<table>
<thead>
<tr>
<th><strong>Title</strong></th>
<th>A Data Mining and Optimization-based Real-time Mobile Intelligent Routing System for City Logistics</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Author(s)</strong></td>
<td>Lin, C; Choy, KL; Pang, GKH; Ng, MTW</td>
</tr>
<tr>
<td><strong>Citation</strong></td>
<td>The IEEE 8th International Conference on Industrial and Information Systems (ICIIS), Peradeniya, USA, 17-20 December 2013. In IEEE International Conference on Industrial and Information Systems Proceedings, 2013, p. 156-161, article no. 6731973</td>
</tr>
<tr>
<td><strong>Issued Date</strong></td>
<td>2013</td>
</tr>
<tr>
<td><strong>URL</strong></td>
<td><a href="http://hdl.handle.net/10722/203996">http://hdl.handle.net/10722/203996</a></td>
</tr>
<tr>
<td><strong>Rights</strong></td>
<td>International Conference on Industrial and Information Systems. Copyright © I E E E.; ©2013 IEEE. Personal use of this material is permitted. However, permission to reprint/republish this material for advertising or promotional purposes or for creating new collective works for resale or redistribution to servers or lists, or to reuse any copyrighted component of this work in other works must be obtained from the IEEE.; This work is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License.</td>
</tr>
</tbody>
</table>
A Data Mining and Optimization-based Real-time Mobile Intelligent Routing System for City Logistics

Canhong Lin, King-lun Choy, Grantham Pang, Michelle T. W. Ng

Abstract—City logistics is facing the challenging problem of providing a quick-response and on-time delivery service in congested urban areas with frequent traffic jams. The dynamically changing traffic conditions make the predetermined best transportation plans suboptimal and consequently cause increased logistics cost and even greater air pollution. To help the driver determine time-optimal routing solutions in order to avoid congestion according to the real-time traffic flow, a Real-time Mobile Intelligent Routing System is designed and deployed on drivers’ Smartphones to help in routing decision making. Data mining techniques are employed to discover the routing patterns from the past cases of routing plans so as to generate case-based routing plans for the drivers. A metaheuristic is used to undertake the optimization of a real-time optimal routing plan based on real-time traffic information. A case study and computational experiments demonstrate the effectiveness of the proposed methods in significantly reducing the traveling time.

Index Terms—Data mining, Intelligent Transportation System, optimization, real-time vehicle routing, Variable Neighborhood Search.

I. INTRODUCTION

The striking development of E-Business in recent years has increased the importance and burden of transportation and logistics in urban areas. With more demanding and time-sensitive customer service requirements as well as the subsequent competition from other logistics companies, the practitioners are facing a more challenging situation of conducting a quick-response and on-time delivery service. Additionally, in the presence of congested urban areas resulting from frequent traffic jams, deliveries are often delayed and thus significantly degrade the level of customer satisfaction. The traditional transportation management system fails to reschedule the dispatching plan in a dynamic traffic environment [1]. The complex city transportation network also seems to frustrate the drivers when they are seeking the best transport plan. Fig. 1 depicts the major two vehicle dispatching problems that logistics companies encounter in their daily operations. The first problem concerns the pre-designed routing plan before performing delivery. This is often based on the driver’s past experience and aims at a shortest path instead of on the shortest traveling time. The second problem is that even though an optimal plan in terms of minimum traveling time is established, it cannot remain optimum as the traffic conditions vary as time goes by. An alternative time-optimal route for vehicles to go on instead of staying on the present road which has bad traffic conditions is difficult to find because traffic information about other roads is invisible to the driver.

Fortunately, based on the large amounts of transportation data generated from the daily operations of distributing cargos, some routing patterns can be discovered through data mining techniques. Fig. 2 illustrates an example of transportation

![Fig. 1. Problems that frequently occur in city logistics](image1)

![Fig. 2. Routing patterns](image2)
segments at around 9:00am on three days. The colors of the roads on the map represent the average traveling speed on the roads (e.g. dark red means very slow traffic flow while light green means fast flow). Because a traffic jam (red path between area A and area C) frequently occurs on the paths from area A to area C at around 9:00am almost every day, the drivers would serve the customers in area B first when they depart from the customer points in area A. Such “common sense” and implicit “area” exists in the experienced drivers’ mind and can be discovered from past routing plans if they are recorded in a database. In other words, the spatial and temporal routing patterns can be discovered from the vast amount of data about past routing plans. These patterns can be transformed into transportation segments with spatial and temporal attributes for generating a case-based feasible routing plan.

