ScienceDirect

Materials Today: Proceedings

Volume 80, Part 3, 2023, Pages 3129-3139

SI-BBA – A novel phishing website detection based on Swarm intelligence with deep learning

Parvathapuram Pavan Kumar, T. Jaya, V. Rajendran

Show more

https://doi.org/10.1016/j.matpr.2021.07.178

Get rights and content

Referred to by 5 NANO 2021 – EXPRESSION OF CONCERN – PART 4

Materials Today: Proceedings, Volume 80, Part 3, 2023, Pages 1704

View PDF

Abstract

Websites phishing is one of several defense coercions to Internet Service Provider. Mainly web phishing focused on stealing private information such as username, password, and credit card details too through imitating a legal creature. Deep learning based Neural Networks are extensively used for phishing detection with high accuracy measures and metrics. In this proposed work, an improved version of Binary Bat namely Swarm Intelligence Binary Bat Algorithm is used for designing the neural network which categorize the network URL websites similar to classification approach. It is utilized for the initial moment in this domain of relevance to the preeminent of our understanding. Our experimental results shows that deep learning based Adam optimizer reaches high

classification accuracy as 94.8% in phishing websites attack detection based on swarm intelligence technique.

Introduction

Phishingattack is a kind of societalproductionassaultfrequently utilized to embezzle user's information, comprising of login testimonials as well as credit card numbers. This happensonce an aggressor, hidden as a faithindividual, dupes a victim into modifyingemail information, such as instantaneous message, or content message. An attack may leads todestructive outcomes. Adebowale et. al [1] work payed attention on the design as well as development of phishing websites clarification which influenced URL and website related images, frames and text. For that, the author proposed hybrid (Intelligent Phishing Detection System) model which are integrated with CNN based algorithms and LSTM. Bo wei et. al [9] introduced light weight deep learning based model to distinguish the malevolent URL also facilitate in real time, power saving phishing URL detection sensor were used. Lakshmi et. al [17] utilized 30 features to identify malicious web pages. Moreover, deep learning based Adam optimizer method was applied for distinguishing malicious web pages from normal web sites. Finally the performances were compared with other conventional machine learning approaches for finding which algorithm generated best outcomes in detecting phishing websites. [33], [34], [35]

The methods used in phishing URL attack are as follows:

- Email Phishing- Mainly harasses is throwing by email.
- Spear Phishing- Another two complicated harass comprising in emails are whaling, Smishing and Vishing.
- Angler Phishing-performed hidden as a customer service financial credit on social media, hopeful to accomplish thedispleased consumer.

Several ways to prevent phishing attacks are as follows

- Distinguish what a phishing trick looks like
- Do not click on that specific link
- Should fix the firewalls to prevent the attackers
- Spin the user passwords frequently

- Find free anti-phishing trappings.
- Do not provide user's information to any sites which is not secure
- Pay attention to the updates regarding sites
- Do not get excited by pop-ups.

The illustration of phishing attack efforts are described as follows

- A spoofed email supposedly from the link (myuniversity.edu) is mass-disseminated to as several faculty members as probable.
- The electronic mail declares that the password of user is going to terminate. Instructions are given to go to myuniversity.edu/renewal to renovate their password within a day (24 h).

The categorizing of phishing attack issues along with its solutions were developed by Benavides et. al [8], [12] using deep learning based algorithms shown in Fig. 1.

Our work focused on detecting attack in the network environment especially in URL websites and categorizing the same into malicious and legitimate [14]. For that, we are implementing novel approach namely SI_BBA for categorizing the network data into legitimate and malicious which may helpful for several organization using network facilities [29].
The main objective of this proposal is

- To train the neural network using a swarm intelligence approach.
- To develop an algorithm called "A novel SI-BBA (Swarm Intelligence – Binary Bat Algorithm) to predict the phishing websites.
- To enhance the performance level of every deep based optimizer approach, measuring has performed as well as compared.

Ram Basnet et. al [25] introduced novel approach namely heuristic based approach to categorize phishing attack as positive and normal mentioned as negative by means of information existing only in URLs. False Positive Rate, and Error rate are the metrics were evaluated to detect the attack depends on dissimilar features in URL. The Fig. 2 illustrates

the general idea about phishing URL attack detection framework using machine learning approaches developed by [25].

