An Integrated Maritime Reasoning and Monitoring System

Kok Soon Oliver Tan, Sian Soo Tng
Defence Science & Technology Agency
Singapore
{tkoksoon, tsiansoo}@dsta.gov.sg

Abstract—Maritime information systems typically offer system features for data collection, information fusion and the automated construction of a basic recognized maritime picture for human operators. However making sense of maritime situations is still very much an intensive human-based cognitive endeavor. Limited human resources face difficult cognitive challenges in making sense of huge amount of data generated daily by dynamic round-the-clock maritime shipping and port activities. An effective information exploitation system for maritime surveillance and monitoring is expected to be able to augment human-based surveillance and monitoring operations with machine-based computing capabilities to interpret and reason about massive amount of situational information, as well as transforming this knowledge into actionable decision indicators for the human decision makers.

The Defence Science & Technology Agency (DSTA) has developed the PACKED model to guide the exploration and development of computing technologies that support human decision making and other cognitive challenges in information rich, complex, and dynamic operational environments. An integrated maritime reasoning and monitoring system illustrates how computing technologies are applied in the areas of Attention, Knowledge, and Comprehension to support situation analysis and monitoring processes for maritime surveillance.

The maritime reasoning and monitoring system consist of Bayesian reasoning, entity network analysis, and movement pattern analysis, to integrate, analyze, interpret and reason about maritime data. The Graylist Network Analysis (GNA) engine probes networks of entity-relationships to uncover “graylisted” entities that are closely associated with blacklisted entities. The Movement Pattern Analysis (MPA) engine uncovers common patterns in vessel movements and monitors real-time for maritime traffic data for anomalous vessel movements. The Bayesian network inference engine combines situation information from data sources and analysis results produced by analytical engines such as the GNA and MPA, with Bayesian knowledge models in order to infer probabilities of occurrences of maritime scenarios.

Keywords – Maritime Monitoring System; Bayesian Reasoning; Scenario Modeling; Relationship Analysis; Trend Analysis; Anomaly Detection

I. INTRODUCTION

Maritime surveillance and monitoring operations have been under increasing pressure to anticipate, detect, and prevent the occurrence of acts of hostility, such as piracy and terrorism. Maritime information systems typically offer system features for data collection, information fusion and automated construction of a basic recognized maritime picture for human operators. However making sense of maritime situations is still very much an intensive human-based cognitive endeavour. Limited human resources face difficult cognitive challenges in making sense of the huge amount of data generated daily by the dynamic round-the-clock maritime shipping and port activities.

An effective information exploitation system for maritime surveillance and monitoring is expected to be able to augment human-based surveillance and monitoring operations with machine-based computing capabilities to interpret and reason about massive amount of situational information, as well as transforming the information into actionable decision indicators for the human decision makers. However, while anomaly detection systems have been used to highlight irregularities in data and identify potentially hostile vessel behaviors, the issue of high false alarm rates remains. There is also the difficulty of forming comprehensive associations between anomalies and scenarios of interest so that follow-on decisions and interdiction actions may be better rationalized.

II. MOTIVATION AND OBJECTIVES

The Defence Science & Technology Agency (DSTA) has developed the PACKED model to guide the exploration and development of computing technologies that support human decision making and other cognitive challenges in information rich, complex, and dynamic operational environments e.g. maritime security. The decision-centric model, shown in Fig. 1, is inspired by the Human Information Processing model [1] but also takes into account the classes of decision support applications that have been deployed in the past, as well as growing demands of advanced computing technologies to provide better support for sensemaking and decision-making.
Briefly, the PACKED model consists of 6 areas of focus for decision support technologies:

1. Perception – This is about technologies that mitigate the data overload challenges with automatic extraction of contextually relevant and meaningful information from raw data sources. Examples of technologies in this area include entity network extractions from unstructured textual sources, video content analysis for object detection and identification, data fusion engines that correlate and combine raw pieces of data into more meaningful information.

2. Attention – These technologies are specifically designed to support attention management by directing decision-makers to known suspicious data patterns whenever they are uncovered in incoming data sources. Attention management also includes automated monitoring of progress of plans or courses of action that follows after decisions are made. Fraudulent activity detection, behavioral anomaly detection, and key performance indicators monitoring are examples of attention management applications.

3. Comprehension – These are analytics-based technologies that help the analysts or decision-makers understand newly acquired information, discover new knowledge, and trigger the formation of new insights by the human analysts. Examples of analytics include movement pattern analysis, scenario analysis, and relationship analysis.

4. Knowledge - The development of insights could also require the tapping of organizational knowledge or harnessing knowledge elicited from other subject-matter experts. Technologies in this area include knowledge management systems and expert systems.

5. Evaluation – This is about generating and evaluating possible courses of actions against organizational goals and constraints. Resource allocation algorithms and simulation-based options evaluation systems are typical technologies that belong to this area.

