TraSeer: A Visual Analytics Tool for Vessel Movements in the Coastal Areas

Guizhen Wang, Abish Malik, Calvin Yau, Chittayong Surakitbanharn and David S. Ebert
Purdue University, 500 Central Drive, West Lafayette, IN 47907,
Phone: 765-496-3747, Fax: 765-494-1028, Email: {wang1908|amalik|yauc|csurakit|ebertd}@purdue.edu

***Abstract*—Vessel movement analysis can reveal offshore shipment activities, such as the workload of sea lines and abnormal movement paths. Discovery of vessel movement information can assist decision makers in understanding trends and identifying patterns. In our work, we propose a trajectory visual analytics tool, TraSeer, to allow users to interactively analyze vessel movements. TraSeer summarizes vessel movement trends and detects anomalous movement paths. Also, TraSeer can integrate external data sources for coastal security analysis. Finally, we provide case studies to highlight the effectiveness of TraSeer in vessel movement pattern detection and latent illegal activity discovery.**

I. Introduction

Each year thousands of vessels move through U.S. coastal waters, including private recreational boats, fishing vessels, cruise ships, ferries and cargo ships. These vessels provide legal recreation, transportation, commerce, but can also support illicit activities, such as illegal fishing and smuggling. All commercial vessels are equipped with a GPS-based Automatic Identification System (AIS) [1] that automatically reports its status information at regular time intervals, including identification, time, position, and movement information.

An in-depth exploration of AIS data logs can assist decision makers in planning resource allocation based on normal patterns of operation and in discovering unusual vessel activities for further investigation [2]. Therefore, vessel activity surveillance is important for safe and efficient vessel operation. However, the massive amount of data and different vessel paths in AIS data can make efficient analysis difficult and often lead to user overload.

Visual analytics has been widely recognized as an effective tool to explore data. Intuitive visual expression and user-friendly interaction have been shown to help end users reduce the burden of exploring complicated data relations. (e.g. [3], [4])

In this paper, we propose TraSeer, a visual analysis system for interactively exploring vessel movement activities. Through TraSeer, users can find vessel movement trends and anomalous movement paths that deviate from regular trends. Unlike common anomalous trajectory detection algorithms [5], [6] that have a prerequisite to use normal vessel movement data to build a behavior model of normal vessel movements, our method can reveal anomalous paths from a dataset mixed with normal and abnormal ones. Features of TraSeer include:

- **Movement pattern abstraction** summarizes vessel movement patterns through their sources and destinations. From the movement pattern summary, users can identify the workload of vessel activities in coastal areas.
- **Anomalous behavior detection** discovers movement paths that deviate from normal vessel trajectories. Locations and times of anomalous occurrences can reveal potential deficiencies of coastal area surveillance.
- **Interactive exploration workflow** provides users with flexible interactions for exploring vessel movement data.

In addition to the trajectory data analysis, TraSeer also has the capability to integrate other data sources to enable a comprehensive situational awareness.

Section II introduces related work of vessel movement analysis. Section III introduces the system overview of TraSeer. Section IV introduces the method to summarize vessel movement trends in this paper. Section V introduces the proposed anomaly detection method in TraSeer. Section VI describes experiments. Section VII discusses the limits of our proposed anomalous vessel movement detection algorithm and Section VIII concludes the whole paper.

II. Related Work

A. Automatic Identification System

The Automatic Identification System (AIS) [1] is an automatic movement tracking system used on ships to identify their locations, speeds, courses and other information. AIS data can be used for vessel collision avoidance, fishing fleet monitoring and traffic services.

Researchers have demonstrated the benefits of using rich vessel movement information contained in the AIS data by outlining anomalous ship behaviors. For example, researchers have used such data to identify vessel behaviors that are different from regular patterns, such as vessel movement paths that deviate from standard sea lines, irregular AIS data reporting frequencies and so on [6]–[8]. In this paper, TraSeer targets AIS data analysis, summarizing vessel movement trends, detecting anomalies, and providing decision makers a comprehensive interactive analysis environment.

B. Vessel movement analysis

Many research efforts have focused on the pattern abstraction of vessel movements. Lee et al. [9] proposed an algorithm to group trajectories into subgroups based on spatial closeness of movement paths. Moreover, some other researchers [10], [11] proposed density-based algorithms to describe the probabilistic spatial distribution of vessel paths.

