Leader based diffusion optimization model in transportation service procurement under heterogeneous drivers' collaboration network

Aghdas Badiee[1], Hamed Kalantari[1,*], Chefi Triki[2,3]

[1]Iran University of Science and Technology, Tehran, Iran
[2]Kent Business School, University of Kent, Canterbury, United Kingdom
[3]Department of Engineering for Innovation, University of Salento, Lecce, Italy

*Corresponding author: hamedkalantari@ind.iust.ac.ir

Abstract

One of the key issues in transportation systems is allocating shipping orders to the most appropriate drivers, in the shortest time and with the maximum profit. Many studies were carried out in the transportation service procurement process for allocating orders, but none of them considered driver-to-driver interactions and applied information diffusion concepts as a framework to maximize the profit, due to the lack of a framework to model the interactions. In this paper, we present a weighted drivers' collaboration network to form the interactions. To predict the behavior of drivers, a new community detection algorithm is developed to extract communities and their leaders, in terms of the speed and the power of receiving and diffusing shipping orders. Also, we present a profit maximization model using information diffusion power of community leaders. The results show the model is able to allocate shipping orders to the most suitable drivers, in the best possible time and with the highest profit. To demonstrate the performance of the developed algorithm, we present a numerical example. Finally, a case study is applied to solve the optimization problem. The results show that the optimized behavior of companies in allocating orders to drivers is based on their risk level, reputation, and the average number of their customers.

Keywords Transportation service procurement; Drivers' collaboration network; Diffusion optimization model; Overlapping community detection; Heterogeneous networks.

1. Introduction

Based on definitions in transportation science, the freight transportation service procurement (TSP) process involves all operations that are essential for acquiring freight services from an external source like planning, purchasing, logistics, payment, monitoring, and inventory management (Lafkihi et al., 2017). One of the important issues in TSP problem is the time management. In order to maximize customer satisfaction, companies try to allocate shipping orders to drivers in the shortest possible time. This poses two major challenges; reducing the profitability of shipping companies and customer dissatisfaction with the shipping experience due to driver inefficiency. First, allocating shipping orders to drivers in the shortest time will be cost a lot for companies. This includes marketing costs to find drivers, as well as allocating the orders to drivers with higher fares, both of which will reduce the company's profit. Second, allocating the orders to drivers in the shortest time can lead to the selection of unsuitable drivers. Drivers who, due to their lack of experience, will lead to customer dissatisfaction with the shipping experience. In this case, in addition to spending a lot of money to quickly allocate the orders to drivers, the company will not have enough knowledge about them, and this will lead to select an inefficient driver. Maybe, it is better for the company to allocate an order to an experienced driver. So, it is necessary to consider the two important factors of "cost" and "drivers appropriateness" along with the factor of "time" management by the companies. Numerous studies have examined the "time" and "cost" in the TSP problem, which will be discussed below, but none of them addresses the role of drivers in improving the performance of the transportation system.

Drivers with common characteristics (e.g., working on a common route, vehicle or trip, etc.) form an interconnected community, which generally share and diffuse information related to their community organically. A transportation manager can first, by carefully discovering the drivers' communities and extracting their requirements, share information or news appropriate to each community or make policies tailored to their needs. For example; for the urban freight drivers' community in Bandar Abbas city in Iran, whose main problem is the lack of diesel fuel stations and also due to the large volume of freight orders because of the presence of industries in the city and also the large number of drivers, has the problem of low and pre-determined fares rather than competitive fares. On the other hand, the communities of drivers in Babol city in Iran, which are mainly faced with the lack of drivers in the city, that led to high fares for shipping orders. Therefore, by discovering and identifying these communities, the transportation manager can design and implement tailored policies and decisions to solve problems of each community.

On the other hand, companies will be able to share their shipping order news among different communities for diffusing them, based on drivers' information diffusion capability within their communities. This will allow freight orders news to be shared at the lowest cost by companies and also organically by drivers' community within it. It is worth mentioning that, drivers of each community will receive information related to their community and diffuse it only within it. In addition to reducing the costs of allocating freight orders, this will increase the efficiency of allocation process and avoid wasting time to find suitable drivers, in the TSP problem.

Also, by identifying the communication power of drivers, it is possible to evaluate the level of communication and also experience of drivers in each community. Drivers who are more communicative usually receive the news earlier than others and are also more likely to spread the news to the others (and vice versa). On the other hand, identifying and analyzing drivers' communication power will help predict their behavior in communities. Drivers react differently when receiving shipping orders news. In a community, the probability that a driver will accept or also diffuse a specific shipping order is different from another driver. Finally, the speed of receiving shipping orders news by drivers are different. By identifying and predicting the behavior of community drivers, companies will be able to use drivers with high ability of the information diffusion when they require a rapid diffusion within a community, or in some cases drivers with a low level of communication or also drivers with the high possibility of accepting shipping orders. Thus, companies will be able to manage the information flow within each community based on their policies, the issue that has not been addressed yet in the TSP problem.

In this paper, we will try to predict the behavior of drivers in communities by modeling driver communications, detecting drivers' communities and extracting the ability of information diffusion of drivers. Also, by developing an optimization model will manage the "time", "cost" and "drivers appropriateness" simultaneously, based on the drivers' ability to diffuse information in the process of allocating shipping orders in TSP problem.

2. Literature Review

The body of literature includes all studies that address the concept of transportation service procurement in a transportation network. Many studies dealt with the subject of transportation service procurement as an essential function in various industries from miscellaneous aspects such as sustainability, auction mechanisms, robust optimization, business constraints, etc. (Xu and Huang, 2013; Zhang et al., 2015). However, different modes of transportation can be considered, including air, rail, road, and sea freight. In this paper, only road freight transportation service procurement is considered.

It is clear that shippers and carriers have conflicting decisions in road transportation service procurement (Yan et al., 2017). A shipper is the person or company who is normally the owner or supplier of commodities. Whereas, a carrier is a person/company that carries goods and is responsible for any possible damage to the goods during the shipping time. Carriers typically involve in some form of auction to bid on distinct lanes of interest. Shippers evaluate bids on lanes separately and then awards lanes to carriers based on numerous factors including price and commercial necessities. Therefore, the shipper seeks to reduce shipping costs and match demand and supply fairly and efficiently while the carrier seeks to increase the total revenue from shipping regarding load driver wages and costs of required fuel, equipment, and insurance. In other words, the content of the literature addressed the problem of road transportation procurement from carrier perspective (bid generation) which seeks to develop an optimal bidding strategy based on the operation cost analysis; and shipper perspective (winner determination problem) which aimed to determine the allocation of lanes to carriers give a set of bids.

Mathematical programming (MP) models are among the most widely used tools in transportation systems. In this regard, many studies have deal with road TSP through MP. For instance, several routing models were developed in order to maximize the carrier's profit from the auction (Wang and Xia, 2005; Yan et al., 2020) or to minimize the total transportation cost of using the company's accessible fleet and engaging external drivers (Triki, 2021). Other modes considered effect of regular and occasional drivers' behaviours on the routing costs for attaining driver fulfillment, and achieve improved customer satisfaction levels due to higher service accessibility for both close and distant customers (Abu Al Hla et al., 2019). Also, a two-stage robust formulation (Remli and Rekik, 2013) or a sampling-based two-stage stochastic programming model (Zhang et al., 2014) were proposed under uncertain shipment volumes.

Besides, a bi-objective branch-and-bound algorithm and eight variants of a multi-objective genetic algorithm were proposed to minimize the total cost and maximize service quality (Buer and Pankratz, 2010). In another study, for the bi-objective combinatorial auction model, a new heuristic solution method was proposed called Pareto neighborhood search (PNS) (Buer and Kopfer, 2014). Also, the carriers in (Yang and Huang, 2020) made the best quantity discount are able to get the greatest chance of being winners for frieght selection through proposing a mixed-integer nonlinear programming (MINP) model for the centralized planning problem (CPP) which is a kind of the winner determination problem (WDP) in transportation service procurement. Moreover, a deterministic model was proposed with a influential and noteworthy theoretical mechanism to overcome the imbalanced issue based on minimizing the service costs in WDP (Yang et al., 2019).

Some researchers incorporated bid generation model with vehicle routing methods and stochastic optimization for carriers to develop a multi-round combinatorial auction of TSP by mixed-integer programming (Chi, 2015). Moreover, a two-stage stochastic integer programming model for the winner determination problem to hedge the shipper's risk under shipment uncertainty was proposed by (Ma et al., 2010), and in a dynamic stochastic distribution context by (Feki et al., 2016).

Moreover, some stochastic bid price-based optimization models were developed (Hammami et al., 2020; Kuyzu et al., 2015; Rekik and Mellouli, 2012; Triki et al., 2014). In this regard, a single objective integer programming model that minimizes the total direct cost and hidden cost in a centralized procurement auction was proposed (Othmane et al., 2019). (Olcaytu and Kuyzu, 2019) contributed to the literature on price estimation in spot truckload markets by the development of load-specific truckload price-estimating methods and measuring their effect on the cost-effectiveness of carriers participating in the transportation network. In another study, bidders (carriers) generate their best bid (package) by using a bundled price to make the most of their utility and enhance the chance of winning the business via mixed integer programming for TSP (Yan et al., 2018b). An integrated multi-round combinatorial auction mechanism was developed for truckload TSP in which a winner determination problem was solved to assign profitable lanes to carriers (Kwon et al., 2005). However, some non-price factors can be identified as influential fields in such problems including on-time performance of the carrier, availability of proper equipment, familiarity with shipper's operation, and even billing accuracy (Jothi Basu et al., 2015b). For instance, we can refer some cases in which the mentioned non-price factors have been considered like a multi-objective model that simultaneously minimizes cost, and maximizes shipper's confidence and marketplace fairness (Ignatius et al., 2014) or sustainability of full truck-load transportation service procurement which was focused in transport logistics via the carrier assignment problem (Jothi Basu et al., 2015a). Also, the pickup and delivery problem with profits, time windows, and reserved requests were within a bid generation problem (Li et al., 2016).

Other studies in this area have considered pickup and delivery problem and service time window of selective request. For example, a mixed integer programming model with single objective function was presented in which delivery lead time is considered (Mamaghani et al., 2019). In addition, for the bi-objective full truckload TSP, a two-phase evolutionary algorithm was developed in which the total transportation costs and transit time were minimized at the same time (Zhang and Hu, 2019). In another study, the carrier's optimal bid generation problem that maximizes the profit was considered under combinatorial auctions (Lee et al., 2007). Also, the bid generation problem for heterogeneous truckload operations was solved by exact and heuristic solution methods (Hammami et al., 2019) and by exact solution with side constraints (Rekik et al., 2017).

There are several studies taking into account bi-level programming to describe relationships between shipper and carrier decisions. For instance, some multi-objective bilevel models which were solved by particle swarm optimization were proposed called as MOPSO in (Yan et al., 2017) and as DBMOPSO-WD in (Yan et al., 2018a). Also, a two-stage stochastic mixed-integer winner determination model was formulated in combinatorial reverse auctions under disruption risks (Qian et al., 2020). Similarly, a bi-objective integer programming model which is solved by a branch-and-bound algorithm was proposed in TSP auction process (Hu et al., 2016). Overall, those who are interested in reviewing the literature on full truckload TSP to identify the gaps from the perspectives of researchers and practitioners can refer to the review (Jothi Basu et al., 2015b).

The classification of the available related literature is reported in Table 1. As can be seen, in none of the above studies, the concepts of social network and the impact of driver's human role in a transportation network have been used to diffuse the information of shipping orders and maximize the profit of TSP system administrators. However, for the first time, the drivers' collaboration networks in two monoplex and multiplex perspectives presented by (Badiee et al., 2020). But, it did not consider the role of drivers in maximizing companies' profit, in TSP problem. In this paper, we will take into account the driver-to-driver interaction for maximizing the profit of companies by developing a new profit maximization model, based on information diffusion power among community leaders. On the other hand, most of the transportation network studies were based on the geographical information of facilities and none of them took into account the drivers' role in TSP problem. These are drivers who distribute the products within a transportation chain, create traffic flows and establish interactions with other actors. Therefore, studying the drivers' role is an important issue in transportation systems, which has been neglected in previous studies. Also, none of the studies used drivers' collaboration networks as frameworks or even a community detection algorithm as a solution method for optimizing information diffusion.

In this paper, we will evaluate the interactions between drivers in transportation systems, in the form of a weighted driver collaboration network, considering the belonging of drivers to more than one community. Also, driver-to-driver communications will be applied to improve the performance of the system. Means, by analyzing these communicatins, dense groups of drivers will be extracted in the form of communities, which can be used to make proper decisions and policies tailored to each community. By modeling driver-to-driver communications, a new algorithm will be developed for predicting the behavior of community drivers, in the form of the speed and the power of each community leader in receiving and diffusing shipping orders. Community leaders have a direct role in diffusing shipping order news within communities and allocating shipping orders to suitable drivers. Finally, by predicting the bahavior of communities, a new profit optimization model will be presented based on information diffusion power of community leaders. Using the optimization model, it will be possible to allocate shipping orders to the most suitable driver, in the shortest possible time and with the highest profit, for companies.

The paper organization is as follows: the problem description is presented in section 3. The methodology, consists of the “network designing”, and the “extended community detection algorithm”, is presented in section 4. Section 5 shows the numerical examples for evaluating the developed algorithm. A case study of multiplex weighted drivers’ collaboration network, the results analysis and the managerial insights are provided in section 6. Conclusion and feature researches are presented in section 7.

