The geographical decision-making chain: formalization and application to maritime risk analysis

Bilal Idiri, Aldo Napoli

The geographical decision-making chain: formalization and application to maritime risk analysis

Bilal Idiri[1], Aldo Napoli[1]

[1] MINES ParisTech, Centre for Research on Risk and Crisis, CS 10207, 06904 Sophia Antipolis Cedex, France

{Bilal.Idiri, Aldo.Napoli}@mines-paristech.fr

Abstract. Maritime traffic monitoring needs tools for spatiotemporal decision support. The operators responsible (e.g. the Coast Guard) must monitor vessels that are represented as objects moving in space and time. Operators use maritime tracking systems to follow the evolution of traffic and make decisions about the risks of a situation. These systems are based on Geographic Information Systems (GIS) and OnLine Transaction Processing (OLTP) approaches, which are prohibitively expensive, very slow and produce operational data unsuited to decision-making. Instead, operators require summarized data that is easier for them to produce and use. Therefore, we propose the definition of a geographical decision-making chain that adds a decision-making dimension to current systems. It consists of a carefully assembled set of tools that can automate the three phases of Business Intelligence, namely data loading, modelling and analysis.

Keywords: Spatiotemporal decision aid, geographical decision-making chain, maritime tracking systems, risk analysis.

1 Introduction

The development of new sensor technologies and the introduction of mandatory requirements for ships to be equipped with transponders have together contributed to the generation of huge volumes of geo-referenced maritime data. This data can be used to improve decision-making in the maritime environment; there are already operational surveillance systems that enable authorities to track vessel movements in near real-time on a display device (screen, touch table, tablet, etc.) and interpret risks on the basis of their kinematics. Two approaches to improving these monitoring systems can be identified in the literature. The first is a post-hoc event analysis that seeks to understand what happened [1], [2] and the second is the real-time identification of abnormal ship behaviour [3], [4], [5], [6].

However, although maritime authorities have expressed their need for decision support systems [7], few studies have focused on the design of these systems [7], [8]. Such systems consist of a set of tools and methods that reduce the time needed to gather, consolidate, model and return data in order to provide a synthesised overview of all maritime activity and thereby help decision-makers to take informed decisions.

In this paper we propose a new paradigm – geo-decision-making – and define the elements of the geo-decision-making chain. This chain provides support for all decision-making functions (from the collection of data from multiple sources to its presentation to end-users) and therefore improves maritime risk analysis. Geo-decision-making adds a mapping dimension to analysis and indicators and therefore enriches decision support systems with a geographic component.

2 Maritime surveillance systems

Maritime surveillance systems consist of a data acquisition infrastructure that captures and transmits ship data (position, course, speed, home port, etc.) and an information processing system that processes, stores and returns information via display devices (see Fig. 1).

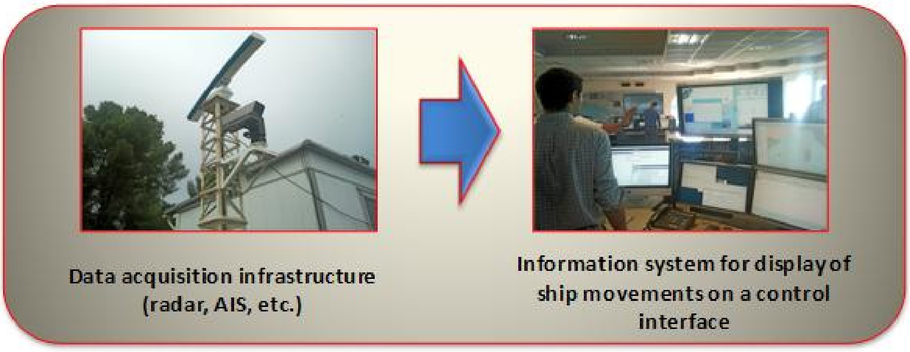

Fig. 1. Demonstration of a new-generation maritime surveillance system implemented in the Maritime Rescue Coordination Centre in France by the I2C project [4].