The problems discussed above call for a decision support system to provide intelligent transportation solutions. To free the decision maker from the arbitrary and complicated routing decision making process, the intelligent system is able to offer case-based feasible routing plans that avoid congestion by connecting a series of transportation segments derived from the extracted routing pattern. Advances in new telecommunication and mobile technologies such as global positioning systems (GPS), geographic information systems (GIS), traffic flow sensors, and Smartphones, make it possible to use real-time traffic information to improve service level, enhance the economy and energy efficiency of logistics [2, 3]. Reflecting the real-time traffic conditions throughout the city, the system can therefore constantly update the time-optimal routing plan during transportation so the goods can be delivered to the customers as soon as possible.

The main objective of this paper is to develop a Smartphone-based Real-time Mobile Intelligent Routing System (RMIRS) for the drivers to make a case-based feasible routing plan to avoid congestion and to dynamically reschedule their routing plans, the aim of which is to reduce the vehicle traveling time and improve the service level. Data mining approaches are employed to derive case-based feasible routing plans. A metaheuristic is used to optimize the real-time routing plans based on the continually received real-time data of the urban traffic flow. The rest of the paper is organized as follows: In Section II, a literature survey on Real-time Routing Problems in addition to the application of data mining techniques in transportation management is presented. Section III describes the architecture of RMIRS and the key modules and engines. Section IV contains an application case study of the implementation of RMIRS and the results of the computational experiment of the case study are presented and discussed. Finally we give concluding remarks in Section V.

II. LITERATURE REVIEW

Real-time Routing Problems arise in real-life scenarios where some information about the components of a transportation network, such as traveling time, customer orders, is continually released or updated over time during the planning period [2]. Time-varying vehicle speeds, due to dynamically changing traffic conditions, are a feature of Real-time Routing Problems where the aim is to minimize the traveling time. This results in different optimal routing plans at different times. In the literature, the most related research work is with regard to the time-dependent vehicle routing problem [4]. In practice, this problem provides a flexible approach to help find feasible routing plans to better utilize the urban transportation infrastructure or avoid congestion, with the implication of reducing logistics cost and overtime expense [5], enhancing the customer service level, and even having a green impact on society [6, 7]. A variety of algorithms have been proposed to deal with the time-dependent vehicle routing problems. In this study, we employ the Variable Neighborhood Search (VNS) to work out such Real-time Routing Problems. VNS is a local search based metaheuristic that was first developed by N. Mladenović, & P. Hansen [8]. Through the systematic change of the neighborhood during the local searching process, VNS stands out in its performance in solving some combinatorial optimization problems. To the best of our knowledge, existing studies that use VNS to tackle Real-time Routing Problems are not found. Hence, it is an interesting research area to explore the performance of VNS in tackling real-time routing problems.

Applying data mining and machine learning techniques in transportation management has engaged researchers’ attention. This is motivated by the fact that a large amount of data is derived from daily operations of traffic and transportation and such data can be extracted and analyzed as potentially useful and insightful information for reaching decisions [9]. This field of research remains at the beginning and related research effort is meager, especially for the study of Real-time Routing Problems. Recently, machine learning techniques like Support Vector Regression (SVR) and data mining methods like clustering are deployed in traffic speed prediction [10], map-matching [11], pedestrian detection [12] for Intelligent Transportation System and retail store clustering for logistics distribution networks [13]. In this study, we attempt to discover the routing patterns at different times of the day from large amounts of past routing information by means of the k-Means clustering algorithm, and make use of the extracted patterns for decision making.

Table I shows a survey on similar routing systems in the literature, including whether they leverage optimization or data mining techniques, GPS, GIS or mobile data capturing techniques or not. It seems that the majority of the current methods for solving Real-time Routing Problems depend largely on optimization-based approach to reach an optimal solution. H.G. Santos et al. have provided a good example of incorporating optimization algorithms with data mining techniques for discovering potential routing patterns from previous best solutions [20]. Compared with these existing
TABLE I