Table 1 Classification of the related literature on the mathematical programming models in the road freight transportation service procurement

Feature			Paper (s)
Perspective	Shipper		(Buer and Pankratz, 2010), (Ma et al., 2010), (Rekik and Mellouli, 2012), (Remli and Rekik, 2013), (Buer and Kopfer, 2014), (Ignatius et al., 2014), (Jothi Basu et al., 2015a), (Jothi Basu et al., 2015b), (Chi, 2015), (Hu et al., 2016), (Feki et al., 2016), (Yan et al., 2017), (Yan et al., 2018a), (Yan et al., 2018b), (Zhang and Hu, 2019), (Othmane et al., 2019), (Yang et al., 2019), (Yang and Huang, 2020), (Triki, 2021), **This paper**
	Carrier		(Wang and Xia, 2005), (Kwon et al., 2005), (Lee et al., 2007), (Zhang et al., 2014), (Triki et al., 2014), (Chi, 2015), (Kuyzu et al., 2015), (Jothi Basu et al., 2015b), (Li et al., 2016), (Rekik et al., 2017), (Yan et al., 2017), (Yan et al., 2018a), (Yan et al., 2018b), (Mamaghani et al., 2019), (Hammami et al., 2019), (Olcaytu and Kuyzu, 2019), (Qian et al., 2020), (Yan et al., 2020), (Hammami et al., 2020)
TSP Type	Bid Generation Problem		(Lee et al., 2007), (Triki et al., 2014), (Kuyzu et al., 2015), (Chi, 2015), (Jothi Basu et al., 2015b), (Li et al., 2016), (Yan et al., 2017), (Rekik et al., 2017), (Yan et al., 2018a), (Yan et al., 2018b), (Hammami et al., 2019), (Mamaghani et al., 2019), (Olcaytu and Kuyzu, 2019), (Yan et al., 2020), (Hammami et al., 2020)
	Winner Determination Problem		(Wang and Xia, 2005), (Kwon et al., 2005), (Buer and Pankratz, 2010), (Ma et al., 2010), (Rekik and Mellouli, 2012), (Remli and Rekik, 2013), (Buer and Kopfer, 2014), (Zhang et al., 2014), (Ignatius et al., 2014), (Jothi Basu et al., 2015a), (Jothi Basu et al., 2015b), (Hu et al., 2016), (Feki et al., 2016), (Yan et al., 2018a), (Zhang and Hu, 2019), (Yang et al., 2019), (Othmane et al., 2019), (Qian et al., 2020), (Yang and Huang, 2020), (Triki, 2021), **This paper**
Objective Function (s)	Economical	Cost	(Wang and Xia, 2005), (Kwon et al., 2005), (Buer and Pankratz, 2010), (Ma et al., 2010), (Rekik and Mellouli, 2012), (Remli and Rekik, 2013), (Buer and Kopfer, 2014), (Ignatius et al., 2014), (Zhang et al., 2014), (Chi, 2015), (Jothi Basu et al., 2015a), (Jothi Basu et al., 2015b), (Hu et al., 2016), (Feki et al., 2016), (Yan et al., 2017), (Yan et al., 2018a), (Yan et al., 2018b), (Zhang and Hu, 2019), (Yang et al., 2019), (Qian et al., 2020), (Yang and Huang, 2020), (Triki, 2021)
		Profit/ Revenue	(Lee et al., 2007), (Triki et al., 2014), (Kuyzu et al., 2015), (Jothi Basu et al., 2015b), (Li et al., 2016), (Yan et al., 2017), (Rekik et al., 2017), (Yan et al., 2018a), (Yan et al., 2018b), (Hammami et al., 2019), (Mamaghani et al., 2019), (Olcaytu and Kuyzu, 2019), (Yan et al., 2020), (Hammami et al., 2020), **This paper**
	Others		(Kwon et al., 2005), (Buer and Pankratz, 2010), (Rekik and Mellouli, 2012), (Ignatius et al., 2014), (Chi, 2015), (Jothi Basu et al., 2015a), (Jothi Basu et al., 2015b), (Hu et al., 2016), (Yan et al., 2017), (Yan et al., 2018b), (Zhang and Hu, 2019), (Othmane et al., 2019), (Mamaghani et al., 2019), (Qian et al., 2020)
Modeling	objective	Single	(Wang and Xia, 2005), (Kwon et al., 2005), (Lee et al., 2007), (Ma et al., 2010), (Rekik and Mellouli, 2012), (Remli and Rekik, 2013), (Zhang et al., 2014), (Triki et al., 2014), (Kuyzu et al., 2015), (Chi, 2015), (Jothi Basu et al., 2015a), (Jothi Basu et al., 2015b), (Li et al., 2016), (Feki et al., 2016), (Rekik et al., 2017), (Yan et al., 2018a), (Yan et al., 2018b), (Mamaghani et al., 2019), (Olcaytu and Kuyzu, 2019), (Hammami et al., 2019), (Yang et al., 2019), (Yan et al., 2020), (Qian et al., 2020), (Hammami et al., 2020), (Yang and Huang, 2020), (Triki, 2021), **This paper**
		Bi/Multi	(Buer and Pankratz, 2010), (Buer and Kopfer, 2014), (Ignatius et al., 2014), (Jothi Basu et al., 2015b), (Hu et al., 2016), (Yan et al., 2017), (Zhang and Hu, 2019), (Othmane et al., 2019)
	Stage	Single	(Wang and Xia, 2005), (Kwon et al., 2005), (Lee et al., 2007), (Buer and Pankratz, 2010), (Rekik and Mellouli, 2012), (Remli and Rekik, 2013), (Buer and Kopfer, 2014), (Ignatius et al., 2014), (Triki et al., 2014), (Kuyzu et al., 2015), (Chi, 2015), (Jothi Basu et al., 2015a), (Jothi Basu et al., 2015b), (Hu et al., 2016), (Li et al., 2016), (Feki et al., 2016), (Yan et al., 2017), (Rekik et al., 2017), (Yan et al., 2018a), (Yan et al., 2018b), (Zhang and Hu, 2019), (Yang et al., 2019), (Olcaytu and Kuyzu, 2019), (Othmane et al., 2019), (Mamaghani et al., 2019), (Hammami et al., 2019), (Yan et al., 2020), (Yang and Huang, 2020), (Triki, 2021), **This paper**
		Two/Multi	(Ma et al., 2010), (Zhang et al., 2014), (Jothi Basu et al., 2015b), (Qian et al., 2020), (Hammami et al., 2020)
	Level	Single	(Wang and Xia, 2005), (Kwon et al., 2005), (Lee et al., 2007), (Ma et al., 2010), (Buer and Pankratz, 2010), (Rekik and Mellouli, 2012), (Remli and Rekik, 2013), (Buer and Kopfer, 2014), (Triki et al., 2014), (Zhang et al., 2014), (Ignatius et al., 2014), (Chi, 2015), (Kuyzu et al., 2015), (Jothi Basu et al., 2015a), (Jothi Basu et al., 2015b), (Li et al., 2016), (Hu et al., 2016), (Rekik et al., 2017), (Zhang and Hu, 2019), (Othmane et al., 2019), (Mamaghani et al., 2019), (Olcaytu and Kuyzu, 2019), (Hammami et al., 2019), (Yang et al., 2019), (Yan et al., 2020), (Qian et al., 2020), (Hammami et al., 2020), (Yang and Huang, 2020), (Triki, 2021), **This paper**
		Bi/Multi	(Jothi Basu et al., 2015b), (Feki et al., 2016), (Yan et al., 2017), (Yan et al., 2018a), (Yan et al., 2018b)
	Type	Integer prog.	(Ma et al., 2010), (Rekik and Mellouli, 2012), (Buer and Kopfer, 2014), (Jothi Basu et al., 2015a), (Jothi Basu et al., 2015b), (Hu et al., 2016), (Zhang and Hu, 2019), (Othmane et al., 2019), (Triki, 2021)
		mixed-integer prog.	(Wang and Xia, 2005), (Kwon et al., 2005), (Lee et al., 2007), (Buer and Pankratz, 2010), (Remli and Rekik, 2013), (Triki et al., 2014), (Zhang et al., 2014), (Ignatius et al., 2014), (Kuyzu et al., 2015), (Chi, 2015), (Jothi Basu et al., 2015b), (Li et al., 2016), (Feki et al., 2016), (Li et al., 2016), (Rekik et al., 2017), (Yan et al., 2017), (Yan et al., 2018a), (Yan et al., 2018b), (Mamaghani et al., 2019), (Olcaytu and Kuyzu, 2019), (Yang et al., 2019), (Hammami et al., 2019), (Yan et al., 2020), (Qian et al., 2020), (Hammami et al., 2020)

		mixed-integer nonlinear prog.	(Yang and Huang, 2020), **This paper**
Conditions	Uncertain	Stochastic	(Ma et al., 2010), (Remli and Rekik, 2013), (Zhang et al., 2014), (Triki et al., 2014), (Chi, 2015), (Kuyzu et al., 2015), (Jothi Basu et al., 2015b), (Feki et al., 2016), (Olcaytu and Kuyzu, 2019), (Qian et al., 2020), (Hammami et al., 2020)
	Uncertain	Fuzzy	(Jothi Basu et al., 2015b), (Yan et al., 2017), (Yan et al., 2018a), (Yan et al., 2018b), **This paper**
	Deterministic		(Wang and Xia, 2005), (Kwon et al., 2005), (Lee et al., 2007), (Buer and Pankratz, 2010), (Rekik and Mellouli, 2012), (Buer and Kopfer, 2014), (Ignatius et al., 2014), (Jothi Basu et al., 2015a), (Jothi Basu et al., 2015b), (Hu et al., 2016), (Li et al., 2016), (Zhang and Hu, 2019), (Othmane et al., 2019), (Mamaghani et al., 2019), (Yang et al., 2019), (Hammami et al., 2019), (Yang and Huang, 2020), (Triki, 2021)
Case Study	Yes		(Chi, 2015), (Yan et al., 2017), (Mamaghani et al., 2019), (Yan et al., 2020), (Yang and Huang, 2020), (Triki, 2021), **This paper**
	No		(Wang and Xia, 2005), (Lee et al., 2007),(Ma et al., 2010), (Buer and Pankratz, 2010), (Rekik and Mellouli, 2012), (Remli and Rekik, 2013), (Buer and Kopfer, 2014), (Zhang et al., 2014), (Triki et al., 2014), (Ignatius et al., 2014), (Kuyzu et al., 2015), (Jothi Basu et al., 2015a), (Jothi Basu et al., 2015b), (Feki et al., 2016), (Li et al., 2016), (Hu et al., 2016), (Rekik et al., 2017), (Yan et al., 2018a), (Yan et al., 2018b), (Zhang and Hu, 2019), (Othmane et al., 2019), (Olcaytu and Kuyzu, 2019), (Hammami et al., 2019), (Yang et al., 2019), (Qian et al., 2020), (Hammami et al., 2020)
Solution Approach	Exact methods	Branch-and-bound alg.	(Lee et al., 2007), (Buer and Pankratz, 2010), (Jothi Basu et al., 2015b), (Hu et al., 2016), (Rekik et al., 2017), (Hammami et al., 2020)
	Exact methods	Constraint generation alg.	(Remli and Rekik, 2013), (Jothi Basu et al., 2015b)
	Exact methods	Linear relaxation	(Kwon et al., 2005), (Lee et al., 2007), (Yang et al., 2019), (Yang and Huang, 2020)
	Exact methods	Goal programming	(Rekik and Mellouli, 2012), (Ignatius et al., 2014), (Jothi Basu et al., 2015b), (Othmane et al., 2019)
	Exact methods	column generation and Lagrangian based tech.	(Lee et al., 2007)
	Exact methods	Social network-based community detection	**This paper**
	Heuristic Methods	Monte carlo approxi.	(Zhang et al., 2014), (Jothi Basu et al., 2015b)
	Heuristic Methods	Scenario-based approxi.	(Qian et al., 2020)
	Heuristic Methods	Dynamic simulation-based	(Feki et al., 2016),
	Heuristic Methods	Decomposition-based	(Triki, 2021)
	Heuristic Methods	Cost comparison-based	(Triki, 2021)
	Heuristic Methods	other	(Wang and Xia, 2005), (Lee et al., 2007), (Triki et al., 2014), (Chi, 2015), (Jothi Basu et al., 2015a), (Jothi Basu et al., 2015b), (Hammami et al., 2019), (Yang et al., 2019), (Yang and Huang, 2020)
	Meta- heuristic Methods	Iterative coordinate search alg.	(Ma et al., 2010), (Kuyzu et al., 2015), (Jothi Basu et al., 2015b), (Olcaytu and Kuyzu, 2019)
	Meta- heuristic Methods	MOPSO/ DBMOPSO-WD /PSO	(Chi, 2015), (Yan et al., 2017), (Yan et al., 2018a), (Yan et al., 2018b), (Yan et al., 2020)
	Meta- heuristic Methods	TPEA	(Zhang and Hu, 2019)
	Meta- heuristic Methods	Improved tabu search alg.	(Mamaghani et al., 2019)
	Meta- heuristic Methods	Adaptive large neighborhood search	(Li et al., 2016)
	Meta- heuristic Methods	Pareto Neighborhood search	(Buer and Kopfer, 2014), (Jothi Basu et al., 2015b)
	Meta- heuristic Methods	Genetic algorithm	(Buer and Pankratz, 2010), (Jothi Basu et al., 2015b), (Chi, 2015)

3. Problem description

The main research problems in this article are as follows:

- **Managing "time", "cost" and "drivers' appropriateness" simultaneously, using information diffusion ability of drivers**

In TSP problem, "cost" and "time" are key issues. The issue of "time" is important because companies, in order to satisfy existing customers or attract new ones, tend to allocate orders to drivers, in the shortest possible time. This will affect the revenue and therefore the profitability of the companies. So, companies have to incur higher costs (such as; advertising, marketing, as well as higher fares) in order to deliver orders to drivers more quickly. In this regard, companies will face the challenge of "cost" management. Means by controlling the allocation costs they have to ensure their profitability. However, reducing the cost of allocating orders will be possible for companies to the extent that the quality of the shipping experience is not seriously compromised. In other words, to avoid facing customer dissatisfaction with the shipping experience, orders need to be allocated to the right and suitable drivers. Therefore, the issue of "drivers' appropriateness" is important, which has not been addressed in previous studies. As a general rule, companies tend to allocate orders to the most appropriate drivers in the shortest possible time at the lowest cost. This issue will be highly depended on the policies of companies in attracting and retaining their customers.

