Signature-based activity detection based on Bayesian networks acquired from expert knowledge

Farzad Fooladvandi  
Saab Microwave Systems  
Training systems and Information Fusion  
Skövde, Sweden.  
farzad.fooladvandi@saabgroup.com

Christoffer Brax  
Saab Microwave Systems  
& University of Skövde  
Skövde, Sweden.  
christoffer.brax@his.se

Per Gustavsson  
Saab Microwave Systems  
Training systems and Information Fusion  
Skövde, Sweden.  
per.m.gustavsson@saabgroup.com

Mikael Fredin  
Saab Microwave Systems  
Training systems and Information Fusion  
Skövde, Sweden.  
mikael.fredin@saabgroup.com

Abstract - The maritime industry is experiencing one of its longest and fastest periods of growth. Hence, the global maritime surveillance capacity is in a great need of growth as well. The detection of vessel activity is an important objective of the civil security domain. Detecting vessel activity may become problematic if audit data is uncertain. This paper aims to investigate if Bayesian networks acquired from expert knowledge can detect activities with a signature-based detection approach. For this, a maritime pilot-boat scenario has been identified with a domain expert. Each of the scenario's activities has been divided up into signatures where each signature relates to a specific Bayesian network information node. The signatures were implemented to find evidences for the Bayesian network information nodes. AIS-data with real world observations have been used for testing, which have shown that it is possible to detect the maritime pilot-boat scenario based on the taken approach.

Keywords: Signature-based detection, Bayesian networks, Knowledge elicitation, Maritime situation awareness, Information fusion.

1 Introduction

The maritime industry is experiencing one of its longest and fastest periods of growth. This phase is due to the past 10 years that has seen an annual growth rate of 3.8% in transport volume, and in the past 3 years this growth rate has almost doubled [12]. Hence, the global maritime surveillance capacity is in a great need of growth as well. This stems according to [5] from the levels of hazardous cargo transports, smuggling of goods and humans, and growth in global terrorism. Due to these activities there is an ongoing implementation of new cooperative systems for ship reporting to meet emerging requirements for detection, identification, and tracking.

The detection of unusual vessel activity is an important civil security maritime domain awareness (MDA) objective. This can be particularly challenging in environments with much vessel traffic.

Maritime organizations involving both the civilian and the military domain often have access to a number of surveillance sources. The ability to make full use of these surveillance systems, e.g., for detecting events and behaviours, is limited due to their inability to fuse data and information from all sources in a timely, accurate, and complete manner. Automated association of sensor information with non-sensor information is an important functionality for surveillance systems, which can help with such tasks as search and rescue, monitoring of specific regions and identifying ship activities that may threaten environment or national security [8]. This is the task of information fusion, which involves combining data and information from multiple sources e.g., sensors and domain experts. Furthermore, the task involves relating this information to achieve improved accuracy and more specific inferences which could not be achieved by the use of a single sensor alone [4].

Information fusion in general and within the military domain in particular contains a high degree of uncertainty. An important technique for uncertainty management is probability theory. A technique such as neural networks is a way of approaching uncertainty, but an alternative to this is Bayesian networks [7]. Given the diverse uncertainty management techniques, this paper will consider a Bayesian network approach to activity detection in the maritime domain. Mainly for the advantages that Bayesian networks possess, such as graphical representation and their suitability for real-time updating when new evidence is produced. There is a need to investigate if different techniques in combination with diverse approaches can help in detecting activities.

2 Background

2.1 Bayesian network

A Bayesian network (BN) is according to [10] and [6] a directed acyclic graph (DAG) consisting of a set of nodes and edges which represents probabilistic (causal) dependences among variables. A BN conveys a joint probability distribution of its variables, which is the product of the local distributions of each node and its parents. The variables for the nodes represent a finite set of mutually exclusive and exhaustive states. The nodes with edges
directed/pointing into them are called “child” nodes and the
nodes which direct/point edges to other nodes are called
“parent” nodes. The DAG represents the structure of
dependencies between nodes and gives the qualitative part
of a BN. The quantitative part consists of conditional
probability tables (CPTs) which are attached to each node.
The relationships between different nodes are quantified by
the conditional probability distribution. The numbers
that are quantified in the conditional probability distribution are
encoded into the BN by using a set of CPTs. The numbers
that a CPT consists of is often facilitated by either experts
determining the numbers or statistical data that is acquired
from real life experiments [7].

