Abstract—Route maintenance through optimal path is the most important feature for efficient packet transfer with minimal computational complexity (i.e., vehicular communication). The packet to be transferred from source end to the destination through intermediate nodes causes congestion using the present single modality system, leading to load imbalance. Moreover, the congestion also occurs on the infeasible route path (i.e., longest route path selection) with temporal data constraints. Recommended road routes for vehicle transfer takes longer hours (i.e., travelling through many intermediate nodes) and fails in maintaining the time and budget constraints. To achieve a single modal road route path with routing efficiency, Transportable Road Route Path Miner (TRRPM) method is proposed in this paper. Initially, the user packet (i.e., vehicle) profiles information is extracted continuously using Secluded Supervision to construct the route path. Next, the profile information about vehicles is grouped together using the cosine similarity based on the frequent movement of different vehicle patterns. The moving temporal data patterns (i.e., vehicle) from the source to destination end with intermediate node points are grouped (i.e., cluster) together based on time factor. Finally, the clustered temporal data patterns are tuned using the Shortest Route Path based K-fold Cross Validation Tuning Method. Temporal data which are taken as sample item set is used to optimize the result on constructed road network route path for vehicular communication. TRRPM method provides the vehicular communication platform with analytical tasks of balanced time and budget factor. Experiment is conducted on factors such as total tuning time, routing efficiency on single modal system, and computational cost factor.

Keywords: Temporal Data Pattern, Shortest Route Path, Single Modality System, K-fold Cross Validation Tuning, Vehicular Communication, Cosine Similarity.

1 INTRODUCTION

Route maintenance is the most critical component for the control and engineering of vehicle transfer in single modal systems. The traffic measurement in order to follow efficient route should make it possible to obtain the spatial flow of traffic. Many research works have been focused on maintaining the route, though with certain flaws. Optimal Route Queries with Forward and Backward Search (ORQ-FBS) [1] evaluated optimal route from the last point to the first and first to last point respectively with the objective of providing optimal routing. However, temporal constraints remained unsolved. Mobile Recommender Systems (MRS) [2] for tourism provided mechanisms for routes and tour recommendations using classification model. But, MRS took long hours and absolutely failed in maintaining the time and budget metrics.

Analyzing the path on the basis of least cost is one of the cheapest means for connecting two different locations that can be obtained by integrating several criteria. Geographic Information System (GIS) tools [3] were introduced with the aid of analytical hierarchy process in obtaining the hybrid route. Frequent trajectory patterns using travel time estimation was applied in [4] with the objective of providing scalability. Moving object pattern mining and trajectory mining were integrated in [5] called as MoveMine that provided scalability and flexibility that provided interactive exploration of mining results. Software Defined Networking (SDN) [6] for traffic engineering was applied to address issues related to traffic engineering problems like fault tolerance and traffic analysis for identification of best route paths.

In the recent years, one of the most widespread measures for route path identification is the design of intelligent transportation systems. Classification of trajectories for road networks was addressed in [7] for effective transportation and traffic planning using frequent pattern based classification. The results obtained proved to be efficient in terms of accuracy, but to classify more sophisticated patterns, it involved much time. Appropriate network distance queries was applied in [8] to solve the issues related to time and provided means for low error. However, traffic congestion remained unsolved. Anonymous query processing was introduced in [9] to address the issues related to traffic congestion by designing edge order anonymization approach by identifying border nodes that also reduced the overhead for identifying route path. Superseding Neighbor Search [10] with the application of superseding graph and superseding relationship identified the efficient path from source node to destination node.
In this paper, we present a Transportable Road Route Path Miner (TRRPM) method with the objective of achieving routing efficiency on a single modal road route path and minimizing the total tuning time. The contributions of TRRPM include the following:

(i) To achieve a single modal road route path with routing efficiency using Transportable Road Route Path Miner (TRRPM) method.

(ii) Efficient construction of route path by continuously extracting the user packet profiles information using Secluded Supervision technique.

(iii) To group the profile information about vehicles using the cosine similarity based on the frequent movement of different vehicle patterns and time factor.

