Associative Learning of Vessel Motion Patterns for Maritime Situation Awareness

Neil A. Bomberger, Bradley J. Rhodes, Michael Seibert, & Allen M. Waxman
Multisensor Exploitation Directorate, Fusion Technology and Systems Division
BAE Systems, Advanced Information Technologies
6 New England Executive Park, Burlington, MA 01803, USA
{neil.bomberger, brad.rhodes, michael.seibert, allen.waxman}@baesystems.com

Abstract - Neurobiologically inspired algorithms have been developed to continuously learn behavioral patterns at a variety of conceptual, spatial, and temporal levels. In this paper, we outline our use of these algorithms for situation awareness in the maritime domain. Our algorithms take real-time tracking information and learn motion pattern models on-the-fly, enabling the models to adapt well to evolving situations while maintaining high levels of performance. The constantly refined models, resulting from concurrent incremental learning, are used to evaluate the behavior patterns of vessels based on their present motion states. At the event level, learning provides the capability to detect (and alert) upon anomalous behavior. At a higher (inter-event) level, learning enables predictions, over pre-defined time horizons, to be made about future vessel location. Predictions can also be used to alert on anomalous behavior. Learning is context-specific and occurs at multiple levels: for example, for individual vessels as well as classes of vessels. Features and performance of our learning system using recorded data are described.

Keywords: Situation awareness, learning, prediction, maritime, neural networks.

1 Introduction

Exploitation algorithms to aid situation awareness aim to effectively provide automated assistance to analysts/operators to achieve their objectives. Often the environments in which these algorithms are to be used are non-stationary and provide limited training data. Our learning approach, inspired by biological and cognitive principles, operates well when confronted with such difficulties. Detection on unusual vessel activity is an important homeland security maritime domain awareness (MDA) objective. This can be particularly challenging in a busy port environment where many vessels are operating concurrently. Vessel activity can be considered at many levels, from atomic events (represented by the current state of a vessel in relation to its environment) through long-term behaviors (which could be conceived of as sequences of events).

One goal of our system is to continuously learn to detect anomalies with little or no operator supervision. We have previously reported successful learning to detect anomalous vessel event behavior [1]. This work is briefly reprised in Section 2 and used to present new results from AIS (Automatic Identification System) data recorded from the Miami Harbor area. Based on this successful application, attention turns toward learning to predict the future behavior of a vessel on the basis of its current behavior. Section 3 describes our methodology for learning to make such predictions and the performance results from the Miami data.

2 Event-level anomaly detection

The objectives are to learn what normal events are and to detect deviations from normalcy. What is normal may differ between contexts—such as class of vessel, weather conditions, or tidal status—so it is crucial that such learning be capable of discovering normalcy for differing contexts. The possible combinations of conditions creating different contexts are numerous enough to defy efforts to manually define (and subsequently use in real-time) a complete set of rules to cover all cases. A continuously learning, adaptive system holds promise for dealing with such complexity without unrealistic involvement of expert input on a frequent basis.

To learn context-sensitive models of vessel behavior, we employed a significantly modified version of the Fuzzy ARTMAP neural network classifier (as has been described more extensively in [1]). Use of this form of learning has also been applied successfully to a variety of image mining and tracking tasks [2-7]. The speed and performance of this learning algorithm makes it suitable for real-time and interactive situations wherein an operator/analyst can help teach the model via simple point and click actions. These reasons also make this technology suitable for event-level learning in MDA. The challenge was to develop this learning-based approach further to satisfy specific MDA needs.

