PREDICTING COGNITIVE EMITTER BEHAVIOR WITH GAIL

Anthony Tai, Ph.D.
Code WXVM
Naval Surface Warfare Center - Crane Division
Crane, IN 47522-5001
anthony.s.tai.civ@us.navy.mil

ABSTRACT

Through interactions with the surroundings, modern cognitive radar systems learn to optimize their decision-making process to intelligently select transmission waveforms and operating parameters. Because of their waveform agility and the ability to respond dynamically, cognitive radars are difficult to track and disrupt. This work aims to explore if Generative Adversarial Imitation Learning (GAIL) can be applied to capture, imitate, and predict the behavior of cognitive radars. We study the basic principles of GAIL, explore its existing applications, and research implementation of the approach for tracking the actions of self-driving cars. We conclude with the feasibility analysis of utilizing GAIL for predicting the behavior of cognitive radar systems.

1 INTRODUCTION

Modern military radar systems equipped with cognitive ability to automate waveform selection and update waveform class parameters are elusive and increasingly difficult to track. In his article (Haykin, 2006) Haykin discussed in details his vision of cognitive radar. In addition to being able to continuously update its receiver via interacting with the environment, the cognitive radar modifies the method for its transmitter to illuminate the environment to gain relevant information. In essence, the radar system forms a closed feedback loop consisting of the environment, receiver, and transmitter. Other researchers also recognized the critical role cognitive radar plays in modern society. Enhanced designs for target detection and tracking (Bell et al., 2015; Yang et al., 2020), and spectrum sensing and prediction (Jalil et al., 2021) were proposed, promising improved tracking performance and detection accuracy with machine intelligence. Subsequently, experiments (Bell et al., 2021; Smith et al., 2015; Oechslin et al., 2017), key metrics (Butterfield et al., 2016), and testbed (Christiansen et al., 2017) were presented for quantifying the progress in cognitive radar development.

Given that modern adversary radar systems with cognitive capabilities may very likely be utilized in a hostile environment, it is crucial to study how recent advances in artificial intelligence and machine learning technology can be applied to learn their intention and predict their behavior. One reasonable approach is to consider adapting human behavior learning in a controlled environment. In particular, various techniques have been suggested for modeling human drivers to simulate realistic driving scenarios to develop advanced intelligent transportation systems. While rule-based methods depend on subject-matter-experts' inputs and assumptions (Asadi and Ghatee, 2015), imitation learning extract important information directly from data (Gindele et al., 2013). Other schemes for learning human behaviors in driving include behavioral cloning (Torabi et al., 2018), hierarchical inverse reinforcement learning (Sun et al., 2018), and deep semantic segmentation (Siam et al., 2017). In this work, we investigate the feasibility of adapting generative adversarial imitation learning (GAIL) (Ho and Ermon, 2016) for predicting the behavior of cognitive radars.

2 GENERATIVE ADVERSARIAL IMITATION LEARNING

For a model (or agent) to duplicate the behavior of a certain subject (or expert) in imitation learning, it needs to learn the policy that guides the subject's decision-making process for action in accordance with the observations it makes in the environment. Contrary to behavioral cloning (BC) which fits a supervised learning model with fixed state-action datasets, generative adversarial imitation learning (GAIL) learns a policy directly from data to generate imitated actions that are indistinguishable from expert actions. It first learns a cost function associated with the subject behavior via maximum causal entropy inverse reinforcement learning. Then reinforcement learning is applied on the cost function to characterize the policy. The core GAIL algorithm exploits generative adversarial training (Goodfellow et al., 2014) to model distributions of expert states and actions.

Generative adversarial network (GAN) models (Goodfellow et al., 2014) have been successfully used in creating realistic photographs of fake aging human faces (Antipov et al., 2017), image-to-image translation (Zhu et al., 2017), 2D to 3D shape induction (Gadelha et al., 2016), and audio waveform generation (Juvela et al., 2018). GANs are known for their ability to mimic any signal that follows a specific distribution and are well suited for generating simulated behavior that closely resembles the ground truth.

