Optimization of Convenience Stores' Distribution System with Web Scraping and Google API Service

Thai Quang LE*, Davar PISHVA**

*International Management, Ritsumeikan Asia Pacific University, 1-1 Jumonjibaru, Beppu, Oita 874-8577 Japan

**Asia Pacific Studies, Ritsumeikan Asia Pacific University, 1-1 Jumonjibaru, Beppu, Oita 874-8577 Japan

quanle11@apu.ac.jp, dpishva@apu.ac.jp

Abstract*—Vehicle Routing Problem (VRP) has never become an obsolete research theme in the field of operations research and supply chain management. Considering that significant number of researchers have already tried addressing VRPs with mathematical modeling and algorithmic approaches, this paper focuses on a practical implementation and employs programming techniques to cope with a particular business problem in convenience stores' distribution system. It optimizes goods distribution process of convenience stores business, which involves lorries delivering products from a warehouse to a network of several convenience stores in a single trip, collecting their garbage, passing by a gas station for re-fueling when needed, and returning back to the warehouse. A mathematical 'network flow model' is initially developed to examine the problem. Geographical data of convenience stores, their associated warehouses, garbage dumpsites and gas stations are subsequently retrieved through programming with the 'web scraping' technique. A prototype of web-based delivery navigation system that utilizes Google API service is then developed to solve the optimal convenience stores' networking problem. Furthermore, a more general perspective of the problem is illustrated with cluster-first-route-second heuristic algorithm and a mobile version of the prototype, which can serve as a real time navigation system for delivery truck drivers, is developed. Validity of obtained results is also examined by other known methods to justify optimality and fast performance of the approach.

***Keyword*—Assignment, Google API Service, Maximal Covering, Modeling, Optimization, Supply Chain, Web Scraping.**

I. INTRODUCTION

It is fair to claim that profit maximization is the ultimate goal of any commercial firms. In a retailing business that comprises of thousands of stores or stations, even a small reduction in product cost can have a big impact on not only the business itself, but also to the nation's economy and social well-being as a whole. Over many decades, there have been many profound researches and proven practices suggesting various ways to make business more lucrative, but most of them focus on effective resources allocation and enhancing productivity, which then lead to production cost reduction. Moreover, in our push-button world where everything changes so quickly along with the high-pace development of technology, machines are gradually replacing the human role in several aspects, such as manufacturing procedures and business administration. Since potential effects of technology on business has not fully been exploited and more and more advanced and pragmatic technology is being invented by an increasing flow of capital investment in research and development (R&D) on a global scale, technology is believed to be one of the decisive keys in business competition today.

Furthermore, apart from profit maximization, more and more businesses are presently paying serious attention to corporate social responsibility (CSR), especially in environmental aspects. It has been a common sense to utilize the perspective of management in different aspects of doing business, and one argument for explaining this is that most business factors such as human capitals, production cost, market sales, etc. are measureable by human beings. However, recently technology has also enabled us to measure different types of environmental factors such as amount of carbon-dioxide emission and so on that are resulted from prolong business process. Data from such measurements force business managers to take environmental factors into consideration and lead the business with a more balance strategy between profit maximization and social responsibility. The challenge here is how to help business stay more lucrative, and at the same time remain friendly to the society and the environment. This is also the question that this paper would like to address, particularly in convenience store retailing business in Japan for which a systematic technological approach is suggested as an answer.

The remainder of the paper is organized as follows: Section 2 addresses delivery route optimizing perspective of convenience stores' distribution problem and some related previous work. Section 3 then delivers a practical solution to enhance the distribution process of a network of convenience stores, which includes network flow modeling, data collection by web-scraping technique, multi-objective route

---

Manuscript received on November 21, 2014. This work is a follow up of an accepted conference paper for the 17th International Conference on Advanced Communication Technology.

Thai Quang LE is a student at Ritsumeikan Asia Pacific University (APU) in the field of Business Administration focusing on Strategic Management and candidate for graduation.

Davar Pishva is a professor in ICT at Ritsumeikan Asia Pacific University (APU) Japan (corresponding author: +81-977-78-1000, fax: +81-977-78-1001, e-mail:dpishva@apu.ac.jp).

optimization with Google Map API and Evolutionary Algorithm in Excel, and an easy to follow Graphic User Interface (GUI) implementation on web-based mobile platform. Section 4 proceeds with detailed discussion and analysis of the research result. Sections 5 subsequently indicate some current limitations and potential future research and conclude the paper.

II. LITERATURE REVIEW

A. Convenience Store Business in Japan

Considering that Japanese people is famous for tireless dedication to working-life, convenience store turns out to be one of the key factors in timesaving for the majority of Japanese and lead to higher overall working productivity of the society. Due to the increasing dependency of Japanese society on convenience stores, growth in this retailing business is also a reflection of the Japanese present living standards [1]. Japan’s convenience stores are well known for catering not only fruitful types of beverages, safe and healthy foods [2], but also to a plenty of in-store services, namely banking, printing, public utility payment (gas, telephone, electricity, etc.), ticket reservation for concerts or sport games and so on [3]. However, it is also claimed that not just the merchandising factor, but also time beneficial factor of convenience stores is important to the Japanese daily life [1]. To illustrate this, since most Japanese visit stores on foot [1], a high-density of roughly over five stores per kilometer, which can be much higher in metropolitan areas, enables anyone to access a store in less than 10 minutes on foot on average. Thus, it is reported that approximately sixty percent of customers visit the same store twice a week or more [2].

Convenience store business is one of the most profitable retailing businesses in Japan, covering 36% of the Japan Retail Food Market in the year 2012 [4-5]. With the increasing total of sales from roughly 8.7 trillion yen in the year 2011 to about 9.4 and 9.8 trillion yen in the year 2012 and 2013 respectively [4]. Regarding the number of stores, there are over 50,000 convenience stores located throughout 47 prefectures of Japan and the number of new establishments has been continuously increasing in recent years; from 43,373 stores in 2011 to 47,801 and 50,234 stores in 2012 and 2013 respectively. Most of the convenience stores are established in the metropolitan Kanto area, with more than 20,000 stores in the year 2013; and Okinawa has the lowest number of convenience stores, about 500 stores in 2013. Since most of the convenience stores learn the system from supermarket management experience, with their limitation of the small store sizes, convenience stores business is adapting many of state-of-art technologies, especially Electronic-Ordering-System (EOS) in collecting and analyzing Point-Of-Sales (POS) data in order to reduce effectively product delivery time from the wholesalers to each store [1, 6].

B. Convenience Store’s Distribution Problem

The price charged for a similar item being sold in convenience store is usually more expensive (by about twenty percent) than in local supermarkets or mass merchandiser [7]. This overhead cost is considered as charges for the convenience values perceived by the customers. To illustrate, one can both purchase foods for dinners and pay the electric bill at a nearby convenience store, instead of passing by both a supermarket and a power utility office, which are not usually located in the same neighborhood. However, from the perspective of business managers, the overhead cost of convenience stores is mostly resulted from overnight operation and expensive merchandise distribution process.