Learning of normal events can be achieved via training with a set of observations that are known to reflect routine activity. Ideally the observations would contain sufficient numbers of exemplars from all the contexts in which the system will be required to operate. However, this is not mandatory because the learning system can adapt at a later point in time, either autonomously or via operator input. Such operator

1 The material in this paper is based upon work supported by the AFOSR under Contract No. F49620-03-C-0022.
interaction is not onerous using our approach since the required input information can be obtained from examples selected on-line by clicking tracks or by confirming/rejecting alerts. As normalcy is learned, new observations can be judged for normalcy. Those events considered unusual can then be flagged as alerts to cue human operator attention. Figure 1 illustrates the learned representation from vessel track data recorded in the Miami Harbor vicinity during August 2004. One aspect of this learned representation is worthy of note here. Panning from west to east (left to right) across the figure the potential locations of vessels become less constrained. In fact, in the east-most section of the region, the learned representation spans the location space. This is in contrast to the more specified routes evident in the learned representation from the New York harbor area reported in [1]. It should also be noted that the great majority of the learned boxes in the east-most area are uniformly pale, an indication that the pattern of travel within this area does not follow particular navigation routes. This pattern of behavior provides insight for analyzing the results of the prediction learning efforts that are the main focus of this paper.

Operators may guide learning by confirming or rejecting alerts raised by the system. Operators may also label events as threatening or innocuous to trigger learning to refine system performance. As activities or contexts change, learning proceeds in a semi-supervised fashion, benefiting from operator experience without intensive interaction. Moreover, the operator has control over the sensitivity level of system alerting to control false alarms.

3 Prediction of future vessel behavior

Event level anomalous behavior detection addresses only certain aspects of situation awareness. For instance, it may be that two events considered in isolation are normal, but their occurrence in a specific order or at a particular temporal interval may warrant closer examination on the part of an operator/analyst. Further system enhancement is necessary to address such issues. The work presented here describes extensions of our learning-based approaches to provide predictions of future vessel location given current vessel location and velocity. The time intervals for prediction are predetermined for this phase of our work and can be selected to suit the operational needs of the users. Although we present examples from a single prediction horizon in this paper (primarily to simplify the exposition), simultaneous learning over multiple horizons does not require any modifications to the approach described herein. It should be noted that the prediction horizons we are pursuing typically extend further than those useful for kinematic tracking hypotheses. The following sections first describe our approach to the problem and then present our results to date.

**Figure 1.** Example using real vessel surveillance data from the Miami harbor area after one month of normalcy model learning (based on 4 dimensions – latitude, longitude, speed, and course) for AIS vessels. **Main:** Map of the relevant region overlaid with learned representation of normal event activities as a set of shaded rectangles. Darker shading is proportional to the number of observations in a box. **Inset:** Learned representation of normal velocity (course and speed dimensions plotted in polar coordinates) of large vessels at a selected location (indicated by the red arrow). This location could be used to compare the velocity of a selected vessel with the range of normal activity for this location.
3.1 On-the-fly learning for prediction of future vessel locations

Our objective is to be able to predict the future position of a vessel given its current behavior (location and velocity). Essentially, this involves learning links between behavioral events. As such, we have leveraged our prior work on learning to discover taxonomic relationships between objects or concepts of interest [8]. It is important that the prediction learning system operates autonomously so as to not make demands on already busy operators. Also essential is that learning occurs incrementally in order to allow the system to take advantage of increasing amounts of data without having to take the system offline. An additional benefit is that the system will be able to adapt to changing behavior patterns automatically.

From recorded AIS data we utilized vessel location (latitude and longitude) and velocity (course and speed). We placed a uniform square grid over the area of interest surrounding the port of Miami so as to discretize vessel location. We also defined a discretization of vessel velocity (speed and direction, as depicted in Figure 2) that enables learning to be contextually specific to the behavior of the vessel. Thus for each vessel report, we were able to place the vessel in a grid location and give it a velocity state.

For purposes of exposition, the chosen temporal prediction horizon is 15 minutes. We buffer vessel reports with a predetermined granularity (say, every minute) for the 15 minute duration of the temporal horizon. When collected data for a vessel spans this duration learning commences, and then proceeds as further reports arrive.

Learning is based on the associative learning algorithm introduced in [8]. Weights between grid locations change via gated Hebbian learning. One set of weights is maintained for each of the velocity states illustrated in Figure 2. The set of weights in which learning takes place is determined by the velocity state of the vessel at the start of the 15 minute temporal prediction window. The velocity state at the end of the window plays no role in the learning process.

