International Journal of Recent Technology and Engineering (IJRTE)
ISSN: 2277-3878, Volume-8 Issue-3, September 2019

Deep Learning Based-Phishing Attack Detection

K. Sumathi, V. Sujatha

Abstract*: Due to the rapid development of the communication technologies and global networking, lots of daily human life activities such as electronic banking, social networks, e-commerce, etc are transferred to the cyberspace. The anonymous, open and uncontrolled infrastructure of the internet enables an excellent platform for cyber attacks. Phishing is one of the cyber attacks in which attackers open some fraudulent websites similar to the popular and legal websites to steal the user's sensitive information. Machine learning techniques such as J48, Support Vector Machine (SVM), Logistic Regression (LR), Naive Bayes (NB) and Artificial Neural Network (ANN) were widely to detect the phishing attacks. But, getting good-quality training data is one of the biggest problems in machine learning. So, a deep learning method called Deep Neural Network (DNN) is introduced to detect the phishing Uniform Resource Locators (URLs). Initially, a feature extractor is used to construct a 30-dimension feature vector based on URL-based features, HTML-based features and domain-based features. These features are given as input to the DNN classifier for phishing attack detection. It consists of one input layer, multiple hidden layers and one output layer. The multiple hidden layers in DNN try to learn high-level features in an incremental manner. Finally, the DNN returns a probability value which represent the phishing URLs and legitimate URLs. By using DNN the accuracy, precision and recall of phishing attack detection is improved.

Keywords* : Artificial Neural Network, Deep learning, Deep Neural Network, Phishing, Machine learning.

I. INTRODUCTION

In the field of computer security, phishing [1] is the criminally fraudulent process of trying to find out the sensitive information such as username, passwords and credit card details by masquerading as a trustworthy entity in an electronic communications. Phishing is a form of cyber-attack that uses counterfeit websites to steal sensitive user information such as credit card numbers, account login credentials, etc. Phishing presents direct risks through the use of stolen credentials and indirect risk to institutions that conduct business on-line through erosion of customer confidence. The damage caused by phishing ranges from denial of access to email to substantial financial loss.

The phishing attacks are broadly classified as deceptive phishing and malware phishing. Deceptive phishing refers to any attack by which fraudsters impersonate a legitimate company and attempt to steal people's personal information or login credentials. Those emails frequently use threats and a sense of urgency to scare users into doing the attackers bidding. Malware-based phishing occurs when an attacker attaches a harmful computer program made to look helpful onto emails, websites and other electronic documents in the internet. The phishing [2] can be implemented in different ways such as follows:

- Email to email: When someone receives an email requesting sensitive information to be sent to the sender.
- Browser-to-website: When some misspelled a legitimate web address on a browser and then referred to a phishing website that has a semantic similarity to the legitimate web address.
- Email-to-website: When someone receives an email embedded with phishing web address.
- Website-to-website: When some clicks on phishing website through a search engine or an online advert.

The detection and mitigation of phishing attacks is a great challenge due to the complexity of current phishing attacks. Machine learning techniques such as J48, Support Vector Machine (SVM), Logistic Regression (LR), Naive Bayes (NB) and Artificial Neural Network (ANN) [3] were used to detect the phishing attacks. However, getting good-quality training data is one of the biggest problems in machine learning because data labeling can be a tedious and expensive one.

So in this paper, a deep learning technique called Deep Neural Network (DNN) is introduced for phishing attack detection. Initially, Uniform Resource Locators (URLs) are given as input to the feature extractor where a 30-dimension vector is constructed. Then the extracted features 30-dimension feature vector is trained in DNN. It consists of one input layer, multiple hidden layers and one output layer. In the input layer, the weight and bias values are initialized for feature vectors. Then in the hidden layer, sigmoid function is applied and then its results are given to the next layer. Finally, one probability value is returned in the output layer which is used to represent the URL is either legitimate URL or phishing URL. By using multiple hidden layers in DNN, neural nets are capable of discovering latent structures within unlabeled data with high accuracy.

The rest of the article is structured as follows: Section II provides the previous researches related to phishing attack detection using data mining techniques. Section III explains the proposed DNN based phishing attack detection in brief. Section IV compares the performance of the proposed technique with the existing technique and Section V concludes the research work.