978-1-5090-6356-7/17/$31.00 ©2017 IEEE

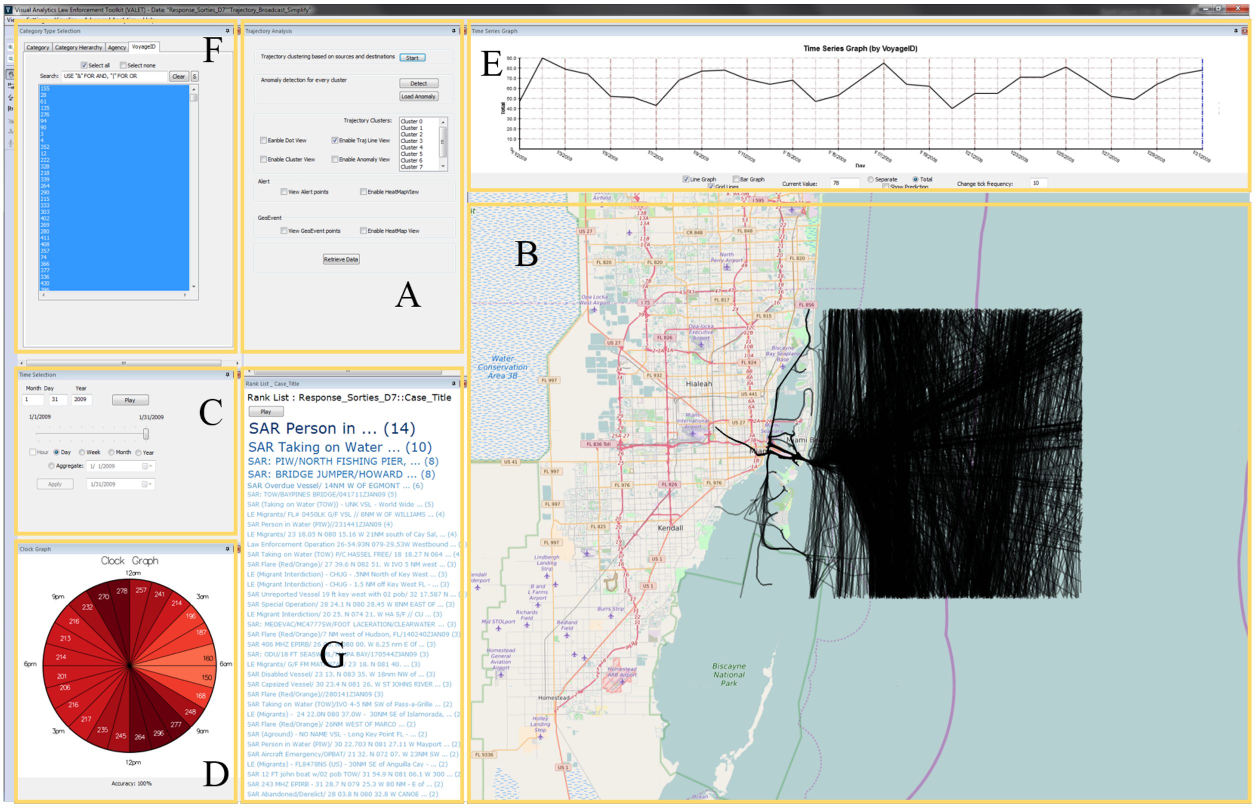

Fig. 1. Interface of the visual analysis system, TraSeer. A) trajectory analysis widget, B) map view, C) time slider, D) clock view, E) time series view, and F) rank list.

Anomaly based trajectory analysis is essential for maritime safety and security missions. In order to support these challenges, researchers have utilized different detection strategies. Anomaly detection algorithms [5], [6] utilize historical AIS records of normal vessel movements to build a behavior model of normal trajectories. Real-time AIS data patterns are compared with the model to identify whether new-coming movements are normal or abnormal. This method requires a training dataset that only contains normal trajectories to build a behavior model. Practically speaking, researchers may not have the prior knowledge to obtain a normal trajectory dataset, especially given that AIS datasets are usually mixtures of normal and abnormal trajectories. Therefore, our system proposes an initial algorithm to detect anomalous movements in a mixed dataset.