Weights between activated nodes change according to an outstar learning law [9]:

$$\Delta w_{ij} = lr \cdot x_{jk} \cdot (x_{ik} - w_{jk})$$

where \( w_{ij} \) is the connection weight from node \( j \) to node \( i \) in the \( k \text{th} \) set of weights (that corresponds to the vessel velocity state at the beginning of the prediction interval, indexed by \( k \)), \( x_{jk} \) and \( x_{ik} \) are the activations of grid locations \( j \) (location at the start of the period—the source location) and \( i \) (location at the end of the period—the target location) respectively, and \( lr \) is the learning rate (i.e., the rate at which the weight will change when nodes \( j \) and \( i \) are co-active). Given the binary activations used in the network, weights are bounded between 0 and 1 and the size of weight changes is solely dependent on the learning rate, which is set to 0.05 for the results presented here. Learning is (presynaptically) gated by activation at the source location. If this location is not active, then no connections from this location to other locations will change their weights. If the source location is active, then links with any active target locations will increase their weights and links with any inactive target locations will decrease their weights.

This form of learning has a number of attractive properties for the current application. First, more frequent combinations of source and target locations are rapidly learned (as indicated by larger weights). Second, random/infrequent combinations will cause learning when they occur but will also be unlearned (through weight decay) when they do not occur. In this way, if a future location is fairly rare from the current source location then the weights will remain quite small—a factor that can be exploited in the prediction process of our system. This property also provides noise tolerance. Third, taken together the first two properties enable the system to automatically track changes in behavior over time (e.g., due to altered operating rules or channel positions). If it is desirable to maintain multiple sets of models for alternating operating conditions, for example, to capture seasonal differences, then a relevant contextual input can be added to separate models for each level of that contextual factor. This approach has been successfully developed and deployed for the event level anomaly detection capability described in Section 2 (and in [1]). Fourth, the learning is entirely unsupervised. It requires no operator intervention.

For each vessel report the system can evaluate current location in light of prior behavior and can make predictions about future vessel location. With respect to predicting future position, weighted links from the current (source) location—and accounting for current velocity state—indicate where vessels have previously been observed to be upon expiration of the temporal horizon. Of course, there may be many non-zero weights emanating from the source location, a large fraction of which may be small (having arisen from only a few observations). This noise can be eliminated by setting a threshold as a lower bound on weight values to be considered for purposes of future location prediction. Using this mechanism enables a relatively small number of higher likelihood predictions to be made based on the current location and velocity of a given vessel. The results that follow are based on this mechanism.

It is also possible to utilize the predictions to enable an alerting functionality similar to that for the anomalous activity alerts described earlier. Consider a particular
vessel that produces a current report that triggers learning. This requires the data buffer to contain a report from that vessel when the temporal horizon window opened. This prior report can be presented to the model for the purpose of generating predicted future locations. If the current location is not among the predicted locations, then an alert can be raised. Such an alert indicates that the vessel is now somewhere it wasn’t expected to be on the basis of its earlier behavior. Although all the results presented here use models generated over all vessels that provided reports, models can be constructed for classes of vessels or even individual vessels as has been successfully demonstrated in the event-level anomaly detection capability already described. Figure 3 provides a visual illustration of the foregoing description using a screenshot from a prototype system.

3.2 Performance results

3.2.1 Procedure

Except where otherwise indicated, the following procedure was used to obtain the results presented in this section. The learning process described in the preceding section (including a learning rate of 0.05 and a prediction horizon of 15 minutes) was presented with five months worth of recorded AIS data. The recorded data was time-stamped upon acquisition and this temporal information provided the basis for data presentation to our system. As a consequence, the results reported here could have been obtained by running our system live in Miami for approximately five months. From the data stream, we sample reports at 60 second intervals: the most recent report for any vessel within the previous minute was presented to the system. After presentation of the entire five months of data was complete, the learned weights comprising the model were frozen. Then, data from a previously unseen month were presented to the system using the same 60 second sampling interval used during learning. It should be noted that this process was used for performance evaluation purposes. In contrast, the system would learn and make predictions concurrently under operational conditions.