Manuscript published on 30 September 2019
* Correspondence Author
K Sumathi*, Department of Computer Applications, CMS College of Science and Commerce, Coimbatore, Tamilnadu, India. Email: sumathiphd2017@gmail.com.
V. Sujatha, Department of Computer Applications, CMS College of Science and Commerce, Coimbatore, Tamilnadu, India.

II. LITERATURE SURNEY

A semi-supervised learning approach called Transductive Support Vector Machine (TSVM) [4] was proposed for detection of phishing websites. Initially, features of web pages were extracted for balancing the drawback of phishing detection based on Document Object Model (DOM).
It included color histogram, gray histogram and spatial relationship between subgraphs. A page analysis based on DOM objects was used to examine the features of sensitive information. TSVM was introduced to train classifier which considered the distribution information that implicitly represented in the large quantity of the unlabeled samples. However, the classifier needs improvement in terms of accuracy.
A new rule-based method [5] was proposed to detect the phishing attacks in internet banking. This method used two novel feature sets to find the website identity. The feature sets include four features to evaluate the page resource identity and four features to determine the access protocol of page resource elements. In order to find the relationship between the content and URL of a page, an approximate string matching algorithm was used in the feature sets. The features sets were processed in the Support Vector Machine (SVM) to classify the legitimate and phishing web pages. However, SVM requires extensive memory for web page classification in many cases.
A fuzzy-rough hybrid system [6] was introduced for detection of phishing in Iranian e-banking. The fuzzy-rough hybrid system was combination of fuzzy logic and rough sets-based data mining. A rough attribute reducer was used to reduce the dimensionality of the data. Then, fuzzy logic was applied to convert the input data into linguistic variables. It generated rules used to detect the phishing websites. However, a membership function used in fuzzy logic greatly influences the performance of phishing detection.
A new fast associative classification algorithm (FACA) [7] was proposed for detecting phishing websites. The FACA was an effective supervised learning approach. It integrated association rule mining and classification into a single process. FACA employed a vertical mining approach called Diffset to discover all frequent itemsets and a new prediction method was used to classify the websites as legitimate websites and phishing websites. However, the computational complexity of this method is high.

A new Hybrid Ensemble Feature Selection (HEFS) framework [8] was proposed for machine learning based phishing detection system. This framework was comprised of two phases. A novel Cumulative Distribution Function gradient (CDF-g) algorithm was used in the first phase of HEFS framework. It generated primary feature subsets and it was given as input to the data perturbation ensemble. It returned secondary feature subsets. A perturbation ensemble function was used in the second phase of HEFS to derive a set of baseline features from the secondary feature subsets. These features were used in Support Vector Machine (SVM), Naïve Bayes, C4.5, JRip and PART classifiers to detect the phishing websites. However, it could be valuable to analysis the impact of feature selection using associative classification.

A novel framework [9] was proposed for detection online phishing email using dynamic evolving network based on reinforcement learning. This framework detected the phishing attacks in the online mode. In the novel framework, a Feature Evaluation and Reduction (FEaR) algorithm was developed to explore the new behavior as to rank a selected list of features. The FEaR algorithm was dynamically changing the number of significant features and extracted them from next email. A Neural Network (NN) was used as the core of the classification model and a Dynamic NN using Reinforcement Learning (DENNuRL) was developed to allow the NN to evolve dynamically and built the best NN able to solve the phishing attack problem. More dataset will be included to the offline dataset to enhance the richness of the model.
An efficient phishing detection model [10] was proposed based on Optimal Feature Selection and Neural Network (OFS-NN). This model solved the overfitting problem of neural network due to the many useless and small influence features used in the neural network. In OFS-NN, a Feature Validity Value (FVV) was introduced to analysis the impact of sensitive features on phishing websites detection. Then, an algorithm was designed based on the new FVV index to select optimal features from the phishing websites. The selected features were trained in the neural network to detect the phishing websites. However more features will improve the classifier performance.