Many previous studies have addressed the issue of improving maritime surveillance. As far as the data acquisition infrastructure is concerned new, smart sensors are available (heterogeneous sensors, aerial drones, sonar networks, etc.). At the information processing level, two approaches can be distinguished. The first is a risk analysis approach based on probabilistic methods ([9] in [10]), statistics ([11] in[10]) and modelling and digital simulation ([12] in[10]). The second is the real-time identification of maritime risk, which can itself be divided into two approaches. The first is based on the identification of risk through knowledge modelling [5], [13], [14], which necessitates knowledge discovery and formalization. The second is based on the identification of risks through visualization [15], [16]. In the latter approach, information is presented in such a way as to be able to extract meaning directly through visualization. Figure 2 summarises the two approaches.

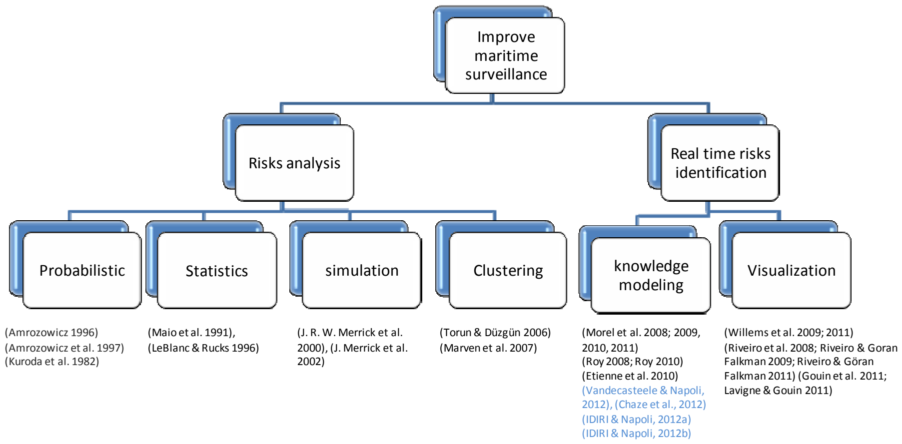

Fig. 2. A summary of current approaches to improving maritime surveillance at the information processing level

3 Approaches to decision support

There are potentially many ways to improve decision support in the maritime domain. For example, approaches have been based on advanced spatial analysis [7], knowledge representation [17] and automatic risk identification [5]. Here, we propose an innovative approach based on knowledge extraction using automatic and semi-automatic data mining techniques. Automatic data mining enables data to be explored for the purpose of knowledge discovery, while semi-automatic data mining gives users the power to interact with the data in order to extract meaning (relationships, models, etc.) through a visual exploration. This user participation in knowledge modelling is particularly interesting as the actors involved are most familiar with the area of application.

To achieve this integration of data mining into maritime surveillance systems, we formalized the use of geo-decision-making tools in the form of a chain (using the analogy of the decision-making chain). This chain structure can support all decision-making functions, from the collection of data from multiple sources to its presentation to end-users. Gouarné [18] describes the following four decision support functions:

Collection: This function supplies the decision support system with data from multiple sources (sensors, operational systems, intelligence, etc.).

Integration: This function ensures data consistency. It provides a unified and normalized data storage model that offers all users the same overview of the data. **Broadcast**: This function distributes data used for decision-making to the various exploration and visualization applications.

Presentation: This function enables the presentation of input information and makes it possible to manage access to the information disseminated.

The aim of our work is not to replace operational maritime surveillance systems with decision support systems. Both systems must evolve in parallel to ensure complementary functionality. Table 1 (below) summarises the contribution of decision support systems to operational systems according to various criteria, namely data, interfaces and queries.

Table 1. Comparison of operational systems and decision support systems

	Operational	**Decision support**
Data	Immediate Detailed Internal to the system Normalised	Historical Aggregated Multiple sources De-normalised
Interface	Complex	Intuitive
Queries	Predefined user queries Slow response to aggregated queries Frequent updates.	Open-ended Rapid response to aggregated queries No updates.