2.2 Activity detection systems

Approaches for detecting activities or anomalies that will
be described here are signature-based detection, anomaly-
based detection and a hybrid variant. The difference
between these three approaches is that a signature-based
detection system identifies patterns in data presumed to be
of particular interest. The patterns are referred to as
signatures which are specific activities/behaviours which
are of interest to detect. An anomaly-based detection (data-
driven) system compares activities/behaviours against a
normal baseline e.g., the normal behaviour of entities [9].
The third approach combines the techniques of the two
detection systems to form a hybrid system.

2.3 AIS

The Automatic Identification System (AIS) is a maritime
safety and vessel traffic system imposed by the
International Maritime Organisation (IMO) [1]. The system
broadcasts position reports and short messages with
information about the ship and the voyage e.g., vessel
identity, position, heading, destination, estimated time of
arrival, nature of cargo, etc. This sort of information can
assist in monitoring and tracking maritime entities for
security reasons.

3 Maritime pilot-boat scenario

The pilot-boat scenario was created with a domain expert
from Saab Microwave Systems. The pilot-boat scenario
involves ships that are escorted by one or more pilot-boats
that have the responsibility to escort a ship to a harbour.
The reason for this is that foreign ships may not be familiar
with how strong currents are in the water and also lack the
practical knowledge of how to dock a ship in a crowded
harbour. Therefore, pilot-boats assist ships in this task so
that accidents and mishaps can be avoided. Figure 1
illustrates the scope of the pilot-boat scenario.

The activities involved in figure 1 are represented
through a set of numbers in the figure. The activities are:

1. Ship waits for pilot-boat(s) within an unspecified waiting
area. The waiting area is denoted by a square.
2. The pilot-boat leaves the docking area.
3. Pilot-boat meets up with the ship and initiation of escort
begins.
4. Pilot-boat travelling to the harbour followed by ship.
This is the actual escort.
5. Ship docks at the docking area belonging to the harbour.
6. Pilot-boat returns to an unspecified docking area.

4 Build and learn a Bayesian network

In this subsection, the construction of the Bayesian network
with the help of expert knowledge will be explained. The
knowledge elicitation session carried out with a domain
expert will also be explained. Finally, both an overview and
a detailed examination of the signature-based detection
software will be given.

4.1 Qualitative part of the Bayesian network

A Bayesian network (BN) has been created of the identified
scenario. The BN model is built up with a TAN (tree-
augmented naive Bayes) structure which is a simple
structure when modelling phenomena from the real world.
The TAN structure follows a diverging topology. Figure 2
illustrates the created BN model for the pilot-boat scenario.

Through the characteristics of the BN model, the child
nodes can be regarded as information variables for the BN,
and the parent node as the query variable. The reason for
this is because the child nodes, given the evidence set on
their values, will determine the outcome of the parent node.
Hence, the calculations made by the child nodes will affect
the probability of the parent node. If a value of the parent
node is known e.g., evidence is present, the child nodes will
not have any affect on each other or on the parent node.
That is, information variables are assumed to be
independent from each other given the evidence about the
query variable. The given explanation exemplifies a
diverging topology. According to [14] the TAN structure
has a linear time complexity for posterior probability

Figure 1. The pilot-boat scenario
calculations and provides very good performance when it comes to problems where one is only interested in identifying the most probable causes of a phenomenon. The pilot-boat scenario is similar in the sense that only certain activities can indicate that the scenario is true. The nodes of the BN model will be described next:

**Scenario in action** is the parent node variable and it has two values, which are true and false. This node is the query variable which represents a belief about whether the scenario is in action or not. If no evidence is present about the parent node, then the probability of the scenario being true or false is dictated by the parent’s child nodes.