THE COMPARISON BETWEEN THE EXISTING ROUTING SYSTEMS

<table>
<thead>
<tr>
<th>Existing routing systems</th>
<th>Methods used for routing vehicles</th>
<th>Real-time/mobile techniques</th>
</tr>
</thead>
<tbody>
<tr>
<td>R. Ruiz et al. [14]</td>
<td>Optimization approaches: two-stage exact methods (stage 1: implicit enumeration algorithm; stage 2: integer programming models).</td>
<td>None</td>
</tr>
<tr>
<td>L. Santos et al. [17]</td>
<td>Optimization approaches: an improved path-scanning heuristic and an ant-colony metaheuristic.</td>
<td>None</td>
</tr>
<tr>
<td>J.E. Mendoza et al. [18]</td>
<td>Optimization approaches: a modified Clark and Wright savings heuristic and two memetic algorithms.</td>
<td>GIS</td>
</tr>
<tr>
<td>L. Santos et al. [19]</td>
<td>Optimization approaches: a modified path-scanning heuristic.</td>
<td>GIS</td>
</tr>
<tr>
<td>H.G Santos et al. [20]</td>
<td>Data Mining and optimization approaches: Genetic Algorithm with using Apriori like algorithms to discover patterns (subtours) commonly found in the best solutions of the population.</td>
<td>None</td>
</tr>
</tbody>
</table>

systems, our proposed system has much more utilization of both optimization and data mining techniques, as well as integrating GIS and smartphone mobile techniques. The idea of applying data mining approach in this study is similar to [20], for mining sequential routing patterns. In addition, a k-Means clustering algorithm is adopted here as an independent module of the system, which is different from [20]. Furthermore, temporal and spatial patterns are specially considered in this study in the context of city logistics. The mobile strategy is featured by using smartphones to undertake the following four tasks: 1) capturing real-time data on traffic condition via GPS or GIS; 2) optimizing transportation routes after receiving real-time traffic information; 3) presenting generated optimal routes for drivers dynamically; 4) passing optimal routing solutions back to the server end in order to store essential data into the database.

III. REAL-TIME MOBILE INTELLIGENT ROUTING SYSTEM (RMIRS)

The architecture of the proposed RMIRS is shown in Fig. 3. It encompasses two major components: the Intranet Server End that is deployed in the company Intranet for data collection, mining and generating case-based feasible routing plans based on past cases of routing plans, and the Mobile Client End which is installed on the driver’s Smartphone to update the real-time optimal routing plans. The communication between these two components is supported by a wireless network infrastructure.

**Intranet Server End.** It consists of a data warehouse and a Case-based Routing Module which includes a Data Mining Engine and Case-based Routes Generation Engine. The above-discussed routing patterns are produced by using the k-Means algorithm to cluster the records in Data Mining Engine and they are stored as transportation segments in the cache database. The Case-based Routes Generation Engine retrieves the transportation segments and determines a case-based feasible routing plan based on the location of the customers and the outbound time of the vehicle. The algorithm of this process is shown in Fig. 4 and Algorithm 1.

![Fig. 3. Architecture of RMIRS](image-url)
**Mobile Client End.** It includes a Real-time Data Receiver for collecting dynamic data of the city traffic flow, a Dynamic Routing Module embedded with a Real-time Routing Optimization Engine for dynamically optimizing the routing plan, and a Route Displayer for showing visual routes for the driver. VNS is used to find an updated time-optimal routing plan when the Real-time Routing Optimization Engine is launched. The new route derived is also sent back to the Data Warehouse as a successful routing case. This component is installed on the driver’s Smartphone as portable software to facilitate the real-time routing decision making process. Fig. 5 and Algorithm 2 shows the operation flow of VNS.

**Algorithm 1: Case-based Routing**

```plaintext
Function Case-based_Routing()
{Clusters} C = k-Means(PastCases P)
{TransportationSegments} T = Transform()
//given a customer set and outbound time
{G(customer)} G = Classify(Customers U)
{InClusters} I = Filter(G)
Clusternext = NULL
Cluster c = depot
Repeat
NextCluster = Retrieve(T, G, c, t)
I.remove(Clusternext)
R.add(Clusternext)
c = Clusternext
T = Time(Clusternext)
Until I is empty
Route r = expand(R, G)
Return r
```

1. Use k-Means algorithms to cluster customers (1)
2. Transform the past routes to transportation segments (subtours) with temporal and spatial attributes (e.g., 8:20 Cluster 1 to Cluster 5) (2)
3. With the customer list to serve, classify the customers to the clusters obtained in (1). The output is denoted as GROUPS. (3)
4. Set the departure time, and the starting point as the depot.
5. Based on the transportation segments obtained in (2), the time, and the current cluster, find the next cluster, add it to the cluster sequence.
6. Remove the current cluster from the GROUPS
7. Set the found next cluster as the current cluster, update the time.
8. Expand the cluster sequence to routes