(iv) To reduce the congestion by obtaining the temporal data as sample item set and

(v) To optimize the results on constructed road network route path using Shortest Route Path based K-fold Cross Validation Tuning Method for vehicular communication.

The rest of the paper is organized as follows: Section 2 reviews the previous work related to route path identification and their limitations. Section 3 formally defines the problem of route maintenance and illustrates its characteristics with the design of transportable road route path miner with the aid of a neat architecture diagram. Section 4 provides with the experimental design to carry our TRRPM. Section 5 discusses in detail with the aid of table and graph form. Finally, Section 6 concludes with the concluding remarks.

2 RELATED WORKS

A trajectory includes data points in sequence without loss of information regarding the routes where each data point objective is to record the information regarding the location and its time stamp. Route Inference based on Collective Knowledge [11] was designed with the objective of providing top-k routes to the users and vehicles using routing algorithm. However, the route preferences of the user were not taken into consideration. Similarity measure for moving objects was used to extract the preferences of users by integrating local mining and cluster ensembling [12]. Temporal data clustering framework was designed in [13] to address representation based temporal data clustering for route path identification using weighted consensus function and partition weighting scheme. Predictive Energy-Efficient Bee-inspired Routing (PEEBR) [14] was designed to provide solutions for routing problem in order to maximize the network lifetime.

In [15], the design of a single route was provided that identified the maximum duration and maximum budget for the design of infrastructure with the objective of identifying efficient route using integer linear programming model. However, several origin-destination pairs were not considered at the same time. Identification of congested routes using road trajectories was addressed in [16] by analyzing the saturation degree of road segment and measuring average moving speed of vehicles in different directions. An adaptive fast path algorithm was designed in [17] using area partition algorithm to obtain several driving pattern mining with the objective of fastest path computation at relatively lesser amount of time.

Though significant amount of research works have been concentrated on sensing and identification of several important places for the users, very little efforts have been made to identify the most efficient routes. A passive route sensing framework was designed in [18] that differentiated and identified different routes using time warping angular speeds and identified the best route from it. However the method had constraints with respect to the movement of different vehicles.

Based on the abovementioned methods, in this work we concentrate on the design of an efficient transportable road route path miner with the objective of improving the routing efficiency and computational cost factor which is briefed in the following sections.

3 DESIGN OF TRANSPORTABLE ROAD ROUTE PATH MINER

In this section, we formulate the problem of mining transportable road route path for efficient vehicular packet communication. This section provides the necessary background regarding optimal route path and the proposed method, transportable road route path miner. The three components involved are (i) vehicle profile information extraction (ii) cosine similarity based clustering and (iii) tuning of route path.

Route path construction using the clustering and tuning procedure in our proposed work helps in performing the vehicular packet communication on the Geographic Information System (GIS). Data mining technique is used in the construction of Transportable Road Route Path Miner (TRRPM) method to easily construct the route map for effective vehicular packet communication. The single modality system (i.e., road network) is taken for route path construction with the temporal data information. Currently, the road network gets congested due to the multiple vehicular users at a particular period of time. This network results in
congestion on the route path. To overcome the issues, data mining techniques are adopted in the proposed work.

Figure 1 shows the occurrence of congestion on the road network and integration of the data mining tools to control the congestion. Initially, the moving streams of the vehicular packets are analyzed and congestion route path is identified. In our proposed work, data mining technique like supervised learning, clustering and tuning procedure are employed to construct the effective route path for single modality system. The congestion occurring vehicular packets are detected and then the profile information about the packet is obtained using the supervised learning.

The supervised learning helps to group the profiles of similar information and improve the clustering efficiency rate. Finally, the tuning procedure is employed in TRRPM for obtaining the shortest route path. The shortest route path constructed with these techniques developed a time effective limited budgeting with the introduction of single modality routing system. The architecture diagram of Transportable Road Route Path Miner (TRRPM) method is shown in Figure 2.