**Ship waiting at area** is the first child node and it takes into account how long a ship has been waiting at an area. The values for this node is, $T$ for time in minutes, $T > 0$ and $T \leq 5$, $T > 5$ and $T \leq 10$, and finally, $T > 10$. Evidence on one of these values will be set, depending on the time that a vessel has been waiting. The discrete values are based on expert knowledge.

**Ship meets pilot-boat** is the second child node which also is the parent of **Ship escorted by pilot-boat**. This node follows the same discretization as the previous child node. With this node the values are based on the distance between a ship and a pilot-boat e.g., $M > 0$ and $M \leq 5$, $M > 5$ and $M \leq 10$, and, $M > 10$, where the unit for $M$ is meters. Evidence will be set on one of the values depending on the distance between a particular vessel and a pilot-boat.

**Ship escorted by pilot-boat** is the third child node which also is the parent of **Ship reach harbour**. This child node takes into account whether a ship is being escorted by a pilot-boat or not. The values for this child node is escorted and not escorted. This node is influenced by the **Ship meets pilot-boat** node, and depending on the evidence set on the values of the **Ship meets pilot-boat** node, the probabilities for **Ship escorted by pilot-boat** are influenced.

**Ship reach harbour** is the fourth child node and this node takes into account whether a ship has reached a harbour or not. This node is also the parent of **Pilot-boat returns to area**. The values for this child node is reach harbour and not reach harbour. The two values are influenced by the evidence set on the **Ship escorted by pilot-boat** node.

**Pilot-boat returns to area** is the fifth and the last child node. This node takes into account if a pilot-boat returns to an unknown area near a harbour after it has escorted a vessel to a harbour. This child node has two values, true and false. It is also influences by the **Ship reach harbour** node.

The BN model resembles the pilot-boat scenario by having the characterizing activities as child nodes e.g., information variables. The structure of the BN model tolerates if an activity is undetected. This means that if the detection of a particular activity is missed for some reason, the BN model can still output an answer. It can be beneficial to have this tolerance due to the uncertainty which resides within each activity.

The advantage of having dependencies between some of the child nodes is to acquire a more realistic representation of the pilot-boat scenario. The choice of incorporating influences between the child nodes allows the nodes to affect each other which results in more realism. For instance, if a ship is being escorted then this would influence the chances of the ship reaching a harbour. The disadvantage of having dependencies between child nodes is that the CPTs become more complex. This may complicate the process of inserting the probabilities into the CPTs and also complicate the knowledge elicitation process. The next part will consider the quantitative part of the BN model.

### 4.2 Quantitative part of the Bayesian network

The quantitative part of a BN involves the CPTs and how to elicit and extract information in a probabilistic form. For this task, a tool was created which involved implementing research material by [2], [3] and [13]. The purpose was to bring forward the graphical characteristics of the verbal-numerical probability scale to a computerized environment, and subsequently compose a means for extracting and eliciting knowledge in a correct fashion. Two parts will be explained here, the computerization of the research material and the knowledge elicitation session which was carried out.

Figure 3 illustrates the knowledge elicitation (KE) tool created. It has mainly two parts, an elicitation session capability and a compilation feature which compiles the probabilities so that they are ready for utilization by the
BN. At the top, the question at hand is displayed and below the question are the values in relation to the BN node. The structure of the inquiring questions can also be viewed at the top. In the middle, the actual verbal-numerical probability scales are visible. As illustrated, the questions can be answered with a label or an associated numerical alternative, depending on how an expert wants to answer the question.

The preparation and accomplishment of the KE session was performed in accordance with the procedure in [11]: 1) select and motivate the expert, 2) train the expert on the elicitation process, 3) structure the questions, 4) elicit and document the expert judgments, and 5) verify the results.

