Discovering Maritime Traffic Route from AIS Network

Po-Ruey Lei
Department of Electrical Engineering
ROC Naval Academy
Zuoying, Taiwan, ROC
Email: cnabarry@gmail.com

Tzu-Hao Tsai
Department of Computer Science
National Chiao Tung University
Hsinchu, Taiwan, ROC
Email: stu8978@gmail.com

Wen-Chih Peng
Department of Computer Science
National Chiao Tung University
Hsinchu, Taiwan, ROC
Email: wcpeng@gmail.com

***Abstract*—The recent build-up network of Automatic Identification System (AIS) equipped on vessels provides a rich source of vessel movement information. AIS is originally designed for automatically exchanging navigation information, such as their unique identification, position, course, and speed, with nearby vessels and terrestrial receivers to affect collision avoidance and safety control. The collected sequences of AIS logs can be considered as maritime trajectory data, i.e., the sequences of location points with timestamps. This vast amount of AIS trajectory data can be collected and employed to achieve an awareness of maritime traffic knowledge.**

This paper is devoted to discovery of maritime traffic route from trajectory data generated by AIS networks. However, AIS trajectory data discovery is a challenging task because of the trajectory data is available with uncertainty. Furthermore, unlike the vehicles' movements are constrained by road networks, there is no such a sea route for vessels to follow in marine areas. To overcome the challenges, we propose a framework of Maritime Traffic Route Discovery (abbreviated as MTRD) to generate pattern-aware routes to achieve an effective understanding of maritime traffic awareness. The proposed framework is evaluated on real AIS data and the experimental results shows that the proposed MTRD is able to extract the marine traffic route effectively and provides a cornerstone of maritime traffic knowledge for traffic management, anomaly detection, and conflict analysis in the future.

Keywords—Maritime Traffic knowledge, AIS System, Trajectory data, Trajectory pattern mining, Traffic Route Discovery

I. INTRODUCTION

Recently, the pervasiveness of Automatic Identification Systems (AIS) has made a large number of the vessels' movement data to be available. In order to affect collision avoidance and safety control, vessels equipped with AIS automatically exchange navigation information, such as their unique identification, position, course, and speed, with nearby vessels and terrestrial receivers of AIS networks to facilitate the tracking and monitoring of vessel location and movement. The collected sequences of AIS logs form maritime trajectory data, i.e., the sequences of location points with timestamps. Due to the trajectory data records vessels' real movements, AIS data is a valuable data source and gives opportunity used for discovering the traffic knowledge from historical AIS trajectories. Accordingly, there is an ever-increasing interest in performing data analysis over AIS data and develop many applications for maritime traffic management [8], [2], [1], anomaly detection [10], [6], and collision risk analysis [13], [12], [18].

In this work, we focus on the problem of discovering the maritime traffic routes from historical trajectories generated by AIS network. However, discovering the maritime traffic routes from AIS trajectories is a non-trivial problem. As numerous researches in the literature have addressed [3], [5], [14], the trajectory data is available with uncertainty. The same situation occurred in trajectory data generated by AIS network. The uncertainty of AIS trajectories would be caused by location sensing techniques and sampling. AIS uses Global Positioning System (GPS) in conjunction with shipboard sensors and digital VHF radio communication equipment to automatically exchange navigation information electronically. The location information of AIS data may be generated with inaccurate GPS measurement or the data may be incomplete caused by radio signal attenuation and loss. Additionally, the transmitters of AIS broadcast data every 2 to 10 seconds depending on a vessel's speed while underway, and every 3 minutes while vessels are at anchor. The irregular and asynchronous sampling would lead to that a vessel's movement may not exactly repeat the same trajectory even if the vessel has similar movement behavior to others.

Moreover, discovering the traffic knowledge from maritime trajectory is made even more difficult due to the maritime area is a free moving space. Unlike the vehicles' movements which are constrained by road networks [14], [16], there is no sea route for vessels to follow in maritime areas. Fig. 1 shows a set of AIS trajectories collected from a maritime area. Obviously, the vessels are moving free and AIS trajectory data is more complex than the trajectories moving along the road network as shown in Fig. 2. Extracting movement behavioral knowledge from uncertain AIS trajectories is a challenging task. Thus, for providing an effective overview of maritime traffic knowledge, we aim to achieve the discovery of common movement behaviors from AIS data and identify the maritime traffic routes.