III. PROPOSED METHODOLOGY

In this section, the proposed Deep Neural Network (DNN) based phishing attack detection is described in detail. Initially, for each URL a 30 dimension vector is constructed by using a feature extractor. Then, the 30-dimension feature vector is processed in DNN to classify the phishing attack signatures. The DNN is a deep learning technique which has the capability of learning features at multiple levels of abstraction that allow a system to learn complex functions mapping the input to the output directly from data, without depending completely on human-crafted features. This classifier is used to classify the URLs as phishing URLs and legitimate URLs.

A. Feature Extraction

A feature extractor gets URLs and web-based code as an input value and returns a vector that consists of thirty features for DNN classifier. These thirty features are comes under URL-based features, HTML-based features and domain-based features categories. Feature 1 to 13 contains the phishing characteristics of URL. Feature 14 to 23 is HTML-based features which are used to detect the anomaly of HTML and JavaScript code. Features 24 to 30 are domain-based features which identifies domain information from the URL. Initially, all URLs in the input message is determined and for each URL a 30-dimension vector is constructed based on 3-types of data. The 30-dimension vector (feature vector) is given as input to DNN classifier.

B. DNN based Phishing Attack Detection

The extracted features are given to train the DNN classifier. DNN is a deep learning which consists of three layers namely input, hidden and output layer. DNN is consists of multiple hidden layers between input and output layers. In DNN, the input layer assigns weights to the input parameters and transfers those to the next layer. Each subsequent layer also assigns weights to their input and generates their output.

At the output layer, the final output value is obtained and error function is calculated to determine how correctly learned those features for identifying the phishing attacks. This training cycle is repeated until the relationship between countermeasures and the extracted features is learned. By using training data, the phishing activities through emails and web-based communication are identified.

The probabilities are denoted as $f(x) = x$ are given to the input layer of neurons. DNN consists of multiple hidden layers which can handle the huge volume of data. Each hidden layer of DNN is defined as sigmoid transfer function which is given as follows:

$$f(x) = \frac{1}{1+e^{-x}} \qquad (1)$$

Each input feature has its own weight values as $w_1, w_2, \ldots, w_n$ and the weighted sum of the inputs is done by the adder function as follows,

$$u = \sum_{i=1}^{n} w_i x_i \qquad (2)$$

The output layer of DNN is described by the following equation.

$$y = f(\sum_{i=1}^{n} w_i x_i + b_i) \qquad (3)$$

In the Eq. (3), y is the output neuron value; $f(x)$ is the transfer function, w_i refers the weight values, x_i denotes input data values and b_i refers to the bias value. Based on the output neuron values, the relationship between countermeasures and the considered parameters is learned which identifies the phishing attack signature. By using this learned model, the existence of phishing attacks is predicted.

Deep Neural Network Algorithm

Input: Training dataset $D = \{(x_1, y_1), (x_2, y_2), \ldots (x_n, y_n)\}$

//x_n is the feature vectors, y_n is the output (phishing or legitimate)

Output: Trained neural network

Initialize all weights and biases in network;

$while(termination\ condition\ is\ not\ satisfied)$

{

$for(each\ training\ parameter\ X\ in\ D)$

{

$for(each\ input\ layer\ node\ j)$

{

$O_j = I_j$ //Output of an input layer

$for(each\ hidden\ or\ output\ layer\ node\ j)$

$H_j = \frac{1}{1+e^{-j}};$

$O_j = f(\sum_{i=1}^{n} w_{ij} x_i + b_j)$

$for(each\ node\ j\ in\ output\ layer)$

$E_j = \frac{1}{2}(t_j^p - o_j^p)^2$

//t_j^p is the desired target output for the p-th observation and the o_j^p is the actual output for the p-th observation.

Update the weight and bias values based on the E_j (error value).

}

}

}

if $O_j < 0.5$

{

return -1 // it is corresponding to legitimate URL

else

return 1 //it is corresponding to phishing URL

}

The above algorithm describes how to identify the phishing URLs in the websites. Initially, the DNN is loaded with extracted feature vectors and then it is attached to the output matrix as the first layer. Then, in every loop of computation, the result will use the matrix multiplication to multiply weight matrix w and then add the bias b. In the multiple hidden layers, sigmoid function is applied and the result is given to the next layer. Based on the error values, the weight and bias values of DNN network is updated. Finally the output layer returns one probability. If the probability value is less than 0.5, then the URL is labelled as legitimate URL. Otherwise, that URL is labelled as phishing URL.

