

WEB PHISHING DETECTION

TEAM ID: PNT2022TMID26256

Mentor:
Mr. M.
Krishnaraj

Team Leader:
Roshan S B

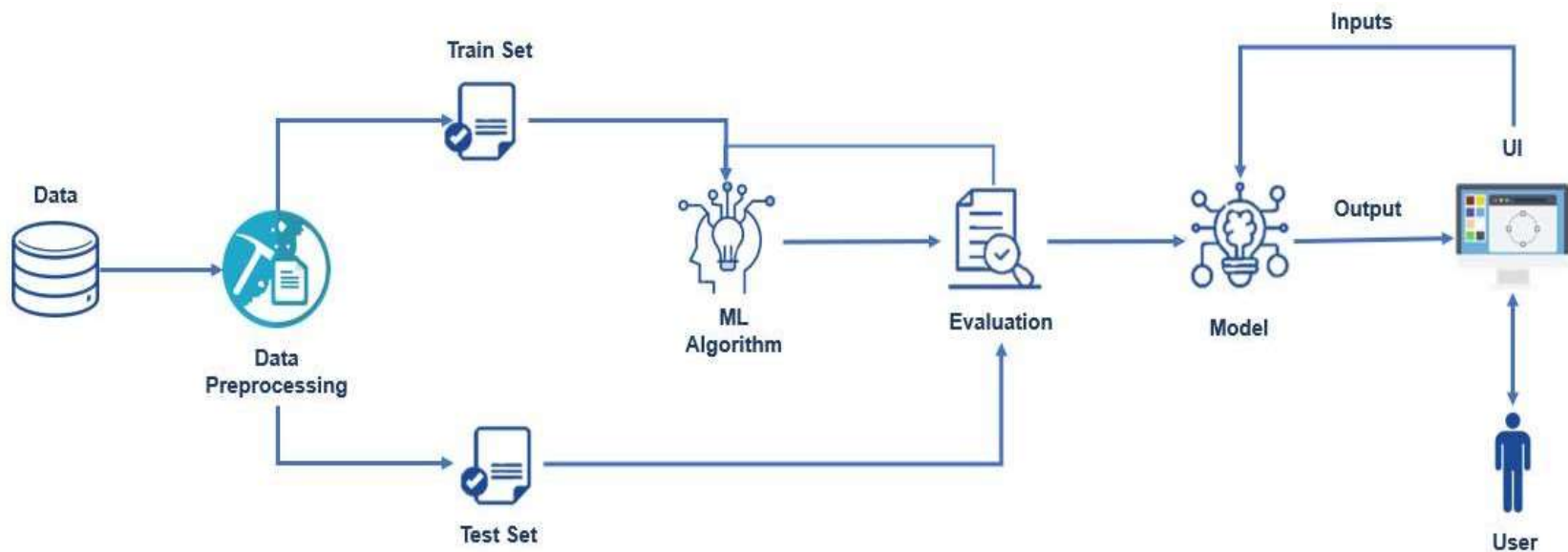
Team Members:
Rohith B
Siva vishva V
Yugendran M

Problem Definition

There are a number of users who purchase products online and make payments through e-banking. There are e-banking websites that ask users to provide sensitive data such as username, password & credit card details, etc., often for malicious reasons. This type of e-banking website is known as a phishing website. Web phishing is one of many security threats to web services on the Internet.

In order to detect and predict e-banking phishing websites, we proposed an intelligent, flexible and effective system that is based on using classification algorithms and techniques to extract the phishing datasets criteria to classify their legitimacy. The e-banking phishing website can be detected based on some important characteristics like URL and domain identity, and security and encryption criteria in the final phishing detection rate.