There are numerous researches that have made efforts to discover movement behavior from trajectory data [3], [5], [17], [4], [7], [14], [16]. However, most of them focus on discovering movement pattern from the trajectory data constrained by road networks. The authors in [5], [7] proposed a partition-and-group framework to detect the trajectory clusters from the hurricane data set and the animal movement data set. The framework generates a representative trajectory of each cluster.

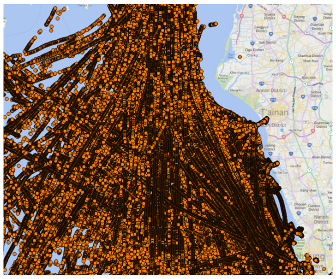

Fig. 1. AIS trajectory data in maritime area

Fig. 2. Taxi trajectory data in road network

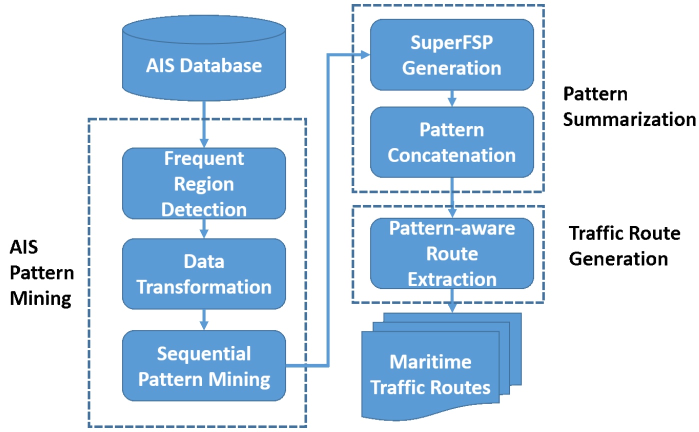

Fig. 3. Framework of maritime traffic routes discovery

Although, in [9], [1], the authors proposed methodology called TREAD automatically learns a synthetic representation of maritime traffic patterns from AIS data. The extracted traffic route is represented in a synthetic route composed of the nodes and segments. Both of them discover the typical movement pattern and exhibited in a compact representation. They did not provide the detail movement route for moving objects and marine vessels.

The purpose of this work is to discover the maritime traffic routes from the trajectory data collected by AIS network. More specifically, we not only discover the trajectory pattern but also extract the movement area of traffic route from those discovered trajectory patterns. We propose a framework of Maritime Traffic Route Discovery(abbreviated as MTRD) to generate pattern-aware routes and achieve a effective understanding of maritime traffic awareness. We leverage the knowledge of movement pattern discovered from a given set of AIS trajectories to generate the traffic routes. The extracted knowledge of maritime traffic routes is able to contribute to many applications such as employing the analysis of route flow and density for maritime traffic management, supporting situation awareness in maritime surveillance, and constructing a normal model of vessel movement behaviors for anomaly detection.

The remainder of the paper is organized as follows. The proposed framework is generally introduced in Section II. Section III technically details the main modules of the proposed framework. In Section IV, we perform an empirical performance evaluation on real data set collected from AIS network. Finally, we summarize our conclusion and future work in Section V.

II. Framework Overview

Given a set of data collected from AIS network, the proposed framework of Maritime Traffic Route Discovery (abbreviated as MTRD) automatically discovers the movement patterns from AIS data in an unsupervised way and then the maritime traffic routes are extracted from the discovered patterns. Figure 3 outlines the functional architecture of MTRD, which comprises of three modules: *AIS Pattern Mining*, *Pattern Summarization*, and *Traffic Route Generation*.

Unlike the traditional maritime surveillance sensors, such as radar or GPS, the AIS system comprehensively represents the identity and properties of a vessel, as well as its behavior. In other words, AIS is able to facilitate the tracking and monitoring of vessel location and movement. Each vessel tracked by AIS is characterized by the properties, including the vessel's static parameters(e.g. name, flag, type), and the current state of its dynamic behavior(e.g. speed, course, location). Based on the vessels' dynamic behavior records originated by the AIS network, the collected sequences of AIS logs can be considered as AIS trajectory data. Each trajectory is represented by a sequence of spatial-temporal points, $T_i = \{(x_1, y_1, t_1), (x_2, y_2, t_2), ..., (x_n, y_n, t_n)\}$, where n is the total number of points.

