Efficiency of routing and scheduling system for small and medium size enterprises utilizing vehicle location data

Tomio Miwa*
Dr. Eng., Associate Professor
Institute of Materials and Systems for Sustainability, Nagoya University, Furo-cho, Chikusa-ku, Nagoya 464-8603, Japan
E-mail: miwa@nagoya-u.jp

Michael G.H. Bell
Ph.D., Professor
Institute of Transport and Logistics Studies, Business School, The University of Sydney, NSW 2006, Australia
E-mail: michael.bell@sydney.edu.au

*Corresponding author. . Tel.: +81-52-789-5018; fax: +81-52-789-5728.

Abstract:
The routing and scheduling for trucks and vans in an urban road network depends critically on the state of the road network. Trucks and vans impose significant costs on other road users and the environment, so improved routing and scheduling benefits more than just the logistics industry. However, small and medium size enterprises (SMEs) in the logistics business cannot justify investment in planning systems. In this paper, an autonomous routing and scheduling system which is available to SMEs is proposed and the efficiency of the system is investigated. The proposed system accumulates vehicle location data in a central server and uses it to generate traffic information. Test simulations using a grid network demonstrate the effects of utilizing and sharing vehicle location data on delivery efficiency. The simulation results show that the improvement of delivery efficiency is mainly due to the reduction of penalty cost for early and late arrival at the customer location. It is also shown that the system leads to the buffer effect from variations in traffic conditions on delivery cost and this effect is enhanced by taking travel time uncertainty into consideration. It is further shown that the presence of measurement periods with insufficient data results in unreliable routing and scheduling. For a reliable system, data collection over a wider area is required rather than dense data in a subset of links.

Key words: Routing, Scheduling, Logistics, Probe vehicles

1. Introduction

The routing and scheduling for trucks and vans in an urban road network depends critically on the state of the road network. Trucks and vans impose significant costs on other road users and the environment, so improved routing and scheduling benefits more than just the logistics industry. Larger logistics companies operating big fleets of vehicles deploy sophisticated tracking and tracing systems to support the implementation of routes and schedules, but these are expensive. Small and medium sized enterprises (SMEs) with smaller fleets of vehicles cannot justify investing in such expensive systems, but are nonetheless numerous and have a large impact on the road network and the urban environment. Improved routing and scheduling for SMEs would bring significant benefits to the businesses themselves, other road users and the urban environment.

Many cities provide web-based journey planners and justify public investment in such systems in terms of improved mobility, promotion of public transport and mitigation of road congestion (e.g., Transport for NSW). These services are increasingly available via the mobile internet and take online traffic and travel information into account. Our study is based on the contention that a similar service for SMEs that operate pickup and delivery tours would be of comparable value by improving logistical efficiency, reducing truck and van traffic, and thereby decreasing congestion, noise and emissions.

In contrast to public transport systems, which offer a certain service level based on a regular schedule, the level of service provided by road transportation – as measured using an index such as travel time over the road network – changes greatly from hour to hour and day to day and also varies among vehicles. To estimate travel times and plan vehicle routing, it is necessary to collect road traffic information. Although stationary roadside sensors such as loop detectors have been traditionally used to collect such information, they do not cover the whole road network for reasons of cost. Furthermore, the real-time use of such information is not trivial as the sensors do not measure travel time directly. To realize a routing and scheduling system that can be made freely available to SMEs, it would be preferable to establish an autonomous system that both collects and uses traffic information. This means it would be beneficial to utilize, as probe data, location information from trucks and vans using the system itself.

In this study, the concept of a web-based routing and scheduling system that can be made freely available to SMEs is presented. The efficiency of such an autonomous routing and scheduling system is then analyzed. Looking at the travel time distribution in the road network, a focus is the analysis of the relationship between the quantity of data that has been accumulated in the historical database and system efficiency. In this way, the influence of fluctuations in traffic conditions on routing and scheduling efficiency is determined. Especially, the contribution of this research is to show the relationship between the historical database being updated day by day and the efficiency of routing and scheduling system.

This paper is organized as follows. Section 2 is a literature review of vehicle routing and scheduling systems for SMEs and the use of vehicle location data in the logistics field. In Section 3, the concept of the proposed autonomous routing and scheduling system is explained. In Section 4, data setting for the test simulations and problem formulation are described. In Section 5, the results of test simulations are presented and the efficiency of the system is discussed. Finally, Section 6 concludes the study and discusses future research.

2. Literature review

2.1. Vehicle routing and scheduling systems and logistics SMEs

The vehicle routing and scheduling problem plays a central role in the fields of physical distribution and logistics (Laporte, 1992). Since the problem was first described by Dantzig & Ramser (1959), hundreds of papers have been published on the topic (Baldacci et al., 2008). Various forms of the problem, such as mixed vehicle types and vehicle capacity (Desrochers & Verhoog, 1991; Baldacci et al., 2008), pickup and delivery on the same tour (Min 1989; Bianchessi & Righini, 2007), customer time windows (Solomon, 1987; Bräysy & Gendreau, 2005), the stochastic nature of travel time and/or customer demand (Gendreau et al. 1996; Kenyon & Morton, 2003; Luo et al., 2016), dynamic optimization of the problem (Haghani & Jung, 2005), multi depot (Mancini, 2016) and so on have so far been considered. The reason for the problem continuing to receive great attention is the mathematical difficulty of its solution as well as the importance of the problem in practice. Although a great variety of planning software is available commercially, including ArcGIS for Transportation Analytics (Institute for Operations Research and the Management Sciences, 2015), such applications are not free to use. Further, the fact that routing and scheduling software should be customized according to customer and delivery characteristics makes widespread adoption of such software difficult. However, most problems faced by SMEs when making deliveries are small optimization problems, so it is only necessary to handle the relatively trivial optimization of small distribution problems in many cases.

Past research indicates that many logistics SMEs are in need of good routing and scheduling plans and that a free or inexpensive system to assist in planning would be beneficial. Golob & Regan (2003) show, by carrying out telephone interviews of 700 logistics companies operating in California, that although almost all find it annoying to re-route because of congestion, only 26% of them use routing and scheduling software. Similarly, Zeimpekis &

Giaglis (2006), in a questionnaire survey of 73 logistics SMEs in Greece, show that only 27% use routing and scheduling software and that almost half of the companies give drivers no instructions, leaving them responsible for daily distribution routing. This research also shows that the biggest constraint on investment in routing and scheduling software is cost, while the second biggest constraint is an unclear return on investment. Grakovski et al. (2008) indicate that logistics SMEs whose work is outsourcing deliveries in Riga, which are responsible for 20% of total city logistics, are unable to provide their vehicles with optimum routes and schedules. Cagliano et al. (2014) developed the urban logistics simulation and optimization system including and the routing and scheduling planning as a component of the system. Past research makes clear that an inexpensive routing and scheduling service would be accepted by SMEs and would make their activities more efficient, benefitting not just the SMEs but also the community as a whole.

One effect of such a system other than routing and scheduling is in the area of real-time information availability and information exchanges among the companies involved. Zeimpekis & Giaglis (2006) indicate that many logistics SMEs desire real-time proof of delivery and real-time re-routing using a PDA or tablet as well as the monitoring of vehicle position to identify bottlenecks in delivery. Grakovski et al. (2008) discuss the importance of information sharing and have developed the internet portal “Riga City Logistics” which coordinates demand and supply of delivery services between customers and SMEs while also optimizing the routing and scheduling problem. Kuswantoro & Rosli (2012) show, through regression analysis using questionnaire results from 120 SMEs in Indonesia, that innovation in information sharing has a big impact on the performance of SMEs. Tummel et al. (2012) show, through a simulation study using German domestic logistics data, that shipment reassignments by logistics SMEs based on information sharing can improve transportation efficiency. Derrouiche et al. (2014) evaluate the delivery collaborations among logistics SMEs through simple case study. Additionally, Gunasekaran et al. (2007) indicate, based on interviews with logistics companies in Hong Kong, that real-time information usage improves the performance of logistics SMEs and can be a major source of logistics productivity. This research demonstrates the importance of real-time information and information sharing among companies.