To support calculation of performance results, the following data were collected from the presentation of each 60 second sample to the system: (1) the sample time of the set of reports—a timestamp with 60 second resolution; (2) the location and velocity state of each vessel report in the sample; (3) the grid locations predicted by the model (using a very low weight threshold so as to enable a range of weight thresholds to be examined in subsequent analysis); and (4) the actual future position of the vessel (available because data is recorded) at the end of the 15 minute prediction horizon is indicated by the diamond. Model predictions of future location are indicated by highlighted grid locations. There are three of them located centrally in the map. The strength of the weight underlying each prediction is indicated by the highlight intensity – small weights are very pale and large weights are darker. Since the actual future location falls within a predicted grid location, this example represents a hit and the vessel marker label is surrounded by asterisks. If none of the predicted grid locations contained the actual future location the ID number would be surrounded by parentheses (representing a miss)—the label of a non-selected vessel (ID 355213000) is an example of such a miss.

Figure 3. Snapshot from Miami harbor surrounds depicting system operation. The location grid is superimposed over an ENC map of the area. Current vessel location is indicated on the map by circular markers and identification numbers. One vessel (ID 31988500) has been selected for prediction display (as indicated by the larger, brighter marker). The actual future position of this vessel (available because data is recorded) at the end of the 15 minute prediction horizon is indicated by the diamond. Model predictions of future location are indicated by highlighted grid locations. There are three of them located centrally in the map. The strength of the weight underlying each prediction is indicated by the highlight intensity – small weights are very pale and large weights are darker. Since the actual future location falls within a predicted grid location, this example represents a hit and the vessel marker label is surrounded by asterisks. If none of the predicted grid locations contained the actual future location the ID number would be surrounded by parentheses (representing a miss)—the label of a non-selected vessel (ID 355213000) is an example of such a miss.
location of each vessel 15 minutes into the future (when available)—as truth data.

These results data were binned into daily sets and the following statistics were calculated for each day. The total number of vessel reports was counted; henceforth this will be referred to as the number of events (#Events). The total number of grid locations predicted was counted—the number of predictions (#Predictions). A fixed weight threshold was used to determine which predictions from the set collected would actually be used in the analysis. The process for determining that threshold will be described presently. There could be zero, one, or more predictions for each event. For each event where one or more predictions was made, the number of ‘event predictions’ (#EventPredictions) was incremented by one. Finally, whether or not the predicted grid locations for a given report contained that actual future location corresponding to that report was determined. In the case of an affirmative answer—a hit—the number of hits was incremented (#Hits).

The following daily summary metrics were calculated. Coverage is the fraction of events for which at least one prediction was made: #EventPredictions/#Events. Recall is the fraction of events for which a correct prediction was made: #Hits/#Events. Precision is the proportion of hits within the set of all predictions: #Hits/#Predictions. Finally, accuracy is the proportion of hits within the set of event predictions: #Hits/#EventPredictions.

Coverage provides a measure of how well the learning has progressed in terms of being able to make predictions for all events presented to the models. Recall and precision are standard information retrieval metrics for assessing model performance. Recall is equivalent to $P_D$ (probability of correct detection) and is an absolute measure of prediction accuracy. Precision counterbalances recall by penalizing rampant over-prediction in order to increase recall. Precision is maximal if one and only one prediction is made for each event and if that single prediction is correct. Precision is related to $P_{FA}$ (probability of false alarms). Accuracy—as defined here—is a relative measure of prediction accuracy in that it measures $P_D$ when a prediction is actually made. In contrast, recall factors in all events irrespective of whether a prediction was made or not.

To find an appropriate weight threshold upon which to base a decision about which links should constitute predictions, we performed a recall/precision sweep across a range of weight thresholds. As the weight threshold increases, fewer weights exceed the threshold so fewer predictions are made. This improves precision, but recall suffers. The crossover point, where recall and precision are equal, determines a break-even operating level that is usually a good choice for further analysis. In the present case, a weight threshold of 0.16 approximated the break-even point and was used for the remainder of the results presented here.