Based on AIS trajectory data, the objective of **AIS Pattern Mining** module is to discover the trajectories with similar movement behavior in form of AIS trajectory pattern. To overcome the problem of uncertainty and moving free within AIS trajectory data, the concept of frequent region is adopted. We use a frequent region to represent a specific location and transform the a point-based trajectory into a sequence of frequent region. The problem of discovering similar trajectories can be processed as a problem of sequential pattern mining. Thus, a AIS pattern is represented as a frequent sequential pattern(FSP).

Additionally, the huge number of FSPs may be generated by the frequent pattern mining process. The unwieldy number of frequent patterns makes the patterns themselves difficult to explore. The **Pattern Summarization** module is developed to summarize the large number of generated patterns using the representative patterns, called summarized pattern.

Finally, the **Traffic Route Generation** module provides a solution to extract traffic routes from each summarized pattern, i.e., to generate the pattern-aware routes. More specifically, the task of the module is to detect a movement channel of a traffic route followed by a group of AIS trajectories those having similar behavior within a summarized pattern.

III. Maritime Traffic Route Discovery

In this section, we technically detail the three main modules of the proposed framework MTRD: AIS Pattern Mining, Pattern Summarization, and Traffic Route Generation.

A. *AIS Pattern Mining*

To deal with the problem of uncertainty and moving free space within AIS trajectory data, the AIS pattern mining module is developed based on the approach of trajectory data mining to explore the vessels' movement behavior from AIS data. Without loss of generality, given a set of trajectories, algorithms of mining movement behaviors will first extract some regions with a certainty degree of popularity, which are referred to as frequent regions. Then, original trajectories are transformed as sequences of frequent regions. With a given sequences of frequent regions, movement behaviors are thus defined as trajectory patterns that frequently appear among sequences of frequent regions. Clearly, trajectory patterns imply that objects usually follow similar movement behaviors.

Thus, the AIS pattern mining module includes three steps: frequent region detection, data transformation, and sequential pattern mining. First, in frequent region detection, the AIS trajectory data is mapped into grid system, and then a cell is detected as a frequent region r_j if the number of trajectories passed the cell has satisfied the user-defined minimum support threshold $MinTs$. More clearly, a frequent region is a grid cell that contains at least $MinTs$ number of trajectory segments passing by the grid cell. Second, the problem of discovering similar trajectories can be processed as a problem of sequential pattern mining by data transformation. Based on the discovered frequent regions, each AIS trajectory is transformed into the region-based trajectory with corresponding frequent regions, i.e., a sequence of frequent region. Note that the points those are not in frequent regions will be regarded as noise. As such, the movement behaviors can be captured by the mobility relations between frequent regions. Finally, the AIS pattern mining is able to realized by the method of sequential pattern mining. In this work, Prefixspan algorithm [11] is applied to mine the frequent sequential patterns(abbreviated as FSPs).

B. Pattern Summarization

After the procedure of AIS pattern mining, the huge number of FSPs can be produced by the frequent pattern mining process. However, the unwieldy number of discovered frequent sequential patterns makes the understanding of generated patterns troublesome. Thus, the pattern summarization module is developed to summarize the large number of generated patterns by using the representative patterns, called summarized patterns(abbreviated as SPs). This module includes two steps: SuperFSP Generation and Pattern Concatenation. The proposed two steps of pattern summarization are developed based on two observations on generated FSPs from AIS trajectory data.

First, we observe that some generated patterns are contained in other super patterns. For example, given three FSPs as $FSP_1 = \{r_1, r_2, r_3, r_4, r_5\}$, $FSP_2 = \{r_1, r_3, r_5\}$, and $FSP_2 = \{r_1, r_3, r_9\}$, FSP_1 is considered as the SuperFSP of FSP_2 due to the elements of FSP_2 could be located with ordered in FSP_1. Intuitively, FSP_1 is not the SuperFSP of FSP_3. In order to promises that each SuperFSP is a unique FSP, The proposed method of SuperFSP generation use the approach for solving the longest common subsequence problem to prune the FSPs those are not super.