In Japan, convenience stores maintain long business hours, usually beginning at seven in the morning and stay open until eleven at night. Because the average salary for a night-work shift (from six until midnight) is sometimes three times higher than the normal working hour, the convenient operating hours raise an additional cost for the business. Additionally, considering limited available land area in its vicinity, convenience store in Japan is usually defined as self-service store with average of 100 to 150 square meter floor space [1]. Because of such small size, convenience stores in Japan normally do not practice inventory, and a well-managed product distribution system is essential to the business development [1], which is called just-in-time delivery. Particularly, all of the products are distributed by trucks from a distribution center (warehouse) to all of the convenience stores on a regular basis. For example, 7-eleven convenience stores implement a combined distribution system by temperature zones by which rice, chilled products are being delivered three times per day, frozen and beverage products are being delivered three to seven times per week [2]. Because products have to be delivered to thousands of convenience store on an average 9 times per day (in the case of 7-eleven stores), the cost for hiring truck drivers and gas consumption rise up. And this expensive distribution process is believed to be the biggest overhead cost associated with the price.

Due to the limitation of floor space, convenience stores do not have the capacity to do inventory. Hence, in order to meet customer’s needs, many of the convenience stores nowadays practice just-in-time distribution system, by which the products are only being delivered when needed. Therefore, on a daily basis, convenience retailers have to deliver goods by trucks from a distribution center to several stores, fill-up the gasoline when necessary, collect all the garbage, and dump the collected garbage at a designated dumpsite; and this distribution patterns have to be repeated many times per day. Throughout the years, convenience store business has found many ways to enhance the situation. For example, 7-eleven convenience stores pioneered the so-called temperature-independent distribution system by which mixed categories are combined before being loaded onto trucks so as to increase payload efficiency and tremendously decrease the replenishments of goods per day [2]. Nevertheless, in order to serve millions of people per day, still thousands of delivery vehicles, mainly trucks and mini vans, at least 3,799 of which belong to 7-eleven, are operating everyday and raising both economic and environmental concerns.

Hence, this research approaches the convenience stores’ distribution problem as a vehicle routing optimization problem, which proposes a systematic way to minimize the total truck delivery time duration in networks of convenience stores, resulting in both financial and environmental savings.

C. Related Works

There have been a considerable number of research investigating vehicle routing problems (VRPs) in various fields, which can be considered as linear optimization with different operational objectives and constrains. In fact, many real-life VRPs comprise of huge number, sometimes hundreds or thousands, of nodes in a networks and it is very time-consuming as well as computationally expensive to obtain an exact optimality in the result. Because of such large scale and demand for multivariable constraints and decision parameters, most of the proposed solutions for VRP are based on heuristics, which can be classified as classical heuristics and meta-heuristic [8]. Just to cite, some of the classical heuristics are Clarke and Wright (1964) [9], Sweep, Petal, etc. Meta-heuristics include evolutionary algorithm [10], Taburoute [11], adaptive memory based tabu search [12], ant colony optimization [13], hybrid cooperating meta-heuristics [14] and so on.

There have also been several researches that implement or extend those algorithms in empirical businesses that practice routing optimization. A decision support system proposed by Lin, Choy, Ho, Lam, Pang, Chin (2014), which also considers level of service as a cost function and has a prototype on mobile-client for courier operation to minimize the total traveling distance [15], is one such example. Unlike convenience stores distribution problem, this problem of courier service involves the continuous alternation in the customer orders, making dynamic routing and flexible time windows as the main concentration of the research. Another recent VRP example is pickup and delivery of customers to the airport, which is examined by Tang, Yu, Li [16]. The research makes use of the CPLEX software to implement an improved label-correcting method to remove infeasible routes to speed up the route search process. A related work had also been carried out by Li, Chen, Sivakumar and Wu [17], which involves optimization of travel time for inventory routing in petroleum business. Due to the nature of the business, new constrains such as hours-of-service regulations of the industry, etc. were considered. Waste collection problem is another type of VRP concentrating on balancing the workload over time by optimizing the decision on container selection and routing, obtaining a cost savings of up to 40% [18]. Due to the uncertainties in traffic congestion and the importance of time management, several researchers have also incorporated time windows as constrains in building the model of the VRP. For example, a research done by Zarandi, Hemmati, Davari, and Turksen (2013) on the location-routing problem with time windows under uncertainty inspects the fuzzy attributes of demands of customers and travel time [19]. Tas, Jabali, Woensel (2014) also introduced a solution procedure to resolve the vehicle routing problem with flexible time windows, in which the customers can be flexibly served before or after a preset tolerable time frame in order to save operational cost [20]. In 2010, in order to tackle with driving hours regulations, a research on departure time optimization for VRP was also proposed by Kok, Hans, and Schutten (2011), which claim to reduce 15% of duty time of truck drivers with the newly introduced algorithm [21]. Regarding convenience store business, Bhusiri, Qureshi, and Taniguchi (2013) proposed a perspective of convenience stores' distribution as a "VRP with soft time windows and simultaneous pickups and deliveries", which focuses on using branch-and-price approach and evolutionary algorithm to optimize the problem with delivery arrival time restriction and tests with a public benchmark test set [22-23].

Because of the limited real business data availability, despite the countless investigations in the VRP, many of them are short of testing the approach on empirical data or real cases. Hence, just few of the related previous works such as [15, 24] can demonstrate a working prototype where the proposed algorithms actually work on practical geographical input data. Furthermore, even though most of the proposed approaches are recognized to have the ability to solve a specific problem on mathematical model, usually the solutions are very complex and expensive in terms of computation.

D. Research Objective

This paper wants to put a forward into the field by introducing a practical methodology to optimize the routing of convenience stores' distribution process, especially in Japan. By utilizing the web-scraping technique to collect geographical data of all of the convenience stores and gas stations in Japan (nearly 100,000 real locations), this research aims to propose a practical solution and benchmark with whole population of real data, rather than mathematical theory and synthetic simulation.

Specifically, the research implements systematic computer software and a working web-based application so as to have the postal code of a distribution network entered by a user, and optimize traveling time of drivers delivering products from a warehouse to the network that includes up-to-8 convenience stores, pass by a gas stations on the route when needed, collect all the garbage and return to the warehouse to complete the process. In this problem, we recognize the practical need of searching for an optimal location of a gas station for the delivery trucks' re-fueling along the trip, which to the best of our knowledge, has not been introduced in the literature so far.

Moreover, the software is further tested with convenience stores networks of over 1500 areas (with different postal codes) in Oita prefecture where the authors reside. Furthermore, industry-proven Google API service is also introduced in order to ensure the retrieval of the real-time traffic data used in the optimization process, as well as to reduce the dependence on high-performing computational infrastructure. Despite the fact that no time window constraints are introduced in the optimization process, the output result already includes all of the estimation of travelling duration under real-time traffic condition, hence facilitating a confident measurement of appropriate departure and arrival point of time.

