WEB PHISHING DETECTION USING A DEEP LEARNING FRAMEWORK

1. Introduction

Web service is a communication protocol and software between two electronic devices over the Internet. Webservices extends the World Wide web infrastructure to provide the methods for an electronic device to connect toother electronic devices. Web services are built on top of open communication protocols such as TCP/IP, HTTP, Java, HTML, and XML. Web service is one of the greatest inventions of mankind so far, and it is also the most profound manifestation of computer influence on human beings. With the rapid development of the Internet and the increasing popularity of electronic payment in web service, Internet fraud and web security have gradually been the main concern of the public.

Web Phishing is a way of such fraud, which uses social engineering technique through short messages, emails, and We Chat to induce users to visit fake websites to get sensitive information like their private account, token for payment, credit card information, and soon. The first phishing attack on AOL (America Online) can be traced back to early 1995. A phisher successfully obtained AOL users personal information. It may lead to not only the abuse of credit card information, but also an attack on the online payment system entirely feasible.

The phishing activity in early 2016 was the highest ever recorded since it began monitoring in 2004.The total number of phishing attacks in 2016 was 1,220,523. This was a 65 percent increase over 2015. In the fourth quarter of 2004, there were 1,609 phishing attacks per month. In the fourth quarter of 2016, there was an average of 92,564 phishing attacks per month, an increase of 5,753% over 12 years. According to the 3rd Microsoft Computing Safer Index Report released in February 2014, the annual worldwide impact of phishing could be as high as $5 billion. With the prevalence of network, phishing has become one of the most serious security threats in modern society, thus making detecting and defending against web phishing an urgent and essential research task.

Web phishing detection is crucial for both private users and enterprises. Some possible solutions to combat phishing were created, including specific legislation and technologies. From a technical point of view, the detection of phishing generally includes the following categories: detection based on a blacklist and white list, detection based on Uniform Resource Locator (URL) features, detection based on web content, and detection based on machine learning. The anti-phishing way using blacklist may be an easy way, but it cannot find new phishing websites. The detection on URL is to analyze the features of URL. The URL of phishing websites may be very similar to real websites to the human eye, but they are different in IP. The content-based detection usually refers to the detection of phishing sites through the page of elements, such as form information, field names, and resource reference.

2. Related Works

Researchers have conducted lot of work in security [12–18], including secure routing [19–21], intrusion detection [22–27], intrusion prevention, and smart grids security. Different from research problems in wireless networks [30–60] and energy networks [61–64], web phishing is the attempt to acquire sensitive information such as usernames, passwords, and credit card details, often for malicious reasons, by masquerading as a trustworthy website on the Internet. Researchers present some solutions to detect web phishing as follows. When we judge whether a specific website is web phishing, the direct way is to use a white list or black list. We may search the URL in some database and decide.

Pawan Prakash e t al presented two ways to detect phishing websites by the blacklist. The first way includes five heuristics to enumerate simple combinations of known phishing sites to discover new phishing URLs. The second way consists of an approximate matching algorithm that dissects a URL into multiple components that are matched individually again sentries in the blacklist. Many well-known browser vendors such as Firefox and Chrome also used a self-built or third-party black-white list, to identify whether the URL is a phishing site. This method is very accurate, but its blacklist or whitelist usually relies on manual maintaining and reviewing. Obviously, these methods are not real time and may cost a lot of time and effort.

3. The Phishing Detection Model Based on DBN

Phishing Feature Extraction and Definition. First, we get real traffic flow from ISP. The data set includes traffic flow for 40 minutes and 24 hours. We construct the graph structure of traffic flow and analyze the characteristics of web phishing from the view of the graph. Each piece of data contains the following fields.

1. AD: user node number.
2. IP: user IP address.
3. TS: access time.
4. URL: Uniform Resource Locator, access web address.
5. REF: request page source.
6. UA: user browser type.
7. $DST-IP$: server address to access.
8. CKE: User Cookie.

A graph is mathematical structures used to model pairwise relations between objects. It is also a very direct way to describe the relationship between nodes in a network. The relationship between the nodes on the Internet can also be expressed through the graph structure. Therefore, we construct a graph to store the real traffic flow data and describe the relationship between the nodes in traffic flow.