Research on interactive visual exploration of vessel trajectories has shown their capabilities to assist decision makers in obtaining comprehensive knowledge about vessel movements. Researchers [10], [11] proposed a heatmap based visualization to show the spatial distribution of vessel movements and use comparative colors to encode statistically significant data value as anomalies. In addition to methods focusing on the statistical information representation, Scheepens et al. [12] use particles to visualize locations of AIS data and design a toolkit to select AIS data through user specified movement directions. Different from visual analysis methods with the main focus on the visual representation, our work concentrates on building an interactive workflow of vessel movement analysis. Trajectory movement pattern abstraction and anomaly detection methods are included in the analysis workflow.

III. System Overview

This section introduces the system design of TraSeer. Through the interface, end users can make use of data analysis models to summarize vessel trends and detect anomalies. Our system also provides multiple views to explore AIS data, as seen in Figure 1. TraSeer includes four main components:

- **Vessel movement data analysis:** The trajectory analysis widget (seen in Figure 1A) provides interactions to summarize the movement trends and detect unusual movement paths. As to the movement trend, the “Trajectory Clusters” list shows clusters of vessels that shares the same origins and destinations.
- **Vessel movement visual representation:** Users can interactively explore trajectory information in the map view (seen in Figure 1B). Each black line shows the movement path of one vessel. Trajectory trending summaries are visualized as groups of lines where colors encode origin-destination pairs. Figure 3 illustrates the movement trend summary around the Miami port in January, 2009. Figure 4 shows the anomaly detection results in one group of trajectories where the darker color encodes anomalous

978-1-5090-6356-7/17/$31.00 ©2017 IEEE

paths that deviate from standard routes. Users can pan, zoom in and out of the map to see trajectories at different spatial granularities.

- **Interaction widgets:** The time slider in Figure 1C allows users to aggregate trajectories at different temporal levels. The clock view in Figure 1D shows the number of vessels sending AIS reports within each hour. The time series graph in Figure 1E shows the number of vessels sending reports over time (e.g. by day, week and month). The widget in Figure 1F allows users to interactively filter the data, using the categorical attributes in the AIS data (e.g. vessel ID and vessel types).
- **Integration with other data sources:** TraSeer can integrate other data source into the system. We show this capability using spatiotemporal based data - the maritime incidents around the coastal area of Florida. TraSeer provides two basic visualizations, the point view and the heatmap view (both seen in Figure 7) to represent the spatial distribution of incident occurrences. Finally, Figure 1G shows a user-selected attribute and a rank list of categories within the attribute in order of decreasing frequency. In our example, "Case_Title" is the user-selected attribute to describe the case type of the maritime incidents in an external data source, Coast Guard search and rescue dataset (SAR) [4], and "SAR Person in Water" is the most frequent type of incidents with fourteen occurrences displayed in the parenthesis.

IV. Vessel movement trend abstraction

Trajectory movement trends can be extracted using different criteria. In our work, we use the origins and destinations of vessels to represent the trends. The trend abstraction process can be divided into two stages: data sampling and data clustering.

The data sampling process selects a subset of AIS data to represent the original large dataset. Sampling is essential to computational performance because every vessel generates data reports at regular temporal periods (e.g. one or two minutes) resulting in large datasets. For example, one vessel may generate over 1,000 reports per day. Therefore, taking all AIS reports into calculation can make the computation a long process. Since every AIS report contains similar information to those within a small temporal vicinity, sampling is a viable method to reduce the computational workload. In this work, we sample the AIS data through a regular temporal interval, (selecting an AIS point every two minutes in the experiments of Section VI).

In the second stage, the data clustering process classifies trajectories into multiple smaller groups with the same origin-destination pairs. Each group represents one movement trend. Every trajectory can be regarded as a sequence of AIS sample points ordered by time. A spatial distance is applied to measure the origin-destination movement proximity of two trajectories. For every two trajectories, their spatial distance is calculated based on information represented by their AIS data sequence, including their movement directions and geospatial distances.