Additionally, the cluster-first-route-second ideology, which has been proposed by many researches in optimizing multi-depot VRP [25-26], can be used to cope up with the convenience stores' distribution problem. Because in real-life practice, prior to optimizing the product delivery routing of a certain convenience stores network, a process of clustering all of the stores located within a region (city or prefecture) into small distribution networks in a way that can minimize comprehensive travelling duration are necessary. Even though the paper concentrates more on the later one in the workflow, which is optimizing a certain convenience store network, a proposed solution using network modeling and Evolution algorithm on Solver add-in of Microsoft Excel [27-28] is also demonstrated to optimize the convenience stores clustering in the first procedure.

III. Methodology

This section examines the distribution problem of convenience stores as a specific case of vehicle routing problems with the objective of finding an optimal delivery route that minimizes gas consumption by reducing the travelling time of trucks distributing goods from the warehouse to several convenience stores, passing by a gas station to refuel and coming back to the warehouse. Different from method approaching the original problem of the Travelling Salesmen Problem, the approach proposed here focuses more on the actual travel time, rather than physical distance optimization.

A. Network Flow Model

1) Graphical Model of the Problem

The convenience stores' distribution optimization problem is examined as a "Network Flow Problem" by assuming each of physical locations in the distributing routes as a node in the network, distance and travelling duration as weights of arches. *The* network's optimization objective is therefore to find the flow connecting every node that associates with the optimal transportation time. A graphical network flow model of a simple convenience store distribution problem involving a warehouse, and 4 convenience stores can be represented as shown in Figures 1, 2 and 3.

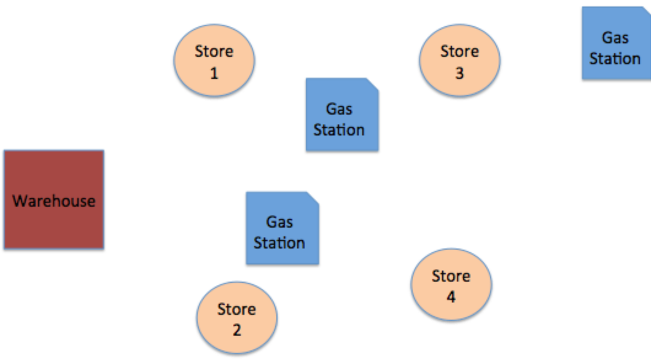

Fig. 1. Each location in the problem is demonstrated as a node in the network flow model.

Fig. 2. The arc connecting two nodes represent the physical distance, or the travelling time needed to move between the two respective locations.

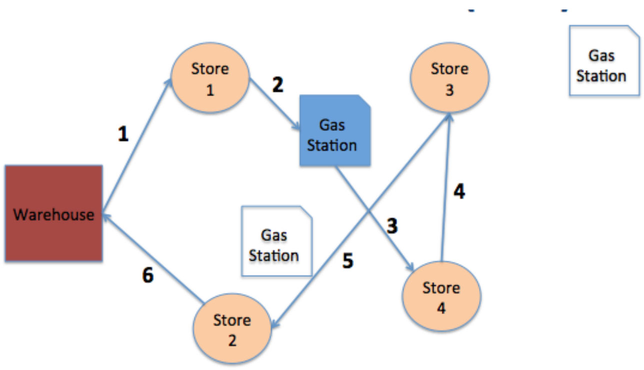

Fig. 3. Blue lines represent the flow connecting nodes and arcs, showing one of many possible routes that can help to achieve the distribution process.

2) Linear Programming Model of the Problem

A linear programming model of the stated convenience stores' distribution problem of a specific stores network can be represented as shown in Figure 4:

Notation for each node:
$a_1, a_2, a_3, .., a_n$

Arc weight between node:
a_i and a_j: $Weight(a_i, a_j)$

Whether exist a flow from node a_i to node a_j:
$Flow(a_i, a_j) = 1(exist) | Flow(a_i, a_j) = 0(non - exist)$

Minimize:
$z = \sum_{i=1}^{n} \sum_{j=1}^{n} Flow(a_i, a_j) Weight(a_i, a_j)$

Fig. 4. Mathematical Representation of Network Flow Model of Convenience Store Problem.

In order to solve this Network Flow optimization problem, collecting information about geographical locations of convenience stores, gas stations, warehouses and garbage dumpsites, as well as the distance and travel time between these locations is indispensable.

B. Data Collection with Web Scraping

To solve the distribution problem of convenience store, i.e., the Network Flow problem above, location data of all of the convenience stores, gas stations, as well as the physical distance and travelling time between those places are necessary. Thus where to find trustworthy sources and how to obtain the data are questions that are discussed in this section of the paper.

1) Data Source

NAVITIME website is established by the NAVITIME Japan Company, a Japanese online portal service focusing on events and places. NAVITIME website presents a friendly GUI (http://www.navitime.co.jp), enabling user not only access detailed information but also navigate routes between thousands of places classified into several categories such as convenience stores, gas stations, restaurants, parking places, schools, clinics and so on. Furthermore, the website also updates newly established spots in Japan, making it more reliable as a references source [29]. As far as data needed to solve the convenience store business problem, the NAVITIME website has lists of almost all of the convenience stores located in Japan, with details such as the store name, address, telephone number, and even the information regarding whether the store sells alcohol, cigarettes or fresh vegetables. Moreover, the website also classifies the convenience stores based on their location in a prefecture or city (e.g., Oita, Beppu) and brand names (e.g., 7-eleven, Lawson, etc.), making it very convenient to find the needed data. Hence, we have selected the NAVITIME website as the main data source for retrieving the geographical information of all of the convenience stores, gas stations in Japan.

2) Collecting Data from NAVITIME Website by Web Scraping

Web scraping or web crawling techniques are usually programmed to interact with a webpage source and extract data stored in it, and "usually, the extracted data might be post-processed, converted in the most structured format and stored for further usage" [30]. This research mainly relies on NAVITIME website for its data source, and web scraping programming technique as the tool to extract the data from the website. Despite the fact that NAVITIME website contains the necessary information for the convenience store distribution problem, the website is only built for normal users who would utilize web browsers to retrieve information on a single physical spot, convenience store or gas stations in this case. In other words, it would require huge amounts of efforts to manually surf the website and collect all of the data for nearly 100,000 locations of convenience stores and gas stations in Japan. Therefore, a systematic and automatic manner of retrieving the data from the website is indispensable, and web scraping technique is undeniably the most promising tool that can help us achieve this.

In order to achieve this, we developed a web-scraping algorithm for crawling data from the NAVITIME website in Ruby programming language and Figure 5 shows its pseudo code. The website categorizes all the convenience stores into different prefectures in Japan, and after accessing each of the link of those prefectures, the website shows all the links, each of which connects to a specific convenience store's information page, sorted into different listing pages. Therefore, a looping algorithm is used to access this hierarchical structure, which simultaneously extract all the needed information of convenience store such as the store name, geographical address, telephone number, and details regarding the availability of fresh vegetables, alcohol or cigarettes, etc., and store them into a local comma-separated values (CSV) file.

```
Web Scraping on NAVITIME Website;
Result: CSV Database File
repeat
    repeat
        repeat
            Access each store's link;
            Extract the store information;
            Append the extracted data into the result CSV file;
            Return to the previous list page;
        until The last store of the list;
        Access the next list page;
    until The last listing page;
    Access the next prefecture listing page
until The last prefecture listing page;
```

Fig. 5. Pseudo code of web-scraping algorithm for the NAVITIME website.