IV. RESULT AND DISCUSSION

In this section, the performance effectiveness of the proposed method is evaluated in in MATLAB 2018a by using the most popular dataset such as Ham, Phishing Corpus and Phishload datasets. The Ham dataset is used for baseline evaluation which consists of 4,150 legitimate e-mail communication and 1897 spam-based e-mails. The Phishing Corpus is used for its complication of phishing e-mail communication. Due to the complication of phishing e-mail data, this dataset is treated as the threat to mitigate. This dataset consists of 4,559 phishing e-mail messages. The Phishload dataset is used for its raw web-based coding structures. It consists of 1,185 legitimate and 3,718 phishing URLs. The effectiveness of existing ANN based phishing attack detection and DNN based phishing attack detection is tested in terms of accuracy, precision, recall and F-measure.

A. Accuracy

It is defined as the fraction of the total number of correct phishing attack detections to the actual dataset size. It measures the overall rate of correctly detected phishing and legitimate URLs.

$$Accuracy = \frac{True\ Positive\ (TP) + True\ Negative\ (TN)}{TP + TN + False\ Positive\ (FP) + False\ Negative\ (FN)}$$

where, TP is the percentage of phishing URLs in the training dataset that are correctly classified as phishing URLs, TN is the percentage of legitimate URLs in the training dataset that are correctly classified as legitimate URLs, FP is the percentage of legitimate URLs that are incorrectly classified as phishing URLs and the FN is the percentage of phishing URLs that are incorrectly classified as legitimate URLs.

Table. 1 shows the comparison of accuracy between DNN and ANN based phishing attack detection for Ham, Phishing Corpus and Phishload datasets.

Table- I: Comparison of Accuracy

Datasets	ANN	DNN
Ham	0.77	0.9
Phishing Corpus	0.78	0.92
Phishload	0.76	0.89

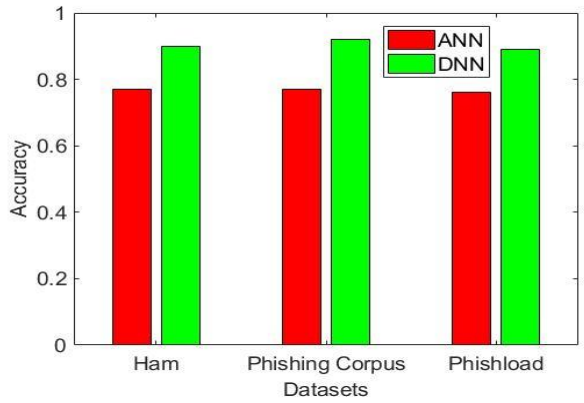

Fig. 1.Comparison of Accuracy

Fig. 1 shows the comparison between ANN and DNN based phishing attack detection in terms of accuracy for three different datasets. The Ham, Phishing Corpus, Phishload datasets are taken in x-axis and the accuracy value is taken in y-axis. For Phishing Corpus dataset, the accuracy of DNN based phishing attack detection is 17.95% greater than ANN based phishing attack detection. From this analysis it is proved that the proposed DNN based phishing attack detection has high accuracy than ANN based phishing attack detection.

B. Precision

Precision measures the exactness of the classifier, i.e., what percentage of URLs that the classifier labeled as phishing URLs and it is given by,

$$Precision = \frac{TP}{TP + FP}$$

Table. 2 shows the comparison of precision between DNN and ANN based phishing attack detection for Ham, Phishing Corpus and Phishload datasets.

Table- II: Comparison of Precision

Datasets	ANN	DNN
Ham	0.76	0.88
Phishing Corpus	0.76	0.89
Phishload	0.75	0.87

Fig. 2 shows the comparison between ANN and DNN based phishing attack detection in terms of precision for three different datasets. The Ham, Phishing Corpus, Phishload datasets are taken in x-axis and the precision value is taken in y-axis. For Phishing Corpus dataset, the precision of DNN based phishing attack detection is 17.11% greater than ANN based phishing attack detection. From this analysis it is proved that the proposed DNN based phishing attack detection has high precision than ANN based phishing attack detection.