In this paper, the main problem is managing the "time", "cost" and "drivers appropriateness" in allocating shipping orders to drivers commensurate with the position, policy and risk level of companies in TSP. To solve this problem, for the first time, we will develop an optimization model based on communication and the "information diffusion" ability among drivers. As mentioned earlier, numerous studies have been conducted on the development of profit maximization models in the TSP problem. But, none of them considers the role of driver-to-driver communication in improving the performance of the transportation system. By identifying the connections between drivers, one can use the information diffusion ability between them. One of the main topics which will be addressed in this article is identifying drivers' communication power and using it to diffuse information in order to help companies allocate the orders faster and cheaper. On the other hand, by identifying the "communication power" and "information diffusion ability" of drivers, the position of them in the transportation network will be determined, which will indicate the level of experience and efficiency and appropriateness of them. Because the driver who has more communication power in the transportation network is known as an experienced and efficient driver. Therefore, this driver has more capabilities in diffusion information. As a result, by using the driver's position companies will be able to diffuse shipping orders news or allocate the orders in accordance with their policies. Also, quantifying the information diffusion ability among drivers requires modeling the driver-to-driver communication and predicting their behavior, in the transportation network. The subject of "predicting the behavior of drivers' community", as the second major research problem, and "modeling driver-to-driver communication", as the third one, are presented as follows:

- **Predicting the behavior of drivers' community**

To predict the behavior of communities, it is necessary to measure and evaluate the two factors, named the "possibility of diffusing information" and the "speed of diffusing information" of drivers in communities. By identifying these two factors, the behavior of each driver when receiving a shipping order news, whether in terms of the possibility or speed of/in diffusing the shipping order news and also the possibility of accepting the order will be discovered and predicted. So, drivers will have different reactions after receiving a shipping order news. Therefore, in order to use drivers' capabilities in diffusing information, it is necessary to identify both "information diffusion possibility" and "information diffusion speed" factors.

Besides, the behavior of any driver is derived from the behavior of influential drivers or leaders within its community. It is required that, shipping order news be first shared among drivers who have the most communication and impact among the others, in a community. One can come to a conclusion that, identifying the community leaders is an important issue. In this article, the following two steps will be taken to solve the problem of predicting the behavior of drivers' community:

- Identifying the community leaders,
- Identifying and evaluating the "information diffusion possibility" and the "information diffusion speed" of drivers.

- **Driver-to-driver communication modeling**

To solve the above two problems, we first need to model the communication between drivers. For this purpose, we will develop a drivers' collaboration network that includes a communication graph of them. The developed drivers' collaboration network should be tailored to the features of the case study used in this paper, which should include different types of communication between drivers, relation intensity among drivers and consider the belonging of a driver to more than one community. Based on the developed network, first, it will be necessary to identify drivers' communities. For this purpose, an overlapping community detection algorithm will be developed. The following two steps will be taken in this regard:

- Developing a drivers' collaboration network
- Developing an overlapping community detection algorithm

Fig. 1 shows the research framework of this article:

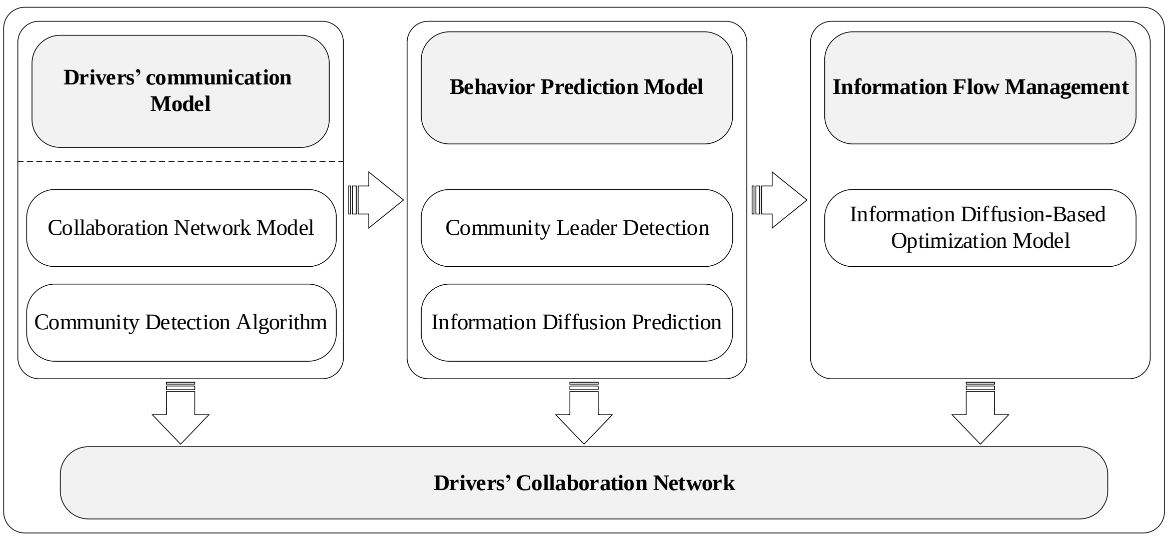

Fig. 1. Research framework

4. Methodology

Based on the research problem issues, in this section we present the optimization model based on information diffusion ability, the dirver-to-driver communication model and also the driver perdiction model.

4.1. Profit maximization model based on information diffusion among community leaders

The profit maximization problem is one of the goals of the transportation system analysis. The profit maximization problem can be examined from two perspectives. First, transportation companies are looking for a suitable driver to allocate freight shipping orders to them, with a minimum wage. Second, companies are trying to reduce advertising costs of allocating freights shipping orders to drivers. In addition to the profit maximization problem, allocating shipping orders to the most suitable drivers in the shortest possible time is another goal. Given that, this issue has a direct effect on the level of satisfaction of freight owners (customers), it can also affect the company's revenue and its total profit. Many studies worked on the profit optimizing problem, but none of them considered taking advantages from the capacities of drivers' collaboration networks. In this paper, we use these capabilities to prepare an optimization model, to solve the profit maximization problems. One of the capabilities is the ability to manage information flows within drivers’ collaboration network, to achieves the network administration goals. Due to the fact that drivers’ collaboration networks are based on relationships and interactions between drivers, this capability can be utilized for diffusing information of shipping orders, in these networks.

In the freight shipping allocation problem, companies are looking for the most suitable drivers to allocate shipping orders, minimize their advertising costs, and maximize their profit. Now, the question is how the companies can manage the information flow to optimize their costs and profits. In this section, a profit maximization model based on information diffusion among community leaders is presented to answer the question. Many studies presented optimization models for the diffusion optimization problem. In these studies, there is no difference among network components for publishing, accepting, or rejecting information. Due to the fact that people have different interests to accept, diffuse or reject information in social networks, overlooking these issues can take the model away from designing real conditions of the problem. In this section, we consider the information diffusion that is proper for each community, which is extracted from the network. This means that, information (e.g., shipping order news) which are not proper for the whole network, will not be diffused all over the network, i.e. the information will be only diffused into its related community.

We consider an important feature of communities, called community leaders, in the process of information diffusion among communities. In this process, information will only be shared with community leaders, for diffusing within the community. Also, we construct a mechanism for diffusing information within the community by these community leaders. So that, information will be diffused from the community leaders to other community members, called followers, which are at different distances from their leaders.

Fig. 2 shows the assumptions of the modeling in terms of the problem, objective, network, and edge type. The dark boxes show the specific assumptions considered in this paper.

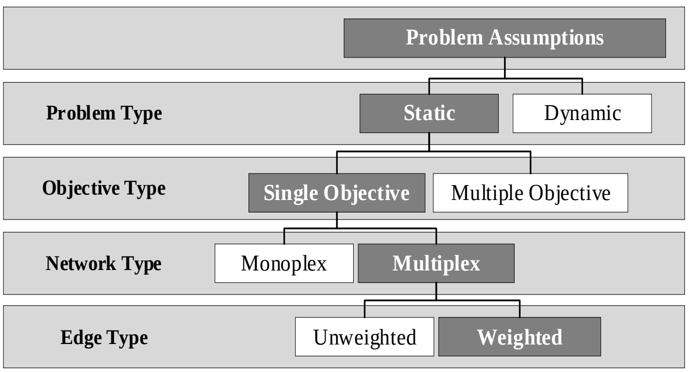

Fig. 2. Modeling assumptions categorization

The symbols and parameters of the proposed model are as follows:

- **Set:**

J_L A set of community leaders in the network.

- **Parameters:**

ρ Percentage of the wage received by a company.

γ_{c_k} Initial diffusion cost for selecting community leader c_k to share freight order news to him/her.

β_k Secondary diffusion cost for delaying in diffusing the shipping order k.

ε_{jc_k}	Threshold for accepting the shipping order k by the driver j in the community c_k.
W_k	Approved wage for shipping order k.
c_k	Community k to which the shipping order k is assigned.
d_{ij}	Distance between selected leader drivers i and j in the network.
X_{ijc_k}	A binary constant, which is 1 when there was a relationship between drivers i and j in the community c_k, o/w is 0.
$\tilde{N}_{ij}$	A possibility value which is a diffusion possibility (DP) from driver i to j.
$\tilde{R}_{jk}$	A possibility value, shows the appropriateness of a company's suggested wage of shipping order k, from the driver j's point of view.
$\mu(\tilde{R}_{jk})$	The membership value of appropriateness of accepting the shipping order k by driver j.
ε_{jk}	The driver j's threshold for accepting shipping order k, which $\varepsilon_{jk} \in [0,1]$.

- **Decision variables:**

Y_{jk}	A binary variable, which is 1 when driver j was assigned to the shipping order k, o/w is 0.
O_{ic_k}	A binary variable, which is 1 when driver i in community c_k was selected as a community leader, o/w is 0.
P_k	Company suggested wage to the shipping order k.

The proposed profit maximization model can be represented mathematically as follows:

$$Max \sum_k \sum_j \left[\rho + (1-\rho) * \left(1 - \frac{P_k}{W_k}\right)\right] Y_{jk} - \sum_k (\sum_i \gamma_{c_k} O_{ic_k} + \sum_j \beta_k \, d_{ij} \, Y_{jk} X_{ijc_k}) \quad (2)$$

Subject to:

$$(2Y_{jk} - 1)[(1 - \min_{i \in \{J_L\}} (1 - \tilde{N}_{ij}) X_{ijc_k}) \mu_{\tilde{R}_{jk}}(P_k) - \varepsilon_{jc_k}] > 0 \ \ \forall k \quad (3)$$

$$\sum_k Y_{jk} \leq 1 \quad \forall j \quad (4)$$

$$\sum_j Y_{jk} \leq 1 \quad \forall k \quad (5)$$

$$\sum_i O_{ic_k} = 1 \ \ \forall k \quad (6)$$

$$0.5 W_k \leq P_k \leq W_k \ \ \forall k \quad (7)$$

$$O_{jc_k}, Y_{jk} \ \in \{0,1\} \quad (8)$$

Equation (2) is the profit maximization objective function of a transportation company with two parts; total revenue and total cost. The first part is the total revenue of the company which contains the company's wages and its incomes, arising from saving costs that should be paid to volunteer drivers.

The second part of the objective function is the total cost, which contains initial cost and secondary cost. The initial cost is the cost of sharing shipping order news to community leaders to diffuse the orders into the communities. Each community has leaders which will be extracted from the community detection algorithm, in the next section. By sharing the freight shipping orders with the community leaders, the leaders will be prepared to diffuse the orders into the community drivers. Also, the cost of allocating shipping orders to volunteer drivers is considered as a secondary cost. Since any delay in allocating shipping orders to volunteer drivers will cause remarkable costs to the company, the distance between drivers to their community leaders is considered. So, drivers who are in a closer distance from the community leaders have a greater chance to get shipping orders allocated to them.

Equation (3) shows that, if the driver j was able to receive shipping order news from at least one driver in the community c_k, and accept the order with the wage P_k and with the possibility value $\mu_{\tilde{R}_{jk}}(P_k)$, which is greater than his "acceptance domain" (*AD*), then the shipping order k will be allocated to driver j in the community c_k. Also, $\tilde{R}_{jk}$ is a fuzzy value of the acceptance affordability of the driver j, in the community k. The *AD* value of each driver, which is a fraction of W_k, is calculated based on his minimum distances from his community leaders, as shown in Table 2. Indeed, community leaders, with *AD=0.9*, have the greatest *AD* value. Means community leaders have the minimum flexibility in accepting shipping orders than to the approved wage W_k, that we refer to as "famous drivers".

Also, the farther we go from the community leaders, the greater possibility the drivers have to accept the shipping orders. In the distance value 3, there exist drivers, that we call "beginner drivers", who are willing to accept shipping orders only with 0.4 of approved wage W_k. The above drivers' information will be extracted at the end of the community detection algorithm. The acceptance affordability is shown in the form of a fuzzy value, in Fig. 3.

Table 2 Acceptance domain of drivers

Minimum distance from community leaders	Acceptance domain
0	0.9
1	0.7
2	0.5
3	0.4

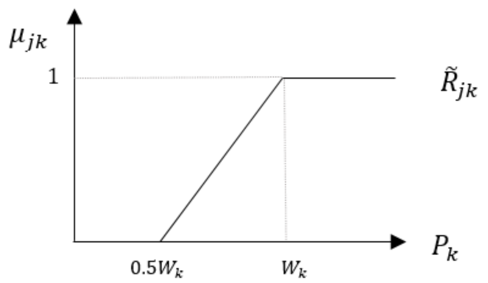

Fig. 3. The fuzzy value of the acceptance affordability $\tilde{R}_{jk}$

Equation (4) shows that the driver j can only accept one shipping order, in all communities. So, if a shipping order is allocated to the driver j in a community, he can not accept any other shipping order, simultaneously. Likewise, equation (5) depicts that the shipping order k can only be allocated to one driver.

