Multinomial HMMs for Intent Recognition in Maritime Domains

Extended Abstract

Logan Carlson, Dalton Navalta
Monica Nicolescu, Mircea Nicolescu*
University of Nevada, Reno

Gail Woodward†
NASA Jet Propulsion Laboratory

ABSTRACT

The need for increased maritime security has prompted research focus on intent recognition solutions for the naval domain. We consider the problem of early classification of the hostile behavior of agents in a dynamic maritime domain and propose a solution using multinomial Hidden Markov Models (HMMs). To enable early detection of hostile behaviors, the proposed approach encodes as observable symbols the rate of change (instead of static values) for parameters relevant to the task. We discuss our implementation of a one-versus-all intent classifier using multinomial HMMs and present the results of our system on three types of hostile behaviors (ram, herd, block) and a benign behavior.

KEYWORDS

Intent Recognition; Maritime; Multinomial HMM

ACM Reference Format:

Logan Carlson, Dalton Navalta Monica Nicolescu, Mircea Nicolescu and Gail Woodward. 2019. Multinomial HMMs for Intent Recognition in Maritime Domains. In *Proc. of the 18th International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2019), Montreal, Canada, May 13–17, 2019,* IFAAMAS, 3 pages.

1 INTRODUCTION

Intent recognition is a classification task with the goal of identifying the high-level intent or goal of another agent or agents, particularly while their actions are unfolding and before their goals have been completed. The information gathered from intent recognition can then be used to understand environments and plan future actions. Challenges in hostile behavior recognition include dynamic-length behavior representations and transitions between distinct behaviors over time. To properly identify such behaviors with continuous observations, the system must be able to quickly detect transitions, while also identifying the behavior signatures. We propose a Hidden Markov Model (HMM) with a sliding observation window, which allows for both behavior predictions and transitions. In this work we focus on detecting three different hostile behaviors: ram, herd and block, as well as on the ability to infer a benign navigation pattern.

*{logancarlson, dnavalta}@nevada.unr.edu
{monica, mircea}@cse.unr.edu
†gail.m.woodward@nasa.jpl.gov

Proc. of the 18th International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2019), N. Agmon, M. E. Taylor, E. Elkind, M. Veloso (eds.), May 13–17, 2019, Montreal, Canada.

2 RELATED WORK

Charniak and Goldman [3] propose plan recognition using Bayesian inference techniques to compute posterior probabilities, though Dynamic Bayesian Networks provide a more compact representation of the observations [13]. Methods for approximate inference can be used in particular domains [9], but still do not achieve real-time performance. Probabilistic context free grammars (PCFGs) [10] have been used in interpreting video sequences [2, 6, 11, 12]. PCFGs typically require that the entire observation sequence be available. [14] uses a Bayesian network to represent PCFG parse trees, demonstrated by [8]. Pynadath and Wellman introduce probabilistic state-dependent grammars (PSDGs) [15] to integrate state and contextual information, exploiting independence properties of PDSG languages for efficient answers to recognition queries.

[18] uses Markov chains for early detection of intelligent agent behaviors. HMMs have also been used in various recognition applications [1, 4, 19, 21–24]. [17] presents an application of HMMs for early pedestrian intent recognition in road transportation. [5, 16, 20] have used sliding observation windows for recognition tasks. In maritime domains, existing research focuses on mitigating piracy, modelling the activities of vessels and evaluating countermeasures [7]. In contrast to previous methods, an HMM approach enables early recognition and efficiently handles continuous data streams for the detection of threatening behaviors in a maritime domain.

3 METHODOLOGY

The behaviors observed in maritime intent recognition can be classified as hostile and non-hostile. Specific hostile behaviors are exhibited by outside agents in the following scenarios:

- *BLOCK*: Intersecting the ship's trajectory.
- *HERD*: Approach and maintain a short distance at a specific angle to reach a desired destination.
- *RAM*: Approaching very quickly from an orthogonal direction.

Velocity, acceleration, and heading are some of the HMM parameters that enable understanding the agent's trajectory and intent.

3.1 Data Collection

The data used was generated using simulation environments from the NASA Jet Propulsion Laboratory. Considering the simulation as a Cartesian plane, the defending vessel begins at $(0, 0)$ and the potentially hostile vessel at unique coordinates in each of the four quadrants. The defending vessel is following the International Regulations for Preventing Collisions at Sea (COLREGs). The simulation collected 1000 frames of data for each scenario, providing multiple behavior executions.

Feature	Enumerations
Relative Angle	00: Facing Toward, 01: Facing Away
Delta Location	10: Moving closer, 11: Moving farther, 12 Stationary
Delta Speed	20: Decelerating, 21: Accelerating, 22: Constant
Delta Angle	30: Turning toward, 31: Turning away; 32: Constant
Delta Relative Heading	40: Increasing, 41: Decreasing, 42: Constant
CPA Time	60: Positive, 61: Negative
CPA Distance Threshold	70: Above, 71: Below

Table 1: The seven features listed with their symbols and the meaning of each symbol.

