coming from the pre-positioned warehouses of the IRC network. Need requests are managed and fulfilled using on-hand inventory. When inventory is missing, supply sources are identified within the IRC network only, and a replenishment order is sent meanwhile available items are delivered by trucks.

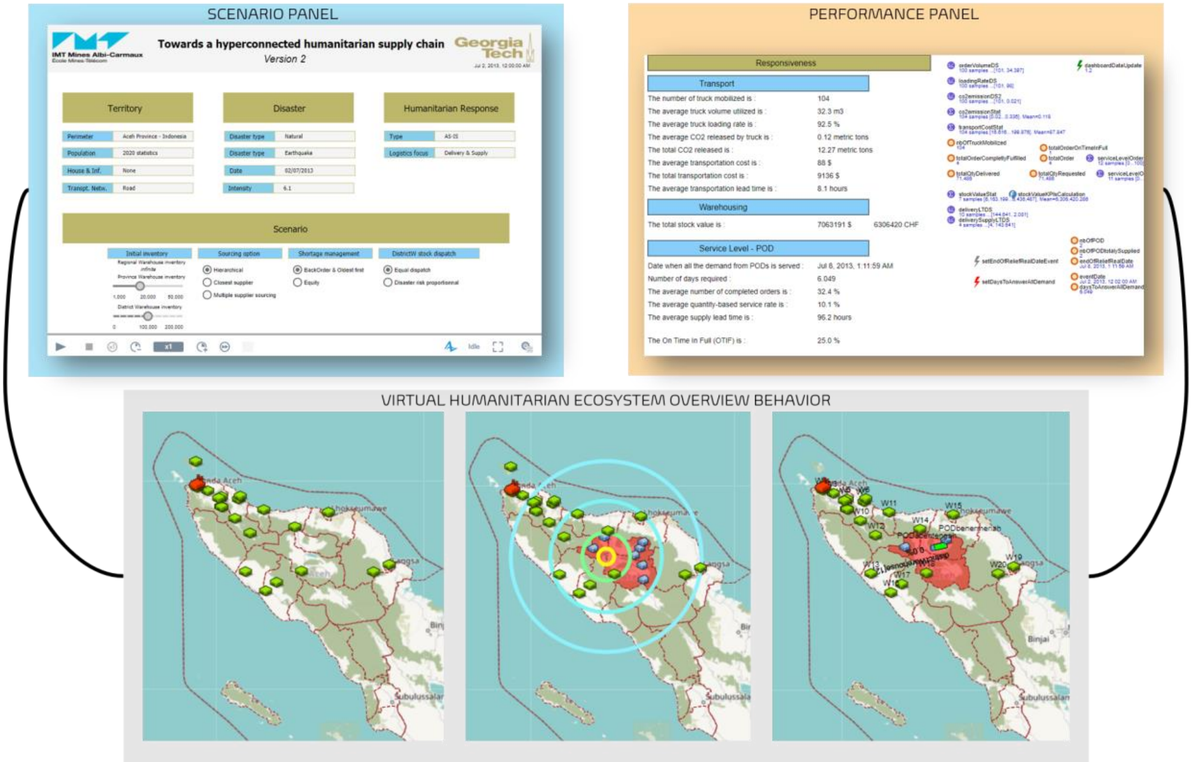


Figure 3 Simulation systems connection and operating overview

Step 3. Construct a computer program and verify. Derived from the previous theoretical model, an agent-based and discrete-event-oriented simulation program has been initiated in Java coding language using the AnyLogic software (see Figure 3). Based on the simulator, the simulation consists of three interrelated components. The first is the scenario system where a user can configure a scenario to test. The configuration consists in a first step of defining the environment and context in which a second system, the virtual humanitarian ecosystem, is about to evolve. The environment is related to a country or region where past disasters, defined by the context, are replayed and kept the same for running an experiment with a set of scenarios. Secondly, the user finalizes setting a scenario by indicating the operational approaches of the humanitarian response system at specific parameters level, called factors of the experiment. The third component is the performance system where performance indicators will be monitored. Regularly, the virtual humanitarian ecosystem feeds the performance system with data to enable real-time performance measurement and monitoring. Therefore, through the Indonesian case study and the run of multiple scenarios, a comparison between the PI-oriented suggestions and the baseline is possible and enables concluding the relevance of the PI approach for the humanitarian sector. Documented tests have been performed to ensure the adequate execution of each programmed component compared to the defined theoretical model.

Step 4. Make pilot runs, observe results, and validate. Our interviews and data analysis could not provide us with all the required parameters and inputs precisely. As a result, in defining the baseline scenario, we made some assumptions and estimated some parameters. To validate these assumptions and estimations, we run pilot tests to ensure that the simulated behavior is close to the historical. Initial results from our pilot studies suggest that our assumptions and estimations are consistent with humanitarian situation reports found. Additionally, workshop sessions with practitioners have also been organized to validate the baseline scenario and associated assumptions such as trucks being dedicated to one organization at a time and returning [to warehouses] empty after delivering items.

Step 5. Experiment design. In this section, the complete design of the experiment intended is presented first and at the end, a portion of it with factors currently implemented in the simulation and used to run a first experiment phase. Once the baseline scenario has been validated, the design of the experiment phase consists of designing scenario variants for making comparisons (Chung, 2004). Alternatives rely on the change of the value of some parameters by the tester. Such input variables are called factors and their values are named levels (Banks, 1998). Since the manipulation of factors has a direct effect on the outputs, some effort should also be spent on defining the performance measurement system for the experiment.

- **Complete experiment design and factor definition**

Experiment design starts with defining factors and associated levels. Since we aim at testing new PI-oriented practices we identified two entry points for change, as triggers of the performance, lying in the network design and in the disaster operations management levels (Grest et al., 2020).

Regarding the network design, we first suggest challenging the node interconnection from a unilateral structure to a multi-directional one. Such a connection allows better spreading of the information within the network while gaining visibility over the nodes' capabilities. As a first step, the connection will be tested within the same network to form a - *physical intranet* - aiming at taking advantage of existing capabilities. Secondly, the connection will be expanded to adjacent networks to form the - *physical internet* - which will offer a larger vision of opportunities for partnerships and/or coordination. Another network design factor to consider is related to asset management. In this case, we wish to test to what extend dynamic capabilities against supply disruptions would bring more agility to the response.

Now regarding the disaster management operations factors, while conceptualizing hyperconnected approaches, we primarily focus on the logistics and related decisions at the sourcing, transporting, and warehousing level. Considering the sourcing, current practices mainly consider a vertical/hierarchical approach and attempt to distribute the relief items and other resources from big central warehouses to smaller ones. Instead, we suggest utilizing multi-directional sourcing to improve response effectiveness with faster deliveries, efficiency with higher resource utilization, and resilience with a more robust sourcing network. Moreover, organizations are currently conducting their transportation operations independently and the consolidation efforts within the organizations are limited. Because of the high volume and frequency of the deliveries and possible limitations on transportation means availability during the disaster response, increasing transportation consolidation both within and between organizations will improve the response performance with higher utilization of transportation means and enable organizations to make faster deliveries. Also, by applying PI concepts and using encapsulation standards, it is possible to achieve more efficient and effective warehousing operations. For example, convenient handling of the relief items will allow organizations to dispatch items faster, utilization of warehouse spaces will increase, and standardized storage units will allow different organizations to share warehouse spaces easily. Further, since the capabilities of humanitarian organizations during the response phase highly depend on their preparedness level, we also considered possible steps to improve the preparedness of the organizations. The main challenge while preparing for natural disasters is the high uncertainty regarding the time, location, and impact of the disaster; at this point, we also wish to study how anticipating potential disasters and their impacts would help organizations in smartly pre-positioning relief items and other resources before disasters.

- **Complete experiment design and performance measurement system**

While measuring the performance of HSCs, our focus will be on four main dimensions: effectiveness, efficiency, resilience and sustainability. Effectiveness is the extent to which the beneficiaries' needs and requirements are met (Beamon and Balcik, 2008). Two main performance indicators related to effectiveness will be response speed (delivery lead times) and response coverage (number of people receiving help). Efficiency measures the rightness of needs evaluation, speed of adaptation to changes, and resource utilization (Neely,1995). As mentioned before, HSCs are working with limited resources and budgets, and must use these limited resources with high utilization rates for accurately assessed needs. Cost of different operations, such as warehousing, transportation and procurement, and the utilization of trucks and warehouses will be used as indicators of efficiency. Resilience can be defined as the ability to resist disruptions or return to the original or a more desirable state after being disturbed (Hosseini, 2016). Considering the continuously changing conditions and potential disruptions on the supply chain networks, resilience plays a key role in smoothly conducting response efforts with minimum impact from unexpected consequences of natural disasters. To measure the resilience of HSCs, we are planning to test their performance under scenarios with different disruption levels. Finally, the UN defines sustainability as “meeting the needs of the present without compromising the ability of future generations to meet their own needs” (UN World Commission, 1987). Since HSCs are conducting very large-scale operations, following green supply chain practices can have a large impact on global warming and climate change issues. In our model, CO_2 emissions resulting from transportation efforts will be the indicator for sustainability. This set of performance indicators can be expanded with more detailed and specific indicators as well; at this level of the research, we are considering very generic indicators.

- **Article experiment description**

As the simulation progresses, factors previously identified are integrated. So far, the programming work mainly focused on the node interconnection at the network design level, as it is a keystone for most of the following PI-oriented factors. Currently, the connection is specific to nodes within the IRC network in the situation of sourcing when selecting a supplier is required. For this factor named - *sourcing* -, three levels are considered: i) the hierarchical approach (i.e., depending on the type of warehouse, another type is preselected as a supplier), ii) the closest supplier approach (i.e., the closest supplier is selected if he can fulfil the received demand) and iii) the multiple sourcing approach (i.e., a list of suppliers is suggested based on proximity and available inventory). A second factor tests two shortage management approaches namely the FIFO that fulfils the first orders received compared to the - *equity* - approach where all demands are at least partially fulfilled by splitting the inventory on hand. Finally, the third tested factor concerns inventory dispatch before the event occurrence and investigates an equal repartition versus a division based on hazard risk. Regarding performance, we focus on the effectiveness of the response through the delivery indicator monitoring the quantity of item distributed over time compared to the demand from beneficiaries and when they match.

Step 6. Make production runs and verify. A first successful experiment run has been initiated based on the experiment design previously depicted. With 3 levels (hierarchical, closest supplier, and multiple sourcing) for the sourcing approach factor and 2 levels (FIFO and equity) for the shortage management factor and 2 levels (equal and risk proportional) for the district warehouses initial stock dispatch, a set of 12 scenarios have been simulated. According to the field data gathered from practitioners, the conditions are the following: initial total district warehouse inventory is set to 100,000, initial province warehouse inventory to 20,000, and runs last for eight virtual days. Verification steps include error message detection and ensuring the correct collection of output data.

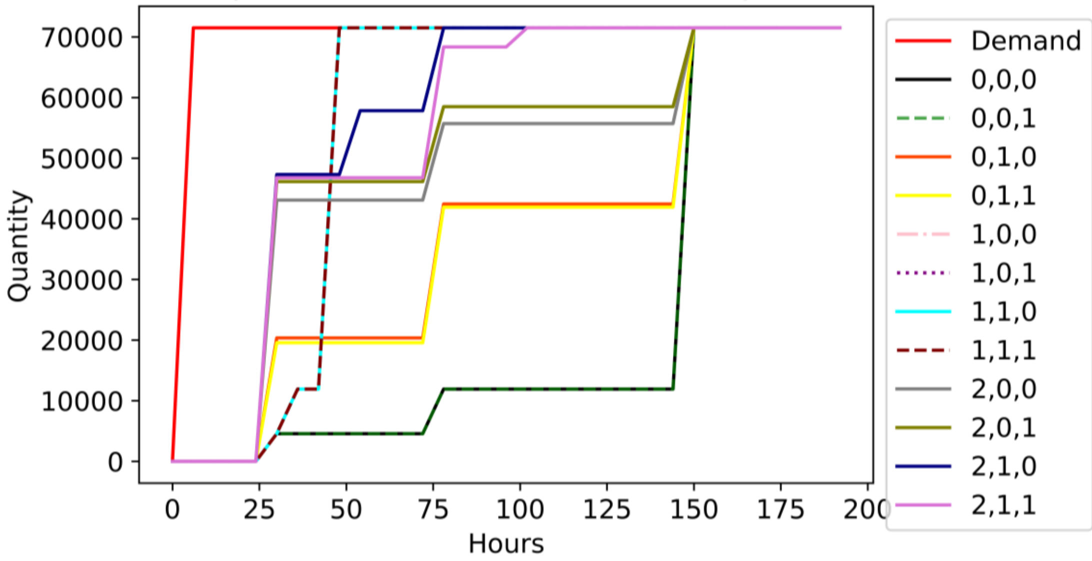


Figure 4 Scenarios performance results in delivering relief items to beneficiaries in the aftermath of a disaster

Steps 7 and 8. Analyze output data, validate, and present results. Once the runs were completed, we gathered performance data from all the scenarios to start analyzing the results. We issued the following graph which shows each scenario’s performance in delivering relief items compared to the cumulative demand over time (see Figure 4, red line). Scenarios are referred to using a tuple format of three figures (X,Y,Z). The first tuple number defines the level for the sourcing factor (0 = hierarchical, 1 = closest supplier, and 2 = multiple sourcing). The second tuple number defines the level for the shortage management factor (0 = FIFO and 1 = equity) while the third tuple number describes the initial district warehouse stock dispatch (0 = equal and 1 = risk proportional). The baseline is the scenario (0,0,0) and the associated performance result is illustrated by the black line. It appears to be the slowest scenario in delivering items. Indeed, it requires almost 6 days to serve all the demand (which is coherent with practitioners’ feedbacks) while over this time only 17% of the demand is satisfied. The introduction of the equity approach for shortage management (scenario (0,1,0) and (0,1,1)) allows serving around 24% more beneficiaries in the same amount of time. The iterations ((2,0,0) and (2,0,1)) using the multiple sourcing approach without the equity also fulfil the total demand in 6 days. However, compared to the baseline, they multiply by ten the number of people served over this period. In contrast, scenarios using the multiple supplier sourcing approach without the equity rule serve even better the demand and, in less time (3.25 days for scenario (2,1,0) and 4.25 days for scenario (2,1,1)). Finally, the fastest scenarios are those using the closest supplier sourcing rule by cutting the delivery time by 68% compared to the baseline. From these observations, it seems the sourcing approach has a significant impact on the supply lead time and quantity served compared to the other factors. Especially, the test of two hyperconnected approaches (closest supplier and multiple sourcing) breaking with the hierarchical structure and allowing horizontal and vertical sourcing connection have seriously improved the response performance results in terms of time and distribution. The equity approach for dealing with the shortage leads to an increase in the number of people served in a timely manner. In contrast, the stock dispatch factor does not seem to have a real impact on this performance indicator. The reason may come from the too slight difference regarding the stock dispatch based on the proportionality defined compared to the equal dispatch. An additional reason may come from the current consideration of a single disaster to respond to.

4 Agenda and Further Research

In this paper, we explained the need for efficient, effective, resilient, and sustainable humanitarian response in response to a disaster. We developed a simulation platform in AnyLogic software and investigated the novel approach of using PI concepts in the HSC realm. In a computational study involving past disaster scenarios that occurred in Indonesia, we simulated operations lead by the baseline and hyperconnected behaviors and showed significant improvement in terms of response speed and coverage. The paper serves as one of the initial investigations for utilizing PI in HSC operations, and as such, it opens several avenues for future research, including optimizing the pre-positioning of the relief items and other resources before disasters to better serve the population. In addition, in the post-disaster time, there is a need for a more efficient and robust work plan for the distribution of the relief items to the PODs. However, the current measures are not sufficient and new indicators to evaluate the operations need to be considered. Those indicators should be more beneficiary-oriented and reflect the life quality of the affected populations in the aftermath of a disaster.

References

- Abdoulkadre, A., Intissar, B.O., Marian, M., and Benoit, M. (2014). Towards Physical Internet Enabled Interconnected Humanitarian Logistics. 1st International Physical Internet Conference 30.
- Abidi, H., de Leeuw, S., and Klumpp, M. (2014). Humanitarian supply chain performance management: a systematic literature review. Supply Chain Management: An International Journal 19, 592–608.
- Banks, J. (1998). Handbook of Simulation: Principles, Methodology, Advances, Applications, and Practice | Wiley (USA: Jonh Wiley & Sons, Inc).
- Beamon, B.M., and Balcik, B. (2008). Performance measurement in humanitarian relief chains. International Journal of Public Sector Management 21, 4–25.
- Chung, C. (2004). Simulation Modeling Handbook: A Practical Approach (CRC Press).
- Cosgrave, J., and Herson, M. (2008). Perceptions of crisis and response: A synthesis of evaluations of the response to the 2005 Pakistan Earthquake.
- Ergun O., L. Gui, J. L.Heier Stamm, P. Keskinocak, J. Swann (2014). Improving humanitarian operations through technology-enabled collaboration. Prod. Oper. Manag. 23(6): 1002–1014.
- Farahani, R., and Rezapour, S. (2011). Logistics Operations and Management: Concepts and Models (Elsevier).
- Grest, M., Lauras, M., and Montreuil, B. (2020). A Humanitarian Supply Chain Maturity Model. p. 613.
- Grest, M., Lauras, M., and Montreuil, B. (2021). Conception d'un modèle de simulation pour l'évaluation de la pertinence de l'Internet Physique pour les chaînes d'approvisionnement humanitaires. (Grenoble), p.
- Haavisto, I., and Goentzel, J. (2015). Measuring humanitarian supply chain performance in a multi-goal context. Jrnl Hum Log and Sup Chn Mnage 5, 300–324.
- Hosseini S., K. Barker, J. E. Ramirez-Marquez. (2016). "A Review of Definitions and Measures of System Resilience." Reliability Engineering & System Safety 145: 47–61.
- İlhan, A. (2011). The Humanitarian Relief Chain. South East European Journal of Economics and Business 6, 45–54.
- Jahre, M., Kembro, J., Rezvanian, T., Ergun, O., Håpnes, S.J., and Berling, P. (2016). Integrating supply chains for emergencies and ongoing operations in UNHCR. Journal of Operations Management 45, 57–72.
- Law, A.M. (2014). Simulation Modeling and Analysis (Dubuque: McGraw-Hill Higher Education).

- Neely A., M. Gregory, K. Platts (1995). "Performance measurement system design". International Journal of Operations & Production Management. Vol. 15 No. 4. pp. 80-116.
- Peretti, U., Tatham, P., Wu, Y., and Sgarbossa, F. (2015). Reverse logistics in humanitarian operations: challenges and opportunities. Journal of Humanitarian Logistics and Supply Chain Management 5, 253–274.
- Sheard, S.A., and Mostashari, A. (2011). Complexity Types: From Science to Systems Engineering. INCOSE International Symposium 21, 673–682.
- Stephenson, J. (2005). Making humanitarian relief networks more effective: operational coordination, trust and sense making. Disasters 29, 337–350.
- Tomasini, R., and Wassenhove, L.V. (2009). Humanitarian Logistics (Springer).
- Van Wassenhove, L.N. (2006). Humanitarian aid logistics: supply chain management in high gear. Journal of the Operational Research Society 57, 475–489.
- United Nations World Commission on Environment and Development. (1987). Article 6 of the United Nations Framework Convention on Climate Change.

Resilient Hyperconnected Logistics Hub Network Design

Onkar Kulkarni[1], Yaarit M. Cohen[1,2], Mathieu Dahan[1,2], Benoit Montreuil[1,3]

1. School of Industrial and Systems Engineering, Physical Internet Center, Supply Chain & Logistics Institute
2. Center for Health and Humanitarian Systems

Georgia Institute of Technology, Atlanta, Georgia 30332, USA

3. Coca-Cola Chair in Material Handling and Distribution

Corresponding author: *onkarkulkarni@gatech.edu*

Abstract: *Logistics networks frequently face disruptions inducing an increase in delivery costs and delays. This paper studies the design of resilient hyperconnected logistics hub networks for the Physical Internet, modeled as an integer programming problem. The objective is to open logistics hubs in order to connect each origin and destination using multiple minimum length edge-disjoint paths. To estimate the resilience of the designed networks, we propose graph-theoretic measures involving (i) the maximum number of edge-disjoint paths connecting each origin and destination, and (ii) the number of short paths traversing each edge. We develop a case study to design a class of parcel delivery networks in China and evaluate the impact of various disruption scenarios on the resulting distance traveled by parcels. Our results show the relevance of the proposed resilience measures and the increased capability of the designed networks to sustain disruptions in comparison to traditional logistics networks.*

Conference Topics: *Interconnected freight transportation, logistics and supply networks*

Keywords: *Hyperconnected Logistics, Parcel Delivery Network, Resilience, Intercity Parcel Delivery, Network Topology, Physical Internet.*

1. Introduction

The recent surge in e-commerce and world trade has led the parcel delivery industry to be one of the fastest growing industries (Jin, 2018). In addition, the fierce competition among courier companies motivates the need for fast and convenient delivery of parcels to customers' doorsteps. This propels the industry to be more asset-intensive for quicker parcel delivery to its customers who are spread out across wide geographical regions (Jin, 2018). As a consequence, the industry requires meticulous planning and proper execution. The planning involves strategic decisions such as logistics hub network design and tactical decisions in parcel and vehicle scheduling for timely and resilient parcel delivery.

An extensive research has already been conducted in the domain of logistics hub network design. Several works have studied the hub-and-spoke configuration for logistics hub networks (O'Kelly and Miller, 1994). The configuration is designed based on three assumptions. First, it assumes no direct delivery between the origin and destination nodes. Second, it considers the availability of all possible transportation links between hubs for travel. Last, it assumes a discount factor to model cost savings for routing parcels through the hubs. In the past, researchers have designed logistics hub networks by relaxing one or more of these assumptions but not all (Lin and Chen, 2008; Gelareh et al., 2010; Ben-Ayed, 2013). It has been shown that such a hub-and-spoke network topology does not perform well in high demand scenarios as they may cause congestion of parcels at hubs during peak delivery times (Tu and Montreuil, 2019). Traditional networks constrain flows through two-tier hub-and-spoke structures or force each hub to be equipped to deal with vast sets of connecting hubs in single-tier point-to-point networks. The lead times, costs, and travel induced by these network topologies are roadblocks toward addressing the efficiency, service capability, resilience, and sustainability challenges faced by the parcel delivery industry (Montreuil et al., 2018). To improve the overall parcel delivery process and to overcome the current limitations, hyperconnected logistics networks are proposed in the Physical Internet (PI) (Montreuil, 2011).

Hyperconnected logistics networks are multi-plane interconnected meshed networks that link open-access hubs present on multiple planes. Together they shape an open network of networks, termed a logistics web (Montreuil et al., 2013, 2018). An initial approach to design multi-plane hyperconnected networks utilized historical demand data and geographical locations as criteria for prospective hub candidates and presented a network flow formulation to select the hubs (Tu and Montreuil, 2019; Ducret, 2014). In PI, these meshed networks can serve as open logistics web infrastructure to be leveraged by the participating players through asset sharing (Montreuil, 2013). Based on this principle, multi-tiered open supply webs have been designed for various purposes such as food distribution, mail delivery, and parcel delivery (Ballot et al., 2012, 2016).

Nevertheless, all logistics networks, including those shaping the PI's logistics web, face disruptions caused by frequent events such as power outages or major traffic jams, as well as low-probability high-impact events such as natural disasters, pandemics, and deliberate attacks. Such disruptions lead to delayed parcel deliveries, increased delivery costs, and excess pressure on functional network components. Considerable efforts have been devoted to gauge the resilience of various logistics and transportation networks in the past. The correlation between structure of the network and its resilience has been showcased through several disruption experiments (Osei-Asamoah and Lownes, 2014). In addition, graph-theoretic measures such as k-shortest path lengths, number of edge-disjoint paths, and node reachability have been utilized to evaluate the resilience of networks (Ip and Wang, 2009, 2011; Herrera et al., 2016). Although these investigations are helpful to assess the resilience of networks, they are rarely used at the network design level (Newman, 2005; Ip and Wang, 2009, 2011; Osei-Asamoah and Lownes, 2014). Some investigations consider disruptions at hubs and transportation links to design a small-scale network (Zhalechian et al., 2018). However, such a small-scale network reveals little to an industry that aims to persistently deliver parcels across a wide geographical region.

This paper proposes an integer programming approach that employs networks' structural properties to design large-scale resilient hyperconnected logistics meshed networks. Specifically, we define the problem of selecting logistics hubs to open in order to connect each origin and destination with multiple edge-disjoint paths of minimum length while ensuring hub hyperconnectivity. This aims to ensure that the network can sustain concurrent edge disruptions that do not induce excessive travel between the origin-destination pairs. In order to estimate the resilience of these networks, we propose two resilience metrics based on network topology that are suitable for the current setting. The first metric analyzes the maximum number of edge-disjoint paths for each origin-destination pair while the second metric studies the number of short paths that traverse each edge. We design multiple resilient logistics networks for the ground transportation and consolidation of parcels in China. In order to evaluate the resilience of the proposed networks, we assess the impact of disruptions on the resulting shortest path lengths in the networks. These disruptions are either random (one or two edges are randomly chosen to be disrupted) or localized (an edge is picked randomly and edges within a specific radius are disrupted). We compare the results of the disruption experiments for the designed networks with those of traditional logistics networks and find that the designed networks have a higher capability to sustain disruptions. Our computational results validate the relevance of the proposed resilience metrics.

The rest of the paper is organized as follows: Section 2 describes the problem setting, formulates the optimization model based on edge-disjoint paths, and presents the resilience measures. In Section 3, we design multiple resilient hyperconnected ground transportation networks across China for parcel delivery and evaluate their resilience using several edge-disruption experiments. Finally, Section 4 provides concluding remarks and avenues for future research.

2. Problem Definition

We consider a logistics company, a group of such companies, or a territorial authority, that seeks to design a resilient hyperconnected intercity logistics hub network to transport commodities between a set of locations that can be origins O or destinations D for different commodities. Let $\mathcal{P} \subseteq O \times D$ denote the set of Origin-Destination (O-D) pairs to be served by opening N logistics hubs among a discrete set of candidate locations, denoted H. We let $G = (O \cup D \cup H, E)$ be the directed graph where E is the set of directed edges $(i,j) \in (O \cup D \cup H)^2$ representing the available transportation links connecting locations i and j. For each O-D pair $p = (s,t) \in \mathcal{P}$, an $s - t$ path $\{s, h_1, \dots, h_n, t\}$ of size $n + 2$ is a sequence of adjacent nodes starting at node s and ending at node t. In other words, an $s - t$ path starts at s, visits logistics hubs in between by traversing network edges to finally reach destination t. The goal is to select a subset of hub locations $H_o \subseteq H$ with $|H_o| \leq N$ so that the subgraph induced by the set of nodes $(O \cup D \cup H_o)$ connects every O-D pair in $\mathcal{P}$, is efficient in terms of transportation (distance traveled), and is resilient against possible edge disruptions. To this end, we next develop an optimization model based on the k shortest edge-disjoint paths between each O-D pair.

2.1 Integer Programming Formulation

The optimization problem we consider aims to select hubs so as to minimize the total length of the k shortest edge-disjoint paths between each O-D pair in the induced subgraph. We say that k paths are edge-disjoint if no edge is traversed by more than one of the paths. The motivation is that an O-D pair that is connected by several edge-disjoint paths is less likely to be fully disconnected after multiple edge disruptions.

We model this problem as an Integer Program (IP) using an edge-based formulation. For every edge $(i,j) \in E$, we denote its length by d_{ij}. We consider the binary variables x_h for all $h \in H$ that represent the opened hubs. In addition, we define for every O-D pair $p \in \mathcal{P}$ and every $(i,j) \in E$ the binary variable $f_{i,j}^{p}$ equal to 1 if (i,j) is traversed by one of the k shortest edge-disjoint paths connecting p. Then, the problem can be formulated as follows:

$$min \sum_{p \in \mathcal{P}} \sum_{(i,j) \in E} d_{ij} f_{ij}^{p} \tag{1}$$

$$subject\ to$$

$$\sum_{h \in H} x_h \leq N \tag{2}$$

$$\sum_{j \in D \cup H | (s,j) \in E} f_{sj}^{p} = k, \qquad \forall p = (s,t) \in \mathcal{P} \tag{3}$$

$$\sum_{i \in O \cup H | (i,t) \in E} f_{it}^{p} = k, \qquad \forall p = (s,t) \in \mathcal{P} \tag{4}$$

$$\sum_{j \in D \cup H | (i,j) \in E} f_{ij}^{p} = \sum_{j \in O \cup H | (j,i) \in E} f_{ji}^{p}, \qquad \forall p = (s,t) \in \mathcal{P}, \forall i \in O \cup D \cup H \backslash \{s,t\} \tag{5}$$

$$2 \cdot f_{ij}^{p} \leq x_i + x_j, \qquad \forall p \in \mathcal{P}, \forall (i,j) \in E \tag{6}$$

$$f_{ij}^{p} \in \{0,1\} \qquad \forall p \in \mathcal{P}, \forall (i,j) \in E$$

$$x_h \in \{0,1\} \qquad \forall h \in H.$$

Objective (1) minimizes the sum over all O-D pairs $p = (s,t) \in \mathcal{P}$ of the lengths of the k shortest edge-disjoint paths between s and t in the subgraph of G induced by the subset of nodes $O \cup D \cup \{h \in H \mid x_h = 1\}$. Indeed, constraint (2) ensures that at most N logistics hubs are opened while constraints (3)-(5) ensure that k edge-disjoint paths connect each O-D pair. Furthermore, constraints (6) prevent a path from traversing an unopened hub. Therefore, at optimality, the binary variables f indeed represent the k shortest edge-disjoint paths between each O-D pair in the induced subgraph.