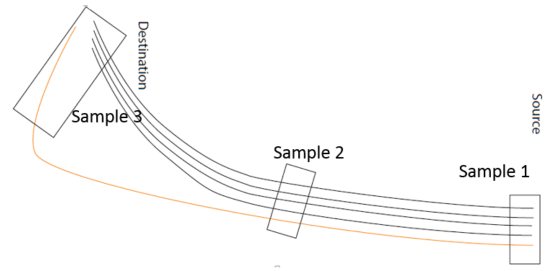

Fig. 2. Illustration of the anomaly detection process. There are four normal trajectories in black and one abnormal trajectory in orange. 3 samples are uniformly picked from every trajectory. In the third set,M_3, the abnormal trajectory can be found since its XY coordinates have a low probabilistic density.

After obtaining the pair-wise distances for all trajectories, the distance based clustering algorithm is adopted to group trajectories. The clustering algorithm detail is introduced by Lee et al. [9].

V. Anomalous vessel movement detection

In this paper, anomalies are defined as vessel paths that deviate from normal trajectories. In this section, we propose a new method to identify trajectories that deviate from normal trends in a dataset mixed with normal and abnormal ones.

Our detection algorithm identifies anomalies by statistically measuring the density distribution of vessel paths. We assume that compared to anomalous paths, normal ones will be dominant in a dataset. Therefore, the vessel movements with high probabilistic density values can be regarded as normal and vice versa. In order to differentiate normal and abnormal paths, we estimate the data distribution and detect unusual paths respectively for every movement trend group obtained in Section IV. Every group has unique movement patterns since its vessels share the same origin and destination and follow the same nautical chart. For example, the pattern of a vessel moving into a port is different from that of a vessel departing the port. To avoid the interference of diverse movement patterns, a density distribution built with data from the same group can improve the distinction of normal and abnormal paths.

In each group, the anomaly detection procedure is divided into two stages. The first stage is to sample AIS data. For each trajectory, N samples are obtained in a uniformed spatial distance. M_i denotes the latitude/longitude coordinate sets of the i-th samples of all the trajectories, where $i = 1, 2, \ldots N$. Therefore, a multi-variate Gaussian distribution is calculated from the latitude/longitude coordinates [13]. The second stage is to find anomalous samples in every set, M_i, from the estimated data distribution. We assume that points of normal trajectories in the same set show similar attribute values since they follow the same movement trend. Therefore, normal AIS points will have higher probabilistic density, and anomalous points will have lower density. Figure 2 illustrates the anomaly detection in one set with $N = 3$. Therefore, paths between two anomalous points will be regarded as anomalous.

VI. Experiments

TraSeer is implemented based on the spatiotemporal visual analytics system, cgSARVA [4]. We present three usage scenarios, showcasing our system for the process of vessel movement analyses. The first scenario is a case study that analyzed vessel movement behaviors in Miami, Florida. The second scenario is a case study that integrated an external maritime incident dataset into TraSeer. The third experiment is a case study that analyzed vessel movement behaviors in Everett, Washington.

A. Case Study

The first demonstrative dataset is the AIS records of more than 1,000 vessels around Miami port from January 1st to January 31st, 2009, provided by the Bureau of Ocean Energy Management and the National Oceanic and Atmospheric Administration [14]. Figure 1 shows an overview of the entire dataset. The map view in Figure 1B shows the movement paths of vessels. The clock view in Figure 1D shows the distribution of active vessels per hour, which reveals there were more active vessels around noon and midnight. The time series graph in Figure 1E shows the daily active vessels around Miami port. Figure 3 shows the top five movement trends with the largest

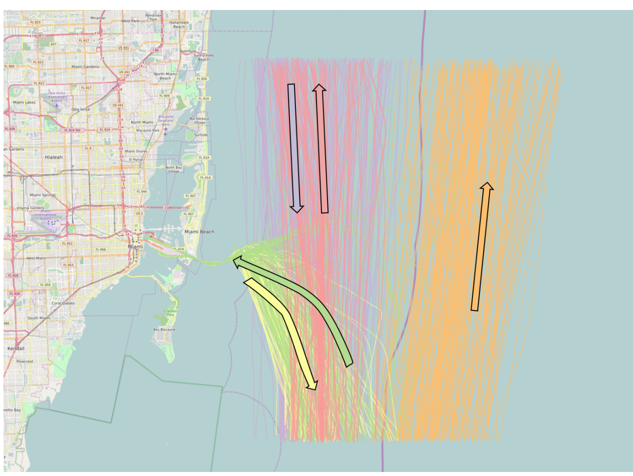

Fig. 3. The top five movement trend abstraction of vessels that moved around the Miami port in January, 2009. Each color encodes one movement trend. Colored block arrows illustrate the movement trend for each corresponding colored vessel paths.