Figure 1 Congestion Control based on Data mining technique

Figure 2 Architecture Diagram of TRRPM Method

As illustrated in Figure 2, congested single modality system is taken and performs the proposed method to provide a congestion control shortest route path construction strategy. To start with, the vehicular packets (i.e., user) information is collected and performs supervised data mining learning procedure. In TRRPM, Secluded Supervision learning technique is adopted. The supervised learned result is used to group similar cosine relationship information. The similar cosine relationship information helps to easily group the similar patterns based on the time factor. Finally, the similar patterns grouped in TRRPM, uses the tuning procedure to identify the shortest road route path for vehicular communication from the source to the destination end through the intermediate points. An elaborate description about the Transportable Road Route Path Miner method is explained in the forthcoming sections.

3.1 VEHICLE PROFILE INFORMATION EXTRACTION

The first component included in the design of Transportable Road Route Path Miner is the extraction of vehicle profile information. The vehicle profile information extraction in TRRPM aims to mine the shortest route path for vehicular packet transformation through road network. The road network connectivity captures the similar vehicle profile information among the vehicles. First, a secluded (i.e.,
distance node isolated points) based supervised learning data mining approach in proposed road network method to extract the user profile information is presented. The supervised learning in data mining procedure uses the past experiments to gain knowledge and insights about the users in order to improve the performance rate.

3.1 SECLUDED SUPERVISION TECHNIQUE

The secluded supervision technique in TRRPM method extracts the profiles of the frequent moving pattern. The objective of secluded supervision technique is to mine the frequent moving patterns from the road network using the prior user knowledge. An effective secluded supervision is a method that learns the relationship between the weakly labeled sources and the destination end data points.

Secluded learning is introduced in our proposed work for easy cross learning of data points for shortest route path identification on single modality road system. In the data points, all the entities (i.e., structure of the data) are identified and with the help of entity tagger, the information is extracted. Using entity tagger, if the vehicle profile information contains two entities, and those entities are in relationship with one another, then it is easiest way to extract the vehicular profile information.

\[
\text{Entity Tagger} \ (U_1, U_2, \ldots, U_n) = \{\text{Entity1} \rightarrow \text{Entity 2}\}
\]  

Let us assume that the source-destination location information and their relationship between entities are gathered (i.e., extracted) in the road network vehicles. In the testing step of TRRPM method, entity relationship is identified using the named entity tagger. Every pair of entities in the vehicular profile user information '\(U_n\)' considers the potential relationship using the secluded learning. Each entity pair of information extracts the user profile information for further grouping of similar user profile vehicles with effective load balancing factor on single modality road system.

3.2 COSINE SIMILARITY BASED CLUSTERING

The second component included in the design of Transportable Road Route Path Miner is the cosine similarity based clustering. In order to decide the relationship between the extracted user vehicular profiles, the clustering operation is carried out. The results obtained using secluded supervised learning technique is used for clustering user profile based on the timely basis. Temporal data are considered in our proposed work for frequent moving pattern clustering. Cosine similarity in TRRPM method is a measure of the similarity between the user profiles information based on the inner product time and space of the vehicles. The cosine similarity based clustering uses cosine angle measurement with the magnitude and orientation of vehicle moving patterns in road network.

\[
\text{Cosine Similarity (CS)} = \frac{\sum_{i=1}^{n} U_i U_{i+1}}{\sqrt{\sum_{i=1}^{n} U_i^2 \sum_{i=1}^{n} U_{i+1}^2}}
\]  

(2)

The cosine similarity of the profile '\(U_i\)' and '\(U_{i+1}\)' is measured based on the orientation (i.e., direction of the vehicle moving pattern) and magnitude (i.e., size of the moving vehicle pattern) in road network. The vehicle moving patterns are also clustered based on the entity tagger with the user profile extracted information. The normalized angle (\(\theta\)) measured for identifying the cosine similarity between the two user profile vectors for measuring the value is defined as,

\[
\theta = \frac{1 - \cos^{-1}(CS)}{\pi}
\]  

(3)

The cosine similarity result of the user profile vehicle with same magnitude and orientation on timely based is grouped together and their measurement is denoted as ‘1’. If the user profiles of the cosine similarity produces different result, then that road user vehicular profile are placed in the neighboring cluster set. Different set of result on comparing the relation of the two user profiles using the cosine similarity produce the result set as [0, 1] in TRRPM method.