2.2 Resilience Measures

To estimate the resilience of logistics hub networks, we consider two graph-theoretic measures. First, we determine for each O-D pair the maximum number of edge-disjoint paths in the network connecting them. This measure indicates the number of simultaneous edge disruptions required to disconnect an O-D pair. In particular, it serves as a proxy to evaluate network robustness: A logistics network with higher number of edge-disjoint paths per O-D pair would be able to handle higher number of concurrent edge disruptions and still maintain its operations.

Second, we define and analyze a new edge-betweenness centrality measure, which computes for each edge $(i,j) \in E$ the fraction of $s-t$ paths, for all $(s,t) \in \mathcal{P}$, of length no more than $(1+\alpha)$ of the shortest $s-t$ path length that traverse (i,j). The parameter α is a nonnegative number that represents the maximum added distance that the logistics company is willing its commodities to travel. More formally, the edge-betweenness of an edge $(i,j) \in E$ can be expressed as follows:

$$\frac{\sum_{(s,t)\in\mathcal{P}} \begin{pmatrix} \#\, s-t \text{ paths traversing } (i,j) \text{ of length at most} \\ (1+\alpha) \times \text{shortest } s-t \text{ path length} \end{pmatrix}}{\sum_{(s,t)\in\mathcal{P}} \begin{pmatrix} \#\, s-t \text{ paths of length at most} \\ (1+\alpha) \times \text{shortest } s-t \text{ path length} \end{pmatrix}}.$$

The premise is that, for an O-D pair $(s,t) \in \mathcal{P}$, paths of length more than $(1+\alpha)$ of the shortest $s-t$ path length induce unnecessary travel and hence will not be utilized for commodity transfer by the logistics company. Therefore, such paths are not relevant in computing the betweenness centrality of an edge. This measure helps us identify the transportation edges that are most critical and that are likely to induce a high increase in travel time if disrupted. A network containing edges with low betweenness centrality is more likely to be resilient to disruptions, as commodities can be rerouted along alternative paths with limited added distance.

3. Computational Analysis

3.1 Case Study Description

We apply the developed methodology to design a resilient hyperconnected intercity parcel logistics hub network to be the backbone infrastructure of China for ground transportation and consolidation of parcels. Core to the Physical Internet spanning China, this network could be leveraged by multiple parcel delivery companies to move numerous millions of parcels every day among Chinese cities. The network is to serve regions that house 93.58% of the Chinese population, are spread across 95.09% of the Chinese inhabitable land, and generate 94.42% of total Chinese GDP (Li et al., 2018). Alternatively, the network topology could also be used by a major logistics provider as an internally shared Physical Intranet.

To design the network, we consider intersections of major highways and existing city-based gateway hubs (inbound/outbound) as candidate locations for intercity hubs (set H). These locations help bypass the intracity traffic and probable unnecessary delays. In addition, due to regulations imposed by the Chinese government, a truck driver can drive for 11 hours per day. Hence, we limit the transportation edges ($(i, j) \in E$) to up to 5.5 hours' drive time between locations to enable truck drivers to return home daily while the parcels keep moving toward their destinations.

3.2 Computational Results

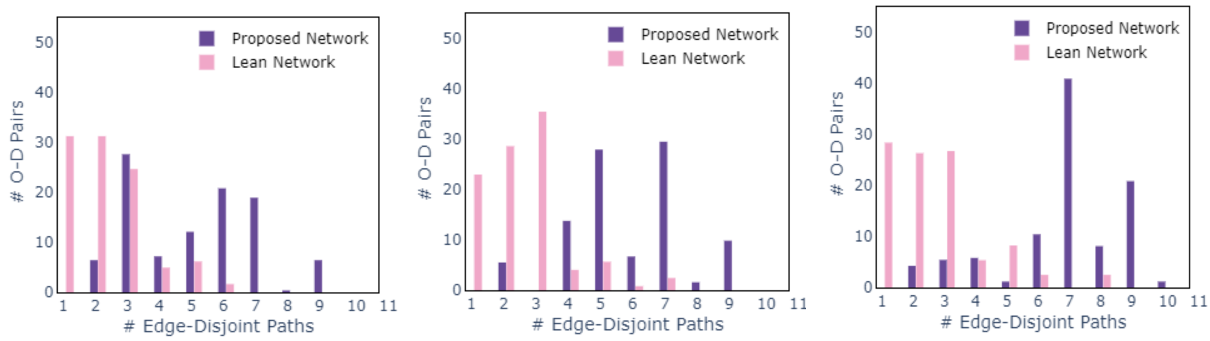
By solving the IP formulated in Section 2, with $k = 2$ edge-disjoint paths for each O-D pair, we designed multiple potential hyperconnected logistics hub networks for different numbers N of logistics hubs to be opened: 70, 80, and 90. The hyperconnected networks with 70 and 90 hubs are shown in Figures 1 and 2, respectively. In these Figures, the yellow asterisks represent regional hubs, and the red lines represent transportation edges between the regional hubs.

Figure 1: IP-based hyperconnected 70-hub logistic network

Figure 2: IP-based hyperconnected 90-hub logistic network

To compare these resilience-optimized hyperconnected logistics networks with traditional ones, we designed lean networks by selecting hubs to open with the goal of minimizing the (single) shortest path length between each O-D pair. Similarly, the transportation edges were limited to 5.5 hours' drive, and the lean networks were generated with the opening of 70, 80, and 90 hubs.

Next, we analyze the resilience measures defined in Section 2.2 for the proposed resilience-optimized and lean networks. The edge-disjoint path distribution over O-D pairs for both types of networks are depicted in Figure 3. We observe that the proposed networks have a greater number of edge-disjoint paths overall compared to the lean networks. In the lean networks, most of the O-D pairs have at most 3 edge-disjoint paths. This implies that 3 or more concurrent edge disruptions can disconnect several O-D pairs if the lean networks were utilized, while they would have minimal impact on the proposed networks. We find that 5 simultaneous edge disruptions are enough to disconnect all O-D pairs in the lean networks, while the 70-, 80-, and 90-hub resilience-optimized networks respectively require 7, 9, and 10 simultaneous edge disruptions. Furthermore, we observe that as the number of opened hubs increases, the number of edge-disjoint paths increases significantly in the proposed networks but is limitedly increased in the lean networks. This suggests that the proposed networks are better prepared to sustain a greater number of simultaneous edge disruptions compared to the lean networks.

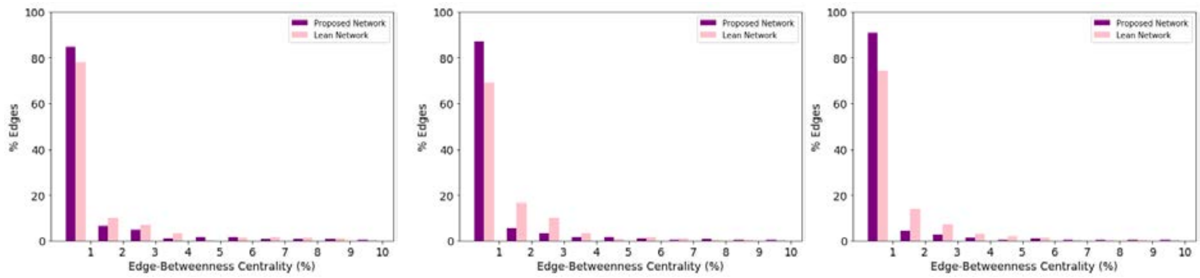


(a) 70-hub networks (b) 80-hub networks (c) 90-hub networks

Figure 3: Edge-disjoint path distribution for the proposed and lean networks

For the edge-betweenness centrality measure, we set $\alpha = 0.2$, i.e., we consider all the paths of length no more than 120% of the corresponding shortest path length. We then compute for each edge the percentage of such paths that traverse that edge. The distribution of this newly defined edge-betweenness centrality measure is depicted in Figure 4. In the 70-, 80-, and 90-hub lean networks, we observe that 78%, 69%, and 74% of the edges, respectively, have betweenness centrality values less than 1%. By comparison, 84%, 87%, and 91% of the edges in the corresponding resilient-optimized networks have betweenness centrality values less than 1%. We note that for the lean networks, the proportion of edges with lower edge-betweenness centrality values remains equivalent as the size of the network increases. In contrast, as we allow more hubs to be opened in the proposed networks, the proportion of edges with low centrality values increases. This suggests that the proposed resilience-optimized networks do leverage the value of opening additional hubs to improve their resilience.

Overall, the lean networks comprise a comparatively greater proportion of edges with high betweenness centrality values. Such edges are critical in keeping the network operational and may cause considerable impact when disrupted. Specifically, the impact includes significant addition in travel time beyond shortest paths, and potential loss of connectivity between O-D pairs. In contrast, the proposed networks have a low proportion of such critical edges, which decreases the chances for high impact during disruptions.



(a) 70-hub networks (b) 80-hub networks (c) 90-hub networks

Figure 4: Edge-betweenness centrality distribution for the proposed and lean networks

In order to validate the resilience of the proposed networks, we analyzed the impact of one or multiple disruptions on the networks. In general, non-adversarial disruptions are either random (i.e., arbitrary set of edges is disrupted) or localized (i.e., geographical regions of varied sizes are impacted, and all the transportation edges of the network within the impacted zone are disrupted). Hence, we ran two sets of experiments: (i) random disruptions, where we examined all possibilities of a single edge disruption and two simultaneous edge disruptions; and (ii) localized disruptions, where we examined all the possibilities of a region centralized in one network edge and impacting edges within an impact radius of 0.5 hours, 1 hour, 1.5 hours and 2 hours of travel time, respectively. The metrics used for comparison quantify the average number of disconnected O-D pairs and the average increase in shortest path length for the O-D pairs that remained connected after each disruption scenario. These metrics provide a comprehensive idea of network performance under disruption. The results of these experiments and the performance comparison with respect to the lean networks are presented in Tables 1 and 2.

The results depict that the lean networks with higher number of hubs maintain better connectivity than lean networks with lower number of hubs. Still, the proposed resilience-optimized networks outperform the lean networks as they guarantee flow of parcels between all the O-D pairs for these disruption scenarios. Importantly the proposed networks achieve this guarantee with lower number of hubs. Furthermore, the additional travel induced by the disruptions in the proposed networks is smaller than in the lean networks. We observe that the increase in the number of simultaneous edge disruptions or in impact radius causes a considerable rise in induced travel and a gradual increase in disconnected O-D pairs for the lean networks. In contrast, the proposed networks worsen at a slower rate as the disruptive impact increases.

The reason for poor performance of lean networks in terms of connectivity of O-D pairs and induced additional travel time can be associated with the presence of edges with higher betweenness centrality values. These edges are critical in nature as they are part of a larger number of short paths for several O-D pairs. When these edges are disrupted, they are most likely to either induce substantial additional travel time or even worse, disconnect O-D pairs. Moreover, the presence of such edges in lean networks of various sizes demonstrates the consistent worse performance for the disruption scenarios. For the proposed resilience-optimized networks, the proportion of higher centrality edges is less, and even lesser as the network size increases. When smaller-centrality edges are disrupted, the impact on the connectivity of O-D pairs and induced additional travel time is insignificant. Hence, the proposed networks are impacted to a smaller extent.

The experiments demonstrate that the proposed network design optimization generates resilient networks that can handle disruptions occurring randomly across their edges, or impacting a localized region, in a better way compared to the lean networks. The results are in tandem with the insights obtained from the topological resilience metrics. Hence, these disruption experiment results validate the proposed resilience metrics.

Table 1: Simultaneous edge-disruption experiment

Disruption Scenario Details	Comparison Metrics	70-Hub Proposed Network	70-Hub Lean Network	80-Hub Proposed Network	80-Hub Lean Network	90-Hub Proposed Network	90-Hub Lean Network
1 Edge disruption	Average # of disconnected O-D pairs	0	0.154	0	0.112	0	0.071
	Average total added length for connected O-D pairs (hrs)	2.370	10.263	5.364	12.917	6.361	9.687
2 Edge disruptions	Average # of disconnected O-D pairs	0	0.316	0	0.228	0	0.146
	Average total added length for connected O-D pairs (hrs)	2.370	10.263	5.364	12.917	6.361	9.687

Table 2: Localized edge-disruption experiment

Impact Radius (hrs)	Comparison Metrics	70-Hub Proposed Network	70-Hub Lean Network	80-Hub Proposed Network	80-Hub Lean Network	90-Hub Proposed Network	90-Hub Lean Network
0.5	Average # of disconnected O-D pairs	0	0.154	0	0.113	0	0.071
	Average total added length for connected O-D pairs (hrs)	5.573	10.763	5.404	13.802	6.767	10.257
1	Average # of disconnected O-D pairs	0	0.155	0	0.119	0	0.073
	Average total added length for connected O-D pairs (hrs)	6.489	13.381	6.672	17.653	6.855	12.880
1.5	Average # of disconnected O-D pairs	0	0.169	0	0.137	0	0.084
	Average total added length for connected O-D pairs (hrs)	7.625	22.702	8.577	26.541	7.404	17.035
2	Average # of disconnected O-D pairs	0	0.184	0	0.151	0	0.167
	Average total added length for connected O-D pairs (hrs)	7.709	25.505	13.372	32.284	8.933	20.131

4. Conclusion

In this paper we motivate the need for resilient logistics hub networks in the realm of Physical Internet, especially in the parcel delivery industry that is rapidly growing. We formulate an integer program to design a resilient hyperconnected logistics hub network which leverages structural properties of resilient networks, such as edge-disjoint paths. The paper also proposes topological measures to assess the resilience of these logistics hub networks: the maximum number of edge-disjoint paths, for each O-D pair; and a new edge-betweenness centrality measure. While the former measure serves as a proxy to evaluate the network robustness, the latter helps us identify the critical transportation edges whose disruption induces O-D travel time increases.

The devised methodology is applied to develop a resilient logistics hub network across China, which can be utilized by the participating logistics providers in the PI. After conducting random and localized disruption experiments for multiple networks, it can be seen that the generated networks are better equipped to sustain possible disruptions than the traditional logistics networks. In particular, they ensure better connectivity between all O-D pairs and comparatively smaller disruption-induced added travel time than lean networks. These disruption experiment results are in tandem with the predictions from the proposed resilience measures and hence validate the resilience measures as well.

This paper serves as one of the initial investigations designing resilient networks in the context of Physical Internet, and as such, it opens several avenues for future research. The proposed approach to design resilient hyperconnected network focuses mainly on network topology and can be extended to consider parcel flows, transportation costs, hub capacity, and potential consolidation opportunities. Moreover, examining the resilience of networks under strategic attacks (disruptions) could help in meaningful ways to develop even more resilient logistics hub networks.

References

Ballot E., O. Gobet, and B. Montreuil (2012): Physical Internet Enabled Open Hub Network Design for Distributed Networked Operations. In Service Orientation in Holonic and Multi-Agent Manufacturing Control (pp. 279–292).

Ben-Ayed O. (2013): Parcel distribution network design problem. Operational Research, 13(2), 211–232.

Ducret R. (2014): Parcel deliveries and urban logistics: Changes and challenges in the courier express and parcel sector in Europe - The French case. Research in Transportation Business and Management, 11, 15–22.

Gelareh S., S. Nickel, and D. Pisinger (2010): Liner shipping hub network design in a competitive environment. Transportation Research Part E: Logistics and Transportation Review, *46*(6), 991–1004.

Herrera M., E. Abraham, and I. Stoianov (2016): A Graph-Theoretic Framework for Assessing the Resilience of Sectorised Water Distribution Networks. Water Resources Management, 30(5), 1685–1699.

Ip W.H., and D. Wang (2009): Resilience evaluation approach of transportation networks. *Proceedings of the 2009* International Joint Conference on Computational Sciences and Optimization, CSO 2009, 2, 618–622.

Ip W.H., and D. Wang (2011): Resilience and friability of transportation networks: Evaluation, analysis and optimization. IEEE Systems Journal, 5(2), 189–198.

Jin, H. W. (2018): Restructuring parcel delivery network by considering dynamic customer demand. Business Administration and Management, 21(2), 83–95.

Lin C., and S. Chen (2008): An integral constrained generalized hub-and-spoke network design problem. Transportation Research Part E: Logistics and Transportation Review, 44, 986–1003.

Li, M., B. He, R. Guo, Y. Li, Y. Chen, and Y. Fan (2018): Study on population distribution pattern at the county level of China. Sustainability (Switzerland), 10(10).

Montreuil B. (2011): Toward a Physical Internet: meeting the global logistics sustainability grand challenge. Logistics Research, 3(2–3), 71–87.

Montreuil B., R.D. Meller and E. Ballot (2013): Physical Internet Foundations, in Service Orientation in Holonic and Multi Agent Manufacturing and Robotics, ed. T. Borangiu, A. Thomas and D. Trentesaux, Springer, p. 151-166.

Montreuil B., S. Buckley, L. Faugere, R. Khir, S. Derhami (2018).: Urban parcel logistics hub and network design: the impact of modularity and hyperconnectivity. 15th IMHRC Proceedings, 19.

Newman M.E.J. (2005): A measure of betweenness centrality based on random walks. Social Networks, 27(1), 39–54.

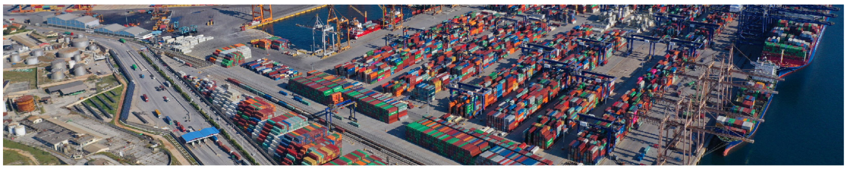
O'Kelly M.E., and H.J Miller (1994): The hub network design problem. Journal of Transport Geography, 2(1), 31–40.

Osei-Asamoah A., and N.E. Lownes (2014): Complex network method of evaluating resilience in surface transportation networks. Transportation Research Record, 2467(2467), 120–128.

Tu D., and B. Montreuil (2019): Hyperconnected Megacity Logistics: Multi-Tier Territory Clustering and Multi-plane Meshed Hub Network Design. 6th International Physical Internet Conference, 159–167.

Winkenbach M., A. Roset, and S. Spinler (2016): Strategic redesign of urban mail and parcel networks at La Poste. Interfaces, 46(5), 445–458.

Zhalechian M., S.A. Torabi., and M. Mohammadi (2018): Hub-and-spoke network design under operational and disruption risks. Transportation Research Part E: Logistics and Transportation Review, 109(May 2017), 20–43.



Modular and Mobile Design of Hyperconnected Parcel Logistics Hub

Sevda Babalou[1], Wencang Bao[1], Benoit Montreuil[1], Leon McGinnis[1,2] ,
Shannon Buckley[1], Ali Barenji[1]
[1]Physical Internet Center, Georgia Tech, Atlanta, GA USA
[2]Keck Virtual Factory Lab, Georgia Tech, Atlanta, GA USA
Corresponding authors: zb.babalou@gatech.edu; benoit.montreuil@isye.gatech.edu

Abstract: *This paper employs modularity and mobility (M^2) for designing recently introduced hyperconnected logistics hubs (HLH) for the Physical Internet, where parcels are encapsulated in modular tote-sized containers arriving in mobile racks, and these totes are consolidated by switching totes in shuffling cells to mobile racks with other totes with shared next destinations. The paper introduces the M^2 framework and its modular standard-sized cells, racks and tote containers. Building on the overall HLH concept, the proposed M^2 hub design is a major step forward with its on-the-fly transformability through operations to adapt to the dynamically changing sizes, mixes, characteristics, and flow of modular containers entering the hub and being consolidated and shipped within a short dwell time target. The paper uses a detailed case study to demonstrate the induced adaptability, adjustability, agility, efficiency, resilience, and scalability, and then it reports on an exploratory simulation experiment contrasting the performance of M^2designs.*

Keywords: *Physical Internet, Hyperconnected Logistics, Parcel Logistics Hub, Modular Hub, Adaptability, Modularity, Resilience, Scalability, Robotic Logistics, Consolidation*

1 Introduction

It is now well known that the Physical Internet induces radical transformation of logistics hubs, notably with its emphasis on universal interconnectivity; standardized modular containers enabling open flow consolidation; inter-hub logistics mesh networks; and fast, efficient, seamless, high-quality, safe, secure, resilient, and sustainable operations [1-4]. Parcel logistics hubs are particularly affected as they are to evolve from strictly performing parcel sorting and consolidation from satellite to main hub to satellite to final destination in as direct shipments as possible, toward consolidating as early as possible the parcels sharing next target hubs into modular containers, and then handling, sorting and consolidating modular containers rather than individual containers [5].

With the mesh network topology in the Physical Internet, the portfolio of active origins and destinations at each hub is greatly reduced as the main flow stems from and to nearby hubs: for example, from and to regional hubs within 4-5 hours to allow drivers to get back home everyday while the modular containers are transferred into other carriers at the hubs to keep on flowing at sustained velocity toward their next destination.

In line with [5], [6] has proposed to reimagine the design of parcel logistics hubs supporting intercity and interregional flows, such reconception affecting their physical and control architecture, and has provided such an innovative design based on shuffling, buffering, and staging cells, with minimal fixed assets and potential reliance on smart mobile robotics for supporting the moves and handlings through the hub.

This paper builds on such an innovative design, and proposes a major step forward from it by its on-the-fly transformation through operations to adapt to the dynamically changing size, mix, and characteristics of modular containers flowing through the hub and being consolidated and shipped within a short dwell time target.

We propose a new approach for designing and laying out logistics hubs that is fundamentally built upon modularity [7,8,9,10] and mobility [11,12], pushing further their application in hub design, so as to improve overall adaptability, adjustability, agility, efficiency, resilience, and scalability under high-velocity service requirements, stochastic demand and flow patterns, and streams of more or less predicatible disruptions. The resulting hubs are referred to as M^2 hubs because of their synergetic leverage of modularity and mobility.

2 Design Logic

The key M^2 concepts are to enforce a standard modular configuration of the cells and circulation spaces so that an underlying grid of potential cell locations can be defined. Because of the grid and cell modularity, and the complete mobility of all internal hub equipment (e.g., mobile robots and mobile racks), the buffer and shuffle cells can be dynamically assigned to cell locations, with the equipment smoothly moved to reflect the new assignments.

A M^2 hub is designed in response to the concept of Hyperconnected Logistics Hub (HLH) [6], including: first, consolidation of totes and racks inside standard size trailer/π-containers. As in the PI concept, π-containers designed to facilitate material handling and storage in the facilities [1], The totes and racks are designed to improve the consolidation in the trailers efficiently. At the M^2 hub, the modular cell designed for the consolidation called *DockCell*.

Consolidating the totes with the exact next destination into racks in HLH is accomplished in four operation steps, unloading the trucks, moving the racks for shuffling, shuffling totes between racks, and moving the racks for loading. *StagingCell* supports loading/unloading, *BufferCell* provides temporary rack storage during the consolidation process, and *ShuffleCell* is where actual consolidation takes place. These standard-sized cells provide adaptability and adjustability of the layouts for increasing/decreasing parcel flow through the smart design and allocation of the cells.

The third M^2 concept is robotized transportations to reduce human involvement in the process in which, the mobile robots are embedded in different cells and circulation. In our design we have been used four types of robots, *LoadBots*, *MoveBots*, *ShuffleBots* ,and *ToteBots*. Even though depicted in their robotized version, all operations done by robots in a M^2 hub can be done by humans with adequate equipment, such as lifting handling devices, and/or with human augmentation, such as augmented reality and artificial reinforcements.

2.1 Modularity of the Cells

The Hyperconnected Logistics Hub, where standard totes are sorted and consolidated by their next destination via robot-centric operations, has fewer types of working cells and thus is suitable for modularity design. An advantage of modularity is that robotic logistic hub can be designed faster via grid-based network, and the locations of working cells can be dynamically adjusted according to daily, even hourly demands. We designed the layouts as a robotic logistics hub with minimum human operation and maximum robotic technology. All the blocks have been designed in standard-sized modules for the parcels in the racks. Each rack is also designed to optimize the trucks' space in HLH. The modules design process impacts the performance of loading/unloading, shuffling, and transferring of the racks inside the facility. We hereafter describe the modular dock, shuffle, staging, and buffer cells, captioned as object-oriented *DockCell, ShuffleCell, StagingCell, and BufferCell.*

2.1.1 DockCell

In this facility, everything has been designed to be standard-sized modules ,and all the space modules are designed based on the dimensions of the racks and operational robots. The design process based on the consolidation of totes and racks started with maximizing utilization of the trucks and defining appropriate design for totes and racks for standard size trailers.

The chosen standard trailer has the dimensions of 47’7”, 9’1”, 9’8”(Figure 1). We considered the side by side orientation of the racks in the trailer relative to the door (Figure 2).

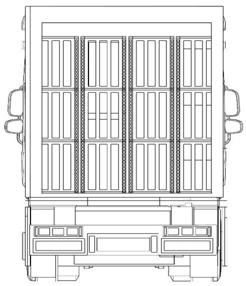


Figure 1: Side-by-Side Orientations

Figure 2: Trailer Dimensions

The design of the totes and racks starts with truck/trailer inside dimensions(47’3”, 8’3”, 9’). Considering clearances inside the truck/trailer, we determine the outside dimensions of the racks. Then we consider how the structure of the racks consumes space to determine the inside dimensions of the rack. Finally, considering clearances inside the racks, determine the actual tote dimensions (Figure 3). We are using the maximum space of the racks for the totes design, considering the ideal tote size 2x2x2 or nominal dimensions of 1x1x1.

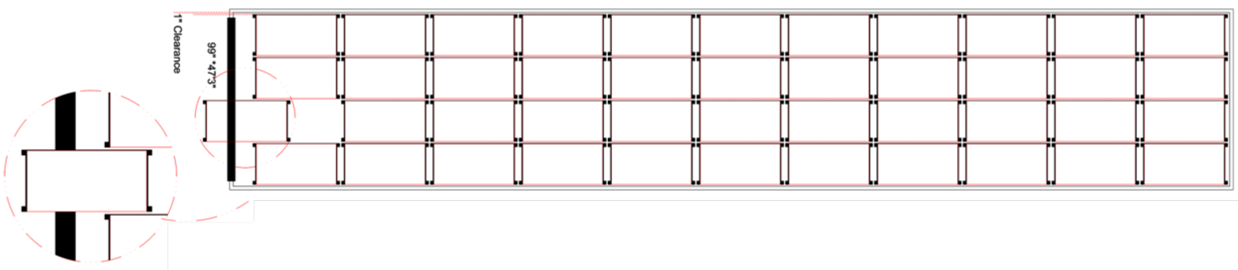


Figure 2: Rack Consolidation Inside a Trailer

The achieved racks have the dimension of 1’10-1/3”x4’x8’4-3/4” and the totes with a maximum size of 1’9”x1’9”x1’9” (Figure 4).

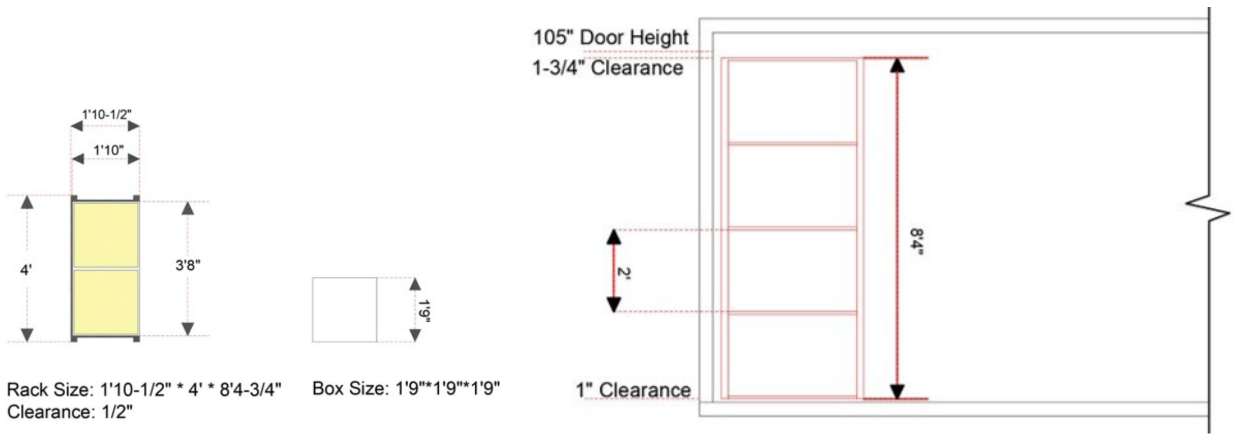


Figure 4: Dimensions of Racks and Totes

2.1.2 ShuffleCell

The shuffling process inside the HLH operates by *SuffleBot* inside the *ShuffleCell*. The process times and space requirements to move the robot arms for the shuffling are estimated through emulation using the Emul8™ software. The emulation results also indicate the optimized number of racks inside the cells. Optimized path planning for *MoveBot*, minimizing the robot's disruption in entering ,and exiting the cell and intersections are considered in all the blocks. As a result, the cell has two separate paths for *MoveBot* and is symmetrically designed to ease configuration and reconfiguration of the layout (Figure 5).

