A Data-based Opportunity Identification Engine for Collaborative Freight Logistics Based on a Trailer Capacity Graph

Authors: Jianlin Luan[a] , Nicolò Daina[b] , Kristian Hegner Reinau[c], Aruna Sivakumar[d], John W. Polak[d]

Corresponding author: Jianlin Luan[1]

[a] School of Maritime Economics and Management, Dalian Maritime University, 1 Linghai Road, 116026, Dalian, China

[b] School of Government and Public Policy, University of Strathclyde, McCance Building, G1 1XQ Glasgow, UK

[c] Freight Transport Research Group, Department of the Built Environment, Aalborg University, Thomas Manns Vej 23, 9220 Aalborg Ø, Denmark

[d] Urban Systems Lab and Centre for Transport Studies, Department of Civil and Environmental Engineering, Imperial College London, Exhibition Road, SW7 2AZ London, UK

Abstract

Logistics operators who participate in horizontal collaboration can gain economic benefits as well as being better placed to meet their environmental goals. Real-time information communication and computation processes are essential if the potential of collaboration is to be fully exploited in practice, however. Such processes come with large computational costs, especially in terms of optimisation of joint route planning, and these costs are a significant hurdle for the implementation of real-time collaborative logistics operations. On the contrary, data-based approaches provide a viable, albeit suboptimal, alternative that can enable real-time collaborative order sharing. Data-based approaches for the identification of collaboration opportunities are typically based on origin-destination (OD) matching of trips from one collaborating company with shipments from the others. This, however, prevents the exploitation of en-route collaboration opportunities. We propose a data-based engine for identifying collaboration opportunities during shipment planning stages thus providing a practical solution for real-time collaborative freight logistics. The engine is based on a multigraph approach, which we refer to as the trailer capacity graph (TCG), which describes the real-time trailer capacity status of all collaborating partners. The TCG approach enables shipments to be matched according to both the OD and the route of trailer trips, and we customise this further with several enhancements to improve performance for real-time operations. Numerical experiments based on real-world data from two logistics companies show that the TCG approach identifies a significantly larger number of opportunities, and provides a higher total distance saving than traditional OD-based matching. The experiments also demonstrate that the implementation of this engine in combination with trailer route approximation and route shape simplification allows trade-offs between computational performance and effectiveness of opportunity identification. This trade-off implies that the engine can be flexibly tailored according to user preferences and requirements.

Keywords: Collaborative freight logistics, Data-based, Real-time, Large-scale, Trailer capacity graph

1 Introduction

1.1 Context

The freight transport industry plays a key role in the economic development of our societies, yet it is subject to significant negative externalities. In the first two decades of the new century, the road freight logistics sector has been under very strong and increasing pressure to reduce its environmental impacts significantly (McKinnon et al., 2015). Along with environmental sustainability, freight transport needs to improve its social sustainability practices, mitigating road accidents and congestion, as well as many other complex multidimensional social impacts, ranging from the welfare of drivers to noise impacts on communities (Kumar & Anbanandam, 2019). In addition to these societal pressures for more responsible road transport, road freight logistics is facing formidable challenges from within the core of its business environment. Traditional challenges such as stronger competition and very low margins are exacerbated by booming e-commerce, which has increased demand volatility and introduced more complexity in value-chains, and extremely short-term fulfilment requirements (Savelsbergh & Van Woensel, 2016).

Internal innovation may not be sufficient to respond to these challenges, or too costly given the low margins of the industry. As a result, providers of freight logistics services are resorting to collaboration. The benefits of collaboration between freight transport carriers are widely recognised. Horizontal collaboration established amongst companies whose transport networks partially overlap has the potential to generate significant shared gains (Cruijssen, Dullaert, et al., 2007; Karam et al., 2020; Leitner et al., 2011). A recent review of ten studies of collaborative logistics by (Allen et al., 2017) reports that horizontal collaborations amongst logistics operators can deliver up to 16% lower distance-based costs, 24% lower environmental costs and a 25% increase in business volume. These benefits, however, are mostly estimated from numerical studies and are subject to considerable variability depending on specific case studies' contexts.

The existing literature on collaborative logistics has placed a strong emphasis on collaborative network design (Audy et al., 2010; Soysal et al., 2018; X. Xu et al., 2021; X. F. Xu et al., 2017), optimal collaborative transport planning with various objectives and constraints (Caballini et al., 2016; Chabot et al., 2018; Hernández et al., 2011), cost and gain sharing mechanisms (Audy et al., 2010; Guajardo & Rönnqvist, 2016; Palhazi Cuervo et al., 2016; Vanovermeire & Sörensen, 2014) and request exchange mechanisms (Gansterer & Hartl, 2016; Houghtalen et al., 2011; M. Lai et al., 2017).

Despite the benefits emerging from theoretical and modelling studies, the potential of collaborative partnerships amongst carriers is yet to be extensively exploited in practice (Creemers et al., 2017). The practical implementation of tools for collaborative logistics is limited to niche applications

(Perboli et al., 2016) or limited to the analysis of historical data to identify partners instead of being applied in the management of collaborative operations (Creemers et al., 2017). The widespread adoption of collaboration practices is prevented by a series of practical barriers and methodological challenges (Cruijssen, Dullaert, et al., 2007; Karam et al., 2021; Mostafa et al., 2019). Amongst the methodological challenges the most significant are the allocation of gains, the identification of a trusted coordinating partner and identification of collaboration opportunities (Cruijssen, Dullaert, et al., 2007). The coordinating partner is usually a trusted third party, with whom fleet or order data is shared. Challenges in finding such trusted third parties are related to the demonstrability that the third party can "coordinate the cooperation in such a way that all participants are satisfied" (Cruijssen, Bräysy, et al., 2007).

The identification of collaboration opportunities is arguably the core of collaborative logistics. This is the focus of this paper. Verdonck et al., (2013) distinguish between order sharing and capacity sharing collaboration opportunities. In order sharing, orders from customers are "shared" or "exchanged" so as to increase efficiency by optimally re-allocating the orders amongst collaboration partners. In the capacity sharing framework, instead of pooling orders, transport companies share assets, specifically vehicle capacity. We note here that this dichotomy between order sharing and capacity sharing approaches is somewhat deceptive because, for order sharing to be possible, the collaborating operators must also share available capacity in their vehicles. Indeed, an order sharing opportunity will involve a company sharing a shipment that is ordered by one of its customers, and another company sharing available capacity on one of its vehicles to fulfil the order delivery. Thus, while the framework we propose in this paper focuses on order sharing, it envisages that each collaborating company also shares its spare capacity to fulfil order shipments originating from other collaborating companies.

Regardless of the collaboration type, the identification of collaboration opportunities, and indeed collaboration partners, can be achieved using operation research (OR) approaches (Cleophas et al., 2019; Gansterer & Hartl, 2018; Vaziri et al., 2019) or data-based approaches (Creemers et al., 2017; Deng, 2014; Suarez-Moreno et al., 2019). Even though methodologies from the OR literature have the advantage in terms of their ability to lead to, or at least seek, optimal solutions, they tend to be NP-hard, which makes them computationally expensive. The computational complexity of such approaches is a significant limitation, as it hinders "real-time information communication processes" that are deemed crucial for effective dynamic planning of horizontal collaborative transport (Pan et al., 2019).

In order to overcome the computational burden characterising traditional OR approaches, data-based, rule-based approaches can be designed for real-time operations and, despite providing suboptimal solutions, can still deliver significant improvements in efficiency compared to a non-collaborative baseline. Hitherto, however, these data-based approaches have typically been based on fully matching the origin-destination (OD) pair of a trailer trip by a company and the OD pair of a shipment order from another company. This full OD matching is very limiting as it excludes the potential of collaboration arising from accepting a shipment from a collaborating partner that requires a limited deviation of the routes of initially planned trailer trip chains before the planned trip chains are executed.

Against this background, our paper introduces a data-based and computationally efficient approach to the real-time identification of “order sharing” collaboration opportunities.The identification of collaboration opportunities is in real-time because during the trip chain plan update stage, dispatchers can use real-time information about the opportunity to dynamically update their initial trip chain plans, by offloading a shipment or receiving a shipment from another company. Our approach is based on a multigraph, which we call the Trailer Capacity Graph (TCG). The TCG makes it possible to match shipments from one company with the available capacity of trailer trips of another company, by allowing en-route pick-ups and drop-offs. Therefore, our proposed approach relaxes the full OD matching requirement. Instead, a shipment from a company can be matched with a trip from another company when the location of the shipment is close enough to the route of the trip. We call this “en-route matching”, even though it happens before the trip chain is executed, because the matching is based on allowing limited deviations on the routes of pre-planned trailer trips to accommodate the additional shipment. En-route matching allows the identification of a larger number of opportunities than traditional data-based approaches that require the full OD match between trailer trips and shipments.

Through numerical experiments based on real-world data from two logistics operators, we demonstrate the superior performance of our TCG-based engine compared to traditional OD matching in terms of the number of collaboration opportunities identified and the total distance savings. While this improvement in operational performance comes at the cost of higher computation times, that increase is not on a par with that seen in OR approaches and does not hinder the possibility of implementation in real-time, real-world operations.

1.2 State-of-the-art for data-based identification of collaboration opportunities

Most of the methodological contributions on collaborative logistics are from the OR literature (Cleophas et al., 2019; Gansterer & Hartl, 2018; Verdonck et al., 2013). While this section reviews studies that use data-based approaches, in order to structure that discussion, we use a classification of order sharing approaches originally developed by Verdonck et al., (2013) for approaches from the OR tradition. In our context, Verdonck et al.’s classification system enables us to categorise the data-based approaches to order sharing in the literature.

Verdonck et al., (2013) identified a number of implementation modes (or types) for order sharing: joint route planning (JRP), auction mechanisms (AM), Bilateral Lane Exchanges (BLE), Information Secured Swapping (ISS) and Shipment Dispatching Policies (SDP). JRP involves pooling orders from all the partners and formulating and solving appropriate vehicle routing problems (VRPs). For an agent (i.e. a carrier or a transport company), order sharing with AM involves first identifying which customer requests should be exchanged, e.g. by solving cost minimisation problems, including route planning, or through some heuristics. Next, the agent informs the cooperating partners that the identified orders are open for bidding. In BLE, full truckloads with specific OD pairs are exchanged. In ISS, transport companies swap orders in an effort to minimise the total travel distance, while making sure that the minimum amount of information is shared. In SDP, a carrier with an expiring shipment deadline, but only partially loaded, picks up appropriate orders from collaborating companies so as to increase its load level.

The discussion of Verdonck et al., (2013) on these order sharing implementation modes focuses particularly on solutions that frame the identification of collaborative opportunities as an optimisation problem. Nonetheless, alternative approaches to optimisation are possible for BLE, ISS or SDP, namely data-based and rule-based approaches that identify temporal and spatial overlaps between the orders of one company and the available capacity of another. Furthermore, hybrid solutions could be applied to JRP, BLE, ISS and SDP cases, in which data-based methods are utilised to simplify the initial optimisation problem. Arguably, while data-based, rule-based approaches may lead to sub-optimal solutions, they do still offer the possibility of significantly improving the operational efficiency of road freight logistics by real-time identification of collaboration opportunities. In the paragraphs that follow, we discuss three data-based or hybrid implementations of order sharing.

Deng (2014) observes that in large alliances or market facilitation systems it is inefficient to search manually for available capacity or shipments, especially in highly dynamic markets characterised by volatile demand, supply and prices. The author addresses this by means of a purely data-based ISS approach for the automated discovery of collaboration opportunities based on applying a hierarchy of rules to match available vehicle capacity and required consignment movements. Specifically, the approach finds, in order of priority, vehicles and shipments that have overlaps in: pick-up and delivery dates, OD, vehicle type and shipment type, available tonnage and shipment weight, available space and shipment size. The algorithm is tested using historical data from transportation companies. While this automated search is likely significantly to cut the cost and times involved in manual searches, Deng does not provide any quantitative estimates for the improvement in the number of matches or the search speed of the hierarchy of rules. The spatial matching of Deng's approach considers a spatial match to be successful when at least a vehicle origin and a shipment origin or a vehicle destination and a shipment destination match but does not consider the potential impacts of any re-routing necessary to fulfil the shipment order when the OD matching is only partial.

Creemers et al., (2017) present a matching procedure to identify potential collaboration partners purely based on the geographical compatibility of respective shipments. This can be viewed as an initial procedure to pre-select a limited number of collaborative partners in an effort to limit the size of subsequent JRP problems. Historical shipment data from logistics companies' databases is homogenised so that it has unique OD pairs. Shipments are clustered so that bundling, round-trip and collect and/or drop opportunities are identified. Potential collaborations are evaluated based on several KPIs, such as the ratio of shared distance to total distance, the ratio of shared volume to total volume, and the ratio of shared tonne-kilometres to total tonne-kilometres. Users may assign weights to those KPIs according to their preferences. Capturing users preferences allows ranking potential collaboration opportunities. In their practical implementation, however, Creemers et al., (2017) do not consider any criteria other than spatial matching and use the procedure only for strategic decisions regarding collaborating partner identification, rather than in operations management.