C. Multi-objective Route Optimization with Google API

1) Google API Service

Google API is a well-known web service that provides online tools for developers who utilize Google's data and platform. Google API consists of several programming libraries serving many areas such as Mapping, Advertising, Cloud storage, etc. Google Direction API is one such service that helps retrieve geographical distance and travelling time from one physical location to another. Moreover, the free version of such service also helps to calculate the optimal route in terms of time to travel through a maximum of 10 physical locations (including the origin and destination locations) [31]. Since Google Direction API operates on the Google server by receiving requests from the users and returning back the results, we can easily utilize the well-established platform of Google server.

Due to the academic nature of this research, we utilized the free version of the Google API service to solve the stated problem despite its limitations. For commercial application, however, the Work version of the service is recommended since in addition to a more generous number of request permits, it also provides the travelling duration and calculates optimal route under real-time traffic conditions, something which can further ensure optimality of the approach.

2) Optimize Convenience Stores Distribution Routing with Google API

In order to use Google API service, a structural request for fulfilling parameter requirements need to be sent to the Google server through a specific URL, which is designated by the service. Consequently, Google API will return the result in a response package. A typical request to Google Direction API to find an optimal travelling route for a trip includes:

- *Origin Point:* The origin location point from where the trip starts: e.g., the warehouse or distribution center.
- *Waypoints:* A list of all geographical points to pass-by during the trip: e.g., all the convenience stores in the network and a nearby gas station.
- *Destination point:* The final destination point where the trip ends: e.g., the warehouse or distribution center.

In order to find the optimal location of a gas station to fulfill the possible re-fueling requirement of the delivery vehicle, an iterative algorithm has been built. Within the collected data of all gas stations in Japan, we first want to find a list of all gas stations located within the same region as the distribution network of convenience stores. This task is obtained by looking up the area name of the distribution network (e.g. prefecture, city and town name) from the input postal code through a web API service (i.e., http://zipcloud.ibsnet.co.jp), and then the result is used as the keywords for searching relevant gas stations in the collected data. After that, each of the returned gas stations is inserted into the request waypoint parameter of Google API service, one by one, and from the returned routing results the optimal one is selected. This route then becomes the optimal one that passes through all the convenience stores in a desired network and at the same time drops by the gas station for re-fueling.

D. Cluster-first-route-second

Cluster-first-route-second is a heuristic method for optimizing large scale VRP, which has two separate phrases, namely clustering and routing. Originally, the algorithm was intended to address problems with vehicle capacity constraints. However, this idea of classifying all the waypoints into different clusters before optimizing the route has been adopted in many practical research questions such as in VRP with Backhauls [32].

In convenience stores business, daily-consumed products have to be delivered to every store on a regular basis. For example, 7-eleven convenience stores have rice and chilled products delivered three times per day, and drinks and frozen foods delivered three to seven times per week [2]. Hence, all of the customers or convenience stores in this case, have the same orders at similar time intervals during a day or a week. This implies that distribution network of a distribution manager who is in charge of hundreds of stores within a city or a prefecture is usually much larger than a chain of convenience stores that can be handled by a single delivery truck. Applying the idea of cluster-first-route-second approach will allow distribution managers to first cluster all of the stores that fall within their delivery network into many small networks in a way that can both satisfy total number of available delivery trucks and optimize the total travel time durations of their respective network. In this section we have utilized free Google Distance Matrix API to obtain the travel time matrix among convenience stores, a sample result of which are shown in Figure 6, and used Evolutionary Algorithm of Solver add-ins on Excel for solution demonstration. As an example, two networks consisting of all the 7-eleven and Family-Mart convenience stores in Beppu City, Oita Prefecture has been selected to illustrate the approach.

Address		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
別府駅	1	0	499	963	440	762	130	250	600	315	864	765	942	604	430	72
大分県別府市石垣東7丁目2055	2	503	0	740	207	374	443	563	207	462	470	633	767	382	286	59
大分県別府市大字鶴見字石田91番1	3	1024	705	0	847	595	967	998	609	1012	418	601	423	873	688	25
大分県別府市新港町4番15号	4	418	216	838	0	443	491	546	304	294	568	751	865	284	344	70
大分県別府市上人西2組1	5	717	312	598	462	0	722	845	226	593	328	791	625	301	551	56
大分県別府市田の湯町8番13号	6	166	494	941	515	762	0	120	595	390	859	743	920	679	425	70
大分県別府市元町21番18号	7	315	609	939	589	897	168	0	720	501	936	784	857	752	477	69
大分県別府市石垣西10丁目4-54	8	603	152	592	294	209	568	688	0	479	322	628	619	320	360	50
大分県別府市北浜3丁目818-166	9	303	478	1018	314	636	359	351	497	0	761	820	1026	478	466	77
大分県別府市大字鉄輪693番4	10	863	412	424	554	302	821	941	316	739	0	609	451	579	603	38
大分県別府市大字鶴見字五反4207-1	11	821	610	595	704	753	764	848	603	809	593	0	824	866	485	35
大分県別府市竹の内1895-15	12	913	693	411	835	583	860	842	597	927	405	820	0	860	732	46
大分県別府市上人ケ浜1412番6	13	544	322	842	320	303	617	672	308	420	572	878	869	0	551	75
大分県別府市石垣西3丁目5番8号	14	445	290	684	310	537	385	505	360	441	602	486	776	529	0	44
大分県別府市大 4組2	15	775	564	241	658	531	718	749	485	763	371	352	470	749	439	

Fig. 6. Travelling time duration matrix of network including a warehouse at Beppu station, and fourteen 7-eleven convenience stores in Beppu City.

The Evolutionary Algorithm with default parameters setting is run on linear programming models manually built in Excel. The optimization objective is set to minimize the total traveling duration in the whole network with a single warehouse. Decision variables show the total stores in each group, the classification of each convenience store, as well as the traveling order. Constraints of the problem are the constant number of available delivery trucks (or the total number of groups), and an assumption that each groups must contain more than two stores. Moreover, the warehouse is assumed to be located at the central station of the city.

For better illustration, visualization Google maps are generated by an online tool (https://mapbuildr.com) to show the clustering and routing result of the real-life networks of 7-eleven (14 stores shown on Figure 7) and Family-Mart (16 stores shown on Figure 8) convenience stores in Beppu City, Oita Prefecture. Moreover, each cluster is marked with distinguish colored dot on the map with the traveling numerical order within the cluster being placed inside, and the big balloon-shaped red mark represents the warehouse.