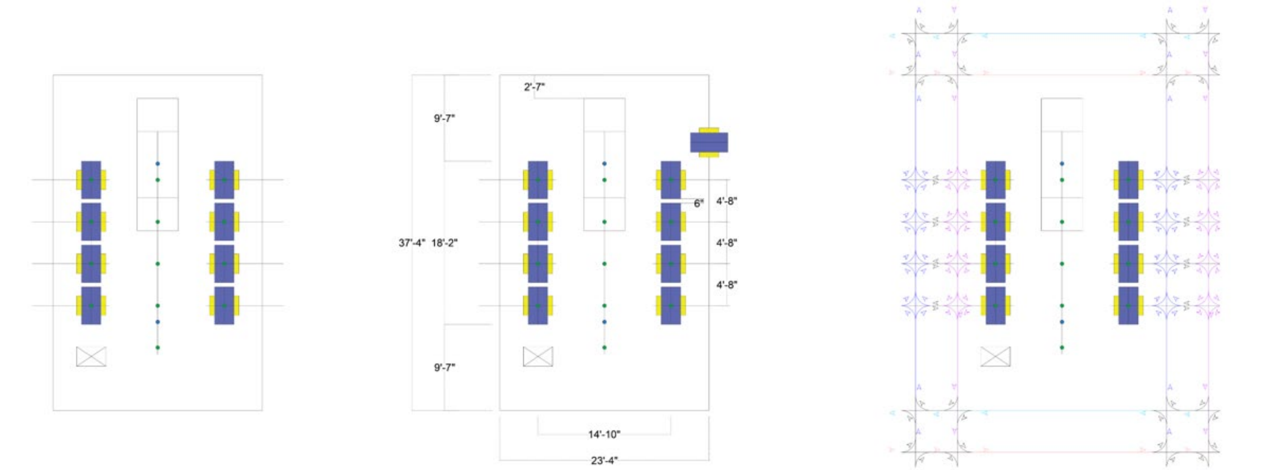


Figure 5: ShuffleCell Where Modular Containers are Shuffled from a Rack to Another

2.1.3 StagingCell

The first step to designing the *StagingCell* is the number of racks (36 racks). This number, along with the required aisle depth to move *LoadBoat* and *MoveBoat*, define the cell area. The orientation of racks inside the trucks specifies which side of the racks is grabbed by *LoadBoat* in the *StagingCell* ,and the dimension of *LoadBoat* with moving rack determines the aisle depth (Figure 6).

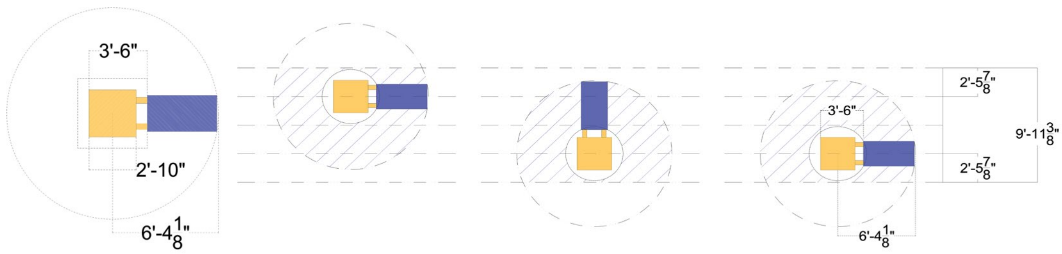


Figure 6: LoadBot and Aisle Dimensions

Figure 7 shows a standard size block for side-to-side rack orientation in the *StagingCell* with the required space for the aisle. Both *LoadBot*, and *MoveBot* uses the paths for unloading the racks and *MoveBoat* for moving the racks to the *Shuffle/BufferCells*.

2.1.4 BufferCell

The *BufferCells* are for the racks waiting to move among the *ShuffleCells*; no specific operation happens in the *BufferCells*. These cells are usually close to the *ShuffleCells* to enable efficient access. The number of these cells is also calculated by the simulation model for the operation's efficiency. *ShuffleCell*, and *BufferCell* are located in the center of HLH. They are designed to be in the same overall size with *SuffleCells* to be movable and exchangeable during the operation (Figure 8).

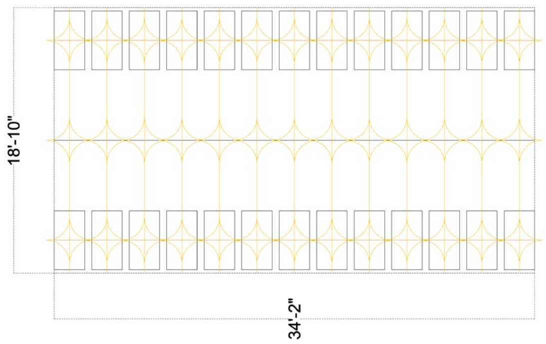


Figure 7: Modularly Dimensioned StagingCell for Mobile Racks after Truck Unloading and Before Loading

2.2 Equipment Mobility

HLH design process emphasis is on the robotic technology and the dynamics of the layouts. The detailed dimensions and behavior of the robots directly impact the design of the blocks. As it is mentioned before, the four different robots are assigned to execute the operations' tasks;

1. *LoadBots* – These robots unload/load racks from the trucks
2. *MoveBots* – These robots transport racks inside the hub between different zones.
3. *ShuffleBots* – These robots shuffle totes between racks in the *ShuffleCells.*
4. *Totebots* – These robots transport totes between the *ShuffleCells*.

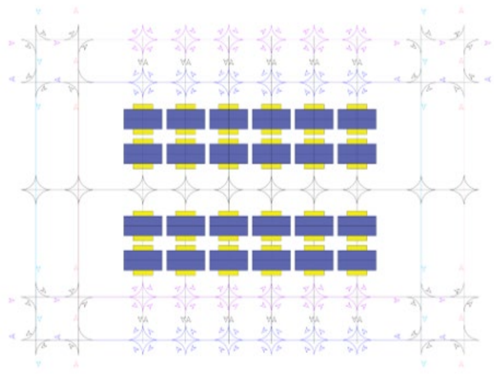


Figure 8: BufferCell

In response to the mobility of the system, the racks should be compatible with robotic technology. The mobile racks are designed with retractable legs to provide space under the rack dedicated for robotic movement. The legs open when *LoadBot* is placing the racks in the *StagingCells* to provide space for *MoveBot*. The underlying space required for *MoveBots*, path planning inside/outside the cells, and the intersections, are illustrated in Figure 9.

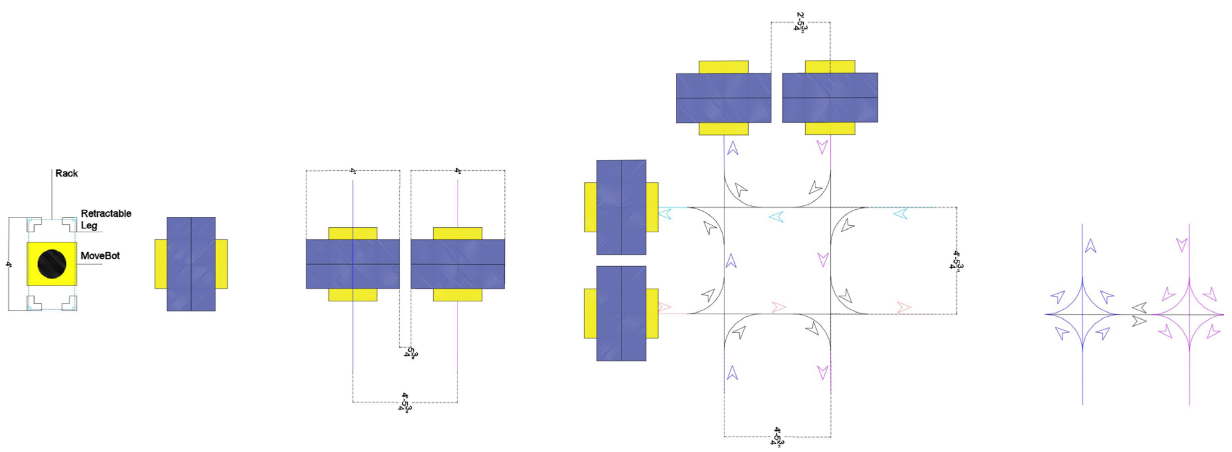


Figure 9: Standard Mobile Racks and MoveBot

Here we include some schematics highlighting some of the transformative impacts. The M^2 hub works exclusively with modular handling containers of tote and box sizes, where the totes can be handled by a human in his/her arms while the boxes are bigger than most pallet or cage sizes. Modular totes are stored and carried in mobile racks as done in goods-to-person systems such as those of Kiva and GreyOrange. Figure 10 depicts mobile racks being loaded in a truck, emphasizing the nice spatial fit of the racks into the trucks.

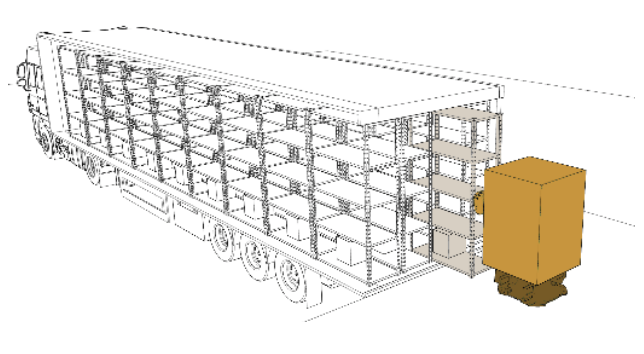


Figure 10: Loading/Unloading of the Racks-DockCell

Figure 11 shows *MoveBots* moving the mobile racks inside the *BufferCells*, it is worth noting that cells are implemented symmetrically and designed to minimize the *MoveBot* disruption considering the clearance between two loaded robots move side-by-side.

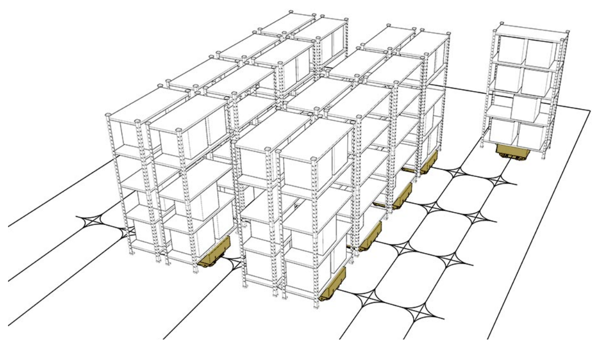


Figure 11: BufferCell with MoveBots

Figure 12 illustrates a *ShuffleBot* in a shuffling cell where totes are moved from their current mobile rack into one where totes sharing the same target next destinations and similar departure times. All constituents of shuffling cells are mobile, and the cells are themselves configured so as to fit within the modular space grid.

Figure 1**3** illustrates a *StagingCell*, where mobile racks are parked after having been unloaded from a truck while waiting for starting processing in *ShuffleCells*. Such cells are dimensioned to fit the modular space grid. Buffer cells have similar functions, yet for accomodating mobile racks between their processing in successive shuffling cells as pertinent.

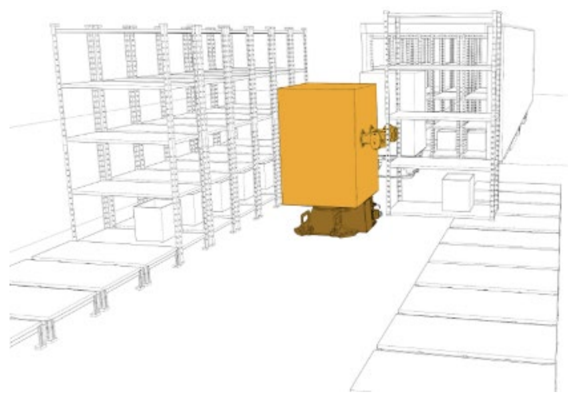


Figure 13: StagingCell for Mobile Racks after Truck Unloading and Before Truck Loading

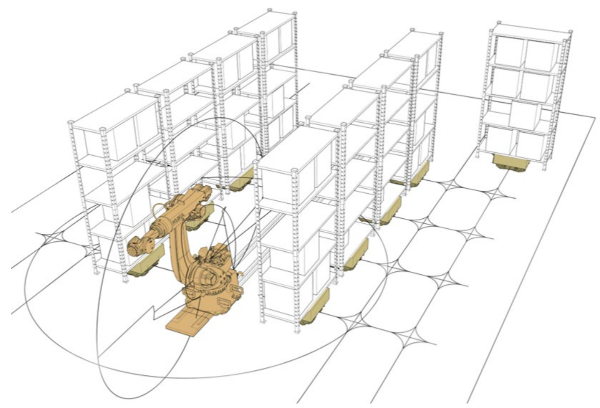


Figure 12: ShuffleBot in ShuffleCell Another

3 Layout Design

The standard size cells design eventually affects the allocation and the overall layout performance. Figure 14 provides a M^2 hub snapshot, emphasizing the spatial modularity of both the cells and the flow paths (strictly virtual, as all moving entities are not bounded by physical tracks). The cells

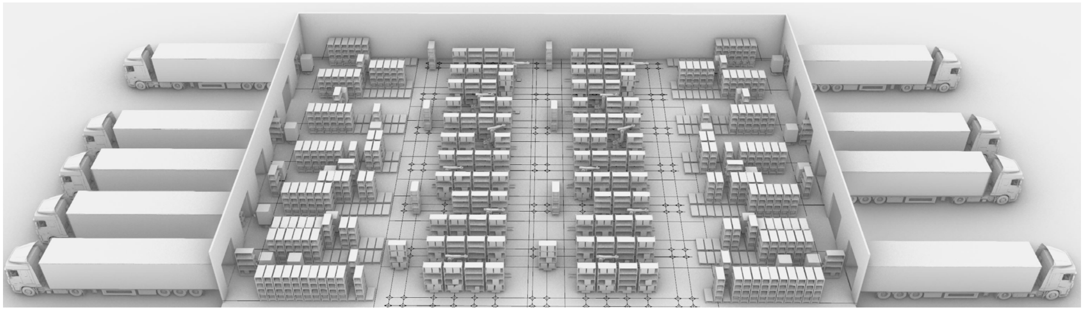


Figure 14: Hub Layout Based on a Modular Space Grid Depicting a Current Deployment of Shuffle, Buffer, and Staging Cells

voluntarily look similar, as they have been designed to be modular, yet each one has a specific mandate for a given time window, then are deactivated or moved to another grid module as pertinent to optimize workflow patterns.

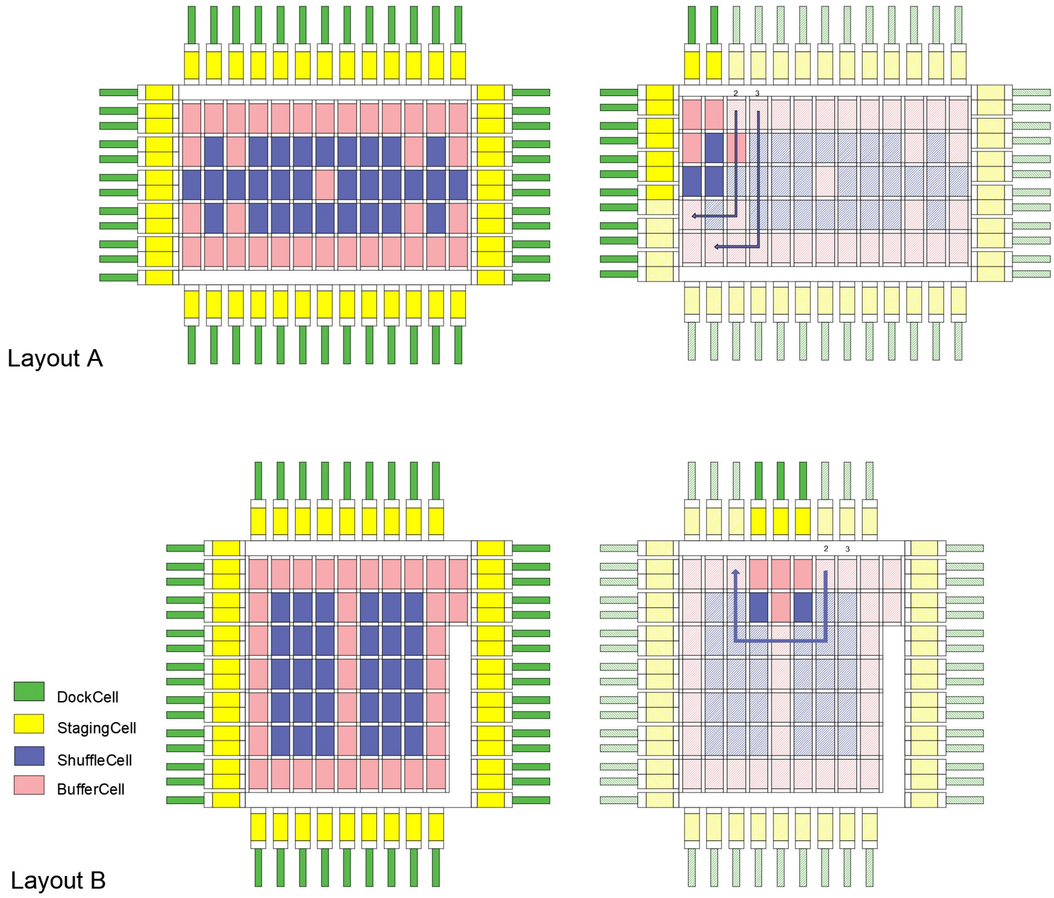


Figure 15: HLH Layouts Depicted with Dynamic Variations of Activated and Disactivated Cells

4 Simulation Experiment

This paper leverages a discrete-event simulation model in AnyLogic™ to assess the potentials results achievable by mobile and modular cell design. The experiment subjects the HLH hub to an average daily demand of 430,701 parcels encapsulated in 71,227 totes loaded on 9350 mobile racks arriving in 622 trucks and departing in 953 trucks. The parcels in a tote share the same target departure time from the hub and at least the next destination hub. In the experiment, the service requirements are such that totes have a maximum dwell time ranging from 2 to 4 hours in the hub. Demand is varying extensively through the day, with a ratio of 18.02-to-1 between maximum and minimum hourly rates, with a large peak toward the end of the day. The HLH hub is expected to meet this demand, accounting for the totes' max dwell times, with a service level of 100%.

The exploratory experiment constrats two alternative scenarios corresponding to distinct HLH hub designs, each considering that cells can be activated and disactivated dynamicall through a day. As shown in Figure 15, layout A is more elongated and puts the buffer cells mostly on the contours while layout layout B is more square and spreads more the buffer cells. This said, both layouts have the same number of activable cells: 50 StagingCells, 30 ShuffleCells, 35 BufferCells. Here, both designs impose a fixed location for each cell, a constraint that will be relaxed in further research. The designs can exploit a maximum of 250 MoveBots, 100 LoadBots, 30 ShuffleBots, and 24 ToteBots. Average speeds are 4 m/s for *MoveBots*, 30 s to move per rack for *LoadBots*, 4 s to take per tote and 5s to trasfer per tote for *ShuffleBots*, and *4 m/s* for *ToteBots*.

In this short paper, we provide empirical simulation-based evidence with key performance indicators (KPIs) on the significant performance gains enabled by M2 hub design.

The results shown in the section correspond to the calibrated version of each design, capable of meeting the service level requirements, while minimizing resource utilization (cells and robots). So KPIs are defined and measured to compare resource efficiency and utilization.

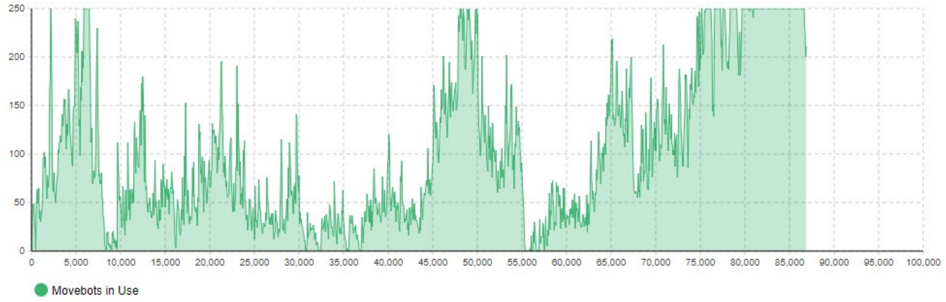
MoveBots population of 250	LayoutA	Layout B
Movement time (Minuts)	148887.4	130390.9

Table 1: Contrasting Robotic Utilization

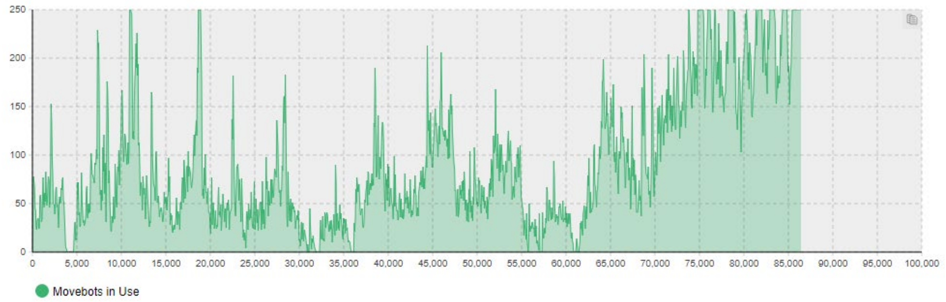
The first three KPIs are the daily robotic movement time, distance travelled, and utilization rate, computed for MoveBots, LoadBots, ShuffleBots, and ToteBots, with results summarized in Table 1. Overall, layout B outperforms layout A.

The fourth KPI is the number of robots and cells of each type concurrently used in the hub. Figure 16 plots this KPI for both designs over a day, focused on the number of *MoveBots* concurrently used to transport racks throughout the hub. Design A induces two boulders of *MoveBot* transport early and mid day, which design B succeeds to avoid, smoothing more the load of *MoveBots* with less 200-250 peaks through the day. Figure 17 similarly plots the number of *ShuffleCells* concurrently in use over time. It exhibits the same induced two-boulder creation in design A, again smoothed out in in design B. Overall, this contributed to design B having better resource utilization than design A, even though both leverage the full available capacity in the latter part of the day.

As discussed in Montreuil et al. (2021), the hub piloting architecture and algorithms also have impact on such performance, so future research is needed, challenged to further reduce these peaks while satisfying service level targets, notably exploiting the specific hub organization and layout, smart dis-activation of cells to minimize inter-cell flows.

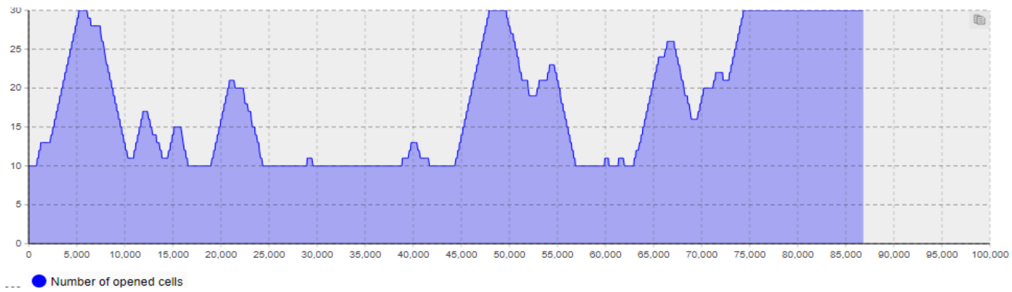


Layout A

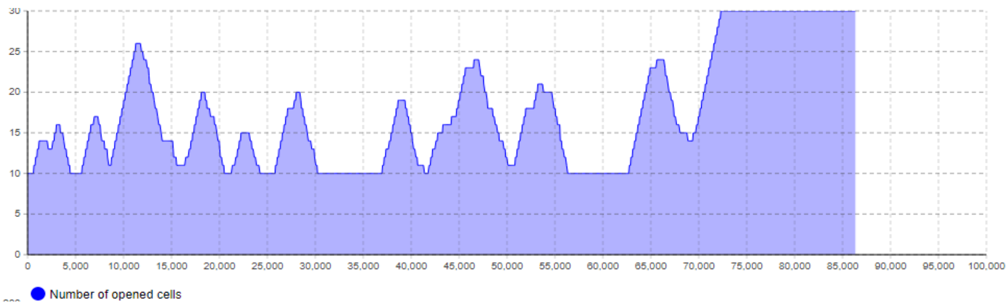


Layout B

Figure 16: Contrasting Movebot Utilization Over a Day



Layout A



Layout B

Figure 17: Contrasting ShuffleCell Utilization Over a Day

5 Conclusion and Future Design

In this paper, we have focused on introducing the modular and mobile (M^2) design of hyperconnected parcel logistics hubs that have been introduced in Montreuil et al. (2021) for the Physical Internet. We have provided an explicit fine-granularity design of such a hub, explaining and visually demonstrating each type of cells, as well as the types of robots used in its fully robotized version. We have put the emphasis on highlighting the modularity and the induced functionalities and capabilities. Given the space constraints of such a short paper, we have selected to provide empirical simulation results that contrasts two alternative designs of M^2 hubs, showing that (1) in the given instance, both succeed to satisfactorily achieve high service performance at 100% with short maximum dwell times at hubs, while doing so in a compact overall space and limited number of cells and robots; and (2) the designs differ in terms of resource utilization, with better peroformance by the design distributing more the cells to leverage the dynamic dis-activation of modular cells so as to minimize flows.

Avenues for further research include broadening the scope of performance criteria and KPIs and shaping more extensive simulation based investigation of M^2 hub capabilities and performance. This requires on one side to develop adequate decision architecture and algorithms so as, for example, better smooth utilization to minimize peaks, smartly decide on dynamic tote and rack assignment to cells, smartly decide on dynamic cell dis-activation, smartly move bots and racks to allow dynamic reconfiguration of the cells throughout the hub. It requires on the other side to extend the simulation modeling capabilities and experiment design to investigate deeper alternative M2 hub configurations and contrasting them with hubs not leveraging hyperconnectivity, modularity and mobility.

References

[1] Montreuil, B., R. D. Meller & E. Ballot (2010). Towards a Physical Internet: the impact on logistics facilities and material handling systems design and innovation, in *Progress in Material Handling Research 2010*, Edited by K. Gue et al., MHIA, p. 305-328.

[2] Ballot, E., B. Montreuil & C. Thivierge (2014). Functional Design of Physical Internet Facilities: A Road-Rail Hub, in *Progress in Material Handling Research Vol. 12*, Ed. B. Montreuil, A. Carrano, K. Gue, R. de Koster, M. Ogle & J. Smith, MHI, Charlotte, NC, U.S.A., p. 28-61.

[3] Meller, R.D., B. Montreuil, C. Thivierge & Z. Montreuil (2014). Functional Design of Physical Internet Facilities: A Road-Based Transit Center, in *Progress in Material Handling Research Vol. 12*, Ed. B. Montreuil, A. Carrano, K. Gue, R. de Koster, M. Ogle & J. Smith, MHI, Charlotte, NC, U.S.A., 347-378.

[4] Montreuil, B., R.D. Meller, C. Thivierge, C., and Z. Montreuil (2014). Functional Design of Physical Internet Facilities: A Unimodal Road-Based Crossdocking Hub, in *Progress in Material Handling Research Vol. 12*, Ed. B. Montreuil, A. Carrano, K. Gue, R. de Koster, M. Ogle, J. Smith, MHI, 379-431.

[5] Montreuil B., S. Buckley, L. Faugere, R. Khir & S. Derhami (2018). Urban Parcel Logistics Hub and Network Design: The Impact of Modularity and Hyperconnectivity, *Progress in Material Handling Research: 2018*, Ed. A. Carrano et al., MHI, Charlotte, NC, U.S.A.

[6] Montreuil B. , L. McGinnis, S. Buckley, S. Babalou, W. Bao & A. Beranji (2021). Modular Design of Hyperconnected Parcel Logistics Hub. Proceedings of IPIC 2021 Conference, submitted for publication.

[7] Balakrishnan, J., & Cheng, C. H. (2005). Dynamic cellular manufacturing under multiperiod planning horizons. *Journal of manufacturing technology management*, Vol. 16 No. 5, pp. 516-530.

[8] Balakrishnan, J., & Cheng, C. H. (2005). Dynamic cellular manufacturing under multiperiod planning horizons. *Journal of manufacturing technology management*, Vol. 16 No. 5, pp. 516-530.

[9] Montreuil, B. (2006). Facilities Network Design: A Recursive Modular Protomodel Based Approach, *Progress in Material Handling Research : 2006*, ed. by R. Meller & al., MHIA, p. 287-315.

[10] Montreuil B., 2007, Layout and location of facilities, *Handbook on Logistics Engineering, Chapter 9*, ed. by G. Donald Taylor, CRC Press, U.S.A.

[11] Marcotte, S. & B. Montreuil (2016). Introducing the Concept of Hyperconnected Mobile Production, *14^{th} IMHRC Proceedings (Karlsruhe,Germany–2016)*.16.