Suarez-Moreno et al., (2019) combine clustering of orders with optimisation to identify opportunities for cargo consolidation. The clustering is based on temporal OD matching of orders and product compatibility. The cargo consolidation optimisation, essentially a JRP problem, is then performed within a cluster in order to limit the computational time needed to identify opportunities.

We should at this point mention that there are also commercial platforms[1] for real-time freight exchanges between logistics operators that are likely to use data-based and optimisation approaches. We do not specifically review such platforms here, because they operate as two-sided online market platforms, which logistics operators *occasionally* use to post available capacity and to offload shipments that they are unwilling or unable to deliver. This modality falls outside the scope of this paper that focuses on longer-term collaborative solutions.

In summary, the limited studies focusing on data-based approaches for the identification of collaboration opportunities are essentially OD-based. That is, they only identify collaboration opportunities when there is a geographical match between the OD pair of a shipment (and therefore order) requested to a company and the OD of a trailer trip from another collaborating company (for ISS or SDP), or between ODs of shipments from collaborating companies (for JRP). Pure data-based approaches for ISS and SDP that focus only on OD matching of shipments and trips miss potential opportunities derived from a partially-loaded trailer transiting in the proximity of orders that could be fulfilled by that trailer with small or null route deviations. In other words, the focus on OD matching in the current data-based approaches to identifying collaboration opportunities neglects en-route opportunities for ISS and SDP order sharing. This is a gap that needs addressing. Low computational time solutions for real-time ISS are of great interest, given the other advantages of ISS over JRP. In

[1] Examples are TIMOCOM (https://www.timocom.co.uk), UBER FREIGHT (https://www.uberfreight.com/), TRUCKSTOP (https://truckstop.com/product/mobile-services/) and others

particular, ISS solutions can minimise both the exchange of potentially commercially sensitive information and negative attention from regulators, who could view JRP as more problematic with respect to competition laws.

1.3 Aim, contributions and significance

This paper aims to contribute to the scientific literature on real-time collaborative freight logistics by addressing a major limitation of current data-based approaches for ISS and SDP order sharing. In fact, current data-based approaches are heavily reliant on OD matching to identify collaboration opportunities: this results in the impossibility to identify and exploit the potential of en-route collaborative opportunities. The solution presented in this paper addresses this limitation, introducing a data-based engine that identifies collaboration opportunities by matching shipments to available trailer capacity along a given route. In this respect, our specific technical contributions are detailed as follows.

First, we introduce a novel approach for collaborative opportunity identification based on the trailer capacity graph (TCG). TCG is a multigraph that describes the spatiotemporal capacity status of a trailer based on an initial trailer trip chain plan. Our TCG approach significantly enhances the capabilities of current data-based approaches in collaboration opportunities identification by including en-route pick-ups and drop-offs of shipments from collaborating companies, relaxing the requirement that OD pairs of trailer trips and shipment need to coincide for the materialisation of a collaboration opportunity.

Second, we show in detail how this novel approach is implemented as an engine that is capable of identifying collaboration opportunities in real-time for real-world operations.

Third, we systematically test the performance of the TCG approach against that of an OD-based approach, using numerical experiments based on empirical data in a two-company collaboration scenario. We show how the TCG approach outperforms the OD-based approach in terms of the number of identified opportunities and the distance savings from the collaboration, with a computational cost that does not hinder its real-time applicability. Since in multi-company scenarios we would expect that costumer locations variability is likely to increase and proximity in routes is likely to be higher than OD pair proximity, our results appear promising also for potential multi-company collaborations.

Finally, thanks to performance enhancement techniques embedded in our opportunity identification engine, namely route approximation and route shape simplification, we demonstrate that operators can determine their preferred trade-off level between operational and computational performance.

The empirical case study in this paper demonstrates that enhanced operational efficiencies are possible using a purely data-based approach for the identification of collaboration opportunities for real-time collaborative logistics applications. Our work, therefore, provides logistics operators with an avenue to achieve further efficiency improvements through effective real-time horizontal collaboration.

Overall, our novel data-based approach and the case study demonstration contribute towards the real-world applicability of ISS frameworks, in recognition of the fact that ISS is attracting attention from both commercial operators and regulators, due to its strengths in respect to commercial data protection and competition law compliance. Indeed, the practical and commercial value in our solution that enables the competition law compliant applicability of an ISS framework was ensured by consultation of industry stakeholders. These included freight logistic companies Freja and Danske Fragtmaand operating competitively over the same area; Project 44 as trusted third-party hosting and operating the matching software (Reinau et al., 2021).

1.4 Structure of the paper

The rest of the paper is structured as follows. In section two, we present the core methodology of the collaboration engine, namely the TCG approach, including the definition of en-route collaboration opportunity and TCG, the opportunity identification constraints and the opportunity identification algorithm. In section three, we demonstrate how the TCG approach is implemented in practice as a real-time collaborative opportunity identification engine for real-world operations. In section four, we compare the performance of TCG implementations with various degrees of performance enhancement against the performance of OD based matching in a two-company collaboration scenario. And finally, in section five, we conclude by summarising the findings from our numerical experiments and highlighting avenues for future improvements of the methodology proposed in this paper.

2 Identification of collaboration opportunities using trailer capacity graphs

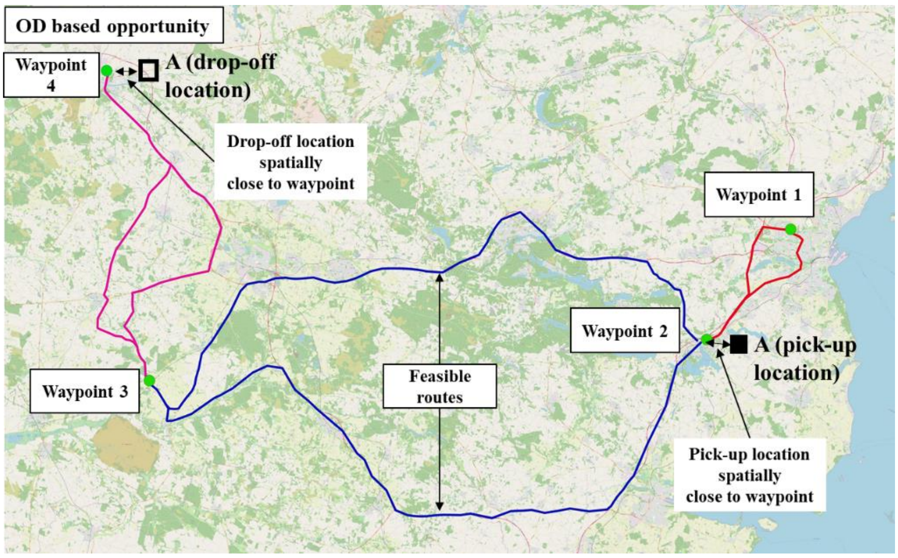
This section describes the definition of en-route collaboration opportunity and the details of the TCG approach. In this approach, a pair of a shipment and a trailer trip chain can be matched as a collaboration opportunity if the pair satisfies both fundamental and additional constraints. More specifically, fundamental constraints are ones that each collaboration opportunity must satisfy, including spatial and temporal proximity of a shipment and a trip chain, and a trailer capacity constraint. Additional constraints are ones arising from the collaborating companies' specific preferences and requirements, for example, types of goods allowed on certain trailers.

2.1 OD-based and en-route collaboration opportunities

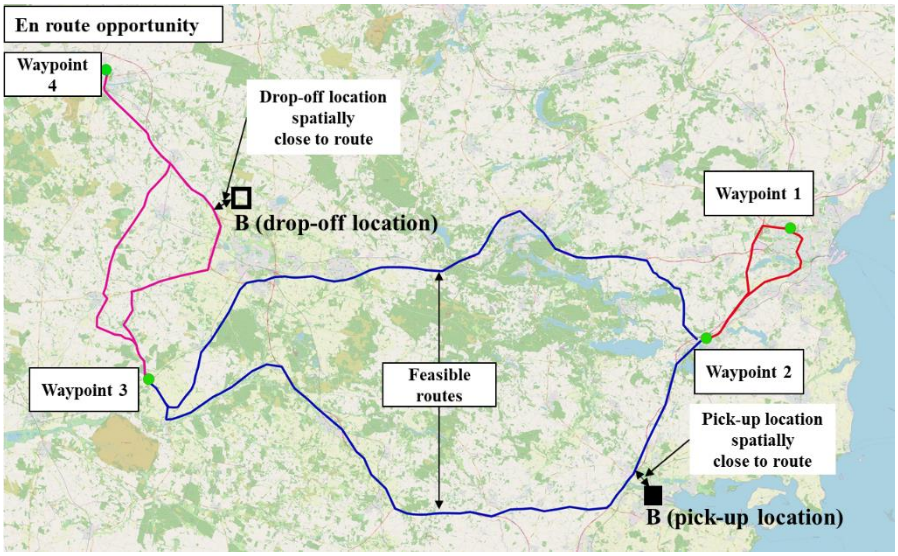
Dispatchers develop the initial plan of their trailer trips based on the information about current orders of their company. Typically, these initial plans are based on regular orders. Initial trailer trips and unassigned orders from all collaborating companies are then analysed to identify collaboration opportunities.

Opportunities can be identified by only looking at matches between ODs of initially planned trailer trips and ODs of shipment orders from the collaborating companies. This OD matching approach is usually adopted in data-based approaches to identify collaboration opportunities. Given an initial plan consisting of a trip chain, OD matching considers only shipments that have origin and destination near stops in this trip chain. Figure 1a shows this situation, where a shipment from a collaborating company can be picked up and dropped off near pre-planned waypoints of a trailer trip chain of another company.

In the present paper, we consider en-route opportunities, i.e. pick-ups and drop-offs at locations that are close to the pre-planned routes of a trailer, but not necessarily coincident with trip ends (i.e. waypoints) in the trip chain, see Figure 1b.



(a)



(b)

Figure 1 – OD-based opportunity (a) and en-route opportunity (b)

Compared with OD-based collaboration opportunities, en-route opportunities do not have strict restrictions on the origins and destinations of trailer trips from one collaborating company and shipments from another. Instead, an en-route opportunity can be identified as long as a pair of shipment order and trailer trip chain can meet a set of predefined spatial, temporal, capacity as well as potential additional constraints. In Figure 1a and Figure 1b, a trailer trip chain has four stopping locations (waypoints, travelling from waypoint 1 to waypoint 4). The curved lines between two consecutive waypoints are geometrical route shapes that the trailer can feasibly take. In Figure 1a, since the pick-up and drop-off locations of shipment A are spatially close to two waypoints, provided that temporal, capacity and additional constraints are satisfied, the trailer can pick up and drop off shipment A with sufficiently small additional travelling distance; hence, shipment A and the trailer trip chain forms an OD-based opportunity. In Figure 1b, since the pick-up and drop-off locations of shipment B are spatially close to the routes of the trips, even though the locations are far away from any waypoints of the trip chain, the trailer can still pick up and drop off shipment B with sufficiently small additional travelling distance. Therefore, shipment B and the trailer trip chain forms an en-route opportunity.

Since different companies are likely to have different customer bases, it is plausible to assume that there is a higher probability of en-route opportunities occurring than that of OD-based opportunities. Therefore, to improve the effectiveness of data-based approaches, there is room for existing approaches to be further enhanced by enabling the identification of en-route opportunities.

2.2 The concept definition of TCG

To identify a collaboration opportunity, we define a TCG, which is a directed multigraph that describes the spatiotemporal status of trailer capacity based on the planned trailer trip data. All the notation that we are using for Section 2 is provided in Table *1*.