E. Web-based Implementation on Mobile Platform for Continuous Routing Updates

Although at the cluster-first-route-second stage, optimal delivery sequence of stores assigned to each delivery truck is calculated, the situation may change during the actual delivery time due to traffic congestion. The ultimate goal of this work is to also build a practical mobile application that can interact with truck drivers in a programmatic manner and enable them to input any references regarding a specific convenience store distribution network. The device should then automatically run the underlying procedures all the way from finding the necessary data to sending requests to Google API service and retrieving the calculated optimal delivery routes through an easy to follow GUI in a real time fashion. This way, the mobile device of the truck drivers can serve as navigation system for the delivery network, calculate optimal delivery route in a continuous fashion and update the truck drivers about a possible re-route option when deemed necessary. To achieve such an objective, a prototype of web-based mobile application and local web server have been established and developed using PHP, HTML5, JavaScript on Angular Framework, and MySQL database management and the following subsections highlights its main features.

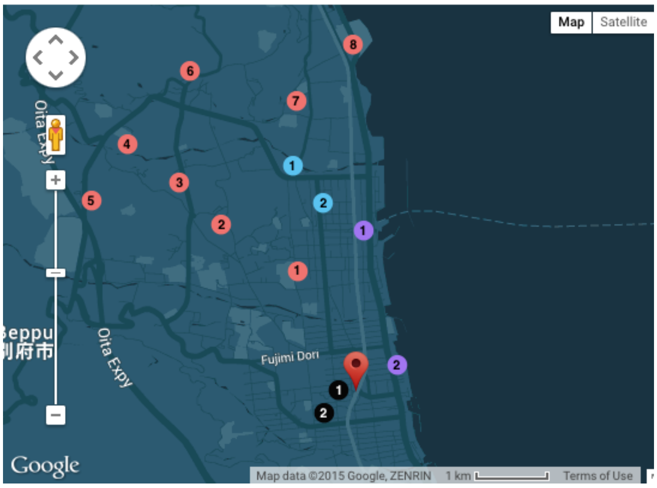

Fig. 7. Clustering with 4 groups (black, purple, blue, and red) and numerical routing order of network consisting of all of fourteen 7-eleven convenience stores located in Beppu City, Oita Prefecture.

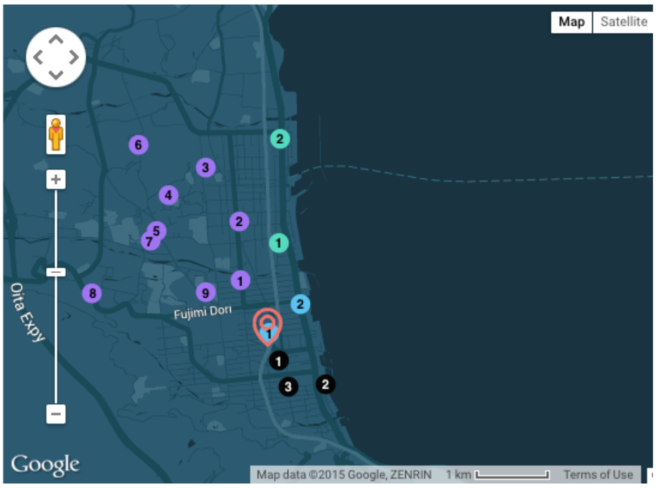

Fig. 8. Clustering with 4 groups (black, blue, green, and purple) and numerical routing order of network consisting of all of 16 Family-Mart convenience stores located in Beppu City, Oita Prefecture.

1) Navigation System Structure

A block diagram of the envisioned system is shown in Figure 9. It is a large online network consisting of a web

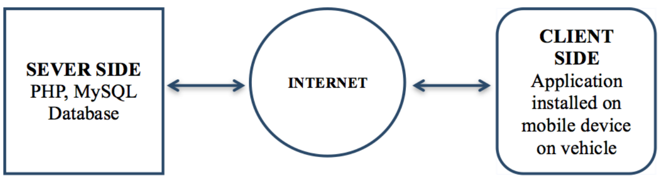

Fig. 9. Structure of the proposed navigation system.

server and various clients, which are mobile devices running the web-based application and available in each delivery truck. The web server's responsibilities include managing the retrieved database and sending geographical location of convenience stores and gas stations located in a requested distribution network, or geographical region. Because of the centralization of the server's database, the system can maintain consistency of response to every client in case there are modifications to any of convenience stores' information. This way, clients will act as bridge to search for convenience stores between the user and central sever. Furthermore, the mobile devices installed in the delivery vehicles will support the drivers by displaying an interactive navigation map, which is based on the familiar Google Map platform, illustrating its familiar perception of up-to-date optimal route for a customizable distribution network.

2) Graphical User Interface & Functions of Client Application

Mobiles are selected as target devices due to their built-in GPS functionality, which can facilitate real-time navigation tracking for the equipped delivery trucks. As such, the process of route optimization can be done not just statically before starting the trip, but also dynamically while the truck is moving. This way, the remaining route gets continuously optimized based on the current traffic condition.

The two figures 10 and 11 demonstrate the developed web-based application running on iOS simulator. Figure 10 shows the page requiring necessary input from the user for searching a list of convenience stores located within a specific region, which is specified by postal code. In this GUI, warehouse location, as well as the intention to re-fuel is also customizable. Figure 11 gives an example of the search result page, letting the user to select up-to 8 convenience stores (7 in the case of passing by a gas station for re-fueling) into their distribution network. Then, Figure 12 shows the result of a navigation map, guiding the distribution route with numbered markers (Gas station marker has "S" suffixed), representing

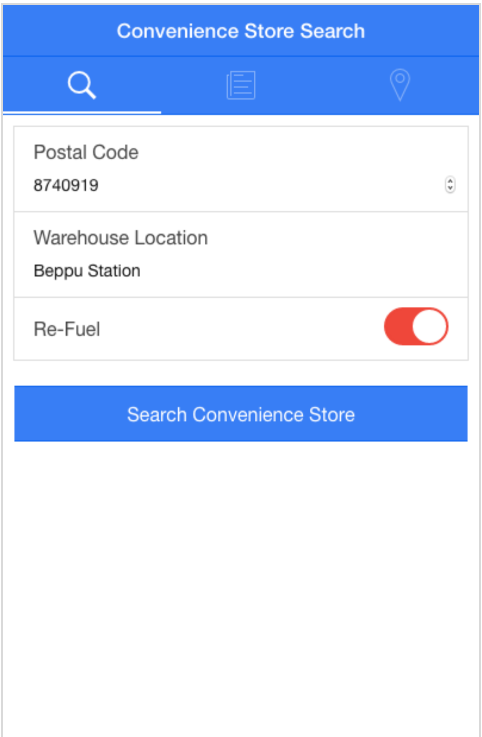

Fig. 10. GUI for showing search result of convenience stores located in an input area (through postal code).

Fig. 11. GUI for showing search result of convenience stores located in an input area (through postal code).

the travelling order, which starts from "0" at 1st delivery point.