Second, Pattern Concatenation is proposed to solve the problem of data uncertainty and incompleteness occurred in AIS trajectory data. The uncertainty and incompleteness in AIS data may cause the incompleteness of generated patterns. By observation, some generated patterns are SuperFSP but have partially similarity with others. For example, given two FSPs as $FSP_1 = \{r_1, r_2, r_3, r_4, r_5\}$ and $FSP_2 = \{r_3, r_4, r_5, r_6, r_7, r_8, r_9\}$. Compared the last three elements of FSP_1 and first three elements of FSP_2, the similar subsequence is $\{r_3, r_4, r_5\}$. FSP_1 and FSP_2 are supposed to have the sequential relations in movement behavior. Those two FSPs are able to be concatenated as the summarized pattern $\{r_1, r_2, r_3, r_4, r_5, r_6, r_7, r_8, r_9\}$. However, the generated patterns is not perfect in real world. If the length of the subsequence comparison is fixed, such pattern concatenation based on last-and-first similarity may fail in many cases. For instance, FSP_1 and FSP_2 are compared by the subsequence length of 2, i.e., $FSP_1.last(2) = \{r_4, r_5\}$ and $FSP_2.first(2) = \{r_3, r_4\}$. They may not be considered as a set of patterns those have the sequential relations in movement behavior. To realize the idea, we propose a last-and-first pattern concatenation and develop a last-and-first similarity (LAF) to be the measurement for pattern concatenation. Given two sequential patterns sp_1 and sp_2, the last-and-first similarity (LAF) based on the compared length ℓ is defined as follows:

$$\text{LAF} = \frac{|LCS(sp_1, sp_2)|}{\ell}, \tag{1}$$

where $|LCS(sp_1, sp_2)|$ is the length of longest common subsequence in sp_1 and sp_2. For effectively pattern concatenation, last-and-first pattern concatenation method dynamically measures the last-and-first similarity and merge the generated patterns those have the max similarity in sequential relations. For example, we evaluate the last-and-first similarity between FSP_1 and FSP_2 by LAF with ℓ from 2 to 4. Then, we obtain LAF(ℓ=2)=1/2, LAF($\ell = 3$)=1, and LAF($\ell = 4$)=3/4. Therefore, FSP_1 and FSP_2 is able to concatenated at the length 3 within the max similarity in last-and-first sequential relations.

The algorithm of pattern summarization is detailed in Algorithm 1. Given a set of AIS patterns (D_{AP}) mined from a set of AIS trajectory data. The algorithm outputs a set of summarized patterns (D_{SP}) according to user-defined minimum last-and-first similarity LAF_{min}. The algorithm includes two steps: SuperFSP Generation and Pattern Concatenation. First, in the step of SuperFSP Generation (line 1 to line 8), the AP_s those enclosed by other patterns are pruned from D_{AP}. Thus, we can promise that each AIS pattern remained in D_{AP} is a unique. In other words, the AIS patterns in $D_{AP.SuperFSP}$ are SuperFSPs(line 8). Then, in the step of Pattern Concatenation (line 9 to line 16), we evaluate the last-and-first sequential relations for each pair of AP_i and AP_j in $D_{AP.SuperFSP}$. Specifically, the last-and-first similarity (LAF) are compared by dynamic length for effectively evaluation. The max range of compared length is bounded by $\ell_{max} == |min(AP_i, AP_j))|/2$ and minimum compared length is required 2. The pair of patterns in $D_{AP.SuperFSP}$ are concatenated if LAF of the pair is satisfied the minimum requirement LAF_{min}. That is those patterns have the last-and-first sequential relations in movement behavior. Finally, the unwieldy number of discovered AIS patterns can be summarized as a representative set of summarized patterns.

C. Traffic Route Generation

Given a summarized pattern, the Traffic Route Generation module extracts a possible traffic route hidden in the summarized pattern. That is, the proposed traffic route is a pattern-

Algorithm 1 : Pattern Summarization

Input: a set of AIS Pattern D_{AP} and the minimum last-and-first similarity LAF_{min}
Output: a set of Summarized Pattern D_{SP}

foreach AP_i and AP_j in D_{AP} **do**
  ℓ_{lcs}=LCS(AP_i,AP_j)
  $\ell_{min}= |min(AP_i, AP_j)|$
  if $\ell_{lcs} == \ell_{min}$ **then**
    remove $min(AP_i, AP_j)$ from D_{AP}
  end
end
$D_{AP.SuperFSP} \leftarrow D_{AP}$
foreach AP_i and AP_j in $D_{AP.SuperFSP}$ **do**
  Max.Compared Length $\ell_{max}= |min(AP_i, AP_j)|/2$
  for ℓ = 2: ℓ_{max}
    if LAF$(AP_i, AP_j, \ell) \geq LAF_{min}$ **then**
      $D_{SP} \leftarrow D_{SP} \cup$ concatenate(AP_i, AP_j)
    end
  end
end