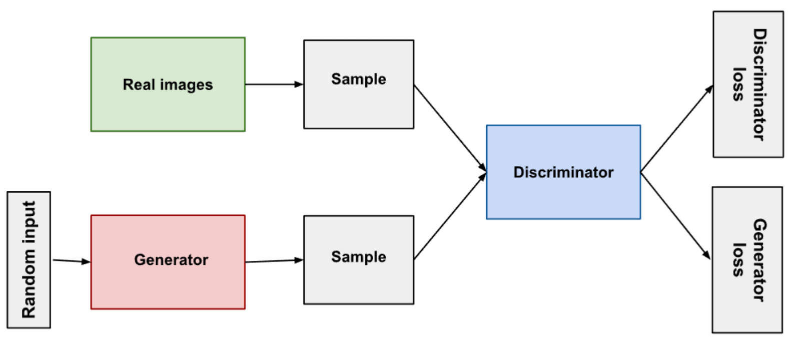

Figure 1: Source: https://developers.google.com/machine-learning/gan/gan_structure. A typical generative adversarial network structure.

Sampling the ground truth expert policy π_E and the simulated policy π_θ, we obtain the expert state-action pairs $X_E(s_t, a_t)_{t=1}^T$ and the simulated state-action pairs X_θ, respectively. Denoting $V(\theta, \psi)$ as the state-value function, D_ψ the Generative Adversarial Network (GAN) discriminator classifier parameterized by the set of parameters ψ, and $D_{\psi(s,a)}$ the output of D_ψ evaluated at state s and action a, the GAIL sigmoid cross entropy objective function is given by

$$\max_{\psi} \min_{\theta} V(\theta, \psi) = \mathbb{E}_{(s,a)\sim X_E}[log\mathcal{D}_\psi(s, a)] + \mathbb{E}_{(s,a)\sim X_\theta}[log(1 - \mathcal{D}_\psi(s, a))]$$

where $E(*)$ represents the expectation function.

The set of parameters ψ characterizing the underlying neural network can be optimized using the above equation with positive samples from X_E and negative samples produced as a result of applying the simulated policy π_θ in the simulation environment. Policy π_θ in turn is determined with a reward function defined as

$$\tilde{r}(s_t, a_t; \psi) = -(log(1 - \mathcal{D}_\psi(s, a)))$$

As sample tuples (s_t, a_t) drawn from X_θ become less distinguishable from the ground truth, the reward values pertaining to a set of policy parameters increase to infinity. Trust region policy optimization (Schulman, 2015) is then used to update the simulated policy π_θ with a state-action space approaching that associated with the ground truth policy π_E.

3 Imitating Driver Behavior with GAIL

In their work (Kuefler et al., 2017), the authors model human driving behavior on the highway using GAIL. Consider a driver following a stochastic policy $\pi(a|s)$ to take certain actions a based on road conditions s. The goal is to find a policy among a class of policies π_θ with parameters θ that approximates human driving behavior in taking sequential decisions in an environment.

The approach presented in their research learns policies through observations to generate non-linear relationship between states and actions, such as large corrections in actions resulting from small changes in the current state. Additionally, the learned policies are products of high-dimension state space corresponding to nearby vehicles and various road conditions. Furthermore, they are stochastic in the sense that different actions may be associated with a given traffic scenario.

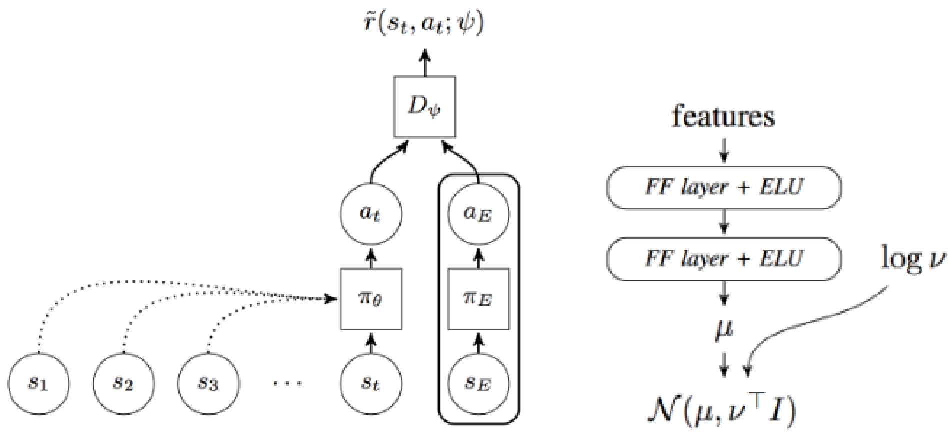

Figure 2: Left panel: GAIL architecture. Right panel: Feedforward multilayer perceptron for driving policy.