As the delivery truck moves, the application automatically verifies whether it has reached a predefined place (e.g. Figure 11) on the route or not. This way, the application has the capability to continuously build a new network, which consists of the current location as new origin, and remaining stores and a gas station (if the truck has not yet passed by one) as the waypoints, without any manual input from the driver. Whenever the driver clicks the "Update" button shown on Figure 12, a new route is also immediately calculated together with the latest traffic condition from Google API service. As

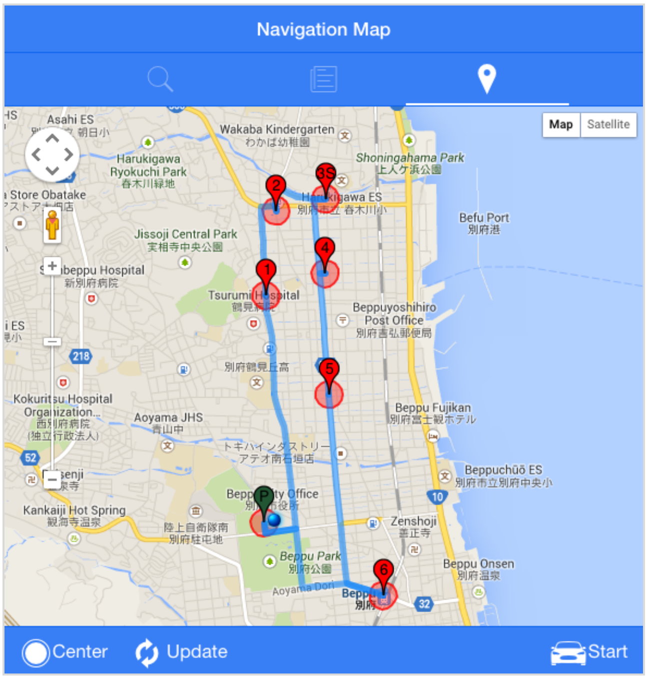

Fig. 12. GUI showing the navigation map with the optimal route calculated and shown. The blue dot represents the current location of the truck, and green marker labeled "P" indicates if the truck has already passed the location.

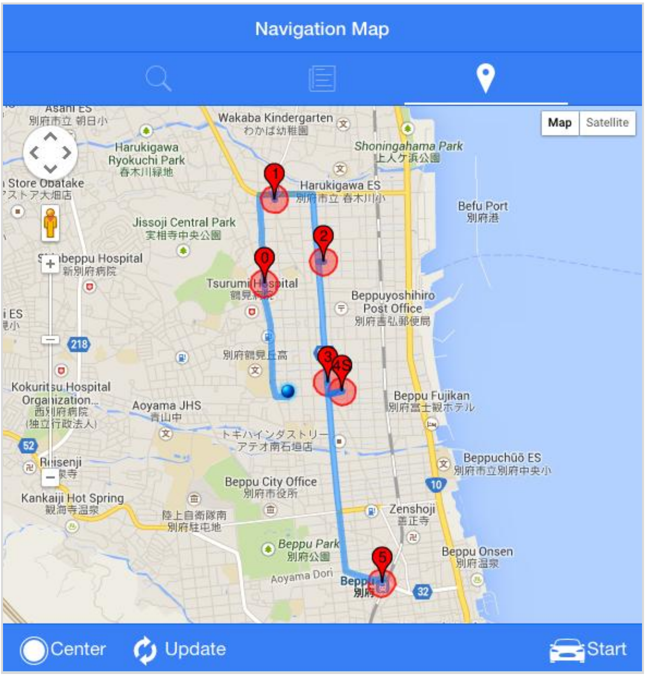

Fig. 13. GUI showing the updated navigation map for optimal route starting from the current location of the truck and passing by all the remaining convenience stores from the initial network.

shown in Figure 13, after the driver has passed the first location and is on the way to the next convenience store (location numbered 1 in Figure 12), a new route is calculated. Please note that in addition to the renumbering of the remaining delivery nodes, optimal gas station location has been altered from previous location labeled "3S" in Figure 12 to a new location labeled "4S" in Figure 13 due to current traffic situation.

IV. Analysis and Discussion

This section provides a thorough analysis of the obtained results and discusses the advantage of implemented methodology in some details. Initially, the envisioned positive impacts of the obtained result are examined from both economical and environmental perspectives. Then some figures illustrating the outcome of the data collection procedure is demonstrated. Several verification test results on the feasibility of the developed computer software are also illustrated. Finally, accuracy and performance of the proposed approach is compared with other well-known methodologies.

A. Commercial and Environmental Perspective

To examine the economical and environmental benefit of our approach, we collected some relevant data on convenience stores in Japan (e.g., the total number of convenience stores, the average number of deliveries per day, etc.) and made several assumption scenarios. For each of the scenario, financial cost saving and carbon dioxide emission prevention were calculated with the aid of online gas consumption calculator tool at http://www.city-data.com/gas/gas.php, which also considers the gas consumption differences between city roads and highways.

Collected Data and Pertinent Assumptions:

- Number of convenience stores in Japan: 56,925
- Average number of deliveries needed to a single store per day: 9
- Average number of convenience stores in a single distribution network, or average number of convenience stores to which a single truck has to deliver a day: 8
- Average hourly salary of a single truck driver: 900 Japanese Yen
- Average fuel cost per gallon: 115 Japanese Yen
- Miles per gallon in city: 30
- Miles per gallon in highway: 40

Based on these assumptions, an example network of eight 7-eleven stores in Beppu City, Oita Prefecture was initially tested with the program. Using the obtained result and by comparing two possible trips with a maximum and minimum travel distance, a heuristic estimation of maximum cost saving and carbon dioxide emission prevention for every convenience stores network in Japan is estimated. Our finding shows that approximately 2.8 billion Japanese Yen can be saved per year, which is equivalent to 56 million tons of carbon dioxide, or roughly 7.8 million homes' electricity for one year. Even though this estimate may not be precise due to location disparity of convenience stores throughout Japan, we can still see the huge impact of the approach in real life practice.

B. Data Collection and Program Testing Result

1) Data Collection

Informational on about 57,000 convenience stores and roughly 40,000 gas stations was collected using web data scraping technique on the NAVITIME website (http://navitime.co.jp). Table 1 shows a summary result of the specific number of physical places, total time needed for data extraction, and actual size of the CSV files in which the extracted data are stored:

TABLE I
Extraction Time of Convenience Stores and Gas Stations' Pertinent Data Using Ruby Script and Nokogiri Library

Data	Convenience Store	Gas Station
Number of Locations	56,925	36,660
Extraction Time	96 minutes	80 minutes
CSV File Size	16.3 MB	14 MB

Considering that our analytical approach depends on the extracted information on all the available convenience stores and gas stations in Japan (nearly 100,000 physical locations in total), the data crawling had to be done on nearly 100,000 web pages. This may give an impression that the above mentioned processing time for data extraction is quite expensive. Nonetheless, because all of the extracted information, e.g., store names, addresses, etc. is almost fixed and does not easily change; the created CSV file can be subsequently used as database for optimization for quite sometimes and enable the system to operate in real time fashion. Therefore, the extraction time of one and a half hours for a periodical update of the database can be considered practically acceptable. Furthermore, despite some limitations of the CSV file utilization for data storage which is discussed in the limitation section, the CSV database file size of about 15 MB is feasible for a typical ruby program to execute any necessary reading or writing operations.