6. Decision – This technology focus here is to support the performance of tasks that arise from decisions made. Intelligent agent-based automation technologies could be designed to takeover the mundane but necessary tasks from the human operators. The execution of decisions in complex operational domains could also require the cooperation and resources of multiple stakeholders. In this case, computer-supported collaboration technologies could be used to assist in the coordination of distributed and interdependent tasks spanning different parties.

DSTA and the Defence Science Organisation National Laboratories have developed an integrated maritime reasoning and monitoring system that incorporates the following technologies [2]:

1. D'Brain
2. Graylist Ranking
3. Dual-HDP Trajectory Analysis

This paper describes how these three computing technologies are applied in the areas of Attention, Knowledge, and Comprehension to support maritime situation analysis and monitoring processes.

III. APPROACH

A. Scenario Analysis

We propose that a Bayesian reasoning system that we call the Scenario Analysis (SA) engine can be used to make sense of information about maritime situations and to augment human operators in surveillance and monitoring operations. The technology objective for the SA engine is to provide human analysts a Knowledge tool that taps on elicited expert knowledge. The SA engine encapsulates the Scenario Analysis process shown in Fig. 2. The SA engine offers a systematic approach to maritime knowledge modelling as well as a machine-assisted platform for monitoring maritime scenarios of interest.

The SA engine consists of 2 components – the D’Brain module and the Software Agents module. The D’Brain module is used for Bayesian-based knowledge modelling and probabilistic situation reasoning [3]. Causal and probabilistic expert knowledge models of scenarios of interest are represented as knowledge fragments. A knowledge fragment is a small network comprising of conditional indicators (nodes) and probabilistic interrelationships (edges). Expert maritime knowledge about entities and causal relationships of events in maritime incidents can be modelled as nodes and edges in these knowledge fragments. These knowledge fragments can be created collectively by a team of subject-matter experts. Each maritime scenario is a combination of knowledge fragments.

The D’Brain module dynamically combines various stored knowledge fragments at run-time by bringing various conditional indicators together and evaluates the likelihood of probable scenario outcomes based on their respective thresholds [4]. Each evaluation is entity-centric i.e. the focus is on identifying entities of interest. This is akin to...
culminating pieces of evidence from multiple data sources in a criminal investigation process.

The Software Agents module is an Attention tool that assists human analysts with automated monitoring for indicators of unfolding maritime scenarios. The Software Agents module is made up of intelligent software-based agents designed to watch for conditional indicators of any probable scenarios described by the knowledge fragments in the D’Brain module. There are 3 categories of conditional indicators:

1. Behavioural indicators (e.g. deviations from known shipping routes, intrusion into prohibited areas, excessive speed of travel)
2. Relationship indicators (e.g. connections between 2 entities, increases in suspicion levels of entities)
3. Environmental characteristics and entity attributes indicators (e.g. time of the day, vessel type)

The following is an example of a simple scenario analysis process:

i. Scenario Identification
   A hypothetic scenario is first conceived by human analysts based on some probable critical maritime events. Very often, the analysts may make reference to any malicious incidents that had happened in other countries. The scenario identification process could also be triggered by information from news articles or intelligence reports.

ii. Scenario Modelling
   The analyst will next use the SA engine to model the scenario as knowledge fragments in the D’Brain module. Measurable conditional indicators are encoded as network nodes in the knowledge fragments. Network nodes could also represent more subjective indicators that require additional human-based assessments.

iii. Computer-aided Scenario Monitoring
   Automated agent-based monitoring of the scenario of interest can be activated after the scenario model is created. Besides data from maritime data sources, the analysis results from the other analytical engines e.g. trend analysis and relationship analysis engines, can also be fed into the Software Agents module. When sufficient indicators for a scenario are detected, the SA Engine informs the human operators via a maritime situation information display. Suspected vessels that would then be subjected to further investigations and deeper research by human analysts.

B. Relationship Analysis

The Graylist Network Analysis (GNA) engine addresses information analysis challenges that fall in the area of Comprehension by building up pertinent relationship models that could be automatically analyzed based on the Relationship Analysis process shown Fig. 3. Relationship models are specific networks of relationships among maritime entities that have been built up and stored in a database through the fusion of incoming structured and unstructured information.

Graylist ranking is part of the relationship analysis process for making sense of associations between entities. The general idea employed in graylist ranking is as follows. First, the network of entities needs to be properly constructed. Entity networks in the maritime domain are heterogeneous in nature with many entity and relationship types. For example, companies may be linked to one another in a variety of manner, e.g. through the chartering of ships, crew, cargo-related contracts etc. Crews may be linked directly to ships and also to the companies they worked for. For correctness, each unique entity should correspond to one vertex in the network.