3.1 Feasibility Analysis

To evaluate the feasibility of adapting GAIL to capture, imitate, and predict the behavior of cognitive radars, we investigate similarities between generic human-driving behavior and general cognitive radar systems. We observe that firstly the system objectives for both cases are well defined. They both perform continuous interactions with the environment. Also, rapid responses to the changes in the environment are often required. Further, efficient feedback mechanisms are needed in both systems to assist informed decision-making. These are supported by constant monitoring and tracking the environment to improve behaviors. Finally, the data needed for making sequential decisions to obtain optimal actions in accordance with the environment conditions are typically received from various levels of sensors in different formats.

Given these strong similarities between learning human-driver behavior and learning cognitive radar system trajectories, we believe that generative adversarial imitation learning may be a good tool for imitating and predicting the behavior of cognitive radar systems.

4 Future Work

Generative AI has achieved many successes in computer vision and natural language processing applications while reinforcement learning has been utilized in training models to learn tasks from scratch. Our future work includes developing and implementing an algorithm based on GAIL and deep reinforcement learning to simulate realistic war-gaming scenarios We would also like to evaluate the feasibility of applying GAIL in a live-virtual-constructive (LVC) environment.

ACKNOWLEDGMENTS

The author gratefully acknowledges the financial support provided by the 2022 Naval Innovative Science & Engineering (NISE) program at the Naval Surface Warfare Center - Crane Division.

REFERENCES

Grigory Antipov, Moez Baccouche, and Jean-Luc Dugelay. Face aging with conditional generative adversarial networks. In *2017 IEEE International Conference on Image Processing (ICIP)*, pages 2089–2093, September 2017. doi: 10.1109/ICIP.2017.8296650. ISSN: 2381-8549.

Reza Asadi and Mehdi Ghatee. A Rule-Based Decision Support System in Intelligent Hazmat Transportation System. *IEEE Transactions on Intelligent Transportation Systems*, 16(5):2756–2764, October 2015. ISSN 1558-0016. doi: 10.1109/TITS.2015.2420993. Conference Name: IEEE Transactions on Intelligent Transportation Systems.

Kristine Bell, Chris Kreucher, and Muralidhar Rangaswamy. An Evaluation of Task and Information Driven Approaches for Radar Resource Allocation. In *2021 IEEE Radar Conference (RadarConf21)*, pages 1–6, May 2021. doi: 10.1109/RadarConf2147009.2021.9455168. ISSN: 2375-5318.

Kristine L. Bell, Christopher J. Baker, Graeme E. Smith, Joel T. Johnson, and Muralidhar Rangaswamy. Cognitive Radar Framework for Target Detection and Tracking. *IEEE Journal of Selected Topics in Signal Processing*, 9(8):1427–1439, December 2015. ISSN 1941-0484. doi: 10.1109/JSTSP.2015.2465304. Conference Name: IEEE Journal of Selected Topics in Signal Processing.

Aaron S. Butterfield, Adam E. Mitchell, Graeme E. Smith, Kristine L. Bell, and Muralidhar Rangaswamy. Metrics for quantifying cognitive radar performance. In *2016 CIE International Conference on Radar (RADAR)*, pages 1–4, October 2016. doi: 10.1109/RADAR.2016.8059298.

Jonas Myhre Christiansen, Graeme E. Smith, and Karl E. Olsen. USRP based cognitive radar testbed. In *2017 IEEE Radar Conference (RadarConf)*, pages 1115–1118, May 2017. doi: 10.1109/RADAR.2017.7944371. ISSN: 2375-5318.

Matheus Gadelha, Subhransu Maji, and Rui Wang. 3D Shape Induction from 2D Views of Multiple Objects, December 2016. URL `http://arxiv.org/abs/1612.05872`. arXiv:1612.05872 [cs].

Tobias Gindele, Sebastian Brechtel, and Rüdiger Dillmann. Learning context sensitive behavior models from observations for predicting traffic situations. In *16th International IEEE Conference on Intelligent Transportation Systems (ITSC 2013)*, pages 1764–1771, October 2013. doi: 10.1109/ITSC.2013.6728484. ISSN: 2153-0017.

Ian Goodfellow, Jean Pouget-Abadie, Mehdi Mirza, Bing Xu, David Warde-Farley, Sherjil Ozair, Aaron Courville, and Yoshua Bengio. Generative adversarial nets. In *Advances in neural information processing systems*, pages 2672–2680, 2014.