2) Program Testing

In order to test correct operation and reliability of a newly developed program, we initially used it to gather various data on roughly 56,925 convenience stores and 39,660 gas stations that are located throughout Japan. We then focused our analytical modeling approach to convenience stores that are located in Oita prefecture. Our finding shows that out of 1,765 different districts in Oita prefecture that have a unique postal code, (a complete list of which can be accessed from the link http://homepage1.nifty.com/tabotabo/pzips/oita.htm), in 754 or 42% of them there exit convenience stores. We then created examples of convenient store networks with up-to 8 stores per network for supply and garbage collection services and in about 95% of them the program could come up with an optimal routing solution in less than 1 minute.

C. Accuracy & Performance

In general, routing optimization problem requires much time and computational resources in order to come up with a feasible solution. In this particular case, since optimization problem consisted of a network having a warehouse, up to 8 convenience stores, and a garbage dumpsite, our algorithm had to examine a total of 40,320 possible routes in determining an optimal solution that could yield the least travel duration or distance. Even though there are many proven algorithms such as Evolutionary or Generic algorithm to solve the routing optimization problem, their analysis in general takes longer time.

This was realized when we validated optimality of our result with that of Evolutionary algorithm built into Excel's Solver Add-in function using its default settings. Although both methods yielded the same result in terms of optimality, performance of our approach which relies on Google API is much faster than that of the Evolutionary algorithm built into Excel's Solver Add-in function, as shown in Table 2. Despite the fact that free version of Google API service is limited to a network of 10 nodes, its enterprise version can handle a network of up to 25 nodes per request.

TABLE II
PERFORMANCE COMPARISON BETWEEN GOOGLE API SERVICE AND EVOLUTIONARY ALGORITHM OF EXCEL SOLVER ADD-IN

Testing Network	Evolutionary Algorithm in Excel Solver Add-in (approximately)	Google API Service
Network of 6 locations	42 seconds	Less than 10 seconds
Network of 7 locations	51 seconds	Less than 10 seconds
Network of 8 locations	52 seconds	Less than 10 seconds
Network of 9 locations	53 seconds	Less than 10 seconds
Network of 10 locations	58 seconds	Less than 10 seconds

D. Cluster-first-route-second Performance

There have been many different ways approaching the cluster-first-route-second heuristic in solving VRP. However, many of them failed to consider the traveling time duration based on real-traffic condition at the time, something that is crucial for true optimization. Our test results shown in Figures 6 and 7 were achieved by Solver add-in in Excel within 1 minute, a performance that is a reasonable in terms of computation time for a network of less than 20 convenience stores within a city. With the usage of Google API for Work to access the distance matrix for real-traffic condition and some further customization of the Evolutionary Algorithm's parameters, this approach can also deal with a large scale network in bigger cities or in a whole prefecture that use multiple warehouses to cater to the needs of convenience stores.

E. Mobile Application Prototype

The developed prototype has been tested on modern browsers, including Chrome mobile on Android and Safari on iOS. Its performance has also been validated with several cases of user input (postal codes, warehouse location, and re-fueling necessity), and with all the cases of distribution network of convenience stores in Beppu City, the application has delivered optimal routing solutions. Moreover, thanks to the mobile device simulation function of desktop Chrome browser, the application has also been tested with different manual GPS locations to simulate continuous movement of vehicles. With the rapid development of silicon technologies, these days portable devices with large screen and high quality display are becoming more affordable. As such, this prototype is developed with the aim of integrating science and technology to come up with a practical solution to a real business problem. Such strategy will be used in our ongoing and future studies and it is hoped that other researchers will also pursue a similar trend.

V. CONCLUSION, LIMITATIONS AND FUTURE RESEARCH DIRECTIONS

This paper raises the awareness of utilizing available technology and programming techniques to attain commercial goal in convenience stores business, and simultaneously take the perspective of CSR into consideration. This research proposes practical methodologies for solving the convenience stores' distribution problem of optimizing product delivering route from a warehouse to several convenience stores, passing by a gas station for re-fueling, and returning back to the warehouse. The approach involves collecting geographical location dataset of about 100,000 real-life convenience stores and gas stations, a web-based mobile application for practical implementing of Google API service and creating database to solve the problem. Specially, a more general view of the convenience stores problem, which recognizes the empirical necessity of dividing a fix number of drivers to many small convenience stores distribution network, is introduced and resolved by using Evolutionary Algorithm of Solver on Excel. Overall, considering its reliability and fast performance, the approach can be extended for practical applications of not only convenience stores' business, but also in other businesses that are looking for advancements in supply chain management.

The followings can be considered limitations and some relevant future research directions.

1) Current web scraping technique strongly replies on the NAVITIME website structure. Specifically, the introduced web scarping technique uses a tree-based algorithm to extract data at pre-indicated tag position. Because of this, when the HTML tags structure of the webpage with the needed data changes, the technique may not work properly. To enhance this, development of an algorithm with artificial intelligence ability, such as the learning-based web wrapper algorithm, may be necessary so as to cope with possible structural changes to the NAVITIME website.
2) The analytical approach of this paper heavily relies on Google API service. Considering the academic nature of the work, however, the developed prototype was only tested with the free version of Google API service which

imposes certain limitation on the number of requests per day. Nonetheless, such limitation can be overcome by utilizing the Google API for Work, which facilitates up to 100,000 requests per day. Moreover, since the Google API for Work also supports calculation of optimal route with waypoints within real-time traffic condition, its usage will ensure a much more reliable result.

3) Although the recommended approach provides a practical solution for the convenience stores' distribution problem with the goal of optimizing the delivery time, which would result in reduction of gas consumption; the problem can also be examined under a pattern learning perspective to have an inner sight into the convenience stores' distribution process. Specifically speaking, apart from travelling time, there are also many other factors that contributes to the gas consumption during the product delivering process, namely; weather, temperature, route terrain, vehicle driving speed, etc. However, not all such factors can be easily controlled and calculated. A promising approach would be to collect data on such factors' impacts on the volume of gas consumption and apply machine-learning techniques so as to understand the patterns of how much each of such factors could contribute to the actual gas consumption. Then, an algorithm that also takes these factors into consideration can be developed. Moreover, because the data is expectedly retrieved in real-time, the learning pattern could be continuously updated to ensure the most pragmatic suggestion. To obtain this, the research promotes further examination into not only pattern learning techniques, but also ubiquitous and sensor network for real-time data collection.

REFERENCES

[1] Akinobu Terasaka, (1998), Development of new store types: the role of convenience stores in Japan, GeoJournal, Volumn 45, Issue 4, pp. 317-325.

[2] 7-Eleven. (n.d.). Convenience Store Operations. Retrieved from http://www.7andi.com/dbps_data/_template_/_user_/_SITE_/localhost/_res/en/ir/library/co/pdf/p24_33.pdf.

[3] Lawson In-Store Services. (n.d.). Retrieved February 18, 2015, from http://lawson.jp/en/product_service/service.

[4] METI. (2014, October 14). Monthly Report on the Current Survey of Commerce. Retrieved from http://www.meti.go.jp/english/statistics/tyo/syoudou_kakuho/index.html.

