Short Term Scientific Mission

Coordination Action COST IC0903 MOVE

FP7 – FET – Open Coordination Action n. 245410.

Report of activities

Fabio Pinelli

October, 2010

1 Introduction

Maritime transportation has a huge impact on the world economy and our everyday lives. Nowadays, the vast majority of world goods is carried by sea by huge container-carriers. The disasters and damages caused in the event of major sea collisions can pose serious threats to the environment and human lives. This stresses the urgent need for the development of maritime navigation systems whose objective will be to contribute to a safer sea. The large availability of maritime positioning data arising in real-time from different heterogeneous sources (sensors, database, Web) with different forms and details opened new opportunities for maritime information systems. Existing works in maritime transportation have been first focusing on tracking systems and simulation techniques in order to provide decision-aided systems. This has been often addressed without taking into account all the benefits such data can provide when analysed appropriately. Amongst many technical and methodological issues to address, there is still a need for spatio-temporal data mining techniques that will favour, from mobile objects raw data, added-value information, understanding and suitable results to maritime control centers as well as marine officers onboard.

Nowadays, in the extremely complex social system of metropolitan areas, the observation of the movements patterns and behavioural models of people is needed for traffic engineers and city managers. Such information are really important for urban planners as well. These analysis become more detailed and precise thank to the availability of new data sources that measure the movements of people more accurately and new analysis techniques, as data mining proposes, that allow the extraction of implicit and hidden knowledge making it available and directly usable.

Therefore, the knowledge provided by the two labs can provide interesting questions which the maritime domain still needs. In this perspective, we collaborate first on the acquisition of sufficient knowledge on the maritime data gathering system infrastructure (Sec. 2); then we applied the existing tools and algorithms for spatio–temporal analysis (Sec. 3). Finally, we plan future works that can provide further answers to the maritime context, (Sec. 4).

2 Maritime data

The availability of maritime data is much larger than in an urban scenario, indeed several different techniques are adopted in order to track ships in the sea, on shore and off shore, and several authorities are disposed in order to gather such data. Different tracking techniques correspond to a different radius of possible monitoring. Satellite ways allow a world wide monitoring of open sea ships. Instead, other techniques are used on on shore environment for allowing harbor authorities to monitor the behavior of ships close to the coast area. The data has been collected using the AIS system: the Automatic Identification System (AIS) is an automated tracking system used on ships and by Vessel Traffic Services (VTS) for identifying and locating Vessels by electronically exchanging data with other nearby ships and VTS stations. An AIS generally integrates a transceiver system, a GPS receiver and other navigational sensors on board such as a gyrocompass and a rate of turn indicator. An AIS transponder runs in an autonomous and continuous mode, regardless its location (e.g., open sea, coastal or inland areas). The transmitted information is broadcasted to a range of 35 miles through VHF communications, to surrounding ships and to maritime Vessel Traffic Systems (i.e., maritime and port authorities).

Several information are carried by means of this technology. Beyond geographic position, time and a unique identification number, the locations are completed by descriptive and meta-data providing either quantitative or qualitative information on the ship trajectory. It can include data on the ship characteristics, route, speed, harbour the ship came from and the destination of the trajectory, material transported, and estimated arrival time. Such trajectories should be ideally tracked and integrated in automated systems that provide the whole range of functionalities to the users, either at the local (i.e., vessel) or global levels (maritime authorities). The technological challenge is multiple: real-time integration of maritime trajectories (a somehow sensor-based issue), integration of these trajectories into distributed databases, mining and knowledge discovery of location data, and development of decision-aided systems.

The dataset covers the Brest bay area in France. The available database contains partial data since October, 2006. The time span covered by the selected subset is almost one year, since 5th February 2009 to 12th December of the same year. It contains 5,756,438 distinguished positions coming from 824 distinct ships.

This is a great opportunity to consider another type of data to analyse which includes both static and kinematic information. This guarantees another strong effort to my scientific background favoured by the MOVE project.