aware route. As the the result of early procedure proposed, a summarized pattern is a sequence of frequent regions followed by many historical AIS trajectories. The idea for pattern-aware route generation is performing statistical analysis on those trajectories within each region of the pattern sequentially. The approach of region-based mobility analysis in our previous work [15] is adopted to materialize the pattern-aware route generation. The movement behavior of a region is extracted from AIS trajectory and represented by a mobility vector. Then, a maritime traffic route can be generated and represented by a sequence of ordered mobility vectors. As shown in Fig 4, the mobility vector $V_m(r_j) = \{D, S_c\}$ is represented by major direction D and crossing-section S_c in a spatial region. The major direction summarizes the movement direction of majority trajectories in the region. The crossing-section defines the spatial borders in the region while the most of trajectories crossed the region. For a region r_j, we extract a set of velocity vectors from each pair of conjunctive trajectory points those exist in the region. Major direction D is computed by the average of the velocity vectors in the region and then all points are projected on the line perpendicular to major direction. Then, based on the projected points, we derive the crossing-section S_c =(*center, left-margin, right-margin*) by statistical analysis. The center point is the average of the projected points. In order to provide a movement channel of a traffic route within which most of trajectories moving along, we define the route margin, i.e., left-margin and right-margin, on each side of mobility vector and remove the outlier. According to empirical rule in statistics, about 95 percent of data are within two standard deviations if a data distribution is approximately normal distribution. Thus, left-margin and right-margin are derived at the distance of two standard deviations away from center point. Sequentially, the traffic route is generated in terms of an ordered sequence of $V_m(r_j)$ discovered from a set of trajectory within a summarized pattern.

Figure 5 is a running example for maritime traffic route generation. Given a summarized pattern, velocity vectors are extracted from trajectories within the pattern as shown in Fig. 5(a) and (b). In Fig. 5(c), based on the velocity vectors,

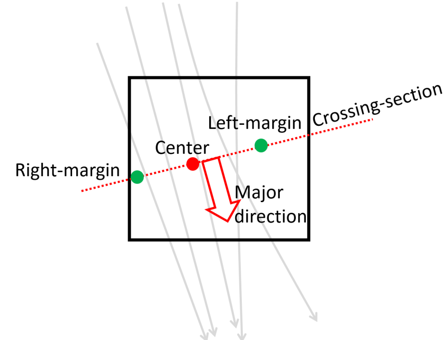

Fig. 4. An example of mobility vector

the traffic route generation module explores the major direction and then retrieves the projected location points on the line perpendicular to major direction. The crossing-section, in terms of the center and both side margins, are computed as red point and green points in Fig. 5(d) according to statistical analysis on projected points. As the result shown in Fig. 5(e), the traffic route is discovered and represented by a sequence of mobility vectors.

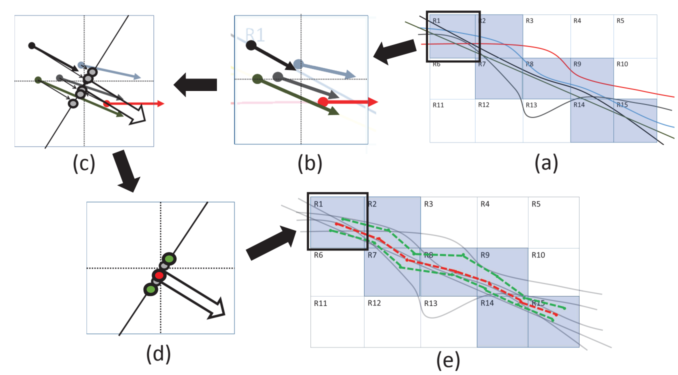

Fig. 5. An running example of maritime traffic route generation

IV. Experiments

The experiments are conducted on a real AIS dataset collected from the AIS network system. The system collects AIS data broadcast by vessels equipped with AIS, including the vessels' unique identification, geo-location, course, speed, and timestamps. An temporal ordered sequence of AIS data can be considered as a AIS trajectory. We extracted a set of AIS trajectory data in a maritime area of 100Km$\times$100Km for five months. The dataset includes 20639 trajectories and 21202212 spatial-temporal points. The proposed framework of Maritime Traffic Route Discovery(abbreviated as MTRD) comprises of three modules: AIS Pattern Mining, Pattern Summarization, and Traffic Route Generation. We first conduct the experiments on the effect of discovery parameter in AIS Pattern Mining and Pattern Summarization. Then, the effectiveness evaluation on the generated traffic route is performed to show the proposed MTRD is able to achive an effective understanding of maritime traffic route. In order to detect the frequent region, we partitioned the area into 5Km$\times$5Km grid-cell size and MinTs is set to be 300 for the frequent region determination.

