AN ACTIVE SELF-LEARNING MODEL FOR DECEPTIVE PHISHING DETECTION

Subhash Niroshan Ariyadasa

188077D

Doctor of Philosophy

Department of Computational Mathematics
Faculty of Information Technology

University of Moratuwa
Sri Lanka

July 2023

AN ACTIVE SELF-LEARNING MODEL FOR DECEPTIVE PHISHING DETECTION

Subhash Niroshan Ariyadasa

188077D

Dissertation submitted in partial fulfillment of the requirements for the degree
Doctor of Philosophy

Department of Computational Mathematics
Faculty of Information Technology

University of Moratuwa
Sri Lanka

July 2023

DECLARATION

I declare that this is my own work and this dissertation does not incorporate without acknowledgement any material previously submitted for a Degree or Diploma in any other University or institute of higher learning and to the best of my knowledge and belief it does not contain any material previously published or written by another person except where the acknowledgement is made in the text.

Also, I hereby grant to University of Moratuwa the non-exclusive right to reproduce and distribute my dissertation, in whole or in part in print, electronic or other medium. I retain the right to use this content in whole or part in future works (such as articles or books).

Signature: _ _ _ _ _ _ _ _ _ _ _ _ _ Date: _ 20-07-2023 _

The above candidate has carried out research for the PhD Dissertation under my supervision.

Name of the supervisor: Dr. K. S. D. Fernando

Signature of the supervisor: _ _ _ _ _ _ _ _ _ _ _ _ _ Date: _ 21/07/2023 _

Name of the supervisor: Prof. M. S. D. Fernando

Signature of the supervisor: _ _ _ _ _ _ _ _ _ _ _ _ _ Date: _ 21-07-2023 _

DEDICATION

This work is entirely dedicated to my respectful parents and my family. Without their constant support, this dissertation would not have been possible. They always inspire me. At the same time, my special thanks also go to my supervisors, who enlightened me with immense academic knowledge and gave me valuable advice whenever I needed it the most.

ACKNOWLEDGMENTS

In a very special way, I would like to thank the following people who have helped me undertake this research:

- My supervisors, Dr. Subha Fernando and Dr. Shantha Fernando, for their enthusiasm for this research, support, encouragement and patience.
- My progress review examiners, Dr. Chamath Keppitiyagama and Dr. Kasun De Zoysa, for their support, encouragement and feedback.
- My progress review evaluation panel, Dr. Thushari Silva, Dr. Thanuja Sandanayake, Dr. Thilini Piyatilake, Dr. C.R.J. Amalraj and Dr. Sagara Sumathipala, for their support, encouragement and feedback.
- Research coordinators at the Department of Computational Mathematics, Dr. Thushari Silva and Dr. Thilini Piyatilake, for their support when arranging my progress evaluations and other PhD program-related tasks.
- Masters program coordinator, Dr. Sagara Sumathipala, for the support received for my course-work completion.
- Dean, Assistant Registrar, and all other officials at the Faculty of Graduate Studies, the University of Moratuwa for all the support I received during my PhD program.
- Director and other officials at the Center for Information Technology Services (CITeS), University of Moratuwa, for their support in my PhD results evaluation process.
- Vice-Chancellor of Uva Wellassa University, Prof. Jayantha Ratnasekera, for allowing me full-time leaves to complete this PhD program.

- Head of the Department of Computer Science and Informatics, Uva Wellassa University, Ms. Harshani Wickramarathna, for her support and encouragement.
- All the relevant officials at the University of Moratuwa and Uva Wellassa University, for their support.
- My respectful parents, my beloved wife and my loving children, for their constant support and encouragement.
- My brothers, respectful parents-in-law, brother-in-law and sisters-in-law, for their support and encouragement.
- All of my colleagues and friends who helped me throughout this PhD program.

I am incredibly grateful for the immeasurable help and support every one of you has contributed to my program. Words alone cannot express my gratitude to you all. Thank you.