Equation (6) ensures that only one leader node, among the community leaders, will be selected to diffuse the shipping order k, in community c_k. Finally, the range of P_k value and the binary nature of the decision variables are shown in equations (7) and (8), respectively.

4.2. Two-layer weighted drivers' collaboration network

The concepts of a drivers' collaboration network was presentedby Badiee et al., (2020), using the graph $G = (V, E)$, in which V is the set of drivers, as graph nodes, and E is the set of edges, showing the "similarities" between nodes. Also, they introduced explanations of the edges, such as: working on a shared vehicle, shared origin, or destination.

In this section, we use the graph $G = \langle V, E, L \rangle$ to present the two-layers weighted drivers' collaboration network, in which V is a set of nodes and E contains a set of $\langle u, v, l \rangle; u, v \in V, l \in L, u \neq v$, that is a type of connection between nodes u and v, in each layer. L depicts the set of layers based on the connection type between nodes. Fig. 4 shows the two-layers weighted drivers' collaboration network. In this figure, E represents how nodes u and v use a common vehicle, in layer 1, and have a common origin or destination in layer 2. In the above network, the "edge weight" shows the degree of connection between two nodes. This means, how much two nodes have interactions with each other in respect to their neighbor nodes. The edge weights of nodes u and v are shown in a binary form (RI_{uv}, RI_{vu}).

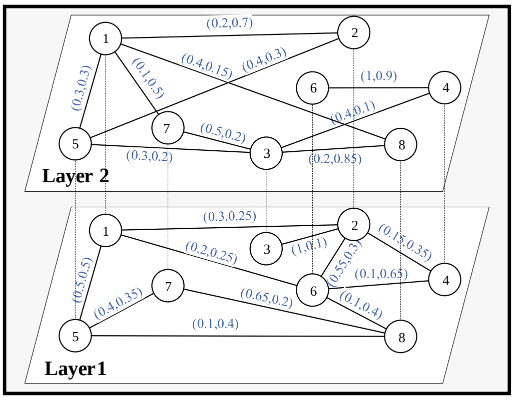

Fig. 4. Two-layers weighted drivers' collaboration network

So, to calculate the edge weights, we introduce the "relationship intensity (RI)" value as follows:

Definition 1: Relationship Intensity:

$$RI_{uv} = {R_{uv}}/{\sum_{i=1\ldots n} R_{ui}} \tag{1}$$

Where R_{ui} represents the total number of observed relationships (e.g., registered bill of lading dataset) between node u and other neighbor nodes, over a specified period of time, in layer l. Also, RI_{uv} is the relationship intensity value of node u with node v.

4.3. Extended overlapping community detection algorithm based on intra-layer expansion, inter-layer merging (Extended-OCDEM)

To solve the profit maximization problem, it is necessary to classify drivers into communities. Community detection algorithms are used to identify the hidden structure of a multiplex network. The multiplex network was presented (Tomasini,

2015) that were grouped into; edge colored, node colored, multiplicity, and temporary networks. The weighed drivers' collaboration network is presented in the form of multiplex network as an edge-colored network.

Based on what was stated in the problem description section, it became clear that the community detection algorithm is a tool to quantify the driver-to-driver communication modeling problem. In fact, what we are looking for in this problem is applying an algorithm to detect communities which should have the feature of overlapping communities, consider different connections between drivers and also different values of relation intensity between them. Therefore, by studying the literature of community detection methods, we select the OCDEM algorithm (Badiee et al., 2020) and develop it by improving the *NCOS* measure, using the relation intensity (*RI*) index for each two nodes, in the network. Some benefits of the OCDEM algorithm are as follows: the algorithm can determine the optimal number of the communities; considering the communities overlapping, so nodes can belong to more than one community; ability to solve edge-colored networks, and finally performing better than other algorithms. Also, the OCDEM algorithm has the ability to save the role of each node in community formation, at each step. In this algorithm, nodes first randomly form local communities. Then, by combining local communities with each other, intermediate communities are created, in each layer. Finally, by merging intermediate communities, final communities are created. Therefore, based on the participation information of each node in each stage of the community formation process, we extend the OCDEM algorithm to identify the leaders of each community. The results of the community leaders' identification algorithm are the main basis for solving the problem of predicting the behavior of drivers' communities.

Finally, since the development of this algorithm requires validations, we compared the extended algorithm with the OCDEM and other algorithms that have the properties of heterogeneous and overlapping.

4.3.1. Preliminaries

Consider the drivers' collaboration network as the graph $G = \langle V, E \rangle$ in which V is a set of nodes and E is a set of edges that represents the relationship between nodes u and v. The notations of the Extended-OCDEM algorithm are as follows:

G	A graph with node set V and edge set E.
$\langle V, E, L \rangle$	Two layers drivers' collaboration network.
A	Adjacency matrix.
$RI_{v_i v_j}$	Relation intensity between nodes v_i ,v_j.
LC^l	Local communities set in layer l.
$Com^l(v_i)$	Layer l's communities, that node v_i belongs to.
Ns^l	Layer l's nodes set.
Es^l	Layer l's sorted edges set.
$N^l(v_i)$	Neighbors of node v_i in layer l.
$N^l(C_k)$	Layer l's communities which have common nodes with community C_k.
u,v	The graph nodes.
M	Modularity measure.
β	Overlapping threshold.
α	Tunable parameter.
γ	adjacency weight of a community.
C_i^l	Community i in the layer l.
V_k	Nodes set of community k.
E_k^{in}	Inner edges of community k.
$\rho_{v_i C_k}$	A binary variable that takes value 1 if node v_i belong to C_k and 0 otherwise.
$\varphi_{v_k C_m}$	Is a binary variable, which is 1 when at least one edge exists between node v_k (of V_k) of C_k and v_m of $C_{m \neq k}$, not only v_k does not belong to C_m but also v_m does not belong to C_k. Otherwise, is 0.
C^l	Communities set in layer l.
$C^{'}$	Inter-layer communities set.
$NC(v_i)$	Communities which hold neighbors of node v_i.
C	The graph's final communities set.
C'^l	The refined communities set (in step 4), in layer l.
C^l	The inter-layer communities set (in step 3), in layer l.
LC^l	The local communities set (in step 2), in layer l.
LCS^{C_i}	The communities set, that make up the final community c_i.

$N^i(u_q)$	The number of repetitions of node u_q in the local communities set, that make up the final community c_i
S^i	The set of ordered nodes in c_i.
V_p^i	Set of volunteer nodes in c_i.
CC_p^i	The closeness centrality value for each volunteer nodes, in c_i.
I^i	The leaders set in c_i.
λ	The coverage coefficient parameter

The pseudo code of the Extended-OCDEM algorithm is as follows:

01. Sort nodes regards to their centrality measure.
02. $Ns^l = (u_1^l, u_2^l, \dots, u_n^l) \forall l = 1,2$
03. Set $Es^l = \phi$ and $Tabu\_list = \phi$.
04. **for** $l = 1{:}2$; **for** $i = 1{:}n$; **for** $j = i + 1{:}n$
05. **if** $u_j^l \notin Tabu\_list$; **if** $A(u_i^l, u_j^l) > 0$
06. $Es^l = Es^l \cup \{e_{ij}^l\}$
07. Update Tabu by putting u_j.
08. **end if**; **end if**; **end for**; **end for; end for**
09. **for** $l = 1{:}2$
10. Set $LC^l = \phi$.
11. **for** $E_i^l \in Es^l \forall (u,v) \in E_i^l$ and $Com(u) \cap Com(v) = \phi$
12. Set $LC_{Temp}^l = \{u, v\}$
13.$NC^l = N^l(u) \cap N^l(v)$
14. **if** $|NC^l| \geq 4$; **for** $node \in NC^l$; **if** $M(LC_{Temp}^l \cup node) > M(LC_{Temp}^l)$
15. Update LC_{Temp} by putting “node”.
16. **end if; end for; end if**
17. Update LC by adding LC_{Temp}.
18. **end for; end for**
19. **for** $l = 1{:}2$
20. $C'^l = LC^l$
21. Set $Tabu\_list = \phi$
22. **for** $C_i^l \in C'^l$; **if** $C_i^l \notin Tabu\_list$
23. $C''^l = C'^l \backslash C_i^l$
24. **for** $C_j^l \in C''^l$; **if** $C_j^l \notin Tabu\_list$
25. **if** $NCWOS(C_i^l, C_j^l) \geq \beta$
26. $Union(C_i^l) = \{C_i^l, C_j^l\}$
27. $C_i^l = Union(C_i^l) \cup C_i^l$
28. Update C''^l by removing $Union(C_i^l)$
29. Update $Tabu\_list$ by adding $Union(C_i^l)$.
30. **end if; end if; end for; end if**
31. **for** $u \in C_i^l$
32. Update $Com^l(u)$
33. **end for**
34. $C'^l = C'^l \backslash C_i^l$
35. **end for; end for**
36. **for** $l = 1{:}2$
37. Set $C'^l = C^l$
38. Set $Outlier = \phi$
39. **for** $u \in V and Com(u) = \phi$; **for** $C_i'^l \in NC(u)$
40. **if** $M_{C_i'^l + \{u\}} > M_{C_i'^l - \{u\}}$
41. Update $C_i'^l$ by adding node $\{u\}$.
42. Update $Com(u)$ by adding node $\{i\}$.
43. **end if; end for**
44. **if** $Com(u) = \phi$
45. Update $Outlier$ by adding $\{u\}$.
46. **end if; end for; end for**
47.Set $C\_init^{l_1} = \{C_1'^1, C_2'^1, \dots, C_n'^1\}$ and $C\_init^{l_2} = \{C_1'^2, C_2'^2, \dots, C_m'^2\}$.
48. Set $Tabu = \phi$
49. **for** $C_i \in C\_init^{l_1}$; **if** $C_i \notin Tabu$; **for** $C_j \in C\_init^{l_2}$
50. **if** $C_j \notin Tabu$; **if** $f_{C_i}^{C_j} > 0$
51. $Union(C_i) = \{C_i, C_j\}$
52. Update C_i by adding $Union(C_i)$.
53. Update $C\_init^{l_2}$ by removing $Union(C_i)$.
54. Update $Tabu$ by adding $Union(C_i)$.
55. **end if; end if; end for; end if**
56. **for** $u \in C_i$
57. update $Com(u)$
58. **end for**
59. update $C\_init^{l_1}$ by removing C_i.
60. **end for**
61. $LCS^{C_i} = \phi$ 62. **for** $C_i \in C(\forall i = 1, .., p)$
63. Find merged $C'^l s$. $(\forall l = 1,2)$
64. **for** $l = 1{:}2$; **for** $C_j'^l \in C'^l (\forall j = 1, .., n)$
65. Find merged $C^l s$.
66. **for** $C_m^l \in C^l (\forall m = 1, .., k)$
67. Find communities merged set LC^l
68. $LCS^{C_i} = \{LCS^{C_i}, LC^l\}$
69. **end for; end for; end for**
70. **for** $nodeu_q \in C_i (\forall q = 1, .., q')$
71. $N^i(u_q) = COUNT(LCS^{C_i}, u_q)$
72. **end for**
73.$S^i = \underset{q=1,..,q'}{SORT}(u_q), by N^i(u_q)$

The Extended-OCDEM algorithm can be described through six steps, as follows:

- **Step 1: Edges probing**

By calculating and sorting the “betweenness centrality” measure for all nodes, the volunteer nodes are addressed, using lines 1 to 8 of the algorithm.

- **Step 2: Extracting local communities**

In lines 9 to 18, using the extracted volunteer edges, the local communities of the network are formed. In this step, the modularity M used to create the local communities. As a result of this step, the local community set $LC = \{LC_1, LC_2, \dots, LC_k\}$ is obtained.

- **Step 3: Intra-layer communities Merging**

Using lines 19 to 35, we evaluate the local communities for being merged. In this step, the node connectivity overlapping score (NCOS) is used. To address the smallest relationship between two communities, we improve the NCOS measure, considering the *RI* value, presented in eq. (1), as a weighted adjacency value between two nodes. The extended NCOS measure will be able to address any relationship between two communities, including the members or non-member neighbor nodes relationship. The extended NCOS measure is as follows:

$$NCOS(C_i,C_j)=\gamma\left[\alpha\frac{|V_i\cap V_j|}{min\{|V_i|,|V_j|\}}+(1-\alpha)\frac{|E_i^{in}\cap E_j^{in}|}{min\{|E_i^{in}|,|E_j^{in}|\}}\right]$$
$$+(1-\gamma)\left[\frac{\sum_{v_k\in V_k}\left(\varphi_{v_kC_{\tilde{k}}}\right)+\sum_{v_{\tilde{k}}\in V_{\tilde{k}}}\left(\varphi_{v_{\tilde{k}}C_k}\right)}{min\{|V_i|,|V_j|\}+\sum_{v_k\in V_k}\left(\varphi_{v_kC_{\tilde{k}}}\right)}\right.$$
$$\left.+\frac{\sum_{v_i\in V_i}\sum_{v_j\in V_j}RI_{v_iv_j}\left(1-\rho_{v_iC_j}\rho_{v_jC_i}\right)}{min\{|E_i^{in}|,|E_j^{in}|\}+\sum_{v_i\in V_i}\sum_{v_j\in V_j}RI_{v_iv_j}\left(1-\rho_{v_iC_j}\rho_{v_jC_i}\right)}\right] \quad (9)$$

In eq. (9), the first part computes the overlapping score between two communities. The next part computes the weighted adjacency level between two communities. Also, the weights are adjusted by γ. As a result, if $NCOS(C_i^l,C_j^l)\geq\beta$, then these local communities should be merged. Also, the smaller the β is, the more communities will be combined.

