COLLEGE OF ENGINEERING AND COMPUTER SCIENCE
FLORIDA ATLANTIC UNIVERSITY

Announces the Ph.D. Dissertation Defense of

Joffrey Leevy

for the degree of Doctor of Philosophy (Ph.D.)

"Machine Learning Algorithms for Predicting Botnet Attacks in IoT Networks"

April 18, 2022, 10:30 a.m.
Virtual Dissertation

Zoom
Meeting ID: 804 093 1436
Passcode: dissert101

DEPARTMENT:
Electrical Engineering and Computer Science

ADVISOR:
Taghi M. Khoshgoftaar, Ph.D.

PH.D. SUPERVISORY COMMITTEE:
Taghi M. Khoshgoftaar, Ph.D., Chair
Xingquan Zhu, Ph.D.
DingDing Wang, Ph.D.
Mehrdad Nojoumian, Ph.D.

ABSTRACT OF DISSERTATION
Machine Learning Algorithms for Predicting Botnet Attacks in IoT Networks

The proliferation of *Internet of Things* (IoT) devices in various networks is being matched by an increase in related cybersecurity risks. To help counter these risks, big datasets such as Bot-IoT were designed to train machine learning algorithms on network-based intrusion detection for IoT devices. From a binary classification perspective, there is a high-class imbalance in Bot-IoT between each of the attack categories and the normal category, and also between the combined attack categories and the normal category. Within the scope of predicting botnet attacks in IoT networks, this dissertation demonstrates the usefulness and efficiency of novel machine learning methods, such as an easy-to-classify method and a unique set of ensemble feature selection techniques. Apart from this work, there are no published studies that focus solely on the individual Bot-IoT categories of attack, namely, *Denial-of-Service* (DoS), *Distributed Denial-of-Service* (DDoS), Reconnaissance, and Information Theft attacks. Since resources and services become inaccessible during DoS and DDoS attacks, this interruption is costly to an organization in terms of both time and money. Reconnaissance attacks often signify the first stage of a cyberattack and preventing them from occurring usually means the end of the intended cyberattack. Information Theft attacks not only erode consumer confidence but may also compromise intellectual property and national security. For the experiment with DoS and DDoS, all the ensemble feature selection techniques, when compared to using no feature selection approach, led to an improvement in performance. Regarding the Reconnaissance experiment, only one of the ensemble feature selection techniques proved to be better than using no feature selection method. In relation to the Information Theft experiment, the ensemble feature selection techniques did not affect performance, positively or negatively. However, the ensemble feature selection approach is recommended because feature reduction eases computational burden and may provide clarity through improved data visualization. For the full Bot-IoT big dataset, an explainable machine learning approach was taken using the Decision Tree classifier. An easy-to-learn Decision Tree model for predicting attacks was obtained with only three features, which is a significant result for big data.

BIOGRAPHICAL SKETCH
Born in Roseau, Dominica
B.S., University of the West Indies, Saint Augustine, Trinidad, 1994
M.S., Nova Southeastern University, Davie, Florida, 2012
Ph.D., Florida Atlantic University, Boca Raton, Florida, 2022

CONCERNING PERIOD OF PREPARATION
& QUALIFYING EXAMINATION

Time in Preparation: 2017 - 2022

Qualifying Examination Passed: Fall 2017

Published Papers:

Leevy, Joffrey L., Taghi M. Khoshgoftaar, Richard A. Bauder, and Naeem Seliya. "A survey on addressing high-class imbalance in big data." Journal of Big Data 5, no. 1 (2018): 1-30.

Leevy, Joffrey L., Taghi M. Khoshgoftaar, Richard A. Bauder, and Naeem Seliya. "The effect of time on the maintenance of a predictive model." In 2019 18th IEEE International Conference On Machine Learning And Applications (ICMLA), pp. 1891-1896. IEEE, 2019.

Hasanin, Tawfiq, Taghi M. Khoshgoftaar, and Joffrey L. Leevy. "A comparison of performance metrics with severely imbalanced network security big data." In 2019 IEEE 20th International Conference on Information Reuse and Integration for Data Science (IRI), pp. 83-88. IEEE, 2019.

Leevy, Joffrey L., and Taghi M. Khoshgoftaar. "A Short Survey of LSTM Models for De-identification of Medical Free Text." In 2020 IEEE 6th International Conference on Collaboration and Internet Computing (CIC), pp. 117-124. IEEE, 2020.

Leevy, Joffrey L., and Taghi M. Khoshgoftaar. "A survey and analysis of intrusion detection models based on cse-cic-ids2018 big data." Journal of Big Data 7, no. 1 (2020): 1-19.

Leevy, Joffrey L., Taghi M. Khoshgoftaar, and Jared M. Peterson. "Mitigating class imbalance for iot network intrusion detection: a survey." In 2021 IEEE Seventh International Conference on Big Data Computing Service and Applications (BigDataService), pp. 143-148. IEEE, 2021.

Leevy, Joffrey L., John Hancock, Taghi M. Khoshgoftaar, and Naeem Seliya. "Iot reconnaissance attack classification with random undersampling and ensemble feature selection." In 2021 IEEE 7th International Conference on Collaboration and Internet Computing (CIC), pp. 41-49. IEEE, 2021.

Leevy, Joffrey L., John Hancock, Taghi M. Khoshgoftaar, and Jared Peterson. "Detecting Information Theft Attacks in the Bot-IoT Dataset." In 2021 20th IEEE International Conference on Machine Learning and Applications (ICMLA), pp. 807-812. IEEE, 2021.

Salekshahrezaee, Zahra, Joffrey L. Leevy, and Taghi M. Khoshgoftaar. "Feature Extraction for Class Imbalance Using a Convolutional Autoencoder and Data Sampling." In 2021 IEEE 33rd International Conference on Tools with Artificial Intelligence (ICTAI), pp. 217-223. IEEE, 2021.

Leevy, Joffrey L., John Hancock, Taghi M. Khoshgoftaar, and Jared M. Peterson. "IoT information theft prediction using ensemble feature selection." Journal of Big Data 9, no. 1 (2022): 1-48.