ABSTRACT

Phishing presents an ongoing and dynamic threat to Internet users, targeting personal and confidential information. Existing anti-phishing solutions encounter challenges in keeping up with the ever-changing nature of these attacks, leading to performance degradation over time. This study aims to develop an autonomous anti-phishing solution that effectively counters evolving phishing threats through continuous knowledge updates. To address the challenge of detecting the latest phishing attacks, SmartiPhish, an autonomous anti-phishing solution with continuous learning support, is proposed. Utilizing a quantitative research approach, data is collected from trusted third parties at multiple time points to create a valid dataset. The primary outcome is a reinforcement learning solution that leverages a novel deep learning model alongside Alexa rank and community decisions. The innovative use of Graph Neural Networks in the anti-phishing domain, combined with Long-term Recurrent Convolutional Networks, enables SmartiPhish to estimate a website's phishing probability using URL and HTML content features. Additionally, the study addresses a crucial research gap by developing a reliable method named PhishRepo for collecting and precisely labelling the latest phishing data. SmartiPhish exhibits positive results, achieving a detection accuracy of 96.40%, an f1-score of 96.42%, and an exceptionally low False Negative Rate (FNR) of 0.029. In real-world web environments, the solution outperforms similar solutions and demonstrates enhanced effectiveness against zero-day phishing attacks. Notably, the integration of continuous learning support facilitates a significant 6% improvement in detection accuracy after six weeks. SmartiPhish's adaptive approach integrates a systematic knowledge acquisition process, enabling dynamic updates of phishing detection features to counter the ever-evolving landscape of phishing attacks. The findings highlight its potential in strengthening cybersecurity measures and provide practical insights for dealing with phishing threats in today's digital world. Continuously updating its knowledge base, SmartiPhish stands as a strong defence, promising improved protection for Internet users.

Keywords: Cyberattack, Deep learning, Graph neural networks, Internet security, Reinforcement learning

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS

ACMR	Absolute Cumulative Majority Relabelling
Adam	Adaptive Moment Estimation
AI	Artificial Intelligence
AIWL	Automated Individual White-List
ANN	Artificial Neural Networks
API	Application Programming Interface
ARMA	Auto-Regressive Moving Average
CA	Certificate Authority
CML	Conventional Machine Learning
CNN	Convolutional Neural Network
ComD	Community Decision
CSS	Cascading Style Sheets
Deque	Double-Ended Queue
DL	Deep Learning
DLM	Deep Learning Model
DNS	Domain Name System
DoS	Denial-of-Service
DQN	Deep Q-learning Network
FIFO	First In, First Out
FNR	False Negative Rate
GAN	Generative Adversarial Networks
GBDT	Gradient Boosting Decision Tree
GCN	Graph Convolutional Networks
GCS	Graph Convolutional Skip
GN	Generator Network
GNN	Graph Neural Network
GSB	Google Safe Browsing
HTML	HyperText Markup Language
HTTPS	Hypertext Transfer Protocol Secure
IPR	Initial Phishing Records
KAP	Knowledge Acquisition Process
k-NN	k-Nearest Neighbour
LightGBM	Light Gradient Boosting Machine
LRCN	Long-term Recurrent Convolutional Network
LSTM	Long Short-Term Memory
MDP	Markov Decision Process
MLP	Multi-Layer Perceptron
MPN	Multilayer Perceptron Network

NLP	Natural Language Processing
NN	Neural Networks
PhaaS	Phishing-as-a-Service
pHash	Perceptual Hashing
PNN	Probabilistic Neural Network
ReLU	Rectified Linear Unit
RESTful	Representational State Transfer
RF	Random Forest
RL	Reinforcement Learning
RLM	Reinforcement Learning Model
RoF	Rotation Forest
SaaS	Software as a Service
SEO	Search Engine Optimisation
SGD	Stochastic Gradient Descent
SMO	Sequential Minimal Optimisation
SMS	Short Message Services
SSL	Secure Sockets Layer
SVM	Support Vector Machine
Tanh	Hyperbolic Tangent
TAP	Target Attack Prevention
TF-IDF	Term Frequency-Inverse Document Frequency
TLD	Top-Level Domain
URL	Uniform Resource Locator
XGBoost	eXtreme Gradient Boosting
XSS	Cross-Site Scripting