Predictive Demand Modeling for New Services in Hyperconnected Urban Parcel Logistics

Zeynab Bahrami-Bidoni and Benoit Montreuil
Physical Internet Center
H. Milton Stewart School of Industrial & Systems Engineering,
Georgia Institute of Technology, Atlanta, GA
Corresponding author: zbahrami@gatech.edu

Abstract: *The rapid growth of demand and fierce competition are encouraging logistics service providers towards expanding their competency in terms of offering novel and faster services and reinventing their logistics system so as to profitably and sustainably gain market shares. However, analyzing customer behavior and the underlying causes of demand variability for new services are complex tasks. This paper is dealing with customer behavior modeling for a service provider who wants to extend its offering system to much faster delivery service than ever done before. To adjust its logistic capacities with future demand, it needs to estimate the volume and geographical distribution of demand for new offered services. By capturing customers' sensitivities to the delivery-time observed in historical sales data and geo-categorization of orders in different time factors, a scenario-based demand generation methodology and tool are introduced for generating a wide range of demand scenarios with probabilistic patterns for customer behavior over all service offers with dynamic pricing. These are used to feed a simulator which models large-scale urban logistics networks service and offerings. In an application, it enables testing service capability improvements achievable by leveraging Physical Internet aligned transformation in a Chinese megacity.*

Conference Topic(s): *Interconnected freight transport, logistics and supply networks; PI Modelling and Simulation; Systems and technologies for interconnected Logistics (3D printing/ Internet of things/ machine learning/ augmented reality/ big data/ artificial intelligence/ blockchain/ cloud computing, digital twins, collaborative decision making); Last mile & City logistics.*

Keywords: *Logistics Service Provider; Scenario-Based Demand Analytics; Customer Behavior Modeling; Physical Internet; Delivery-Time Sensitivity; Urban Parcel Logistics; Last-Mile Delivery Network; Hyperconnected Logistics; New Service; Artificial Intelligence.*

1 Introduction

In fast-developing industries with increasing global competition in the introduction of new and innovative services, proper assignment of marketing and operational resources, and best use the available data to explain the demand and sales dynamics of a new service are crucial in success and profitability of new offerings. The research leading to this paper stems from a large-scale collaborative industry-university project (see Campos et al., 2021), notably its work package dealing with demand prediction and customer behavior modeling. The partnering company is a leading Chinese parcel logistic service provider (LSP) that is facing rapid market growth and endeavors to improve efficiency and extend service offers to the less than few hours intra-city delivery services. Such fast delivery is way beyond their current market offers. Therefore, for future demand prediction, it is required to model potential/current customer behaviors and preferences to estimate what portion of current customers may shift to order with new faster services, how many potential customers are to join to consider this company because of its new offers, and how customers are to react if their promised offers are not satisfactorily fulfilled. In this paper, we propose approaches to model customer behavior for new urban parcel logistic services. A key target application of these approaches relates to the comprehensive scenario-based probabilistic modeling of demand and customer behavior for parcel logistics, as introduced by Bahrami-Bidoni & Montreuil (2021). Such modeling notably allows to generate demand to feed the urban parcel logistics simulator developed through the project and to test the simulation performance of alternative logistic systems and offerings, and to monitor the level of customer satisfaction for multiple scenarios (Kaboudvand et al., 2021).

The paper is structured as follows. Section 2 provides a brief literature review. Section 3 focuses on historical demand and behavior data analytics. Section 4 tackles demand and customer behavior modeling for new services. Section 5 introduces a creative methodology inspired by generalized Bass model to analytics of price-based demand response. Section 6 presents a proposed scenario-based user-interactive application and its uses. Finally, section 7 concludes the paper through a synthesis of the contribution, limitations, and avenues for further research.

2 Literature review

The detailed taxonomy illustrated in Figure 1 provides a comprehensive application-oriented overview of the existing literature on demand and customer data analytics. Analytics is known as the scientific process of transforming data into insights for making better decisions, and at a conceptual perspective level, it is commonly divided into four stages (Wang et al., 2018): (1) descriptive data summarization and visualization for exploratory purposes, (2) explanatory diagnostic models that estimate relationships between variables and allow for hypothesis testing, (3) predictive models that enable forecasts of variables of interest and simulation of the effect of marketing control settings, and (4) prescriptive optimization models that are used to determine optimal levels of control variables. Some of the techniques and methodologies adopted or developed to address each application have been summarized in Figure 1. Moreover, Figure 1 shows that the size and degree of structure in data increases from right to left, and the feasibility of the higher level of analysis decreases as a function of big data dimensions. It illustrates that the information value of the data grows as its volume, variety, and velocity increase, but that the decision value derived from analytical methods increases at the expense of increased model complexity and computational cost. Wang et al. (2018) and Chicco G. (2016) in their review papers provided a critical examination of analytics methods, with a marketing focus, by tracing their historical development, examining their applications to structured and unstructured data generated within or external to a firm, and reviewing their potential to support marketing decisions.

The application of smart meter data analytics on customer behavior (Wedel & Kannan (2016)) are particularly revealing in regard to our paper due to its emphasis on customer behavior. As summarized in Hong & Fan (2016), there are three ways to modify the workflow to generate probabilistic forecasts: 1) generating multiple input scenarios to feed to a point forecasting model; 2) applying probabilistic forecasting models, and 3) augmenting point outputs to probabilistic outputs by imposing simulated or modeled residuals or making ensembles of point forecasts. The scenario generation method was also used to develop a probabilistic view of power distribution system reliability indices (Black et al. (2018)).

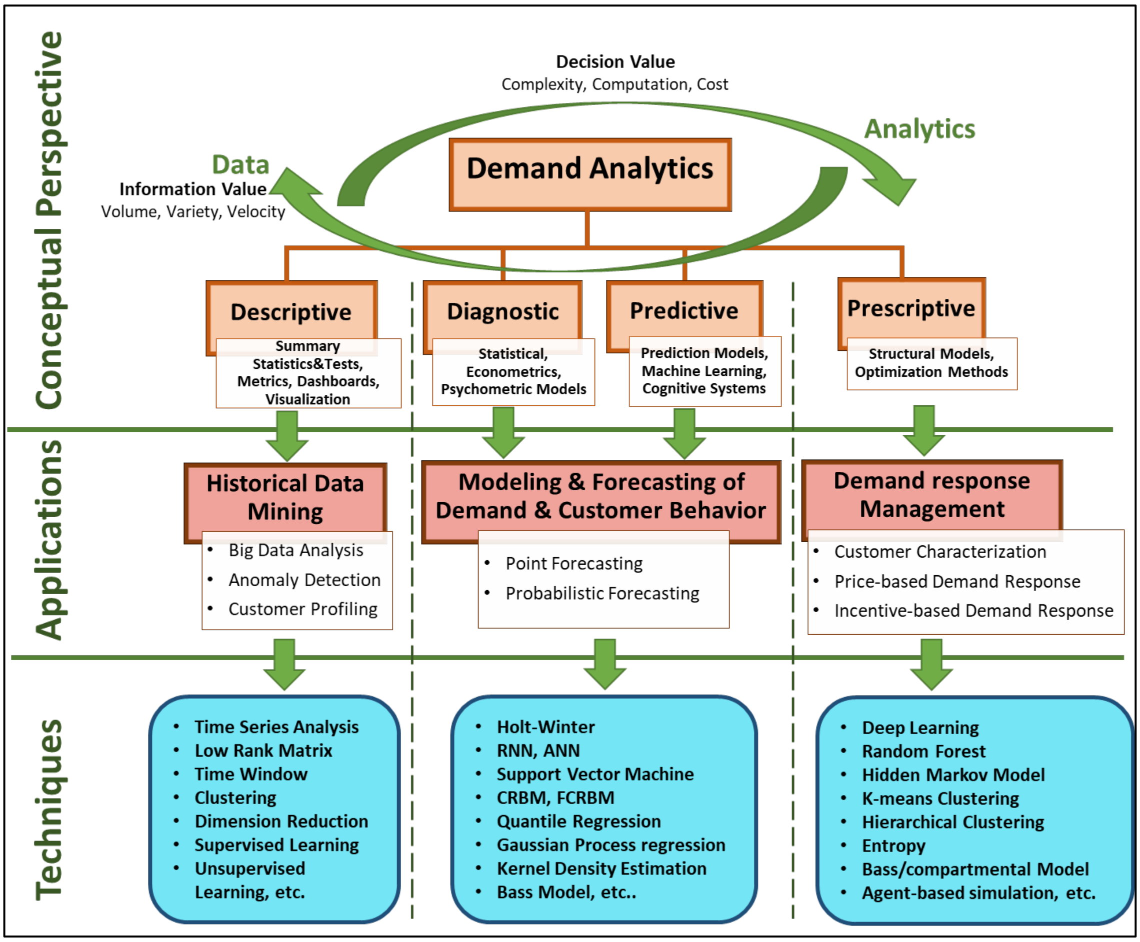


Figure 1: Taxonomy of literature review over predictive demand and customer data analytics

The Bass model was first to forecast demand for new products which have been widely used in new product forecasting mainly because of its concise conceptual framework, parameter setting and good explanatory ability (Bass, 1969). It was defined using a differential equation in which purchases are initiated by mass communication and further driven by word of mouth (WOM) from past purchases. Bass model and its various extensions (e.g. Ramírez-Hassan & Montoya-Blandón, 2020) typically explain the customer purchase process at a highly aggregate level of a market which does not allow for capturing customer heterogeneity. Moreover, they assume that all customers can communicate with all the others (perfect mixing), an assumption often violated in observed network of customers. Consequently, their estimates are highly biased and hard to relate to customer behavior (Abedi, 2019). Bass model and its extensions require the accuracy and completeness of sales data to ensure prediction accuracy, and this type of models is also difficult to adapt to the complex and changeable market environment.

Following the Bass model and its variations to model adoption of a new product, a variety of methods have been proposed: agent-based simulation models (e.g. Rand & Rust, 2011; Kiesling et al., 2012); equation-based models, particularly, compartmental modeling (e.g. Rahmandad & Sterman, 2008; Hariharan et al., 2015). These have in common modeling the adoption pattern at the level of each compartment. Compartmental models add spatial and psychographic segments to Bass-type models, allowing modelers to approximate the heterogeneity, clustering, and communication network of customers, while retaining an analytical structure. These models have been mainly explored in the context of demand forecasting and describing customer behavior, but their suitability as a scalable modeling tool to support large scale marketing and/or operational decision making is not explored.

Abedi (2019) proposed a flexible compartmental model and assessed its suitability in terms of use of data, adherence to micro-level customer behavior, and use in large scale decision making. The paper reveals that compartmental models result in estimates that are less biased and can predict the shape of the adoption curve significantly better than the Bass model. A hybrid new product demand forecasting model proposed by Yin et al. (2020) combines clustering analysis, using a fuzzy clustering-rough set method, and deep learning, using a long short-term memory (LSTM) neural network model. First, they make a primary prediction based on classical Bass model, where fuzzy clustering and rough set methods are used to obtain the attributes of new product through the historical sales data analysis of related products. Then, they correct the prediction error by a LSTM model to form robust predictive results. Temporal sales pattern are often dynamically influenced by word of mouth, previous experience, and loyalty as well as by marketing/advertising and distribution support (Abedi, 2019). Demand estimates for new products in the existing methodologies mostly rely on adoption patterns through customers' network connections and products' similarities. Until now, existing prediction models for new services have rarely been fully validated in terms of quality and consistency. Yet the need for such models is significant in parcel logistic systems, as making such predictions is part of critical tasks faced by LSP managers on a regular basis.

To address this problem, we constructed three connected models hereafter presented. First is a model that predicts total customers who consider the LSP to order (customers who come to the system-centers or website- and receiving offers). This is a scenario-based mathematical model that has been proposed for hourly long-term demand, which consists of a nested combination of three subsections using 2-D representations for modeling specific events and the hourly variation within a week. Second is a model of customer behavior and sensitivities to the delivery time. This model, for any scenario assumption, will determine the probability of selecting among offers by different types of customers. Third is a model providing probabilistic parameter values of the above two models in different scenarios. In our reported research experimentation, the first two models were implemented inside the parcel routing simulator. For the third model, we provided an AI-based App with a user interface to compute the value for input parameters of above two models based on the specific assumed scenario. Then we fed values of these parameters to the comprehensive parcel routing simulation model (in CSV files) and investigated the impact of customer preferences on demand. During simulation runs, we could see how the source and destination of orders is generated in the area and how the customers will select among offers. We also tested the performance of simulated logistic system for that specific scenario. In the rest of paper, four stages of analytics in this study have been presented.

3 Historical data analytics

Predicting demand faces many information challenges. For instance, estimating potential intracity demand for parcel logistics in an overall megacity market depends on a lot of factors such as population and microeconomic growth factors to name a few. Most challenging is the lack of knowledge about potential customers who approach diverse competitors and those who reject all offers. In most cases, the only direct information one can get comes from the historical sales data of one or a few LSPs and is about those customers who approached and ordered from a given LSP. Rarely are logged rejected offers at sale time, nor the fact that the customer considered several LSPs.

3.1 Big data analysis and anomaly detection

The growth of customers in this sector has resulted in the use of big data analytics to understand customers' behavior in predicting the demand of items. It uses a complex process of examining large amount of data to uncover hidden patterns in the information. It is established on the basis of finding correlation between various parameters that are recorded, understanding purchase patterns and applying statistical measures on collected data. For this purpose, a big real-world Chinese megacity database has been studied as a bench test including one-year waybills and Barcode Scanning streaming (in terabytes volume) as well as some multivariate data fusion such as Latitude/longitude geographical information about hubs and customers, Socio-demographic info and so on which all have been used for profiling different type of customers and clustering them in terms of their preferences over services and their sensitivities on delivery-time. Cleaning this big data, detecting anomalies/outliers, and estimating Null/missed data was a challenging part of this project in the first steps.

3.2 Customer profiling

Customers' diversity can be modeled through geo-categorization. For instance, consider a city as a grid of small units such as in Figure 2. These units can for example represent zip codes. Based on the type of buildings, facilities or departments, it is possible to categorize customers who are living in each unit. Grid units can be categorized as entirely or partially be business, industrial, residential, etc. Thus, to model customer behavior we tried to categorize potential and loyal customer's preferences based on their order types and geographical-sociodemographic characteristics and probabilistically modeling their behaviors as presented in section 4.3 to see how they choose among offers (or reject them).

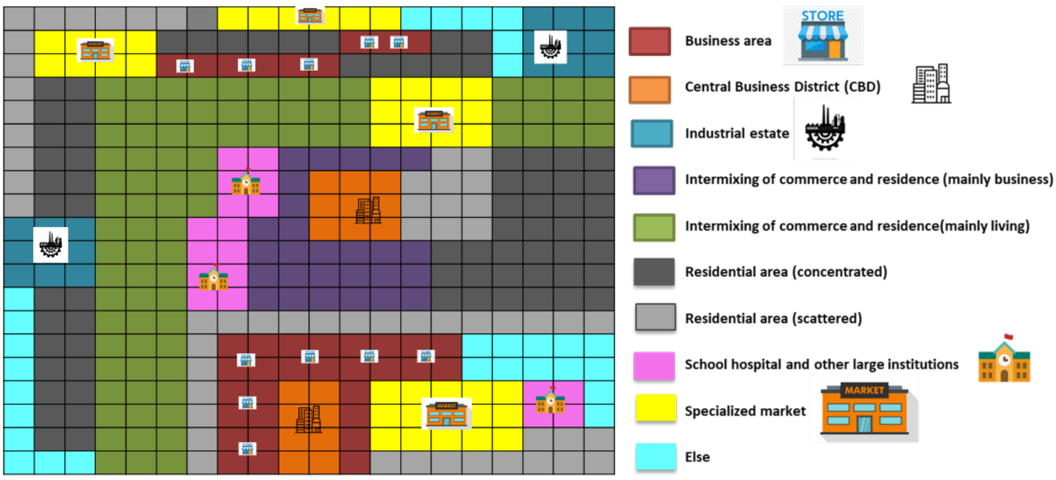


Figure 2: City as a grid of units with the corresponding type of area

4 Modeling demand and customer behavior for new services

4.1 General perspective of proposed framework

The proposed approach for implementing the demand and customer behavior model as well as feeding its outputs to the given simulated logistic system have been summarized in Figure 3. In this project, we have used an urban logistics simulator developed using the Anylogic tool to simulate the parcel routing logistics with offering service levels (Kaboudvand et al. 2021). Probabilistic modeling and analytics, based on multivariate data, customer profiles, and historical sales trend over past time intervals, enables generating probabilistic demand patterns, customer preferences, and price-based demand response patterns into the simulated virtual world which are explained in sections 4.2, 4.3, and 5. This sets the stage for the agents operating the logistic system in the simulator.

Generation of potential orders by the demand generator requires knowing the feasible service types and also estimating their costs to find the feasible domain area for the offers. As an LSP may have some predefined routing cycles in the logistic model, and an estimate of how long will take each trip between different logistics hubs during different times of day, it becomes possible to estimate feasible piecewise trajectory trips, and their duration between source and destination of an order. For doing this, the offering manager collaborates with the logistic model to get a real-time robust estimate on the shipping time of the parcel from pickup point to the delivery point through the system. Based on this information, the LSP is to filter offers by removing those that are infeasible with respect to the fastest possible delivery time. After LSPs provide the set of feasible service offers, the customer chooses between them or reject them all, in a way that can be estimated through the probabilistic model of customer behavior computed by models in section 4.3 and 5. If the customer's preference is not in the list of LSP offers, the customer will choose or not the next best fitting feasible offer, which can be modeled based on cumulative probability function of corresponding order category.

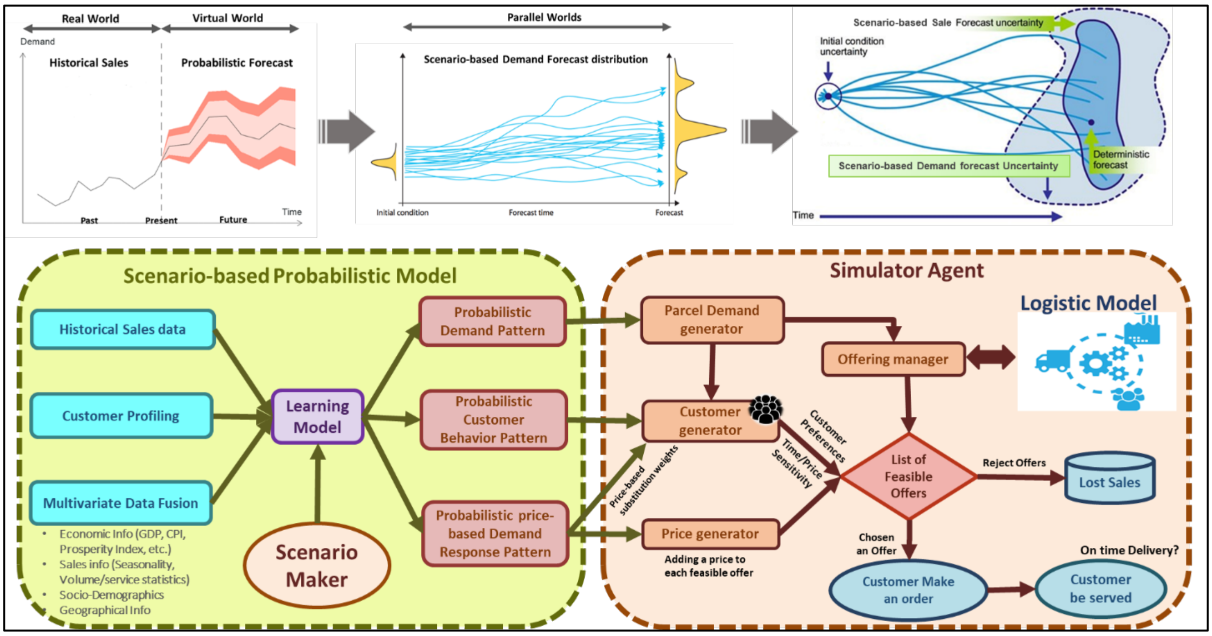


Figure 3: Scenario-based predictive modeling of uncertain demand

After enough simulation iterations over intervals and scenarios, this leads to adequately estimating the distribution of scenario-based forecast for potential demand, sales, and lost sales. The reference Kaboudvand et al. (2021) provides more details regarding how the simulator parameters are obtained from probabilistic customer behavior models as inputs, how they interact with each other, and how verification and validation of the simulator are conducted.

4.2 Modeling probabilistic pattern for scenario-based demand volume and geographical source/destination distribution

To generate the patterns for demand volume, we use a mathematical model inspired by the methodology proposed in Filik et al. (2011) which contains a nested combination of three subsections using weekly residual load variations and two-dimensional representations for modeling specific events and the hourly variations within a week. A data-driven learning and fitting surface functions to this template structure allow us getting the base pattern of demand volume with lower/upper bounds on the probabilistic range at the hourly granularity over a given forecast period. We have introduced and explained this proposed approach with more detail in Bahrami-Bidoni & Montreuil (2021).

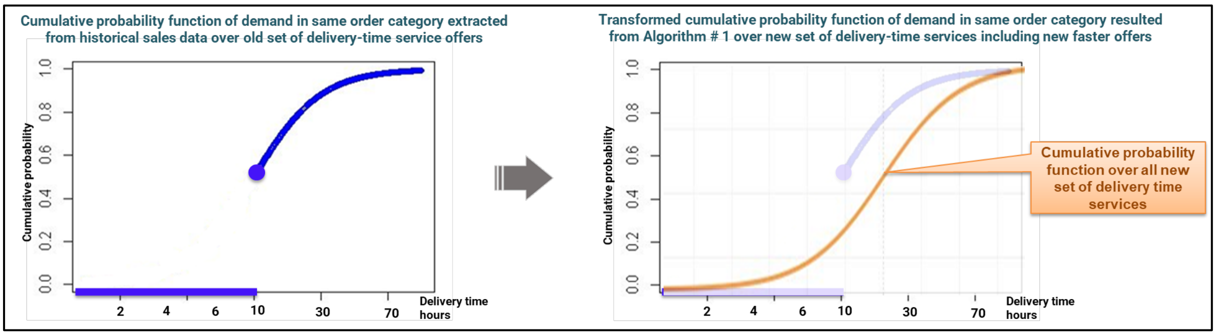


Figure 4: Scenario-based cumulative probability demand estimation for new faster services (example)

D: *Average daily demand forecasted by first model presented in section 4.2.*

P: *Total percentage of potential customers seeking new faster services (one of scenario assumption).*

$O^{Old}=\{t_1^{Old},t_2^{Old},\dots,t_s^{Old}\}$ *: The set of all promised delivery-times in the old service offering system.*

$O^{New}=\{t_1^{New},t_2^{New},\dots,t_r^{New}\}$ *: The set of all faster delivery-times in the new service offering system.*

$V^c=\left(v^c_{t_1^{Old}},v^c_{t_2^{Old}},\dots,v^c_{t_s^{Old}}\right)$, $\sum_{i=1}^{s} v^c_{t_i^{Old}}=1$ *: Probability vector of demand over the set of old offering services.*

Algorithm #1:

Goal: *Computing continues cumulative probability function of category c for selecting over all delivery-time services.*

Stage 1: Compute the vectors $K^c=\left(K_0^c,K_1^c,K_2^c,\dots,K_s^c\right)$,

where $K_0^c=P*D*w_{pd}^c$ and

$K_i=(1-P)*D*w_{pd}^c*v^c_{t_i^{Old}}$, $i=1,\dots,s$

Stage 2: Normalize K^c to get the vector $K^{c'}=\frac{1}{\sum_{i=0}^{s}K_i^c}K^c$

where its first component $K_1^{c'}$ is the cumulative probability of demand over all new offering services.

Stage 3: $a=K_1^{c'}\big/\left(t_r^{New}\right)^2$, $L=K_1^{c'}$

Do for $i=1:r$
$\hat{F}(t_i^{New})=a*(t_i^{New})^2$
End

and

Do for $i=1:s$
$L=L+K_{i+r}^{c'}$
$\hat{F}(t_i^{Old})=L$
End

Stage 4: *Fitting below function inspired by Bass diffusion Model* $F_{p,q}(t)=\left(1-\exp(-(p+q)t)\right)\big/\left(1+(q/p)\exp(-(p+q)t)\right)$ *and compute the optimal* p^*,q^* *coefficients with lower function error to the data sets* $T=\{t,\ t\in(O^{New}\cup O^{Old})\}$, $\hat{F}(T)=\{\hat{F}(t),\ t\in T\}$

Return $F_{p^*,q^*}(t)$ *As the cumulative probability function of category c for selecting over all delivery-time services.* $(t\in T=O^{New}\cup O^{Old})$

Figure 5: Algorithm for simulation of customer preference over all potential continuous delivery-time

4.3 Modeling probabilistic customer behavior pattern for demanding services

First, a category based on four attributes (*h, s, r, d*) has been considered for each potential order: *h* is the time window of placing order (which hour of a day), *s* and *r* are category types of sender and receiver of corresponding parcel (type means geo-categorization in section 3.2), and d is the categorical distance variable between order's source and destination (e.g. *d1< 10 km, 10 km<=d2<40 km, 40 km <=d3*). For instance, a potential order category could be a parcel from a business to a residential area with *d2* distance at 2 PM. Moreover, the proportional demand weight for faster services (w_{pd}^c) is computed for order category *c* by $w_{pd}^c=w_{TS}^c*w^c_{fd,t_1^{old}}$, where the delivery-time sensitivity weights (w_{TS}^c) obtained by equation (1).

$$w_{TS}^c=R^c(t_1^{old})\big/R^c(t_1^{old})+R^c(t_2^{old}) \tag{1}$$

where $R^c(t_1^{old})$ and $R^c(t_2^{old})$ are the average demand rates in category *c* for the first and second fastest delivery-time offered on historical sales data, and $w^c_{fd,t_1^{old}}$ is the fractional historical demand volume weight for the t_1^{old} in category *c* among all other categories.

The cumulative probability function over previously available offering set of delivery times will be extracted from historical sales data as a base (e.g. see left diagram of Figure 4). Then, based on assumptions in given scenario and using the algorithm in Figure 5, a new piecewise cumulative probability function will be simulated over all the new sets of delivery time services and finally lead to completing the probabilistic customer behavior model over a new set of service levels for a given scenario (e.g. see right diagram of Figure 4). In the first two steps of algorithm, the cumulative probability demand over all new offers is obtained and demand probability over old offer set calibrate based on value of *P* in scenario assumption and *d* from 4.2). In stage 3, a rough estimate for *F(t)* over new delivery-time services is computed and then the optimal value of *p* and *q* parameters will be obtained such that continues Bass model function fitted to the estimated discrete cumulative probability in corresponding order category.

5 Price/Incentive-based probabilistic demand response analytics

Research shows that customer satisfaction, advertisements, word of mouth (WOM), and incentive offers/plans have a positive effect on increasing the potential customers. Thus, modeling the customer sensitivities to the price changes will help to construct a dynamic pricing system that optimizes profits while increasing demand market share and keeping high the level of satisfaction. Here, the functional framework of a generalized BM, presented by Ramírez-Hassan & Montoya-Blandón (2020) incorporating market effort into the BM, have been used for modeling demand adoption diffusion rate due to internal/external influences over the delivery-time axis. *F(t)* and *f(t)* are the cumulative and non-cumulative proportions of demand at offered delivery-time t, and Y(t) is the total number of potential customers demanding faster services up to but not including offers with delivery time t. The coefficient (*p*) is the rate of spontaneous demand adoption, and the coefficient (*q*) is the rate of imitation of demand adoption that the optimal *p* and *q* parameters for any order category computed by the algorithm in Figure 5. *T* is the set of all available delivery-time services which offered.

$$\frac{f(t_i)}{1-F(t_i)} = (p+qY(t_i))x(t_i), \quad x(t_i) = 1+\alpha_1 \frac{P(t_i)-P(t_{i-1})}{P(t_{i-1})} + \alpha_2 \frac{\max\{0, A(t_i)-A(t_{i-1})\}}{A(t_{i-1})}, \ t_i \in T \qquad (2)$$

Where $x(t_i)$, $P(t_i)$, and $A(t_i)$ are respectively the market effort, the price and the advertising for the i^{th} service offer with t_i promised delivery-time. These variables enter the market effort equation as percentage increases. The sale on offered service with delivery-time t_i will be computed by $S(t_i) = F(t_i) - F(t_{i-1}) + e$, where *e* is an additive normally distributed error term with variance σ^2 and *F(t)* is given by equation (3).

$$F(t_i) = \frac{1-\exp\left\{-\bar{X}(t_i)(p+q)\right\}}{1+(q/p)\exp\left\{-\bar{X}(t_i)(p+q)\right\}}, \qquad \bar{X}(t_i) = t + \alpha_1 Ln\left(\frac{P(t_i)}{P^{\min}}\right) + \alpha_2 Ln\left(\frac{\tilde{A}(t_i)}{A(0)}\right), \qquad t_i \in T \qquad (3)$$

Where $\bar{X}(t_i)$ is the cumulative market effort, found by transforming equation (2) into continuous delivery-time and integrating from 0 to t_i, and $\tilde{A}(t_i)$ is the last value for which there was a positive change in advertising ($P^{\min}$ is the minimum price that was offered among available services). Moreover, the time of peak sales defined as the time of the highest diffusion rate *S(t)* can be calculated by using the equation $t^* = (\ln q - \ln p)/(p+q)$ which means that decreasing/increasing in price, advertisement, and any other market effort on offering a service with promised delivery-time by t^* would have highest impact to increase demand on the corresponding order category.