Table 1 – TCG Notation

Variable	Definition
c, t	A company and a trailer, respectively.
TCG_c	The TCG of company c.
V_c	A set of vertices that represents the waypoints of all the trip chains of trailers belonging to company c.
$v_n^{t,c}$	The nth vertex (waypoint) of a trip chain of trailer t belonging to company c.
$l_n^{t,c}$	The geographical location of vertex $v_n^{t,c}$.
$\boldsymbol{h_n^{t,c}}$	The expected time window when trailer t is at $v_n^{t,c}$.
$k_n^{t,c}$	The available capacity of trailer t at $v_n^{t,c}$.
N^t	The total number of waypoints of a trailer t.
u_t^c	A trip chain of trailer t of company c.
T_c	The set of all trailers t belonging to company c.
E_c	The set of edges that represents all trailer trips of company c.
$e_{n-1,n}^{t,c}$	A directed edge of trailer t that connects $v_{n-1}^{t,c}$ and $v_n^{t,c}$.
$f_{n-1,n}^{t,c}$	The available trailer capacity over edge $e_{n-1,n}^{t,c}$.
$\boldsymbol{r_{n-1,n}^{t,c}}$	The set of maximum M geometrical route shapes ($r_{n-1,n}^{t,c,m}$) of edge $e_{n-1,n}^{t,c}$, $\boldsymbol{r_{n-1,n}^{t,c}} = \{r_{n-1,n}^{t,c,m} \mid m \in [1, M]\}$.
$\boldsymbol{\alpha_{n-1,n}^{t,c}}$	The vector of additional attributes of edge $e_{n-1,n}^{t,c}$ that correspond to additional constraints.
S_c	The set of all shipments belonging to company c.
I^c	The total number of shipments belonging to company c.
s_i^c	A shipment i belonging to company c.
o_i^c, d_i^c	The geographical pick-up and drop-off locations of shipment s_i^c, respectively.
$\boldsymbol{h_{i,o}^c}, \boldsymbol{h_{i,d}^c}$	The pick-up and drop-off time windows of shipment s_i^c, respectively.
g_i^c	The load of goods of s_i^c.
$\boldsymbol{\alpha_i^c}$	The vector of additional attributes of shipment s_i^c that correspond to additional constraints.
$e_{p-1,p}^{t,y}, e_{q-1,q}^{t,y}$	The pick-up edge and drop-off edge for trailer t from company y, respectively.
$\tau_{a,b}^{t.y}$	The estimated travel time for trailer t from the location a to location b.

$\epsilon^{t,y}_{q-1,q,i}$	The estimated additional travel time for trailer t to pick up s^x_i on edge $e^{t,y}_{q-1,q}$.
π^o_i	The estimated time needed for a trailer to handle shipment s^x_i at o^x_i.
$\boldsymbol{h}'^{t,y}_{\boldsymbol{l}}$	The new estimated time windows of trailer t assuming that t will pick up or drop off shipment i.
φ	A pre-specified distance difference threshold.
$D(\alpha, \beta)$	The diversion distance between a location α and an edge route β.

For each company, c, a TCG is defined as:

$$TCG_c = (V_c, E_c) \tag{1}$$

$$V_c = \left\{v^{t,c}_n\left(l^{t,c}_n, \boldsymbol{h}^{t,c}_{\boldsymbol{n}}, k^{t,c}_n\right) \middle| n \in [1, N^t], v^{t,c}_n \in u^c_t, t \in T_c\right\} \tag{2}$$

$$E_c = \left\{e^{t,c}_{n-1,n}\left(v^{t,c}_{n-1}, v^{t,c}_n, f^{t,c}_{n-1,n}, \boldsymbol{r}^{t,c}_{\boldsymbol{n-1,n}}, \boldsymbol{\alpha}^{t,c}_{\boldsymbol{n-1,n}}\right) \middle| n \in N^t, t \in T_c\right\} \tag{3}$$

More specifically, all the $v^{t,c}_n$ of trailer t are ordered primarily based on the $\boldsymbol{h}^{t,c}_{\boldsymbol{n}}$, from oldest to newest. In the case where the precision of $\boldsymbol{h}^{t,c}_{\boldsymbol{n}}$ is not sufficient to order all $v^{t,c}_n$, heuristics based on characteristics of different companies' data (e.g. the order of shipment and trip IDs) can be applied instead to order $v^{t,c}_n$. Moreover, a trailer t can only pick up and/or drop off shipment goods at each $v^{t,c}_n$; hence, the available capacity of a trailer will only change at $v^{t,c}_n$, and the available trailer capacity of an edge is set as that of its origin, namely:

$$f^{t,c}_{n-1,n} = k^{t,c}_{n-1} \tag{4}$$

An edge in the TCG is a trailer trip that connects a pair of consecutive vertices; therefore, the TCG essentially converts all the trips of a trailer into a single trip chain that consists of all the trips. Furthermore, the definition given by Equation 3 indicates that there can be multiple route shapes for an edge. Intuitively, an edge with multiple route shapes is equivalent to multiple edges connecting the same pair of vertices. Hence, a TCG is essentially a multigraph. The need for "multiple edges" is because it is likely that the trailer trip plan data does not include the actual route and route shape information for a trip; hence, the actual route and route shape need to be estimated. To increase the probability of correctly estimating the trailer route, and thus the probability of matching shipments and trips, multiple feasible route alternatives are estimated for an edge in the TCG. Nevertheless, a trailer can only choose one of the routes of an edge. Finally, even though vertices of different trailers may share the same location, Equation 3 defines that the vertices of different trailers are not connected.

Usually, multigraphs are applied for vehicle route planning with attributes, such as the travel time/ cost of an edge (Andelmin & Bartolini, 2019; D. S. W. Lai et al., 2016; Soriano et al., 2020). The main difference between TCG and the other multigraphs applied for transport planning is that an additional attribute, the geometrical route shape of each edge, is added in a TCG. A geometrical route shape of an edge describes what roads/ locations on the map that the trailer can pass by. Given the locations of a shipment and this geometrical route shape of an edge, the straight-line or driving distance between the shipment locations and a trailer trip can be easily computed; and this, in turn, makes it possible to evaluate if the pick-up or drop-off locations of shipments are spatially close enough to any trips of the trip chain. It is this that gives the TCG the advantage over conventional OD-based approaches in terms of identifying collaboration opportunities. That is, it allows the identification of en-route collaboration opportunities when the locations of the shipments are only sufficiently close to the geometrical shapes of reasonable routes of trailer trips (i.e. the TCG edges). In contrast, OD-based approaches can only identify opportunities where the shipment locations are in the proximity of the starting and endpoints of trailer trips.

2.3 Constraints for matching shipments using TCG

In order to match shipments with the trailer trips described by a TCG, shipments of company c are defined as:

$$S_c = \left\{ s_i^c\left(o_i^c, d_i^c, \boldsymbol{h_{i,o}^c}, \boldsymbol{h_{i,d}^c}, g_i^c, \boldsymbol{\alpha_i^c}\right) \middle| i \in I^c \right\} \tag{5}$$

Given the above definition of TCG and shipments, collaboration opportunities can be identified by matching shipments and TCG edges (trips) based on the fundamental and additional constraints. We first focus on matching based on fundamental constraints, namely, the spatial and temporal proximity of a shipment and a trip, and the trailer capacity constraint.

A shipment s_i^x, belonging to company x, and a trailer trip chain u_t^y, belonging to company y, form a collaboration opportunity if s_i^x, and at least a pair of edges of u_t^y, $\left(e_{p-1,p}^{t,y}, e_{q-1,q}^{t,y}\right)$, where $p \leq q$, satisfy the following constraints:

$$\begin{cases} \boldsymbol{h_{p-1}^{t,y}} + \tau_{p-1,o}^{t.y} \cap \boldsymbol{h_{i,o}^x} \neq \emptyset \\ \boldsymbol{h_{p-1}^{t,y}} + \tau_{p-1,o}^{t.y} + \tau_{o,d}^{t.y} + \pi_i^o \cap \boldsymbol{h_{i,d}^x} \neq \emptyset \quad if\ p = q \\ \boldsymbol{h_{q-1}^{t,y}} + \epsilon_{p-1,p}^{t,y} + \tau_{q-1,d}^{t.y} + \pi_i^o \cap \boldsymbol{h_{i,d}^x} \neq \emptyset \quad if\ p \neq q \end{cases} \tag{6}$$

$$\boldsymbol{h_l^{t,y}} \cap \boldsymbol{h_l^{\prime t,y}} \neq \emptyset, \forall l \geq p \tag{7}$$

$$\begin{cases} D\left(o_i^x, r_{p-1,p}^{t,y,a}\right) \leq \varphi \\ D\left(d_i^x, r_{q-1,q}^{t,y,b}\right) \leq \varphi \end{cases} , \exists j, k \in [1, M], p \leq q \tag{8}$$

$$a = b,\ \ if\ \ p = q \tag{9}$$

$$k_n^{t,y} \geq g_i^x \quad , \forall n \in [p-1, q-1] \tag{10}$$

Equation 6 defines the temporal constraint applied to a collaboration opportunity. Namely, the pick-up/drop-off time window of a shipment must overlap with the estimated time window of a trailer's arrival at the shipment location. More specifically, it is assumed that the trailer's arrival time window at the pick-up location is not affected by the handling time of the other shipments matched with the trailer at the same time since the other shipments are not yet accepted by the dispatchers. The constraint on the drop-off time window considers two situations. When the pick-up and drop-off edges are the same, the trailer's arrival time window is the sum of its departure time window of the edge, the estimated travel time from the start of the edge to the pick-up location, the estimated travel time from the pick-up location to the drop-off location, and the estimated shipment handling time at the pick-up location. When the pick-up and drop-off edges are different, the trailer's arrival time window is the sum of its departure time window of the drop-off edge, the estimated additional travel time for the trailer to pick up the shipment on the pick-up edge, the estimated travel time for the trailer from the start of the drop-off edge to the drop-off location, and the estimated shipment handling time at the pick-up location. Equation 7 specifies that trailers must still be able to comply with their original time plans if they fulfil the other companies' shipments. Equation 8 specifies that the diversion distance between the pick-up/drop-off location of a shipment and the pick-up/drop-off edge must be shorter than a distance threshold φ. Equation 9 specifies that when the pick-up edge and the drop-off edge is the same, the matched route must also be the same; thus, eliminating cases where a trailer and a shipment is matched even though the trailer can only either pick up or drop off the shipment. Equation 10 defines the trailer capacity constraint. Specifically, the available trailer capacity at any vertex between the start of the pick-up edge and the start of the drop-off edge must not be smaller than the load of a shipment.

Additional constraints are important to identify realistic opportunities. They must be satisfied by opportunities when they exist. However, they do not necessarily exist and also change from case to case. Our method is flexible in that when different numeric (e.g. maximum weight of a single shipment that a specific trailer can take) or non-numeric (e.g. type of goods allowed) additional constraints exist, they are used them as an another layer of filter when identifying opportunities.

As shown in Equation 10, the trailer capacity constraint essentially compares a numerical attribute of TCG edges (available trailer capacity) and the corresponding numerical attribute of a shipment (shipment load). Hence, as long as an additional constraint can be represented as a numerical attribute of an edge and a shipment, TCG can account for it by formulating it in the same way as Equation 10. In respect to non-numerical additional constraints, meanwhile, as long as they can be represented as attributes of edges and shipments, they can be formulated as comparing categorical or literal values, which can be generalised as follows:

$$\alpha_i^x \equiv \alpha_{n-1,n}^{t,y} \tag{11}$$

For example, if $\alpha_{n-1,n}^{t,y}$ of edge $e_{n-1,n}^{t,c}$ is [goods_type_allowed: food] and α_i^x of shipment i is [goods_type_shipment: food], then Equation (11) compares 'if goods_type_shipment= goods_type_allowed.

Figure 2 shows an intuitive example of matching shipments using a TCG, where there are three shipments from a company (shipment A, B and C) and the TCG of a trailer from another company that consists of three edges (straight arrow lines A, B and C). The first entry in the parentheses next to each pick-up location (filled square) or drop-off location (empty square) is the ID of the shipment, and the second entry is the load of the shipment. Each curved line is a geometrical route shape of the corresponding edge and the number over each curved line is the available capacity on the edge. Moreover, the temporal constraints are assumed to be satisfied by all shipments and edges. In this example, shipment A and edges A and C form an en-route opportunity, where edge A is the pick-up edge and edge C is the drop-off edge. Shipment B and the TCG cannot be matched since the capacity on edge C (five) is smaller than the load of shipment B (six). Shipment C and the TCG cannot be matched since its pick-up location and drop-off location are spatially close to different route shapes of the same edge (edge B).

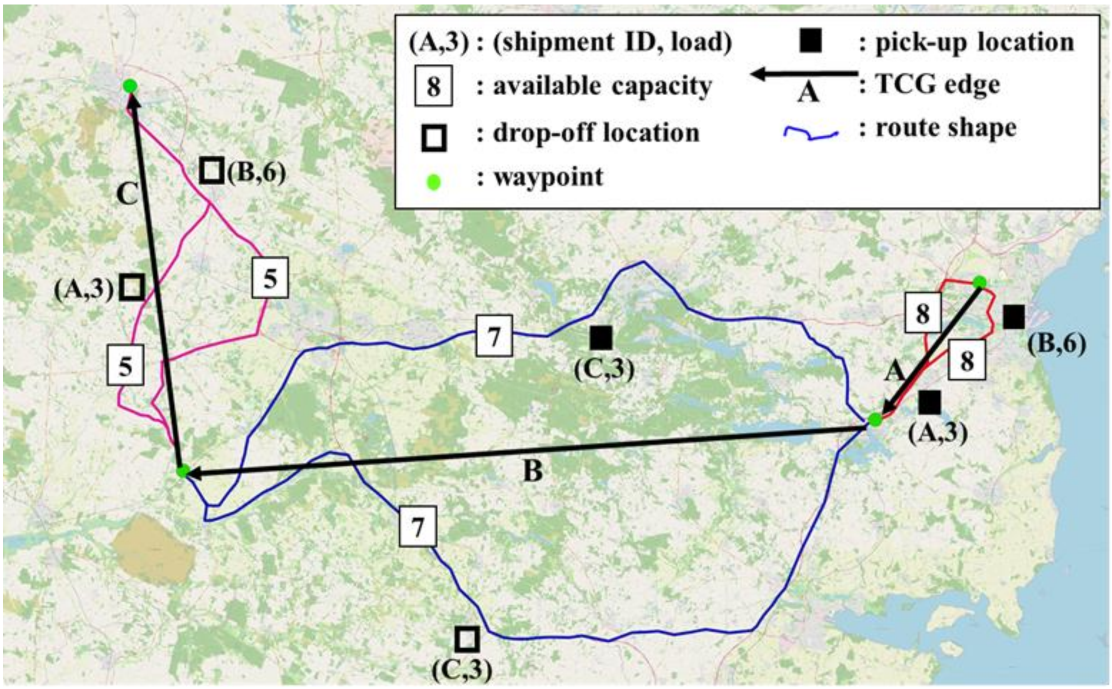


Figure 2- An example of matching shipments using the TCG

2.4 Opportunity identification algorithm

Based on the definition of TCG and the aforementioned matching constraints, an algorithm is proposed for efficiently identifying opportunities by matching shipments with the TCG; a pseudo code for this algorithm is described in Figure 3.