Give an undirected graph $G = (V, E)$, where V includes two kinds of node:
(i) user node AD;
(ii) access URL and REF.

$E \subset V \times V$ denotes an access relationship between REF, AD, and URL. The vertices of the graph $G = (V, E)$ are as follows:

(i) User node VAD has one attribute: total access times (vertex out-degree).

(ii) User node $VURL$ has two attributes: total accessed times (vertex in-degree) and website registration time.

The edges of the graph $G = (V, E)$ are as follows:

(i) The number of visits: which corresponds to the number of occurrences of the edge, the number of times an AD may have access to a URL, or the number of direct links between two URLs, depending on the corresponding vertex type.

(ii) Cookie: the cookie field in the access record.

(iii) UA: User Agent in the access record.

Feature Definition.

We define two kinds of features to detect web phishing, and they are an original feature and interactive feature.

Original Feature.

There are some features in the phishing URL, such as special characters. We definite these features in URL as an original feature as follows:

(i) $O1$: there are special characters in URL, such as @, Unicode, and so on. Those special characters are not allowed in a normal URL.

(ii) $O2$: there are too many dots or less than four dots in normal URL.

(iii) $O3$: the age of the domain is too short. For example, the age of the normal domain is more than 3 months.

Interaction Feature.

There are some features in graph $G = (V, E)$, such as access frequency. We define these features

through a node relationship as interaction feature as follows:

(i) $I1$: in-degree of URL node from REF is very small. In general, the normal websites do not link to phishing sites. The phishing sites are directly accessed.

(ii) $I2$: out-degree of URL node is very small. In order to get personal private information, the phishing sites are usually terminal websites and do not link to the other sites.

(iii) $I3$: the frequency of URL from AD is one. In general, one user accesses the phishing site only one time and the user cannot access the phishing site more than one time.

(iv) $I4$: when AD accesses URL, user browser type UA is not the main browser. Well-known browser vendors often have a built-in filtering phishing site plug-in. A user who uses unknown browsers is more likely to access the phishing sites.

(v) $I5$: there is no cookie in user. The phishing site does not leave its cookie in user.

4.ALGORITHM.

Require: Visible Layer V = {V1, ..., Vm}, Hidden Layer H = {h1, ..., hn}
Ensure: Gradient Approximation $\Delta\theta \leftarrow$ Δwij, Δai, Δbj for i in {1...n}, j in {1...m}
1: **for** i in {1...n}, j in {1...m} **do**
2: Initialize $\Delta wij = \Delta ai = \Delta bj = 0$
3: **end for**
4: **for** Each V in V **do**
5: V0 $\leftarrow$ V
6: **for** t in {0...$k-1$} **do**
7: **for** i in {1...n} **do**
8: Sample ht
$i \sim p(hi|\mathrm{V}t)$
9: **end for**
10: **for** j in {1...m} **do**
11: Sample Vt
$j \sim p(\mathrm{V}j|ht)$
12: **end for**
13: **end for**
14: **end for**
15: **for** i in {1...n}, j in {1...m} **do**
16: $\Delta wij \leftarrow \Delta wij + p(hi|\mathrm{V}0)$ V0
$j - p(hi|\mathrm{V}k)$ Vk
j
17: $\Delta ai \leftarrow \Delta ai + p(hi|\mathrm{V}0) - p(hi|\mathrm{V}k)$
18: $\Delta bj \leftarrow \Delta bj +$ V0
$j -$ Vk
j
19: **end for**

4.1. ALGORITHM 2:

Require: Period T, Learning Rate η, Momentumρ, Visible Layer V, Hidden Layer H, Number of visible and hidden layer units nV, nh, Offset Vector a, b, Weight Matrix$W$
Ensure: θ = {W, a, b}
1: Initialize W, a, b
2: **for** $i \in$ {1...T} **do**
3: Calling CD-k to generate $\Delta\theta$ = {ΔW, Δa, Δb}
4: $W \leftarrow \rho W + \eta((1/n\mathrm{V})\ \Delta W)$
5: $a \leftarrow \rho a + \eta((1/n\mathrm{V})\ \Delta a)$
6: $b \leftarrow \rho b + \eta((1/n\mathrm{V})\ \Delta b)$
7: **end for**