During our observation of the system learning and making predictions, we noted that often the actual future location of a vessel was in a grid location adjacent to a predicted location (or within two grid locations of a prediction). So, in addition to determining prediction correctness on the basis of direct hits (henceforth $d0$), we calculated recall, precision, and accuracy on the basis of near misses. In one case, $d1$, we extended each prediction grid location to include all eight adjacent locations. If the future vessel location fell within this zone, a hit was counted. In a second case, $d2$, we further extended $d1$ by including all 16 grid locations adjacent to it, making a total of 25 grid locations within which a future location could fall in order to register a hit. These more lenient criteria were for exploratory purposes to assess model performance, and are justifiable on the basis that a uniform grid size across the entire region of interest (including river, inner harbor, navigation channel, and open sea regions) is suboptimal in some part(s) of that region.

The event-level learned representation presented in Figure 1 suggests that vessel behavior differs across the region of interest. In order to further characterize the performance of the location prediction model in different parts of that region, we created four zones as illustrated in Figure 4. Essentially Zone 1 covers the Miami River in the west-most portion of the region of interest. Zone 2 covers the Miami Harbor proper, Zone 3 covers the controlled approach area east of the harbor, and Zone 4 covers open water. The issue of different vessel operating characteristics in different zones will be returned to during the discussion of the results and in our directions for future work.

### 3.2.2 Results

Table 1 presents summary statistics (mean ± standard deviation) from the daily statistics. The top set of results covers global performance across all zones in the Miami Harbor region of interest for each of the prediction hit neighborhoods $d0$, $d1$, and $d2$. It should be noted that coverage does not change with increasing prediction hit neighborhood size, but recall, precision, and accuracy all increase due the larger number of correct predictions occurring as the decision criterion is relaxed. Below the global results are those for Zones 1 and 2 (combined due to their similarity) and Zones 3 and 4 (presented in combination and separately).

Over 55% of all events in the region of interest generated at least one predicted future location. As the prediction hit neighborhood increased from $d0$ to $d2$, the accuracy of these predictions increased from about one in three to about two in three. Against all events, the rate of correct prediction increased from about one in five at $d0$ to about two in five at $d2$. Precision rates matched those of recall (which is unsurprising since the weight threshold setting was based on the global recall/precision break-even point). When evaluating the adequacy of these results, it should be taken into account that the data comprising each event report is very limited and that there is zero level of operator intervention required for model learning. However, these global results do not provide a comprehensive picture.

To gain more insight, we exploited inhomogeneity of vessel behavior in different parts of the region as identified from Figure 1. Our zone-based analysis reveals differential levels of performance. Coverage for Zones 1 and 2 is very high at about 83% of all events, which is in stark contrast to very poor coverage (less than 30%) in
Zones 3 and 4. However the latter result masks a considerable difference between Zones 3 and 4. Clearly Zone 4 performance is quite weak—a consequence of the apparently random vessel behavior that occurs in the open water east of Miami Harbor. Since the pattern is the same across prediction hit neighborhoods, we will use the most liberal neighborhood, \( d_2 \), to examine zone-based differences in recall, accuracy, and precision. A recall performance gradient decreasing from Zones 1 and 2 through Zone 3 to Zone 4 mirrors that of coverage. Accuracy exceeds 70% for Zones 1 and 2 and Zone 3, but is only 54% for Zone 4. Precision is independent of zone, a consequence of how we selected the weight threshold for determining predictions. Clearly, the more constrained vessel behaviors in Zones 1 and 2 are more readily learned by our system when compared to the essentially random looking behavior in Zone 4.