Section snippets

Related work

Somesha et. al [18] developed several models such as deep based Neural Network, Long short term Memory, CNN for detecting phishing URL websites. These models achieve accuracy as 99.5% for Neural Networks, 99.6% for Long Short Term Memory, and 99.4% for CNN. This proposed model makes the model vigorous to malfunction and enhances the phishing recognition speed. Suleiman Y. Yerima et. al [28] and I Saha et. al [15] introduced deeplearning based CNN approach to obtain high accuracy classification

Phishing data acquisition

The benchmark dataset phishing.csv is downloaded from the following link

https://www.kaggle.com/akashkr/phishing-url-eda-and-modelling/data ↗. Here, we have taken 4898 samples from legitimate websites and phishing websites samples as 6157.

Preprocessing

The primary processing of data in order to prepare it for primary processing or for further analysis. It eliminates the features that contain missing values or null values.

3.3 Feature extraction

The relevant features related to phishing websites URL are extracted through this phase.

Proposed algorithm coding using python language

We are proposing a novel deep learning based Swarm Intelligence-Binary Bat Algorithm for finding the phishing URL websites attack occur in the network surroundings also categorizing the attack websites from normal one.

Output: Deep Neural Network model based on hyper-parameter tuning

Step 1: Initialization of models SI-BBA

Step 2: While termination condition not meet

- **Do**

Step 3: Solution=SI-BBA_best_model ();

Step 4: Epoch=pattern_epoch();

Step 5: Batch=pattern_batch ();

Step 6:

Features in dataset

The dataset comprises of several features of URL such as user id, IP address, length, port HTTP tokens etcfor detecting and classifying the phishing websites attacks in the network environment. The features utilized in the datasets for distinguishing attack and normal are described in Fig. 5.

Dataset classification

The number of samples we have taken for phishing websites detection as 11055. The samples are splitted into training and testing phase samples for evaluate the model characteristics and also better

Comparison of existing method with proposed algorithm

Table 4 illustrates the comparison of existing work and proposed work in finding the phishing URL detection based on accuracy metrics.

Conclusion

In this paper we proposed a deep learning model based SI-BBA algorithm for the recognition of phishing websites and also performed classification of phishing websites from legitimate websites. The deep learning based Adam optimizer algorithm achieves the classification accuracy as 94.8% with 0.2 loss value. In future, by adjusting certain key manipulated features such as number of epochs, learning rate and batch size, we will achieve more accuracy so as to mutually consequence in finest

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Special issue articles Recommended articles

References (35)

Adebowale, M.A.,Lwin, K.T.andHossain, M.A.(2020), “Intelligent phishing detection scheme using deep learning...

Adriana-Cristina Enache, Valentin Sgârciu and Alina Petrescu-Niţa “Intelligent Feature Selection Method rooted in...

Aksu D., Turgut Z., Üstebay S., Aydin M.A. (2019) Phishing Analysis of Websites Using Classification Techniques. In:...

Alloghani M., Al-Jumeily D., Hussain A., Mustafina J., Baker T., Aljaaf A.J. (2020) Implementation of Machine Learning...

Arun Kulkarni, Leonard L. Brown, “Phishing Websites Detection using Machine Learning”, (IJACSA) International Journal...

A. Basit *et al.*
A comprehensive survey of AI-enabled phishing attacks detection techniques
Telecommunication System (2021)

A. Begum *et al.*
A Study of Malicious URL Detection Using Machine Learning and Heuristic Approaches

Benavides E., Fuertes W., Sanchez S., Sanchez M. (2020) Classification of Phishing Attack Solutions by Employing Deep...

Bo Wei, Rebeen Ali Hamad, Longzhi Yang, Xuan He, Hao Wang, Bin Gao and Wai Lok Woo “A Deep-Learning-Driven Light-Weight...

Cuzzocrea, A., Martinelli, F., &Mercaldo, F. (2018).Applying Machine Learning Techniques to Detect and Analyze Web...

View more references

Cited by (0)

View full text