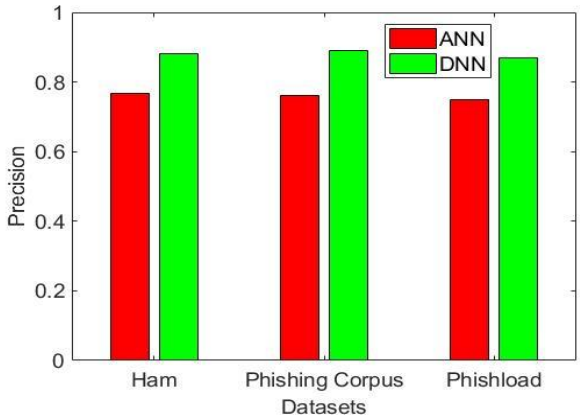

Fig. 2.Comparison of Precision

C. Recall

Recall measures the completeness of the classifier results, i.e., what percentage of phishing URLs did the classifier label as phishing, and is given by

$$Recall = \frac{TP}{TP + FN}$$

Table. 3 shows the comparison of recall between DNN and ANN based phishing attack detection for Ham, Phishing Corpus and Phishload datasets.

Table- III: Comparison of Recall

Datasets	ANN	DNN
Ham	0.77	0.87
Phishing Corpus	0.77	0.88
Phishload	0.76	0.88

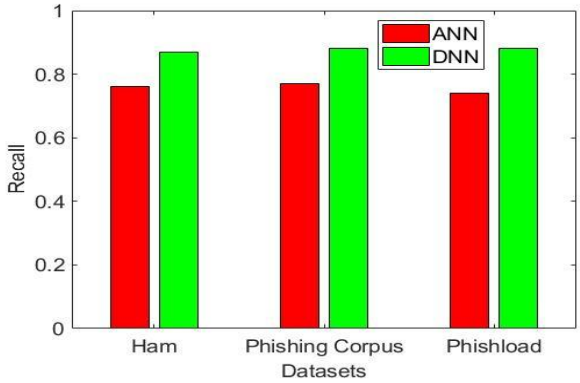

Fig. 3.Comparison of Recall

Fig. 3 shows the comparison between ANN and DNN based phishing attack detection in terms of recall for three different datasets. The Ham, Phishing Corpus, Phishload datasets are taken in x-axis and the recall value is taken in y-axis. For Phishing Corpus dataset, the recall of DNN based phishing attack detection is 14.29% greater than ANN based phishing attack detection. From this analysis it is proved that the proposed DNN based phishing attack detection has high recall than ANN based phishing attack detection.

D. F-measure

F-measure is computed as the harmonic mean of the precision and recall. It is calculated as,

$$F - measure = 2 \times \frac{Precision \times Recall}{Precision + Recall}$$

Table. 4 shows the comparison of f-measure between DNN and ANN based phishing attack detection for Ham, Phishing Corpus and Phishload datasets.

Table- IV: Comparison of F-measure

Datasets	ANN	DNN
Ham	0.77	0.87
Phishing Corpus	0.76	0.885
Phishload	0.755	0.878

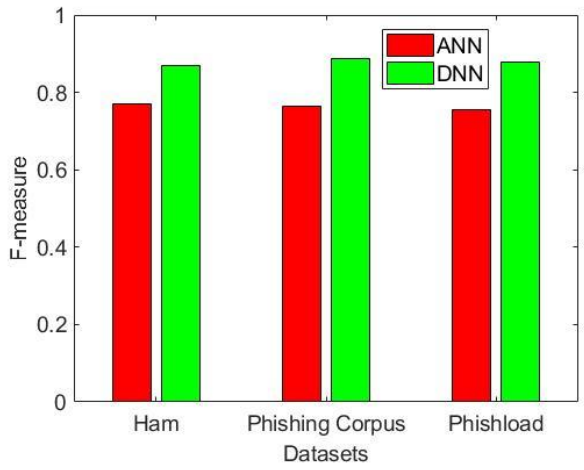

Fig. 4.Comparison of F-measure

Fig. 4 shows the comparison between ANN and DNN based phishing attack detection in terms of f-measure for three different datasets. Ham, Phishing Corpus, Phishload datasets are taken in x-axis and the f-measure value is taken in y-axis. For Phishing Corpus dataset, the recall of DNN based phishing attack detection is 16.45% greater than ANN based phishing attack detection. From this analysis it is proved that the proposed DNN based phishing attack detection has high f-measure than ANN based phishing attack detection.