- **Step 4: Communities Refining**

After merging local communities in Step 3, it can happen that some nodes in the network exist that have not been assigned to any community. Thus, using communities refining, lines 36 to 46 of the algorithm, these nodes will be evaluated and assigned to the communities. In this step, we calculate the modularity value of the community after having assigned each node to it and compare this value with the modularity value of the non-assignment case. So, if the newer modularity value is increased, then the node will be assigned to the community.

- **Step 5: Inter-layer communities aggregating**

We have communities in two layers that should be combined to extract the final communities. Lines 47 to 60 show the inter-layer community aggregating step. Badiee et al., (2020) proposed the “community fitness” measure, $f_{C_i}^{C_j}$, as follows:

Definition 2: Community fitness measure

The fitness value of community C_i is defined as follows:

$$f_{C_i}^{C_j}=\rho_{C_i+\{C_j\}}-\rho_{C_i-\{C_j\}} \quad (10)$$

Where C_i and C_j are two communities from different layers. $\rho_{C_i+\{C_j\}}$ is the redundancy value of community C_i when community C_j was added to it and $\rho_{C_i-\{C_j\}}$ is the redundancy value of the community C_i without adding the community C_j. If the fitness value is positive, the two communities C_i and C_j should be aggregated to create the final community. The redundancy value of community C can be calculated as bellow:

$$\rho_c=\sum_{(u,v)\in\overline{P_c}}\frac{|\{l:\exists(u,v,l)\in E\}|}{|L|\times|P_c|} \quad (11)$$

- **Step 6: Community leaders’ extraction**

We design the community leader extraction section, lines 61 to 73, as an extension of the OCDEM algorithm, to address each community leader in order to share shipping order news into the communities. As mentioned before, only the leaders who are able to diffuse the shipping order news into the communities. To evaluate which nodes are community leaders, we analyzed the Extended-OCDEM algorithm from step 5 to step 2. For each node u in the final community i, the participation degree of the node in the construction of the final community i (step 5), its reconstructed communities (step 4), its inter-layer communities (step 3) and finally its local communities (step 2) are calculated. Also, the sum of the above values makes the overall participation degree of the node u. Sorting nodes based on their overall participation degree, calculating the closeness centrality measure for each one and reordering them, will result community leaders.

4.3.2. Time complexity evaluation

If we split the algorithm into six steps, the overall complexity of the algorithm is the highest complexity of each of them. Table 3 shows the time complexity of each step when applied to solve an instance having n nodes and m edges.

Table 3 The complexity of the algorithm

Steps	Order	Descriptions
1	$O(2n+m)$	
2	$O(2md)$	d is the average degree of a node and $d\ll m$.

3	$O(2k_1^2)$	k_1 is the number of local communities and $k_1 < n$.
4	$O(2nd)$	
5	$O(k_2k_2')$	k_2 and k_2' are the numbers of final communities in two layers and $k_2, k_2' \ll n$.
6	$O(k_1k_2k_3)$	k_3 is the number of final communities and $k_1k_2k_3 \ll n^2$.
Overall	$O(2n^2)$	

5. Numerical example

To evaluate the performance of *Extended-OCDEM* algorithm, we use two methods called "*Layer Aggregation*" (Suthers et al., 2013), "*Ensemble Clustering*" (Fern and Brodley, 2004) and "*mux-licod*" which is applied on *LICOD* algorithm (Yakoubi and Kanawati, 2014). To extract the best results from these algorithms, the parameter value σ is set to 0.9. The OCDEM algorithm (Badiee et al., 2020) is also applied with the parameters $\alpha = 0.6$, $\beta = 0.35$ and $\gamma = 0.9$.

5.1. Real network datasets

We use six real two-layer networks[1] *BKOFF, BKFRAT, FFTWYT, DBLP-ppc, Friendfeed-ita* and "*Autonomous Systems network*[2]", as reported in Table 4, to evaluate the performance of the algorithms.

Table 4 Real networks details

ID	Networks	Number of nodes	Total number of edges
1	*BKOFF*	40	2,034
2	*BKFRAT*	58	5,240
3	*Autonomous Systems*	6,474	13,895
4	*FFTWYT*	6,407	74,762
5	*DBLP-ppc*	108,408	222,510
6	*Friendfeed-ita*	21,006	573,600

5.2. Evaluation measures on real network

To the best of our knowledge, supervised measures cannot be used to evaluate the performance of the algorithms, because multiplex networks with ground-truth partitions into communities do not exist. So, unsupervised estimation must be used, as bellow:

- ***EQ***: The *EQ* measure is as follow:

$$EQ = \frac{1}{2m}\sum_{c=1}^{k}\sum_{u,v\in C_c}\frac{1}{O_uO_v}\left(A_{uv} - \frac{d_ud_v}{2m}\right) \tag{12}$$

Where O_u is the number of communities that hold node *u*, d_u is the degree of node, also *m* is the number of edges. For a larger *EQ* value, the better community detection result will be obtained.

- **Redundancy**: The redundancy measure is as follow:

$$\rho_{Network} = \frac{1}{|C|}\sum_{c\in C}\rho_c \tag{13}$$

Where, ρ_c is calculated as *Eq*. 11. Using this measure, the redundancy value will be greater, when the edge contains more layers.

5.3. Evaluation results

Table 5 shows the comparison of the considered algorithms. In this table, the community size (|C|) of Extended-OCDEM is greater than the OCDEM algorithm, in all networks. This because of using the *RI* value in step 3, which obtains the smaller value for *NCOS* measure between two communities, and consequently merging less communities to each other.

The extended OCDEM algorithm had the highest value of the *EQ* measure. This means that the detected communities of this algorithm were separated the drivers with the highest accuracy. Also, compared to other algorithms, the drivers in these communities have a lot of internal communication and at the same time have the least amount of communication with drivers in other communities. Therefore, the extended OCDEM algorithm has been able to identify communities more accurately than other algorithms. Also, the extended OCDEM algorithm has a better performance in grouping drivers into communities than the OCDEM algorithm. The reason is improving the *NCOS* measure by using the *RI* index, in this algorithm. Using *RI* index allowed the extended algorithm to consider the smallest number of communications between drivers and the identified communities to be separated from each other more accurately.

[1] https://networkdata.ics.uci.edu/

[2] C. L. DuBois and P. Smyth, "UCI network data repository," *Web page http//networkdata. ics. uci. edu*, 2008.

Also, the *RI* index considered the relations of each community driver with other drivers in its neighboring communities, and this caused the identified communities to have relatively more overlapping than the OCDEM algorithm. As a result, the *Redundancy* measure for the extended OCDEM algorithm, in all real networks except the BKFRAT network, has higher values than the other algorithms. It should be noted that, the large overlapping of drivers in the communities for the extended OCDEM algorithm does not contrast with the high EQ value. This is because, the extended OCDEM algorithm controls the accuracy of community detection by grouping drivers into more communities than other algorithms. Therefore, the number of communities in the extended algorithm is more than the OCDEM algorithm as well as the others.

Table 5 The comparison of the algorithms

Network ID	Layer Aggregation			Ensemble Clustering			mux-licod			OCDEM			Extended-OCDEM		
	$\|C\|$	EQ	ρ	$\|C\|$	EQ	ρ	$\|C\|$	EQ	ρ	$\|C\|$	EQ	ρ	$\|C\|$	EQ	ρ
1	7	0.36	0.09	6	0.32	0.15	7	0.36	0.09	8	0.41	0.33	9	0.42	0.34
2	10	0.28	0.05	9	0.26	0.17	10	0.28	0.05	13	0.44	0.34	15	0.47	0.33
3	261	0.29	0.07	287	0.36	0.18	243	0.39	0.09	265	0.35	0.31	248	0.43	0.35
4	1613	0.25	0.08	1622	0.33	0.23	1613	0.25	0.08	1698	0.38	0.47	1719	0.41	0.47
5	26598	0.14	0.1	26602	0.35	0.18	26598	0.14	0.1	26681	0.37	0.42	26706	0.41	0.45
6	4439	0.19	0.11	4419	0.31	0.21	4439	0.19	0.11	4446	0.35	0.43	4489	0.39	0.44

6. Case study

Given that the main approach of the paper is to solve the TSP problem and extract practical insights, in the transportation system and in a situation where there is only information about social interactions between drivers, an attempt was made to use a study related to this area and also complies with the features of the paper. To solve the information diffusion among community leaders-based profit maximization model in the road freight drivers' collaboration network, a case study as reported in Table 6 (Badiee et al., 2020) is applied. In this study, it is possible to model the communication of drivers in the form of drivers' collaboration network and also to extract the relation intensity among them within the network. This table shows the information of issued bills of lading from Iran freight road transportation system, in one month. Based on the details of Table 6, we design 2-layers weighted drivers' collaboration network as a graph $G = \langle V, E, L\rangle$. Edges set in layer one shows that two nodes *u* and *v* work on a shared vehicle at least once a month. Also, edges set in the second layer indicate that two nodes share at least 80 percentage of their paths during a month. The information of the weighted drivers' collaboration network is shown in Table 7.

Table 6 Information of issued bills of lading

Topic	Number
Total companies	4,109
Total number of bills of lading	1,202,756
freights with two drivers	314,638
Total number of drivers	209,315
Unique vehicles	103,357
Origins	1,390
Destinations	3,176

Table 7 Weighted drivers' collaboration network

Items	Layer #1	Layer #2
Nodes	1862	1416
Edges	4696	4061

Given that, the network developed in this paper is designed in a weighted form, by calculating the RI_{uv}, *RI* value between two nodes (by means of Eq. 1) we obtain the edge weights and make the weighted network. Due to the fact that, the *RI* value between two nodes v and u can be different in two layers, the *RI* is defined as the fuzzy number $\widetilde{RI}_{ij} = (a, b, b)$, where the value a is the minimum *RI* value between two nodes v and u and the value b is the biggest one, in the two layers. Also, the *RI* value has a direct effect on *DP* value $\widetilde{N}_{ij}$. Consequently, by extracting the *RI* value between two nodes u and v, the *DP* value $\widetilde{N}_{uv}$ will be extracted.

6.1. Result analysis

First, the problem of detecting communities needs to be solved and the results must be used to address the profit maximization model, through the developed model, as follows:

6.1.1. Solving Extended-OCDEM algorithm

- **Steps 1 to 5 results: final communities**

After creating the multiplex weighted drivers' collaboration network, it is necessary to solve the Extended-OCDEM algorithm. The results of solving steps 1 to 5 of the algorithm are as summarized in Table 8.

Table 8 Results of the Extended-OCDEM algorithm for the weighted drivers' collaboration network

Community statistics	% Overlapping nodes in communities	Redundancy measure	Number of Communities	Number of Overlapping Nodes	Non-overlapping Nodes	Number of Outliers
C_{Max_Size}	83.2	0.281	**49**	**366**	**704**	**792**
C_{mean_Size}	69.0	0.191				
C_{min_Size}	75.0	0.397				
Avg.	**47.7**	**0.283**				

According to Table 8, the total number of final communities in the weighted drivers' collaboration network is 49. The average percentage of overlapping nodes in the communities is 47.7, i.e. only 47.7 percentage, of nodes belong to more than one community. Also, there are 704 nodes in the final communities that belong to only one community. Correspondingly, 792 nodes, called outliers, do not belong to any community. According to the redundancy measure, for those communities with the largest size, the redundancy value is equal to 0.281, for communities with an average size the redundancy value is equal to 0.191 and for the communities with the minimum size this value is equal to 0.397. On average, the redundancy value of the entire network is equal to 0.283.

The distribution of active nodes in the final communities is shown in Fig 5. According to the figure, about 50 percentage of the overlapping nodes in the weighted drivers' collaboration network belong to only two communities, 23 percentage of these nodes to three communities, 10 percentage of the nodes to 4 communities and 14 percentage of the nodes belong to more than 4 communities. Also, the maximum number of communities in which overlapping nodes belong to is 10.

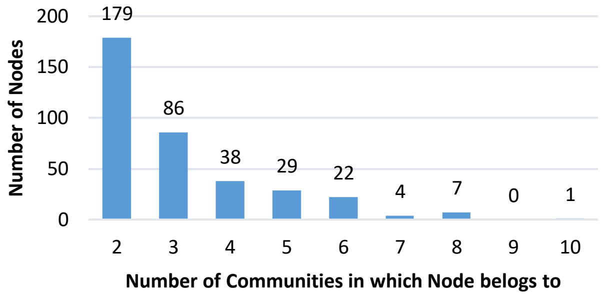

Fig. 5. Dispersion of active nodes in final communities

Table 9 shows the EQ values obtained from steps 2 to 5. The EQ value of step 2 has a very small value. This is because creating each local community is only based on a volunteer edge and expanding the local community through merging only two common nodes of this volunteer edge, without considering any other common node.

Table 9 Comparison of EQ values

Layers	EQ			
	Step 2	Step 3	Step 4	Step 5
L_1	0.077	0.159	0.269	**0.350**
L_2	0.064	0.237	0.299	

By combining the intra-layer communities, the EQ value in step 3 has increased significantly, in two layers, compared to step 2. Finally, using the inter-layer aggregation in step 5, the EQ value becomes equal to 0.35. Based on the results, it can be concluded that the Extended-OCDEM algorithm has a very satisfactory performance in achieving the final communities.

One of the most important applications of community detection algorithms in transportation systems is diffusing shipping orders using drivers' communities in order to find an appropriate driver, in a shortest *possible* time. To do this, we first need to extract community leaders, calculate the distances between each follower node and its leaders and finally evaluate the followers' diffusion probability values by distributing the shipping order news among the leaders.

- **Step 6: Extracting communities' leaders**

Since steps 1 to 5 of the Extended-OCDEM algorithm divided the network into dense meaningful groups, the purpose of step 6 is to identify the community leaders to diffuse shipping order news into the communities. We consider the coverage value λ=0.6. Fig. 6 displays the leader size of each community. The number of leaders in each community is identified as a binary value. For example, the binary value (7,14) indicates that in the community 7, there exist 14 drivers that are identified as the leaders.