[5] USDA Foreign Agriculture Service. (2013, December 12). Japan Retail Food - Retail Sector 2012.

[6] Arai, Y. and Yamada, H. (1994) Development of Convenience Store System in Japan, 1970's-1980's. Terasaka, A. and Takahashi, S. eds., Comparative Study on Retail Trade: Tradition and Innovation. Ryutsu Keizai University.

[7] Broda Christian, Ephraim Leibtag, and David E. Weinstein. (2009). The Role of Prices in Measuring the Poor's Living Standards. Journal of Economic Perspectives, 23(2), pp. 77-97. Cordeau

[8] J-F Cordeau, M. Gendreau, G. Laporte, J-Y Potvin, F. Semet. (May 2002). A Guide to Vehicle Routing Heuristics, The Journal of the Operational Research Society, Vol. 53, No, pp. 512-522.

[9] G. Clarke and J. W. Wright. (1964). Scheduling of vehicles from a central depot to a number of delivery points, Operations Research, vol. 12, pp. 568–581.

[10] Back, T. (1996). Evolutionary algorithms in theory and practice. Oxford Univ. Press.

[11] Gendreau, M., Hertz, A., & Laporte, G. (1994). A tabu search heuristic for the vehicle routing problem. Management science, 40(10), pp. 1276-1290.

[12] Taillard, É., Badeau, P., Gendreau, M., Guertin, F., & Potvin, J. Y. (1997). A tabu search heuristic for the vehicle routing problem with soft time windows. Transportation science, 31(2), pp. 170-186.

[13] Dorigo, M., Maniezzo, V., & Colorni, A. (1996). Ant system: optimization by a colony of cooperating agents. Systems, Man, and Cybernetics, Part B: Cybernetics, IEEE Transactions on, 26(1), pp. 29-41.

[14] Dariusz Barbucha. (2014). A Hybrid Cooperative and Reactive Search Approach for Vehicle Routing Problem with Time Windows, Intelligent Information and Database Systems, pp. 207-216.

[15] Canhong Lin, K.L. Choy, G.T.S. Ho, H.Y. Lam, Grantham K.H. Pang, K.S. Chin. (November 2014). A decision support system for optimizing dynamic courier routing operations, Expert Systems with Applications, Volume 41, Issue 15, pp. 6917-6933, ISSN 0957-4174.

[16] Jiafu Tang, Yang Yu, Jia Li. (January 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: Logistics and Transportation Review, Volume 73, pp. 114-132, ISSN 1366-5545.

[17] Kunpeng Li, Bin Chen, Appa Iyer Sivakumar, Yong Wu. (October 2014). An inventory–routing problem with the objective of travel time minimization, European Journal of Operational Research, Volume 236, Issue 3, pp. 936-945, ISSN 0377-2217.

[18] Mes, M., Schutten, M., & Rivera, A. P. (2014). Inventory routing for dynamic waste collection. Waste management, 34(9), pp. 1564-1576.

[19] Zarandi, M. H. F., Hemmati, A., Davari, S., & Turksen, I. B. (2013). Capacitated location-routing problem with time windows under uncertainty. Knowledge-Based Systems, 37, pp. 480-489.

[20] Duygu Taş, Ola F, Tom Van Woensel. (December 2014). A Vehicle Routing Problem with Flexible Time Windows, Computers & Operations Research, Volume 52, Part A, pp. 39-54, ISSN 0305-0548.

[21] A.L. Kok, E.W. Hans, J.M.J. Schutten. (May 2011). Optimizing departure times in vehicle routes, European Journal of Operational Research, Volume 210, Issue 3, pp. 579-587, ISSN 0377-2217.

[22] Narath Bhusiri, Ali Gul Gureshi, Eiichi Taniguchi. (2013). Vehicle Routing and Scheduling Problems for Convenience Store Industry Considering Soft Time Windows, Proceedings of the Eastern Asia Society for Transportation Studies, Volume 9.

[23] M.M. Solomon. (1987). Algorithms for the vehicle routing and scheduling problems with time window constraints, Oper. Res., vol. 35, pp. 254-265.

[24] András Király, János Abonyi. (February 2015). Redesign of the supply of mobile mechanics based on a novel genetic optimization algorithm using Google Maps API, Engineering Applications of Artificial Intelligence, Volume 38, pp. 122-130, ISSN 0952-1976.

[25] Rodolfo Dondo, Jaime Cerdá. (February 2007). A cluster-based optimization approach for the multi-depot heterogeneous fleet vehicle routing problem with time windows, European Journal of Operational Research, Volume 176, Issue 3, pp. 1478-1507, ISSN 0377-2217.

[26] Ren Chunyu; Wang Xiaobo. (2006). Research on VRP Optimizing Based on Hierarchy clustering and IGA under Common Distribution, Computational Intelligence and Security, 2006 International Conference on, vol.1, pp. 143-146.

[27] VBA for Modelers: Developing Decision Support Systems With Microsoft Office Excel S. Christian Albright (South-Western Pub 2011/08/26).

[28] Frontline Systems Inc., Solver URL: http://www.solver.com.

[29] NAVITIME. (n.d.). Retrieved February 19, 2015, from http://www.navitime.co.jp.

[30] Marcos D. Assunção, Rodrigo N. Calheiros, Silvia Bianchi, Marco A.S. Netto, Rajkumar Buyya. (August 2014). Big Data computing and clouds: Trends and future directions, Journal of Parallel and Distributed Computing, ISSN 0743-7315.

[31] Google Developers. (n.d.). The Google Directions API - Google Maps API Web Services. Retrieved from https://developers.google.com/maps/documentation/directions/.
Gulcin Dinc Yalcin and Nihal Erginel. (July 2012). A Heuristic Based on Integer Programming for the Vehicle Routing Problem with Backhauls. Proceedings of the 2012 International Conference on Industrial Engineering and Operations Management.

Thai Quang LE was born in Khanh Hoa Province, Vietnam in 1993. He is presently pursuing his studies at Ritsumeikan Asia Pacific University (APU) in the field of Business Administration focusing on Strategic Management and candidate for graduation. He has received numerous scholarships including, APU Tuition Reduction Scholarship from 2011-2015, JASSO (Japan Student Services Organization) Scholarship from 2011-2012, and Oita Prefecture Scholarship in 2012-2013, 2014-2015. His research interests include applied programming in business, modeling for decision-making and machine learning, on which he has carried out a few presentations. The theme of his thesis is retail activity optimization and he has become an IEEE member since 2014.

Davar Pishva is a professor in ICT at the College of Asia Pacific Studies, Ritsumeikan Asia Pacific University (APU) Japan. In teaching, he has been focusing on information security, technology management, VBA for modelers, structured decision making and carries out his lectures in an applied manner. In research, his current interests include biometrics; e-learning, environmentally sound and ICT enhanced technologies. Dr. Pishva received his PhD degree in System Engineering from Mie University, Japan. He is Secretary General of IAAPS (International Association for Asia Pacific Studies), Senior Member of IEEE, and a member of IEICE (Institute of Electronics Information & Communication Engineers), IAAPS and University & College Management Association.