\[
\begin{array}{c}
\text{Index} \\
\text{Similar Cosine Values (1)} \\
\text{Secluded Supervised Learning} \\
\text{Cluster} \\
\text{Dissimilar Cosine Values (0)}
\end{array}
\]

Figure 3 Clustering of Vehicle Profile Using Cosine Similarity

Figure 3 shows the application of cosine similarity to group similar vectors based on the orientation and magnitude with the resultant information obtained through secluded supervised learning information. The cosine values are measured and their information is indexed for cross validating. The clustering of the profiles helps to easily cross validate the profiles and identify the shortest road network route path with minimal computational cost factor. The advantages point on
the cosine similarity in TRRPM method is the angular distance and timely basis of clustering.

3.3 Tuning of Route Path

Finally, the last component in the design of Transportable Road Route Path Miner is the tuning of route path based on the K-fold cross validation method. Tuning data mining approach developed in our proposed work help to facilitate the route path selection with the shortest intermediate node points. Tuning procedure is employed to attain high quality result on single modality road network construction. The temporal frequent data moving pattern are used widely for tuning the data set and identify the shortest route path. The route path provides the destination specific optimization based on the K-fold cross validation tuning method.

3.3.1 K-Fold Cross Validation Tuning Method

K-fold cross validation tuning method is a technique used to select the shortest path on single modality road network. The single modality road network in TRRPM method uses the indexed cluster information for K-fold tuning of information and selects the shortest path. The K-fold cross validation tuning method over different road path is formalized as,

\[ S_K(p_1,p_2...pn) = \sum_{i=1}^{k} \text{Min(CS } (p_1,p_2...pn)) \]  

(4)

The shortest path represented by \( S_k \) in (4) denotes the shortest path over the fold of different path in road network. The minimum path distance attained using the cosine similarity clusters measure value. The algorithmic pseudo code for producing the accurate shortest path is administered and it is described as,

\begin{verbatim}
Input: User Profile Information \{Ur1, Ur2, ..., Urn\}. Different paths \{p1, p2, ..., pn\}. Shortest path over K fold \( S_k = S_1, S_2, ..., S_n \}\{i, j, k, \alpha, \beta, ...\} \{Entity_a = Entity_1, Entity_2, ..., Entity_n\}

begin
1: for each \( Ur_i \)
2: extraction of user vehicle information
3: Entity Tagger (\( Ur_1, Ur_2, ..., Ur_n \)) = \{Entity1 \rightarrow Entity 2\}
4: analyze temporal packet moving patterns
5: identify cosine similarity measurement for \( Ur_i \) and \( Ur_{i+1} \)
6: perform clustering for identifying the cosine similarity between the two user profile vectors
7: measure normalized angle
8: evaluate K-fold cross validation using \( S_k(p_1, p_2 ... pn) \)
9: end for
end

Return \( SPR(S_k(p_i)) \)
\end{verbatim}

The above algorithmic point concern with the tuning procedure in our proposed work helps to identify the shortest path with high routing efficiency. The cross validation K-fold based routing for single modality road system improves the load balancing factor with minimal processing time. The total tuning time is reduced as per the K-fold based shortest path routing for the vehicular frequent pattern movement from the source node to destination point. The routing procedure further produces high performance result factor using TRRPM method.

4 Experimental Evaluation

Transportable Road Route Path Miner (TRRPM) method performs the experiment in JAVA platform using the Weka Tool. JAVA platform uses this tool to develop the road network for effective data transfer from the source end to the destination using the Bike Direct Network Dataset. Bike Direct is a strategic network of routes across the Adelaide metropolitan region in shape file format for the experimental evaluation. This data set is taken from the South Australian Government Data directory link to perform the experimental work. TRRPM method takes this dataset to map the route for the road network.

The vehicular communication on road network (i.e., single modality) develops the route map for communication between different types of vehicle. Bike Direct Network set information are used to compare the result of the TRRPM method with the existing general Optimal Route Queries with Forward and Backward Search (ORQ-FBS) [1] and Mobile Recommender Systems (MRS) [2] for tourism. Experiment is conducted on factors such as routing efficiency on single modal system, computational cost factor and total tuning time.