5. Test and Analysis

Test Data and Evaluation Criterion. The test data come from ISP and are composed of two data sets. The small data set includes real traffic flow for 40 minutes. The big data set includes real traffic flow for 24 hours. After pretreatment, we get record sum, unique IP, unique AD, and unique URL as in Table 1. This paper belongs to a classical binary classification model application. In the binary classification model, the results are usually marked as Positive (P) or Negative (N). In this paper, the corresponding node is either a phishing site or not a phishing site. Then with the classification results with a priori facts, there will be the following four categories:

(i) True Positive (TP): is actually P and the classification is also P

(ii) False Positive (FP): is actually N and the classification is also P

(iii) True Negative (TN): is actually N and the classification is also N

(iv) False Negative (FN): is actually P and the classification is also N

The above classification data can generate four categories of evaluation criterions with details as follows:

(i) Accuracy (ACC): $ACC = (TP+TN)/(TP+TN+FP+FN)$

(ii) True Positive Rate (TPR, Recall): $TPR = TP/ (TP +FN)$

(iii) False Positive Rate (FPR, Fall-Out): $FPR = FP/ (FP +TN)$

(iv) Positive Predictive Value (PPV, Precision): $PPV =TP/ (TP + FP)$

SVM model can be seen as a shallow feature extraction (with a hidden layer). DBN selects at least two layers in order to relatively enhance the feature selection effect, and too many layers will lead to overfitting. DBN main module declaration is as in Listing 1. Some parameters are explained as follows:

(i) layers: the number of nodes per layer. Here, as the visible layer has a total of 10 different variables as a

set of features, select 10 as the number of visible layer nodes.

(ii) inputs: the matrix to be trained.

(iii) initial Learning Rate: learning rate.

(iv) momentum: learning rate correction momentum. Select the default value.

(v) use Binary Values Visible Reconstruction: whether to use the binary value to reconstruct the visible layer.

Select the initial value false.

(vi) std Weights: the upper and lower bounds of the weight matrix are initialized.

6. Conclusions

In this paper, we analyze the features of phishing websites and present two types of features for web phishing detection: original feature and interaction feature. Then we introduce DBN to detect phishing websites and discuss the detection model and algorithm for DBN. We train DBN and get the appropriate parameters for detection in the small data set. In the end, we use the big data set to test DBN and TPR is approximately 90%.

Data Availability

The test data used to support the findings of this study have not been made available because these data belong to the ISP (Internet Service Provider).

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

This work is supported by National Natural Science Foundation of China (61571290, 61831007, and 61431008), the NSFC Zhejiang Joint Fund for the Integration of Industrialization and Informa ionization under Grant U1509219, and Shanghai Municipal Science and Technology Project under Grants16511102605 and 16DZ1200702 and NSF Grants 1652669 and 1539047.

7.References

[1] https://en.wikipedia.org/wiki/Web service.

[2] O. Adam, Y. C. Lee, and A. Y. Zumaya, "Stochastic resource provisioning for containerized multi-tier web services in clouds," *IEEE Transactions on Parallel and Distributed Systems*, vol. 28, no. 7, pp. 2060–2073, 2017.

[3] T. Bujlow, V. Carela-Espanol, J. Sole-Pareta, and P. Barlet-Ros, "A survey on web tracking: Mechanisms, implications, and defenses," *Proceedings of the IEEE*, vol. 105, no. 8, pp. 1476–1510, 2017.

[4] H.-C. Huang ,Z.-K. Zhang,H.-W. Cheng, and S.W. Shieh, "Web application security: Threats, countermeasures, and pitfalls," *The Computer Journal*, vol. 50, no. 6, pp. 81–85, 2017.

[5] https://en.wikipedia.org/wiki/WeChat.

[6] K. Rekouche, *Early phishing*, 2011.

[7] http://www.antiphishing.org/.

[8] Microsoft, "20% Indians are victims of online phishing attacks: Microsoft," *IANS*, 2014, http://news.biharprabha.com/.

[9] L.Wu,X.Du, andJ.Wu, "Effectivedefense schemes for phishing attacks on mobile computing platforms," *IEEE Transactions on* Vehicular Technology, vol. 65, no. 8, pp. 6678–6691, 2016.