A. Effect of Discovery Parameters

Based on the collected AIS trajectories in terms of sequences of frequent regions, we discover the AIS trajectory pattern by the frequent sequential pattern mining. We first evaluate the effects of various MinSup on frequent sequential pattern mining. MinSup is a user-defined minimum support threshold for determining a sequence to be a frequent sequential pattern (FSP) in Prefixspan algorithm [11]. Figure 6 shows the results of FSP mining while MinSup is varied from 200 to 400 by Prefixspan algorithm. Meanwhile, the number of SuperFSP generation for pattern summarization is also shown in Fig 6. As can be seen, the experimental results show that the number of FSP and SuperFSP decreased as MinSup increased. Notice that, the proposed SuperFSP pruning method can effectively reduce the number of patterns.

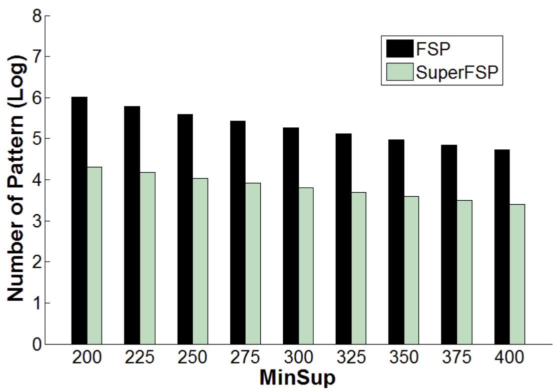

Fig. 6. Effect of MinSup for FSP mining and SuperFSP generation

To evaluate the effectiveness of the proposed SuperFSP pruning, we use the data compression ratio to measure the effectiveness of the proposed method. Data compression ratio is defined as the ratio between the uncompressed size and compressed size and the measure is defined as

$$\text{Data compression ratio} = \frac{UncompressedSize}{CompressedSize}, \quad (2)$$

As shown in Fig. 7, the experimental result shows that the proposed SuperFSP pruning can give high compression ratio.

Moreover, we measure the space savings of data for storage size evaluation on SuperFSP pruning. The data space savings is defined as the reduction in size relative to the uncompressed size and the measure is defined as

$$\text{Data space savings} = 1 - \frac{CompressedSize}{UncompressedSize}, \quad (3)$$

We vary the value of MinSup from 200 to 400 and plot the data space savings after SuperFSP pruning in Fig. 8. The data space savings can be achived over 90 %.

Figure 9 shows the number of summarized pattern generated as the LAF_{min} similarity is varied. LAF_{min} is minimum last-and-first similarity, which is the measurement for pattern concatenation. The result of Fig. 9is exactly what we expect. After the SuperFSP pruning, each FSP should be unique while $LAF_{min} = 1$. In other words, the number of patterns do not reduce as $LAF_{min} = 1$. While LAF decreases, the number of patterns decreases due to the pattern concatenation. The experimental result implies that the patterns those have the sequential relations in movement behavior can be concatenated by the proposed last-and-first pattern concatenation method.