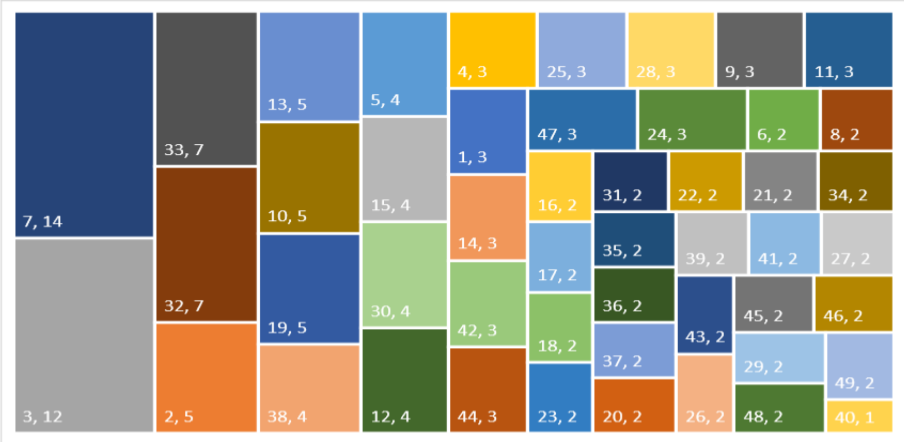

Fig. 6. Size of the community leaders in weighted drivers' collaboration network.

By identifying the community leaders, one can easily investigate the impressibility of each community follower by its community leaders. Measuring the impressibility can be done in two different ways. One of these methods is to calculate the distance between each community follower and its leaders. The second one is to analyze the possibility value of receiving shipping order news by each follower from its community leaders. The results of these two issues are explained as follows:

- **Distance from community leaders**

We calculate the time distance needed by each follower to receive news from its community leaders. It is obvious that, if the time distance between followers and their community leaders was shorter, then the overall time of allocating the shipping orders to the appropriate nodes (followers or leaders) would be even more reduced. We use the "shortest path" method to extract the distance between followers and their leaders.

The dispersion of followers based on their shortest path values from the leaders is shown in Fig. 7. Averagely, for 72 percent of community followers, the shortest path value is one, which means that when the community leaders diffuse the shipping orders news into the community, 72 percent of followers, will be first informed before others, in the community. Also, for 27 percent of community followers the shortest path value is two, and only one percent of the followers have the shortest path value three. For communities 2, 6, 25, 26, 43, and 48 the shortest path value is one, which means that all the community followers will be informed, at the first time.

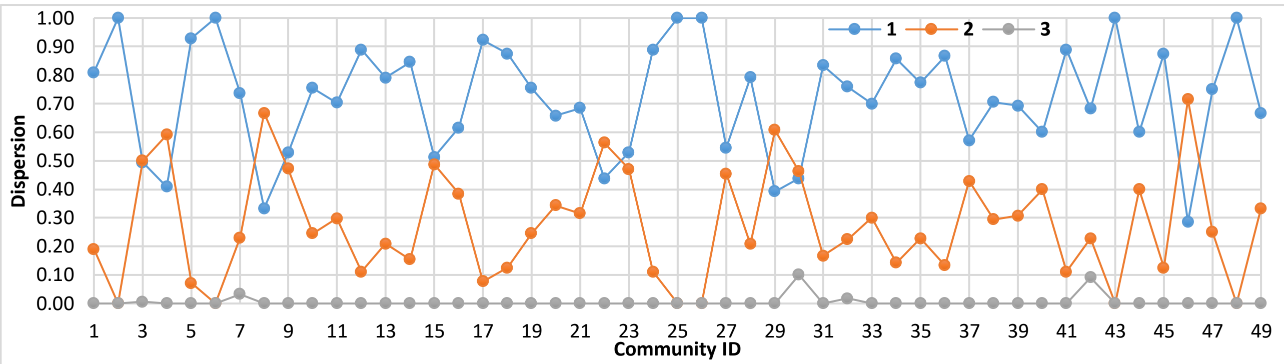

Fig. 7. Dispersion of followers based on their shortest paths (1, 2 and 3) from their leaders in each community

- **Followers' information diffusion possibility**

Another way to analyze the impressibility of community followers is to calculate their DP value. The DP value is calculated using the RI (as per Eq. 1) of the nodes with the "product" operator, as a T-norm function. Fig. 8 displays the distribution of community followers based on their DP values. The DP values are explained in the form of z-number. For example, for community #6, the z-number value is as bellow:

$$(6, [0.6,0.7], 0.58) = Probability\big(Possibility\big(Community(\#6)\big) \in [0.6,0.7]\big)\ is\ 0.58$$

This means that, 58 percent of followers have a DP value between 0.6 and 0.7.
Also, for community #26 we have the z-number value as follows:

$$(26, [0.9,1], 0.63) = Probability\big(Possibility\big(Community(\#26)\big) \in [0.9,1]\big)\ is\ 0.63$$

Which means that 63% of followers in community #26 have a DP value between 0.9 and 1.

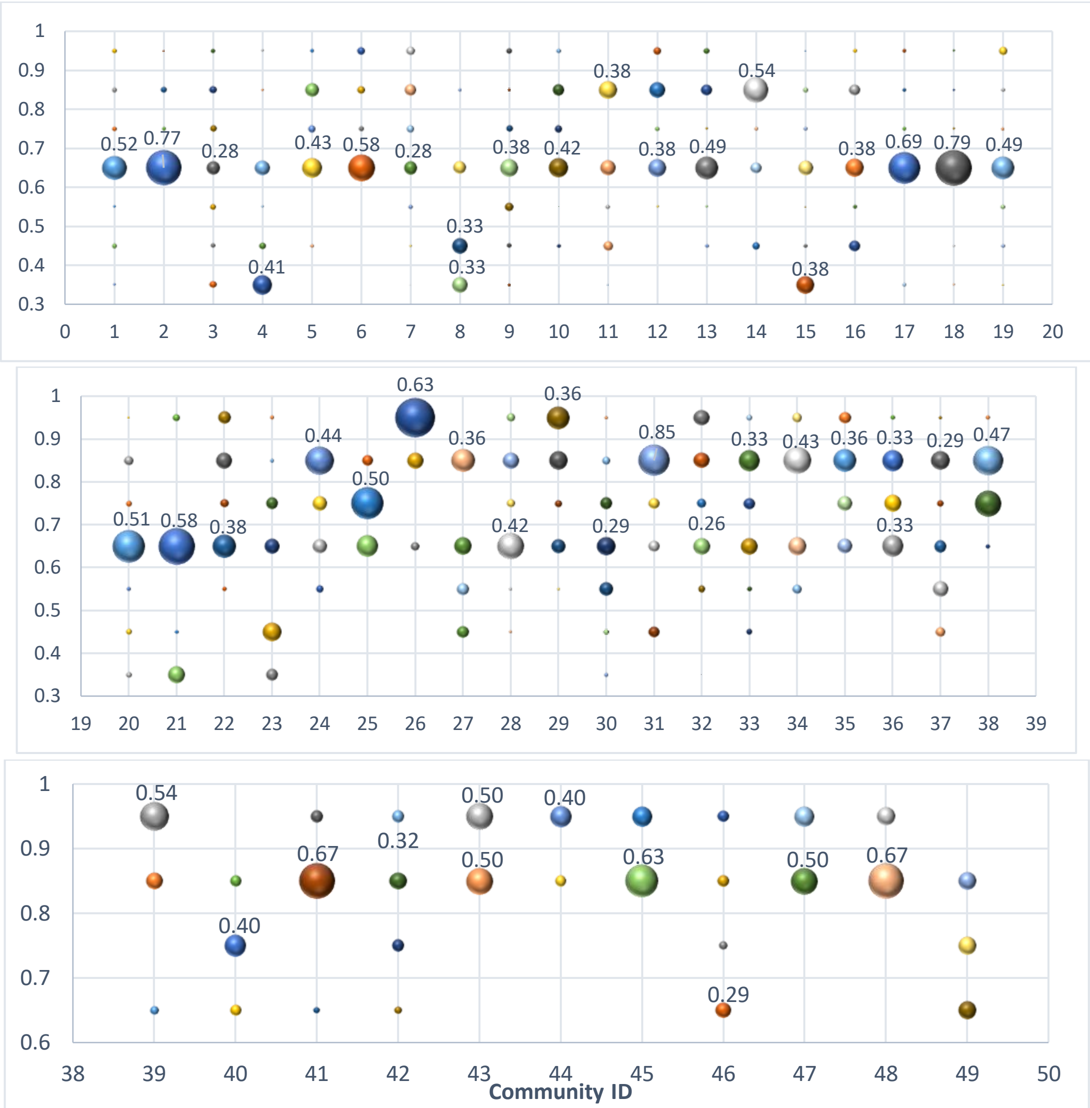

Fig. 8. Dispersion of followers based on their DP values

The average DP value has a normal distribution function with $\mu = 0.73$ and $\sigma = 0.09$, for all 49 communities, that means that with a DP value of 0.73, all followers will diffuse the received shipping order news between each other.

Consider a situation that a company wants to diffuse the shipping order news to only communities whose followers have great dispersions ($\geq 50\%$) in large DP values (≥ 0.8). Therefore, the desired z-number value for a community is as bellow:

$$(Community, possibility \geq 0.8, probability > \%50)$$

The dispersion of followers in each community when the DP value is smaller or larger than 0.8 is shown in Fig. 9.

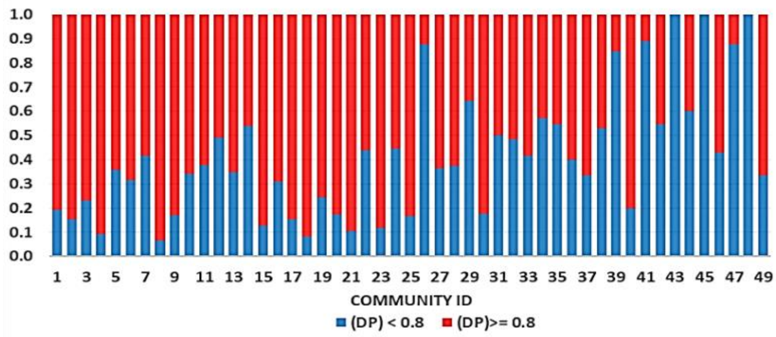

Fig. 9. Community Followers' Dispersion

The results are extracted as follows:

- Communities that have their z-number value $(Com.id, \alpha \geq 0.8, \beta \geq \%50)$ are as below:

$$\{14,26,29,34,35,38,39,41,42,43,44,45,47,48\}$$

i.e. this set consists of communities that over 50 percentages of their followers have $DP \geq 0.8$. So, for these communities, the company can easily distribute the shipping order news to their leaders, in order to find the best drivers.

- Communities with the z-number value $(Community\ id, \alpha < 0.8, \beta > \%50)$:

$$\left\{\begin{matrix} 1,2,3,4,5,6,7,8,9,10,11,12,13,15,16,17,18,19,20 \\ 21,22,23,24,25,27,28,30,31,32,33,36,37,40,46,49 \end{matrix}\right\}$$

Which consists of communities that over 50 percentages of their followers have $DP < 0.8$. In this case, using community leaders to share shipping order news is not advisable. So, it is just proposed to allocate shipping orders to these communities' leaders.

6.1.2. Solving profit maximization model

After identifying the community leaders and calculating the DP value of each community node, in this section we use the diffusion capacity of community leaders to select the most appropriate driver for allocating the shipping order. We solve the profit maximization model using information diffusion among community leaders. The results, summarized in Table 10 for three different values of objective function weights, report the outcomes of solving the maximization model in which the information of allocated shipping orders to each community is explained. Selected drivers can be community leaders or followers. If the selected driver was a follower, then its community leader, who diffuses shipping order to it, is specified. Fig. 10 shows that for the optimum freight shipping wage P_k, the total utility decreases when the revenue part weight decreases. This is due to the increase of the cost part weight of the objective function. For a particular weight of the revenue part, the utility function value varies in different communities, due to the existence of follower nodes at different distances from their leaders, in that community. Accordingly, there are three types of results in the utility function value, as follows:

- **Peak:** Communities in which shipping orders are allocated to the followers that have the shortest path "3" from their own leaders. Given such an orders-followers allocation, followers having lower wages are selected (with respect to those having the shortest path "1" and "2") and the company's revenue increases. So, the objective function is in its maximum value. In this case, the results include communities 3, 7, 30, 32, and 42.
- **Valley:** Communities in which the shipping orders are allocated to the follower nodes that have the shortest path "1" from their leaders. Due to the fact that the acceptance threshold of these followers is greater than the followers with the shortest path "2" and "3", the shipping orders are accepted with a higher wage and, thus, the objective function is in its minimum value. This case involves communities 2, 6, ..., 25, 26, 43, and 48.
- Other utility values that belong to the communities which have the followers with the shortest path "1" and "2" and the shipping orders are allocated to the followers with the shortest path "2".

However, if the weight 0.5 is chosen for both revenue and cost parts in the objective function, the optimum follower node will not be obtained at farther level, i.e. with the shortest path "2" or "3". As a result, the presence of followers with the shortest path "2" or "3" will not lead, in this case, to higher revenue.

