Proceedings of the 4th International Conference on Statistics: Theory and Applications (ICSTA'22)
Prague, Czech Republic – July 28- 30, 2022
Paper No. 112
DOI: 10.11159/icsta22.112

Reinforcement Learning Based Optimal Adversarial Pathway Estimation Using Remotely Sensed Spectral-Terrain Data and Human Value Assessment

Josef Affourtit[1], and Nicholas Scott[2]
[1]Improbable LLC
4100 N. Fairfax Dr., Suite 500, Arlington, VA, 22203, USA
josef.affourtit@gmail.com; nscott@riversideresearch.org
[2]Riverside Research Institute
2640 Hibiscus Way, Beavercreek, OH, 45431, USA

Extended Abstract

Geo-intelligence organizations are often faced with the need to determine optimal pathways that adversaries may take based on various types of information including remotely sensed imagery and human geo-intelligence. The mobile enemy problem, where the objective is to predict the pathway that a mobile enemy may take, is considered here as a way to develop a statistical/signal processing formulism to assist leadership in making better decisions about how to estimate the whereabouts of an adversary. A two-tier processing pipeline utilizing feature extraction and reinforcement learning-based optimal pathway estimation was created to demonstrate how human/machine learning teaming can be exploited to address a geo-intelligence problem. The information used in the processor development consists of an open-source hyperspectral imagery (HSI) data set [1].

A strip map of terrain HSI was divided into 32 x 32 pixel image chips where principal component analysis [2] was used to reduce the dimension and decrease the noise of the hyperspectral signatures. Spectral dictionary endmembers [3] were estimated from the denoised HSI data using the unsupervised learning algorithms of k-means clustering [4] and automatic target generator processing [3]. This substage was necessary in order to perform image chip value estimation. In this evaluation stage, five different algorithms were used to calculate different value fields. Each technique used a feature extraction method designating the relative value of each image chip comprising the complete HSI scene. The first algorithm used for HSI image chip value estimation consisted of abundance estimation via nonnegative constrained least squares matched filtration [3] along with a Mahalanobis distance metric [5]. The second algorithm used was abundance estimation via orthogonal matching pursuit [6] along with a Euclidean distance metric. The final three algorithmic methods consisted of threshold-based spatial HSI spectral gradient estimation, Laplacian eigenmap kurtosis [7] estimation, and finally human value assessment of HSI image chips.

Value field estimation for each of the five algorithms was quantitatively possible but HSI data noise necessitated robust ways to obtain smoother and physically sensible value fields. A linear combination of all five value fields was proposed as a way to accomplish this. However, since human value assignment is often considered the most important value field component in many geo-intelligence problems, it was used in the development of the second processing part - optimal pathway estimation.

Q-learning and State-Action-Reward-State-Action (SARSA) learning [8,9], trial and error reinforcement learning algorithms fueled by human assigned value fields, were used to estimate optimal pathways that an adversary would take over the HSI scene. Human assigned reward fields delineate useful thoroughfares for reaching a goal state providing a utility function Q(s,a) representing the discounted cumulative reward for taking a specific action from a specific state. Q-learning and SARSA learning both utilized the maximum action policy and Dijkstra's algorithm [10] to estimate the optimal pathway from the Q function which quantifies how good an action is given a certain state. Preliminary results show that Q-learning and the maximum action policy with a learning rate of $\gamma = 0.8$ provides an agent or adversarial pathway which varies widely over the HSI scene. Q-learning used with Dijkstra's algorithm causes the optimal pathway to vary less widely over the HSI data consistent with the least step principle policy. SARSA learning used in conjunction with Dijkstra's algorithm shows a similar trend of less variation over the scene but with an optimal pathway which does not

strictly follow the peaks of reward field. The explicit reason for this is not clear. Preliminary results suggest that the human-computational formulism is a viable platform for future development of a robust geo-intelligence processor.

References

[1] D. Snyder, J. Kerekes, I. Fairweather, R. Crabtree, J. Shive, and S. Hager, "Development of a Web-based Application to Evaluate Target Finding Algorithms," *Proceedings of the 2008 IEEE International Geoscience and Remote Sensing Symposium (IGARSS)*, Boston, MA, 2008, vol. 2, pp. 915-918.

[2] W. L. Martinez, and A. R. Martinez, *Computational Statistics Handbook with Matlab*. London, UK: Chapman and Hall/CRC, 2001.

[3] Chein-I Chang, *Hyperspectral Data Processing: Algorithms Design and Analysis.* Hoboken, NJ: John Wiley and Sons, Inc., 2013.

[4] S. Theodoridis, A. Pikrakis, K. Koutroumbas, and D. Cavouras, *Introduction to Pattern Recognition: A Matlab Approach*. Oxford, UK: Academic Press, 2010.

[5] S. Theodoridis, and K. Koutroumbas, *Pattern Recognition, Second Edition*. San Diego, CA: Academic Press, 2003.

[6] J. J. Thiagarajan, K. N. Ramamurthy, P. Turaga, and A. Spanias, *Image Understanding Using Sparse Representations*. Williston, VT: Morgan and Claypool Publishers, 2014.

[7] A.Vathy-Fogarassy, and J. Abonyi, *Graph-Based Clustering and Data Visualization Algorithms.* London, UK: Springer, 2013.

[8] S. Marsland, *Machine Learning: An Algorithmic Perspective, Second Edition*. Boca Raton, FL: Chapman and Hall/CRC, 2015.

[9] C. Szepesvari, *Algorithms for Reinforcement Learning.* Williston, VT: Morgan and Claypool Publishers, 2010.

[10] H. O.-Arranz, D. R. Llanos, and A. G.-Escribano. *The Shortest-Path Problem, Analysis and Comparison of Methods*. Williston, VT: Morgan and Claypool Publishers, 2014.

.