2.2. Use of vehicle location data as probe data in the logistics field

Many trucks involved in delivery services are equipped with a GPS device for operation and management (Wang et al. 2015). Additionally, location data from mobile phones or tablets can be utilized as a probe vehicle data (Herrera et al., 2010). Lin et al. (2008) estimate nationwide link speed distributions from truck location data collected in the United States. Comendador et al. (2012) propose efficiency indicators, based on GPS data from delivery trucks, providing a simple fleet management tool for SMEs. Fazekas et al. (2014) analyze emergency events involving trucks, such as abrupt braking, using data from on-board vehicular safety systems with location information. Simroth & Zähle (2011) develop an algorithm for predicting long-distance travel times, which are relevant for logistics fleets, utilizing truck location data. Their algorithm predicts and updates travel time en route to a destination using histrical data and travel time experienced from the origin to the truck's location. Kim et al. (2011) develop a GIS-based system for generating updated routing instructions for vehicles carrying hazardous materials and show the efficiency of real-time information within the vicinity of the vehicle's current location. It is clear from this literature that many attempts have been made to use location data from trucks and that location data can be utilized as probe data for estimating travel times for vehicle routing and scheduling.

3. Concept of autonomous routing and scheduling system

As mentioned above, this research focuses on the concept of a service for SMEs that operate pickup and delivery tours. The vision is a routing and scheduling service delivered to truck and van drivers via the mobile internet. Drivers would access the service via a standard smartphone or tablet computer. Therefore, it is important to design user friendly interface to emphasize the system effect. Key to the vision is a web-based diary kept by the SME indicating the time windows within which pickups and deliveries need to be made, the venues and the type of loads. The drivers would be fed guidance via a smartphone or tablet computer in the vehicle, which in turn would feed the system with the current location of each vehicle and proof of delivery. Here, the SMEs participating in the system are assumed to agree to share the location information of their trucks.

Figure 1 shows the components of the conceptual web-based dynamic routing and scheduling system for SMEs. As already noted, one of the key components of the system is the utilization of location data from trucks as probe vehicle data. Even though each SME operates a small fleet, the data gathered from numerous SMEs will become a huge data source. Sharing the information among SMEs will have a significant impact on the accuracy of routing and scheduling plan. Location data is transmitted from each vehicle through the mobile internet to a system server, where it is stored in a historical database and also used to confirm delivery progress. Although more concrete system architectures which consider various data format, mobile cellular networks, personal security, and so on, are proposed in literature (Wang, F.Y., 2010; Zhang et al., 2011; Bin et al., 2013; Pan et al., 2014), it is out of focus of this research.

The historical database is conceptually the same as those described in the literature (Ehmke et al., 2012; Miwa et al., 2015). Efficient routing and scheduling depends on the availability of historical travel time data; a single and static travel time value per link only poorly represents the traffic situation. A city logistics routing system requires time-dependent travel times that capture changes in traffic conditions on each network link (Ehmke et al., 2012). The importance of time-dependent travel times for improving routing performance is widely recognized in other literature (Ichoua et al., 2003; Potvin et al., 2006; Figliozzi, 2009; Maden et al., 2010).

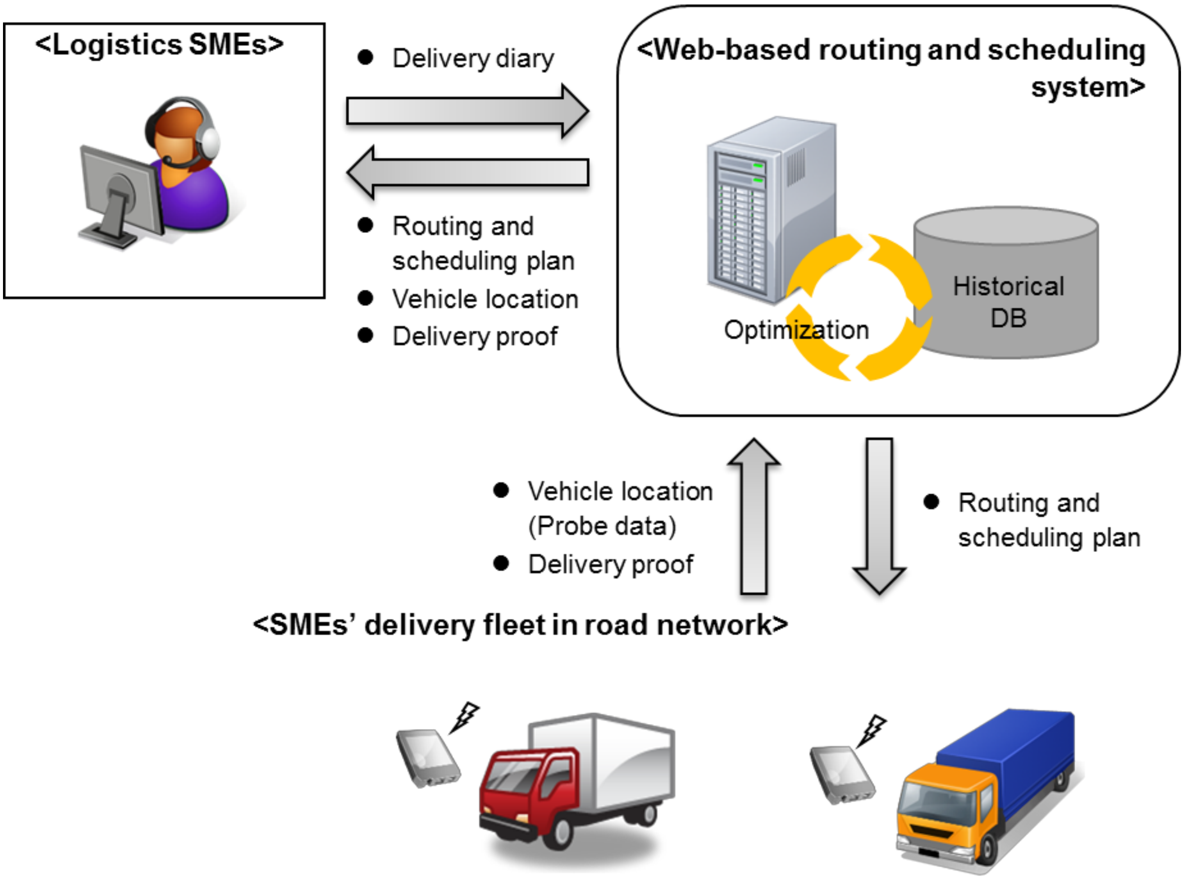


Fig. 1. Components of web-based dynamic routing and scheduling system for SMEs

4. Simulation study on routing and scheduling system utilizing vehicle location data

4.1. Simulation framework

A test simulation is implemented so as to obtain a perspective of the proposed system. The focus is the evaluation of the efficiency of utilization of vehicle location data. Figure 2 shows the simulation framework. In the test simulation, the parameters of link travel time distributions are set as input information in regard to traffic conditions and path travel time distributions are derived from them. Trucks depart from their depots following a planned schedule and take the planned path. It is assumed that truck location data is collected and processed as probe vehicle data. As a result, link travel times are obtained and accumulated in the historical database. Utilizing this database, the routing and scheduling system plans the order in which customers are visited, the paths between customers and the departure time from the depot so as to minimize total cost. The historical database is updated every day, so the accumulation of newly collected data will affect planning on and after the following day. The accuracy of historical database becomes higher as the operating day goes on.

In the historical database, link travel times are accumulated and averaged over each 5-minute period (the measurement period). The information saved in the database is the mean link travel time, the variance of link travel time and the number of passes for each measurement period. The database has an initial value of link travel time that is used if there is no data point for that measurement period along a link. On the first day's operation, for example, all measurement periods on all links are occupied by the initial value. The initial value of link travel time can be set on an empirical basis, such as by calculating link travel time from the speed limit. The variance is set to zero when there are no data points.

