COLLEGE OF ENGINEERING AND COMPUTER SCIENCE
FLORIDA ATLANTIC UNIVERSITY

Announces the Ph.D. Dissertation Defense of

Richard Zuech

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

"Machine Learning Algorithms for the Detection and Analysis of Web Attacks"

November 1, 2021, 10:30 a.m.
Virtual Dissertation

Zoom
Meeting ID: 978 7337 3152
Passcode: 811913

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.
Bassem Alhalabi, Ph.D.
Hanqi Zhuang, Ph.D.

ABSTRACT OF DISSERTATION
Machine Learning Algorithms for the Detection and Analysis of Web Attacks

The Internet has provided humanity with many great benefits, but it has also introduced new risks and dangers. E-commerce and other web portals have become large industries with big data. Criminals and other bad actors constantly seek to exploit these web properties through web attacks. Being able to properly detect these web attacks is a crucial component in the overall cybersecurity landscape. Machine learning is one tool that can assist in detecting web attacks. However, properly using machine learning to detect web attacks does not come without its challenges. Classification algorithms can have difficulty with severe levels of class imbalance. Class imbalance occurs when one class label disproportionately outnumbers another class label. For example, in cybersecurity, it is common for the negative (normal) label to severely outnumber the positive (attack) label. Another difficulty encountered in machine learning is models can be complex, thus making it difficult for even subject matter experts to truly understand a model's detection process. Moreover, it is important for practitioners to determine which input features to include or exclude in their models for optimal detection performance. This dissertation studies machine learning algorithms in detecting web attacks with big data. Severe class imbalance is a common problem in cybersecurity, and mainstream machine learning research does not sufficiently consider this with web attacks. Our research first investigates the problems associated with severe class imbalance and rarity. Rarity is an extreme form of class imbalance where the positive class suffers extremely low positive class count, thus making it difficult for the classifiers to discriminate. In reducing imbalance, we demonstrate random undersampling can effectively mitigate the class imbalance and rarity problems associated with web attacks. Furthermore, our research introduces a novel feature popularity technique which produces easier to understand models by only including the fewer, most popular features. Feature popularity granted us new insights into the web attack detection process, even though we had already intensely studied it. Even so, we proceed cautiously in selecting the best input features, as we determined that the "most important" Destination Port feature might be contaminated by lopsided traffic distributions.

BIOGRAPHICAL SKETCH
Born in Milwaukee, Wisconsin, USA
B.S., Georgia Institute of Technology, Atlanta, Georgia, 1996
M.S., Florida Atlantic University, Boca Raton, Florida, 1999
Ph.D., Florida Atlantic University, Boca Raton, Florida, 2021

CONCERNING PERIOD OF PREPARATION
& QUALIFYING EXAMINATION

Time in Preparation: 2013 - 2021

Qualifying Examination Passed: Fall 2012

Selected Published Papers:

R. Zuech, J. Hancock, and T. M. Khoshgoftaar. "Investigating Rarity in Web Attacks with Ensemble Learners." Journal of Big Data, vol. 8, no. 1, 2021, pages 1–27.

R. Zuech, J. Hancock, and T. M. Khoshgoftaar. "Detecting Web Attacks in Severely Imbalanced Network Traffic Data." Information Reuse and Integration for Data Science (IRI), 2021 IEEE 22nd International Conference on. IEEE, 2021, pages 267–273.

R. Zuech, J. Hancock, and T. M. Khoshgoftaar. "Detecting SQL Injection Web Attacks Using Ensemble Learners and Data Sampling." Cyber Security and Resilience (CSR), 2021 IEEE International Conference on. IEEE, 2021, pages 27–34.

R. Zuech, J. Hancock, and T. M. Khoshgoftaar. "Feature Popularity Between Different Web Attacks with Supervised Feature Selection Rankers." Machine Learning and Applications (ICMLA), 2021 20th IEEE International Conference on. IEEE, December 2021.

J. L. Leevy, J. Hancock, R. Zuech, and T. M. Khoshgoftaar. "Detecting Cybersecurity Attacks Using Different Network Features with LightGBM and XGBoost Learners." Cognitive Machine Intelligence (CogMI), 2020 IEEE Second International Conference on. IEEE, 2020, pages 190–197.