volume of vessels based on their origin-destination pairs. From these clusters, we find that vessels in the green group went into the Miami port from the Southeast; vessels in the yellow group departed from the Miami port; vessels in the purple group went by the port and moved toward the south; and vessels in both the orange group and in the red group went by the port and moved toward the north.

Figure 4 shows the anomaly detection results of one group of vessels. Here, light green trajectories encode the normal paths, and dark green trajectories encode anomalous paths. We find that these vessels were trying to move into the port. Vessels that have anomalous behaviors would go on a path

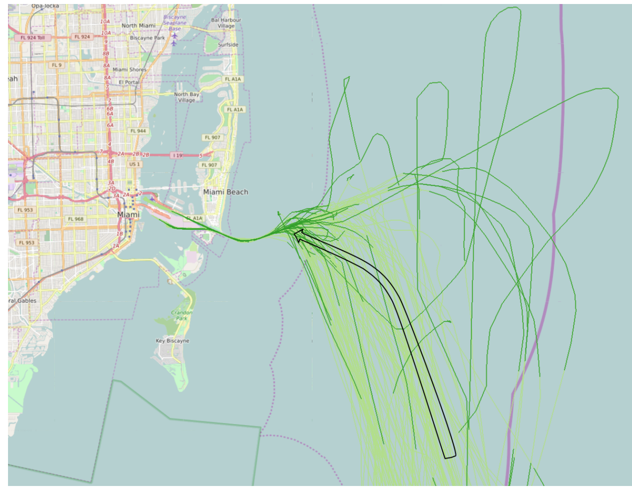

Fig. 4. Anomaly detection results of vessels that tried to move into the Miami port from the Southeast in January, 2009. The block arrow illustrates the normal vessel paths. Darker green encodes anomalous paths.

that is quite different from the normal paths. Unusual paths detected in Figure 4 may assist decision makers to explore potential illegal activities around the coastal area in Miami.

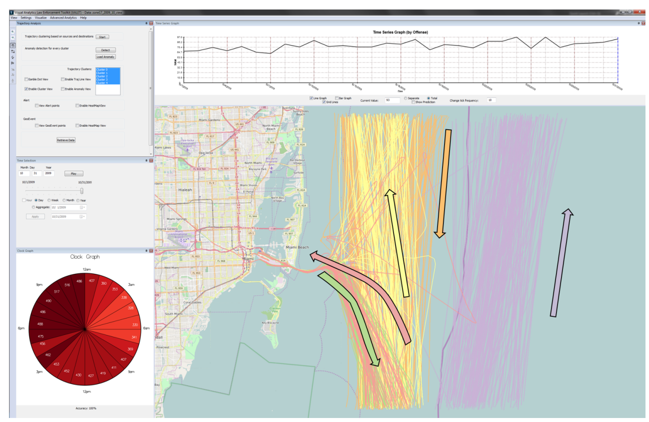

Fig. 5. The top five movement trend abstraction of vessels that moved around the Miami port in October, 2009. Each color encodes one movement trend. Colored block arrows illustrate the movement trend for each corresponding colored vessel paths.

The second demonstrative dataset contains the AIS records of more than 1,000 vessels in the same region, but for a different portion of the year: from October 1st to October 31st, 2009 [14]. Comparing the clock views in Figure 1 and Figure 5, we can see that the hourly distribution of active vessels was different in January, 2009 tan in October, 2009. In January, both the morning and the night had the largest number of active vessels. In October, the hours from 3PM to 12AM had the largest number of vessels. Figure 5 shows the top five movement trends having the largest vessel volume based on the same measurements in Figure 3. We can see that the top five movement trends were the same in both January and October. Figure 6 shows the anomalous paths in the movement trend of vessels going into the Miami port in October, 2009. We can