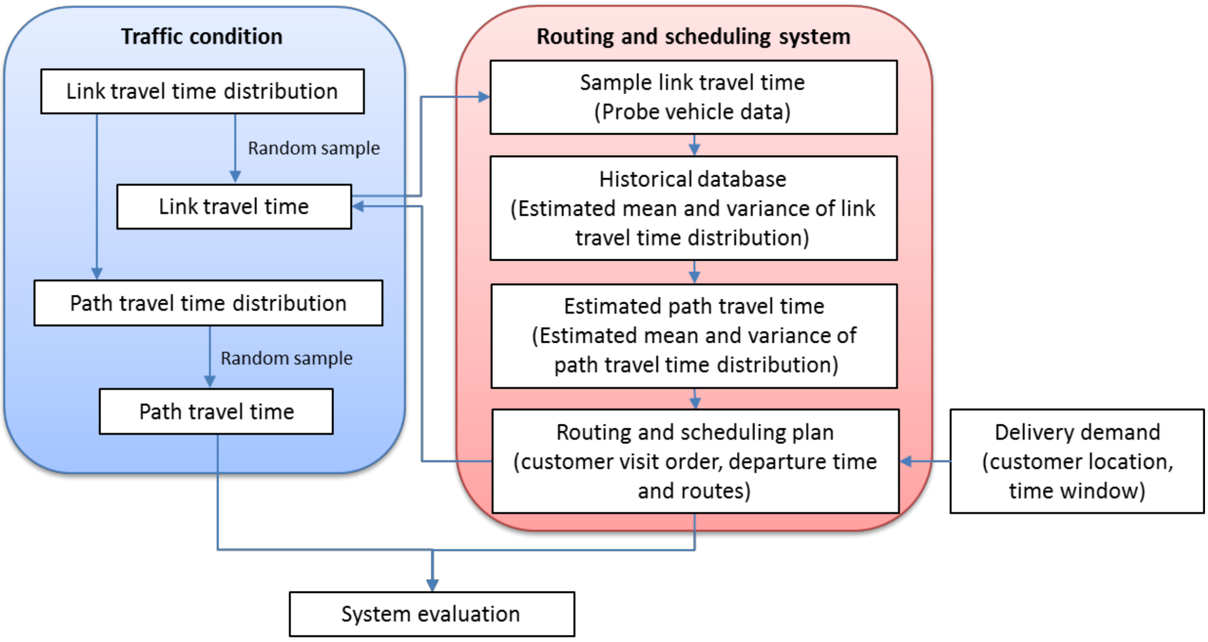


Fig. 2. Data flow in the test simulation model

It should be noted that two aspects of the system are dynamic. One is the routing and scheduling plan, which is continuously updated as new orders are entered during the delivery or pickup tour and as traffic conditions change (as in Haghani & Jung, 2005). The other is the day-to-day dynamics relating to daily changes in traffic conditions. In the case of the routing and scheduling method proposed in this paper, the day-to-day dynamics also relate to the evolution of the historical database and the convergence of system efficiency. The day-to-day dynamics in autonomous routing and scheduling systems is not investigated up to now as far as authors know.

In this study, following past research (see, for example, Chen et al., 2007), travel time is assumed to follow a log-normal distribution. Rakha et al. (2006) pointed out the existence of strong correlations among link travel times and empirically demonstrated the most accurate method for calculating the variance of path travel time. In their method, it is assumed that the path's coefficient of variance is set equal to the mean coefficient of variation over all links along the path. Using this method, the mean and variance of path travel time are expressed by the following equations.

$$\bar{T}(d,k) = \sum_{l \in L_k} \bar{t}_l\big(d + \bar{T}(d,k_{-l})\big) \tag{1.a}$$

$$var\{T(d,k)\} = \frac{\{\bar{T}(d,k)\}^2}{|L_k|^2} \cdot \left[\sum_{l \in L_k} \frac{\sqrt{var\{t_l\big(d + \bar{T}(d,k_{-l})\big)\}}}{\bar{t}_l\big(d + \bar{T}(d,k_{-l})\big)}\right]^2 \tag{1.b}$$

where $T(d,k)$ is the travel time of path k at departure time d, $\bar{T}(d,k)$ is its mean value, $\bar{T}(d,k_{-l})$ is the mean travel time of path k from the origin node to the inflow node of link l, $t_l\big(d + \bar{T}(d,k_{-l})\big)$ is the travel time of link l on path k at departure time d, $\bar{t}_l\big(d + \bar{T}(d,k_{-l})\big)$ is its mean value, L_k is the set of links which compose path k, and $|L_k|$ is the number of links in L_k. Random values of link and path travel time are drawn from the log-normal distribution. In the case of path travel time, the probability density function is expressed as follows,

$$f\big(T(d,k)\big) = \frac{1}{\sqrt{2\pi}\sigma_{k,d} T(d,k)} exp\left[-\frac{\big\{ln\big(T(d,k)\big) - \mu_{k,d}\big\}^2}{2{\sigma_{k,d}}^2}\right] \tag{2.a}$$

$$\mu_{k,d} = ln\left[\frac{\bar{T}(d,k)^2}{\sqrt{var\big(T(d,k)\big) + \bar{T}(d,k)^2}}\right] \tag{2.b}$$

$$\sigma_{k,d} = \sqrt{ln\left\{1 + \frac{var\big(T(d,k)\big)}{\bar{T}(d,k)^2}\right\}} \tag{2.c}$$

where $\mu_{k,d}$ and $\sigma_{k,d}$ are the mean and standard deviation of the natural logarithm of path travel time $T(d,k)$, respectively.

It should be noted that the path travel times are drawn from above probability density function rather than from samples of link travel times, since the latter are implicitly correlated with each other. Therefore, in this study, random samples of link travel times, which will be accumulated in the historical database, and random samples of path travel times are treated independently. This means that the test simulation conducted here evaluates not information accuracy but efficiency of utilization of location data and the day-to-day evolution of the routing and scheduling system.

4.2. Data setting and problem formulation

As this study addresses SMEs, the system will handle a large number of small optimization problems, so the challenge lies not in solving large and complex optimization problems, but rather in efficiently managing data feeds and the relatively trivial optimization (and reoptimization) of small distribution problems. For very small problems, an exhaustive search may be feasible; in the case of larger problems, recourse to the saving method (Clarke & Wright, 1964) or some other suitable heuristics might be required to achieve the required speed (Gayialis & Tatsiopoulos, 2004).

The test simulation model uses a grid network which is useful for getting a fundamental knowledge free from particular city conditions. It is shown by Figure 3, which consists of 80 links and 25 nodes. Two road types, major and minor, are assumed, with major roads consisting of a ring road and other major roads. All links have the same length of 5 km and link travel speeds change with time considering traffic congestion. Figure 4 shows the speed profiles of each type of road, which are set by reference to Kaparias et al. (2007). The speed profiles show the mean value of link travel speed in each measurement period. The average values of link travel speed for whole business period, which is set as 8:30–17:30, are 36.6 km/h for major roads and 30.1 km/h for minor roads. The initial values of link travel time are set by these average speeds. All paths between any pair of nodes are set in advance.