The questions were structured according to the order of the BN nodes e.g., from left to right. Each question took into account whether the parent node was true or false. When it came to the child nodes that had dependencies from both the parent node and another child node, the question became somewhat more complicated.

In conjunction with the question, the values of the actual BN node are aligned with each scale. The purpose of this is to bring attention to the values at hand and to allow the expert to have a visual overview of the values that need to be assessed. Table 1 to table 6 illustrates the CPTs for each BN node. The probabilities of the CPTs are based on the expert knowledge which was acquired with the KE tool.

<table>
<thead>
<tr>
<th>Scenario in action</th>
<th>True</th>
<th>False</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.14</td>
<td>0.86</td>
<td></td>
</tr>
</tbody>
</table>

Table 1. CPT for the parent node Scenario in action.

Table 2. CPT for the child node Ship waiting at area

<table>
<thead>
<tr>
<th>Ship waiting at area</th>
<th>Scenario in action</th>
</tr>
</thead>
<tbody>
<tr>
<td>T &gt; 0 and T &lt;= 5</td>
<td>True 0.06</td>
</tr>
<tr>
<td>T &gt; 5 and T &lt;= 10</td>
<td>True 0.2</td>
</tr>
<tr>
<td>T &gt; 10</td>
<td>True 0.74</td>
</tr>
</tbody>
</table>

Table 3 shows the probabilities of how close a ship needs to be to a pilot-boat before it is considered to be involved in a pilot-boat scenario. The CPT in table 3 follows the same structure as the CPT in table 2. Hence, there is a 78% probability that the distance between the ship and the pilot-boat is within the interval of 0 to 5 meters, given that the pilot-boat scenario is true. Neither of the CPTs in table 2 and table 3 are affected by other child nodes. From the appearance of the CPT in table 4, one can see that it is more complex in comparison to the CPTs in table 2 and table 3. This is due to the influence by another child node. If the BN was absent from influences between the child nodes, table 4 would be similar to table 2 and table 3. The CPT in table 4 presents the probabilities of a ship being escorted, depending on the distance between the ship and the pilot-boat, and whether or not the pilot-boat scenario is true. It is possible to see that the CPT in table 4 brings about more realism, but at the cost of complexity.

Table 3. CPT for the child node Ship meets pilot-boat.

<table>
<thead>
<tr>
<th>Ship meets pilot-boat</th>
<th>Scenario in action</th>
</tr>
</thead>
<tbody>
<tr>
<td>M &gt; 0 and M &lt;= 5</td>
<td>True 0.78</td>
</tr>
<tr>
<td>M &gt; 5 and M &lt;= 10</td>
<td>True 0.19</td>
</tr>
<tr>
<td>M &gt; 10</td>
<td>True 0.03</td>
</tr>
</tbody>
</table>

The CPT in table 5 follows a similar structure as in table 4. The child node Ship reach harbour is influenced by another child node which explains the complexity of this CPT too. The CPT in table 5 presents the probabilities of a ship reaching a harbour, depending on whether or not the ship have been escorted and if the pilot-boat scenario is true or false.

The final CPT in table 6 presents the probabilities of a pilot-boat returning to an area, depending on whether or not the assisted ship have reached a harbour and if the pilot-boat scenario is true or false.
Table 4. CPT for the child node Ship escorted by pilot-boat.

<table>
<thead>
<tr>
<th>Scenario in action</th>
<th>Escorted</th>
<th>Not escorted</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ship meets pilot-boat</td>
<td>M &gt; 0 and M &lt;= 5</td>
<td>M &gt; 5 and M &lt;= 10</td>
</tr>
<tr>
<td>M &gt; 0 and M &lt;= 5</td>
<td>0.84</td>
<td>0.78</td>
</tr>
<tr>
<td>M &gt; 5 and M &lt;= 10</td>
<td>0.86</td>
<td>0.69</td>
</tr>
</tbody>
</table>

Table 5. CPT for the child node Ship reach harbour.