3.2 Feature Engineering

Continuous observations as positional coordinates, velocity and acceleration are collected during simulation and used to create features. Discrete features in Table 1 are codified as strings and concatenated, forming the observable symbols as HMM input.

3.3 Implementation

This sequential benefit of HMMs makes action recognition tasks straightforward because the length of a sequence is known before it is analyzed. However, intent recognition requires that we identify an intended action before it is completed, so a *sliding window* of most-recent frames is used for each inference. An inference cannot be made until the given number of frames have been gathered.

The other vessel may completely leave and reenter the range of our detection, which is managed using a threshold value (10 frames) that differentiates between erratic or temporary sensor information loss and the vessel leaving observation range completely. For each frame for which a classification will be attempted, the sliding window is tested against all of the HMMs. The classifier returns the intent associated with the HMM with the highest probability.

4 EVALUATION

Our classification system was evaluated based on 1) accuracy, 2) the number of prediction switches, 3) the number of frames before a hostile prediction is made, and 4) the number of frames before the correct behavior is detected. We trained HMMs using 276 simulated scenarios of each intent. These were used to test 8 scenarios of each intent. A varying number of internal HMM states were tested, with 5 states performing optimally. Varying sliding window lengths were tested, and 30 frames was found to be optimal. The average overall accuracy of our classifier with 5 internal states and a 30 frame sliding window was $\sim$ 67%, with individual behavior accuracy shown in Figure 1. Consider that behaviors are similar in the early stages of each intent, so some misclassification is expected.

As seen in Figure 2, the classifier is able to consistently detect the primary behaviors in less than fifty frames, where the first prediction is made at frame 29 due to the sliding window. Considering this delay, the algorithm is able to predict the behaviors within 20 predictions on average.

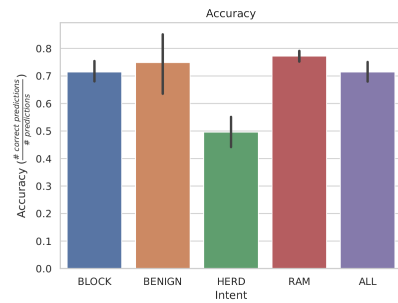

Figure 1: Accuracy averages with standard deviations across all behaviors.

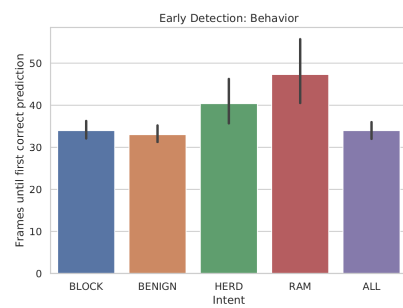

Figure 2: Bar plot of the early detection: behavior when 5-state HMMs are tested with 30-frame sliding windows.

The *Early Detection: Hostile* metric is consistently 29 frames due to the enforced sliding window. The correct behavior may not be identified, but some hostile behavior is immediately identified once the sliding window is fully populated. The number of transitions between primary behaviors for each intent provided encouraging results, with consistently less than 15 transitions for 1000 frame scenarios. The most commonly observed symbol for HERD behaviors was also common in RAM scenarios, and the second most common HERD symbol was the most common overall, explaining the low HERD accuracy and the oscillation in behavior predictions.

5 CONCLUSION

Applications of intent recognition in maritime domains are limited despite successes in other domains. HMMs offer a solution to this disparity and we present our solution to the problem of identifying hostile intents using multinomial HMMs. The models are capable of detecting multiple hostile intents during the early stages of the behaviors, giving ample time for evasive maneuvers. We achieved promising results and will continue efforts through changes to the models and the incorporation of on-water data.

6 ACKNOWLEDGEMENTS

This research has been supported by Office of Naval Research award N00014-15-1-2212.

REFERENCES

[1] Palwasha Afsar, Paulo Cortez, and Henrique Santos. 2015. Automatic Human Action Recognition from Video Using Hidden Markov Model. In *2015 IEEE 18th International Conference on Computational Science and Engineering*. 105–109. https://doi.org/10.1109/CSE.2015.41

[2] Matthew Brand. 1997. The "Inverse Hollywood Problem": From Video to Scripts and Storyboards via Causal Analysis. In *Proceedings of the Fourteenth National Conference on Artificial Intelligence and Ninth Conference on Innovative Applications of Artificial Intelligence (AAAI'97/IAAI'97)*. AAAI Press, 132–137. http://dl.acm.org/citation.cfm?id=1867406.1867427

[3] Eugene Charniak and Robert P. Goldman. 1993. A Bayesian Model of Plan Recognition. *Artif. Intell.* 64, 1 (Nov. 1993), 53–79. https://doi.org/10.1016/0004-3702(93)90060-O

[4] Ming-Hsu Cheng, Meng-Fen Ho, and Chung-Lin Huang. 2008. Gait analysis for human identification through manifold learning and HMM. *Pattern Recognition* 41, 8 (2008), 2541 – 2553. https://doi.org/10.1016/j.patcog.2007.11.021

[5] Tom Fawcett and Foster Provost. 1997. Adaptive Fraud Detection. *Data Mining and Knowledge Discovery* 1, 3 (01 Sep 1997), 291–316. https://doi.org/10.1023/A:1009700419189

[6] Y. A. Ivanov and A. F. Bobick. 2000. Recognition of visual activities and interactions by stochastic parsing. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 22, 8 (Aug 2000), 852–872. https://doi.org/10.1109/34.868686

[7] Michal Jakob, Ondřej Vaněk, Ondřej Hrstka, and Michal Pěchouček. 2012. Agents vs. Pirates: Multi-agent Simulation and Optimization to Fight Maritime Piracy. (2012), 37–44.

