

ABSTRACT

Phishing attempts continue to be a serious concern to people as well as organisations at an age when cyber threats are constantly changing. Traditional techniques for identifying phishing URLs frequently cannot keep up with the ever-changing tactics used by illicit organisations. This article presents Phish Not, a state-of-the-art cloud-based machine-learning system that has been painstakingly designed to improve the precision and quickness of phishing URL detection.

Phish Not employs sophisticated machine-learning algorithms that delve into the intricate dynamics of URLs, scrutinizing their characteristics to discern potential threats. Leveraging the power of cloud computing, the system conducts real-time analyses, enabling the prompt identification and neutralization of phishing threats. Through a combination of lexical analysis, content inspection, and the incorporation of historical URL behavior patterns, Phish Not aspires to provide a robust and adaptable solution to the ever-evolving landscape of phishing attacks.

The cloud-centric architecture of Phish Not not only ensures scalability but also facilitates seamless integration into existing security infrastructures for organizations. Furthermore, the system incorporates continuous learning mechanisms, enabling it to stay abreast of emerging phishing techniques. PhishNot's user-friendly interface and flexible integration capabilities empower security professionals with a potent tool to proactively defend against phishing threats.

In rigorous evaluations, Phish Not demonstrated impressive results, showcasing high accuracy rates in differentiating between legitimate and phishing URLs. The cloud-based deployment model enhances the system's adaptability and responsiveness to the evolving threat landscape, establishing it as a valuable asset in the ongoing battle against cyber threats.

ACRONYM

SDN	Software Defined Networks
DPI	Deep Packet Inspection
URL	Uniform Resource Locator
NLP	Natural Language Processing
HTML	Hyper Text Markup Language
ReLU	Rectified Linear Unit
SVM	Support Vector Machine
CNN	Convolutional Neural Network
DTOF	Decision Tree and Optimal Features
ANN	Artificial Neural Network
NLP	Natural Language Processing
CBR-PDS	A case-based reasoning Phishing detection system

TABLE OF CONTENTS

	Page No.
ABSTRACT	i
ACRONYMS	ii
TABLE OF CONTENTS	iii
LIST OF FIGURES	iv
CHAPTER 1	INRTRODUCTION
1.1	Introduction 01
1.2	Objectives 02
1.3	Phishing 03
CHAPTER 2	LITERATURE REVIEW
2.1	Existing System 06
2.2	Drawbacks 06
CHAPTER 3	PROPOSED SYSTEM
3.1	Pre-Deployment Stage 09
3.2	Deployment Stage 10
3.3	Key Features 10
3.4	Dataset 11
CHAPTER 4	DESIGN
4.1	UML Diagram 13
4.1.1	Use Case Diagram 13
4.1.2	Activity Diagram 14
CHAPTER 5	IMPLEMENTATION
5.1	Algorithms 15
5.2	Software Requirement Specification 24
5.2.1	Software Requirements 24
5.2.2	Hardware Requirements 25
CHAPTER 6	EVALUATION DETAILS
6.1	Sample Code 26
CHAPTER 7	FUTURE WORK
CHAPTER 8	CONCLUSION
CHAPTER 9	REFERENCES

LIST OF FIGURES

Fig No	Name of Figure	Page No
1.1	Phishing	5
3.1	PhishNot System Overview	9
3.2	Phising attack features	12
4.1	Use case Diagram	13
4.2	Activity Diagram	14
6.1	Evaluation Process	26
6.2	Dataset	27
6.3	Correlation Heatmap	28
6.4	Data Pie Chart	29
6.5	Training & Testing Accuracy for Random Forest	31
6.6	Training & Testing Accuracy for Logistic Regression	33
6.7	Training & Testing Accuracy for Decision Tree	35
6.8	Training & Testing Accuracy for Gaussian Naïve Bayes	37
6.9	Training & Testing Accuracy for MLP	39
6.10	ML Model Comparisons	41
6.11	Various ML Models Results	41