Firstly, the pre-processing step decomposes a shipment s_i into shipment origin, so_i $\left(o_i, \boldsymbol{h_{i,o}}, g_i, \boldsymbol{a_i}\right)$, and shipment destination, sd_i $\left(d_i, \boldsymbol{h_{i,d}}, g_i, \boldsymbol{a_i}\right)$. Secondly, in the matching step, for s_i, the algorithm finds a set of TCG edges ($e^t_{n-1,n}$) that occur on the same day as so_i and that satisfy the spatial constraint. Afterwards, for s_i, it finds edges ($e^t_{m-1,m}$) occurring on the same day as sd_i, belonging to the same trailer as $e^t_{n-1,n}$, happening at the same time or after $e^t_{n-1,n}$, and satisfying the spatial constraint. This step yields a set of edge pairs $E_i^2 = \left[\left(e^t_{n-1,n},\ e^t_{m-1,m}\right)\right]$ for each s_i. In step 3 (filtering step), $E_i^2 = \left[\left(e^t_{n-1,n},\ e^t_{m-1,m}\right)\right]$ for each s_i is first filtered based on the capacity and additional constraints and then further filtered based on temporal constraints. This step yields the final edge pairs for each s_i, and each pair of edges contains an edge for picking up s_i and an edge for dropping off s_i. Each pair of edges and s_i then forms an opportunity.

In general, this algorithm applies a multi-level filtering strategy to search for pairs of shipments and TCG edges that satisfy the fundamental and additional constraints. The main aim of this algorithm is to reduce the number of evaluations of the spatial and temporal constraints since these are among the most computationally expensive processes when searching for qualifying pairs of shipments and TCG edges. Firstly, the destination of a shipment is only matched spatially with TCG edges belonging to the trailers that have edges that are spatially matched with the shipment origin, which reduces the number of evaluations of the spatial constraint. Secondly, since the evaluation of the temporal constraint involves updating the time windows of trip edges after the pick-up edge, it is conducted in the last step so as to minimise the number of time window updates needed.

Step 1 (Pre-processing):

For each shipment s_i , decompose s_i into shipment origin so_i $\left(o_i \ , \boldsymbol{h_{i,o}}, g_i \ , \boldsymbol{a_i}\right)$ and shipment destination sd_i $\left(d_i \ , \boldsymbol{h_{i,d}}, g_i \ , \boldsymbol{a_i}\right)$

Step 2 (Matching): For each s_i ,

Step 2.1: match so_i with each edge of TCG ($e^t_{n-1,n}$) belonging to the other companies based on the following constraints:

$$\boldsymbol{h_{i,o}}.date = [\boldsymbol{h^t_{n-1}}, \boldsymbol{h^t_n}].date$$
$$dis\left(o_i \ , \boldsymbol{r^t_{n-1,n}}\right) \leq \varepsilon_d$$

Output: $E^1_i = \left[e^t_{n-1,n}\right]$

Step 2.2: match sd_i with each edge of TCG based on the following constraints:

$$m \geq n \qquad m, n \in N^{to}$$
$$\boldsymbol{h_{i,d}}.date = [\boldsymbol{h^t_{m-1}}, \boldsymbol{h^t_m}].date$$
$$dis\left(d_i \ , \boldsymbol{r^{t1}_{m-1,m}}\right) \leq \varepsilon_d$$

Output: $E^2_i = \left[\left(e^t_{n-1,n}, \ e^t_{m-1,m}\right)\right]$

Step 3 (Filtering):

Step 3.1: filter each pair $\left(e^t_{n-1,n}, \ e^t_{m-1,m}\right)$ in E^2_i based on the capacity constraint and additional constraints:

$$cv^t_p \geq g_i \quad , \forall p \in [n-1, m-1]$$
$$a_i \ \equiv a^t_{p-1,p}$$

Output: $E^3_i = \left[\left(e^t_{n-1,n}, \ e^t_{m-1,m}\right)\right]$

Step 3.2: filter each pair $\left(e^t_{n-1,n}, \ e^t_{m-1,m}\right)$ in E^3_i based on the temporal constraint:

$$\begin{cases} \boldsymbol{h^{t,y}_{p-1}} + \tau^{t.y}_{p-1,o} \cap \boldsymbol{h^x_{i,o}} \neq \emptyset & \\ \boldsymbol{h^{t,y}_{p-1}} + \tau^{t.y}_{p-1,o} + \tau^{t.y}_{o,d} + \pi^o_i \cap \boldsymbol{h^x_{i,d}} \neq \emptyset & if \ p = q \\ \boldsymbol{h^{t,y}_{q-1}} + \epsilon^{t,y}_{p-1,p} + \tau^{t.y}_{q-1,d} + \pi^o_i \cap \boldsymbol{h^x_{i,d}} \neq \emptyset & if \ p \neq q \end{cases}$$

Final output: $E^*_i = \left[\left(e^t_{n-1,n}, \ e^t_{m-1,m}\right)\right]$

Figure 3 - Pseudo code of the opportunity identification algorithm

3 The Opportunity Identification Engine

This section describes the development of the opportunity identification engine that implements the TCG approach. This implementation includes approximation techniques that, if triggered, speed up the identification of collaboration opportunities, albeit, as we shall see in the numerical experiments in Section 4, at the cost of identifying fewer opportunities.

3.1 Engine description for practical implementation

For the practical implementation of the approach, the opportunity identification engine is hosted by a trusted third party, which receives input data from collaborating companies and sends them notices when collaboration opportunities are identified. The collaborating companies need to agree to participate in horizontal collaboration as medium-term partners and to share the necessary data with the trusted third party.

To identify collaboration opportunities, for each time interval (for example, every five minutes), the engine creates in real-time a TCG based on each company's live trailer trip data and matches the live shipments of companies with the TCG edges of other companies. The matched shipment and trips will be presented to relevant companies as collaboration opportunities via notifications, which is the output of this engine. A collaboration agreement is achieved when a presented opportunity is accepted by dispatchers from both companies. This dispatchers' decision on presented opportunities, either acceptance or rejection, will be sent back to the engine. For accepted opportunities, the resultant exchange of the shipments between companies and the change of the trip plans are updated in the engine. Opportunities rejected, if still valid, will be sent to dispatchers in the next round since dispatchers may change their decisions and such opportunities may become acceptable.

To respect the confidentiality of trading data, firstly, participating companies have the freedom to choose which shipment and trailer trip data are shared with the engine. Hence, shipments and trailer trips considered as secrets will not be shared with the engine. In addition, when sending notices to dispatchers about the opportunities, only information about the opportunity that is essential for dispatchers to evaluate the opportunities is included. For example, the dispatcher from the owner of the shipment can only see the distance saving from the other company, without knowing how the trailer from the other company is routed. Also, the accurate shipment locations will only be provided to the dispatchers of the trailers after an opportunity is accepted.

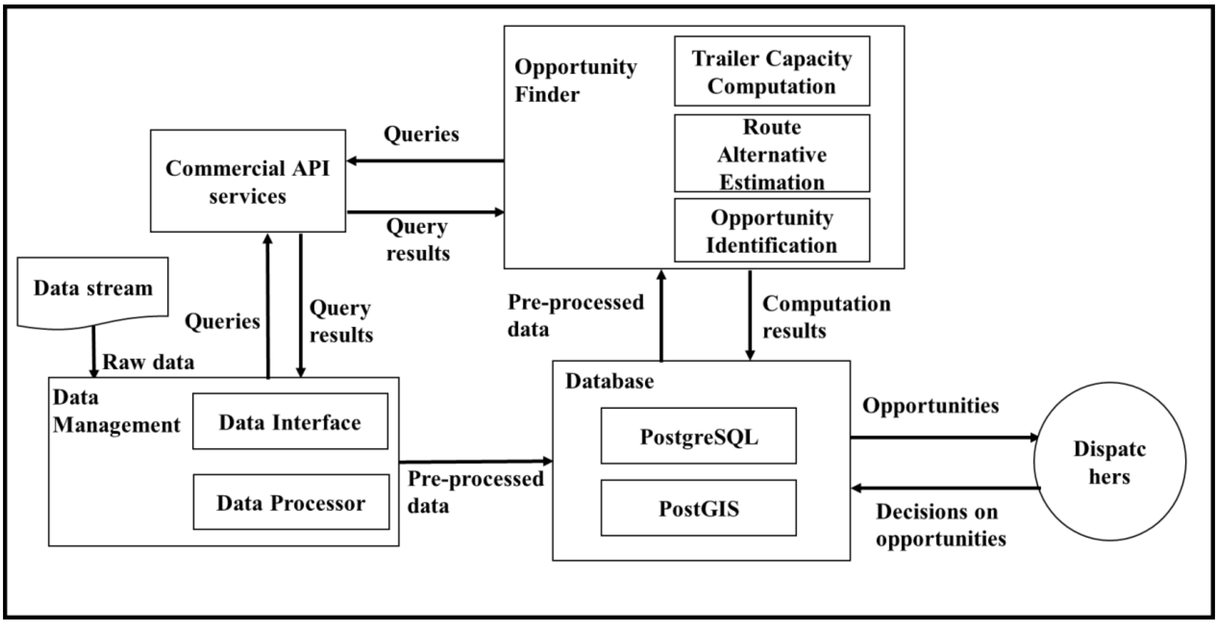


Figure 4 - Engine architecture

Figure describes the system architecture of this engine. In general, there are three main modules: a data management module, an opportunity finder module and a database. In the data management module, there are two sub-modules. The data interface receives live shipment and trailer trip data from different companies, and the data processor processes the raw data to be ready for the opportunity finder module to use. In the opportunity finder module, there are three sub-modules. The trailer capacity computation sub-module extracts trailers' en-route capacity from trip plan data. The route alternative estimation sub-module obtains the geographical information for the route alternatives of each edge in the TCG, Finally, the opportunity identification sub-module matches shipments and trailer trips from different companies using the TCG algorithm. The third main component of the engine is the database module. This stores all the data needed for identifying opportunities, including shipment data, trailer trip plan data and geographical data. This engine could in principal be applied for collaborations between several companies since this engine emphasises algorithm and operation efficiency, the details of which will be provided in the following sub-sections. However, as shown in Section 4, the demonstration of its effectiveness for identifying opportunities and computational performance focuses on a two-company collaboration scenario due to data resource limitation.

3.2 Data Management Module

The data management module handles live shipment and trailer trip data from different companies. The output of this module is shipment and trailer trip data that conforms to a consistent data format that can be efficiently consumed by the opportunity finder module. Thus, the data management module deals with the problem of diversity in data formats in the data received from the various collaborating companies.

Currently, our engine is being implemented using data from two large logistics companies operating over the entire national territory of a European country. The trusted third-party running the engine is a logistic data service operator. For reasons of commercial sensitivity, we will refer hereafter simply to Company 1 (C1) and Company 2 (C2). C1 and C2 have very different data schemas for the live data they provide. This significant difference in data schema, especially in the representation of a shipment and trailer trip, makes it difficult and inefficient simply to use the raw data to identify opportunities using a TCG. Hence, unified formats need to be designed if the engine is to handle different companies' shipment and trip data.

Table *2* shows the unified data formats designed for shipment data and trailer trip data. For shipments, the data covers the information on the pick-up and drop-off locations and expected time of a shipment action (pick-up or drop-off), and the information about its goods (e.g. goods weights). For trailer trips, the data covers the identity of the trailer, information about the waypoints of the trailer's trip, the change in the trailer capacity at each waypoint and the company running this trip.