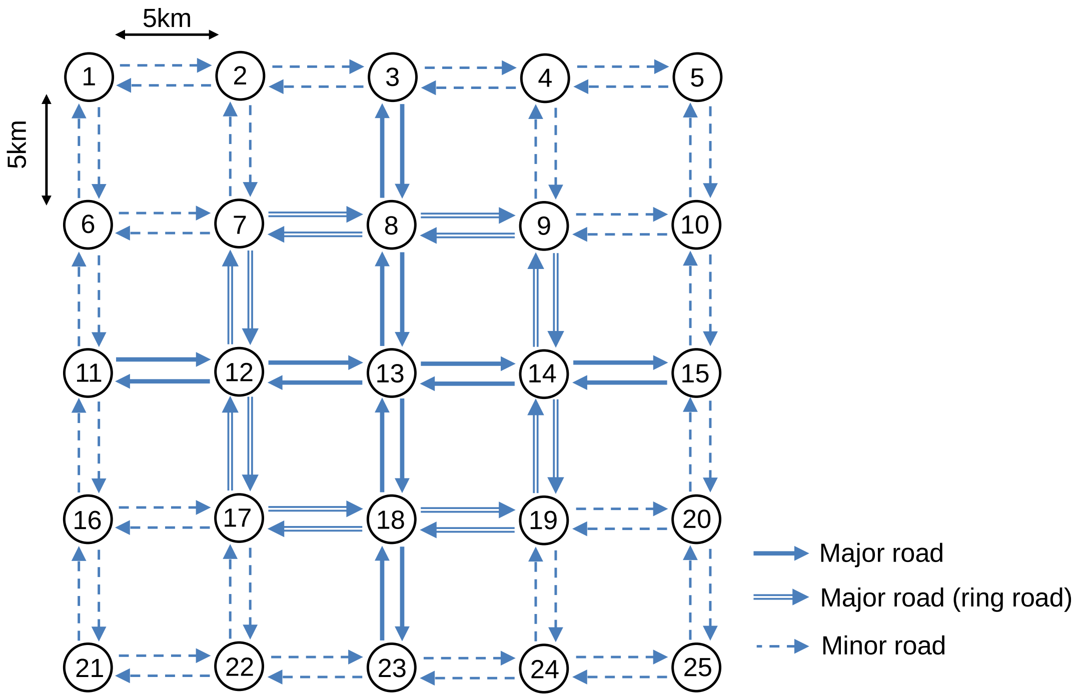


Fig. 3. Grid network for test simulation

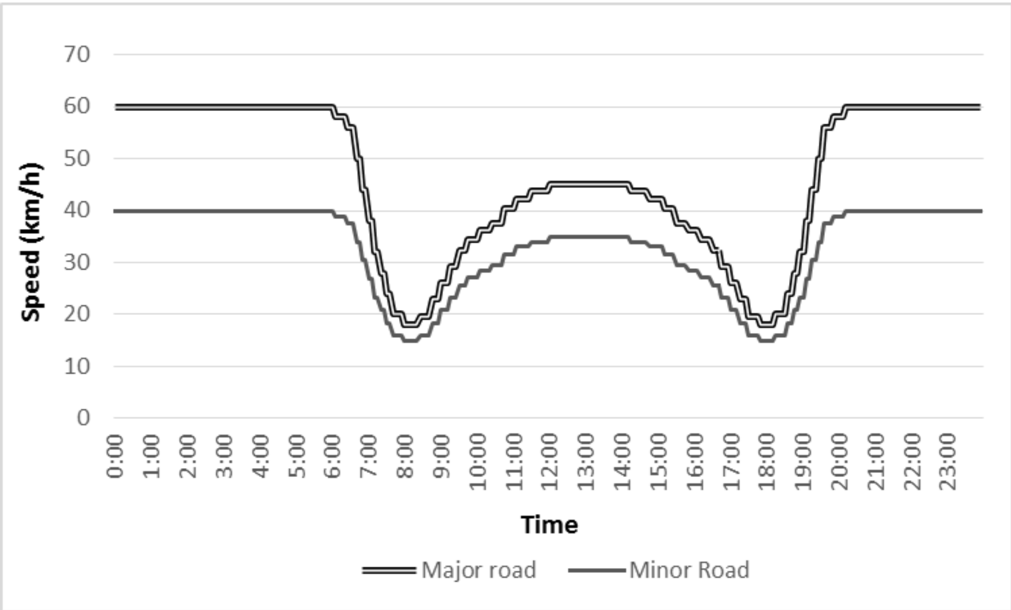


Fig. 4. Link travel speed profiles

Link travel speeds are assumed to follow a normal distribution (Kaparias et al., 2008). According to past research (Yamamoto et al., 2009), the variance is defined by the mean speed as follows,

$$v_l(h) \sim N[\bar{v}_l(h), \{\beta_l \cdot \bar{v}_l(h)\}^2] \tag{3}$$

where, for link l and for measurement period h, $v_l(h)$ is the link travel speed, $\bar{v}_l(h)$ is the mean link travel speed, and β_l is a parameter expressing the coefficient of variance. This parameter might, according to past research, be set to 0.21. In this test simulation, it is set at 0.05 for the ring road and 0.3 for all other roads. This yields an average value of coefficient of variation of 0.25. The conversion from link travel speed to link space speed is handled by the method shown in Rakha et al. (2006).

Fifty SMEs are assumed to be located evenly over the network. That is, two SMEs are located at each node. All SMEs have three customers and customer locations are set randomly. Customer locations for each SMEs are not changed through simulation study. Each SME is assumed to own only one truck and load capacity is not considered. The reason for this choice of only one vehicle with unlimited capacity is that, generally, an SME has a small number of trucks and delivery volumes are not large. It should be noted that even if elaborate settings about location of SME, number of trucks and capacity limit are considered, the findings will not be changed significantly. It will require calculating many cases for averaging the influence of variety of settings and will influence only on the number of operating days until convergence. That is, if the more complex settings are considered, the larger (smaller) number of operating days will be needed for convergence. Each customer has a 1–3 hour time window, with the size of the time window set randomly within this range. The time window size was determined by reference to the past research (Taniguchi & Shimamoto, 2004; Tang et al., 2015). Trucks start their tours during business hours. The departure time is considered at five-minute interval and sojourn time at customer is set 10 minutes.

In this case, the objective function of the problem is formulated to minimize the total cost of delivery. The total cost consists of operating cost and the penalty cost for early and late arrival at the customer location. Since the problem has been simplified by the assumption that each SME has only one vehicle with unlimited capacity, the objective function can be expressed as follows,

$$\min C(d_0, \mathbf{x}, \mathbf{k}) = C_t(d_0, \mathbf{x}, \mathbf{k}) + C_p(d_0, \mathbf{x}, \mathbf{k}) \tag{4.a}$$

$$C_t(d_0, \mathbf{x}, \mathbf{k}) = c_t \sum_{j=1}^{n} T\left(d_{j-1}, x_{j-1}, x_j, k_{j-1,j}\right) \tag{4.b}$$

$$C_p(d_0, \mathbf{x}, \mathbf{k}) = \sum_{j=1}^{n} \begin{Bmatrix} p_{early} \cdot \max\left[tw_{start,j} - d_{j-1} - T\left(d_{j-1}, x_{j-1}, x_j, k_{j-1,j}\right), 0\right] \\ +p_{late} \cdot \max\left[d_{j-1} + T\left(d_{j-1}, x_{j-1}, x_j, k_{j-1,j}\right) - tw_{end,j}, 0\right] \end{Bmatrix} \tag{4.c}$$

subject to

$$x_0 = x_n = D_0 \tag{4.d}$$

$$\prod_{j=1}^{n} \left(x_j - D_i \right) = 0 \quad \forall i = 1,2,\cdots,n \tag{4.e}$$

$$t_s \leq d_0 \leq t_e \tag{4.f}$$

where d_0 is the departure time from the depot, $\mathbf{x}$ is the vector of the customer visit order, x_j is an element of vector $\mathbf{x}$ and represents the jth customer visited, $\mathbf{k}$ is the vector of the planned path, $k_{j-1,j}$ is an element of vector $\mathbf{k}$ and represents the path from x_{j-1} to x_j. Also, $C(d_0, \mathbf{x}, \mathbf{k})$ is the total cost, $C_t(d_0, \mathbf{x}, \mathbf{k})$ is the operating cost, $C_p(d_0, \mathbf{x}, \mathbf{k})$ is the penalty cost due to early and/or late arrival, n is the number of customers, d_j is the departure time from the jth customer, $T\left(d_{j-1}, x_{j-1}, x_j, k_{j-1,j}\right)$ is the travel time on path $k_{j-1,j}$ from x_{j-1} to x_j at departure time d_{j-1}, c_t is the operating cost per minute, p_{early} and p_{late} are the penalty costs per minute for early arrival and late arrival, respectively. Additionally, $tw_{start,j}$ and $tw_{end,j}$ are the start and end times of the delivery time window for the jth customer, D_0 is the depot, D_i is the ith customer in the delivery list, and t_s and t_e are the start and end of business hours.

An alternative formulation for incorporating the penalty cost is proposed by Ando & Taniguchi (2006). In their method, the uncertainty of path travel time is considered and the expected penalty cost is calculated based on the path travel time distribution.

$$\begin{aligned} & C_p(d_0, \mathbf{x}, \mathbf{k}) \\ & = \sum_{j=1}^{n} \left\{ \begin{aligned} & \int_0^{tw_{start,j} - d_{j-1}} f\left(\omega | d_{j-1}, x_{j-1}, x_j, k_{j-1,j}\right) \cdot p_{early} \cdot \left(tw_{start,j} - d_{j-1} - \omega\right) d\omega \\ & + \int_{tw_{late,j} - d_{j-1}}^{\infty} f\left(\omega | d_{j-1}, x_{j-1}, x_j, k_{j-1,j}\right) \cdot p_{late} \cdot \left(d_{j-1} + \omega - tw_{end,j}\right) d\omega \end{aligned} \right\} \end{aligned} \tag{5}$$

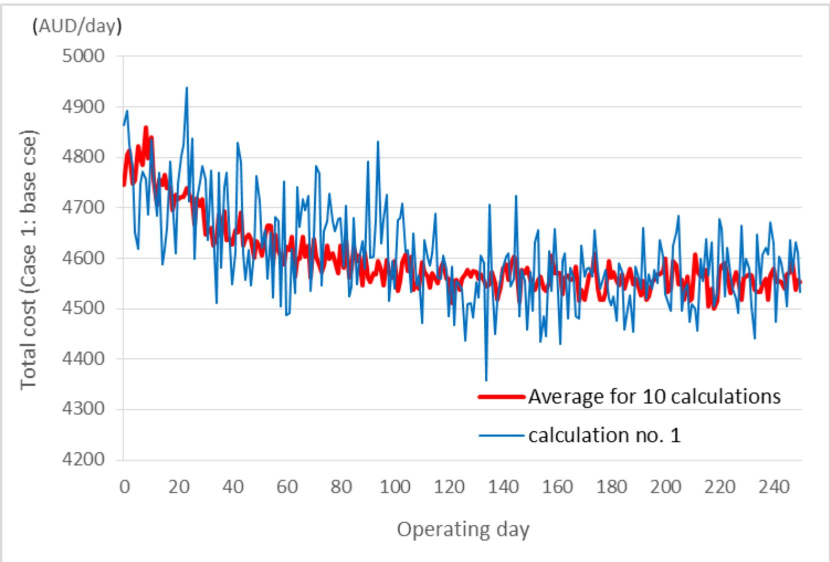
where $f(\cdot)$ is the probability density function of path travel time and is given by equation (2). In this research, the above two methods of considering penalty cost are compared.