Table 10 The shipping orders' information assigned to each community

Community number	W=0.5				W=0.7				W=0.9			
	optimum wage	leader	assigned driver		optimum wage	leader	assigned driver		optimum wage	leader	assigned driver	
			distance	driver id			Distance	driver id			distance	driver id
1	0.93	-	0	15	0.78	15	2	66	0.78	15	2	66
2	0.93	-	0	22	0.86	22	1	108	0.86	22	1	108
3	0.93	-	0	12	0.71	12	3	563	0.71	12	3	563
4	0.93	-	0	12	0.78	12	2	41	0.78	12	2	41
5	0.93	-	0	228	0.78	228	2	452	0.78	228	2	452
6	0.93	-	0	281	0.86	281	1	221	0.86	281	1	221
7	0.93	-	0	79	0.71	79	3	17	0.71	79	3	17
8	0.93	-	0	381	0.78	381	2	46	0.78	381	2	46
9	0.93	-	0	130	0.78	130	2	727	0.78	130	2	727
10	0.93	-	0	309	0.78	309	2	232	0.78	309	2	232
11	0.93	-	0	472	0.78	472	2	1016	0.78	472	2	1016
12	0.93	-	0	6	0.78	6	2	312	0.78	6	2	312
13	0.93	-	0	29	0.78	29	2	71	0.78	29	2	71
14	0.93	-	0	169	0.78	169	2	40	0.78	169	2	40
15	0.93	-	0	226	0.78	226	2	191	0.78	226	2	191
16	0.93	-	0	170	0.78	170	2	59	0.78	170	2	59
17	0.93	-	0	210	0.78	210	2	209	0.78	210	2	209
18	0.93	-	0	103	0.78	103	2	186	0.78	103	2	186
19	0.93	-	0	347	0.78	347	2	453	0.78	347	2	453
20	0.93	-	0	428	0.78	428	2	175	0.78	428	2	175
21	0.93	-	0	21	0.78	21	2	177	0.78	21	2	177
22	0.93	-	0	2	0.78	2	2	17	0.78	2	2	17
23	0.93	-	0	325	0.78	325	2	75	0.78	325	2	75
24	0.93	-	0	244	0.78	244	2	357	0.78	244	2	357
25	0.93	-	0	244	0.86	244	1	1031	0.86	244	1	1031
26	0.93	-	0	182	0.86	182	1	800	0.86	182	1	800
27	0.93	-	0	222	0.78	222	2	401	0.78	222	2	401
28	0.93	-	0	5	0.78	5	2	378	0.78	5	2	378
29	0.93	-	0	126	0.78	126	2	307	0.78	126	2	307
30	0.93	-	0	371	0.71	371	3	129	0.71	371	3	129
31	0.93	-	0	150	0.78	150	2	192	0.78	150	2	192
32	0.93	-	0	102	0.71	102	3	1503	0.71	102	3	1503
33	0.93	-	0	421	0.78	421	2	16	0.78	421	2	16
34	0.93	-	0	144	0.78	144	2	572	0.78	144	2	572
35	0.93	-	0	600	0.78	600	2	86	0.78	600	2	86
36	0.93	-	0	385	0.78	385	2	1198	0.78	385	2	1198
37	0.93	-	0	1074	0.78	1074	2	379	0.78	1074	2	379
38	0.93	-	0	195	0.78	195	2	164	0.78	195	2	164
39	0.93	-	0	207	0.78	207	2	539	0.78	207	2	539
40	0.93	-	0	33	0.78	33	2	443	0.78	33	2	443
41	0.93	-	0	110	0.78	110	2	373	0.78	110	2	373
42	0.93	-	0	155	0.71	155	3	798	0.71	155	3	798
43	0.93	-	0	185	0.86	185	1	595	0.86	185	1	595
44	0.93	-	0	47	0.78	47	2	1213	0.78	47	2	1213
45	0.93	-	0	480	0.78	480	2	66	0.78	480	2	66
46	0.93	-	0	343	0.78	343	2	388	0.78	343	2	388
47	0.93	-	0	35	0.78	35	2	72	0.78	35	2	72
48	0.93	-	0	14	0.86	14	1	500	0.86	14	1	500
49	0.93	-	0	935	0.78	935	2	936	0.78	935	2	936

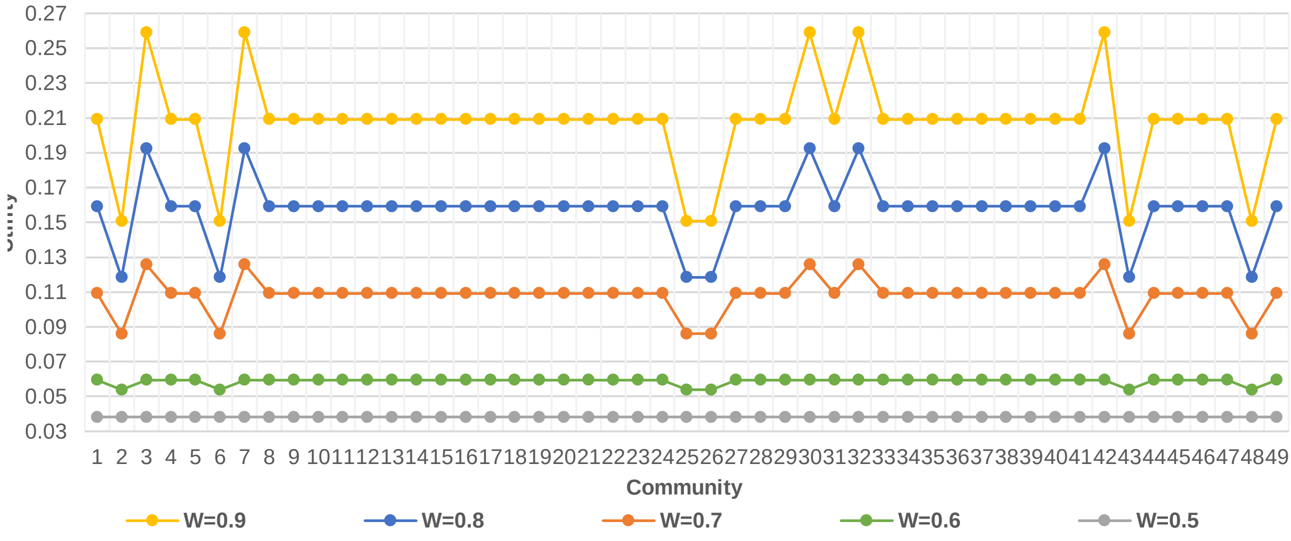

Fig. 10. Utility values base on Revenue weights W

Fig. 11 shows the optimum wage values for each community base on different revenue weights. Whenever the company is well-known and has its own customers, it will be too risky to spend a relatively long time for finding the most suitable driver. So, the company tends to investigate deeper into the community to find drivers who accept the shipping orders with the lowest wages. As a result, the company will consider higher value for the revenue weight, with respect to the cost part. This is clearly reflected in Fig. 11 in which the optimum wage values are lower than the other weights when considering higher values of revenue weight.

However, if the company is a novice and does not have consolidated customers, it prefers to allocate the received shipping orders to the driver, as soon as possible. So, it will avoid any risk to spend a long time to find the most suitable driver and will tend to look for drivers at the community surface, which means community leaders, in the shortest possible time. Thus, as shown in Fig. 11 (red line) the optimum wage P_k is identified with higher values, which leads to minimum revenue for the company.

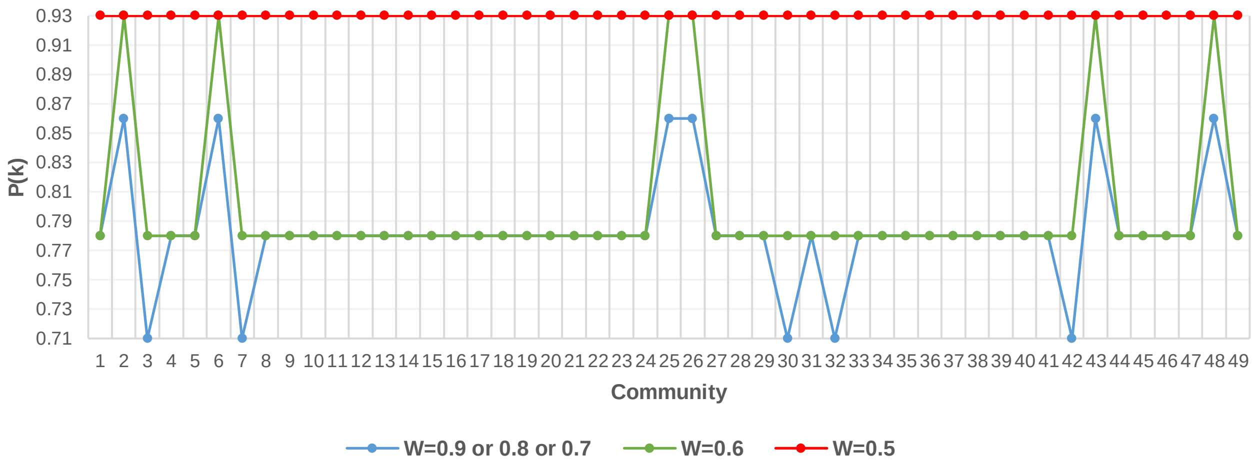

Fig. 11. Community's optimum wages base on different revenue weights W

Fig. 12 shows that, in a risk-aversion scenario, i.e. when the company is a novice and does not have many customers, by choosing the revenue weight in a minimum value, the company tends to allocate shipping orders to well-known drivers, who are at the community surface and not in a community's depth. For example, for the revenue weight value of 0.6, the company allocates 88 percent of its orders to followers with the shortest path "2" and 12 percent of its orders to communities' leaders. Also, for a revenue weight of 0.5 the company allocates 100 percent of its orders to the communities' leaders.

On the other hand, if the company is well established and considers a large value for the revenue weight it tends to accept the risk of prolonging the shipping order allocating time, to look for drivers in the depths of communities who are willing to accept shipping orders with the lowest wage. In this situation, none of the community leaders will be involved in accepting shipping orders. For example, for revenue weight values of 0.7, 0.8, and 0.9, the company allocates 12 percent of its orders to followers with the shortest path “1” and 78 percent of the orders to followers with the shortest path “2” and 10 percent to followers with the shortest path “3”.

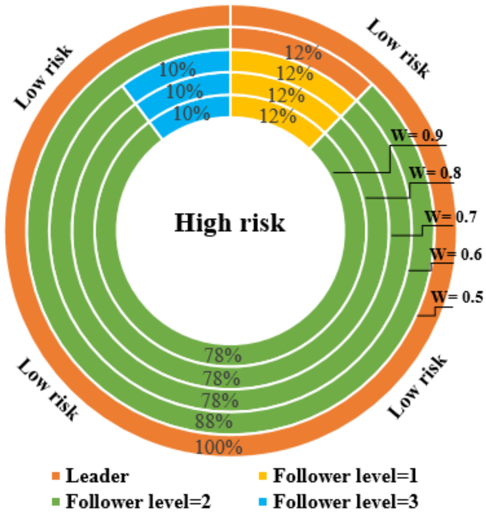

Fig. 12. Combination of the selected drivers base on different revenue weights (W)

6.2. Managerial insights

Based on the results obtained from solving the community detection algorithm and the profit maximization model, the managerial insights are as follows:

- **Driver appropriateness**
  - **In communities' level, existing experienced drivers**: Drivers with a high level of communication and also community leaders (with the shortest path value of 0 or 1) are at the community level. Due to their communications and experiences, these drivers accept shipping orders at a higher fare than other drivers in the community. Also, the possibility value of diffusing shipping order news for these drivers is very high.
  - **In depth of the communities, novice drivers exist:** Drivers who are at the depth of the communities (with shortest value of 3) have a low level of communication and experience compared to others. Therefore, they are willing to accept shipping orders at a lower fare than others. On the other hand, due to the small communication of these drivers, they publish the shipping order news with a less possibility when they receive it.
- **Time Management**
  - **Assign shipping orders to drivers at the level of communities, with the shortest time**: If the company wants to deliver shipping orders to drivers in the shortest possible time, it must look for drivers whose exist at the level of the community (with the shortest path value of 0 or 1).
  - **Assign shipping orders to drivers in the depth of communities, with the most time:** If the time factor is not important for company in allocating shipping orders, it must allocate the orders to drivers in the depth of communities (with the shortest path value of 3).
- **Cost management**
  - **Allocating shipping orders in depths of communities, at the lowest cost:** If the company wants to allocate shipping orders to drivers with the lowest cost, it should look for drivers in depths of communities (with the shortest path value of 3).
  - **Allocating shipping orders at the level of communities, at the highest cost:** If the cost factor is not important in the orders allocation, for the company, it tends to deliver the orders to drivers at the community level (with shortest path value of 0 or 1).

  By combining the above insights, the following two scenarios are derived for companies in allocating their shipping orders based on their brand position and risk level:

- **High-risk scenario for well-known companies**

  Those companies that have a good brand position and have a larger market share, have a higher level of risk for losing customers due to their large number of customers. Therefore, they tend to allocate shipping orders to drivers whose are at the depth of the community. These drivers accept the orders at lower fares due to their being novice. As a result, the profits of these companies are at their highest value.

- **Low-risk scenario for novice companies**

  Novice companies that have few customers and are not willing to accept the risk of prolonging the time of allocating the orders, tend to deliver the orders to drivers at the level of communities in order to maximize customers satisfaction in their shipping experience. These drivers, who are generally community leaders, will accept the orders at a higher fare than others. Therefore, the profit of companies from allocating the orders will be at its lowest value.

Finally, all companies, depending on their brand position and their risk level, can choose a combination of the above two scenarios in order to optimize their profit from allocating shipping orders to drivers.

7. Conclusion

In this paper, first, by developing a two-layer weighted driver collaboration network, the communication between drivers was modeled. The network considered different types of communication between drivers as well as the different values of relation intensity between drivers. To answer the problem of managing the "time", "cost" and "driver appropriateness" in allocating the shipping orders, for the first time we developed a profit maximization model based on the information diffusion ability of leaders and also followers of communities. Also, in order to take advantage of the information diffusion ability among drivers in the optimization model, we developed an algorithm for predicting the behavior of drivers (Step 6) and based on it, we extend the OCDEM algorithm. By proposing the extended OCDEM algorithm, alongside identifying the community leaders, the behavior of each community driver at the time of receiving order news was identified and evaluated in terms of the possibility of accepting the orders and also the power and the speed of diffusing order news.