[8] K. M. Kitani, Y. Sato, and A. Sugimoto. 2005. Deleted Interpolation Using a Hierarchical Bayesian Grammar Network for Recognizing Human Activity. In *2005 IEEE International Workshop on Visual Surveillance and Performance Evaluation of Tracking and Surveillance*. 239–246. https://doi.org/10.1109/VSPETS.2005.1570921

[9] Neal Lesh and James Allen. 1999. Simulation-based Inference for Plan Monitoring. In *Proceedings of the Sixteenth National Conference on Artificial Intelligence and the Eleventh Innovative Applications of Artificial Intelligence Conference Innovative Applications of Artificial Intelligence (AAAI '99/IAAI '99)*. American Association for Artificial Intelligence, Menlo Park, CA, USA, 358–365. http://dl.acm.org/citation.cfm?id=315149.315325

[10] C. D. Manning and H. Scutze. 2003. *Foundations of Statistical Natural Language Processing*. The MIT Press.

[11] D. Minnen, I. Essa, and T. Starner. 2003. Expectation grammars: leveraging high-level expectations for activity recognition. In *2003 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2003. Proceedings.*, Vol. 2. II–II. https://doi.org/10.1109/CVPR.2003.1211525

[12] Darnell Moore and Irfan Essa. 2002. Recognizing Multitasked Activities from Video Using Stochastic Context-Free Grammar. In *In Proc. AAAI National Conf. on AI*. AAI, 770–776.

[13] Richard E. Neapolitan. 2003. *Learning Bayesian Networks*. Prentice-Hall, Inc., Upper Saddle River, NJ, USA.

[14] David V. Pynadath and Michael P. Wellman. 1996. Generalized Queries on Probabilistic Context-free Grammars. In *Proceedings of the Thirteenth National Conference on Artificial Intelligence - Volume 2 (AAAI'96)*. AAAI Press, 1285–1290. http://dl.acm.org/citation.cfm?id=1864519.1864577

[15] David V. Pynadath and Michael P. Wellman. 2000. Probabilistic State-dependent Grammars for Plan Recognition. In *Proceedings of the Sixteenth Conference on Uncertainty in Artificial Intelligence (UAI'00)*. Morgan Kaufmann Publishers Inc., San Francisco, CA, USA, 507–514. http://dl.acm.org/citation.cfm?id=2073946.2074005

[16] Ning Qian and Terrence J. Sejnowski. 1988. Predicting the secondary structure of globular proteins using neural network models. *Journal of Molecular Biology* 202 (1988), 865–884. Issue 4.

[17] Raúl Quintero, Ignacio Parra, J. Lorenzo, David Fernández Llorca, and Miguel Ángel Sotelo. 2017. Pedestrian intention recognition by means of a Hidden Markov Model and body language. *2017 IEEE 20th International Conference on Intelligent Transportation Systems (ITSC)* (2017), 262–268.

[18] Riccardo Sartea and Alessandro Farinelli. 2018. Detection of Intelligent Agent Behaviors Using Markov Chains. (July 2018), 2064–2066.

[19] Björn Schuller, Gerhard Rigoll, and Manfred Lang. 2003. Hidden Markov model-based speech emotion recognition. In *Acoustics, Speech, and Signal Processing, 2003. Proceedings.(ICASSP'03). 2003 IEEE International Conference on*, Vol. 2. IEEE, II–1.

[20] Terrence J. Sejnowski and Charles R. Rosenberg. 1987. Parallel Networks that Learn to Produce English Text. *Journal of Complex Systems* 1 (1987), 145–168. Issue 1.

[21] Thad Starner and Alex Pentland. 1997. *Real-Time American Sign Language Recognition from Video Using Hidden Markov Models*. Springer Netherlands, Dordrecht, 227–243. https://doi.org/10.1007/978-94-015-8935-2_10

[22] Christian Vogler and Dimitris Metaxas. 1999. Parallel Hidden Markov Models for American Sign Language Recognition. (1999).

[23] Jie Yang, Yangsheng Xu, and Chiou S Chen. 1994. Hidden markov model approach to skill learning and its application to telerobotics. *IEEE transactions on robotics and automation* 10, 5 (1994), 621–631.

[24] Byung-Jun Yoon. 2009. Hidden Markov Models and their Applications in Biological Sequence Analysis. *Current Genomics* 10 (2009), 402–415. Issue 6. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2766791/