S. Haykin. Cognitive radar: a way of the future. *IEEE Signal Processing Magazine*, 23(1):30–40, January 2006. ISSN 1558-0792. doi: 10.1109/MSP.2006.1593335. Conference Name: IEEE Signal Processing Magazine.

Jonathan Ho and Stefano Ermon. Generative Adversarial Imitation Learning. In *Advances in Neural Information Processing Systems*, volume 29. Curran Associates, Inc., 2016. URL `https://papers.nips.cc/paper/2016/hash/cc7e2b878868cbae992d1fb743995d8f-Abstract.html`.

Syed Qaisar Jalil, Stephan Chalup, and Mubashir Husain Rehmani. Cognitive Radio Spectrum Sensing and Prediction Using Deep Reinforcement Learning. In *2021 International Joint Conference on Neural Networks (IJCNN)*, pages 1–8, Shenzhen, China, July 2021. IEEE. ISBN 978-1-66543-900-8. doi: 10.1109/IJCNN52387.2021.9533497. URL `https://ieeexplore.ieee.org/document/9533497/`.

Lauri Juvela, Bajibabu Bollepalli, Junichi Yamagishi, and Paavo Alku. Waveform generation for text-to-speech synthesis using pitch-synchronous multi-scale generative adversarial networks, October 2018. URL `http://arxiv.org/abs/1810.12598`. arXiv:1810.12598 [cs, eess, stat].

Alex Kuefler, Jeremy Morton, Tim Wheeler, and Mykel Kochenderfer. Imitating driver behavior with generative adversarial networks. In *2017 IEEE Intelligent Vehicles Symposium (IV)*, pages 204–211, June 2017. doi: 10.1109/IVS.2017.7995721.

R. Oechslin, U. Aulenbacher, K. Rech, S. Hinrichsen, S. Wieland, and P. Wellig. Cognitive radar experiments with CODIR. In *International Conference on Radar Systems (Radar 2017)*, pages 1–6, October 2017. doi: 10.1049/cp.2017.0386.

John Schulman. Trust Region Policy Optimization. *ICML*, 2015.

Mennatullah Siam, Sara Elkerdawy, Martin Jagersand, and Senthil Yogamani. Deep semantic segmentation for automated driving: Taxonomy, roadmap and challenges. In *2017 IEEE 20th International Conference on Intelligent Transportation Systems (ITSC)*, pages 1–8, October 2017. doi: 10.1109/ITSC.2017.8317714. ISSN: 2153-0017.

Graeme E. Smith, Zach Cammenga, Adam Mitchell, Kristine L. Bell, Murali Rangaswamy, Joel T. Johnson, and Christopher J. Baker. Experiments with cognitive radar. In *2015 IEEE 6th International Workshop on Computational Advances in Multi-Sensor Adaptive Processing (CAMSAP)*, pages 293–296, December 2015. doi: 10.1109/CAMSAP.2015.7383794.

Liting Sun, Wei Zhan, and Masayoshi Tomizuka. Probabilistic Prediction of Interactive Driving Behavior via Hierarchical Inverse Reinforcement Learning. *arXiv:1809.02926 [cs, stat]*, September 2018. URL `http://arxiv.org/abs/1809.02926`. arXiv: 1809.02926.

Faraz Torabi, Garrett Warnell, and Peter Stone. Behavioral Cloning from Observation. *arXiv:1805.01954 [cs]*, May 2018. URL `http://arxiv.org/abs/1805.01954`. arXiv: 1805.01954 version: 2.

Chaoqun Yang, Xiaofeng Wang, Heng Zhang, and Yu Zheng. Waveform Design For Track-Before-Detect-Based Cognitive Radars. In *2020 IEEE 11th Sensor Array and Multichannel Signal Processing Workshop (SAM)*, pages 1–5, June 2020. doi: 10.1109/SAM48682.2020.9104384. ISSN: 2151-870X.

Jun-Yan Zhu, Taesung Park, Phillip Isola, and Alexei A. Efros. Unpaired Image-To-Image Translation Using Cycle-Consistent Adversarial Networks. pages 2223–2232, 2017. URL `https://openaccess.thecvf.com/content_iccv_2017/html/Zhu_Unpaired_Image-To-Image_Translation_ICCV_2017_paper.html`.