Department of Computer Science
11000 University Parkway
Pensacola, FL 32514 -5750

March 31, 2025

Nomination Statement

Title of Paper: Multiclass SVM: OvO vs. OvA on UWF-ZeekDataFall22 Dataset

Rocio Krebs[1*], Sikha S. Bagui[1**], Dustin Mink[2], Subhash C. Bagui[3]
[1]Department of Computer Science
[2]Department of Cybersecurity
[3]Department of Mathematics and Statistics
University of West Florida, Pensacola, FL USA

***Student author*
***Corresponding author*

This paper, with a lead student co-author, and three faculty members the Departments of Computer Science, Cybersecurity and Mathematics/Statistics respectively, lies at the confluence of these three fields, to create the ideal data science team. Published in October of 2024 in *Electronics*, a peer-reviewed Scopus indexed journal, the paper has already started gaining attention by researchers, attested by the fact that it has already started receiving citations. Our student co-author, part of our Cyber Analytics Research Group (CAR) at the University of West Florida, is one of the many students we have guided over the past few years – for more informaiton on the work of CAR, please visit datasets.uwf.edu

Uniqueness of our Work (Paper):

Effective multi-class classification is a major problem when dealing with network intrusion detection. This paper addresses the technical challenges of applying Support Vector Machines (SVM) for multi-class classification in network intrusion detection using the UWF-ZeekDataFall22 dataset, a newly created modern network attack dataset, created using crowd-sourced Zeek log data, labeled based on the MITRE ATT&CK framework. This is the first dataset available leveraging the MITRE ATT&CK framework.

In this work, Zeek, an open-source traffic analyzer, that provides text versus binary based network data needed to tackle today's toughest network challenges in the enterprise, cloud, and industrial computing environments, was used to collect data. This one-of-a-kind new modern dataset, *UWF-ZeekData22*, was labeled using the MITRE Adversarial Tactics, Techniques, and Common Knowledge (ATT&CK) framework. The MITRE ATT&CK framework is a globally accessible knowledge base of adversary tactics and techniques used to accomplish specific objectives. This work, specifically using Zeek's Connection (Conn) Log files from the *UWF-ZeekData22* dataset, identifies the MITRE ATT&CK adversary tactics, Reconnaissance, Discovery, Resource Development, Defense Evasion, Initial Access, Command and Control, Lateral Movement, Persistence and Collection.

Another key challenge lies in handling imbalanced classes in complex attack patterns, which is inherent in network intrusion detection data. This work addresses the difficulties in implementing SVMs for multi-class classification using One-vs.-One (OvO) and One-vs.-All (OvA) methods. Issues discussed are: (i) scalability, due to the large volume of network traffic logs; (ii) the tendency of SVM to be sensitive to noisy data and class imbalance. Preprocessing techniques were applied to improve feature selection and reduce noise in the data, and SMOTE was used to address the class imbalance issue. The unique structure of network traffic data, with overlapping patterns between attack vectors, posed significant challenges in achieving accurate classification. Our model reached an accuracy of over 90% with OvO and over 80% with OvA, demonstrating that despite these challenges, multi-class SVMs can be effectively applied to complex intrusion detection tasks when combined with appropriate balancing and preprocessing techniques.

Anthony G Pinto

Anthony G Pinto
Department of Computer Science
University of West Florida
Pensacola, FL 32514