Given that the above model was an optimization model and the results of solving this model for the case study were exact, it can be concluded that for all 49 communities obtained from the network, the shipping orders were allocated in the shortest possible time, with the lowest cost and to the most suitable drivers. By solving the optimization model, the values of the objective function can be defined in the following two limits:

- **Maximum value of the objective function:** Includes those communities where the orders allocated to drivers with a minimum distance of 3 from their community leaders. In this case, due to the fact that the order allocation to novice drivers was conducted with the lowest fare, the profit of the company was at its highest value.
- **Minimum value of the objective function:** Includes those communities where the orders assigned to only one community leader. In this case, due to the fact that the order was allocated to the community leader with the highest amount of fare, the profit of the company was at its lowest value.

We also analyzed the performance of the extended OCDEM algorithm with 4 other known algorithms in 6 real networks for 2 evaluation criteria. The results show that the extended OCDEM algorithm, in addition to being unique in identifying community leaders, performed better in detecting communities (*EQ* measure) than the others, due to the equipment of the NCOS score with the *RI* index. The *RI* index empowered the extended OCDEM algorithm to consider the smallest relation between driver and the others in its neighbor communities.

Finally, by applying the diffusion optimization model on a case study of freight road drivers' collaboration networks, these two significant insights were obtained:

- **Well-known companies choose high-risk scenario:** Companies who are well-known and have their wide set of customers, tend to allocate shipping orders to drivers in depths of the communities, who can accept the orders with a minimum wage due to being novice. So, the total profit of these companies will be in a greater value.
- **Novice companies choose low-risk scenario:** Novice companies try to allocate shipping orders to communities' leaders, to avoid the risks of spending a long time to find the most suitable driver. So, the total profit of these companies will be in a lower value, but the shipping orders will be assigned in a shortest possible time.

As can be noticed, the real behaviors of companies in assigning drivers to shipping orders were appropriately shown from the results of the optimization model. Therefore, the model outcomes give valuable visions to the transportation companies on how to allocate the orders to drivers based on their brand position, strategies and risk-level in order to improve their performance.

The following issues are suggested for future research:

Developing an optimization model to manage the risk of losing customers, developing a model with credit rating of drivers to manage the allocation of shipping orders to them, in a competitive environment, developing a multi-objective optimization model by considering environmental, economic and social sustainability factors.

References

Abu Al Hla, Y., Othman, M., Saleh, Y., 2019. Optimising an eco-friendly vehicle routing problem model using regular and occasional drivers integrated with driver behaviour control. J. Clean. Prod. https://doi.org/10.1016/j.jclepro.2019.06.156

Agha Mohammad Ali Kermani, M., Badiee, A., Aliahmadi, A

., Ghazanfari, M., Kalantari, H., 2016. Introducing a procedure for developing a novel centrality measure (Sociability Centrality) for social networks using TOPSIS method and genetic algorithm. Comput. Human Behav. 56, 295–305. https://doi.org/10.1016/j.chb.2015.11.008

Badiee, A., Ghazanfari, M., 2019. Development of a monopoly pricing model for diffusion maximization in fuzzy weighted social networks with negative externalities of heterogeneous nodes using a case study. Neural Comput. Appl. 31, 6287–6301.

Badiee, A., Ghazanfari, M., 2018. A monopoly pricing model for diffusion maximization based on heterogeneous nodes and negative network externalities (Case study: A novel product). Decis. Sci. Lett. https://doi.org/10.5267/j.dsl.2017.8.001

Badiee, A., Kalantari, H., Ghazanfari, M., Fathian, M., Shahanaghi, K., 2020. Introducing drivers' collaboration network: A two-layers social network perspective in road transportation system analysis. Res. Transp. Bus. Manag. https://doi.org/10.1016/j.rtbm.2020.100532

Buer, T., Kopfer, H., 2014. A Pareto-metaheuristic for a bi-objective winner determination problem in a combinatorial reverse auction. Comput. Oper. Res. https://doi.org/10.1016/j.cor.2013.04.004

Buer, T., Pankratz, G., 2010. Solving a bi-objective winner determination problem in a transportation procurement auction. Logist. Res. https://doi.org/10.1007/s12159-010-0031-8

Chen, Y., Frei, A., Mahmassani, H.S., 2014. From Personal Attitudes to Public Opinion. Transp. Res. Rec. J. Transp. Res. Board 2430, 28–37. https://doi.org/10.3141/2430-04

Chi, C., 2015. Mathematical Programming Approaches to Home Healthcare Nurse Routing Problem and Truckload Transportation Procurement via Combinatorial Auctions.

Feki, Y., Hajji, A., Rekik, M., 2016. A hedging policy for carriers' selection under availability and demand uncertainty. Transp. Res. Part E Logist. Transp. Rev. https://doi.org/10.1016/j.tre.2015.11.011

Fern, X.Z., Brodley, C.E., 2004. Solving cluster ensemble problems by bipartite graph partitioning, in: Proceedings, Twenty-First International Conference on Machine Learning, ICML 2004. https://doi.org/10.1145/1015330.1015414

Hammami, F., Rekik, M., Coelho, L.C., 2020. The Combinatorial Bid Construction Problem with Stochastic Prices for Transportation Services Procurement. Faculté des sciences de l'administration.

Hammami, F., Rekik, M., Coelho, L.C., 2019. Exact and heuristic solution approaches for the bid construction problem in transportation procurement auctions with a heterogeneous fleet. Transp. Res. Part E Logist. Transp. Rev. https://doi.org/10.1016/j.tre.2019.05.009

Harris, J.K., Wong, R., Thompson, K., Haire-Joshu, D., Hipp, J.A., 2015. Networks of collaboration among scientists in a center for diabetes translation research. PLoS One 10, e0136457.

Hu, Q., Zhang, Z., Lim, A., 2016. Transportation service procurement problem with transit time. Transp. Res. Part B Methodol. https://doi.org/10.1016/j.trb.2016.01.007

Ignatius, J., Hosseini-Motlagh, S.M., Goh, M., Sepehri, M.M., Mustafa, A., Rahman, A., 2014. Multiobjective combinatorial auctions in transportation procurement. Math. Probl. Eng. https://doi.org/10.1155/2014/951783

Jothi Basu, R., Bai, R., Palaniappan, P.L.K., 2015a. A strategic approach to improve sustainability in transportation service procurement. Transp. Res. Part E Logist. Transp. Rev. https://doi.org/10.1016/j.tre.2014.10.015

Jothi Basu, R., Subramanian, N., Cheikhrouhou, N., 2015b. Review of Full Truckload Transportation Service Procurement. Transp. Rev. 35, 599–621. https://doi.org/10.1080/01441647.2015.1038741

Kalantari, H., Ghazanfari, M., Fathian, M., Shahanaghi, K., 2020. Multi-objective optimization model in a heterogeneous weighted network through key nodes identification in overlapping communities. Comput. Ind. Eng. https://doi.org/10.1016/j.cie.2020.106413

Kivelä, M., Arenas, A., Barthelemy, M., Gleeson, J.P., Moreno, Y., Porter, M.A., 2014. Multilayer networks. J. Complex Networks. https://doi.org/10.1093/comnet/cnu016

Kuyzu, G., Akyol, Ç.G., Ergun, Ö., Savelsbergh, M., 2015. Bid price optimization for truckload carriers in simultaneous transportation procurement auctions. Transp. Res. Part B Methodol. https://doi.org/10.1016/j.trb.2014.11.012

Kwon, R.H., Lee, C.-G., Ma, Z., 2005. An integrated combinatorial auction mechanism for truckload transportation procurement. Univ. Toronto, http//www. mie. utoronto. ca/labs/ilr/MultiRound. pdf, Work. Pap.

Lafkihi, M., Pan, S., Ballot, E., 2017. Mechanisms for freight transportation service procurement: A literature-based analysis, in: The 47th International Conference on Computers & Industrial Engineering (CIE47).

Lee, C.G., Kwon, R.H., Ma, Z., 2007. A carrier's optimal bid generation problem in combinatorial auctions for transportation procurement. Transp. Res. Part E Logist. Transp. Rev. https://doi.org/10.1016/j.tre.2005.01.004

Li, Y., Chen, H., Prins, C., 2016. Adaptive large neighborhood search for the pickup and delivery problem with time windows, profits, and reserved requests. Eur. J. Oper. Res. https://doi.org/10.1016/j.ejor.2015.12.032

Lim, A., Qin, H., Xu, Z., 2012. The freight allocation problem with lane cost balancing constraint. Eur. J. Oper. Res. https://doi.org/10.1016/j.ejor.2011.08.028

Lim, A., Wang, F., Xu, Z., 2006. A transportation problem with minimum quantity commitment. Transp. Sci. https://doi.org/10.1287/trsc.1050.0123

Ma, Z., Kwon, R.H., Lee, C.-G., 2010. A stochastic programming winner determination model for truckload procurement under shipment uncertainty. Transp. Res. Part E Logist. Transp. Rev. 46, 49–60.

Mamaghani, E.J., Chen, H., Prins, C., Demir, E., 2019. An improved tabu search algorithm for a multi-period bid generation problem with the consideration of delivery lead time. IFAC-PapersOnLine. https://doi.org/10.1016/j.ifacol.2019.11.599

Olcaytu, E., Kuyzu, G., 2019. Location-based distribution estimation for stochastic bid price optimization. Transp. Lett. 1–15.

Othmane, I. Ben, Mellouli, S., Rekik, M., 2019. Reputation-Based Winner Determination Problem in Centralized Combinatorial Auctions for the Procurement of Transportation Services. CIRRELT.

Qian, X., Chan, F.T.S., Yin, M., Zhang, Q., Huang, M., Fu, X., 2020. A two-stage stochastic winner determination model integrating a hybrid mitigation strategy for transportation service procurement auctions. Comput. Ind. Eng. https://doi.org/10.1016/j.cie.2020.106703

Rekik, M., Desaulniers, G., Elhallaoui, I., Saddoune, M., 2017. An exact solution approach for bid construction in truckload transportation procurement auctions with side constraints. GERAD, École des hautes études commerciales.

Rekik, M., Mellouli, S., 2012. Reputation-based winner determination problem for combinatorial transportation procurement auctions. J. Oper. Res. Soc. https://doi.org/10.1057/jors.2011.108

Remli, N., Rekik, M., 2013. A robust winner determination problem for combinatorial transportation auctions under uncertain shipment volumes. Transp. Res. Part C Emerg. Technol. https://doi.org/10.1016/j.trc.2013.07.006

Suthers, D., Fusco, J., Schank, P., Chu, K.H., Schlager, M., 2013. Discovery of community structures in a heterogeneous professional online network, in: Proceedings of the Annual Hawaii International Conference on System Sciences. https://doi.org/10.1109/HICSS.2013.179

Tomasini, M., 2015. An introduction to multilayer networks. BioComplex Lab. Florida Inst. Technol. Melbourne, USA 1–14.

Triki, C., 2021. Using combinatorial auctions for the procurement of occasional drivers in the freight transportation: A case-study. J. Clean. Prod. https://doi.org/10.1016/j.jclepro.2021.127057

Triki, C., Oprea, S., Beraldi, P., Crainic, T.G., 2014. The stochastic bid generation problem in combinatorial transportation auctions. Eur. J. Oper. Res. https://doi.org/10.1016/j.ejor.2013.06.013

Wang, X., Xia, M., 2005. Combinatorial Bid Generation Problem for Transportation Service Procurement. Transp. Res. Rec. J. Transp. Res. Board. https://doi.org/10.1177/0361198105192300120

Xu, S.X., Huang, G.Q., 2013. Transportation service procurement in periodic sealed double auctions with stochastic demand and supply. Transp. Res. Part B Methodol. https://doi.org/10.1016/j.trb.2013.07.015

Yakoubi, Z., Kanawati, R., 2014. LICOD: A Leader-driven algorithm for community detection in complex networks. Vietnam J. Comput. Sci. https://doi.org/10.1007/s40595-014-0025-6

Yan, F., Chen, K., Xu, M., 2020. A bid generation problem for combinatorial transportation auctions considering in-vehicle consolidations. Asia Pacific J. Mark. Logist. https://doi.org/10.1108/APJML-09-2019-0556

Yan, F., Ma, Y., Feng, C., 2018a. A bi-level programming for transportation services procurement based on combinatorial auction with fuzzy random parameters. Asia Pacific J. Mark. Logist.

Yan, F., Ma, Y., Xu, M., Ge, X., 2018b. Transportation Service Procurement Bid Construction Problem from Less Than Truckload Perspective. Math. Probl. Eng. 2018, 1728512. https://doi.org/10.1155/2018/1728512

Yan, F., Xu, M., Ma, Y., Yu, H., 2017. Price optimization for transportation service procurement with fuzzy random shipments: from shipper's perspective. Transp. Lett. 9, 258–275.

Yang, F., Huang, Y.H., 2020. Linearization technique with superior expressions for centralized planning problem with discount policy. Phys. A Stat. Mech. its Appl. 542, 122746. https://doi.org/10.1016/j.physa.2019.122746

Yang, F., Huang, Y.H., Li, J., 2019. Alternative solution algorithm for winner determination problem with quantity discount of transportation service procurement. Phys. A Stat. Mech. its Appl. 535, 122286. https://doi.org/10.1016/j.physa.2019.122286

Zhang, B., Ding, H., Li, H., Wang, W., Yao, T., 2014. A Sampling-Based Stochastic Winner Determination Model for Truckload Service Procurement. Networks Spat. Econ. https://doi.org/10.1007/s11067-013-9214-6

Zhang, B., Yao, T., Friesz, T.L., Sun, Y., 2015. A tractable two-stage robust winner determination model for truckload service procurement via combinatorial auctions. Transp. Res. Part B Methodol. https://doi.org/10.1016/j.trb.2015.03.019

Zhang, M., Hu, Q., 2019. An effective evolutionary algorithm for the biobjective full truckload transportation service procurement problem. Comput. Ind. Eng. https://doi.org/10.1016/j.cie.2018.11.036