**Algorithm 2: VNS**

```plaintext
Function VNS()
Solution S_{best} = Case-based_Routing()
NeighborhoodStructure (N_{i}, i=1,2,...,k)
Repeat
i = 1
Repeat
//generate a random solution S_{ran} in N_{i}(S_{best})
S_{ran} = Shaking(S_{best})
//S_{opt} is the obtained optimum
S_{opt} = LocalSearch(S_{ran})
If (S_{opt} is better than S_{best}) then
S_{best} = S_{opt}
i = i + 1
Else
i = i + 1
Until i > k
Until stop-condition met
```

1. Generate the initial solution from the case-based routing algorithm, denote it as S_{best}; Select a stopping criterion; Define the set of neighborhood structure N_{i} (i=1,2,...,k)
2. Shaking: generate a random solution S_{ran} in the i_{th} neighborhood of S_{best}
3. Local search: find the optimal solution S_{opt} near S_{ran}
4. Is S_{opt} better than S_{best}?
5. Is stopping criterion met?

**IV. APPLICATION CASE STUDY**

**A. Industrial background**

ABC is a Third-party logistics (3PL) company in Shenzhen, China. The daily logistics operation is to distribute requested goods to the customers scattered at different places in the urban area. In this case, a single vehicle has to serve 46 customers and finally returns to the depot. RMIRS is employed in this company with the purpose of improving its transportation efficiency and quick-response delivery service level. Fig. 6 depicts the steps of implementing RMIRS in this case study.
B. Results and discussion

The main objective of the computational experiment is to evaluate the feasibility, effectiveness and performance of the proposed algorithm and system. The algorithm and key engines are coded in Java, and run on a personal computer equipped with an Intel Core i5 1.8GHz processor with 4GB RAM.

Step 1: Clustering past data of customers and routes. Fig. 7 and Table II show the result of clustering the 80 customers of the existing records in the database, based on their geographical location. The starting and ending places of each record are then transferred to the corresponding clusters.

Step 2: Generating case-based routes. Given the data of the 46 customers, these customers are classified into a certain cluster shown in Fig. 7. The 46 customers are classified into these clusters in Table II: 0, 1, 2, 3, 5, 7, 10, 12, 15, 16, 17, 18, 19, 21, 23, 24, 25, 26, 28. When the departure time of the vehicle is defined, the most similar transportation segments in the database are retrieved, in terms of starting cluster and time, to constitute a feasible case-based route for the driver. Fig. 7 shows an example of connecting 4 transportation segments, in which the clusters include some of the 46 customers. Table III gives the derived customer visiting sequence in the form of clusters and customers of 2 scenarios where the vehicle leaves the depot at different times.

Step 3: Receiving real-time traffic information. The real-time information of traffic conditions is gained by using mobile technologies such as GPS, traffic flow sensors, etc. Such information is conveyed to the Smartphone in order to determine a real-time optimal routing plan in Step 4.

Step 4: Dynamically optimizing incumbent route. After receiving the real-time traffic information, the Real-time Routing Engine is able to employ VNS to seek another optimal route. Table IV shows the performance of VNS in solving the Real-time Routing Problem. It is noteworthy that the CPU time is short enough such that the decision making process can be swiftly achieved during the transportation, which is very important in a dynamic environment.

In this case, the driver updates the routing plan every time he completes serving a customer. The driver obtains the new route at the place of the incumbent customer and travels to the next customer. Table V presents all the yielded decisions which the clusters include some of the 46 customers. Table III gives the derived customer visiting sequence in the form of clusters and customers of 2 scenarios where the vehicle leaves the depot at different times.

Step 5: Displaying route and storing optimized routes. As shown in Fig. 8, the derived optimal route is displayed on Google Maps. The driver can even adjust the optimal route according their experience. Meanwhile, the information of this route is transmitted back to the data warehouse.