Technical Architecture



Paper 1

Authors:

Zuochao Dou, Issa Khalil, Abdallah Khreishah, Ala Al-Fuqaha and Mohsen Guizani

Title:

Systematization of Knowledge (SoK): A Systematic Review of Software-Based Web Phishing Detection

Published Journal:

IEEE Communications Surveys & Tutorials
(Volume: 19, Issue: 4, Fourthquarter 2017)

Published Date:

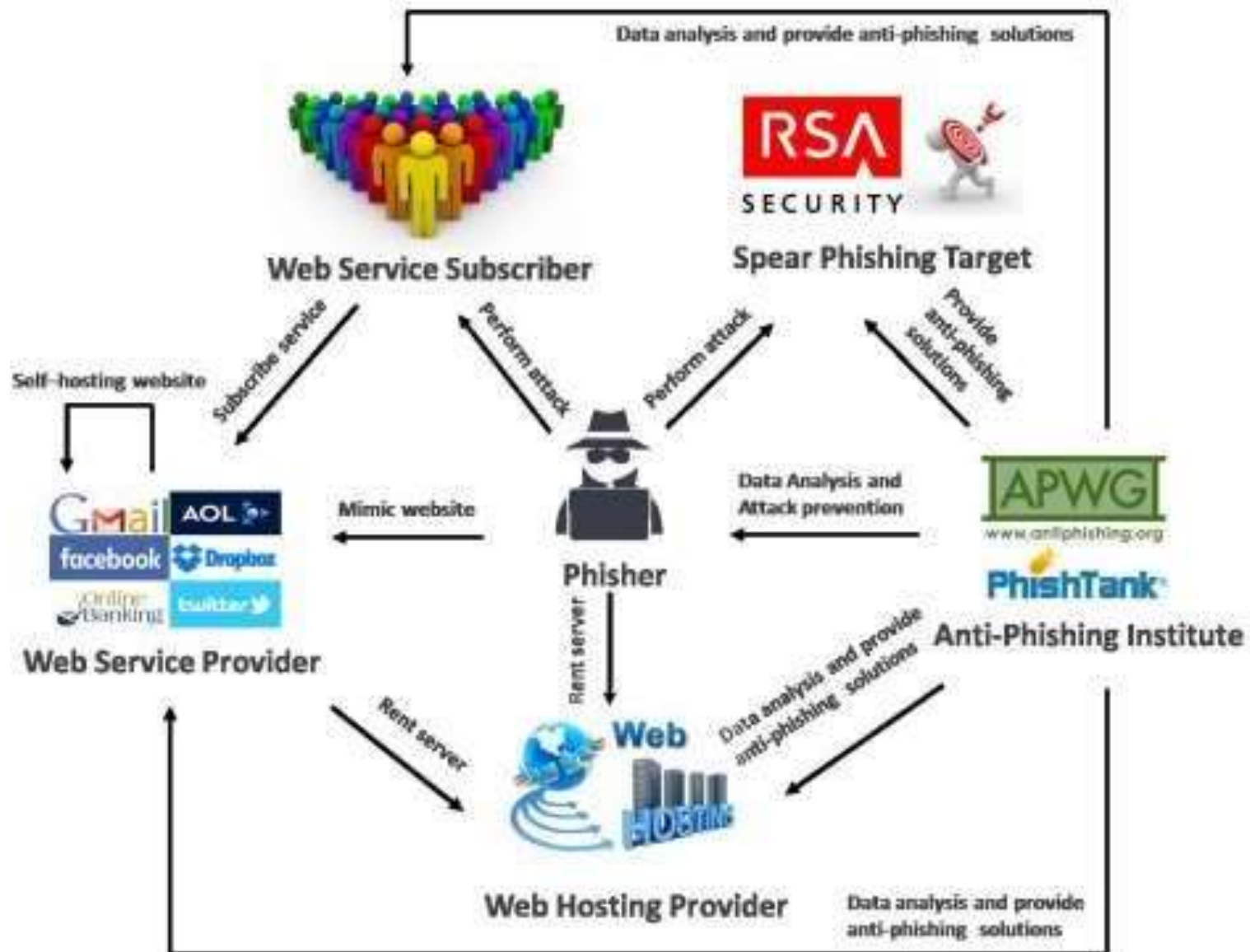
13 September 2017

Objective:

The objective of this paper is to provide a systematic understanding of existing phishing detection studies and provide a comprehensive way to evaluate phishing detection approaches from different perspectives in order to guide future developments and validations of new or upgraded anti-phishing techniques. Software-based phishing detection includes many important aspects like the phishing detection life cycle, taxonomy of phishing detection schemes, evaluation datasets, detection features, and evaluation metrics and strategies.