More specifically, the change in the trailer capacity at a waypoint depends on the maximum capacity of the trailer and the goods loaded and unloaded at that waypoint, the details of which will be described in Section 3.3.1. Also, since, in practice, there are different measurements of shipment load, namely loading metres[2] and weight, the shipment load and trailer capacity are not always directly comparable. Hence, loading meters and weight are further unified into a *converted weight* measure:

$$\omega = MAX(\theta \times \gamma, w) \tag{12}$$

where ω is the converted weight; γ is the loading metre; w is the weight in kilograms, and θ is a coefficient representing maximum weight in kilogram per loading metre. Since the weight of a loading metre varies from shipment to shipment, θ is chosen as the maximum ratio. This helps to avoid the case where the weight of a shipment is underestimated or the available capacity of a trailer is overestimated, which will lead to matches that do not satisfy the trailer capacity constraint.

Table 2 - Unified shipment data and trailer data formats

Shipments		**Trailer Trips**	
Column name	**Definition**	**Column name**	**Definition**
Shipment ID	A unique ID of a trailer	Trailer ID	A unique ID of a trailer
Goods ID	The ID of the goods	Waypoint ID	A unique ID of a waypoint of the trailer
Converted weight	The shipment goods load measured by kilogram	Waypoint order	The order of the waypoint
Action	Either pick-up or drop-off	Position timestamp	The expected timestamp when trailer is at the waypoint
Action window	The expected window when the action happens	Load changed in converted weight	The trailer capacity measured by kilogram
Action location geometry	The geographical position of the waypoint	Waypoint geometry	The geographical position of the waypoint
Company	The company of the trailer	Waypoint address	The literal address of the waypoint
–	–	Company	The company of the trailer

[2] A loading metre is the standard unit of measurement for truck transportation, and refers to one metre of loading space of a truck's length (Group Legero, 2019). Here, the common European measure is used; hence, a loading metre is approximately 2.4 square meters.

3.3 Opportunity Finder Module

The opportunity finder module implements the TCG approach for finding collaboration opportunities. Figure describes the workflow of the opportunity finder module, the details of which are described in the following subsections.

In general, the trailer capacity computation sub-module first computes available trailer capacity at each vertex in the TCG based on the unified trailer data, which is formatted as vertices (waypoints). After that, the route alternative estimation sub-module will first attempt to find approximated routes between two consecutive vertices from the route shape table. If approximated routes are found, they will be fed into the opportunity identification sub-module. If this is not successful, it will invoke an external routing API to get the complete route shapes, which will be simplified by the route simplification algorithm. The simplified route shapes will then be combined with vertices with the capacity to complete the TCG. The TCG and the unified shipment data will then be fed into the opportunity identification sub-module, which seeks to match the shipments and the TCG based on the opportunity identification algorithm. In this paper, we only focus on the fundamental constraints. Nevertheless, as shown in Section 2.2 and Section 2.3, additional constraints can also be integrated into this engine if needed.

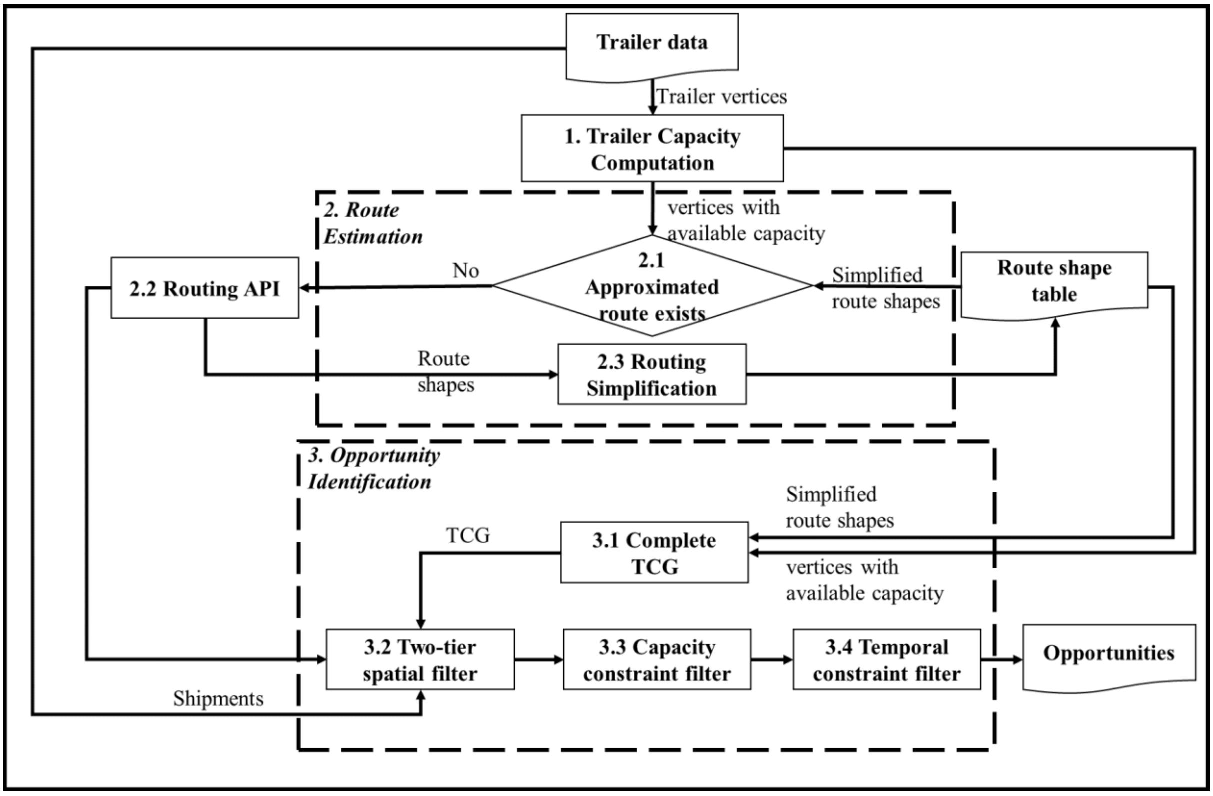


Figure 5 - Workflow of the opportunity finder module

3.3.1 Trailer capacity computation

Since companies C1 and C2 only record which shipments are loaded and unloaded at a waypoint of a trip, as shown in Table *2*, the unified trailer trip data only records the change of trailer load at each waypoint. Hence, to create a TCG, the trailer capacity available at each vertex needs to be computed.

The available capacity of a trailer at a waypoint can be computed based on the load change accumulated until the current waypoint. Since the load change is positive when a trailer picks up shipments and negative when it drops off shipments, the available capacity of trailer t belonging to company c at waypoint $v_n^{t,c}$ can be computed using Equation 13:

$$k_n^{t,c} = \rho_t^{max} - \sum_i \sigma_i^{t,c} \quad i \in \{i \leq n\} \tag{13}$$

where $\sigma_i^{t,c}$ is the load change of trailer t at its waypoint $v_i^{t,c}$; and ρ_t^{max} is the maximum capacity of trailer t.

3.3.2 Route alternative estimation

For estimating route alternatives of each edge in a TCG, external routing API services are used to search for M feasible and significantly different route alternatives between the start and end vertices of each edge, based on the geographical location (coordinates) of each vertex. It is worth noting that it is not guaranteed that M route alternatives can be found because, sometimes, feasible route alternatives significantly overlap with each other, and hence, there are not enough significantly different and feasible route alternatives. Moreover, since estimating route alternatives and measuring the distance between a shipment location and a route alternative are the two most time-consuming processes in this engine, the following two approaches are adopted to reduce the engine running time.

3.3.2.1 Trailer route approximation (RA)

When estimating routes of edges, it is reasonable to assume that if the start and end locations of two edges are each close to each other, it is likely that the routes of the two edges are largely overlapping. Furthermore, in order to measure the spatial distance between a shipment location and an edge with a route r, another route that largely overlaps with r is likely to produce a distance with a negligible error compared with the spatial distance threshold and the length of route r. Given those assumptions, for a pair of locations whose routes have not been searched before, if there is a close pair of locations that has already been searched, the routes of this searched pair can be directly used for the distance measurement for the unsearched pair.

Hence, to further reduce the number of routing API requests, this engine approximates route alternatives based on the proximity between the start and end locations of two edges. More specifically, before searching for the route alternatives for an edge using external APIs, the engine first searches for the closet pair of start and end vertices in the route shape table using the following rule:

$$\begin{cases} d\left(v_{n-1}, v_{n-1}^{*}\right) \leq \eta \\ d\left(v_{n}, v_{n}^{*}\right) \leq \eta \\ d\left(v_{n-1}, v_{n-1}^{*}\right) + d\left(v_{n}, v_{n}^{*}\right) \leq d\left(v_{n-1}, v_{n-1}'\right) + d\left(v_{n}, v_{n}'\right), \ \ \forall v_{n-1}', v_{n}' \end{cases} \tag{14}$$

where $d(v_1, v_2)$ is the straight-line distance between v_1 and v_2; v_{n-1} and v_n is the start and end vertex of an edge to search for route alternatives; v_{n-1}^{*} and v_n^{*} is the closet pair start and end vertex in the route shape table; v_{n-1}' and v_n' are a pair of start and end vertices in the route shape table; η is a distance threshold. If more than one pair of vertices is found in the route shape table, the pair with the shortest total distance to the edge being assessed is selected. The routes of the selected pair of vertices are used as the routes for edge (v_{n-1}, v_n).

3.3.2.2 Route shape simplification (RS)

A significant issue with the obtained route shapes of an edge is that their shapes usually consist of a large number of map coordinate points that seek to describe the curvature of roads in detail. These closely packed coordinate points will not significantly improve the accuracy of distance computation, but can significantly slow down the distance computation since the computational time increases with the complexity of the route shape.

To improve the efficiency when measuring distance between a shipment location and a route, a clustering algorithm, DBSCAN (Ester et al., 1996), is applied to simplify the shape of a route alternative in the route shape table. DBSCAN is a density-based clustering algorithm that can group points and their close neighbours. One of the main advantages of DBSCAN is that it is highly scalable and has a relatively low computational cost. Moreover, unlike most other clustering algorithms, it does not require the number of groups (clusters) to be pre-specified, which is difficult to determine because of the diversity of route shapes. Instead, DBSCAN requires a pre-specified parameter (λ) that defines the maximum distance between two samples for them to be considered as in the same cluster. A higher λ will mean that more coordinate points will be included within the same group, and hence the route shape will become simpler but less accurate. DBSCAN also requires a parameter that defines the minimum number of points necessary to form a cluster. This is set as one in this study so that all the points of a route can be classified into a cluster.

Here, the general idea is to start by grouping tightly concentrated coordinate points and select one point from that group (for our purposes the one closest to the centroid of the cluster) to represent the cluster. Then, the shape consisting of these representative points is the simplified shape of a route alternative, which will be used to compute the distance between shipment locations and routes. Intuitively, this simplification can be treated as zooming out a route shape.

Since this route shape simplification only runs once for each route, the computational cost of the DBSCAN algorithm can be further offset when a route is used multiple times for distance computation. Moreover, because the simplification of each route shape is independent, parallel computing technology is applied to simplify several route shapes simultaneously, which can dramatically reduce the running time of the route simplification task. Additionally, less complex route shapes can also reduce the time needed to store the route shapes in the route shape table, which further improves the system performance.

3.3.3 Opportunity identification

The opportunity identification sub-module implements the opportunity identification algorithm to find opportunities. In order to increase efficiency, a two-tier spatial filter is developed to reduce the computational cost of evaluating the spatial constraint. The estimated time windows, based on the estimated travel time of this filter, are further used to evaluate the temporal constraint.

3.3.3.1 Two-tier spatial filter

For spatial constraints, although the most accurate spatial distance measurement between a shipment location and a trailer route is the diversion distance ($D(\alpha, \beta)$), namely, the travel distance that a trailer diverts from the route to get to the shipment location and back to the route, the measurement of this diversion distance is costly since it requires using external routing APIs to calculate the shortest routes. In our system, therefore, a two-tier spatial filter is developed to accommodate the high computing cost of ascertaining diversion distances. The general idea of this two-tier filter is as follows: First, a less precise but much more efficient spatial constraint is used to filter out a potentially large number of shipment locations and trailer trip pairs. Then, actual diversion distances are obtained only for the remaining locations and pairs.

Hence, in the first tier, the straight-line distance between a shipment location and a route, which can be directly computed efficiently within the system, is applied as a spatial constraint. The straight-line distance is the minimum travelling distance between the shipment location and the route. If the straight-line distance is longer than a threshold φ, the actual travelling distance must be longer than φ. Hence, satisfying the straight-line distance constraint is a necessary condition for the spatial constraint to be satisfied.

In the second tier, for the shipment location and trailer trip pairs selected from the first tier, the diversion distance between the shipment location and trip is estimated using an external routing API. Specifically, since the main purpose of this system is to utilise residual trailer capacity in daily operations, it is necessary that the capacity-sharing operations should not interfere significantly with the original operations of the trailers. Hence, when estimating the diversion distance, the order of trailer travel through shipment locations, and the origin and destination of its route, are defined using the following rules:

(1) A trailer's trip section/ route must start from its origin waypoint and end at its destination waypoint;

(2) When a trip section passes the first tier filter with both the pick-up location and drop-off location of the same shipment, the pick-up location must be passed by the trailer before the drop-off location.