---

Fig. 6. Implementation flow

---

**TABLE II**

<table>
<thead>
<tr>
<th>Cluster</th>
<th>Center of 30 Clusters</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>22.717955</td>
</tr>
<tr>
<td>2</td>
<td>22.7090244</td>
</tr>
<tr>
<td>3</td>
<td>22.5265369</td>
</tr>
<tr>
<td>...</td>
<td>...</td>
</tr>
<tr>
<td>27</td>
<td>22.6811689</td>
</tr>
<tr>
<td>28</td>
<td>22.6873086</td>
</tr>
<tr>
<td>29</td>
<td>22.5475502</td>
</tr>
<tr>
<td>30</td>
<td>22.6522884</td>
</tr>
</tbody>
</table>

**TABLE III**

<table>
<thead>
<tr>
<th>Time</th>
<th>Clusters</th>
<th>Route</th>
</tr>
</thead>
<tbody>
<tr>
<td>8:20</td>
<td>23-18-17-10-12-15-24-3-21-5-1-7-26-3</td>
<td>0-33-42-26-28-36-45-6-8-5-37-32-39-3-1-4</td>
</tr>
<tr>
<td></td>
<td>0-2-28-19-25-16</td>
<td>-30-35-46-38-10-17-25-29-44-7-9-11-12-14</td>
</tr>
<tr>
<td></td>
<td>15-20-43-0</td>
<td>-15-20-43-0</td>
</tr>
<tr>
<td>16:30</td>
<td>18-21-7-5-25-2-1-1</td>
<td>0-26-28-36-1-4-2-13-16-18-21-23-19-22-24-34</td>
</tr>
<tr>
<td></td>
<td>3-19-17-24-16-3</td>
<td>3-42-10-17-25-29-44-45-32-39-9-11-12-14-</td>
</tr>
<tr>
<td></td>
<td>15-20-43-3-0</td>
<td>15-20-43-3-0</td>
</tr>
</tbody>
</table>

---

**TABLE IV**

<table>
<thead>
<tr>
<th>Iteration</th>
<th>Best result(min)</th>
<th>Average result(min)</th>
<th>CPU(s)</th>
</tr>
</thead>
<tbody>
<tr>
<td>6000</td>
<td>398.85</td>
<td>428.98</td>
<td>8</td>
</tr>
<tr>
<td>8000</td>
<td>391.97</td>
<td>424.96</td>
<td>10.99</td>
</tr>
<tr>
<td>10000</td>
<td>390.42</td>
<td>421.37</td>
<td>13.76</td>
</tr>
</tbody>
</table>
TABLE V  
REAL-TIME OPTIMAL TRAVELING TIME

<table>
<thead>
<tr>
<th>t</th>
<th>Pre-route(min)</th>
<th>Opt-route(min)</th>
<th>Improvement (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>472</td>
<td>168</td>
<td>64.41</td>
</tr>
<tr>
<td>1</td>
<td>1037</td>
<td>194</td>
<td>81.29</td>
</tr>
<tr>
<td>2</td>
<td>1155</td>
<td>251</td>
<td>78.27</td>
</tr>
<tr>
<td>3</td>
<td>1139</td>
<td>186</td>
<td>83.67</td>
</tr>
<tr>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
</tr>
<tr>
<td>41</td>
<td>395</td>
<td>293</td>
<td>25.82</td>
</tr>
<tr>
<td>42</td>
<td>289</td>
<td>289</td>
<td>0.00</td>
</tr>
<tr>
<td>43</td>
<td>293</td>
<td>293</td>
<td>0.00</td>
</tr>
</tbody>
</table>

Fig. 8. Google Maps-based route displayer

V. CONCLUSION AND FUTURE WORKS

In this study we develop a prototype of a Real-time Mobile Intelligent Routing System in order to assist in making case-based feasible routing plans and solving Real-time Routing Problems for the drivers who drive in congested cities. Our aim is to help them avoid congestion and reduce their traveling time. The case study and experiment results show the feasibility and effectiveness of the proposed clustering and optimization algorithms. The proposed Smartphone-based intelligent transportation system would be useful for the drivers in their daily operations. More importantly, it improves the service level and has a green impact by reducing delayed delivery and prevents delivery vehicles being trapped in traffic jams. It is expected that this system will enable researchers to consider more operational constraints on Real-time Routing Problems in a more practical way, involving such aspects as service time windows, multiple and heterogeneous vehicles, etc.

Although the use of k-Means and VNS algorithms for tackling the issues of city logistics in this study is effective and promising, it remains to be explored if the proposed method significantly outperforms other approaches. Future study is proposed to include more quantitative evaluation of the operational performance of RMIIRS, and more comparison with other existing methods in terms of the solution quality.

ACKNOWLEDGMENT

The authors wish to thank the Research Office of The Hong Kong Polytechnic University for supporting the project (Project Code: RT6B).

REFERENCES