Since a routing and scheduling problem cannot be linearized, it has no efficient solving method. In this study, for both formulations, an exhaustive method is applied to search for the optimum solution. That is, all feasible combination of departure time from the depot, customer visit order and paths are calculated and compared. The Trapezoidal rule is used for integration calculation in equation (5). In the formulation given by equation (4.c), as with many practical routing and scheduling systems, the penalty cost is calculated based on the average arrival time. This method of penalty calculation is called an “average arrival time based penalty” (AP) in this study. On the other hand, the alternative method expressed by equation (5) is called the “expected penalty” method (EP). In the following test simulation, by reference to the literature (Ando & Taniguchi, 2006), c_t is set to 0.3 AUD/min, p_{early} and p_{late} are set 0.6 AUD/min and 1.2 AUD/min, respectively. In addition, if the truck arrives back at the depot after business hours, a late arrival penalty is applied.

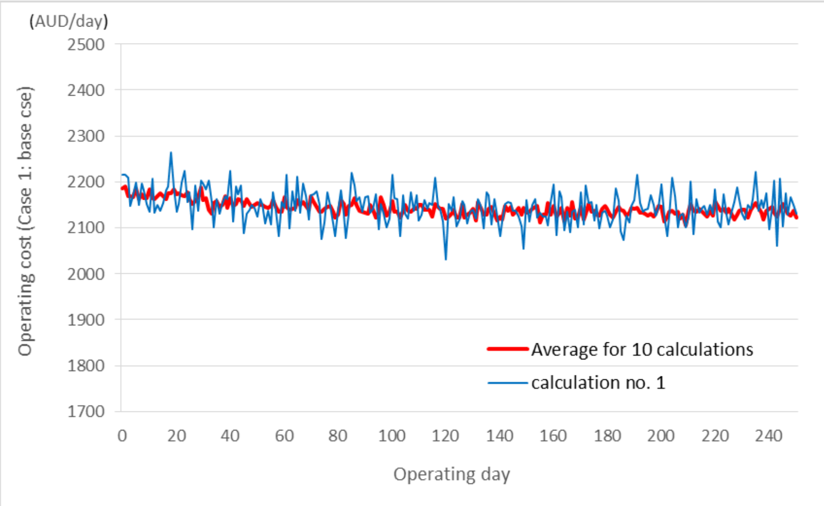
5. Test simulation results and discussion

5.1. Evolution of cost reduction attained by utilizing vehicle location data

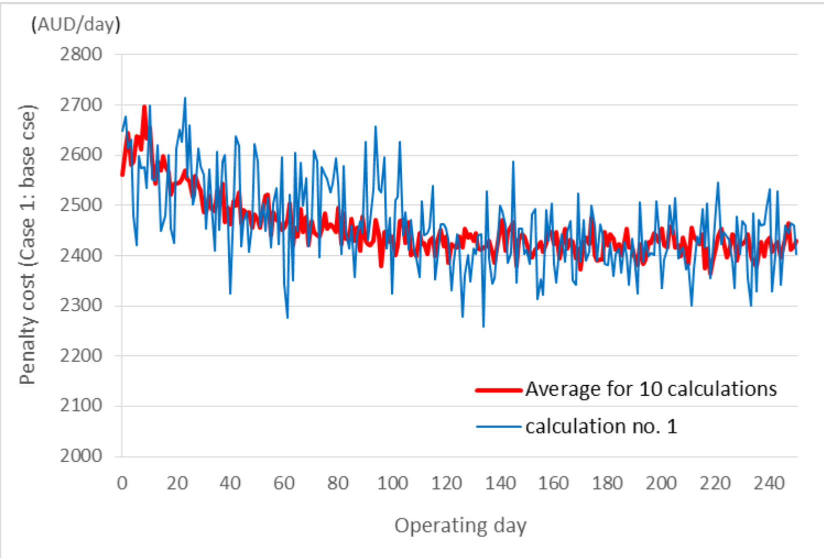
Since the test simulation is influenced by the randomness of the path and link travel times, it was conducted ten times with different random seeds. The log-normally distributed random variables were obtained using the Box-Muller method (Box & Muller, 1958). Figure 5 shows the day-to-day changes in (a) total cost, (b) operating cost and (c) penalty cost for one simulation using the AP method (blue lines). The average value for the ten simulations is shown by the bold red lines. It is clear that as data collection continues, the total cost decreases gradually. This is due to the falling penalty cost; operating cost decreases only marginally. In addition, day-to-day variations of penalty cost and total cost are large, while operating cost varies little.



(a)



(b)



(c)

Fig. 5. Day-to-day changes in cost based on arrival time penalty (AP): (a) total cost, (b) operating cost, (c) penalty cost.

Figure 6 shows the average total costs for the 10 simulations by both AP and EP. This figure shows that this routing and scheduling system utilizing location data has no effect for the first few days (until about the 10th operating day). This is due to the randomness of the link and path travel times. The rate of total cost reduction drops at around the 100th operating day in this test situation, and is near convergence by around the 200th operating day. Taking the convergence value of total cost to be the averaged cost from the 201st to the 250th operating days, the overall reduction achieved is -4.0% for the AP method and -4.5% for the EP method. These reductions are comparable to those shown in the past research which adopted the similar cost function (for example, Ando & Taniguchi, 2006). Although these reductions are small, this is partly because the initial values of link travel time are set based on average link travel speeds during business hours. If speed limits are applied (60km/h for major roads and 40km/h for minor roads), the percentage reductions become -10.4% and -10.8%, respectively. Finally, the converged total cost by EP is slightly lower than that by AP.

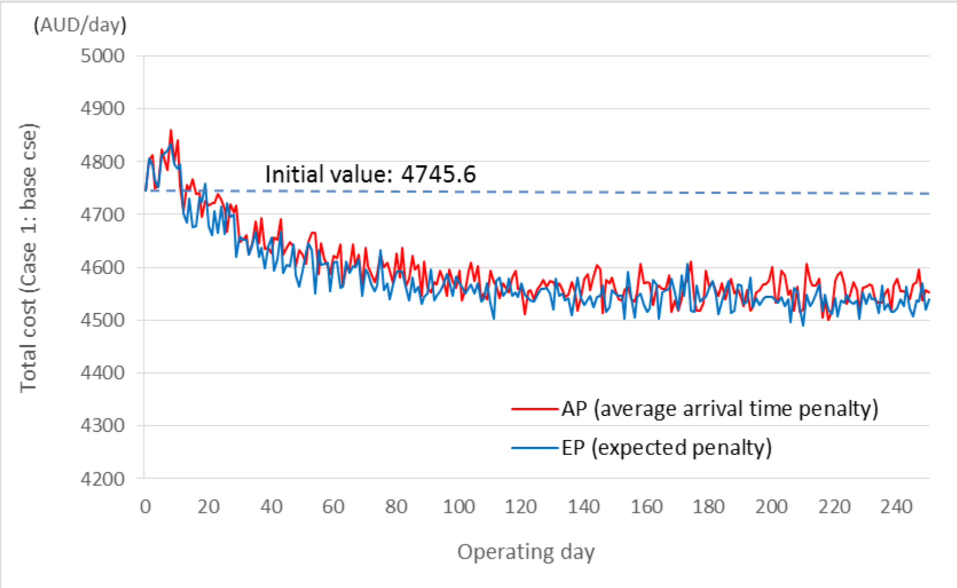


Fig. 6. Comparison of total cost by AP and EP formulations

Table 1 summarizes the converged value of total cost for six different scenarios (again with 10 simulations for each). All converged total costs are average values from the 201st to the 250th operating day. The variances of total cost from the 201st to the 250th operating day are also shown. In case 1, the base case, the data settings as explained in section 4.2 are applied. In case 2, the time windows for all customers are set to 1 hour. Case 3 is divided into two cases according to the magnitude of link speed changes. In case 3a, link speed reductions related to regulation speeds are increased by 1.5 times. On the other hand, in case 3b, the link speed reductions are decreased to one half (0.5 times). The former represents a situation with more congestion while the latter represents moderate congestion. Case 4 also consists of two cases. Case 4a considers a higher penalty cost situation, in which the penalty cost is increased by 1.5 times. On the other hand, case 4b considers a moderate penalty cost, in which the penalty cost is decreased to one half (0.5 times). Percentage figures shown in parentheses in the rightmost column are the reduction from the average total cost without location data (cost at 0th operating day).