<table>
<thead>
<tr>
<th>Scenario in action</th>
<th>Escorted</th>
<th>Not escorted</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ship reach harbour</td>
<td>M &gt; 0 and M &lt;= 5</td>
<td>M &gt; 5 and M &lt;= 10</td>
</tr>
<tr>
<td>M &gt; 0 and M &lt;= 5</td>
<td>0.98</td>
<td>0.9</td>
</tr>
<tr>
<td>M &gt; 5 and M &lt;= 10</td>
<td>0.02</td>
<td>0.1</td>
</tr>
</tbody>
</table>

Table 6. CPT for the child node Pilot-boat returns to area.

<table>
<thead>
<tr>
<th>Scenario in action</th>
<th>True</th>
<th>False</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ship reach harbour</td>
<td>Reaching</td>
<td>Not reaching</td>
</tr>
<tr>
<td>Reaching</td>
<td>0.93</td>
<td>0.79</td>
</tr>
<tr>
<td>Not reaching</td>
<td>0.07</td>
<td>0.21</td>
</tr>
</tbody>
</table>

The complete signature-based detection software consists of a number of significant methods. These methods are important when detecting the signatures of each activity in the pilot-boat scenario. Next, an overview of the signature-based detection software will be presented.

4.3 Overview of the signature-based detection software

Figure 4 presents an overview of the signature-based detection software and illustrates the two primary sources of data that the software can utilize. One is the expert knowledge and the second is the AIS-data.

The first part in figure 4 involves the AIS-data preprocess, which filters and associates observations to vessels. The filtering is meant to consider observations that are only within a predefined coastal area.

The next part considers the actual signatures that resemble the pilot-boat scenario. The signature-based detection software is built up of methods that are related to specific BN child nodes. The preprocessed AIS-data is utilized by each of the BN related methods and the main task of each method is to find evidence for the BN child nodes. When the methods have been executed, evidence is passed to the next part which handles the BN.

The BN in figure 4 can be viewed as the point where two sources of processed data will be fused into information. The first source is the evidence acquired by the signature-based detection software. The second source is the expert knowledge which is elicited by the knowledge elicitation tool. The BN query node will deliver a probability as a result based on the two data sources. The probability will then indicate if a pilot-boat scenario has occurred or not. An operator can at this stage be notified of any detected pilot-boat scenarios.

4.4 Implementation of the signature-based detection software

The signature-based detection software is built up of five methods that are specifically created and calibrated to detect certain signatures. These five methods are of importance for the success of detecting the signatures, which together resemble the pilot-boat scenario. The performed calculation of each method will generate and collect evidence for a specific BN node, which the specific method is built for. Each of the methods that are bound to a specific BN child node will be described next:

isShipWaitingAtArea attempts to detect how long a vessel has been waiting in an area. A vessel can be moved by currents and waves while waiting, therefore a number of observations are examined and if the distance from the first observation to the last is less than for instance five meters, the vessel is regarded as waiting. The thresholds for the waiting periods are based on the values of the BN child node Ship waiting at area. When one of these values e.g., T > 0 and T <= 5, T > 5 and T <= 10, and, T > 10, are fulfilled, the information is stored for later methods and also for the Ship waiting at area child node, which will use this information as evidence.

isShipNearPilotBoat is the next method and attempts to find relations between a vessel and a pilot-boat. This method recovers evidence for the values of the second BN child node Ship meets pilot-boat which are M > 0 and M <= 5, M > 5 and M <= 10, and, M > 10. The main task here is to calculate the distance between two sets of latitude and longitude coordinates. Any results that reside within the intervals that the values advocate are stored. The results will indicate whether or not a vessel has been in contact with a pilot-boat. The stored results are then utilized as evidence for the Ship meets pilot-boat child node.