4 Formalization and support of the geo-decision-making chain in the maritime domain

The concept of the decision-making chain is widely used in the field of Business Intelligence. It can be defined as the set of tools that lead to the production of actionable information. For example the software vendor SAS[1], which has been specialized in the field since 1976 uses the term "data production line" or the "decision-making chain". Gouarné [18] uses the term "data provision chain" and Sandoval [19] uses the term "supply chain information". However, we have not been able to find equivalent concepts in the realm of geo-decision-making despite the existence of specific tools for the processing of geo-decision-making information.

[1] http://www.sas.com/

4.1 Formalization of the geo-decision-making chain

Using the analogy of the decision-making chain, we define the geo-decision-making chain as the set of tools that form the processing chain for geographic information, from data collection to its presentation to decision-makers. Specifically, it consists of a collection of tools that are carefully structured in order to ensure the four decision support functions previously discussed. These tools are shown in Figure 3 and include: Spatial Extract, Transform and Load (Spatial ETL), Spatial Data Warehouse (SDW), Spatial OnLine Analytical Processing (SOLAP), Spatial Data Mining (SDM) and the DashBoard (DB).

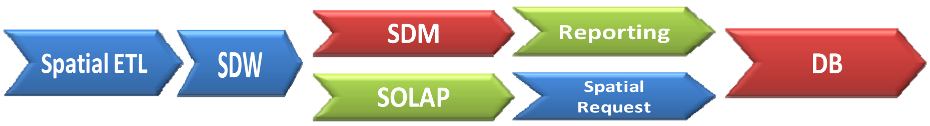

Fig. 3. The tools used in the geo-decision-making chain

The construction of this chain presents a major challenge in terms of the interoperability of the various constituent tools. Careful thought is required to arrive at a system where each individual tool fulfils a role that supports the global function, namely the generation of geo-decision-making information. The links that make up the geo-decision-making chain are described in more detail below:

Spatial Extract, Transform and Load tools make it possible to extract geographic data from heterogeneous sources (a database, flat files, business applications, etc.) and to transform and load it into the Spatial Data Warehouse (SDW). These tools can be re-run each time data sources are updated, or on demand for the production of actionable information.

The **Spatial Data Warehouse** is a non-volatile spatial database. Each entry into the database is time stamped in order to follow its evolution over time. The database is the sole repository for all data used by decision support tools.

Spatial Data Mining is the non-trivial extraction of implicit and potentially useful knowledge from data supplied by the spatial database [20].

The **Spatial OLAP** has been defined by Bédard as, "a visual platform specifically designed to support a rapid and efficient spatiotemporal analysis through a multidimensional approach that includes mapping, graphical and tabular levels of aggregation" [21].

Reporting groups together tools for the automatic preparation of decision support reports or output states based on the data stored in the data warehouse or post-analysis results.

Spatial request are intuitive interfaces for making queries.

The **Dashboard** enables activities to be controlled through the monitoring of key indicators.

4.2 The post-hoc analysis of maritime risks

These decision support tools that include a geographic component make it possible to establish the state of the maritime traffic situation. The system integrates data from multiple sources related to the same situation with the aim of understanding the causality between events. It automates the processing and preparation of raw data before its analysis. The data repository (or SDW) that records the evolution of maritime activity serves as a tool for automatic (SDM) data mining and semi-automatic (SOLAP) data mining that leads to knowledge discovery.

The analysis of maritime phenomena requires the ability to represent information cartographically. For example, the left-hand side (Part 1) of Figure 4 shows a graphical and tabular representation of the number of vessel accidents according to maritime zone. This graphical representation suggests that accidents are uniformly distributed across zones. However, a cartographic display of the accident distribution reveals a positive correlation with a median tanker trajectory (Part 2, Figure 4).