Phishing Detection Techniques:

- Data mining algorithms
- Sequential Minimal Optimization (SMO)
- Logistic Regression (LR)
- Neural Network (NN)



Paper 2

Authors:

Said Salloum, Tarek Gaber, Sunil Vadera and Khaled Shaalan

Title:

A Systematic Literature Review on Phishing Email Detection Using Natural Language Processing Techniques

Published Journal:

IEEE Access (Volume: 10)

Published Date:

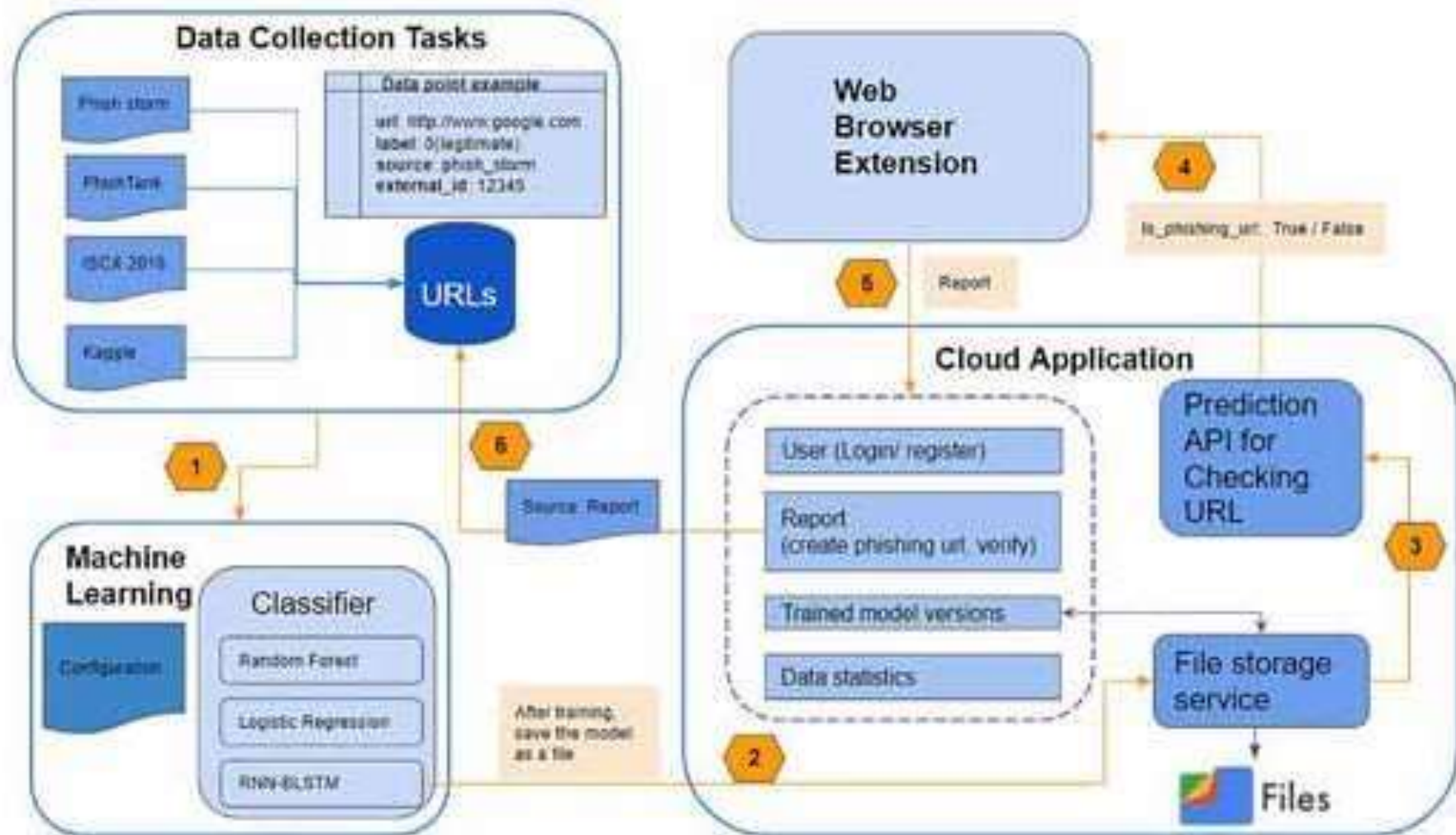
14 June 2022

Objective:

The most common phishing attack vectors include communication channels like emails and other messaging apps. This paper offers a comprehensive literature review of studies that aim to utilize natural language processing (NLP) and machine learning (ML) methods for detecting phishing emails. The work is aimed to survey the work published using NLP and ML for detecting phishing emails but has not systematically reviewed all published papers in the last 10 years.

Phishing Detection Techniques:

- Natural Language Processing
- Adam optimizer
- Convolutional Neural Network (CNN)
- Latent Semantic Analysis (LSA)



Paper 3

Authors:

Jian Mao, Wenqian Tian, Pei Li, Tao Wei and Zhenkai Liang

Title:

Phishing-Alarm: Robust and Efficient Phishing Detection via Page Component Similarity

Published Journal:

IEEE Access (Volume: 5)

Published Date:

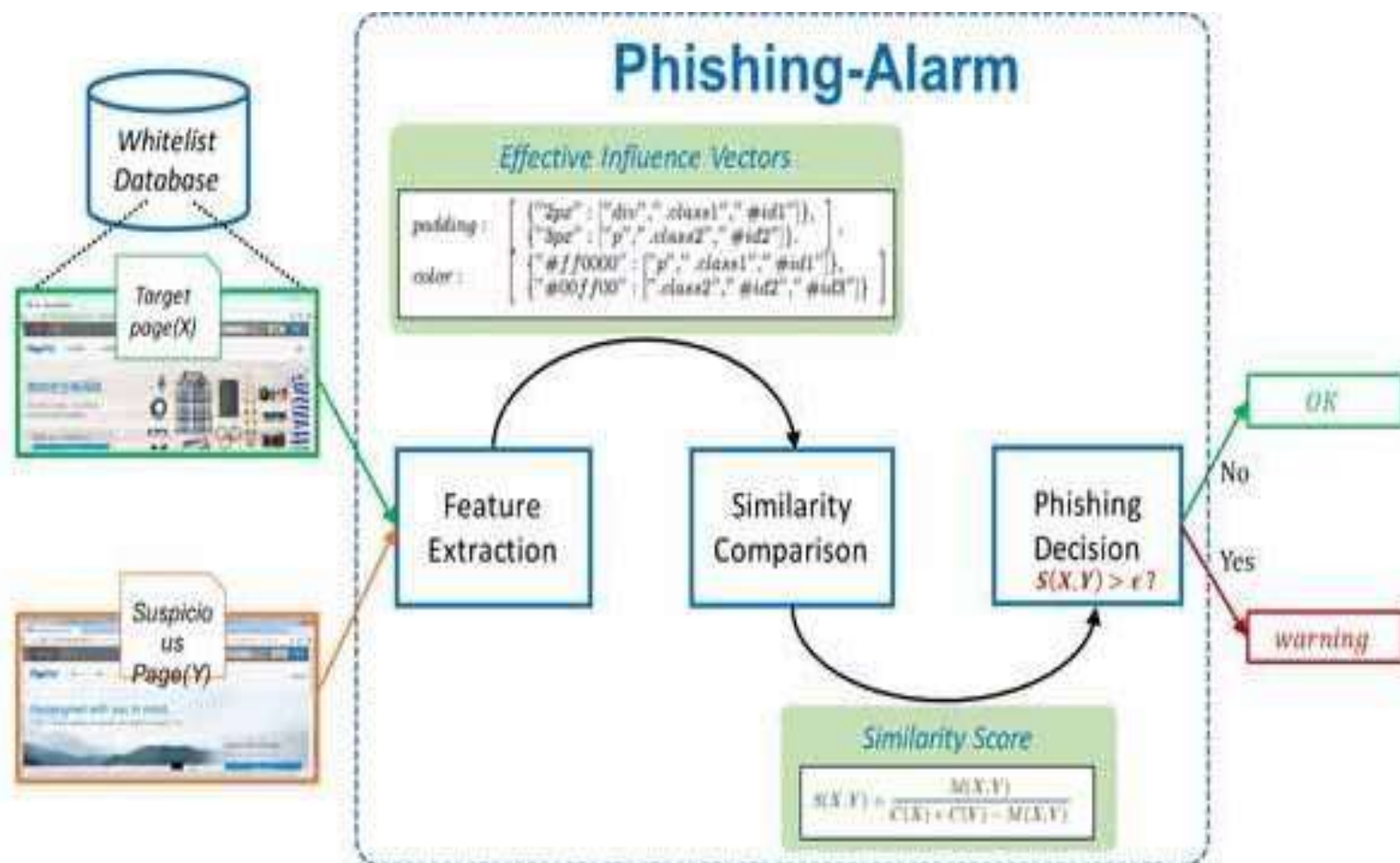
23 August 2017

Objective:

The paper proposes a robust solution to identify phishing pages according to the visual similarity of web page components, which are difficult to be evaded by attackers. The authors developed techniques to select the effective features on a web page, and propose an efficient method for page similarity detection according to these features. Their approach was prototyped and evaluated using a large set of phishing pages. The results illustrate that the approach is efficient and effective.

Phishing Detection Techniques:

- Phishing – alarm
- Black / White list based detection
- URL based detection
- Content based detection



Paper 4

Authors:

Peng Yang, Guangzhen Zhao and Peng Zeng

Title:

Phishing Website Detection Based on Multidimensional Features
Driven by Deep Learning

Published Journal:

IEEE Access (Volume: 7)

Published Date:

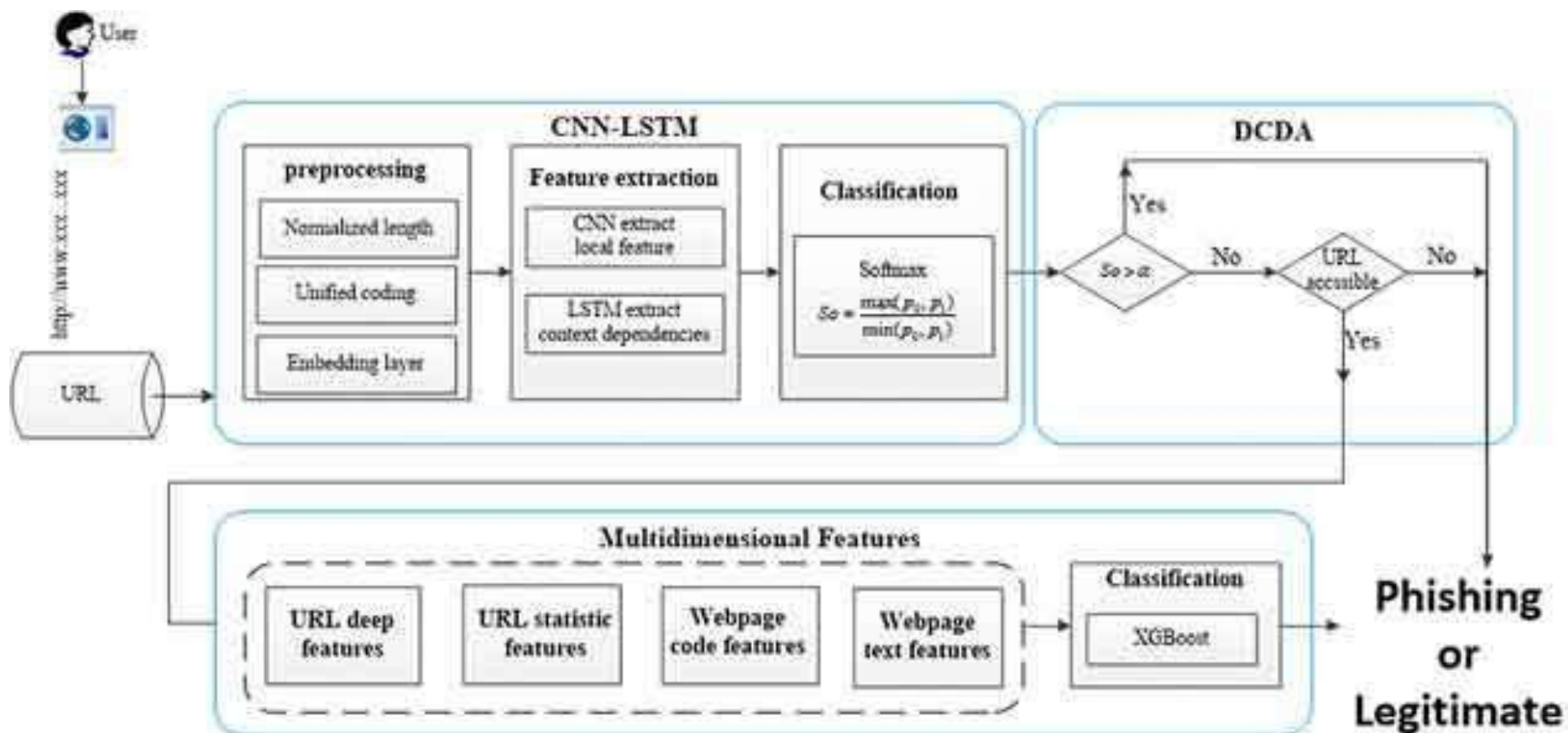
11 January 2019

Objective:

The APWG (Anti-Phishing Working Group) defines phishing as a criminal mechanism employing both social engineering and technical subterfuge to steal personal identity data and financial account credentials of consumers. A multidimensional feature phishing detection approach was proposed based on a fast detection method by using deep learning (MFPD). A dynamic category decision algorithm (DCDA) is proposed. By revising the output judgment conditions of the softmax classifier in the deep learning process and setting a threshold, the detection time can be reduced.

Phishing Detection Algorithms:

- The CNN-LSTM Algorithm
- The Multidimensional Feature Algorithm
- The Dynamic Category Decision Algorithm



Paper 5

Authors:

Saad Al-Ahmadi, Afrah Alotaibi and Omar Alsaleh

Title:

PDGAN: Phishing Detection With Generative Adversarial Networks

Published Journal:

IEEE Access (Volume: 10)

Published Date:

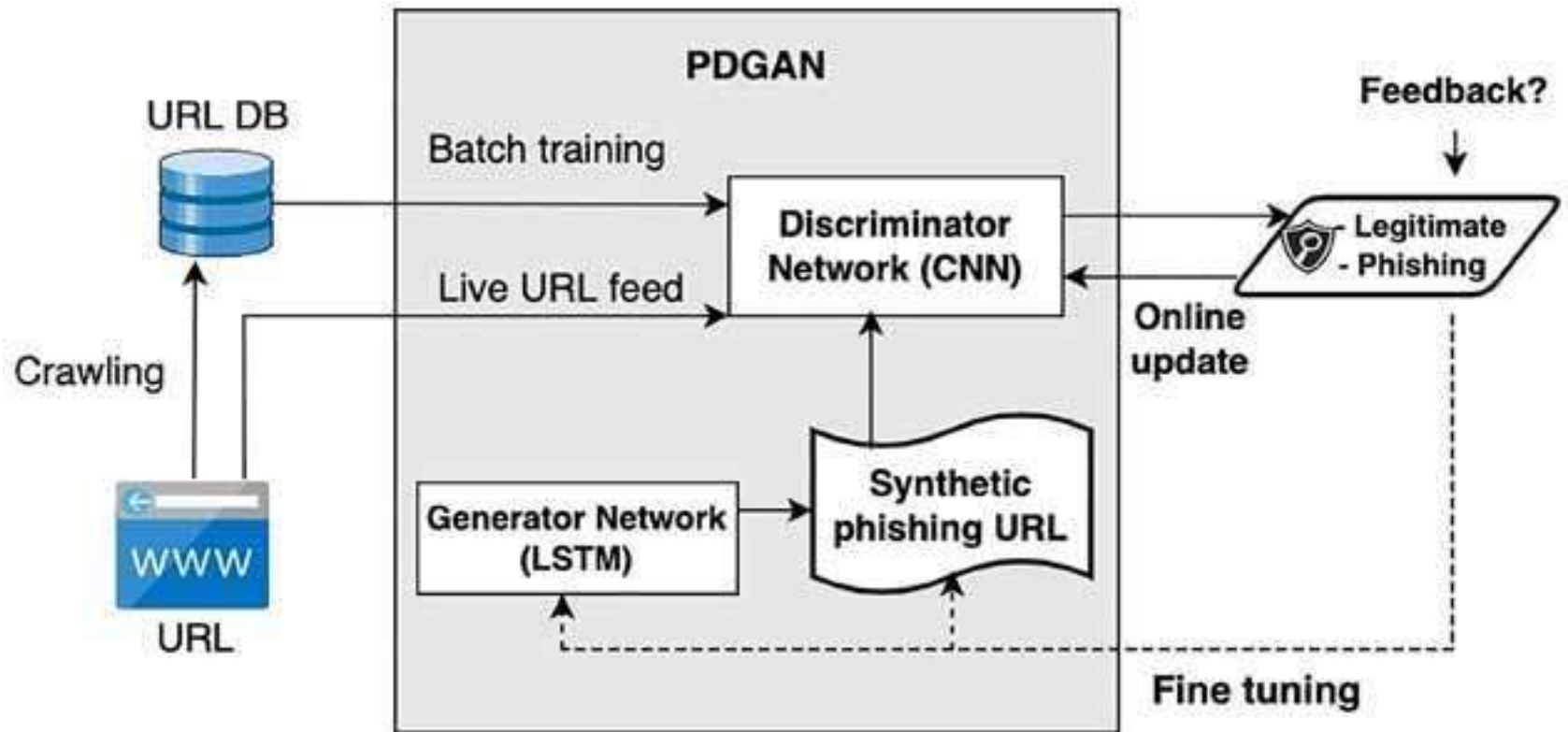
18 April 2022

Objective:

A website is classified as phishing in a machine learning approach if the tested website results match the predefined feature set. The performance of this approach depends on the feature set, training data, and classification algorithm. Using machine learning algorithms can enable unseen URLs to be easily detected. A phishing website detection approach PDGAN, which does not depend on webpage content but rather only on a webpage's URL is designed. PDGAN uses a deep learning model, namely a GAN, whose adversarial process allows the model to learn different variations in phishing features and produce a final model that provides better detection results.

Phishing Detection Techniques:

- Deep learning
- Deep Neural Network
- Long short-term memory network (LSTM)



Thank You 😊