978-1-5090-6356-7/17/$31.00 ©2017 IEEE

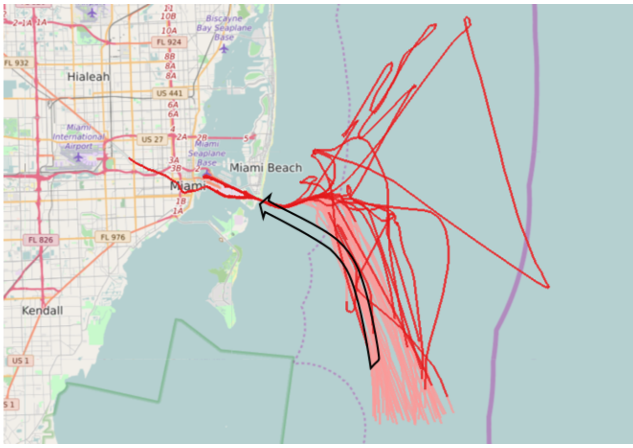

Fig. 6. Anomaly detection results of vessels that tried to move into the Miami port from the Southeast in October, 2009. The block arrow illustrates the normal vessel paths. Darker red encodes anomalous paths.

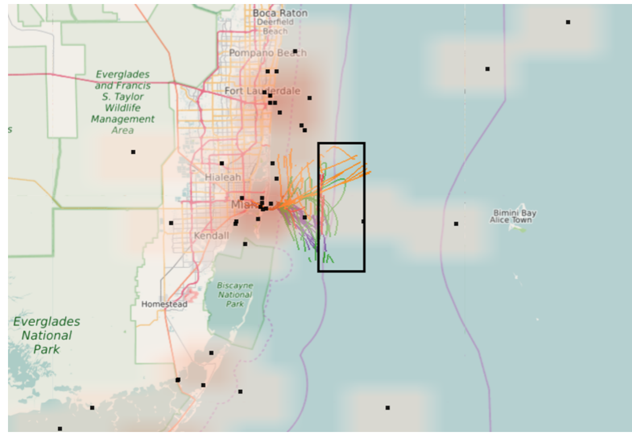

Fig. 7. Integration of maritime safety incidents shown both as a heatmap and points showing specific incident locations.

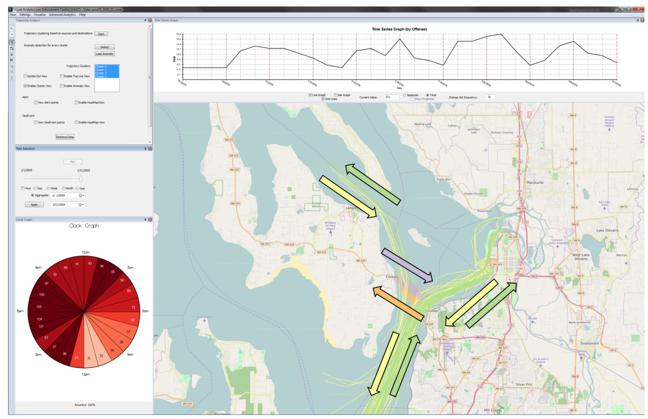

Fig. 8. Overview of the AIS data around Everett, WA from January 1st to January 31st, 2009 in TraSeer. The map view shows vessel movement trend abstraction . Each color encodes one movement trend. Colored block arrows illustrate the movement trend for each corresponding colored vessel paths.

find that the anomalous paths in October were quite similar to that in January (seen in Figure 4). The anomaly in common is that several vessels detoured toward the north before going into the Miami port, which may reveal an unusual phenomenon in the Miami port related with the maritime safety surveillance.

B. Integration with external data sources

In order to show that TraSeer has the capability to integrate multiple coastal security related data sources, we used the Coast Guard search and rescue dataset (SAR) [4] of maritime incident reports in the Atlantic area. Vessel movement data is the same as Section VI-A. Both AIS data and SAR data were selected from January 1st to January 31st, 2009. In Figure 7, black points and the heatmap show the spatial distribution of incidents, and colored lines are detected vessel movement anomalies around the Miami port. Both the two data provide information to assist maritime risk assessment. However, the SAR data identifies regions with high incident occurrences. Also, anomalies detected from the AIS data show potential suspicious movement paths of vessels. The combination of AIS data and SAR data can help evaluate a more comprehensive situational awareness program for coastal area safety and security.