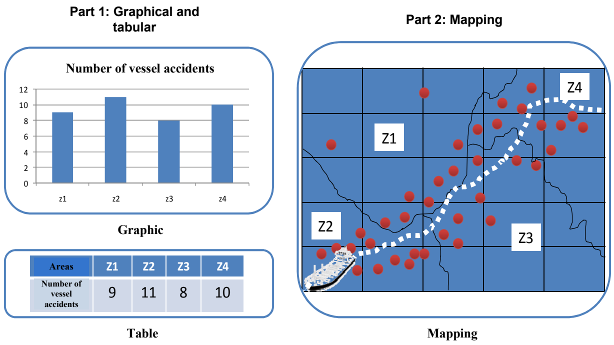

Fig. 4. Example of accident distribution according to maritime zone demonstrating the usefulness of the SOLAP analysis

4.3 Real-time identification of maritime risk

Geo-decision-making is not suited to the real-time identification of marine risk. On the other hand, it makes it possible to model knowledge that can be used to improve risk identification. The two data mining components (SDM and SOLAP) make it possible to develop knowledge models that lead to a better understanding and perhaps anticipation of risks.

5 Application to the maritime domain

Our work has focussed on the application of one component of the geo-decision-making chain (the SDM) to the maritime domain. The application of data mining techniques to historical maritime accident data made it possible to automatically generate knowledge. For example, the application of association rules[2] enabled the generation of knowledge in the form of implications such as, "If *Incident_Type = Capsize/Listing* then *Location_Of_Accident = Coastal-waters*; *confidence*[3] *= 77%*" [22]. This rule means that if there is a capsize-type accident then, in 77% of cases it will be located in coastal areas. In addition, the application of spatial clustering[4] techniques made it possible to model zones with a high density of accidents. The areas identified in this way can be used as zone patterns that enable maritime authorities to focus on specific areas and to identify the vessels that frequent them, in order to better target surveillance activities.

6 Conclusion

Our aim is to meet the need for decision support in the domain of maritime surveillance. Therefore, this paper proposes a transversal approach that supports all decision-making functions, from data collection to its presentation to decision-makers. It is based on a new paradigm that we call geo-decision-making. We describe the formalization of the geo-decision-making chain and how it can be applied to the domain of maritime surveillance. We have paid particular attention to two components of the chain (Spatial Data Mining and Spatial OLAP) which enable the discovery of knowledge for the analysis of maritime risks. The application of spatial data mining techniques to historical maritime accidents makes it possible to discover interesting knowledge that may be used to automate risk identification.

It is important to note that we are not suggesting that operational maritime surveillance systems be replaced by decision support systems; on the contrary, the idea is to improve decision support in the domain of maritime surveillance. Both systems must evolve in parallel in order to provide complementary functionality.

---

[2] Association rule learning is an unsupervised data mining method that makes it possible, from a set of objects that frequently appear together in a database, to extract knowledge rules.

[3] This is an indicator of the confidence of the rule. It indicates the frequency of transactions in the data set which contain the itemset.

[4] Clustering is a data mining method that makes it possible to automatically group objects in clusters according to their degree of similarity or difference.