The routing efficiency on single modal system measures the efficiency of the network through which the routing can be performed. The computational cost factor measures the time taken to perform the computation based on the magnitude and orientation of vehicles in road network. It is measured in terms of milliseconds (ms). The mathematical formulation for computational cost is given as below:

\[ CC = Time \times \text{Cosine Similarity (CS)} \]
The computational cost as given above refers to the time taken to perform the cosine similarity of the user profile information ‘Uᵢ’ and ‘Uᵢ₊₁’ in terms of orientation and magnitude. The total tuning time using TRRPM method is the time taken to perform the tuning procedure on single modality road network construction. It is measured in terms of milliseconds (ms). The total tuning time is given as below

\[ Tuning time = \text{Time} \left( \min(p₁ + p₂ + \ldots + pₙ) \right) \]  

The total tuning time given above is the summation of total time taken to identify the intermediate node points over different road path p₁, p₂, pₙ in road network.

5 RESULTS ANALYSIS OF TRRPM

In order to analyze the characteristics and functionality of the TRRPM method, we quantitatively accessed the performance on the basis of Bike Direct Network Dataset by comparing the outcomes to the results achieved with the Shortest Path Routing (SPR) algorithm. The Transportable Road Route Path Miner (TRRPM) method is compared against the existing general Optimal Route Queries with Forward and Backward Search (ORQ-FBS) [1] and Mobile Recommender Systems (MRS) [2] for tourism. The experimental results using JAVA are compared and analyzed with the help of values obtained during experimentation which is provided in the table and graphical form as given below. To support transient performance, in Table 1 we apply Shortest Path Routing (SPR) algorithm to obtain the routing efficiency and comparison is made with two other existing methods, ORQ-FBS and MRS.

Table 1 Routing efficiency with respect to user profile information

<table>
<thead>
<tr>
<th>User Profile Information (U)</th>
<th>Routing Efficiency (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>TRRPM</td>
</tr>
<tr>
<td>3</td>
<td>55.92</td>
</tr>
<tr>
<td>6</td>
<td>61.35</td>
</tr>
<tr>
<td>9</td>
<td>68.22</td>
</tr>
<tr>
<td>12</td>
<td>62.45</td>
</tr>
<tr>
<td>15</td>
<td>65.33</td>
</tr>
</tbody>
</table>

Figure 4 Measure of routing efficiency

Figure 4 shows the measure of routing efficiency with respect to user profile information obtained from the user or vehicles in the range of 3 to 21 using Bike Direct Network Dataset. From the figure 4, we can observe that the routing efficiency using the Transportable Road Route Path Miner (TRRPM) method is high when compared to ORQ-FBS [1] and MRS [2]. This is because of the application of secluded supervision that obtains the user profile information in an efficient manner.

This is due to the fact that using TRRPM method similar vehicle profile information is captured based on supervised learning data mining approach called the secluded supervision technique. The secluded supervision technique efficiently isolates the distance node points using the past experiments with the aid of prior user knowledge. In addition, with the help of entity tagger, the source-destination location information and their relationship between entities are gathered using TRRPM method, improving the routing efficiency on single modal system by 15 – 21 % compared to ORQ-FBS and 22 – 32 % compared to MRS respectively.
Table 2: Computational cost involved during the design of TRRPM, ORQ-FBS and MRS

<table>
<thead>
<tr>
<th>User Profile Information (U)</th>
<th>Computational Cost (ms)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>TRRPM</td>
</tr>
<tr>
<td>3</td>
<td>0.0133</td>
</tr>
<tr>
<td>6</td>
<td>0.0142</td>
</tr>
<tr>
<td>9</td>
<td>0.0139</td>
</tr>
<tr>
<td>12</td>
<td>0.0143</td>
</tr>
<tr>
<td>15</td>
<td>0.0148</td>
</tr>
<tr>
<td>18</td>
<td>0.0141</td>
</tr>
<tr>
<td>21</td>
<td>0.0144</td>
</tr>
</tbody>
</table>

The comparison of computational cost is presented in table 2 with respect to the varying user profile information in the range of 3 – 21 collected at different time periods between 7 am and 7 pm. With increase in the number and size of user profile information, the computational cost to obtain similar user profile information is also increased.