The foregoing analysis included all vessel velocity states. Figure 4 illustrates model performance when only westward and eastward velocity states are considered. Since westward tracks are approaching the harbor and approaching more heavily constrained zones, performance is substantially better than when all states are considered. Coverage is 94% for Zones 1 and 2, 90% for Zone 3, and nearly 50% for Zone 4. For the \( d_2 \) prediction hit neighborhood, recall exceeds 68% for Zones 1 through 3 and is over 31% for Zone 4. The corresponding accuracy is 72% for Zones 1 and 2, 77% for Zone 3, and almost 69% for Zone 4. These results are particularly encouraging since they represent the cases where vessels are approaching the harbor and river areas. These are precisely those cases in which higher prediction performance is most desirable for security and law enforcement purposes. When eastward velocity states are concerned, performance in Zones 1 and 2 (the river and harbor areas) is also very good. Coverage is 85% and \( d_2 \) recall and accuracy are 58% and 68% respectively. Performance degrades dramatically for Zone 3 and Zone 4, but this is perhaps not as crucial because vessels are leaving the main areas being secured.

Online incremental learning quickly produces the levels of performance described herein. Figure 5 illustrates how coverage and recall develop over increasing learning durations. For each learning duration plotted, the prediction model learned up to that point was used to make predictions (using the same data set used to produce the results already presented). These results are inclusive of all velocity states and use the \( d_2 \) prediction hit neighborhood. Both global and Zones 1 and 2 results are presented for comparative purposes. Coverage increases rapidly over the first month of learning. It continues to increase, albeit at a slower rate, over four more months of exposure to recorded data. The development of both global and Zone 1 and 2 coverage follow very similar profiles over time. Recall performance changes over time mirror those for coverage results: initial rapid improvement followed by much slowed increases in performance. Interestingly, the accuracy at each stage stays within the very narrow range around 70% as indicated in Figure 5. Our system thus exhibits rapid learning and slow refinement. It can quickly reach a relatively high level of performance and then continues to improve as it is exposed to increasing amounts of data.

### 4 Conclusions

We have described an enhancement to our learning-based maritime domain awareness system to produce predictions of future vessel location on the basis of current vessel behavior. Prediction models are learned autonomously over one or more pre-specified temporal horizons. We showed that our system learns quickly to achieve reasonable levels of performance. Incremental learning drives continued performance improvement as the system continues monitoring and evaluating vessel behavior.