6 Proposed scenario-based user-interactive application

We have conceived and developed a scenario maker application in MATLAB allowing us to set scenario assumptions through an interactive user interface, and to run the model under these assumptions to generate scenario-based probabilistic distributions of customer behavior. This includes probability distribution of customer preferences on different available offers for all categories and substitution probability for each offer when is not feasible. Aggregated, this results in scenario-based probabilistic patterns for hourly total demand volume for intracity, inbound, and outbound flows, to be transposed into orders in a simulated logistics system.

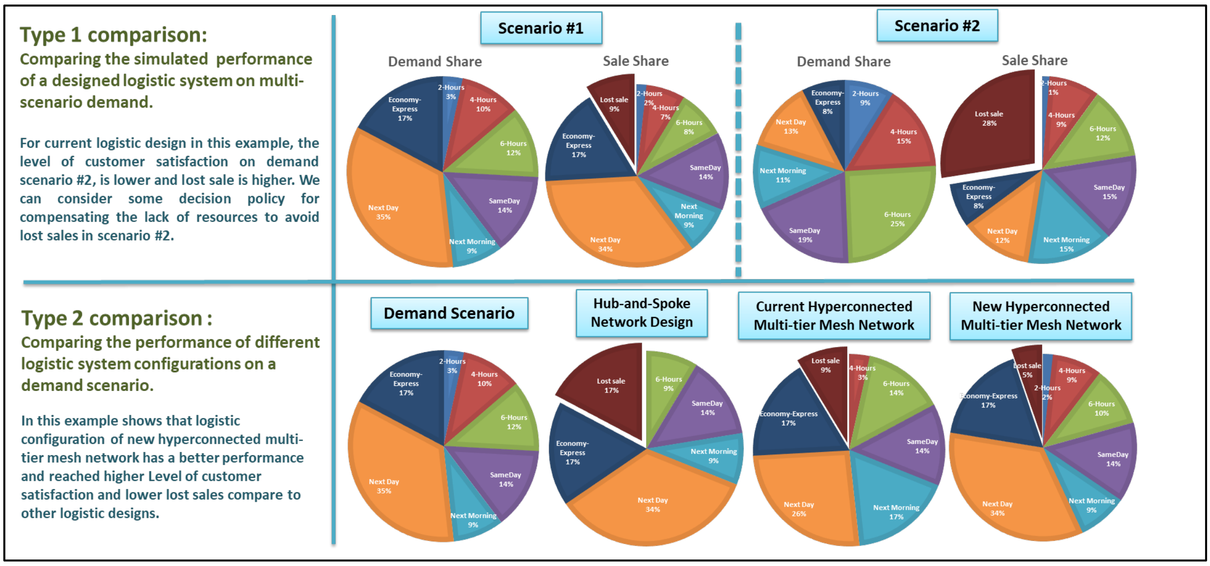


Figure 6: Beneficial outcomes of analysis over simulator results by feeding scenario-based probabilistic patterns for demand and customer behavior.

Using the scenario-based probabilistic models presented in sections 4.2, 4.3, and 5, and the probability distributions for source/destination of parcels on different time factors, enables to generate future scenario-based forecasted demand logs for the given time horizon which includes every single parcel request, as part of a demand scenario to be simulated. As depicted in Figure 3, alternative future logs based on alternative scenarios can be simulated in parallel worlds, and the results from these simulations combined to reveal and analyze logistics outcome and performance distributions. By using the demand and customer behavior models as drivers for the logistics simulator, it allows performing simulations jointly enabling to compare, beyond logistics costs and environmental impacts, the sales and customer satisfaction outcomes from different scenarios. Figure 6 provides examples of how feeding probabilistic patterns for demand and customer behavior helps holistically evaluating logistics system performance.

7 Conclusion

This study is a part of an industry-university project for a large urban parcel logistic system, focused on modeling and forecasting the intracity demand and within-city customer behavior to extend service offers for faster delivery. We proposed a scenario-based probabilistic modeling approach for demand generation to feed the parcel routing simulation model used for testing the performance of current/alternative logistic systems and monitoring the level of customer satisfaction. Moreover, a scenario-based application with an interactive user-interface has been described to make a wide range of various demand scenarios and to simulate customer preference on new services. This tool is resulting in scenario-based probabilistic patterns for demand and customer behavior which are feedable to the logistic simulator for testing performance and sales attributes as well as providing scenario-based forecasted demand logs for any arbitrary duration. Comparing multiple scenarios by analysis on demand logs is providing deep managerial insights on demand shape and its geographical distribution in terms of volume and service types, which is helpful to be prepared for future risks/challenges with suitable policies. Also, multi-scenario comparison by analysis on simulated sales after feeding probabilistic patterns leads to a better understanding of sale shape and its geographical distribution in terms of volume and service types and a better understanding about required capacities or lack of resources in different locations or different hours of the day which result to lost sales. Moreover, it helps comparing logistic performance for different hub/routing designs and the level of on-time delivery and customer satisfaction. Beyond the need for further empirical investigation of the paper's contribution, a key avenue for further research with high potential for model improvement is to explicitly account for customer satisfaction and its positive/negative impacts on future demand.

References

- Abedi, V. S. (2019): Compartmental diffusion modeling: Describing customer heterogeneity & communication network to support decisions for new product introductions. *Physica A: Statistical Mechanics and its Applications*, *536*, 120964.
- Bahrami-Bidoni Z., B. Montreuil (2021): Enabling Scientific Assessment of Large Scale Hyperconnected Urban Parcel Logistics: Scenario-based Demand and Customer Behavior Modeling. Proceedings of IISE 2021 Conference.
- Bass, F. M. (1969): A new product growth for model consumer durables. *Management science*, *15*(5), 215-227.
- Campos M., B. Montreuil, L. McGinnis, S. Kaboudvand, S. Kwon, Z. Bahrami-Bidoni, L. Faugere, S. Buckley (2021): Enabling Scientific Assessment of Large Scale Hyperconnected Urban Parcel Logistics: System Configuration and Assessment. Proceedings of IISE 2021.
- Chicco G. (2016): Customer behavior and data analytics. in 2016 International Conference and Exposition on Electrical and Power Engineering (EPE). IEEE, 2016, pp. 771–779.
- Filik, Ü. B., Gerek, Ö. N., Kurban, M. (2011): A novel modeling approach for hourly forecasting of long-term electric energy demand. *Energy Conversion and Management*, *52*(1), 199-211.
- Hariharan, V. G., Talukdar, D., Kwon, C. (2015): Optimal targeting of advertisement for new products with multiple consumer segments. *Internat. J. of research in marketing*, *32*(3), 263-271.
- Hong T., S. Fan (2016): Probabilistic electric load forecasting: A tutorial review. International Journal of Forecasting, vol. 32, no. 3, pp. 914–938, 2016.
- Kaboudvand S., M. Campos, B. Montreuil (2021): Enabling Scientific Assessment of Large Scale Hyperconnected Urban Parcel Logistics: Agent-Based Simulator Design", IISE 2021 Conference.
- Kiesling, E., Günther, M., Stummer, C., Wakolbinger, L. M. (2012): Agent-based simulation of innovation diffusion: a review. *Central European Journal of Operations Research*, *20*(2),183-230.
- Rahmandad H., Sterman J. (2008): Heterogeneity & network structure in the dynamics of diffusion: Comparing agent-based and differential equation models. *Management Science*, *54*(5), 998-1014.
- Rand W., R. T. Rust (2011): Agent-based modeling in marketing: Guidelines for rigor. *International Journal of research in Marketing*, *28*(3), 181-193.
- Ramírez-Hassan, A., Montoya-Blandón, S. (2020). Forecasting from others' experience: Bayesian estimation of the generalized Bass model. International Journal of Forecasting, 36(2), 442-465.
- Wang Y., Q. Chen, T. Hong, C. Kang (2018): Review of smart meter data analytics: Applications, methodologies, and challenges," IEEE Transactions on Smart Grid, v10, no3, 3125–3148.
- Wedel M., P. Kannan (2016): Marketing analytics for data-rich environments," Journal of Marketing, vol. 80, no. 6, pp. 97–121, 2016.
- Yin, P., Dou, G., Lin, X., & Liu, L. (2020): A hybrid method for forecasting new product sales based on fuzzy clustering and deep learning. *Kybernetes*.

Smart Track 4 Waterway – Regulatory Analysis

Gaëlle Bonjour and Valérie Bailly-Hascoet

IPIC 2021
8th International Physical Internet Conference
June 14-17, 2021
Virtual

IDIT Interreg North-West Europe ST4W

INTRODUCTION

Smart Track 4 Waterway is a European project (Interreg NWE) whose main objective is to promote the modal transfer from the road to the river by monitoring the loading from the first to the last mile, enabled by the hierarchical tracking of packaged goods.

ST4W objectives are:

- Modal shift facilitation from road to waterway thanks to a seamless synchronisation between actors,
- Enabled by a cloud-oriented management system for shipment by inland waterway
- Fed with data from hierarchical traceability of shipment including its geo-localization.

To achieve these objectives the project required a regulatory analysis of the Inland Waterway Logistics processes and participants to determine if there are current hurdles to facilitating either the desired seamless synchronization between actors and systems. And, if so, to define them.

OBJECTIVES

The objectives of the Regulatory Analysis (undertaken by IDIT) are:

- Analysis of the inland waterway transport regulatory framework for the project (Electronic Data Regulations, Reporting Requirements/Customs, Regulations for Road Transport and Inland Waterway Transport).
- Stakeholder mapping for the different types of data collected and transferred during operations.
- Analysis of the governance structure for Electronic Data and identifying legal challenges for innovative tracking and tracing technologies.
- Recommendations to overcome trust issues and promote the uptake of electronic documents while maintaining security of data.

METHODOLOGY

IDIT interrogated the legal and regulatory hurdles to waterway transport adoption via 4 key tasks:

1. Definition of legal frameworks by theme. For each theme, the relevant International/ European, national and regional regulations were identified to allow further analysis of the legal impacts on the development of waterway transport. Consultation with the project partners determined the regulations that had the largest impact on facilitating modal shift.
2. Identification of the various stakeholder relationships involved in the required data collection and exchange processes via two common transport scenarios. Interviews with the stakeholders provided feedback on practical data sharing issues.
3. Analysis of Electronic Data Governance for the four different types of data (Personal data, Commercial data , Goods data and Vehicle data) identifying the legal and practical issues for each.
4. Consultation with a working group of stakeholders to determine conclusions and recommendations on the key issues identified via the above tasks.

RESULTS

1. The regulations that have the greatest impact on the efficiency of inland waterway transport are:
   - Electronic Data Regulations
     - Personal Data & Privacy
     - Commercial Data
     - Electronic Commerce
     - Transport Info Systems
     - Cyber Security
   - Reporting Formalities
     - Single Maritime Window
     - Customs
2. The desired/required innovations were found to be hampered by regulations as follows:

The development/uptake of....	Is hampered/limited by
Geo-tracking of Vehicles	- General Data Protection Regulations - Personal privacy protections in Inland Waterway Transport - Rules on Geo-tracking in road transport
Electronic Transport Documents	- Lack of current obligation to accept e-docs - Fragmented national regulations
Automated Data Sharing in Transport Information Systems	- Risk of Cyber Security and - Risk of Commercial Espionnage

DISCUSSION

This project confirmed that the number of interconnected and inter-reliant regulations and rules for multimodal transport (including inland waterway transport) are too numerous to be practical both for research purposes as well as for operations.

Sharing data through the ST4W application poses certain legal and business issues; Regarding privacy, liability and commercial sensitivity.

Two major problems emerge from the analysis of the stakeholder interviews:

- a lack of confidence in the sharing of personal data and ;
- an administrative burden linked to the lack of interoperability between systems.

FUTURE WORK

The Interreg NWE program has granted (Capitalisation Call) the ST4W an extension until 09/2022 to demonstrate how the cloud-oriented management solution can link to sustainable last-mile delivery solutions in 5 pilot sites. With a focus on data and responsibility transfer and the delivery interfaces, IDIT will compare legal and urban planning regulations for the pilot sites and recommend improvements for last mile operations using IWT.

It is also recommended that blockchain or DLT be investigated as a potential solution for secure, controlled and compliant data storage and sharing amongst stakeholders in these complex intermodal corridors.

ACKNOWLEDGEMENTS

IDIT would like to acknowledge the help and support of our partners in the ST4W project: Multitel, CIRCOE, Logistics in Wallonia, Technische Universiteit Eindhoven, Stichting Bureau Telematica Binnenvaart, Stichting Projecten Binnenvaart, Port Autonome du Centre et de l'Ouest, Port de Bruxelles - Haven van Brussel, Universität Duisburg-Essen (UDE), Blue Line Logistics nv, Inlecom System Ltd
https://www.nweurope.eu/projects/project-search/st4w-smart-tracking-data-network-for-shipment-by-inland-waterway/#tab-1

Spatio-temporal Arrival Prediction over Hyperconnected Logistics Networks

Yujia Xu, Yao Xie, Benoit Montreuil[1]

IPIC 2021
8th International Physical Internet Conference
June 14-17, 2021
Virtual

1 Georgia Tech Supply Chain and Logistics Institute
Interdisciplinary Research Center
Physical Internet Center
Supply Chain & Logistics Institute

Introduction

- Arrival prediction is a vital component in supply chain and logistics.
- Planning and operational decisions depend on predictions.
- Hyperconnected logistics enable a new opportunity for prediction by capturing interaction and correlation between different locations and over time in the network.
- Arrivals at one location may have a non-homogeneous influence on future arrivals at other nearby locations.

To capture the temporal dependence of past events, we aim to
- Introduce a simple arrival distribution prediction approach.
- Propose a novel method to model and predict arrival events from spatio-temporal sequential data based on a spatio-temporal interactive Bernoulli process, which can capture the spatio-temporal correlations and interactions without assuming time-decaying influence.
- Make arrival predictions for any locations at any future time.

Methodology

1. Arrival distribution prediction

Given a path, assume the travelling time X_i on segment i (from hubs to hubs) is independent random variable that is normally distributed,

$$X_i \sim N(\mu_{X_i}, \sigma_{X_i}^2)$$

Then the total travelling time Y of n segments is also normally distributed,

$$Y = \sum_{i=1}^{n} X_i$$

$$Y \sim N(\sum_{i=1}^{n} \mu_{X_i}, \sum_{i=1}^{n} \sigma_{X_i}^2)$$

2. Spatio-temporal interactive Bernoulli process model

- We aim at building a parametric probabilistic model on spatial-temporal arriving events given specified memory depth. We assume that for $t \geq 1$, the conditional probability of $\omega_{tk} = 1$, is specified as

$$\mathbb{P}[\omega_{tk} = 1 | \boldsymbol{\omega}_{t-1-d}^{t-1}] = \beta_k + \sum_{s=1}^{d} \sum_{l=1}^{K} \beta_{kl}^{s}\, \omega_{(t-s)l}, 1 \leq k \leq K.$$

- Then we elaborate on the Least-Square (LS) and the Maximum Likelihood (ML) method for parameter estimation. For example, using LS method, we can recover the collection of parameters by minimizing the sum-of-square errors between conditional probabilities and observed arrival events:

$$\Psi_{\omega^N}(\hat{\beta}) = \frac{1}{2N} \sum_{t=1}^{d} \sum_{k=1}^{K} ((\beta_k + \sum_{s=1}^{d} \sum_{l=1}^{K} \beta_{kl}^{s}\, \omega_{(t-s)l}) - \omega_{tk})^2$$

- As stated in **Juditsky et al. (2020)**[1], the problem can be solved efficiently in polynomial time as a constrained convex optimization problem. Then the recovered collection of parameters β can be used to make predictions for future arrivals for any locations at any future time.

T	Time horizon (total number of hours)
K	Total number of locations
x_{tk}	Arrival observation at time t (hourly) and location k
ω_{tk}	$\begin{cases} = 1, \text{ if containers arrives at time t (hourly) and location } k \\ = 0, \text{otherwise} \end{cases}$
$\boldsymbol{\omega}_{t-d}^{t}$	$= \{\omega_{sk}, t-d \leq s \leq t, 1 \leq k \leq K\}$, a set of observations given memory depth d.
β_k	Birthrate, denotes the probability of arrival at k-th location.
β_{kl}^{s}	The magnitude of interactive influence on location k and time t of the event occurred at time $t-s$ and location l.

Table of Notations

Reference: [1] Juditsky, A., Nemirovski, A., Xie, L. and Xie, Y., 2020. Convex recovery of marked spatio-temporal point processes. arXiv preprint arXiv:2003.12935.

Real-data Simulation

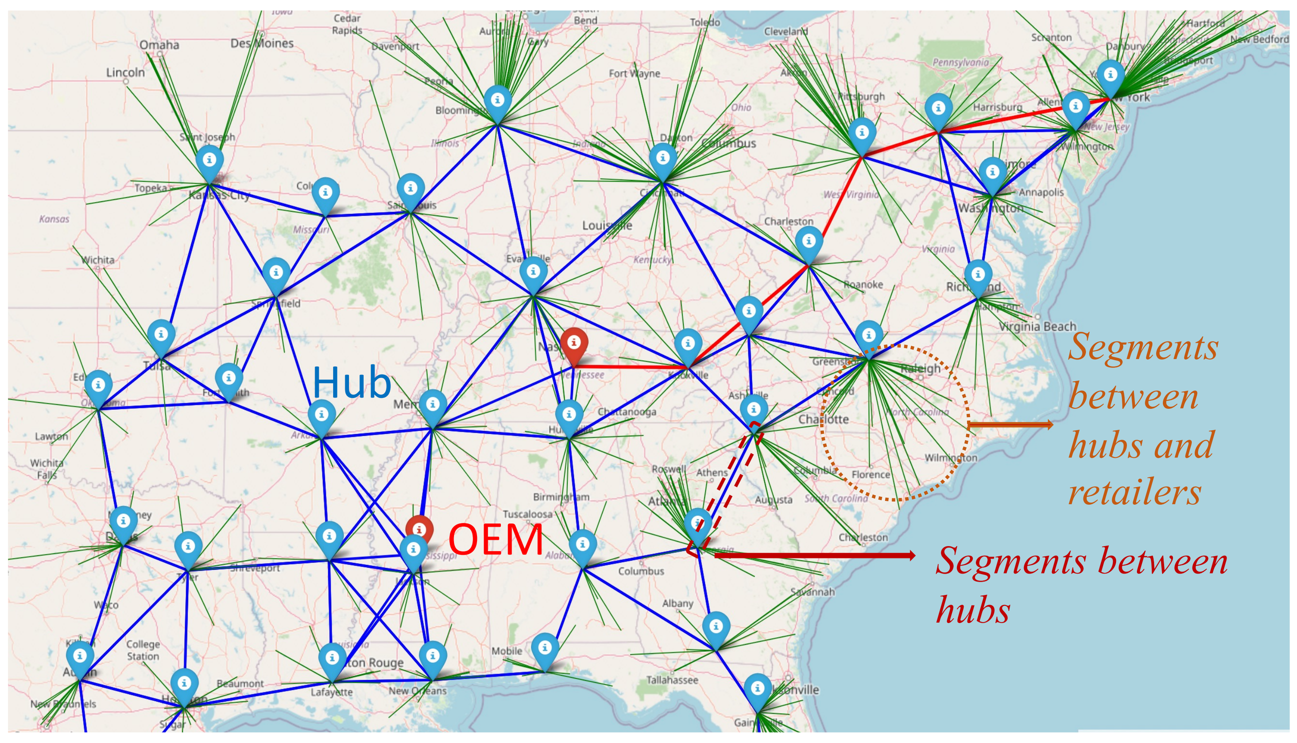


Figure 1

The simulation is created to generate the arrival events based on:
1. Demand data from OEMs
2. A brand new hyperconnected network designed by the Physical Internet center

Containers are delivered from OEMs to retailers via retailers' nearest hub by shortest path. Travelling time on segments is stochastically generated based on historical knowledge. The simulation is used to enact testbed scenario instances.

Results

1. Arrival distribution prediction

We generate 10000 containers travelling on the path marked in the red lines in Figure 1, which has 5 segments. By assuming the travelling time on each segment is normally distributed, the total travelling time is also normal, as shown in Figure 2.

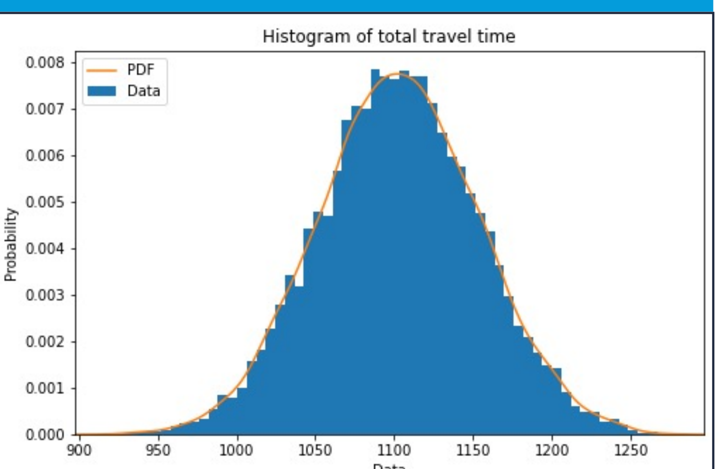


Figure 2

2. Spatio-temporal interactive Bernoulli process model

We use 360-day training data and 360-day testing data. Given discrete t (hours), space indices k (hubs), and memory depth $d = 24$ hours, we recover the collection of parameters β through observations ω^N from the simulation using the LS and ML methods, as shown in Figure 3(a), (b) and (c). The problems is solved within a couple of minutes. Then we predict the arrivals using the conditional probability with recovered parameters and compare the frequency of arrivals to the hubs with those in the testing data. As shown in Figure 3(d), the predicted frequencies calculated by the two methods are very close to the testing frequencies. It is also possible to update predicted arrivals dynamically based on real-time data with the precalculated parameters.

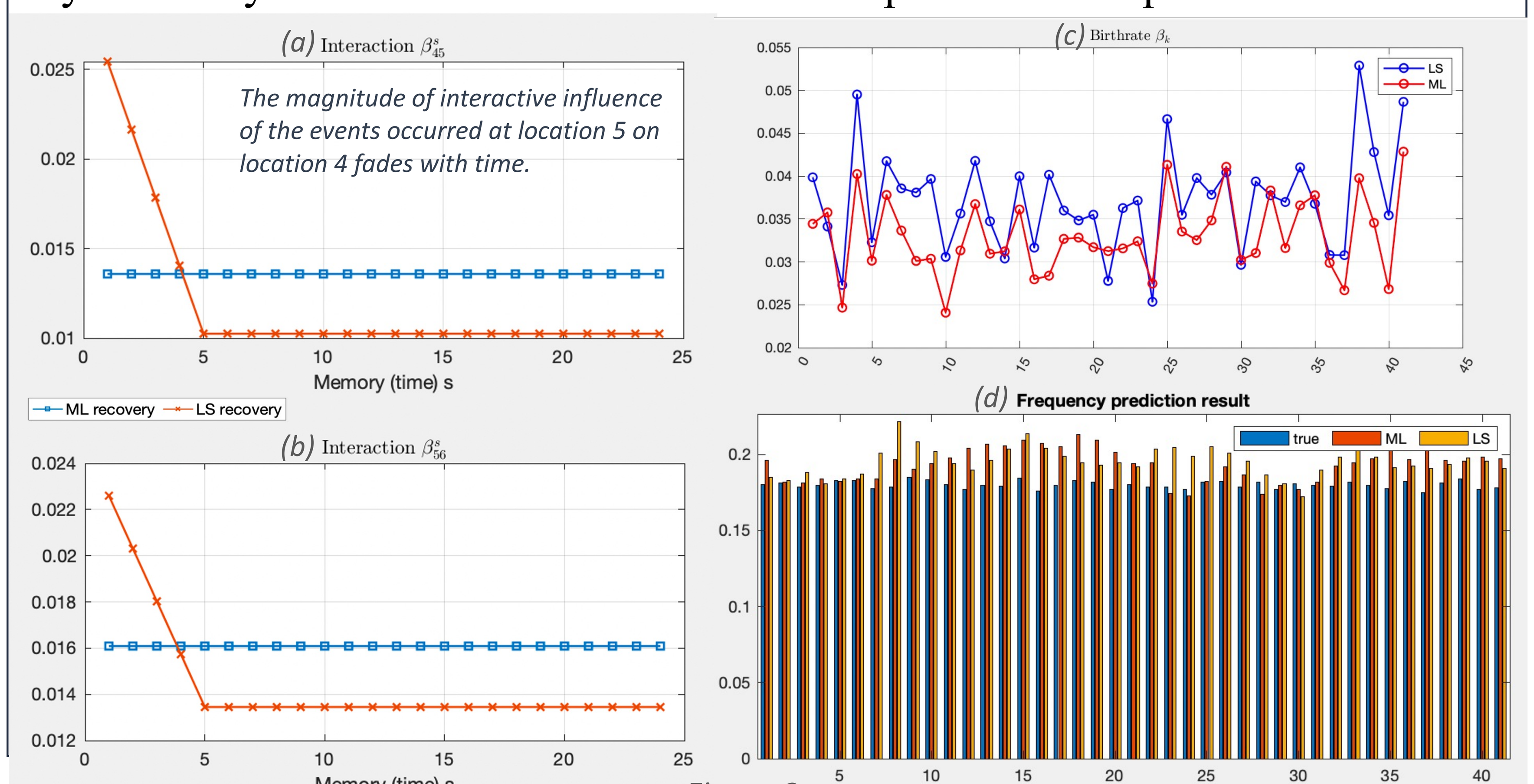


Figure 3

Conclusion

We propose two novel methods to model and predict the spatio-temporal arrival events over hyperconnected logistics networks and demonstrate the excellent performance of our approaches in a simulated environment. Future research is needed on enhancing predictive capabilities leveraging live environmental and traffic information.