References

1. Torun, A., Düzgün, S.: Using spatial data mining techniques to reveal vulnerability of people and places due to oil transportation and accidents: a case study of Istanbul strait. ISPRS Technical Commission II Symposium. pp. 12 – 14. , Vienna (2006).
2. Marven, C., Canessa, R., Keller, P.: Exploratory Spatial Data Analysis to Support Maritime Search and Rescue Planning. In: Li, J., Zlatanova, S., and Fabbri, A.G. (eds.) Geomatics Solutions for Disaster Management. pp. 271–288. Springer Berlin Heidelberg, New York (2007).
3. Morel, M.: SisMaris: "Système d'Information et de Surveillance MARitime pour l'Identification des comportements Suspects de navire" in French. première Conférence Méditerranéenne Côtière et Maritime CM 2. pp. 261–264. , Hammamet - Tunisie (2009).
4. Morel, M., Broussolle, J.: I2C, Interoperable sensors & information sources for Common detection of abnormal vessel behaviours and collaborative suspect events analysis. MAST 2011 Conference Session. MAST Events Ltd , Marseille, France (2011).
5. Idiri, B., Napoli, A.: The automatic identification system of maritime accident risk using rule-based reasoning. IEEE Internationale Conference on System of Systems Engineering (SOSE 2012). , Genoa, Italie (2012).
6. Chaze, X., Bouejla, A., Napoli, A., Guarnieri, F., Eude, T., Alhadef, B.: The Contribution of Bayesian Networks to Manage Risks of Maritime Piracy against Oil Offshore Fields. International Workshops Information Technologies for the Maritime Sector (ITEMS 2012), Database Systems for Advanced Applications, Lecture Notes in Computer Science Volume 7240. pp. 81–91. Springer Berlin Heidelberg, Busan, South-Korea (2012).
7. Claramunt, C., Devogele, T., Fournier, S., Noyon, V., Petit, M., Ray, C.: Maritime GIS: From Monitoring to Simulation Systems. IF&GIS 2007. pp. 34–44. , St. Petersburg, Russia (2007).
8. VanDeCasteele, A., Napoli, A.: "La place du géodécisionnel dans les systèmes de surveillance maritime de nouvelle génération Apport du géodécisionnel dans la surveillance maritime" in French. Conférence internationale de Géomatique et d'Analyse Spatiale (SAGEO 2011), in Proceedings Espace géographique: exploration, extraction, modélisation, analyse et représentation des données et des connaissances. pp. 1–15. , Paris, France (2011).
9. Amrozowicz, M.D., Brown, A., Golay, M.: A Probabilistic Analysis Of Tanker Groundings. 7th International Offshore and Polar Engineering Conference. pp. 1–19. , Honolulu, Hawaii (1997).
10. Uluscu, O., Ozbas, B., Altiok, T., Or, I.: Risk analysis of the vessel traffic in the strait of Istanbul. Risk analysis: an official publication of the Society for Risk Analysis. 29, 1454–72 (2009).
11. LeBlanc, L.A., Rucks, C.T.: A multiple discriminant analysis of vessel accidents. Accident Analysis & Prevention. 28, 501–510 (1996).
12. Merrick, J., Van Dorp, J., Mazzuchi, T., Harrald, J., Spahn, J., Grabowski, M.: The PrinceWilliam Sound risk assessment. 25–40 (2002).

13. Roy, J.: Rule-based expert system for maritime anomaly detection. SPIE, the International Society for Optical Engineering. p. 12p. , Orlando, Florida, USA (2010).
14. Etienne, L., Devogele, T., Bouju, A.: Spatio-temporal trajectory analysis of mobile objects following. the International Symposium on Spatial Data Handling (SDH). , Hong Kong (2010).
15. Willems, N., van de Wetering, H., van Wijk, J.J.: Evaluation of the Visibility of Vessel Movement Features in Trajectory Visualizations. Computer Graphics Forum. 30, 801–810 (2011).
16. Riveiro, M., Falkman, G.: The role of visualization and interaction in maritime anomaly detection. In: Chung Wong, P., Park, J., Hao, M.C., Chen, C., Börner, K., Kao, D.L., and Roberts, J.C. (eds.) Proceedings of the SPIE. p. pp. 78680M–78680M–12 (2011).
17. Vandecasteele, A., Napoli, A.: Enhancement of Ontology with Spatial Reasoning Capabilities to Support Maritime Anomaly Detection. System of Systems Engineering (SoSE). , Genoa, Italie (2012).
18. Gouarné, J.-M.: "Le Projet décisionnel: enjeux, modèles et architectures du Data Warehouse" in French. (1997).
19. Sandoval, V.: "L'informatique décisionnelle" in French. (2007).
20. Krzysztof, K., Han, J.: Discovery of Spatial Association Rules in Geographic Information Databases. In: Egenhofer, M.J. and Herring, J.R. (eds.) Advances in Spatial Databases: 4th International Symposium SSD '95. pp. 47–66. Springer-Verlag, Portland, ME, USA (1995).
21. Bédard, Y.: Spatial OLAP. Forum annuel sur la R-D, Géomatique VI: Un monde accessible. pp. 13–14 (1997).
22. Idiri, B., Napoli, A.: "Découverte de règles d'association pour l'aide à la prévision des accidents maritimes" in French. 12ème Conférence Internationale Francophone sur l'Extraction et la Gestion de Connaissance (EGC 2012), Revue des Nouvelles Technologies de l'Information RNTI-E-23 Éditions Hermann. pp. 243–248. , Bordeaux, France (2012).