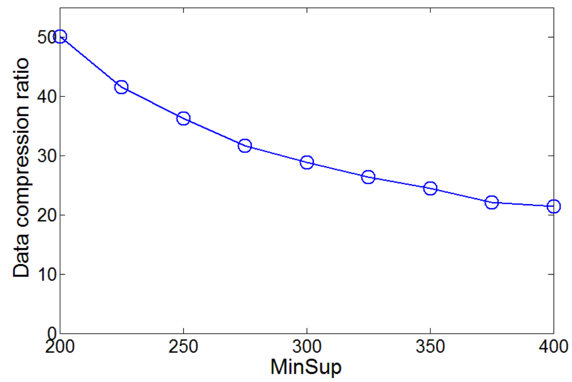

Fig. 7. Data compression evaluation by SuperFSP pruning method

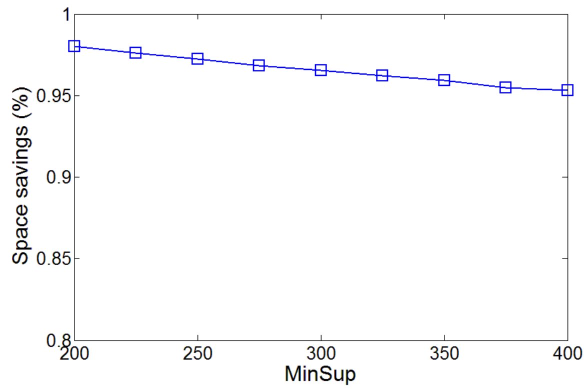

Fig. 8. Space savings evaluation by SuperFSP pruning method

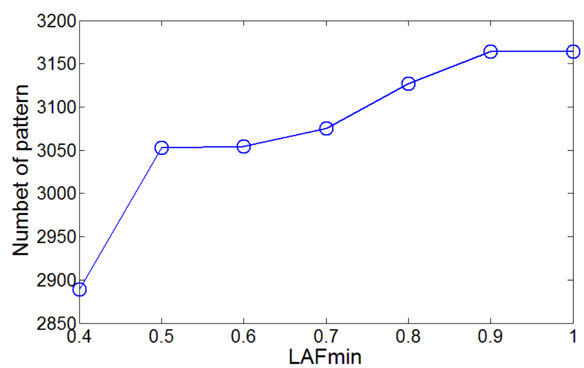

Fig. 9. Effect of LAF similarity on traffic pattern generation

B. Effectiveness Evaluation

In the proposed MTRD, the final objective is to discover the maritime traffic route from trajectory data generated by AIS networks. Thus, we conduct the effectiveness evaluation on traffic route generation. The effectiveness of the proposed MTRD is evaluated by average coverage rate. Given a discovered traffic route, i.e., a pattern-aware route, the average coverage rate of the traffic route discovered from the AIS data is measured by

$$\text{Average Coverage Rate} = \frac{\sum_{k=1}^{n} \frac{P_c(r_k)}{P(r_k)}}{n}, \quad (4)$$

where $P_c(r_k)$ is the number of points contained in the route area contained of the region r_k, $P(r_k)$ is the total number of points contained in the region r_k, and n is the length of region-sequence in the traffic route.

As shown in Fig. 10, the average coverage rate is up to 76% while MinSup is varied from 200 to 400. The result indicates that the proposed MTRD is able to discover the maritime traffic route from AIS trajectory data effectively.

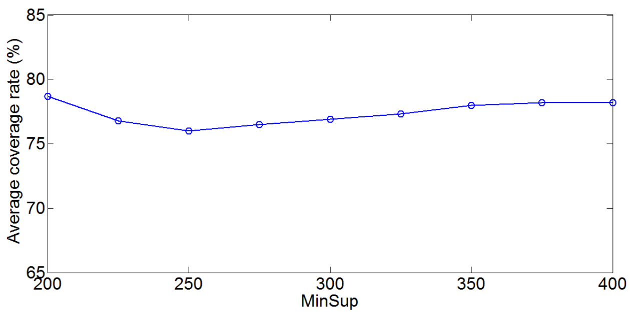

Fig. 10. Effectiveness evaluation of traffic route generation

V. Conclusion and Future Work

This work presents a framework of Maritime Traffic Route Discovery (abbreviated as MTRD) to discovers the maritime traffic route from AIS trajectory data generated by AIS network. We leverage the knowledge of movement pattern discovered from a given set of AIS trajectories to generate the traffic routes. To overcome the problem of data uncertainty and free moving marine space, MTRD includes three modules. AIS Pattern Mining module discovers the trajectory patterns from AIS data in an unsupervised way and then the discovered patterns are summarized by SuperFSP pruning and pattern concatenation in Pattern Summarization module. Finally, the maritime traffic routes are extracted from the discovered patterns by statistical approach using in traffic route generation module. More specifically, a discovered traffic route is represented in form of a movement channel followed by a group of AIS trajectories those having similar behavior within a summarized pattern. The proposed MTRD not only discovers the trajectory pattern but also extracts the movement area of traffic route from those discovered trajectory patterns. Obviously, MTRD provides a better awareness of maritime traffic route discovered from AIS trajectory data. Based on real AIS data, the experimental results show that the proposed MTRD is able to effectively discover the traffic routes from AIS trajectories. In the future, this work will be applied as a cornerstone for researching the problem of traffic management, anomaly detection, and conflict detection in the maritime domain.