The table shows that cost reductions are high in case 3a and low in case 2. This means that the system is particularly effective in a highly congested situation and less effective when scheduling constraints are strict. For all cases, the cost reductions achieved by EP are slightly better than those by AP. This means that taking travel time uncertainty into account contributes to the minimization of total cost, but the difference is not significant on an average. More noteworthy is the reduction of day-to-day variances of total costs by AP and EP from those of total costs without location data use. These reductions reach around -50% in many cases. Therefore, the one of the biggest merit of introduction of routing and scheduling system is the buffer effect from variations in traffic conditions on delivery cost. It is also noteworthy that there are significant differences in day-to-day variances of total costs between AP and EP. In all cases, the variances of total costs by EP are much smaller than those by AP. This means that taking travel time uncertainty into account can buffer the effect of day-to-day variations in traffic conditions. To look more deeply into this point, Table 2 shows the results of statistical tests (significant difference test) of total cost variances between AP and EP. It is found that the total cost variance by EP is smaller than that by AP with statistical significance in many cases. A reasonable explanation for the lack of significant difference of total cost variance in case 4a is that in the case of so high penalties, routing and scheduling cost is sensitive due to penalty cost. So the influence of traffic condition variance on planning cannot be reduced even if travel time uncertainty is considered.

Table 1. Summary of convergence value of total cost*[1]

			Calculation number										Average
			1	2	3	4	5	6	7	8	9	10	(% reduction*[2])
<Case 1> Base case	Total cost without location data	Mean	4735	4749	4726	4736	4,710	4,720	4,748	4,737	4,759	4,743	4,736
		Variance	7,408	9,932	6,395	11,203	9,862	10,963	11,192	8,179	14,839	7,787	9,776
	Average Arrival Time Penalty	Mean	4,570	4,554	4,529	4,585	4,547	4,547	4,555	4,545	4,539	4,565	4,554 (-3.8%)
		Variance	4,005	5,540	4,576	5,921	5,649	3,273	5,035	6,244	5,425	5,874	5,154 (-47.3%)
	Expected Penalty	Mean	4,539	4,529	4,534	4,546	4,539	4,532	4,530	4,528	4,525	4,532	4,533 (-4.3%)
		Variance	2,760	3,473	4,087	3,411	3,166	3,851	4,241	3,635	4,834	2,195	3,565 (-63.5%)
<Case 2> Strict time window	Total cost without location data	Mean	6,160	6,164	6,135	6,158	6,125	6,123	6,179	6,154	6,178	6,150	6,153
		Variance	12,413	12,279	8,241	14,145	10,223	13,776	14,502	12,668	18,914	10,617	12,778
	Average Arrival Time Penalty	Mean	5,988	5,960	5,972	5,985	5,985	5,952	5,990	5,956	5,995	5,962	5,974 (-2.9%)
		Variance	10,473	9,626	7,903	6,730	6,269	9,159	8,801	4,807	11,280	7,104	8,215 (-35.7%)
	Expected Penalty	Mean	5,932	5,968	5,957	5,959	5,933	5,962	5,960	5,962	5,965	5,977	5,958 (-3.2%)
		Variance	6,192	7,438	5,574	6,400	5,935	4,990	5,637	6,883	6,582	7,054	6,269 (-50.9%)
<Case 3a> High congestion	Total cost without location data	Mean	5,725	5,732	5,689	5,737	5,678	5,699	5,732	5,712	5,747	5,732	5,718
		Variance	16,923	21,439	12,228	26,717	20,029	23,748	25,806	17,805	32,267	21,435	21,840
	Average Arrival Time Penalty	Mean	5,395	5,419	5,384	5,427	5,414	5,382	5,393	5,394	5,392	5,393	5,399 (-5.6%)
		Variance	15,522	17,461	13,040	16,085	9,579	12,927	18,948	11,125	15,599	13,374	14,366 (-34.2%)
	Expected Penalty	Mean	5,375	5,340	5,338	5,354	5,342	5,328	5,385	5,384	5,353	5,355	5,355 (-6.3%)
		Variance	13,730	9,073	11,400	11,699	6,872	8,166	13,998	15,364	11,011	11,808	11,312 (-48.2%)

Table 1 (cont.). Summary of convergence value of total cost

			Calculation number										Average
			1	2	3	4	5	6	7	8	9	10	(% reduction*[2])
<Case 3b> Low congestion	Total cost without location data	Mean	4,144	4,160	4,153	4,154	4,137	4,146	4,161	4,155	4,162	4,147	4,152
		Variance	2,588	4,502	3,348	3,069	2,585	2,970	3,245	4,266	5,113	2,276	3,396
	Average Arrival Time Penalty	Mean	4,069	4,067	4,056	4,066	4,059	4,052	4,053	4,053	4,055	4,054	4,058 (-2.3%)
		Variance	1,437	2,057	1,485	1,613	1,590	1,391	1,437	1,343	1,354	996	1,470 (-56.7%)
	Expected Penalty	Mean	4,038	4,057	4,045	4,048	4,048	4,033	4,034	4,049	4,051	4,045	4,045 (-2.6%)
		Variance	919	1,021	1,129	1,721	1,293	1,133	1,458	1,220	723	968	1,159 (-65.9%)
<Case 4a> High time penalty	Total cost without location data	Mean	3,445	3,452	3,431	3,447	3,433	3,436	3,449	3,442	3,459	3,451	3,445
		Variance	2,991	3,375	2,165	4,488	3,662	4,635	4,006	3,571	5,253	2,756	3,690
	Average Arrival Time Penalty	Mean	3,324	3,339	3,326	3,341	3,323	3,329	3,320	3,322	3,325	3,339	3,329 (-3.4%)
		Variance	1,373	1,905	1,946	1,361	1,777	1,566	1,676	1,568	1,468	2,335	1,698 (-54.0)
	Expected Penalty	Mean	3,321	3,323	3,322	3,326	3,317	3,315	3,313	3,324	3,337	3,319	3,322 (-3.6%)
		Variance	1,323	1,468	1,756	1,075	1,209	1,943	1,089	2,334	1,907	2,327	1,643 (-55.5%)
<Case 4b> Low time penalty	Total cost without location data	Mean	6,038	6,060	6,042	6,046	6,007	6,018	6,066	6,045	6,072	6,042	6,044
		Variance	16,222	23,827	12,714	21,939	18,075	22,837	21,544	17,595	31,094	16,531	20,238
	Average Arrival Time Penalty	Mean	5,767	5,777	5,782	5,741	5,733	5,769	5,774	5,754	5,745	5,735	5,758 (-4.7%)
		Variance	9,800	10,895	13,052	10,615	9,665	7,165	9,684	9,445	12,548	10,625	10,349 (-48.9%)
	Expected Penalty	Mean	5,751	5,738	5,743	5,733	5,737	5,742	5,752	5,758	5,749	5,752	5,745 (-4.9%)
		Variance	12,773	7,034	8,446	5,199	6,270	9,177	9,738	8,931	10,150	9,522	8,724 (-56.9%)

*[1] Convergence values of total costs are averaged values from the 201st to the 250th operating day.

*[2] Percentage reductions are calculated against the average total cost without location data.

Table 2. Statistical tests of total cost variance

		Total cost variance (n = 10) Mean	s.d.	Significant difference test of mean values (t-stat, d.f. = 18)
<Case 1> Base case	Average Arrival Time Penalty	5,154.2	940.8	4.16**
	Expected Penalty	3,565.3	756.0	
<Case 2> Strict time window	Average Arrival Time Penalty	8,215.2	2018.1	2.86*
	Expected Penalty	6,268.6	753.5	
<Case 3a> High congestion	Average Arrival Time Penalty	14,366.0	2,879.5	2.46*
	Expected Penalty	11,312.1	2,679.0	
<Case 3b> Low congestion	Average Arrival Time Penalty	1,470.4	267.1	2.52*
	Expected Penalty	1,158.6	285.2	
<Case 4a> High time penalty	Average Arrival Time Penalty	1,697.6	302.5	0.30
	Expected Penalty	1,643.1	479.6	
<Case 4b> Low time penalty	Average Arrival Time Penalty	10,349.3	1,658.6	1.89(*)
	Expected Penalty	8,724.1	2,153.3	

** $p < 0.01$, * $p < 0.05$, (*), $p < 0.1$

5.2. Effect of sharing of location information

Figure 7 shows the comparison of average total costs between the system with the location information sharing and that without sharing among SMEs. From the figure, it can be found that the sharing of location information can accelerate the efficiency, although the convergence costs are almost the same. Obviously, the reason for the difference in efficiency is that the data points in each measurement period can be accumulated fast by using location information of other SMEs. The relatively big differences are shown around early operating days when enough data points are not accumulated.