Moreover, since different routes have different travel distances, instead of defining φ as a fixed distance value, it is defined using a percentage of the travel distance of a route and a maximum distance threshold:

$$\varphi = MAX(\mu \times d_r, \varphi_{max}) \tag{15}$$

where μ is a pre-specified percentage; d_r is the travel distance of route r; and φ_{max} is the maximum distance threshold between a matching shipment location and a trailer's route.

3.3.3.2 Implementing the temporal constraint

In terms of the temporal constraint, different implementations of the constraints described by Equation (6) and Equation (7) can be applied based on the characteristics of the data provided. When the data provides detailed time windows of shipment actions and trailer trip waypoints, the temporal constraint can be evaluated based on the provided time windows, the estimated dwell and handling time for a shipment from other companies, and the estimated trailer travel time obtained from the external routing API that was used in the second tier of the spatial filter.

When an opportunity is accepted by both the shipment company and the trailer company, this acceptance will be reflected in the system. Firstly, the shared shipment will be moved from the original company to the company that accepts this shipment, and the corresponding trailer trip will be removed from its TCG if the trip is not needed anymore. Secondly, for the company that accepts the shipment, this shipment will be added into the plan of the trailer that conducts it, thus updating that trailer's TCG.

4 Numerical Experiments

In this section, experiments were conducted to compare the effectiveness and the computational performance of our proposed approach against a traditional OD-based approach based on real-world data collected from two road freight logistics companies, C1 and C2. The following two metrics were used to quantify the effectiveness of the approaches:

(1) The number of collaboration opportunities identified;

(2) The reduction in total distance travelled with respect to the non-collaborative case.

In this paper, we do not attempt to quantify the collaboration outcome in terms of system cost or company-specific economics benefits, because the data shared by the company does not include information about shipment order revenue and trailer travel cost information. We acknowledge that these quantities play an important role in accepting or rejecting an identified opportunity, and in quantifying the overall benefit that can be accrued through collaboration by each company. This is out of the scope of this paper, however, since our primary purpose is to demonstrate an effective data-based approach for opportunity identification that significantly improves on the currently prevalent OD-based method. Due to data limitations, the experiments are conducted based on a two-company collaboration. While the reason why this proposed approach is expected to work for collaboration with more than two companies is analysed in Section 4.5, we aim to further validate the effectiveness of this method in multi-company scenarios when more data becomes available in future work.

The computational performance of the TCG approach is also compared with the OD-based approach in terms of critical processing time (CPT), route shape processing time (RSPT) and the opportunity identification time (OIT). RSPT is the sum of the time cost of RA, RS and route shape storing in each engine run. OIT is the time cost of the first-tier spatial filter in each engine run. CPT is the sum of RSPT and OIT.

The results presented in this section include TCG performance in the absence and presence of route shape simplification (RS) and route approximation (RA). This makes it possible to show the trade-off between effectiveness metrics and processing time as the levels of RS and RA are varied.

4.1 Input data and assumptions

The data used for the experiments covers the period from 2 September 2019 to 16 September 2019, excluding weekends. The temporal resolution of the data provided by the two companies is a day. At the end of 16 September 2019, the two companies recorded a total of 10584 trailer trip sections and 14136 shipments, with a total trip travel distance of around 831844.26 kilometres. In the experiment, μ is set as 10% and φ_{max} is set as 25 kilometres. θ is set as 1850 kilograms per loading metre. The planning horizon is set as one day. To avoid small shipments being identified, which is not appealing for dispatchers, the minimum converted weight of shipment that is allowed to be identified is set as 5550 kilograms (equivalent to 3 loading metres). The default value of π_i^o is set as 10 minutes.

To reduce the complexity of the analysis and thus experiment running time, the results in this section are those when all the data during the experiment period is collected. Since the live data is collected in an accumulative manner, the running time in this section represents the maximum engine running time.

Traditional OD-based approaches use the same temporal and capacity constraints as the TCG approach; however, it only matches trailer trips' waypoints with shipments' ODs; hence, We set φ to 25 kilometres for the benchmark OD-based approach.

The total trip travel distance when C1 and C2 collaborate is the combined result of the opportunities identified and dispatchers' decisions about whether or not to accept those opportunities. Hence, an automated decision process for the acceptance and choices of opportunities was adopted in the engine in order to approximate the decision-making process of the dispatchers:

(1) All opportunities are considered on a first-come-first-served basis; which means that if a trailer's capacity has already been occupied by its own shipments and shared shipments, no further shared shipments will be accepted;

(2) When receiving a matched shipment, the dispatcher of the matched trailers will select the trailer with the minimum diversion distance and send the acceptance of this opportunity to the dispatcher of the matched shipment;

(3) When receiving the acceptance of the opportunity, the dispatcher of the matched shipment will compare the diversion distance of the selected trailer and the travel distance of his own trailer to fulfil the matched shipment, if the former is less than the latter, then he will accept this opportunity. Note that, if there are actions associated with other shipments at a location of a matched shipment, the travel distance of his own trailer for the matched shipment is considered zero, since the trailer must in any case travel to that location for those other shipments.

Furthermore, it should be remembered that the acceptance of some opportunities may result in other opportunities being cancelled, since if one opportunity is accepted, the originating company's corresponding trailer trip for that opportunity's matched shipment may be cancelled, and this would mean that any other opportunities associated with that cancelled trip could no longer be accepted. These opportunities that have such knock-on effects are excluded from this analysis for two reasons. Firstly, most of the opportunities (around 91%) do not belong to this category. Secondly, deciding whether or not to accept these opportunities requires balancing between the monetary benefit of accepting one opportunity and that of the opportunities that such acceptance would necessarily cancel. Such a judgement relies on information about shipment pricing and trailer travel cost that is unavailable for this research.

4.2 Results

Table *3* presents the effectiveness metrics and computational time performance results for the numerical experiments. These results are presented for the original TCG (without RA or RS), for the TCG approach with increasing levels of RA and RS, and for the OD-based approach. The metric in

Table *3* is the number of identified opportunities, the associated distance savings in kilometres compared to the non-collaborative case (in absolute and relative terms), RSPT, OIT and CPT.

The original TCG approach, i.e. the TCG approach without RS and RA, identifies more than three times as many opportunities as the OD-based approach. These deliver total distance savings of 17219 km compared to 7403 km for the OD-based approach. The original TCG approach therefore more than doubles the distance reduction achieved by the traditional OD-based approach. One should note that the fact that distance savings are a relatively small proportion of the total distance in the non-collaborative case is specific to the input data used in the numerical analyses. That is, the distance savings will always depend greatly on the two specific companies and the level of their spatiotemporal operational overlap over the period considered in a particular case study. The main indication of our result is that the original TCG approach is significantly more effective than the traditional OD-based approach in finding collaboration opportunities and thereby in generating distance savings.

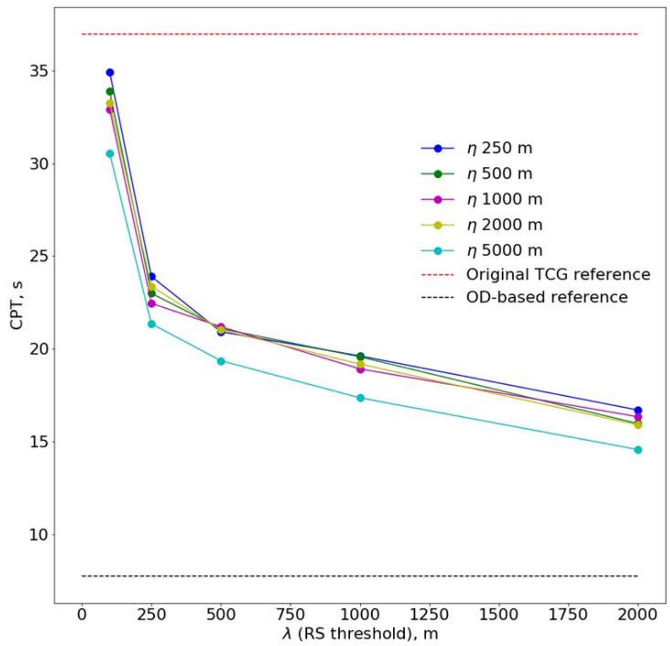
Table 3 Numerical Experiment Results

Approach	η Approximation threshold for RA	λ Cluster threshold for RS	Identified opportunity	Total distance saving (km)	Total distance saving %	Route shape processing time (Secs)	Opportunity identifying time (Secs)	Critical processing time (Secs)
Original TCG	–	–	2452	17219.18	2.07%	15.02	21.95	36.98
OD-based	–	–	765	7403.41	0.89%	1.05	6.7	7.75
TCG-based engine	250	100	2408	15968.73	1.92%	16.2	18.72	34.92
		250	2142	15224.61	1.83%	9.12	14.78	23.9
		500	1989	13755.52	1.65%	7.71	13.21	20.92
		1000	1718	13283.63	1.60%	7.46	12.16	19.62
		2000	850	9065.42	1.09%	7.21	9.49	16.7
	500	100	2222	15995.81	1.92%	15.5	18.38	33.88
		250	2149	16029.35	1.93%	8.61	14.38	22.99
		500	1991	14336.53	1.72%	7.29	13.8	21.09
		1000	1702	14630.6	1.76%	7.15	12.42	19.57
		2000	794	9065.42	1.09%	7.03	8.94	15.97
	1000	100	2208	14801.06	1.78%	14.79	18.13	32.92
		250	2179	16029.35	1.93%	8.12	14.33	22.45
		500	2002	13926.26	1.67%	7.45	13.74	21.19
		1000	1715	14144.61	1.70%	7.1	11.82	18.92
		2000	796	8844.41	1.06%	6.95	9.39	16.34
	2000	100	2375	15955.72	1.92%	14.61	18.66	33.27
		250	2181	14861.68	1.79%	8.38	14.99	23.37
		500	1988	14309.45	1.72%	7.3	13.72	21.02
		1000	1721	13679.1	1.64%	7.03	12.15	19.18
		2000	800	8848.12	1.06%	6.88	9.02	15.9
	5000	100	2210	15514.26	1.87%	11.72	18.82	30.54
		250	2084	15559.93	1.87%	6.93	14.42	21.35
		500	1998	13729.11	1.65%	5.9	13.46	19.36
		1000	1704	14168.79	1.70%	5.58	11.78	17.36
		2000	797	8900.11	1.07%	5.4	9.17	14.57

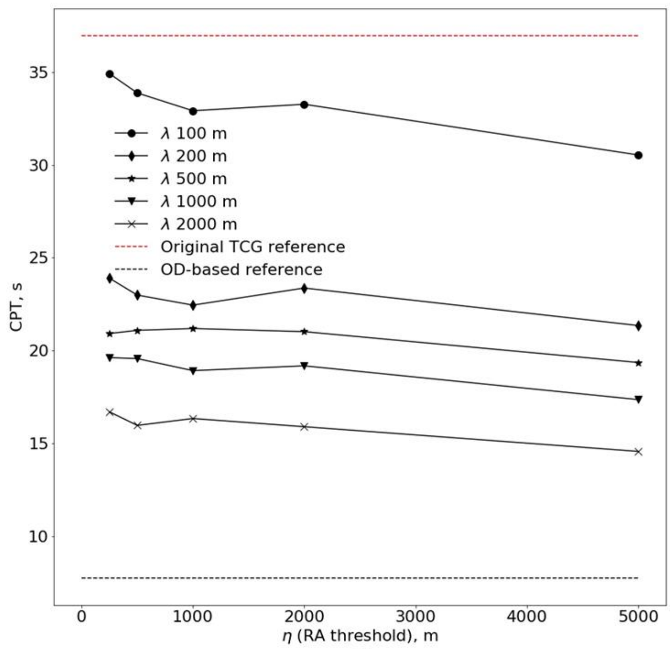
4.3 Reducing critical processing time with RA and RS

The superior effectiveness of the original TCG approach, however, comes at the expense of the time needed to process the data and identify the collaboration opportunities. The total processing time for identifying all the opportunities in the case study is 37s for the TCG and 8s for the OD-based approach. Both these values are compatible with the dispatcher planning operation of the two case study companies, but in applications with several companies, any gain in processing time is important. Accordingly, critical processing time (CPT) savings can be achieved for the TCG approach if RA and RS are enacted.