Dynamic Consolidation and Packing of Parcels in Modular Physical Internet Handling Containers

Nidhima Grover, Benoit Montreuil

IPIC 2021
8th International Physical Internet Conference
June 14-17, 2021
Virtual

Georgia Tech Physical Internet Center
Supply Chain & Logistics Institute

Motivation

Problems:
1) Emptiness in trucks: They are 40-60% full
2) Non standardization in unit loads
3) Unreliable service levels

Objectives:
1) Select appropriate sized containers
2) Ensure timely departure of parcels
3) Improve packing in vehicles and containers

Modular Containers

- Standardized Physical Internet containers
- Uniquely identifiable, snap-interlock design, eco-friendly
- Three tiers: Packaging Handling, Transport
- We focus on consolidation inside handling containers

Image: Handling container encapsulated in transport container

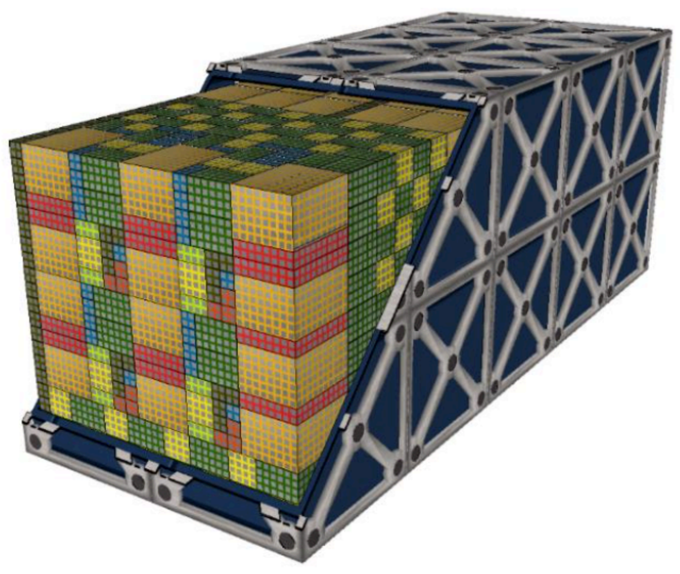


Image source: Montreuil, Benoit; Ballot, Eric; and Tremblay, William, "Modular Design of Physical Internet Transport, Handling and Packaging Containers" (2014). *13th IMHRC Proceedings (Cincinnati, Ohio. USA – 2014)*. 1

Dynamic Containerized Parcel Consolidation

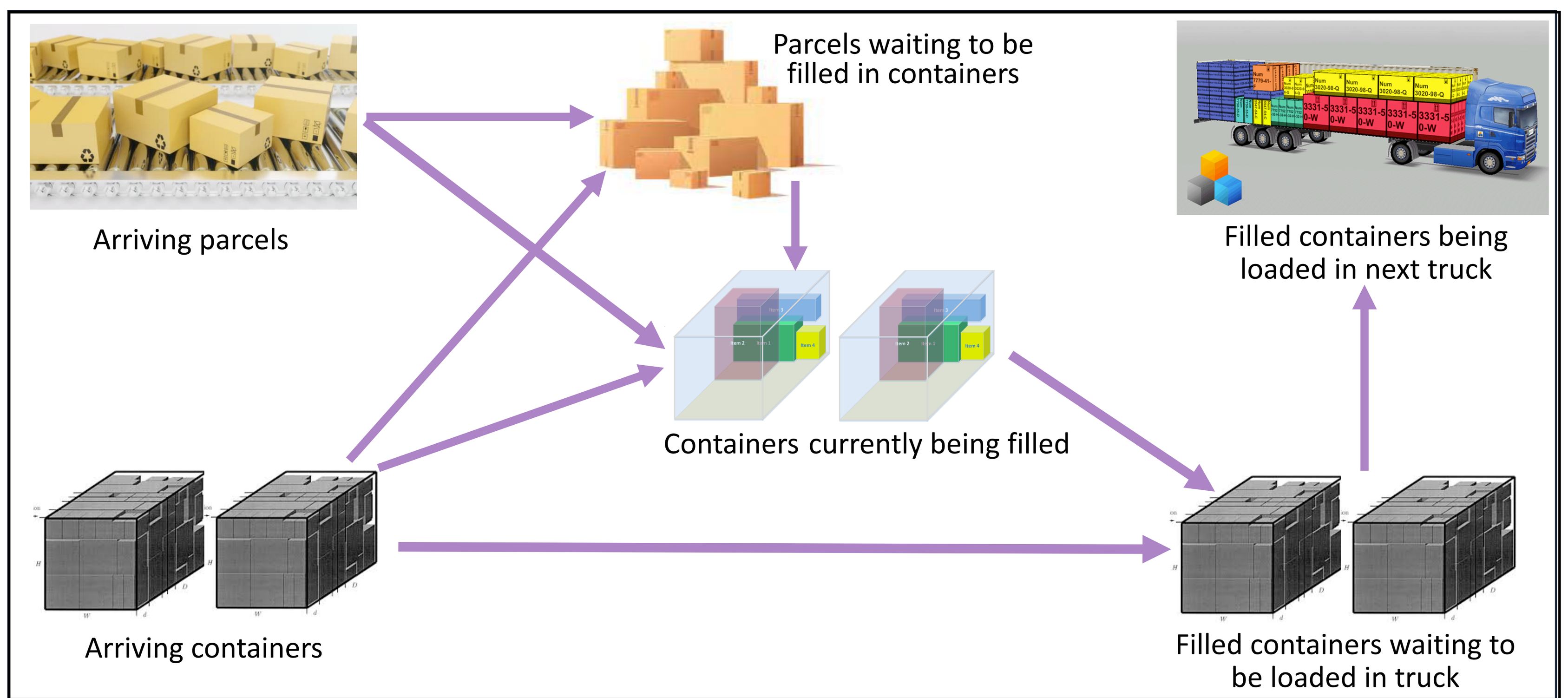
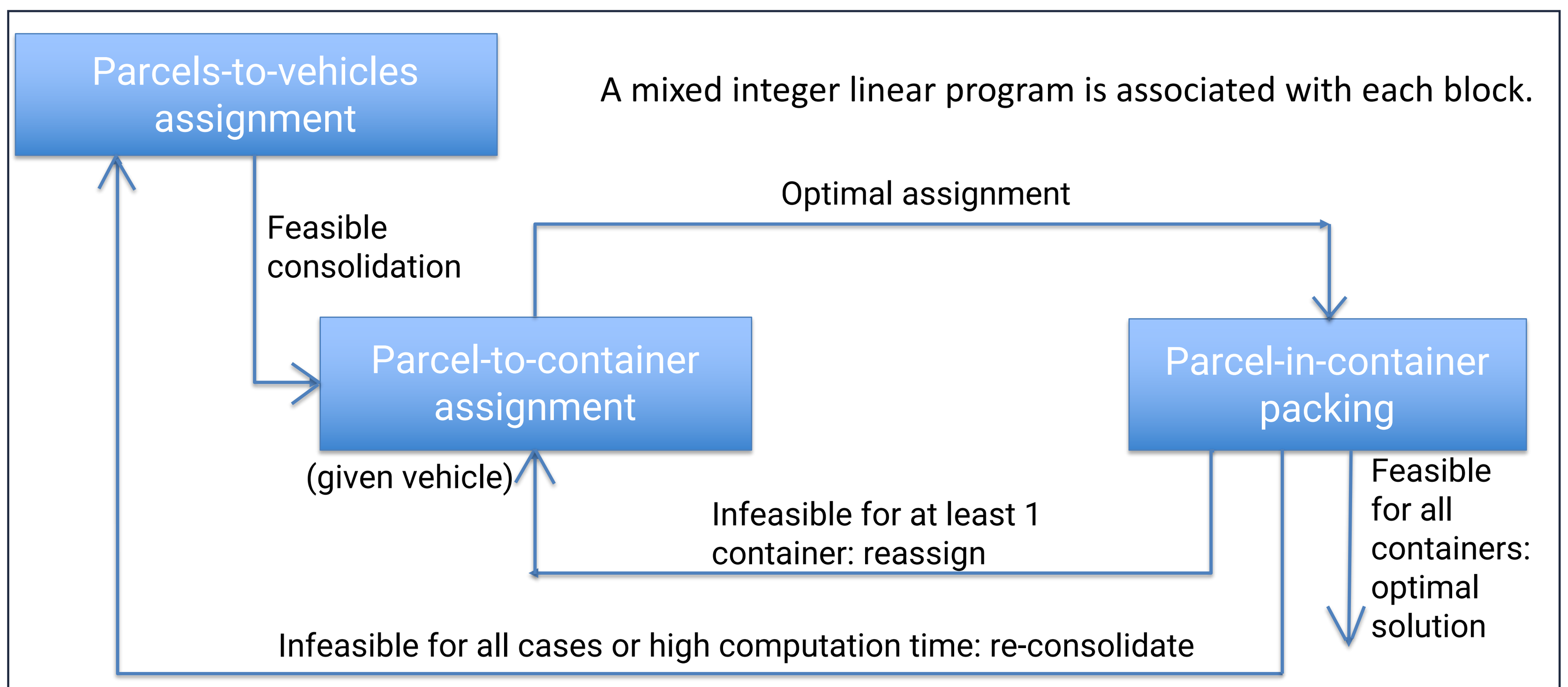


Image sources: Alamy.com, gettyimages.ae; vectorstock.com; Ha, Chi Trung, et al. "An online packing heuristic for the three-dimensional container loading problem in dynamic environments and the Physical Internet." *European Conference on the Applications of Evolutionary Computation*. Springer, Cham, 2017.; Pisinger, David. "Heuristics for the container loading problem." *European journal of operational research* 141.2 (2002): 382-392.

Method



Parcel-to-vehicle Assignment Results

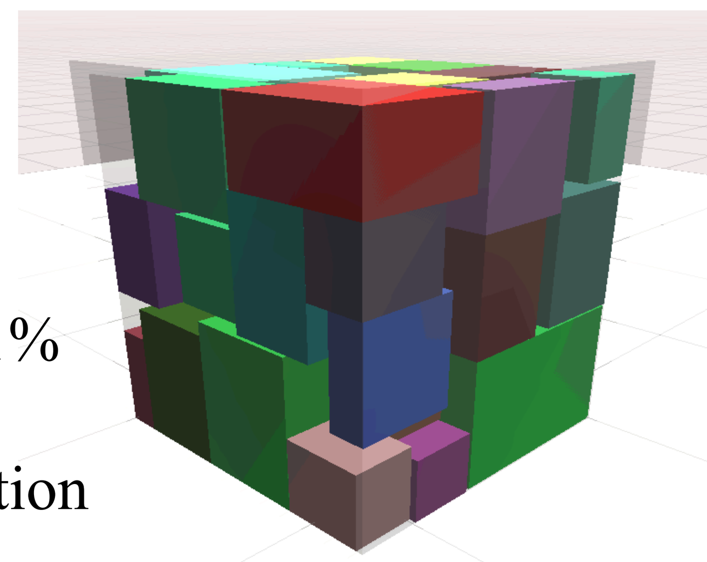
- Vehicle Data: 64 potential departure times for vehicles going toward targeted next hub in one day with varying capacities
- Parcel Data: One-day path, arrival time, target time of departure, dimensions, weight of 1712 parcels
- Following are the preliminary results from the parcels-to-vehicle assignment for parcels going towards the same hub

Vehicle ID	Vehicle departure time (mins)	Vehicle capacity (m^3)	Number of parcels	Total parcel volume	Parcel Fill %
59	69.75	7.6	154	5.28	69.5
58	147.1	18.7	382	12.36	66.1
57	166.03	18.7	389	13.03	69.7
56	189.82	18.7	408	13.37	71.5
53	272.07	18.7	379	12.71	68

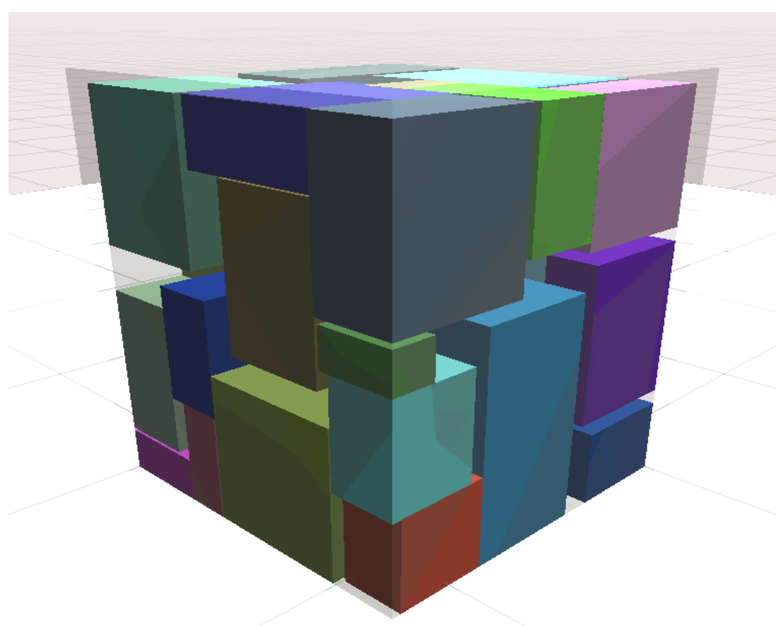
Parcel-in-container Packing Results

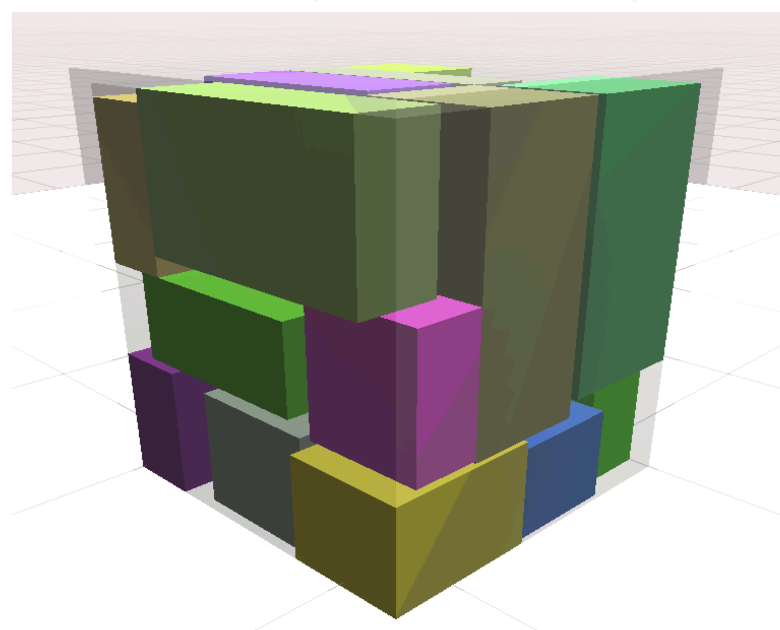
Model is run for all parcels assigned to a vehicle
- 100 parcels; Max fill = 82%
- Container sizes available: 2x2x2, 4x2x2, 4x4x2, 4x4x4 (feet)

Preliminary illustrative results:
- Computation time = 66 seconds
- Selected containers; 4x4x4 (3)
- Fill = 79%, 81.5%, 82%; Net = 81%



Images: Container Packing Visualization
The 3D models are made in Unity





Future Research

- Integration of 3 models and simulation-based experiments
- Acceleration of computation time for solving packing model
- Weight and loading stability considerations
- Ordering for packing fragile parcels
- ML-based updating fill percentage in each iteration
- Live-case considerations for grouping parcels based on sequence of hubs traveled

Autonomous and interconnected technologies for the container supply chain: The MOSES Concept

Nikolaos P. Ventikos[1], Konstantinos Louzis[1], Ioannis Kanellopoulos[1] ,Eleni Krikigianni[2], Evangelia Latsa[2], Gerco Hagesteijn[3], Hans Cozijn[3], Hans van den Broek[4]

IPIC 2021
8th International Physical Internet Conference
June 14-17, 2021
Virtual

1 2 3 4

Introduction & Concept

The maritime branch of the EU container supply chain is configured as hub-and-spoke networks where large containerships transport cargo to Deep-Sea Shipping (DSS) ports and a limited number of Short Sea Shipping (SSS) feeder lines and container trucks on RoPax vessels distribute cargo to smaller ports for further transhipment. However, economically, environmentally, and socially unsustainable land-based transportation still covers most of the cargo transhipment from Hub ports. MOSES aims to significantly improve the modal shift to SSS by creating sustainable feeder services to small ports that have limited or no infrastructure (Figure 1). This will be achieved by developing the following highly automated/autonomous technologies and integrating them in existing operational processes: i) a digital collaboration and matchmaking platform (**MOSES platform**), ii) an **innovative, green feeder vessel** with a **robotic container handling system**, iii) a manoeuvring and docking scheme where **autonomous tugboats** collaborate with an automated mooring system.

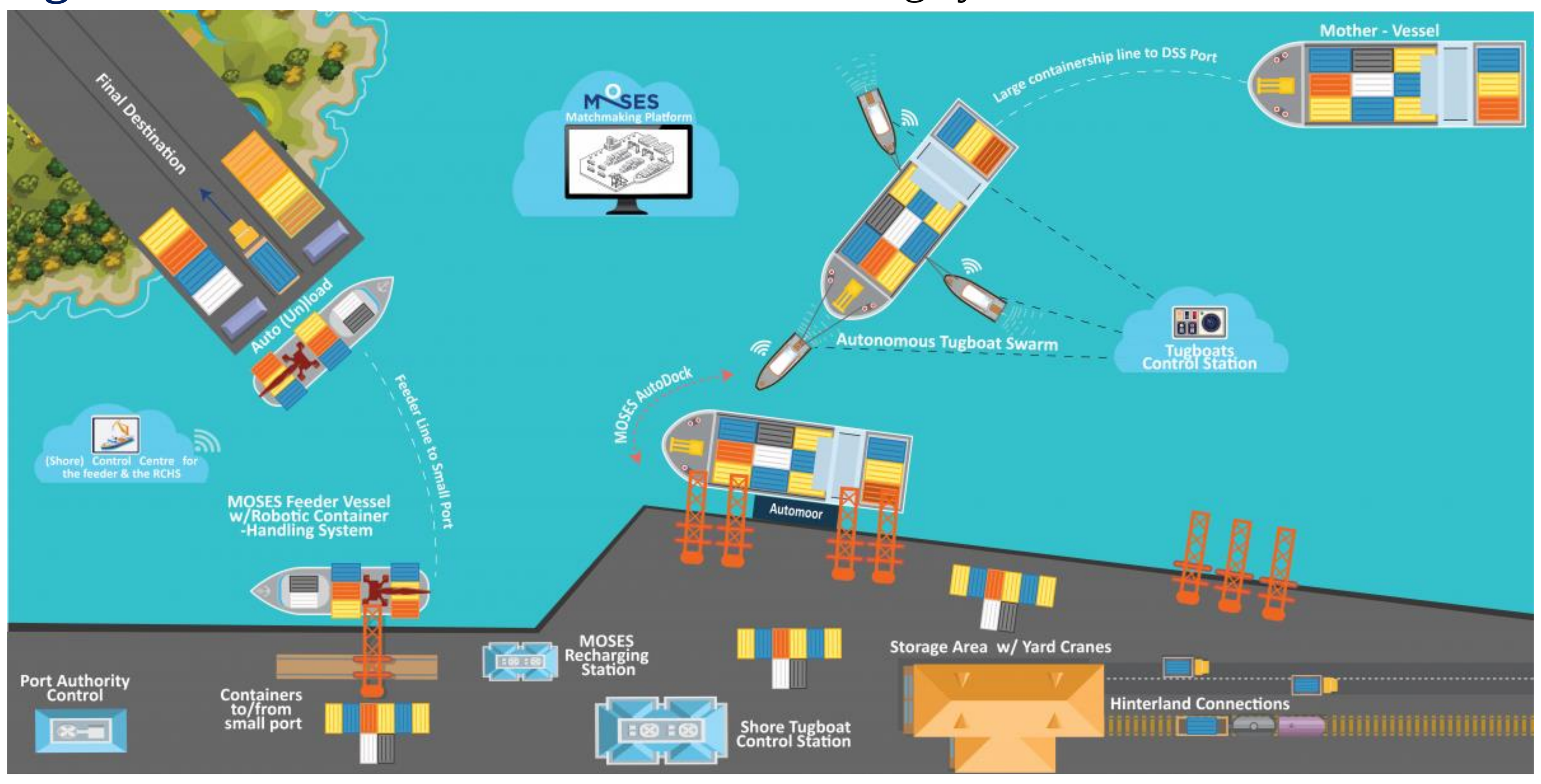


Figure 1: MOSES Concept and innovations

MOSES Matchmaking platform proposes a solution where data feeds from various sources allow logistics stakeholders to optimise and adjust routing plans, bringing the PI one step closer to the targets set for 2030 and 2040.

Objectives

- Create sustainable, efficient SSS services that are competitive to land-based transportation (road/truck and rail) by enabling interconnected and horizontal collaboration between shippers and port operators, as well as other logistics stakeholders.
- Stimulate synchromodality by exploiting the interoperability and interconnectedness of Physical Internet systems, which provide efficient and more environmentally friendly cargo transport and induce a positive modal shift to SSS.
- Improved efficiency and end-to-end delivery times of SSS mode, by balancing the backhaul traffic through sustainable feeder services and automated tugboats that will reduce lead and waiting times for large containerships in DSS ports.

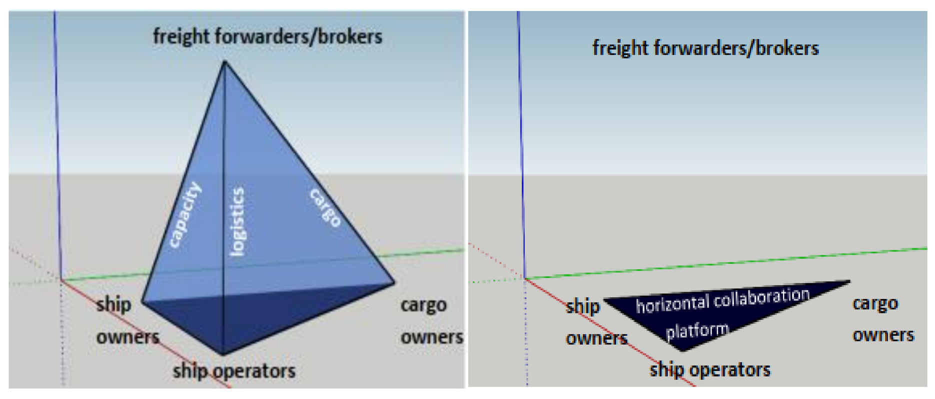
MOSES Matchmaking Platform

MOSES matchmaking platform is an advanced digital collaboration platform, that aims to foster the multidisciplinary horizontal collaboration (Figure 2) and interaction between key actors of the maritime supply chain by offering scenario–building capabilities that consider as-is and to-be costs, energy efficiency, and environmental footprint. The platform will also encompass a specific module for sharing information on empty containers.

The platform will get as an input relevant cargo data offers and respective flows in both parts of the SSS route from the respective stakeholders' systems, besides a comprehensive map of alternative transport modes' traditional flows. Using a Machine Learning matchmaking logic and data driven-based analytics for the best cargo allocation, the platform will aggregate the information over multiple shippers and make the data visible to logistics service providers by capitalising on both sea & hinterland transport modes (road/truck, rail and inland waterways).

The MOSES platform will be supported by suitable governance models and will enable shippers and carriers to consolidate flows for both directions, addressing planned deliveries/spot capacity, whilst triggering inter-firm collaboration, which is critical in the innovation process.

Figure 2: Left: the current status of collaboration, Right: The expected level of collaboration



Innovative Feeder Vessel with the Robotic container-handling system

MOSES innovative feeder will be a new automated ship design with reduced environmental footprint that will enable optimized and automated/autonomous operation during sea passage between large and small ports and will be designed with the outlook of fully autonomous navigation in the future. The innovative feeder services will be supported by the capabilities of the MOSES matchmaking platform in terms of the available and required cargo flows. MOSES innovative feeder will be outfitted with an autonomous system for cargo handling (Figures 3 & 4). The **MOSES robotic container-handling system (RCHS)** will be designed as a fully self-supporting system, without requiring human intervention, and it will be capable of safely (un)loading containers from the innovative feeder to the quay side. To detect containers, their relative position, and other relevant operational parameters, an advanced sensor suite is necessary for building an accurate 3D-world model of the operating environment. This 3D-world model will also detect people and objects other than containers (e.g., trucks) for maintaining safety. The RCHS will be coupled with a human operator, who will supervise the operation from a Shore Control Station (SCS).

Figure 3: Graphical representation of MOSES feeder with the RCHS

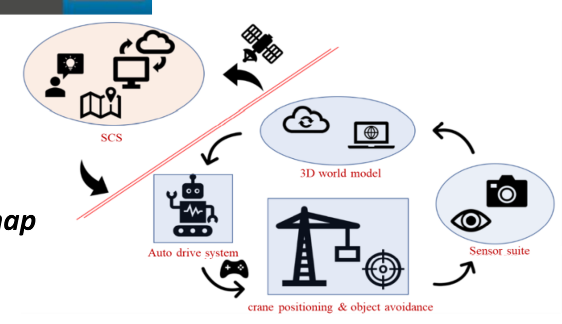


Figure 4: Concept map of MOSES RCHS

MOSES Autonomous Tugboats

The **MOSES Autonomous tugboats swarm** will assist a large containership approaching the DSS port with the manoeuvring process. MOSES will develop technologies that can be retrofitted to existing tugboats that include hardware units, AI and navigational control software, and sensors. These will enable them to work cooperatively in a swarm formation. The MOSES autonomous tugboats automate the docking process of a large mother vessel by collaborating with a re-engineered version of the Trelleborg's AutoMoor system (MOSES AutoDock). The tugboats swarm is monitored through a remote-control station located at the DSS port (MOSES Shore Tugboat Control Station), which continuously monitors and gathers information about the process until the vessel is safely docked and communicates with the Terminal Operation System (TOS) to initiate stevedoring operations. The container owners can subsequently use the MOSES matchmaking platform to determine the best connection (i.e., hinterland or SSS**)** based on parameters such as Turn around time (TAT), cost, and environmental footprint.

Conclusions & Ambitions

MOSES automation technologies will minimise safety risk in manoeuvring, berthing, and cargo handling in seaports. The modal shift and optimisation of MOSES will also reduce the environmental footprint per transported TEU through optimisation of next-leg deliveries and modal shift to SSS and rail. MOSES innovations include the deployment of multiple data producing devices that will contribute to the development of logistics applications and an automated, interconnected, multimodal transportation system, in line with the ALICE Physical Internet Roadmap.

MOSES project has received funding from the European Union's Horizon 2020 research & innovation programme under grant agreement No. 861678.

Conference Sponsor-Logo Area

The Benefits and Price of Hyperconnectivity in last mile delivery

Johan LEVEQUE, Walid Klibi, Gautier STAUFFER

IPIC 2021
8th International Physical Internet Conference
June 14-17, 2021
Virtual

KEDGE BUSINESS SCHOOL

LA POSTE

INTRODUCTION

We investigate the impact of mutualizing the delivery network design of several delivery actors.
While a lot of companies are redesigning their urban delivery network by adding optimally positioned micro-hubs to city inner centers, a lot of them are including new delivery methods such as cargo-bike. Contrary to vans, cargo bikes have the possibility to effectuate several delivery shifts per day, making them much more agile in a hyperconnected delivery network. Sharing the benefits of mutualizing several delivery actor's network provides a better quality of service, such as same day delivery, by creating a denser network. On the other hand, optimally designing a shared network leads to additionnal hidden costs due to sub-optimality.

HISTORICAL DELIVERY NETWORKS

Urban delivery networks are most of the time built on the idea of massifying the parcel flow. Thus, those flows are sorted in big peri-urban industrial platforms that implies a sub-optimal travel of inner-city flows. This historical design leads to a loss in term of quality of service due to the time needed to treat one local parcel as seen in the following figure. On the meantime, a lot of national-wide delivery actors are operating such networks sometimes even fulfilled by the same peri-urban platforms.

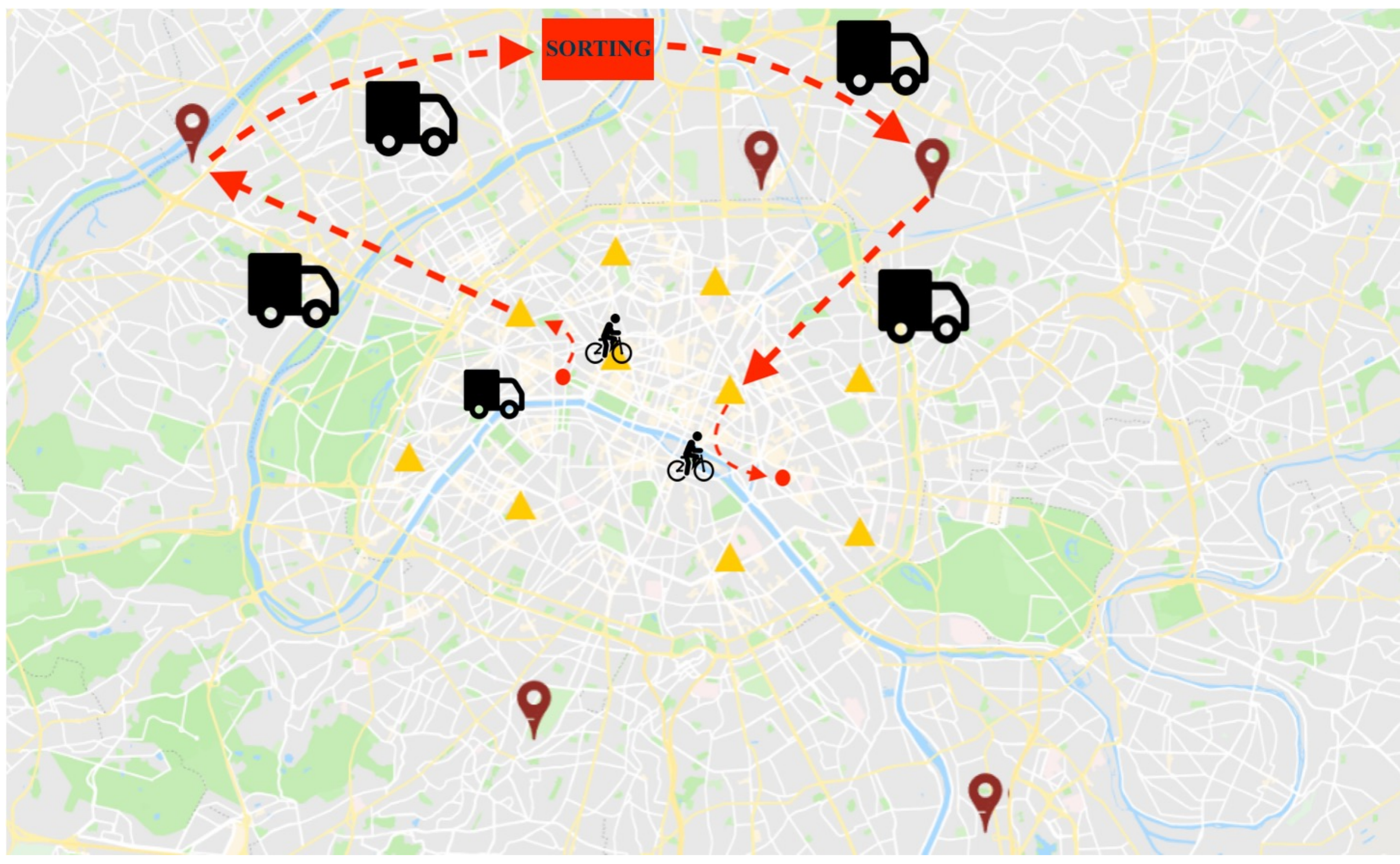


Figure 1 : *Travel of a inner-city flow in a historical single actor massified urban delivery network.*

METHODOLOGY

First, we give ourself a list of delivery actors giving us a list of sharable micro-hubs locations by merging their pool of potential micro-hubs. We then divide our city into clusters in which we estimate the daily delivery price using density based routing length estimation methods.