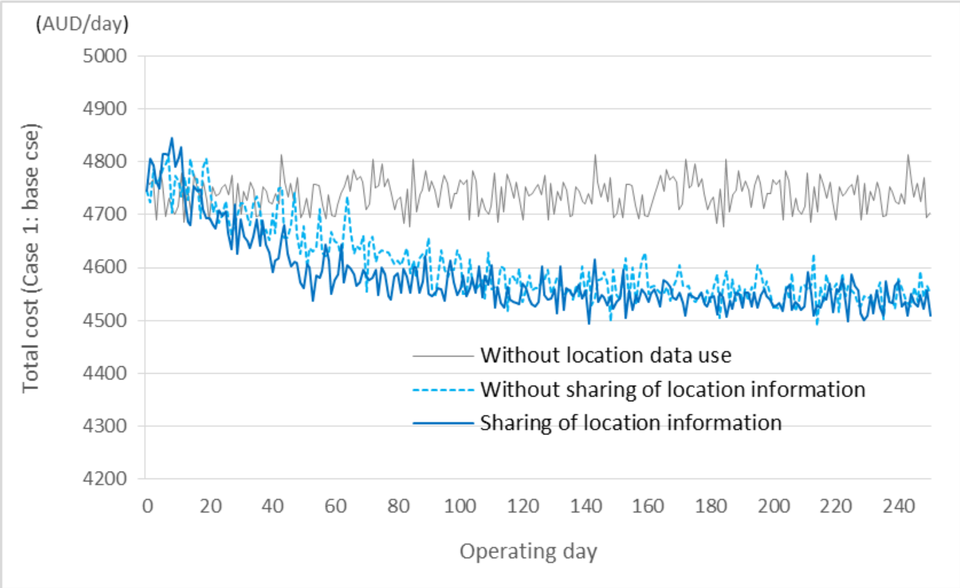


Fig. 7. Sharing of location information and system efficiency

5.3. Relationship between amount of historical data and system efficiency

Although the results given in the previous subsection demonstrate the evolution of routing and scheduling along operating days, the relationship between the amount of data in the historical database and system efficiency is not demonstrated. To investigate this, Figure 8 shows the relationship between the number of data points in the historical database and system efficiency. In the analysis here, the routing and scheduling problem was calculated while increasing the number of data points in each measurement period one by one. Since all simulations showed similar results, the result of only case 1 is shown in the figure.

It is clear from the figure that just five new data points in each measurement period gives good system efficiency, while system efficiency is close to convergence with thirty data points. The average total costs from the 101st to the 150th data points are 4,537.3 for AP and 4,527.2 for EP, respectively, representing percentage reductions of -4.8% for ATP and -5.0% for EP, respectively. This means that the convergence total costs indicated in the previous subsection are not true convergence values. It can be thought that this difference between the convergence values shown here and those shown in the previous subsection is due to the unevenness of data points in the historical database. That is, the presence of measurement periods with sparse data can lead to unreliable routing and scheduling. Therefore, to realize a more reliable and efficient system, even data collection in wider area is required than dense data in a part of links.

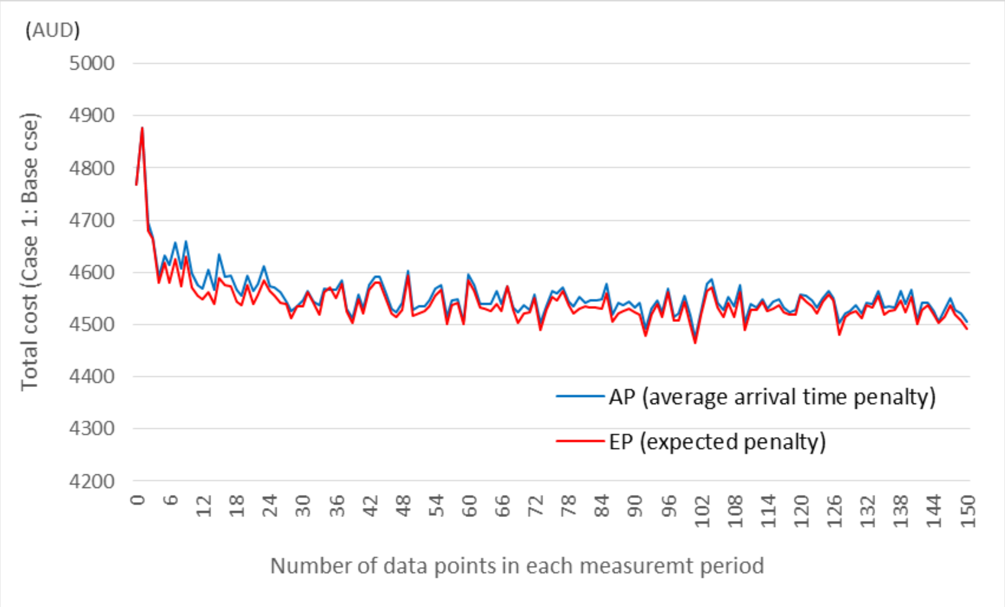


Fig. 8. Number of accumulated data points and system efficiency

6. Conclusion and future research

This paper describes a concept for an autonomous routing and scheduling system for SMEs and offers a perspective on its effectiveness through test simulations. A preliminary literature review indicates that an inexpensive routing and scheduling system would find wide acceptance among SMEs with logistics tasks and would increase the efficiency of their deliveries. It is proposed that, to establish such a system, utilizing vehicle location data as traffic information is of considerable importance. This leads to a conceptual system in which SMEs feed information to the system on a daily basis through the web while the system returns routing and scheduling assistance. A driver would be fed guidance via a smartphone or tablet computer in the vehicle, which in turn would feed the system with the current location of each vehicle and also proof of delivery. The location data fed back to the server via the mobile internet is used as probe vehicle data.

Test simulations demonstrate the effect of utilizing vehicle location data on delivery efficiency. In particular, the evolution of the system as location data is gathered day-by-day is analyzed. The results show that the system leads to improved efficiency mainly as a result of reduced penalty costs and buffer effect from variations in traffic conditions on delivery cost. The buffer effect is enhanced by taking travel time uncertainty into consideration. Additionally, the presence of measurement periods with sparse data points results in unreliable routing and scheduling. For a more reliable system, data collection over a wider area is required rather than dense data in a subset of links.

Future research on this topic should consider the algorithm for finding a path with minimum penalty cost. In the case that there is no effective algorithm, k shortest path algorithm can be used (for example, Yen, 1971). A further area for improvement would be to make use of vehicle location data in real time, which has been shown in past research to be beneficial. This would entail the real-time use of vehicle location data to carry out dynamic optimization of the routing and scheduling problem. Finally, more realistic outputs would be obtained if the other problems and systems surrounding the routing and scheduling problem, such as diffusion of mobile information service, interactions between trucks and other road users, and so on, were included in the simulation (Wang, 2010; Cagliano et al., 2014).

References

Ando, N. & Taniguchi, E. (2006). Travel Time Reliability in Vehicle Routing and Scheduling with Time Windows. *Network and Special Economics* 6, 293-311.

Baldacci, R., Battarra, M. & Vigo, D. (2008). Routing and Heterogeneous Fleet of Vehicles, In Golden et al. (eds.), The Vehicle Routing Problem (3-27): Springer.

Bianchessi, N. & Righini G. (2007). Heuristic algorithm for the vehicle routing problem with simultaneous pick-up and delivery. *Computers & Operations Research* 34, 578-594.

Box, G.E.P. & Muller, M.E. (1958). A note on the generation of random normal deviates. *The Annals of Mathematical Statistics* 29, 610-611.

Bin, X., Xiaohong, C., Hangfei, L. & Chao, Y. (2013). Decision Oriented Intelligent Transport Information Platform Design Research – Case study of Hangzhou City. *Procedia – Social and Behavioral Sciences* 96, 2230-2239.

Bräysy, O. & Gendreau, M. (2005). Vehicle Routing Problem with Time Windows, Part I: Route Construction and Local Search Algorithms. *Transportation Science* 39(1), 104-118.

Cagliano, A.C., Gobbato, L., Tadei, R. & Perboli, G. (2014). ITS for E-grocery Business: the Simulation and Optimization of Urban Logistics Project. *Transportation Research Procedia* 3, 489-498.

Chen, B.Y., Lam, W.H.K., Sumalee, A. & Li Z. (2012). Reliable shortest path finding in stochastic networks with spatial correlated link travel times. *International Journal of Geographical Information Science* 26(2), 365-386.