6a and 6b show that implementing RS leads to a more pronounced reduction in the critical processing time, but the rate of this reduction decreases as the cluster threshold for route simplification λ increases. The effect of RA in reducing the critical processing time is more moderate and occurs at a fairly constant rate as the approximation threshold η increases.



a)



b)

Figure 6 – Effect of RA and RS thresholds on the critical processing time

In summary, RA and RS reduce the processing time of the overall opportunity identification process. In the specific two-company case of this experiment, it might not be necessary to implement RA and RS because a CPT of 37s is still manageable for real-time operations. In cases where several companies participate in the collaboration, however, the CPT of original TCG scales up dramatically, and in such circumstances, the CPT reduction enabled by RS and RA can help to make TCG viable for real-time operations.

4.4 The effectiveness cost of RA and RS

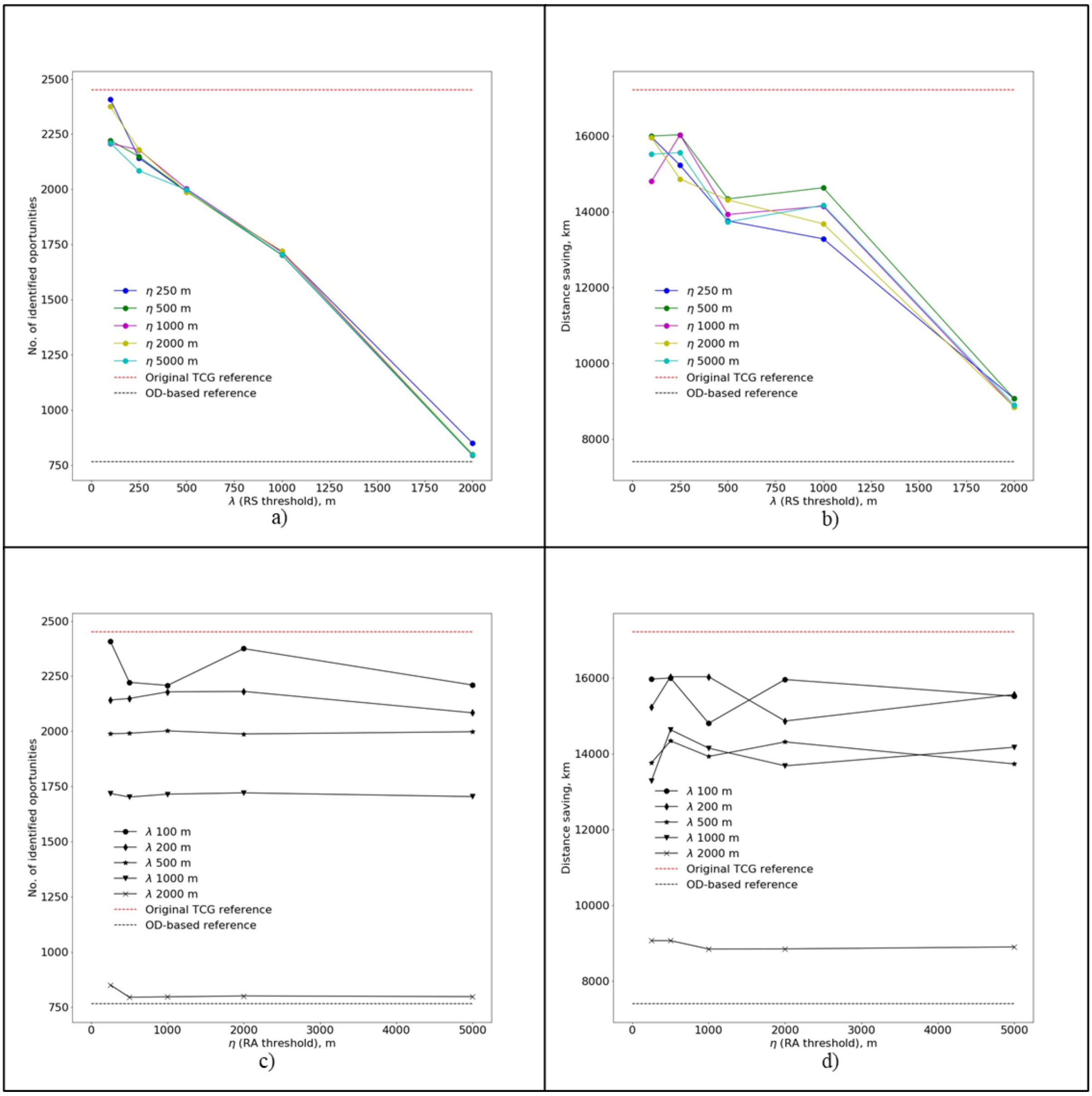


Figure 7 – Effect of RA and RS thresholds on opportunity identification and distance savings

When RA and RS are applied, the performance of the TCG changes depending on the level of the approximation threshold η and the cluster threshold λ. In particular, as 7a shows, the number of identified opportunities are particularly sensitive to the RS cluster threshold λ (indeed, the number of identified opportunities decreases (almost) linearly as λ increases). This reduction in opportunities leads to a decrease in distance savings, although the relation between λ and distance savings is not as markedly linear (Figure 7b). Interestingly, neither the number of opportunities nor distance savings are very sensitive to η (Figure 7c and Figure 7d).

The reason for the impact of λ and η on the number of identified opportunities is that they influence the measurement accuracy of the distance between the shipment locations and the route shapes of the trailer trips. As λ increases, the route shape of a trailer trip will be formed by a smaller number of coordinate points; and hence, the route shape will become less precise. This leads to less accurate distance measurement between shipment locations and route shapes, which results in some opportunities being missed and false opportunities being identified but rejected. On the other hand, when λ and η are relatively small, the approximated and simplified route shape tends to be similar to the original route shape; therefore, the error in the distance between shipment locations and route shapes is relatively small, which will not lead to a large number of missed opportunities and false opportunities.

Overall, Figure 7 demonstrates that RA and RS can generally preserve the TCG's advantage over the OD-based approach. This is because the approximated and simplified route shapes are still more detailed than just their starting and ending points. Hence, some shipments that are far from the start and end points can still be matched with trailer trips by using the approximated and simplified route shapes. Indeed, the OD-based approach can be viewed as a special case of the TCG-based engine, where the route shape of each trip section is aggressively simplified to consist of only the starting and ending points of the route shape. Hence, the OD-based approach is not capable of identifying the opportunities where shipments are sufficiently far from both the starting points and ending points of trailer trips. In contrast, the TCG allows new opportunities to be identified as long as the locations of the shipments are sufficiently close to reasonable routes of trailer trips.

4.5 Discussion

In essence, when the operating areas of different companies overlap sufficiently, and when there are more shipments with locations far from the starting or ending points of trips of other companies' trailers, but close to the routes of the trailer trips, the margin between the number of opportunities identified by this engine and those identified by the OD-based approach will increase. On the other hand, when more shipment locations are close to the starting or ending points of other companies' trailer trips, that margin will decrease. In multi-company scenarios we would expect that costumer locations variability is likely to increase and proximity in routes is likely to be higher than OD pair proximity. This suggests that our approach would be more effective in collaboration opportunity identification than OD-based approaches also in multi-company collaboration scenarios. Hence, notwithstanding that out numerical experiment focused on a two-company collaboration due to data availability, our results appear promising also for potential multi-company collaborations. However, as we point out in Section 5.2, the multi-company scenario will need to be validated empirically in future research based on data from more than two companies.

The results presented above show that the original TCG and the OD-based approaches represent two extremes. The original TCG identifies the most opportunities and thus generates the largest distance savings compared to the non-collaborative case, but this comes at the cost of the highest running time. On the contrary, the OD-based approach brings the smallest number of identified opportunities but costs the least time to run.

The results also show that to reduce the running time cost, trailer route data can be processed through RS and RA. Setting the level of RA and RS (i.e. choosing values of η and λ) when implementing the TCG approach leads to a trade-off between running time cost and the effectiveness of the collaboration engine. Figure 8 further highlights this trade-off. The figure shows the CPT savings relative to the original TCG case versus the number of opportunities identified. Each plotted point represents a TCG implementation. The OD-based approach is also represented in this graph. Due to the different impact profiles of η and λ, there are clearly combinations of these two parameters that are Pareto efficient. The result in Figure 8 implies that the implementation of RA and RS can provide the flexibility, depending on the specific application case, to identify a suitable compromise between the effectiveness of opportunity identification and the running time. Indeed, a multi-objective optimisation procedure can potentially be conducted for this engine to achieve the optimal compromise.

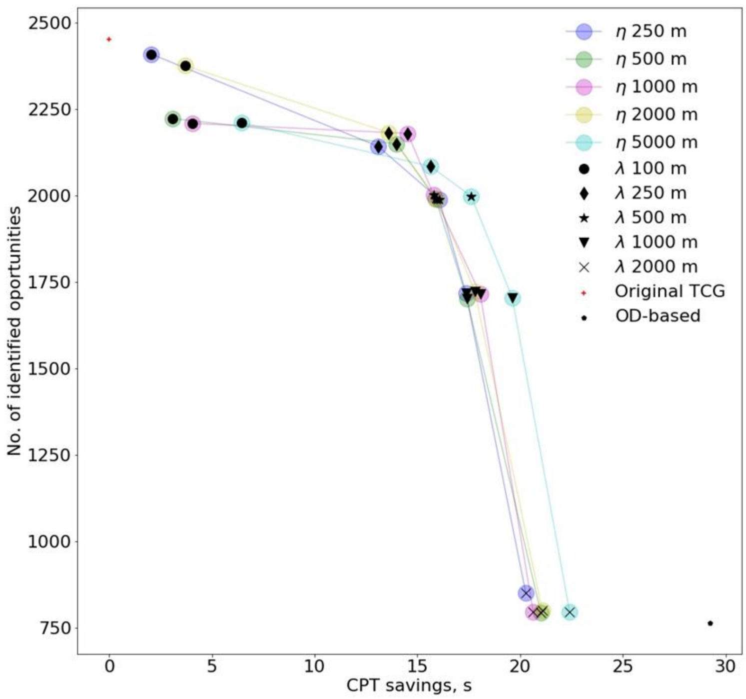


Figure 8 – CPT – Number of identified opportunities trade-off

In summary, the greater flexibility of the TCG approach, and its superiority in terms of identifying opportunities, make it overall a more effective tool for data-based collaboration identification than the traditional OD-based approach.

It should be noted that the absolute magnitude of the performance indicators resulting from our numerical experiments depends on the specific macroscopic characteristics of the empirical setting, e.g. the level of spatial and temporal overlap of operations between companies. Notwithstanding this, the comparison between the TCG and OD-based approach presented in this paper is fair since it is based on the same underlying input data. Furthermore, we deem that some general results of our comparison are valid more broadly than the specific empirical setting. Indeed, with any degree of operational overlap, the TCG approach will provide at least as high an operational performance as the OD-based approach. Nonetheless, the extent to which different macroscopic characteristics might affect the overall performance of the TCG approach remains an open question for further research.

5 Conclusions and Future Work

5.1 Conclusions

Real-time information communication and computation processes are deemed essential in the practical implementation of collaborative road freight logistics (Pan et al., 2019). The high computational cost of operations research methods for order sharing makes it difficult to implement in real-time collaborative logistics operations, however. On the contrary, data-based approaches provide a viable, albeit suboptimal, alternative that can enable the exploitation of real-time collaborative opportunities.

By and large, data-based collaborative engines rely on OD matching to identify collaboration opportunities. Yet OD matching fails to identify potential en-route collaboration opportunities if it is implemented for order sharing via information secure swapping (ISS) or shipment dispatching policies (SDP) procedures, in which a shipment of a collaborating company is matched with a trailer trip of another. This leads to significant unexploited collaboration potential.

This paper proposes a data-based engine for the identification of real-time collaborative opportunities that enables en-route matching by means of a different approach, the trailer capacity graph (TCG), overcoming the main limitation of OD-based collaborative engines. Furthermore, our engine based on the TCG approach is enhanced by route approximation and route shape simplification that can cut computational times if lower performance in respect to opportunity identification is deemed acceptable. Hence, the configuration of this engine in terms of trailer route approximation and route shape simplification enables a trade-off between computational performance and the effectiveness of opportunity identification. This trade-off implies that the engine can be flexibly tailored according to user preferences and requirements.

We test our TCG approach against the OD-based approach in numerical experiments based on real-world shipment and trailer trip data from two logistics companies. The results from the numerical experiments show that the TCG approach identifies a significantly larger number of opportunities than traditional OD-based matching because it enables en-route matching. The numerical experiments also demonstrate that when the effectiveness of opportunity identification of our approach is compromised in order to achieve better computational time, it can still generally preserve its advantage over traditional OD-based matching. While the effectiveness of our approach is validated in a two-company collaboration scenario, this approach is expected to work for multi-company collaboration scenario.

Overall, the analyses presented in this paper demonstrate that higher operational efficiencies are achievable in real-time with our approach that moves beyond traditional OD-matching for the identification of collaboration opportunities. This is a significant result as it enables a practice-ready solution that can be easily implemented for ISS. Ultimately, our solution enables collaboration frameworks, such as ISS, that can protect commercially sensitive data and are competition law compliant, to operate real-time with improved efficiencies. This can improve the attractiveness of such frameworks and support wider uptake of collaborative logistics among operators, at a time when it is paramount to quickly provide solutions for cleaner and more economically sustainable freight transport.