C. Case Study 2

The third experiment uses the AIS records of nearly 698 vessels around Everett, Washington from January 1st to January 31st, 2009 [14].

Figure 8 shows the vessel movement trend abstraction of AIS data collected on from January 1st to January 31st, 2009. The clock view shows the hourly distribution of the number of vessels in Everett, Washington. The time series view shows the daily distribution of the number of vessels. Our system identified four movement trends in this dataset. Paths in yellow show the vessels moving from the Northeast toward the Southwest; paths in green show the vessels moving from the Southwest toward the Northeast; paths in purple shows the vessels moving along the ferry route from the Clinton terminal to the Mukilteo terminal and paths in orange show the inverse ferry route. Figure 9 illustrates the detected anomaly in the movement trend from the Southwest to the Northeast. There are two types of unusual paths, marked in darker green. The two unusual paths (annotated with 1) deviated from the normal paths that moved closer to the island. The other six unusual paths (annotated with 2) also deviated from the normal paths closer to the island in the bottom right. Vessels moving from the Southwest bifurcated at the region in orange and the majority of them, almost 70 vessels, continued to move along the island in the right bottom. However, a smaller group of six vessels moved forward along the island on the left side, deviating from the other 70 paths.

VII. Discussion

Section V outlines initial efforts to address anomaly detection in the AIS data. The accuracy of our method is affected by three factors. The first factor is the data size. In essence, the detection algorithm in Section V estimates probabilistic

978-1-5090-6356-7/17/$31.00 ©2017 IEEE

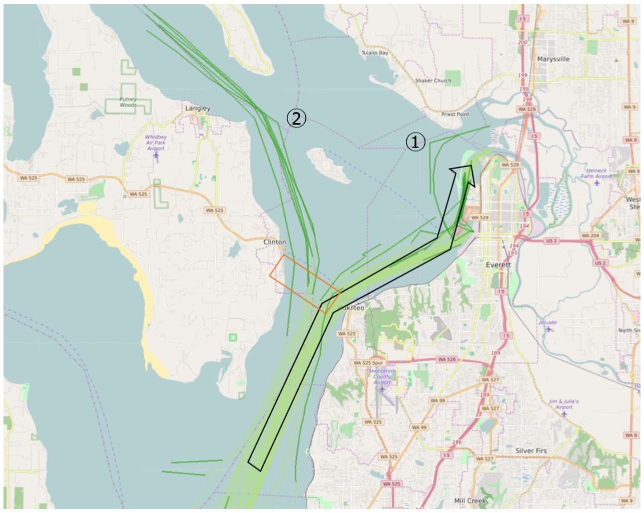

Fig. 9. Anomaly detection results of vessels that moved from the Southwest toward the Northeast. The block arrow illustrates the normal vessel paths. Brown encodes detected anomalous paths.

density distributions of trajectories. The larger the number of vessel movement trajectories, the better the estimation result. The case study in Section VI-A uses an AIS dataset of more than 1,000 vessels to estimate the density distribution of various movement behaviors. But a dataset with fewer trajectories, (e.g. a dataset with only 100 vessel movement paths), lacks adequate information to capture the practical distribution of normal vessel movement patterns. The second factor is the number of normal trajectories versus abnormal trajectories. The detection algorithm has a prerequisite condition that normal trajectories are dominant in the dataset compared to abnormal ones. Only if a pattern had a lower density in the dataset, will one path be regarded as anomalous. Therefore, our proposed algorithm cannot be guaranteed to be accurate in a dataset that violates the prerequisite condition. The last factor is the trajectory data sampling method. Currently, we uniformly sample locations of vessels along their movement paths. Diverse starting and ending locations of vessels will enlarge dissimilarities of selected samples, and further disturb normal movement trend estimation in a smaller dataset.

Although the capability of our anomaly detection method has been demonstrated in Section VI, some more robust methods [15], [16] using methodologies, such as the Bayes rule and neural networks, have been proposed to detect anomalies in a dataset mixed with normal and abnormal vessel movement behaviors. Our method is simpler and more efficient but not as accurate in detecting and classifying anomalies. Compared to their computation complexities, the advantage of our method is simplicity and efficiency. In the future, we will explore integrating components of these algorithms to improve the robustness of our detection method.