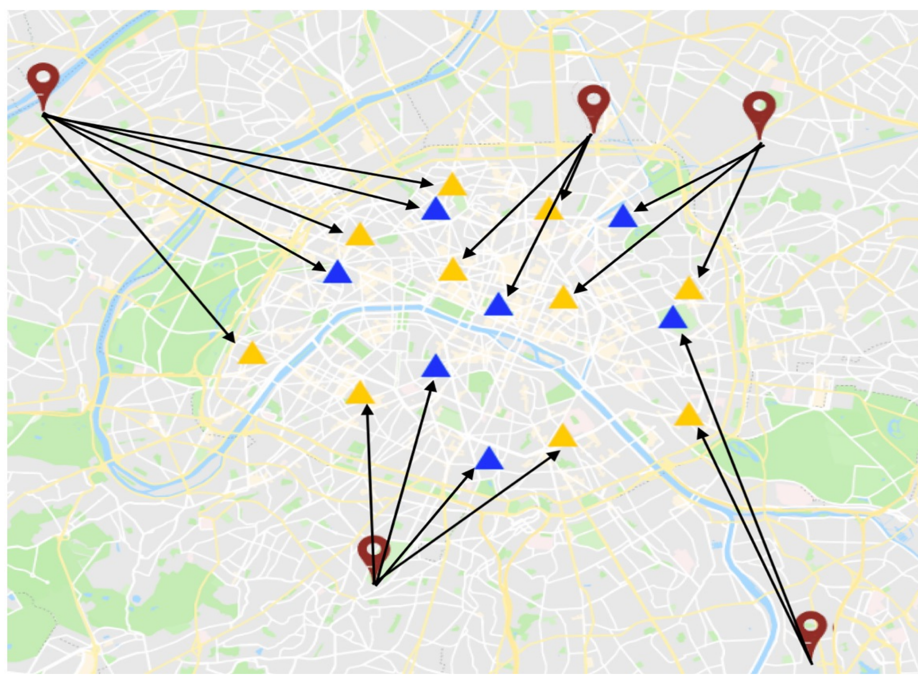


Figure 2 : *Micro-hub Potential locations for Multi-actor delivery network*

Second we build a multilayer mathematical model in which we consider a multi actor delivery network and try to minimize the total price by contraining the maximum transit time for each parcel flow, increasing the service level as wanted. We enforce connectivity by opening transversal connections between active micro-hubs.

PRICE AND BENEFITS OF HYPERCONNECTIVITY

The results of a our model give us a mutualized multi-actors delivery network design allowing us to use the agility of the cargo-bikes to provide a higher service level increasing the benefits. However, this service level increase comes with the investment of creating sub-optimal mono-actor networks when taken individually.

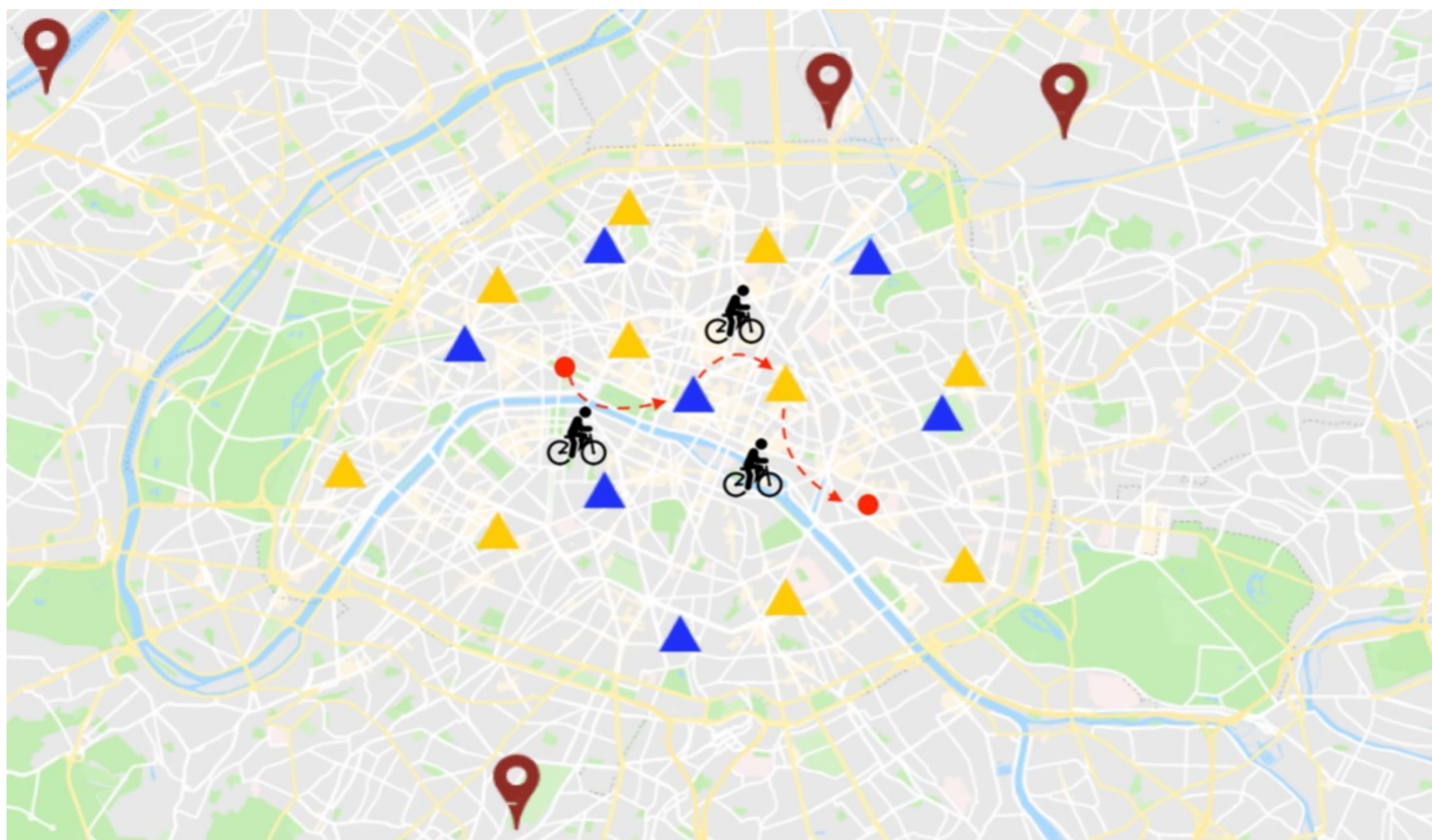


Figure 3 : *Travel of a inner-city flow in optimized mutualized hyperconnected urban delivery network.*

Building on Synergies between Freight Logistics and People Mobility in Urban Areas

Walid Klibi, Olivier Labarthe, Ghazaleh Ahmadi, Jean-Christophe Deschamps and Benoit Montreuil

IPIC 2021
8th International Physical Internet Conference
June 14-17, 2021
Virtual

KEDGE BUSINESS SCHOOL · ims · université de BORDEAUX · Georgia Tech · Physical Internet Center

OBJECTIVES

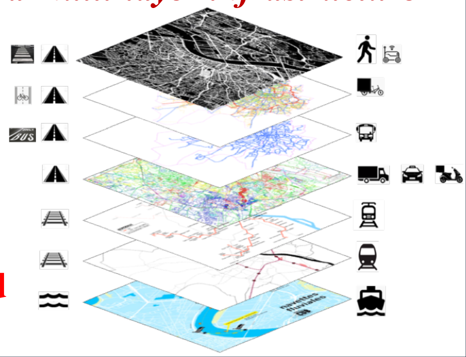
- Investigates the feasibility of **goods transshipment** with a joint usage of **public mobility** and **freight urban vehicles**.
- Assess the **potential benefit** of a **joint mobility system** for goods delivery **in urban areas.**

CONTEXT : URBAN AREAS

A multimodal mobility on a multi-layer infrastructure

Transportation of goods : **lower volumes** per shipment and **higher number** of on-demand shipments.

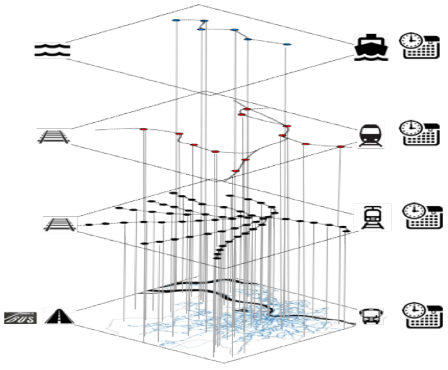
How to transship goods based on the joint use of public transport modes and on-demand freight modes ?



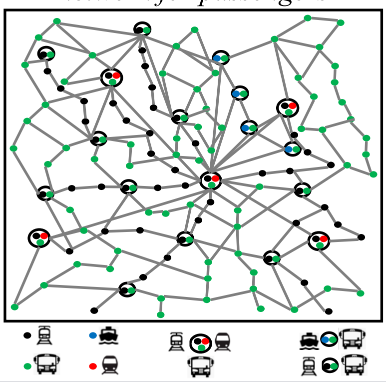
METHODOLOGY

It builds on **Hyperconnectivity** and **Synchromodality** concepts and on a **data-driven** simulation-optimization framework.

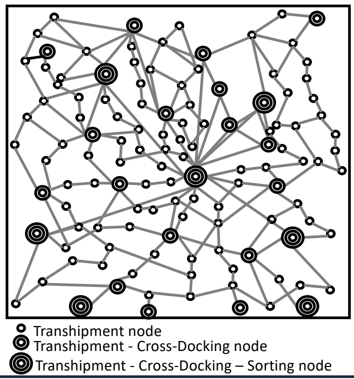
Multi-layer representation of the scheduled public transportation network



Urban area interaction network for passengers



Hyperconnected City Logistics based on people mobility networks

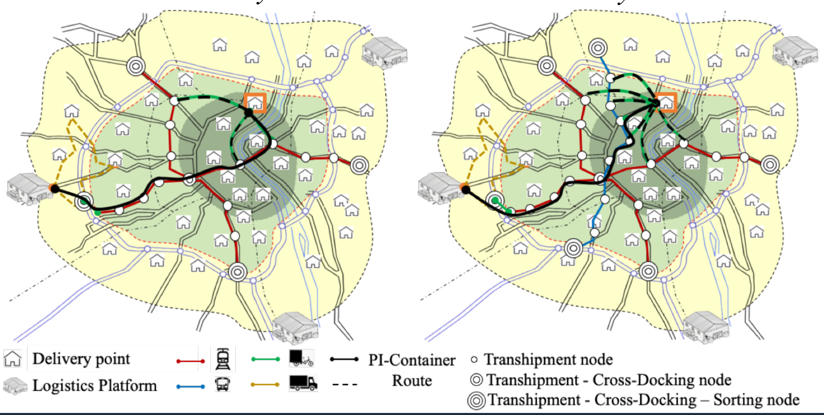


ILLUSTRATIVE CASE

A showcase is built within Bordeaux (France) with **two alternative scenarios** (itinerary and modes)

- Ensuring adequate **delivery service using public transportation** instead of dedicated on-demand vehicles.
- **Reducing** the **waiting time** and **parcels time-space footprint** within the urban area.
- Increasing the **societal benefit** with **less** on-demand vehicles and **congestion** on the shared transportation infrastructure.

PI-Containers by Tram · *PI-Containers by Tram and buses*



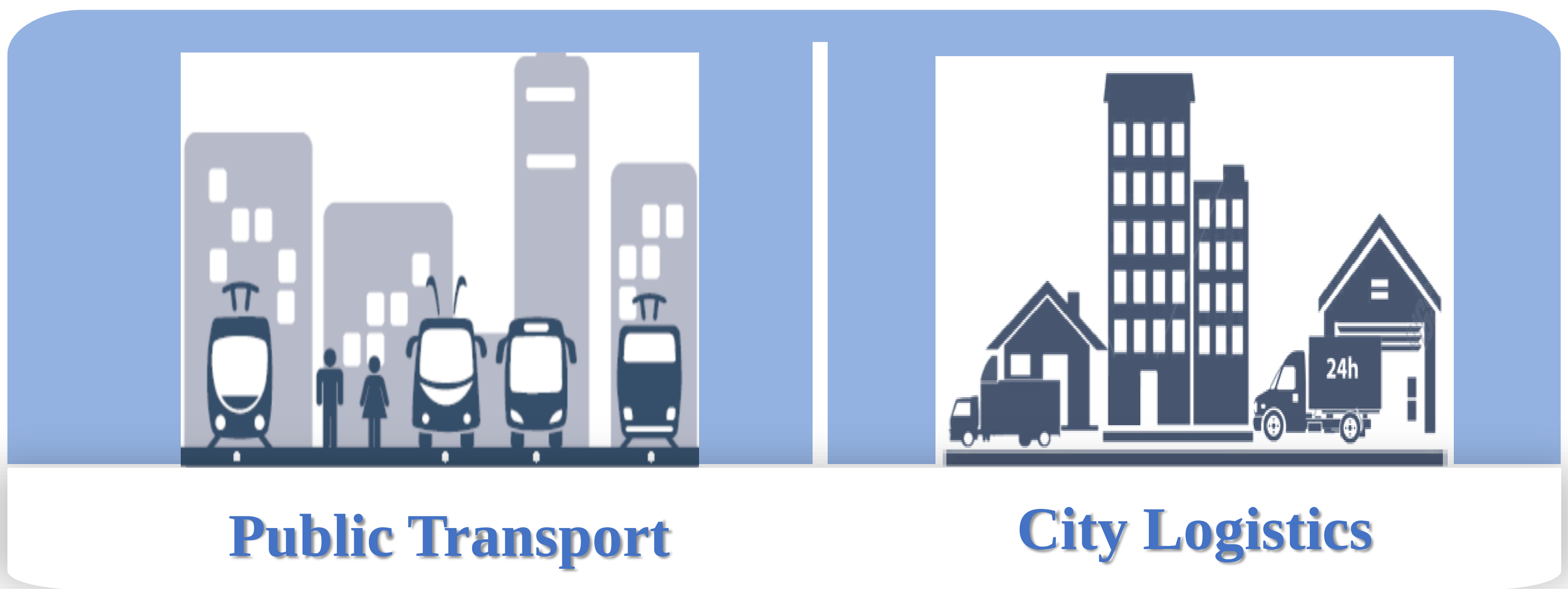
Conference Sponsor-Logo Area

SYNCHRONIZING PUBLIC TRANSPORT CO-MODALITY NETWORKS FOR CITY LOGISTICS

Xuan Yang, Hao Luo and Ming Jiang

IPIC 2021
8th International Physical Internet Conference
June 14-17, 2021
Virtual

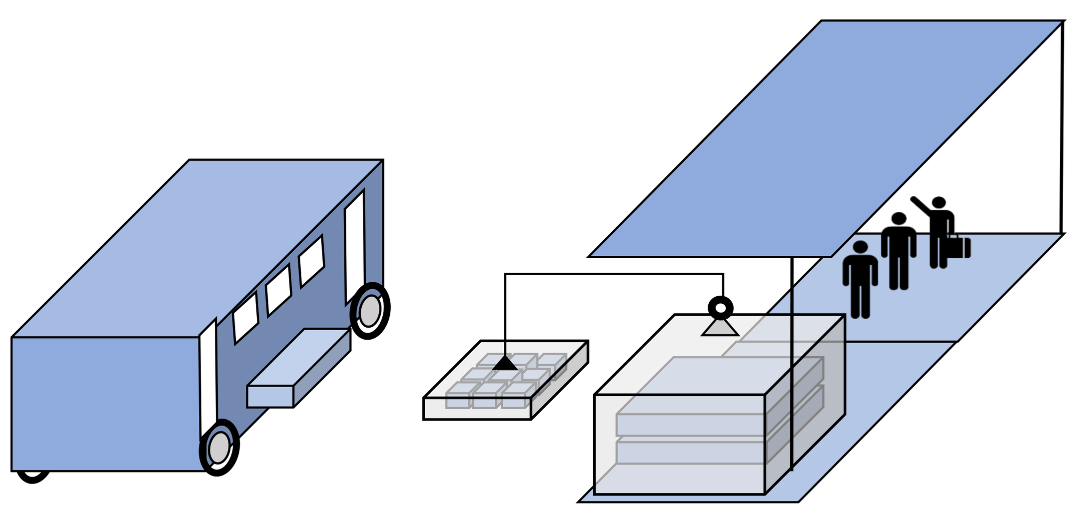
Introduction



- Public transport is widely considered as an ideal to travel around.
- City logistics focuses on improving the efficiency of urban freight transportation while reducing congestion, emission and noise.

Public transport co-modality aims to combine city logistics with urban public transport to ensure a more efficient and sustainable urban transportation system. Most public transport systems operate service that has some common characteristics such as pre-determined vehicle cycle, fixed route and limited capacity. Based on these characteristics, how to make packages delivery plan to synchronize freight flows and transit networks becomes a significant challenge in public transport co-modality.

Co-modality Solution



Illustrative example of PI container exchange between bus and bus stop.

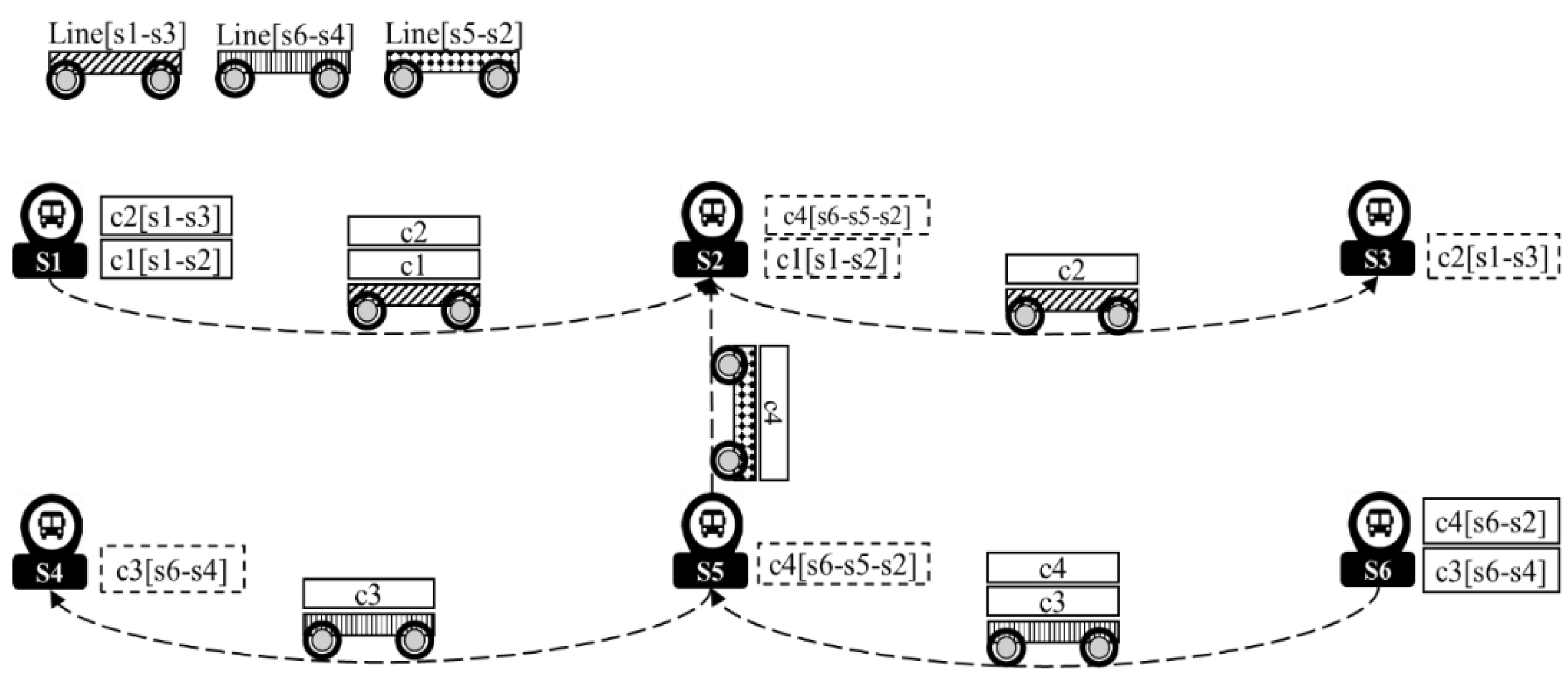
Assuming that design decisions on the infrastructure are already made and thus a network is already implemented within an urban area. In this network, the bus plays the role of transit for both passengers and express packages at the same time. At bus stops, smart containers used for temporary storage of packages make it possible to exchange packages between bus and bus stop automatically with no disturbance to the movements of passengers.

The bus stop area is divided into two areas: passenger area and freight area. All packages are stored in the physical internet containers which is designed to achieve autonomous loading/unloading.

Objective

This paper considers conceptual and mathematical models in which passengers and express parcels are handled in the same bus network. The problem aims to combine logistics couriers and bus network to respond to a number of packages transport requests in an urban area without any delays, while minimizing the total service time.

Passenger and Freight Sync



A demonstration of goods transportation in public transport network.

Freight on Transit:
Fixed route and frequency

Constraints
Time Space

Objective
To minimize the total freight delivery time

$$Min \sum_{r \in R} \left(\sum_{l \in L} y^{l}_{r,d(r)} \cdot a^{l}_{d(r)} - \lambda(r) \right)$$

Conclusion

An efficient memory-based solution approach is proposed and validated. The effect of the density of freight bus stops on the packages delivery efficiency of the bus network is examined. Numerical studies are conducted to illustrate the analytical observations and provide further insights.

Production-Intralogistics Synchronization in Physical Internet-enabled Manufacturing Systems

Mingxing Li and George Q. Huang

IPIC 2021
8th International Physical Internet Conference
June 14-17, 2021
Virtual

香港大學
工業及製造系統工程系
Department of Industrial and Manufacturing Systems Engineering
The University of Hong Kong

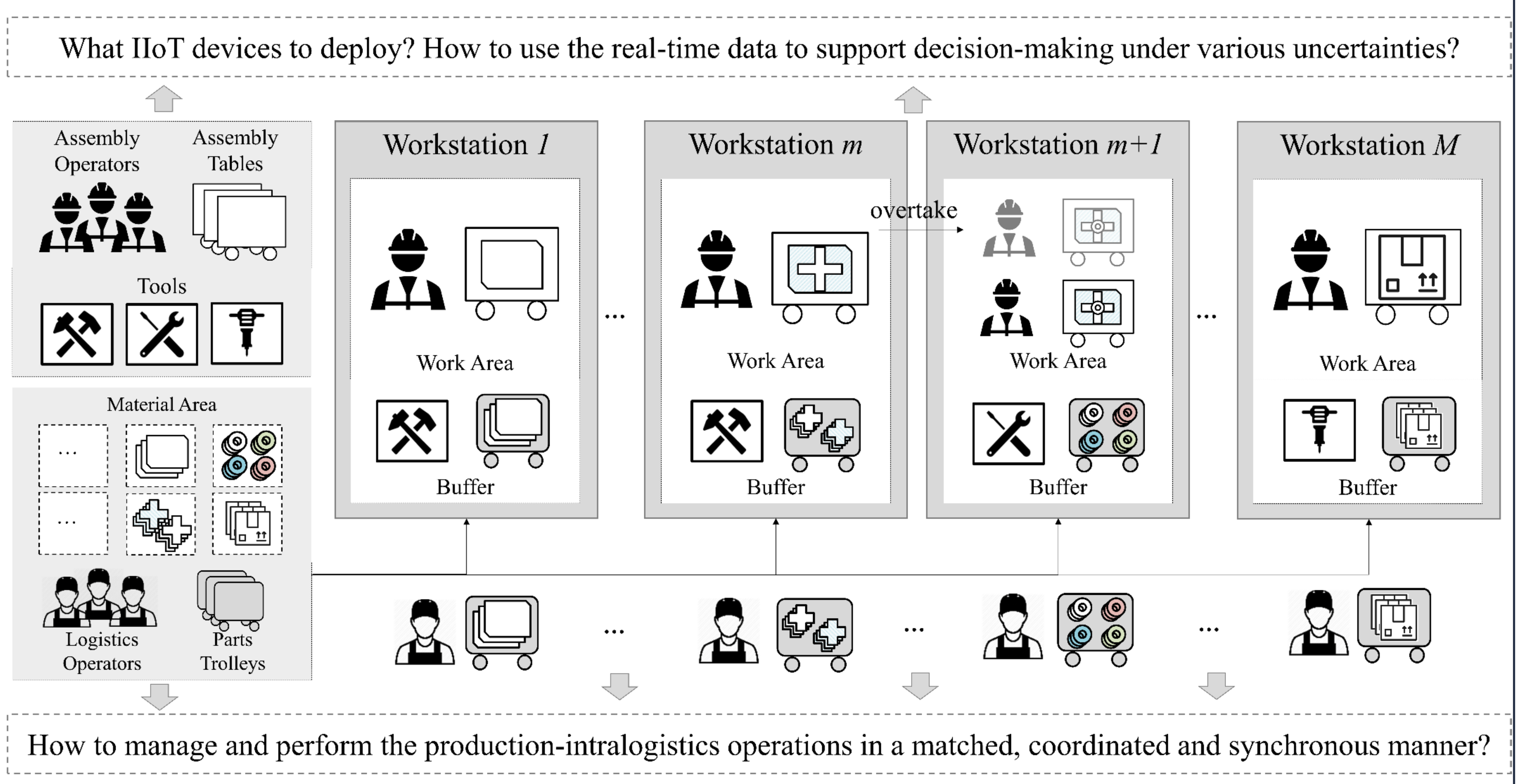
Introduction

The widespread adoption of Physical Internet (PI) technologies has promoted data and information sharing, real-time communication, and networking in the industry, which transforming the operations are managed and performed in many fields such as manufacturing. For example, the production operations and intralogistics operations in a shop floor are inherently coupled and entangled by physical flow (raw materials, components, sub-assemblies, or work-in-progresses). In traditional management mode, the production and intralogistics processes are managed separately by different departments without considering global benefits because the information cannot be timely collected and shared among the department to make informed decisions. Nowadays, advanced Industry 4.0 technologies such as Industrial Internet-of-Things (IIoT), digital twin, and cloud computing are gradually adopted by manufacturers to upgrade their factories. The sheer amount of data are real-timely collected, transmitted, and analyzed so that the information barriers among different department of a single factory are removed. Therefore, it is possible to manage the production and intralogistics processes in a synchronized manner by leveraging the strengths of real-time data, to improve the overall production efficiency and resource utilization.

The Flexible Assembly Line

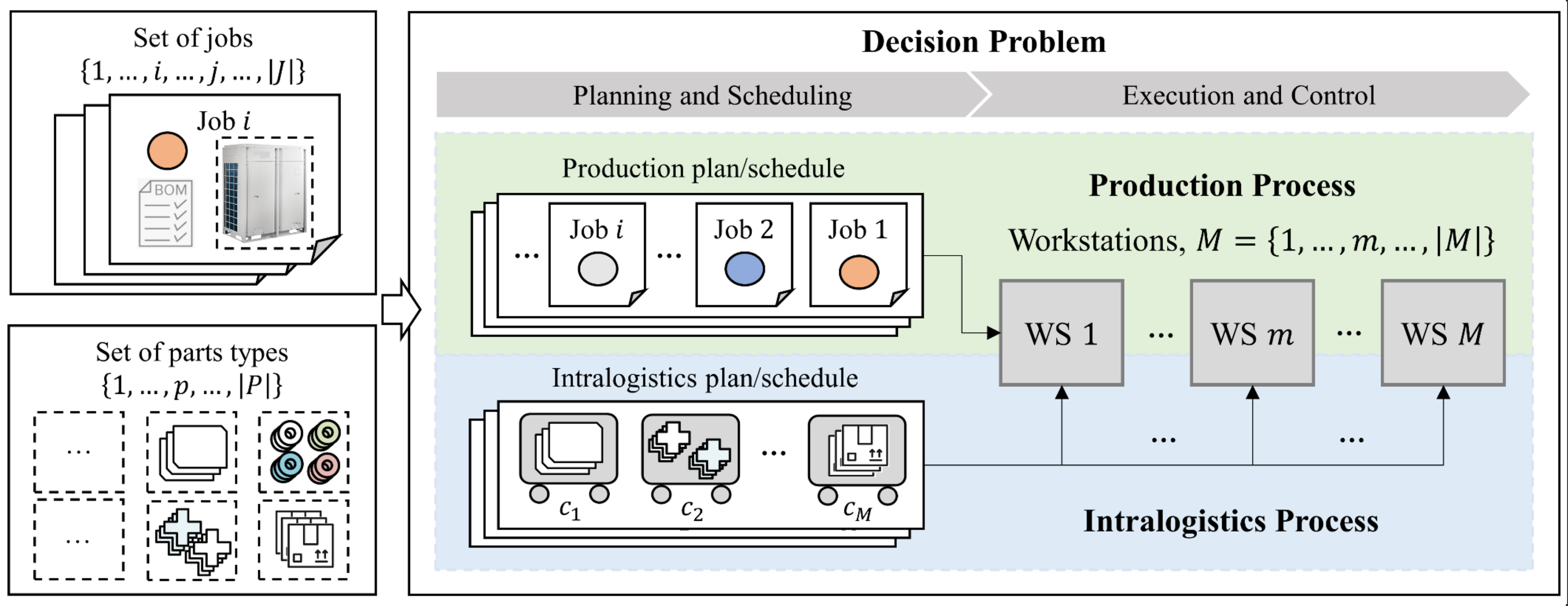
The investigated workshop is called flexible assembly line (FAL). FAL combines the features of flexible/hybrid flow shop (FFS) and traditional assembly lines. The FAL integrates the feature of parallel operations in FFS into the assembly line. As the figure depicts, there are a series of workstations that contain a work area for operations and a buffer area for placing the required tool and parts corresponding to the specific operation. Mobile assembly tables and part trolleys are adopted to replace the conveyor belt and line side stocking in traditional assembly lines. Each product is carried by an assembly table to go through all workstations in a unidirectional flow. All operations of a single product are performed at the table. Parallel operations are allowed at the workstation, which means one job can overtake another if necessary. Required parts are consolidated in the part trolleys, and logistics operators take the loaded trolleys to the designated workstation as scheduled. Now the manufacturer is faced with two new problems:

- What IIoT devices to deploy and what data to capture? How to utilize the real-time data for decision-making under various spatiotemporal uncertainties (e.g., equipment failure, stochastic operational time, new job arrivals) and minimize their adverse effects?
- How to manage and perform the production-intralogistics operations in a matched, coordinated, and synchronous manner to ensure a smooth workflow and high throughput with improved customer satisfaction?