5.2 Limitations and future work

The magnitude of the performance indicators obtained in the numerical experiments is clearly dependent on the specific macroscopic characteristics of the empirical setting, e.g. the level of spatial and temporal overlap of operations between the two companies in the case study over the period covered by the input data. While this does not affect the fairness of our comparison between the TCG approach and our benchmark OD-based approach, the extent to which different macroscopic characteristics might affect the overall performance of the TCG approach should be investigated in future research. Our numerical experiments have shown that the approach has been tested, validated, and proved valuable for a two companies' collaboration but its application for multi-company collaboration will require further validation in future works.

The engine demonstrated in this paper filters opportunities based on fundamental constraints only. Hence, further research will also need to focus on the development of methods that filter opportunities based on additional constraints. In particular, a data-driven filter that models users' preferences is currently under development because finding suitable partners and opportunities rests not only on "tangible" factors (geographical, temporal, capacity and shipment type compatibilities) but also on latent "non-tangible" factors, e.g. trust between partners (Creemers et al., 2017) and operator knowledge, which are typically unobservable in databases (Ilie-Zudor et al., 2015).

In addition, currently the engine only ranks the identified opportunities based on the distance saving; hence, it does not have a sophisticated approach to select the best opportunities for the dispatchers when multiple shipments are matched with the same trailer or multiple trailers are matched with the same shipment. Therefore, a future research direction is to develop a real-time opportunity selection approach that can relieve the dispatchers from the complicated opportunities selection task.

As a final remark, we would like to point out that given the promising numerical results presented in this paper the engine, as described in Section 3.1, has been partially piloted in a two-company application for the two companies that provided the data for the numerical experiments, with Gatehouse Logistics acting as third-party collaboration engine host. The pilot trial, described in (Reinau et al., 2021), demonstrated that poor data quality available from the data management systems of the two company has proved a to be significant implementation challenge and for the real-world implementation of the engine data reporting and management need to be addressed for a beneficial implementation of the engine.

6 Acknowledgements

This work was supported by Innovation Fund Denmark [project number 6156-00001B]. We would also like to thank the two anonymous large logistics companies for providing the data.

7 References

Allen, J., Bektaş, T., Cherrett, T., Friday, A., McLeod, F., Piecyk, M., Piotrowska, M., & Zaltz Austwick, M. (2017). Enabling a freight traffic controller for collaborative multidrop urban logistics: Practical and theoretical challenges. *Transportation Research Record*, *2609*(1), 77–84. https://doi.org/10.3141/2609-09

Andelmin, J., & Bartolini, E. (2019). A multi-start local search heuristic for the Green Vehicle Routing Problem based on a multigraph reformulation. *Computers and Operations Research, 109,* 43–63. https://doi.org/10.1016/j.cor.2019.04.018

Audy, J. F., D'Amours, S., Lehoux, N., & Rönnqvist, M. (2010). Generic mechanisms for coordinating operations and sharing financial benefits in collaborative logistics. *IFIP Advances in Information and Communication Technology*, *336 AICT,* 537–544. https://doi.org/10.1007/978-3-642-15961-9_64

Caballini, C., Sacone, S., & Saeednia, M. (2016). Cooperation among truck carriers in seaport containerized transportation. *Transportation Research Part E: Logistics and Transportation Review, 93,* 38–56. https://doi.org/10.1016/j.tre.2016.05.007

Chabot, T., Bouchard, F., Legault-Michaud, A., Renaud, J., & Coelho, L. C. (2018). Service level, cost and environmental optimization of collaborative transportation. *Transportation Research Part E: Logistics and Transportation Review*, *110*, 1–14. https://doi.org/10.1016/j.tre.2017.11.008

Cleophas, C., Cottrill, C., Ehmke, J. F., & Tierney, K. (2019). Collaborative urban transportation: Recent advances in theory and practice. *European Journal of Operational Research, 273*(3), 801–816. https://doi.org/10.1016/j.ejor.2018.04.037

Creemers, S., Woumans, G., Boute, R., & Beliën, J. (2017). Tri-vizor uses an efficient algorithm to identify collaborative shipping opportunities. *Interfaces, 47*(3), 244–259. https://doi.org/10.1287/inte.2016.0878

Cruijssen, F., Bräysy, O., Dullaert, W., Fleuren, H., & Salomon, M. (2007). Joint route planning under varying market conditions. *International Journal of Physical Distribution and Logistics Management*, *37*(4), 287–304. https://doi.org/10.1108/09600030710752514

Cruijssen, F., Dullaert, W., & Fleuren, H. (2007). Horizontal cooperation in transport and logistics: A literature review. In *Transportation Journal* (Vol. 46, Issue 3, pp. 22–39). https://doi.org/10.2307/20713677

Deng, J. X. (2014). Matching model for scattering collaborative logistics business. *Applied Mechanics and Materials*, *556–562*, 6500–6505. https://doi.org/10.4028/www.scientific.net/AMM.556-562.6500

Ester, M., Kriegel, H.-P., Sander, J., & Xu, X. (1996). A Density-Based Algorithm for Discovering Clusters in Large Spatial Databases with Noise. *Proceedings of the 2nd International Conference on Knowledge Discovery and Data Mining*, *96*(34), 226–231. https://doi.org/https://dl.acm.org/doi/10.5555/3001460.3001507

Gansterer, M., & Hartl, R. F. (2016). Request evaluation strategies for carriers in auction-based collaborations. *OR Spectrum*, *38*(1), 3–23. https://doi.org/10.1007/s00291-015-0411-1

Gansterer, M., & Hartl, R. F. (2018). Collaborative vehicle routing: A survey. *European Journal of Operational Research*, *268*(1), 1–12. https://doi.org/10.1016/j.ejor.2017.10.023

Group Legero. (2019). *Loading Meter (LDM), Width & Height*. Legero Group. https://www.legerogroup.com/tools-extras/loading-meter-ldm-definition/

Guajardo, M., & Rönnqvist, M. (2016). A review on cost allocation methods in collaborative transportation. *International Transactions in Operational Research*, *23*(3), 371–392. https://doi.org/10.1111/itor.12205

Hernández, S., Peeta, S., & Kalafatas, G. (2011). A less-than-truckload carrier collaboration planning problem under dynamic capacities. *Transportation Research Part E: Logistics and Transportation Review*, *47*(6), 933–946. https://doi.org/10.1016/j.tre.2011.03.001

Houghtalen, L., Ergun, Ö., & Sokol, J. (2011). Designing mechanisms for the management of carrier alliances. *Transportation Science*, *45*(4), 465–482. https://doi.org/10.1287/trsc.1100.0358

Ilie-Zudor, E., Ekárt, A., Kemeny, Z., Buckingham, C., Welch, P., & Monostori, L. (2015). Advanced predictive-analysis-based decision support for collaborative logistics networks. *Supply Chain Management*, *20*(4), 369–388. https://doi.org/10.1108/SCM-10-2014-0323

Karam, A., Hussein, M., & Reinau, K. H. (2021). Analysis of the barriers to implementing horizontal collaborative transport using a hybrid fuzzy Delphi-AHP approach. *Journal of Cleaner Production*, *321*, 128943. https://doi.org/https://doi.org/10.1016/j.jclepro.2021.128943

Karam, A., Tsiulin, S., Reinau, K. H., & Eltawil, A. (2020). An Improved Two-Level Approach for the Collaborative Freight Delivery in Urban Areas. In *LISS2019* (pp. 775–787). https://doi.org/10.1007/978-981-15-5682-1_56

Kumar, A., & Anbanandam, R. (2019). Development of social sustainability index for freight transportation system. *Journal of Cleaner Production*, *210*, 77–92. https://doi.org/https://doi.org/10.1016/j.jclepro.2018.10.353

Lai, D. S. W., Caliskan Demirag, O., & Leung, J. M. Y. (2016). A tabu search heuristic for the heterogeneous vehicle routing problem on a multigraph. *Transportation Research Part E: Logistics and Transportation Review*, *86*, 32–52. https://doi.org/10.1016/j.tre.2015.12.001

Lai, M., Cai, X., & Hu, Q. (2017). An iterative auction for carrier collaboration in truckload pickup and delivery. *Transportation Research Part E: Logistics and Transportation Review*, *107*, 60–80. https://doi.org/10.1016/j.tre.2017.09.006

Leitner, R., Meizer, F., Prochazka, M., & Sihn, W. (2011). Structural concepts for horizontal cooperation to increase efficiency in logistics. *CIRP Journal of Manufacturing Science and Technology*, *4*(3), 332–337. https://doi.org/https://doi.org/10.1016/j.cirpj.2011.01.009

McKinnon, A., Browne, M., Whiteing, A., & Piecyk, M. (2015). *Green logistics: Improving the environmental sustainability of logistics*. Kogan Page Publishers.

Mostafa, A. K. A., Reinau, K. H., Daina, N., Luan, J., Østergaard, C. R., & Preisler, U. (2019). A preliminary analysis of main barriers to implement collaborative freight transportation using a DEMATEL method. In *7th IEEE International Conference on Advanced Logistics and Transport (IEEE ICALT 2019)*.

Palhazi Cuervo, D., Vanovermeire, C., & Sörensen, K. (2016). Determining collaborative profits in coalitions formed by two partners with varying characteristics. *Transportation Research Part C: Emerging Technologies*, *70*, 171–184. https://doi.org/https://doi.org/10.1016/j.trc.2015.12.011

Pan, S., Trentesaux, D., Ballot, E., & Huang, G. Q. (2019). Horizontal collaborative transport: survey of solutions and practical implementation issues. In *International Journal of Production Research* (Vol. 57, Issues 15–16, pp. 5340–5361). https://doi.org/10.1080/00207543.2019.1574040

Perboli, G., Rosano, M., & Gobbato, L. (2016). Decision support system for collaborative freight transportation management: A tool for mixing traditional and green logistics. *ILS 2016 - 6th International Conference on Information Systems, Logistics and Supply Chain*.

Reinau, K. H., Østergaard, C. R., Luan, J., Mostafa, A. K. A., Daina, N., Svendsen, O., Jensen, A. J. K., Nielsen, J. G., & Brodersen, J. (2021). *Collaborative logistics and fierce competition-the path to sustainable freight transport*. Institut for Byggeri, By og Miljø (BUILD), Aalborg Universitet.

Savelsbergh, M., & Van Woensel, T. (2016). City logistics: Challenges and opportunities. *Transportation Science*, *50*(2), 579–590. https://doi.org/10.1287/trsc.2016.0675

Soriano, A., Vidal, T., Gansterer, M., & Doerner, K. (2020). The vehicle routing problem with arrival time diversification on a multigraph. *European Journal of Operational Research*, *286*(2), 564–575. https://doi.org/10.1016/j.ejor.2020.03.061

Soysal, M., Bloemhof-Ruwaard, J. M., Haijema, R., & van der Vorst, J. G. A. J. (2018). Modeling a green inventory routing problem for perishable products with horizontal collaboration. *Computers and Operations Research*, *89*, 168–182. https://doi.org/10.1016/j.cor.2016.02.003

Suarez-Moreno, J. D., Garcia-Castillo, J., Castaneda-Velasquez, A. M., & Cardenas-Hurtado, A. F. (2019). Making horizontal collaboration among shippers feasible through the application of an ITS. *Proceedings of the 2nd Latin American Conference on Intelligent Transportation Systems, ITS LATAM 2019*, 1–6. https://doi.org/10.1109/ITSLATAM.2019.8721342

Vanovermeire, C., & Sörensen, K. (2014). Integration of the cost allocation in the optimization of collaborative bundling. *Transportation Research Part E: Logistics and Transportation Review*, *72*, 125–143. https://doi.org/https://doi.org/10.1016/j.tre.2014.09.009

Vaziri, S., Etebari, F., & Vahdani, B. (2019). Development and optimization of a horizontal carrier collaboration vehicle routing model with multi-commodity request allocation. *Journal of Cleaner Production*, *224*, 492–505. https://doi.org/10.1016/j.jclepro.2019.02.043

Verdonck, L., Caris, A. N., Ramaekers, K., & Janssens, G. K. (2013). Collaborative Logistics from the Perspective of Road Transportation Companies. In *Transport Reviews* (Vol. 33, Issue 6, pp. 700–719). https://doi.org/10.1080/01441647.2013.853706

Xu, X. F., Hao, J., Deng, Y. R., & Wang, Y. (2017). Design optimization of resource combination for collaborative logistics network under uncertainty. *Applied Soft Computing Journal*, *56*, 684–691. https://doi.org/10.1016/j.asoc.2016.07.036

Xu, X., He, Y., & Ji, Q. (2021). Collaborative logistics network: a new business mode in the platform economy. *International Journal of Logistics Research and Applications*, 1–23. https://doi.org/10.1080/13675567.2021.1926948