isShipBeingEscorted attempts to use the information compiled from the former method to detect if a vessel is being escorted by a pilot-boat. The method observes how long the vessel and the pilot-boat have had contact.
Figure 4. Overview of the signature-based detection software

e.g., being close to each other. The contact between the vessel and the pilot-boat must also involve movement to indicate that the vessel and the pilot-boat are travelling together. The close vicinity and the movement will decide whether or not the vessel is being escorted. This information is stored for the next methods and is also used as evidence for the Ship escorted by pilot-boat BN child node values escorted and not escorted.

isShipNearHarbour is the method that examines whether or not a vessel has reached a harbour. It is considered that a vessel is near or has reached a harbour if the vessel at hand has travelled within a harbour zone represented by a set of latitude and longitude coordinates. When a vessel and a pilot-boat are travelling together, and are detected to be within a harbour zone, the event will be stored. This information will then be used by later methods and will also be used for evidence by the Ship reach harbour BN child node. The child node values that this node finds evidence for is reach harbour and not reach harbour.

isPilotBoatReturningToArea attempts to detect if a pilot-boat, after escorting a vessel to a harbour, deviates from the vessel. The main task here is to detect when a pilot-boat is leaving a vessel. The method utilizes the information stored by the former method to determine when the activity of escorting has been completed. An examination will then be performed by the method to detect whether or not the pilot-boat is deviating from the vessel. If the pilot-boat is deviating from the vessel, then the method will observe the pilot-boat’s movement until it has stopped. When no more movement is detected, the pilot-boat is considered to have returned to an area. This information is then used as evidence for the Pilot-boat returns to area BN child node. The child node values that this method finds evidence for is true and false.

When all of the methods have been executed on each vessel, the evidence on each child node will dictate what the probability of the parent node Scenario in action will be. The probability of a pilot-boat scenario taking place or not, will be determined by the generated belief of the parent node Scenario in action. The belief is in probabilistic form and indicates how probable it is that the pilot-boat scenario has taken place. Based on the values of the parent node which are true or false, the signature-based detection software will be informed of whether or not the pilot-boat scenario was detected.

5 Testing the signature-based detection software

Two kind of tests where carried out on the complete software. The first one considered a manually created pilot-boat scenario to be inserted in the AIS-data. The manually created scenario had the purpose of testing if the signature-based detection software possesses the capabilities of detecting a pilot-boat scenario based on the identified pilot-boat activities/signatures, which the domain expert advocated. Figure 5 illustrates the setup for the manually created pilot-boat scenario. The pilot-boat scenario was inserted in AIS-data which resembled a ship that had been waiting for a pilot-boat for more than ten minutes. The distance between the ship and the pilot-boat was set to between 0 and 5 meters, and the specified interval was also considered when the ship was being escorted by the pilot-boat. That is, the ship and the pilot-boat would have a distance between 0 and 5 meters during the escort. The ship would reach the specified harbour depicted with a rectangle in figure 5. Finally, the pilot-boat would deviate from the ship and return to the area from which it started. The green (dark) dots in figure 5 represent a ship and the yellow (light) dots represent a pilot-boat. The black rectangle displays the harbour which the ship desires to dock at and the arrows represent the starting direction of each vessel. The result of this test will only reflect the software’s capability of detecting signatures in the absence of uncertainty. With real world observations, there are other factors that can affect the detection. Factors such as observation uncertainty, vessels behaving strange, environment conditions e.g., waves, currents and weather, can all affect the detection of pilot-boat scenarios. The second test applies a complete set of available AIS-data on the signature-based detection software. The purpose of this
The test is to examine the software executing on real world observations and also to analyse if any pilot-boat scenarios can be detected based on real world data. The complete set of available AIS-data covers the time period from the 1st to the 18th of January 2008. The location of the gathered AIS-data is the west coast of Sweden. There are more than 63 million observations available in the area within the time period. Only the observations residing within and close to the harbours of Gothenburg were considered. The areal choice is based on the mixture of both open waters and coastal area.