This is followed by the identification and ranking of suspicious entities. The aim is to rank suspects higher than non-suspects and there are many powerful ranking algorithms from the literature that we can apply.

A popular algorithm and perhaps a natural choice for validating our relationship analysis process is the PageRank algorithm that drives the popular Google search engine [5]. This algorithm assigns a numerical weight to a hyperlinked set of documents with the purpose of "measuring" its relative importance within the set. It uses the web’s hyperlink structure as an indicator of an individual page's value [6]. In essence, a link from page A to page B is interpreted as a vote
by page A for page B. The page that casts the vote is also analyzed. Votes cast by pages that are themselves "important" weigh more heavily and help to make other pages "important". The PageRank of a page is defined recursively and depends on the number and PageRank values of all incoming links. A page that is linked with many pages with high PageRank values receives a high rank itself. If there were no links to a web page, this would mean that there was no support for that page.

**Figure 3.** Relationship analysis process.

Modification of the original PageRank algorithm is required for the purpose of finding suspicious maritime entities. The modified algorithm is called Constrained PageRank. Given that the PageRank algorithm is propagating the rank values during the iteration process, we start with a set of vertices which are already known to have high rank values (i.e. representing the “blacklisted” entities) and use these vertices to evolve the rank values (i.e. known as the “grayness” scores) for the other vertices. The basic idea is to first assign high and fixed rank values for key vertices and allow these vertices to propagate rank values to the rest of the vertices throughout the network.

**C. Movement Pattern Analysis**

Another analytical tool that we have developed to support Comprehension is the Movement Pattern Analysis (MPA) engine. In the proposed movement pattern analysis process shown in Fig. 4, analysts first identify maritime regions of interest based on general statistics of historical vessel movement data. This is followed by the selection of relevant data subsets and desired features for clustering by the MPA engine. The MPA engine examines vessel movements for common activity patterns. The Dual-HDP model for trajectory analysis is applied to learn shipping activity patterns from observed vessel traffic data features that include vessel positions, courses and speeds [7]. Trend models that comprise of multiple learned activity patterns are finally generated by the engine.

Next, the generated trend models are used for anomaly detection. The MPA engine monitors vessel movements in real-time. Vessels are assigned anomaly scores based on the probability of their association with the learned trend models that represent normally expected vessel behaviours. The higher the associations with the trend models, the lower will be the anomaly scores for the vessels. The vessels are then ranked according to their anomaly scores for further investigation by human analysts.

**Figure 4.** Movement pattern analysis process.

**IV. RESULTS**

The results of the scenario analysis, relationship analysis and movement pattern analysis processes are briefly illustrated with the following examples.

**A. Scenario Analysis**

Fig. 5 shows a possible scenario in maritime security. The scenario consists of multiple knowledge fragments that are combined at run-time by the D’Brain module. The input nodes, or the indicators, are monitored by software agents in the Software Agents module. The human analysts are alerted when sufficient indicators are detected to infer a likely occurrence of a node of interest. For example the leaf node “suspected floating bomb attack” is the node of interest in this maritime security scenario that involves indicators such as “ship deviation trend”, “ship speed”, “ship grayness rank value” etc.
B. Relationship Analysis

Fig. 6 shows an entity relationship network with a list of entities displayed to the right (note: the entity names have been blurred out for confidentiality reasons). A list of pre-identified blacklisted entities is used by the GNA engine to produce a grayness score for every entity in the relationship network. Grayness scores are visualised with varying shades of gray where darker shades correspond to higher scores.

C. Movement Pattern Analysis

The MPA engine was used to process maritime traffic data collected from the Singapore Territorial Waters as well as from the Arabian Sea. Figure 7 shows two activity patterns identified by the MPA engine using ferry movement data from the Singapore region. Figure 8 shows two activity patterns of shipping vessels in the Arabian Sea region.
V. CONCLUSION

We have assembled a maritime reasoning and monitoring system, comprising Bayesian reasoning, entity network analysis, and movement pattern analysis, to integrate, analyze, interpret and reason about maritime data. The Graylist Network Analysis (GNA) engine probes networks of entity-relationships to uncover “graylisted” entities that are closely associated with blacklisted entities. The Movement Pattern Analysis (MPA) engine uncovers common patterns in vessel movements and monitors real-time maritime traffic data for anomalous vessel movements. The Bayesian inference engine combines situation information from data sources and analysis results produced by analytical engines, such as the GNA and MPA, with Bayesian knowledge models in order to infer probabilities of occurrences of maritime scenarios. These scenario occurrence probabilities could also be used as early warning indicators by other maritime surveillance and monitoring systems for more comprehensive situation reasoning functions.

REFERENCES


Figure 8. Vessels activity pattern in the Arabian Sea region (Feature: Position, Course, Speed).