This is because of the application of secluded supervision technique.

The computational cost factor is reduced due to the clustering of vehicle profile information obtained from secluded supervision through cosine similarity based clustering. By applying cosine similarity based clustering, based on timely basis, inner product time and space of the vehicles are obtained that efficiently measures the similarity between the user profiles information and reduces the computational cost factor by 8 – 9 % when compared to ORQ-FBS. Besides, the cosine angle measurement with the magnitude and orientation of vehicles in road network, based on the entity tagger is grouped on timely basis helps in easily cross validation of profiles with minimal computational cost factor by 15 – 17 % compared to MRS.

Table 3: Tuning time with respect to different number of paths

<table>
<thead>
<tr>
<th>Paths (p)</th>
<th>Tuning time (ms)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>TRRPM</td>
</tr>
<tr>
<td>1</td>
<td>2.32</td>
</tr>
<tr>
<td>2</td>
<td>2.75</td>
</tr>
<tr>
<td>3</td>
<td>2.99</td>
</tr>
<tr>
<td>4</td>
<td>3.12</td>
</tr>
<tr>
<td>5</td>
<td>3.15</td>
</tr>
<tr>
<td>6</td>
<td>3.18</td>
</tr>
</tbody>
</table>

The tuning time for TRRPM method is elaborated in table 3. We consider the method with varying number of paths acquired from Bike Direct Network Dataset at different time period for experimental purpose using JAVA.

Figure 5 Measure of computational cost

To ascertain the performance of the computational cost, comparison is made with two other existing methods, general Optimal Route Queries with Forward and Backward Search (ORQ-FBS) [1] and Mobile Recommender Systems (MRS) [2] for tourism. In Figure 5, the sizes of user profile information are varied between 3 and 21. From the figure it is illustrative that the computational cost is lesser using the proposed Transportable Road Route Path Miner (TRRPM) method when compared to the two other existing methods.

Figure 6 Measure of tuning time
In figure 6, we depict the tuning time attained using seven different paths for experimental purpose and applied in JAVA. From the figure, the value of tuning time achieved using the proposed Transportable Road Route Path Miner (TRRPM) method is lower when compared to two other existing methods namely, Optimal Route Queries with Forward and Backward Search (ORQ-FBS) [1] and Mobile Recommender Systems (MRS) [2] for tourism. Besides we can also observe that by increasing the number of paths, the tuning time is also increased using all the methods. But comparatively, it is lower in TRRPM method because of the application of K-fold Cross Validation Tuning method that performs the tuning of clustered profile in an efficient manner.

By applying K-fold Cross Validation Tuning method, shortest path on single modality road network is selected that drastically reduces the total tuning time by 9 – 13 % compared to ORQ-FBS. In addition, the Transportable Road Route Path Miner (TRRPM) method selects the shortest path by applying indexed cluster information for K-fold tuning of information over different fold in road network. This improves the load balancing factor with minimal total tuning time by 16 – 22 % compared to MRS.

6 CONCLUSION

A Transportable Road Route Path Miner (TRRPM) method has been designed to achieve robustness in terms of routing efficiency on single modal system and to minimize the computational cost factor using the cosine similarity function. The Shortest Path Routing (SPR) algorithm is adopted that uses the entity tagger and extracts the vehicle profile information of those entities that are in relationship with one another with minimal tuning time. The Transportable Road Route Path Miner (TRRPM) method is adaptive because of the introduction of Secluded Supervision that applies supervised learning data mining approach to isolate the distance node points. In addition, the cosine similarity based clustering uses magnitude and orientation of vehicle moving patterns in road network for grouping vehicle information on timely basis. Finally, the tuning of route path is performed efficiently by applying K-fold cross validation that selects the shortest path on single modality road network. Experimental evaluation is conducted with the Bike Direct Network Dataset to analyze the robustness with respect to routing efficiency on single modal system and measures the performance in terms of computational cost, tuning time. Performances results reveal that the proposed TRRPM method provides higher level of routing efficiency and minimize the tuning time and computational cost on single modal system. Compared to the existing route identification methods, the proposed Transportable Road Route Path Miner (TRRPM) method is comparatively better than compared to state-of-art works.