5.1 Test results

The first test which involved the manually created activities that resembled a pilot-boat scenario was detected. Hence, the test showed that each of the activities can be detected with the absence of uncertainty in the AIS-data. The test also showed that each of the methods detected the signatures that were present in the manually created pilot-boat scenario. Based on the evidences produced by the methods there was a 91% probability that the pilot-boat scenario took place.

The second test applied real world observations on the signature-based detection software. The result of this test showed that the software can detect ships involved in a pilot-boat scenario. Due to the lack of information on the actual number of pilot-boat scenario occurrences in the AIS-data, the accuracy of the BN and the signatures could not be set. The test showed that during 14 days of AIS-observations, the amount of detected pilot-boat scenarios was 74. These detected pilot-boat scenarios were then examined with the purpose of determining whether the detections were true positive or false positive. The examination showed that out of 74 potential pilot-boat scenarios, there were 45 true positives and 29 false positives. Hence, this justifies that the software has the necessary capability for detecting pilot-boat scenarios. The examination also showed that the software detects scenarios which are slightly similar to the pilot-boat scenario. But the performed examination only considered detected scenarios that where in accordance with the pilot-boat scenario described in this paper. There are approximately 3 pilot-boat scenarios occurring each day that are in accordance with the signatures specified by the domain expert.

Figure 5, Manually created pilot-boat scenario.

Figure 6. Number of pilot-boat scenarios detected per day.

The reasons for the 29 false positive detections can for instance be due to a reverse pilot-boat scenario where a ship is escorted from a harbour. Ships may also be involved in scenarios where pilot-boats are assisting in some other activities. For instance, a pilot-boat captain may board the ship and manually assist it to reach the harbour. Any scenarios which are similar to the one in this paper may trigger the signature-based detection software to notify that a potential pilot-boat scenario has been detected, even if it slightly deviates from the identified pilot-boat activities. Figure 6 displays the total number of detected pilot-boat scenarios over 14 days of AIS-data. The figure also illustrates the number of true positives and false positives for each observed day. Figure 7 illustrates a ship involved in a pilot-boat scenario which was detected based on real world AIS-data.

Figure 7. A pilot-boat scenario detected from AIS-data.

Figure 7 illustrates an actual pilot-boat scenario taking place. The green (dark) dots represent the ship, and the yellow (light) dots represent the pilot-boat. At the left of figure 7 one can see how the two vessels end up with each other. After that the actual escorting of the ship is taking
place. At the right the pilot-boat and the ship has reached the harbour and the pilot-boat returns continues to a waiting area.

6 Conclusions

It has been shown that Bayesian networks acquired from expert knowledge can detect activities with a signature-based detection approach. The TAN structure of the created BN makes it possible to freely modify the network. Hence, adding more detail to resemble the real world more accurately is possible. The BN model which represents the pilot-boat scenario can be used as a basis for representing other scenarios. In order to achieve this, child nodes e.g., information variables, have to resemble activities which are part of the scenario. Based on a collection of child nodes resembling all the scenario’s activities, the parent node e.g., query variable, have to represent the scenario. This sort of setup with information variables and a query variable made it possible to detect the pilot-boat scenarios with success.

An important factor for the detection of the pilot-boat scenario is the required evidences for each information node. To be able to acquire evidence, the signatures representing each information node need to clearly be defined. The signatures need to represent the activities which the scenario is based on. The reason for this is because the signatures will dictate in what procedure the underlying methods and mechanisms need to function.

The knowledge elicitation tool helped to elicit probabilities for the BN model representing the pilot-boat scenario. The knowledge elicitation tool was used in an elicitation session which based on the expert’s review was a smart and an efficient way of approaching a task of acquiring probabilities. This was compared to the procedure of manually inserting probabilities in the CPTs of the BN. The combination of verbal and numerical choices for answering questions made the process of eliciting probabilities faster and also comfortable for the expert.

References