Acknowledgment

Po-Ruey Lei was supported in part by the the Ministry of Science and Technology of Taiwan, Project No. MOST-104-2221-E-012-003 and MOST-105-2221-E-012-003.

References

[1] V. F. Arguedas, G. Pallotta, and M. Vespe. Automatic generation of geographical networks for maritime traffic surveillance. In *Information Fusion (FUSION), 2014 17th International Conference on*, pages 1–8, July 2014.

[2] L. Cazzanti, L. M. Millefiori, and G. Arcieri. A document-based data model for large scale computational maritime situational awareness. In *Big Data (Big Data), 2015 IEEE International Conference on*, pages 1350–1356, Oct 2015.

[3] F. Giannotti, M. Nanni, F. Pinelli, and D. Pedreschi. Trajectory pattern mining. In *Proc. of KDD*, pages 330–339, 2007.

[4] H. Jeung, M. L. Yiu, X. Zhou, and C. S. Jensen. Path prediction and predictive range querying in road network databases. *VLDB Journal*, 19(4):585–602, 2010.

[5] J.-G. Lee, J. Han, and K.-Y. Whang. Trajectory clustering: A partition-and-group framework. In *Proc. of SIGMOD*, pages 593–604, 2007.

[6] P. Lei. A framework for anomaly detection in maritime trajectory behavior. *Knowl. Inf. Syst.*, 47(1):189–214, 2016.

[7] Z. Li, J.-G. Lee, X. Li, and J. Han. Incremental clustering for trajectories. In *Proc. of DASFAA*, pages 32–46, 2010.

[8] F. Mazzarella, V. F. Arguedas, and M. Vespe. Knowledge-based vessel position prediction using historical ais data. In *Sensor Data Fusion: Trends, Solutions, Applications (SDF), 2015*, pages 1–6, Oct 2015.

[9] G. Pallotta, M. Vespe, and K. Bryan. Traffic knowledge discovery from AIS data. In *Proceedings of the 16th International Conference on Information Fusion, FUSION 2013, Istanbul, Turkey, July 9-12, 2013*, pages 1996–2003. IEEE, 2013.

[10] G. Pallotta, M. Vespe, and K. Bryan. Vessel pattern knowledge discovery from AIS data: A framework for anomaly detection and route prediction. *Entropy*, 15(6):2218–2245, 2013.

[11] J. Pei, J. Han, B. Mortazavi-Asl, J. Wang, H. Pinto, Q. Chen, U. Dayal, and M.-C. Hsu. Mining sequential patterns by pattern-growth: the prefixspan approach. *IEEE Transactions on Knowledge and Data Engineering*, 16(11):1424–1440, 2004.

[12] P. Silveira, A. Teixeira, and C. G. Soares. Use of ais data to characterise marine traffic patterns and ship collision risk off the coast of portugal. *Journal of Navigation*, 66(06):879–898, 2013.

[13] O.-V. E. Sormunen, F. Goerlandt, J. Häkkinen, A. Posti, M. Hänninen, J. Montewka, K. Ståhlberg, and P. Kujala. Uncertainty in maritime risk analysis: Extended case study on chemical tanker collisions. *Proceedings of the Institution of Mechanical Engineers, Part M: Journal of Engineering for the Maritime Environment*, page 1475090213515640, 2014.

[14] L.-Y. Wei, Y. Zheng, and W.-C. Peng. Constructing popular routes from uncertain trajectories. In *Proc. of KDD*, pages 195–203, 2012.

[15] Y. T. Wen, C. H. Lai, P. R. Lei, and W. C. Peng. Routeminer: Mining ship routes from a massive maritime trajectories. In *2014 IEEE 15th International Conference on Mobile Data Management*, volume 1, pages 353–356, July 2014.

[16] A. Y. Xue, R. Zhang, Y. Zheng, X. Xie, J. Huang, and Z. Xu. Destination prediction by sub-trajectory synthesis and privacy protection against such prediction. In *Proc. of ICDM*, 2013.

[17] Y. Ye, Y. Zheng, Y. Chen, J. Feng, and X. Xie. Mining individual life pattern based on location history. In *Proc. of MDM*, pages 1–10, 2009.

[18] M. Zhou, J. Chen, Q. Ge, X. Huang, and Y. Liu. Ais data based identification of systematic collision risk for maritime intelligent transport system. In *Communications (ICC), 2013 IEEE International Conference on*, pages 6158–6162. IEEE, 2013.