### Table 1. Performance results over all velocity states.

<table>
<thead>
<tr>
<th>Zone(s)</th>
<th>Prediction Hit Neighborhood</th>
<th>Coverage #EvPreds/#Events</th>
<th>Recall #Hits/#Events</th>
<th>Precision #Hits/#Predictions</th>
<th>Accuracy #Hits/#EvPreds</th>
</tr>
</thead>
<tbody>
<tr>
<td>All</td>
<td>( d_0 )</td>
<td>0.56±0.03</td>
<td>0.20±0.04</td>
<td>0.20±0.03</td>
<td>0.36±0.05</td>
</tr>
<tr>
<td></td>
<td>( d_1 )</td>
<td>0.56±0.03</td>
<td>0.33±0.04</td>
<td>0.33±0.03</td>
<td>0.60±0.05</td>
</tr>
<tr>
<td></td>
<td>( d_2 )</td>
<td>0.56±0.03</td>
<td>0.39±0.04</td>
<td>0.39±0.03</td>
<td>0.69±0.05</td>
</tr>
<tr>
<td>1 &amp; 2</td>
<td>( d_0 )</td>
<td>0.83±0.03</td>
<td>0.35±0.06</td>
<td>0.23±0.03</td>
<td>0.41±0.06</td>
</tr>
<tr>
<td></td>
<td>( d_1 )</td>
<td>0.83±0.03</td>
<td>0.52±0.06</td>
<td>0.34±0.03</td>
<td>0.63±0.06</td>
</tr>
<tr>
<td></td>
<td>( d_2 )</td>
<td>0.83±0.03</td>
<td>0.59±0.06</td>
<td>0.38±0.03</td>
<td>0.71±0.06</td>
</tr>
<tr>
<td>3 &amp; 4</td>
<td>( d_0 )</td>
<td>0.30±0.03</td>
<td>0.07±0.02</td>
<td>0.13±0.02</td>
<td>0.22±0.05</td>
</tr>
<tr>
<td></td>
<td>( d_1 )</td>
<td>0.30±0.03</td>
<td>0.15±0.03</td>
<td>0.31±0.04</td>
<td>0.51±0.07</td>
</tr>
<tr>
<td></td>
<td>( d_2 )</td>
<td>0.30±0.03</td>
<td>0.19±0.03</td>
<td>0.40±0.03</td>
<td>0.66±0.07</td>
</tr>
<tr>
<td>3</td>
<td>( d_0 )</td>
<td>0.44±0.07</td>
<td>0.12±0.03</td>
<td>0.14±0.03</td>
<td>0.27±0.05</td>
</tr>
<tr>
<td></td>
<td>( d_1 )</td>
<td>0.44±0.07</td>
<td>0.25±0.06</td>
<td>0.31±0.03</td>
<td>0.58±0.07</td>
</tr>
<tr>
<td></td>
<td>( d_2 )</td>
<td>0.44±0.07</td>
<td>0.33±0.06</td>
<td>0.40±0.03</td>
<td>0.75±0.06</td>
</tr>
<tr>
<td>4</td>
<td>( d_0 )</td>
<td>0.21±0.03</td>
<td>0.03±0.01</td>
<td>0.11±0.04</td>
<td>0.15±0.06</td>
</tr>
<tr>
<td></td>
<td>( d_1 )</td>
<td>0.21±0.03</td>
<td>0.09±0.02</td>
<td>0.29±0.05</td>
<td>0.41±0.09</td>
</tr>
<tr>
<td></td>
<td>( d_2 )</td>
<td>0.21±0.03</td>
<td>0.11±0.03</td>
<td>0.39±0.05</td>
<td>0.54±0.09</td>
</tr>
</tbody>
</table>
Prediction performance was not uniform over the entire region of interest around Miami Harbor. Where vessel navigation was relatively constrained (in the Miami River and in the harbor proper) prediction performance was very good. Elsewhere, especially in open water, performance was quite poor. Furthermore, performance was better when vessels travel westward toward the harbor. From a port awareness and security perspective, prediction performance is best for the most important situations—when vessels are in or approaching critical infrastructure regions.

We believe the poorer performance in Zones 3 and 4 are partially due to the fine grid used in those regions. While appropriate for inner harbor and river navigation, this grid resolution was not conducive to effective learning where navigation was less constrained. To address this issue, and to improve performance against such vessel behavior, future work will investigate differential grid resolutions in different zones. We anticipate that utilization of larger and fewer grid locations in Zones 3 and 4 will facilitate the event link learning mechanism employed in this paper as the basis for making predictions. The current approach of manually defining the location discretization grids within the region of interest is suboptimal. Thus, another future research goal is to actually learn the spatial discretization on-the-fly. Finally, we will investigate performance of this system at other port sites to establish its general utility and look for other mechanisms with which to further enhance its capability to provide maritime situation awareness.

**Figure 4.** Comparison of model performance between westward versus eastward source velocity states broken down by zone (as delineated by the vertical purple lines). Performance for zones 1 and 2 is combined due to similarity. Results for each of the prediction hit neighborhoods are shown for the Zone 1 & 2 case. Coverage does not differ, but each of the other measures improves with prediction hit neighborhood for both eastward and westward velocity states. This trend also holds for the Zone 3 and the Zone 4 case (even though the results are not presented here). For an equivalent hit neighborhood, performance in Zone 1 & 2 is superior to that of Zone 3 which, in turn, is superior to performance in Zone 4. It can be seen that performance for westward states is always superior to that for eastward states. The differences are smallest for Zone 1 & 2 and increase for Zone 3 and Zone 4. Performance for eastward courses in Zone 4 reflects the almost random nature of vessel trajectories away from Miami once open water is reached.

**Figure 5.** Coverage and recall performance as learning progresses over many months. Both performance measures increase rapidly early, after which improvements are more gradual. Across the entire learning process, prediction accuracy stays within the narrow zone indicated in the bottom panel.
References