Chen, K., Yu, L., Guo, J., & Wen, H. (2007). Characteristics analysis of road network reliability in Beijing based on the data logs from taxis. Proceedings of the 86th annual meeting of the Transportation Research Board, Washington, DC.

Clarke, G. & Wright, J.W. (1964). Scheduling of Vehicles from a Central Depot to a Number of Delivery Points. *Operations Research* 12(4), 568-581.

Comendador, J., Lopez-Lambas, M. E., Monzon, A. & Hengliang, X. (2012). Seeking Efficiency in Urban Freight Distribution: A Comparative Analysis of Courier and Food Distribution Companies. Proceedings of the 91st annual meeting of the Transportation Research Board, Washington, DC.

Dantzig G.B. & Ramser, J.H. (1959). Truck dispatching problem. *Management Science* 6(1),

80-91.
Derrouiche, R., Moutaoukil, A. & Neubert, G. (2014). Integration of Social Concerns in Collaborative Logistics and Transportation Networks. In: Camarinha-Matos, L.M. & Afsarmanesh, H. (eds.), Collaborative Systems for Smart Networked Environments (730–738): Springer.

Desrochers, M. & Verhoog, T.W. (1991). A new heuristic for the fleet size and mix vehicle routing problem. *Computers & Operations Research* 18(3), 263-274.

Ehmke, J.F., Meisel, S. & Mattfeld, C. (2012). Floating car based travel times for city logistics. *Transportation Research Part C* 21, 338-352.

Fazekas, Z., Gáspár, P., Biró, Z. & Kovács, R. (2014). Driver Behavior, Truck Motion and Dangerous Road Locations – Unfolding Form Emergency Braking Data. *Transportation Research Part E* 65, 3-15.

Figliozzi, M.A. (2009). A Route Improvement Algorithm for the Vehicle Routing Problem with Time Dependent Travel Times. Proceedings of the 88th Transportation Research Board Annual Meeting, Washington, DC.

Gayialis, S.P. & Tatsiopoulos, I.P. (2004). Design of an IT-driven decision support system for vehicle routing and scheduling. *European Journal of Operational Research* 152, 382-398.

Gendreau, M., Laporte, G. & Séguin, R. (1996). Stochastic vehicle routing. *European Journal of Operational Research* 88, 3-12.

Golob, T.F. & Regan, A.C. (2003). Traffic Congestion and Trucking Managers' Use of Automated Routing and Scheduling. *Transportation Research Part E* 39, 61-78.

Grakovski, A., Kabashkin, I. & Ressin, A. (2008). Intelligent Transport System for Intra-City Logistics Based on WWW Technologies. *Transport and Telecommunication* 9(3), 30-38.

Gunasekaran, A., Ngai, E.W. & Cheng, T.C.E. (2007). Developing an E-Logistics System: A Case Study. *International Journal of Logistics* 10(4), 333-349.

Haghani, A. & Jung, S. (2005). A Dynamic Vehicle Routing Problem With Time-Dependent Travel Times. *Computers & Operations Research* 32, 2959-2986.

Herrera, J.C., Work, D.B., Herring, R., Ban, X., Jacobson, Q. & Bayen, A.M. (2010). Evaluation Of Traffic Data Obtained Via GPS-Enabled Mobile Phones: The Mobile Century Field Experiment. *Transportation Research Part C* 18, 568-583.

Ichoua, S., Gendreau, M. & Potvin, J.Y., (2003). Vehicle Dispatching With Time-Dependent Travel Times. *European Journal of Operational Research* 144, 379-396.

Institute for Operations Research and the Management Sciences, Vehicle Routing Software Survey. http://www.orms-today.org/surveys/Vehicle_Routing/vrss.html. (accessed July 19, 2015).

Kaparias, I., Bell, M.G.H & Belzner, H. (2008). A New Measure of Travel Time Reliability for In-Vehicle Navigation Systems. *Journal of Intelligent transportation Systems* 12(4), 202-211.

Kaparias, I., Bell, M.G.H., Chen, Y. & Bogenberger, K. (2007). ICNavS: A Tool For Reliable

Dynamic Route Guidance. *IET Journal Intelligent Transport Systems* 1(4), 225-233.
Kenyon, A.S. & Morton D.P. (2003). Stochastic Vehicle Routing with Random Travel Times. *Transportation Science* 37(1), 69-82.
Kim, M., Miller-Hooks, E. & Nair, R. (2011). A Geographic Information System-Based Real-Time Decision Support Framework for Routing Vehicles Carrying Hazardous Materials. *Journal of Intelligent Transportation Systems* 15(1), 28-41.
Kuswantoro, F. & Rosli, M.M. (2012). Logistics Efficiency and Firm Performance: Evidence from Indonesian Small and Medium Enterprises. *American International Journal of Contemporary Research* 2(6), 102-111.
Laporte, G. (1992). The Vehicle Routing Problem: An overview of exact and approximate algorithms. *European Journal of Operational Research* 59, 345-358.
Luo, Z., Qin, H., Zhang, D. and Lim, A. (2016). Adaptive large neighborhood search heuristics for the vehicle routing problem with stochastic demands and weight-related cost. *Transportation Research Part E* 85, 69-89.
Maden, W., Eglese, R. & Black, D. (2010). Vehicle routing and scheduling with time-varying data: A case study. *Journal of the Operational Research Society* 61, 515-522.
Mancini, S. (2016). A real-life Multi Depot Multi Period Vehicle Routing Problem with a Heterogeneous Fleet: Formulation and Adaptive Large Neighborhood Search based Metaheuristic. *Transportation Research Part C* 70, 100-112.
Min, H. (1989). The multiple vehicle routing problem with simultaneous delivery and pick-up points. *Transportation Research Part A* 23(5), 377-386.
Miwa, T., Ishiguro, Y., Yamamoto, T. & Morikawa, T. (2015). Allocation planning for probe taxi devices aimed at minimizing losses to travel time information users. *Journal of Intelligent Transportation Systems* 19(4), 399-410.
Pan, H.H., Wang, S.C. & Yan, K.Q. (2014). An integrated data exchange platform for Intelligent Transportation Systems. *Computer Standards & Interfaces* 36, 657-671.
Potvin, J.Y., Xu, Y. & Benyahia, I. (2006). Vehicle Routing and Scheduling with Dynamic Travel Times. *Computers & Operations Research* 33, 1129-1137.
Rakha, H., El-Shawarby, I., Arafeh, M. & Dion, F. (2006). Estimating Path Travel-Time Reliability. ITSC 2006 – 9th International IEEE Conference on Intelligent Transportation Systems, Toronto, Canada, 236-241.
Simroth, A. & Zähle, H. (2011). Travel Time Prediction Using Floating Car Data Applied to Logistics Planning. *IEEE Transactions on Intelligent Transportation systems* 12(1), 243-253.
Solomon, M.M. (1987). Algorithms for the Vehicle Routing and Scheduling Problems with Time Window Constraints. *Operations Research* 35(2), 254-265.
Taniguchi, E. & Shimamoto, H. (2004). Intelligent Transportation System Based Dynamic Vehicle Routing And Scheduling With Variable Travel Times. *Transportation Research Part C* 12, 235-250.

Tang, J., Yu, Y. & Li, J. (2015). An exact algorithm for the multi-trip vehicle routing and scheduling problem of pickup and delivery of customers to the airport. *Transportation Research Part E* 73, 114-132.

Transport for NSW. Trip Planner. http://www.transportnsw.info/en/index.page?. (accessed May 25, 2016)

Wang, F.Y. (2010) Parallel Control and Management for Intelligent Transportation Systems: Concepts, Architectures, and Applications. *IEEE Transactions on Intelligent Transportation systems* 11(3), 630-638.

Wang, Z., Goodchild, A. & McCormack, E. (2015) Measuring Truck Travel Time Reliability Using Truck Probe GPS Data. *Journal of Intelligent Transportation Systems* 20(2), 103-112.

Yamamoto, T., Miwa, T., Takeshita, T. & Morikawa, T. (2009). Updating Dynamic Origin-Destination Matrices Using Observed Link Travel Speed By Probe Vehicles. *Transportation and Traffic Theory 2009*, 723-738.

Yen J.Y. (1971). Finding the K Shortest Loopless Paths in a Network. *Management Science* 17(11), 712-716.

Zeimpekis, V. & Giaglis, G.M. (2006). Urban Dynamic Real-Time Distribution Services; Insights from SMES. *Journal of Enterprise Information Management* 19(4), 367-388.

Zhang, J., Wang, F.Y., Wang, K., Lin, W.H., Xu, X. & Chen, C. (2011). Data-Driven Intelligent Transportation Systems: A Survey. *IEEE Transactions on Intelligent Transportation systems* 12(4), 1624-1639.