VIII. Conclusion and Future Work

TraSeer is a visual analytics tool for AIS data based vessel movement analysis. This tool combines visualization, data analysis and interactive exploration to help users identify potential maritime security issues. Experiments show its capability to summarize movement trends and identify latent trajectory anomalies, as well as integrate other maritime data sources. In the future, we will improve our system to provide end users with a sophisticated visual analytics environment. One possible direction will be to integrate the spatial granularity into the vessel movement analysis workflow to provide data analysis based on the granularity specified by end users (e.g. the national level or the city level).

Acknowledgment

This work was funded by the U.S. Department of Homeland Security VACCINE Center under Award Number 2009-ST-061-CI0007.

References

[1] "Automatic identification system overview," http://www.navcen.uscg.gov/?pageName=AISmain.

[2] R. Laxhammar, "Anomaly detection in trajectory data for surveillance applications," Ph.D. dissertation, Örebro University, 2011.

[3] J. T. Stasko, C. Görg, and Z. Liu, "Jigsaw: supporting investigative analysis through interactive visualization," *Information Visualization*, vol. 7, no. 2, pp. 118–132, 2008.

[4] A. Malik, R. Maciejewski, B. Maule, and D. S. Ebert, "A visual analytics process for maritime resource allocation and risk assessment," in *IEEE Conference on Visual Analytics Science and Technology*, Oct 2011, pp. 221–230.

[5] R. Laxhammar, G. Falkman, and E. Sviestins, "Anomaly detection in sea traffic-a comparison of the gaussian mixture model and the kernel density estimator," in *12th International Conference on Information Fusion*. IEEE, 2009, pp. 756–763.

[6] G. Pallotta, M. Vespe, and K. Bryan, "Vessel pattern knowledge discovery from ais data: A framework for anomaly detection and route prediction." *Entropy*, vol. 15, no. 6, pp. 2218–2245, 2013.

[7] R. O. Lane, D. A. Nevell, S. D. Hayward, and T. W. Beaney, "Maritime anomaly detection and threat assessment," in *13th Conference on Information Fusion (FUSION)*. IEEE, 2010, pp. 1–8.

[8] B. Ristic, B. La Scala, M. Morelande, and N. Gordon, "Statistical analysis of motion patterns in ais data: Anomaly detection and motion prediction," in *11th International Conference on Information Fusion*. IEEE, 2008, pp. 1–7.

[9] J.-G. Lee, J. Han, and K.-Y. Whang, "Trajectory clustering: A partition-and-group framework," in *Proceedings of the ACM SIGMOD International Conference on Management of Data*. ACM, 2007, pp. 593–604.

[10] O. D. Lampe and H. Hauser, "Interactive visualization of streaming data with kernel density estimation," in *IEEE Pacific Visualization Symposium*. IEEE, 2011, pp. 171–178.

[11] R. Scheepens, N. Willems, H. van de Wetering, G. Andrienko, N. Andrienko, and J. J. van Wijk, "Composite density maps for multivariate trajectories," *IEEE Transactions on Visualization and Computer Graphics*, vol. 17, no. 12, pp. 2518–2527, 2011.

[12] R. Scheepens, C. Hurter, H. van de Wetering, and J. J. van Wijk, "Visualization, selection, and analysis of traffic flows," *IEEE Transactions on Visualization and Computer Graphics*, vol. 22, no. 1, pp. 379–388, 2016.

[13] P. Filzmoser, R. G. Garrett, and C. Reimann, "Multivariate outlier detection in exploration geochemistry," *Computers & Geosciences*, vol. 31, no. 5, pp. 579 – 587, 2005.

[14] "Marinecadastre.gov," http://marinecadastre.gov/ais/.

[15] W. Hu, X. Xiao, Z. Fu, D. Xie, T. Tan, and S. Maybank, "A system for learning statistical motion patterns," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 28, no. 9, pp. 1450–1464, Sept 2006.

[16] D. Garagic, B. J. Rhodes, N. A. Bomberger, and M. Zandipour, "Adaptive mixture-based neural network approach for higher-level fusion and automated behavior monitoring," in *IEEE International Conference on Communications*, June 2009, pp. 1–6.

978-1-5090-6356-7/17/$31.00 ©2017 IEEE