V. CONCLUSION

Phishing detection in social network platform is considered to be a recent ever growing process that focuses on attaining higher values of detection accuracy. In this paper, a Deep Neural Network (DNN) is introduced for phishing attack detection. DNN gets 30 feature vector based on URL-based data, HTML-based data and domain-based data. Then, the weight and bias values are initialized in the input layer. Sigmoid functions are processed in the hidden layers and it is given as input to the next layer. Multiple hidden layers are used in the DNN. Finally, the output layer of DNN classifies the URLs as either legitimate or phishing URL. By using deep learning, the high-level features are learned effectively in an incremental manner which improves the accuracy, precision, recall and f-measure. The experimental results show that the proposed DNN method has high accuracy, precision, recall and f-measure than the ANN method for Ham, Phishing Corpus and Phishload datasets.

REFERENCES

1. R. S. Rao, and S. T. Ali, "PhishShield: a desktop application to detect phishing webpages through heuristic approach," *Procedia Comput. Sci.* vol. 54, 2015, pp. 147-156.
2. O. A. Akanbi, I. S. Amiri and E. Fazeldehkordi, "Cyber security: A machine learning approach to phishing," 2014.
3. T. Chin, K. Xiong, and C. Hu, "Phishlimiter: A Phishing Detection and Mitigation Approach Using Software-Defined Networking," *IEEE Access*, vol. 6, 2018, pp. 42516-42531.
4. Y. Li, R. Xiao, J. Feng, and L. Zhao, "A semi-supervised learning approach for detection of phishing webpages," Optik, vol. 124, no. 23, 2013, pp. 6027-6033.
5. M. Moghimi, and A. Y. Varjani, "New rule-based phishing detection method," *Expert syst. appl.*, vol. 53, 2016, pp. 231-242.
6. G. A. Montazer, and S. ArabYarmohammadi, "Detection of phishing attacks in Iranian e-banking using a fuzzy–rough hybrid system," Appl. Soft Comput., vol. 35, 2015, pp. 482-492.
7. W. E. Hadi, F. Aburub, and S. Alhawari, "A new fast associative classification algorithm for detecting phishing websites," Appl. Soft Comput., vol. 48, 2016, pp. 729-734.
8. K. L. Chiew, C. L. Tan, K. Wong, K. S. Yong, and W. K. Tiong, "A new hybrid ensemble feature selection framework for machine learning-based phishing detection system," *Inf. Sci.*, vol. 484, 2019, pp.153-166.
9. S. Smadi, N. Aslam, and L. Zhang, "Detection of online phishing email using dynamic evolving neural network based on reinforcement learning," *Deci. Support Syst.*, vol. 107, 2018, pp. 88-102.
10. E. Zhu, Y. Chen, C. Ye, X. Li, and F. Liu, "OFS-NN: An Effective Phishing Websites Detection Model Based on Optimal Feature Selection and Neural Network," *IEEE Access*, vol. 7, 2019, pp. 73271-73284.

AUTHORS PROFILE

K. Sumathi, is a student of CMS college of science and commerce, affiliated to Bharathiar university, Coimbatore, Tamilnadu, India. She is pursing Ph.D in Computer Science. She is doing research in the area of information security.

V. Sujatha, has 16 years of teaching experience and 2 years of IT Industrial experience. Her area of specialization is web mining, IoT and Big Data Analysis. She has published 24 research articles in National and International Journals and also presented papers in several National Conferences, Seminars and Workshops. She is currently guiding M.Phil and Ph.D Scholars. She has an ideal knowledge in programming languages, DOT NET frameworks and has developed two live projects using Visual programming. She also sets question papers for universities in TamilNadu.