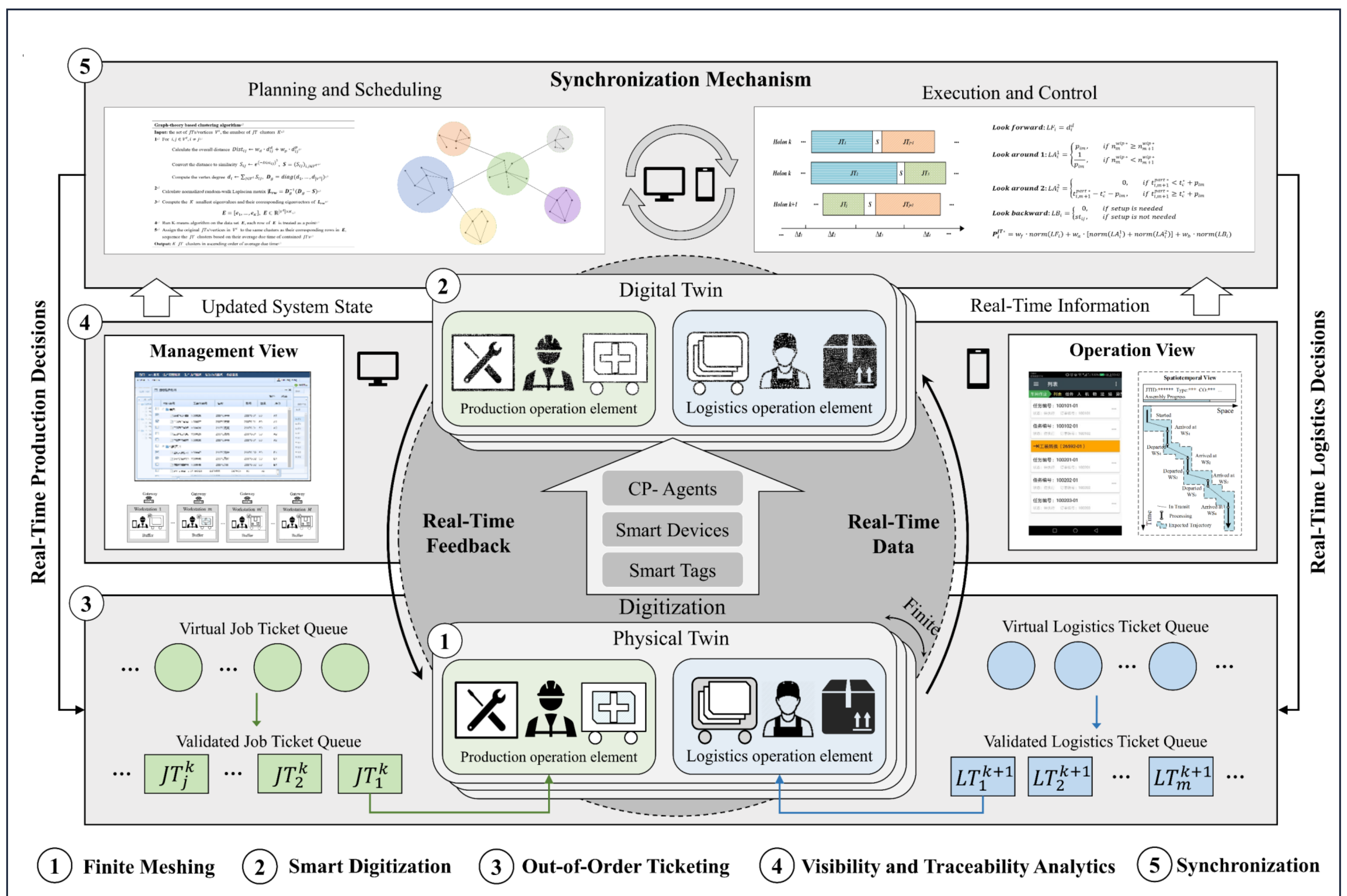


Production-Intralogistics Synchronization in FAL

The structure of FAL is flexible to achieve dynamic line balancing; thus, the focus of the production management in FAL has to switch from balancing the line to managing the production-intralogistics operations in a matched, coordinated, and synchronous manner to ensure a smooth workflow and high throughput. The paper defines it as the PiL synchronization problem. As shown in the figure, The decisions are to make a production plan/schedule for assembling request products and corresponding intralogistics plan/schedule for each workstation to fulfill the demand and optimize selected measures regarding customer satisfaction, throughput, and the PiL synchronization degree.



5-Phase Implementation of PI-enabled Synchronizd PiL System



Conclusion

This study investigates the production-intralogistics (PiL) synchronization problem in a new assembly system. A 5-phase framework is proposed to implement the synchronized PiL in the manufacturing system. It has made pioneer work in investigating the production-intralogistics synchronization problem in PI-enabled manufacturing environments and contributes to the methodological research in utilizing real-time data to facilitate operations management in the context of Industry 4.0.

A dynamic routing protocol with payments for the Physical Internet: A simulation with learning agents

Martin Briand, Rod Franklin

IPIC 2021
8th International Physical Internet Conference
June 14-17, 2021
Virtual

MINES ParisTech | PSL | THE KLU KÜHNE LOGISTICS UNIVERSITY

Introduction: need for dynamism

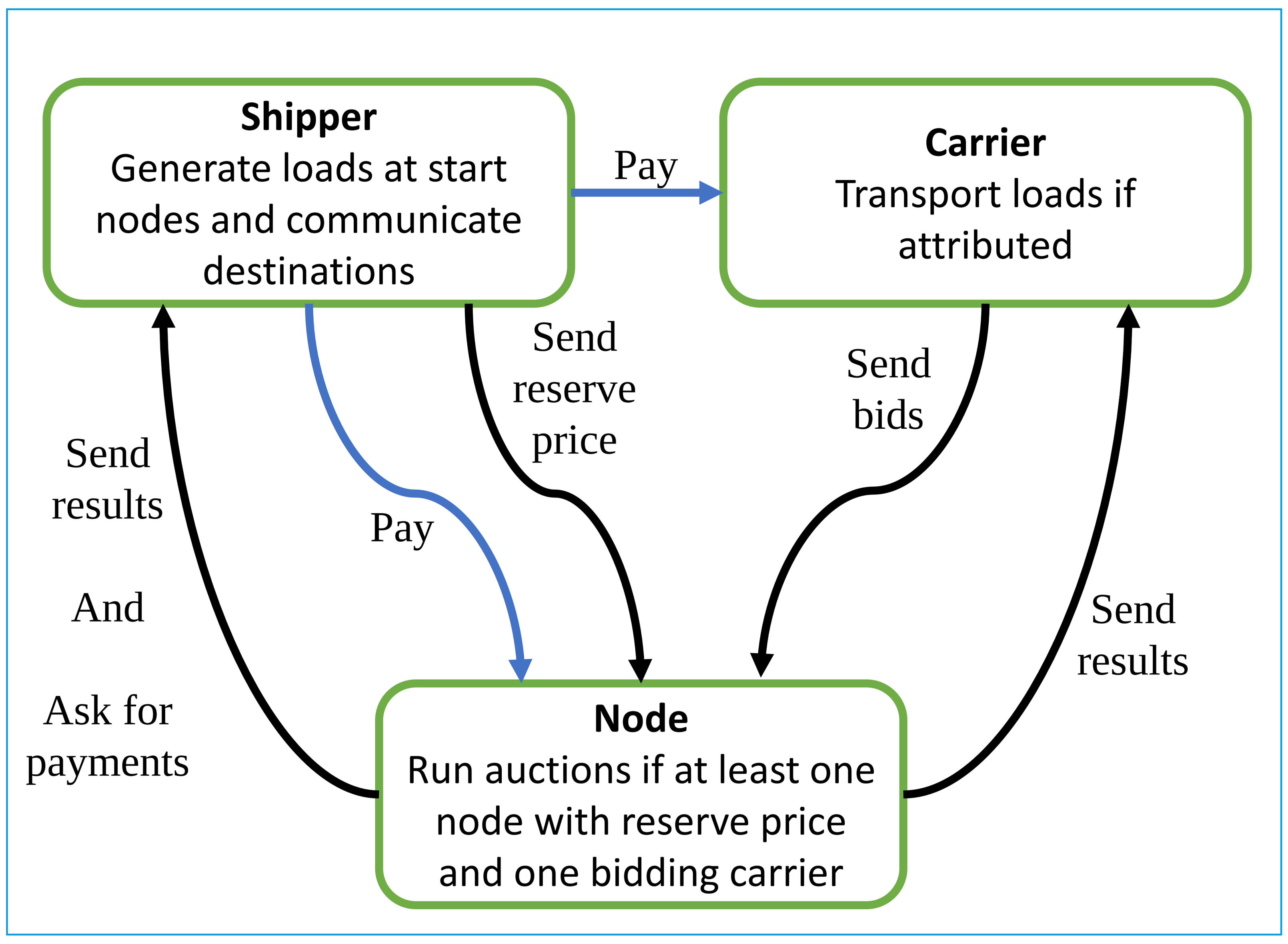
- The **Physical Internet** aims to route loads **dynamically** over a logistic **network**.
- Each times a **Load** arrives at a **Node**, a decision must be **quickly** taken to jointly select the **Next node**, the **Carrier**, and the **price** for this new segment.
- Hence the need for a **Digital Internet like protocol**. Something **fast** and **reliable**.
- This poster presents what we believe is the **first PI dynamic routing protocol taking payments into account**. It also present an **environment to test** the new protocol.

Protocol: Presentation

Because the freight transport market is highly competitive, we can use a **game theoretic** approach to design our protocol: the key component is an **auction** (see Auction).

It involves three agents: **Shippers** to generate Loads, **Carriers** to transport them, and **Nodes** to materialize the network and make the routing decisions.

Each time a Load and a Carrier are waiting at a load, an auction is run. Note that the case when we have a single carrier bidding should be rare in the context of a **Few-To-Many market**.



Protocol: Auction

Generate:

- **Reserve Price** for **Shippers** by answering: "How much am I ready to pay to get this loads from this node to its destination?" → R
- **Bids** for **Carriers** by answering: "What is the minimum price I am ready to be paid to transport the load to each of this next node?" → b_l^i for each carrier i and each node l
- **Penalties** for **Nodes** by answering: "What is the expected price to deliver the load after reaching a next node?" (sort of DI *Administrative Distance*) → w_l

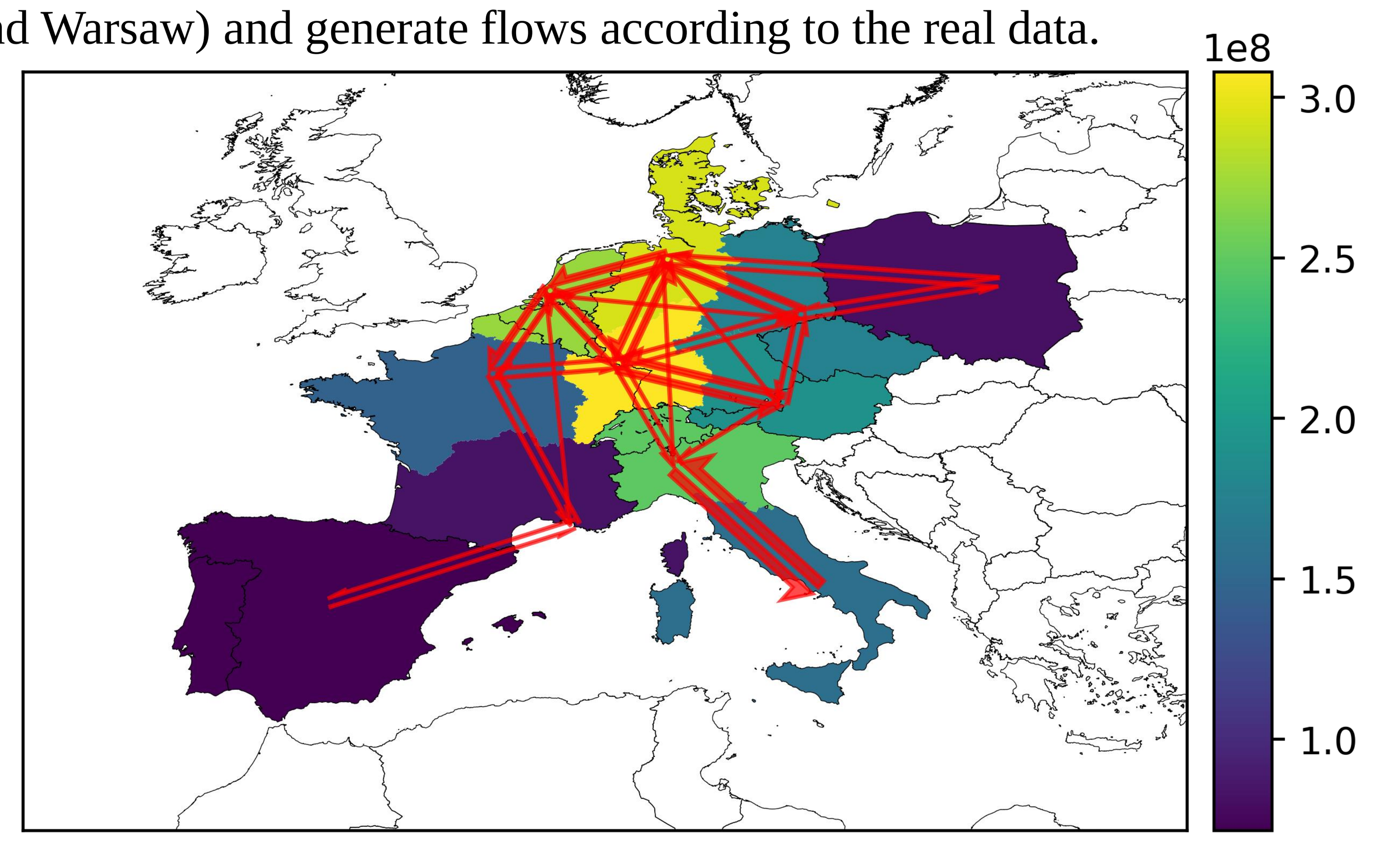
Select $(i^*, l^*) = argmin(w_l + b_l^i)$ and $B = min_{i \neq i^*}(w_l + b_l^i)$.
Carrier i^* must transport load **to node** l^* and is **paid** $B - w_{l^*}$.

The auction is ***IR***, ***IC***, ***BB***, and almost always ***AE*** (in a Few-To-Many market).

Simulation: Network

We make a simulation of the **Western European freight transport market** using the ETIS+ database.

We aggregate volumes of 11 regions to 11 single nodes (Bremen, Dresden, Madrid, Marseille, Milan, Naples, Paris, Rotterdam, Saarbrücken, Salzburg, and Warsaw) and generate flows according to the real data.



Regions for each node and Main cross-region flows.
Legend: total tonnages (imports + exports, excluding internal transports)

Simulation: Agents

Carriers try to **bid** in a **profitable** manner to **offset the following costs:**

$$C(t) = OnTheRoad(t) \times RoadC + FarFromHomeC(t)$$

where

$$FarFromHomeC(t) = NotAtHome(t) \times DriversC + AdminC(t)$$

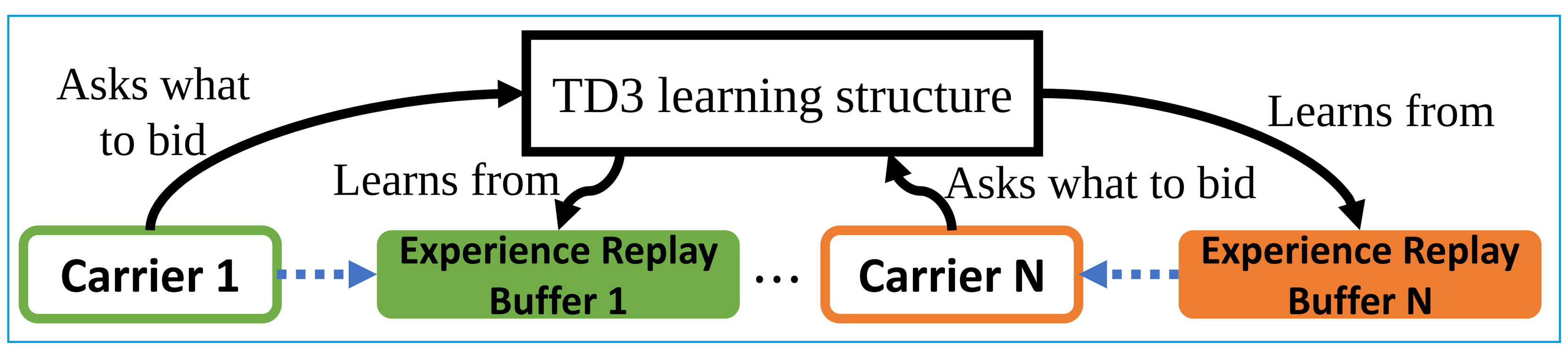
$RoadC$ and $DriversC$ are constant for each carrier and $AdminCosts(t)$ is an exponential function doubling the total cost in 10 days.

Shippers bid high to make sure Reserve Prices are not involved.

Nodes charge zero and then we increase this price.

Simulation: Carriers' bids

We simulate Carriers using **Reinforcement Learning**. To increase the learning speed and avoid to relearn the model for each new agent, we use a **centralized learning TD3 structure** for all carriers.



Can Physical Internet support the integrated management of last-mile and first-mile (reverse) logistics? An exploratory analysis

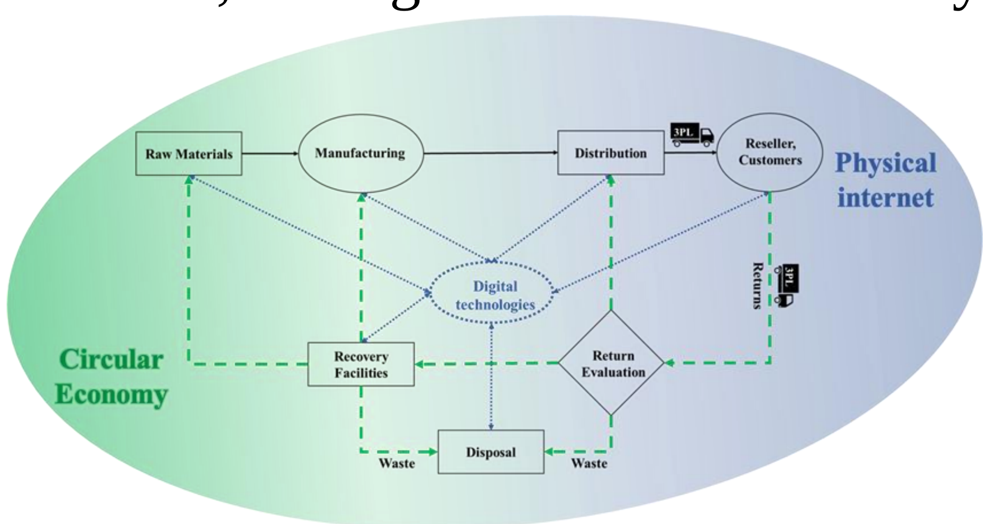
Agnusdei G.P., Sgarbossa F.

IPIC 2021
8th International Physical Internet Conference
June 14-17, 2021
Virtual

NTNU

INTRODUCTION

Closed-loop supply chain (CLSC) management developed within the CE paradigm and consists mainly in managing forward and reverse flows within a supply chain simultaneously. Integrating forward and reverse supply chains means that "last mile" logistics activities within the forward flow are strictly interconnected with the first-mile logistics activities of the reverse flow. In this context, Physical Internet (PI) paradigm represents an opportunity to increase logistics management efficiency, because it attempts to transform the way physical objects are handled, moved, stored, realized, supplied and used, by applying concepts from internet data transfer to real-world shipping processes, aiming towards sustainability.

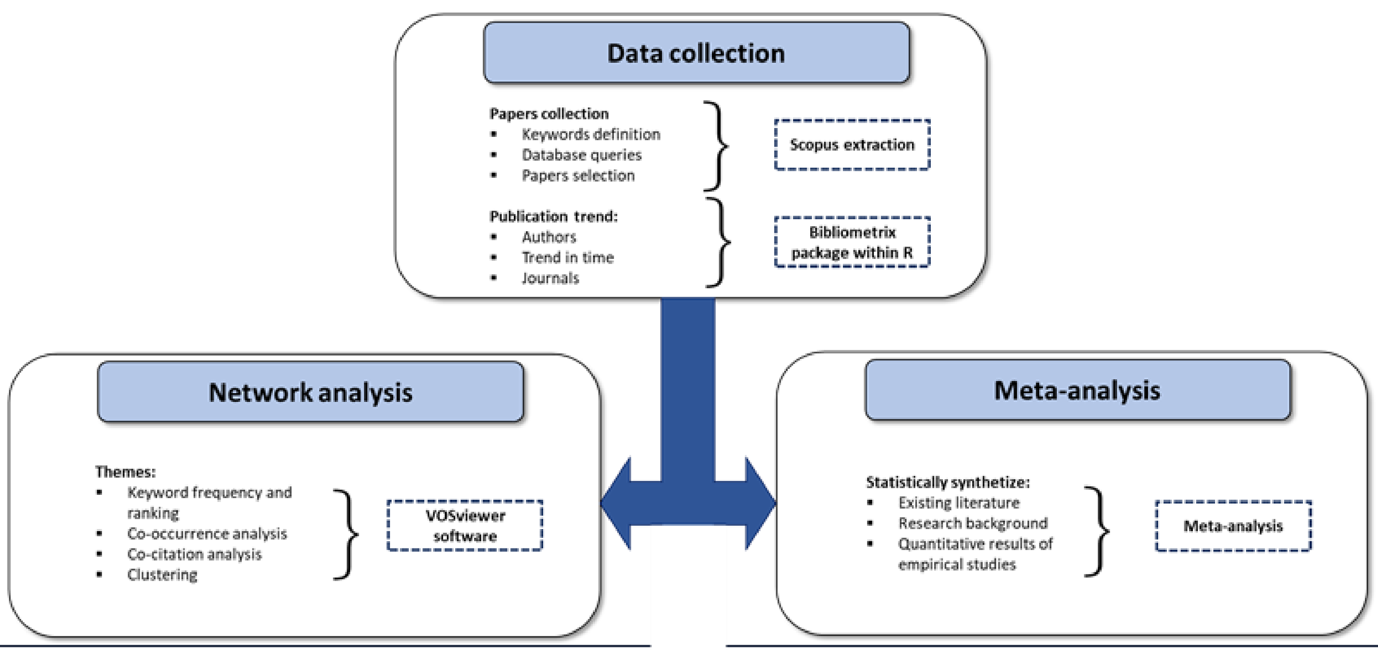


RESEARCH AIM

The aim is to investigate the opportunities of the PI application in the last mile and first mile (reverse) logistics in order to identify new frameworks for developing a Sustainable Supply Chain Management (SSCM) based on the circular economy approach.

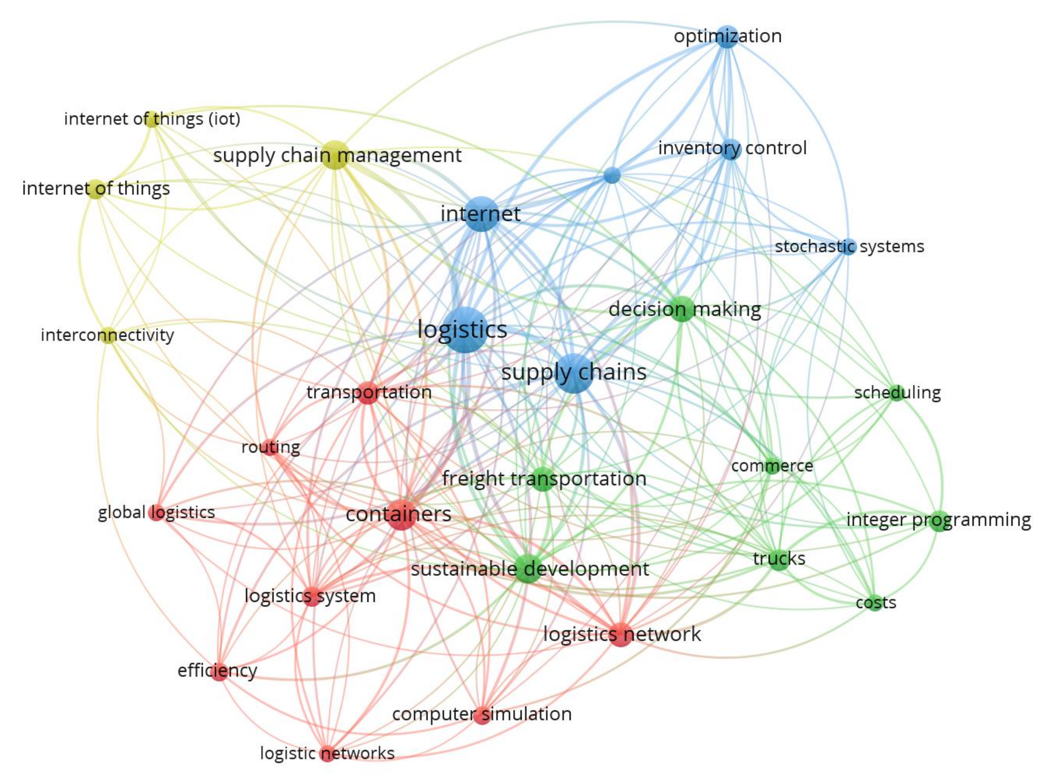
METHODOLOGY

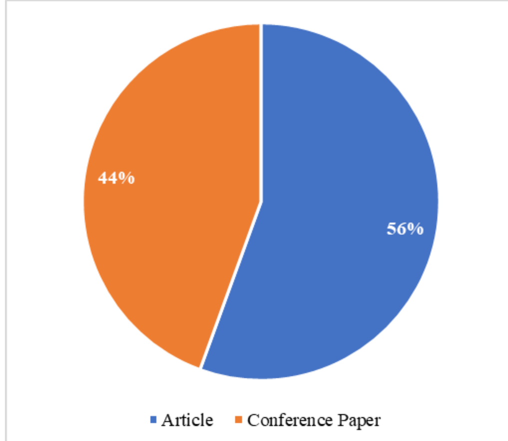
The research design adopted for this study is based on three steps: (i) the collection of studies about current PI applications in last-mile and first-mile (reverse) logistics activities; (ii) the network analysis to identify established and emerging research clusters; (iii) a meta-analysis to statistically synthetize existing literature and visualize the research background by combining and assessing the quantitative results of empirical studies.

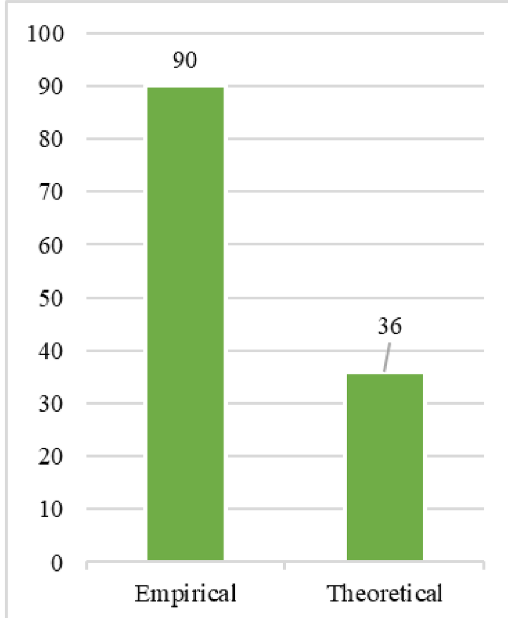


PRELIMINARY RESULTS

126 documents were extracted from Scopus. Only 90 documents are empirical. 56% are articles and 44% are conference papers. The network analysis allows for the identification of four clusters within the research field. The green cluster refers to transportation and decision making, the red one refers to container and logistics network, the yellow cluster concerns IoT and the blue one regards inventory control and optimization in SC.







CONCLUSIONS

The study offers multiple opportunities to managers who must choose management options for optimizing the integration of last-mile and first-mile logistics in their everyday work. Commodities can be handled and stored through standard, modular and smart boxes, which will enable a shift toward a much more distributed, hyperconnected and efficient logistics system. The results of the conducted analyses serve as a tool for managers to assess the current state of the PI implementation in last-mile and first-mile (reverse) logistics and to identify their future needs so that they may decide whether to invest and improve current physical, digital and operational interconnectivity, thus enhancing their firm SSCM performance.

FUTURE WORK

Since meta-analysis is useful to statistically synthetize existing literature by verifying some hypotheses, it will be carried out to identify areas which need further investigation.