AN ARGMAX ONE-VS-ALL APPROACH FOR MULTI-CLASS ANOMALY-BASED NETWORK INTRUSION DETECTION SYSTEM

OWOKA, EMMANUEL OLUSOLA
(20PCG02184)

AUGUST, 2022

AN ARGMAX ONE-VS-ALL APPROACH FOR MULTI-CLASS ANOMALY-BASED NETWORK INTRUSION DETECTION SYSTEM

BY

OWOKA, EMMANUEL OLUSOLA
(20PCG02184)
B.Sc Computer Science, University of Benin, Benin-City

A DISSERTATION SUBMITTED TO THE SCHOOL OF POSTGRADUATE STUDIES IN PARTIAL FULFILMENT OF THE REQUIREMENTS FOR THE AWARD OF MASTER OF SCIENCE (M.Sc) DEGREE IN MANAGEMENT INFORMATION SYSTEMS IN THE DEPARTMENT OF COMPUTER AND INFORMATION SCIENCES, COLLEGE OF SCIENCE AND TECHNOLOGY, COVENANT UNIVERSITY.

AUGUST, 2022

ACCEPTANCE

This is to attest that this dissertation is accepted in partial fulfilment of the requirements for the award of the degree of MASTER of Sciences in Management Information Systems in the Department of Computer and Information Sciences, College of Science and Technology, Covenant University, Ota, Nigeria.

Mr. Taiwo B. Erewunmi
(Secretary, School of Postgraduate Studies) **Signature and Date**

Prof Akan B. Williams
(Dean, School of Postgraduate Studies) **Signature and Date**

DECLARATION

I, **OWOKA, EMMANUEL OLUSOLA (20PCG02184),** declare that this research was carried out by me under the supervision of Dr. Aderonke A. Oni of the Department of Computer and Information Sciences, College of Science and Technology, Covenant University, Ota, Nigeria. I attest that the dissertation has not been presented either wholly or partially for the award of any degree elsewhere. All sources of data and scholarly information used in this dissertation are duly acknowledged.

OWOKA, EMMANUEL OLUSOLA

Signature and Date

CERTIFICATION

We certify that this dissertation titled **“AN ARGMAX ONE-VS-ALL APPROACH FOR MULTI-CLASS ANOMALY-BASED NETWORK INTRUSION DETECTION SYSTEM”** is an original research work carried out by **OWOKA, EMMANUEL OLUSOLA (20PCG02184)** in the Department of Computer and Information Sciences, College of Science and Technology, Covenant University, Ota, Nigeria under the supervision of Dr. Aderonke A. Oni. We have examined and found this work acceptable as part of the requirements for the award of Master of Science in Management Information Systems.

Dr. Aderonke A. Oni
(Supervisor) **Signature and Date**

Prof. Olufunke O. Oladipupo
(Head of Department) **Signature and Date**

Prof. Olufunke R. Vincent
(External Examiner) **Signature and Date**

Prof. Akan B. Williams
(Dean, School of Postgraduate Studies) **Signature and Date**

DEDICATION

I dedicate this work to the Almighty God, for His infinite wisdom, grace, and love over my life. Also, this work is dedicated to my loving parents who have both worked exceptionally hard to set me up for success.

ACKNOWLEDGEMENTS

I appreciate God almighty for the strength, grace, wisdom and understanding to execute this study. Many thanks to the head of department, Prof. Olufunke Oladipupo, and the entire members of the faculty for this opportunity. I appreciate my supervisor, Dr. Aderonke Oni for her guidance, support, and encouragement. I also appreciate my family, friends and colleagues for their help and support.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS

AI	Artificial Intelligence
A-NIDS	Anomaly Network Intrusion Detection Systems
AWS CLI	Amazon Web Service Command Line Interface
BiDLSTM	Bidirectional Long Short-Term Memory
CES-CIC-IDS	Communications Security Establishment and the Canadian Institute for Cybersecurity Intrusion Detection System
CNN	Convolution Neural Network
CRISP-DM	CRoss Industry Standard Process for Data Mining
CSV	Comma Separated Value
DBN	Deep Belief Network
DDoS	Distributed Denial of Service
DL	Deep Learning
DNN	Deep Neural Network
DoS	Denial of Service
DVWA	Damn Vulnerable Web App
EFC	Energy-Based Flow Classification
ELM	Extreme Learning Machine
EOT	Edge-of-Things
FAR	False Alarm Rate
FN	False Negative
FP	False Positive
FTP	File Transfer Protocol
HOIC	High Orbit Ion Cannon
IDS	Intrusion Detection System
Interpol	International Criminal Police Organization
IoT	Internet of Things

IR	Imbalance Ratio
JSON	JavaScript Object Notation
KNN	K-Nearest Neighbours
LASSO	Least Absolute Shrinkage and Selection Operator
LightGBM	Light Gradient Boosting Machine
LSTM	Long Short-Term Memory
ML	Machine Learning
MLP	Multilayer Perceptron
MSE	Mean Squared Error
NB	Naïve Bayes
NIDS	Network Intrusion Detection
NIST	National Institute of Standards and Technology
NMUS	NearMiss Under-sampling
NN	Neural Network
PCA	Principal Component Analysis
PKI	Public Key Infrastructure
R^2	Coefficient of Determination
R2L	Root to Local
RBM	Restricted Boltzmann Machines
RF	Random Forest
RNN	Recurrent Neural Network
SC	Silhouette Coefficient
SCADA	Supervisory Control and Data Acquisition
SFSDT	Sequence Forward Selection algorithm with Decision Tree
SMO	Sequential Minimal Optimization
SMOTE	Synthetic Minority Over-sampling Technique
SSH	Secure Shell

SVM	Support Vector Machine
TN	True Negative
TP	True Positive
U2R	User to Root
WELM	Weighted Extreme Learning Machine
XGBoost	eXtreme Gradient Boosting
XML	eXtensible Markup Language
XSS	Cross-Site Scripting

ABSTRACT

The internet is advancing at a fast pace, and it is very essential to individuals and organizations. Also, there are a lot of malicious actors on the internet and a successful attack on a victim can be very devastating. Hence, the growing need for cybersecurity. Network security helps protect computer networks from attackers and this can be achieved with the help of intrusion detection systems (IDS). Over the years researchers have proposed improvements to IDSs, however, the problem of low detection rate especially towards minority classes within the available datasets plagues the research area. This study builds and evaluates an ensemble anomaly-based network intrusion detection system for multi-class classification using an argmax one-vs-all approach. The Communications Security Establishment and the Canadian Institute for Cybersecurity Intrusion Detection System 2018 dataset (CSE-CIC-IDS2018), referred to as CICIDS2018, was used in this study. The eXtreme Gradient Boosting (XGBoost) was used for feature selection and the Minority Oversampling Technique (SMOTE) alongside cost-sensitive learning were utilized to address the imbalanced nature of the CICIDS2018 dataset. The Multilayer Perceptron (MLP), Random Forest (RF), and XGBoost were used to build the ensemble model. A one-vs-all approach was adopted to design an ensemble of the classifiers tailored to detecting a specific class within the dataset. This means that the feature selection process was done for each class, producing multiple datasets based on the number of classes within the dataset. The results of the classifiers are then combined and aggregated using the argmax function. Finally, the proposed model was evaluated against other models, existing works in literature and unknown attacks to see how well the model performs. The results showed that the proposed approach performs better than other approaches achieving a better macro average F1-score of 83.50% and an improved classification of the minority classes, attaining an F1-score of 29.95% and 75.98% in the infiltration and web classes respectively. The infiltration class was seen to be hard to decipher from the benign class and so approaches to properly separate and oversample the infiltration class should be taken to improve the detection of the class.

Keywords: ***Intrusion Detection System, CICIDS2018, Cyber Security, Machine Learning, Deep Learning***