3 M-atlas at sea

M-Atlas is a stand alone application that provides several techniques in order to perform different spatio-temporal analysis on the top of data describing moving object trajectories. During the visiting period, we perform several analysis by means of M-atlas in order to get more useful information about the mobility behaviours at sea. We conduct two parallel analysis one considering different dataset for each month, and the other takes into account the whole dataset covering almost one year. The goal of this multiple analysis is to highlight common and uncommon behaviours considering a different time granularity.

Initially, we use the temporal analysis furnished as a simple tool in M-Atlas. We conduct this analysis in order to individuate if there is a match between the annual and the months temporal distribution. As aspected, a global perspective furnishes results that in certain periods present small deviations and differences, but in general the temporal behaviours of ships in the covered area are quite similar. The further step of our investigation regards a validation test of algorithms and methodologies included in M-Atlas. Indeed, we want to discover well know group of trajectories, such as the commuter ships that bring people from Brest port to work and vice-versa. Therefore, we proceed to extract clusters of trajectories that follow a similar route, we execute the T-clustering algorithm included in M-Atlas using the route similarity distance. Moreover, we continue the analysis looking at the temporal distribution of starting and arrival points comparing them with the arrival and departure timetables provided by the company. The results of such analysis are very good in most of the cases, considering both monthly and yearly datasets, and also the results are qualitatively comparable in both cases.

The execution of the T-clustering algorithm highlights interesting results in terms of regularity of spatial patterns, for instance the clusters performing commuter trips, but also by means of a monthly analysis we are able to detect some interesting new cluster that appear only in certain months. For instance, we recognize groups of cargo vessels in a different waiting areas in the Brest Bay, as well as we discover cluster of trajectories with the same ship type performing almost the same path. Moreover, we discover areas where rescue&save ships wait to execute their operations in difficult days. All these results are obtained considering each month as a separated dataset. While we are considering the complete dataset, the tool presents more difficulties to detect behaviours that deviates significantly from the typical ways in such sea area.

During this deep analysis of maritime data, we have the chance to provide a serious test of M-Atlas and its tools, and thus we are able to point out positive and negative remarks. In our experiments we test the performances of different tools and they show very good results in terms of execution time considering the huge volume of data when we analyse the year dataset. The main drawback we found is the building trajectory procedure, where we can use some temporal and spatial thresholds, but sometime they cannot be the same for different type of ships. This is due to the fact that M-Atlas has been developed for working mainly on urban environment whit only one type of vehicle.

4 Conclusion and future works

The results we obtain we the previous analysis allow us to plan some future works for enhancing specific tools for mobility analysis at sea. A first improvement of the existing tools is the integration of static data in the analysis. The data coming from urban domain do not contain any information about the driver, the car type and so on, instead the AIS data we analyse include this information and it is very useful for improve the results we can get with existing tools.

The integration with static information consent us to create a specific algorithm for trajectory classification based on clustering results. Each ship type presents distinguishable spatio-temporal behaviours in itself (e.g. the passenger ships perform different line covering a different area and in different moments). Thus, the application of a cluster algorithm can not produce high quality results if it does not take into account this information. Indeed, our idea is to consider trajectories with the same ship type, after that we cluster them in order to get their regularities applying the T-clustering algorithm, and finally we generate the set of classes we want to use in our classifier. The classification process is improved with the usage of a network with the nodes represent significant regions in the sea area, such as harbors, waiting areas, and crossing areas.

Another open research issue that the collaboration between the two laboratories allows a deeper investigation is the detection of anomalies and outlier behaviors within the discover of dangerous behaviours. The results provided with the T-Clustering algorithm can still provide a starting point for this research exploration. Indeed, we consider the outliers as ships which strongly deviate from the typical behaviours, and we can highlight them by means of T-Clustering algorithm.

We conclude saying that the visiting period allows a first meaningful analysis of maritime data using advanced spatio-temporal data mining techniques, and moreover it opens interesting research scenario's for further and fruitful scientific collaboration for the development of specific data mining techniques to be applied in a maritime domain. In fact, specific tools available now in M-Atlas are dedicated to an urban context, some other specific tools are necessary to perform a deeper analysis of vessel trajectories because the data are different as well as the needs of maritime authorities.

The exchange of knowledge and experiences between two laboratories gives us the opportunity to improve our relative scientific backgrounds, and this is promoted by the MOVE project.