Survey on Phishing attack Detection and Mitigation

Jibi Mariam Biju[1], Anju J Prakash[2]

[1]M.Tech, CSE Department, Sree Buddha College of Engineering, Kerala, India
[2]Assistant Professor, CSE Department, Sree Buddha College of Engineering, Kerala, India

---***---

Abstract - *Phishing is one of the riskiest social engineering techniques that crush users where attackers gain access to our critical information system. It is an easy and efficient method to deceit people using internet. Phishing can be done through the use of e-mail communication with an embedded hyperlink. Due to the intricacy of the current phishing attack, the detection and extenuation of phishing attack is a grand challenge. Static detection rules that are earlier used are not effective in the real world due to the dynamics of phishing attacks. Later a new technique for Deep Packet Inspection (DPI) and then it is leveraged with Software-Defined Networking (SDN) to identify phishing activities through e-mail and web-based communication. Those methods are mentioned in this paper.*

Key Words: Phishing, Software Defined Network (SDN), Artificial Neural Network (ANN).

1. INTRODUCTION

An act of attempting to retrieve user details such as username, password and credit card information as a responsible entity in an electronic communication. Amongst a daunting pile of unread e-mail messages lies a precarious hyper link to a malicious website, waiting for a user's click. This message sent by a threat actor potentially compromise an organization or computing system. Many organisms conduct programs to inform users to be more careful about such unsure emails, but many end users fail to adhere to such policies that lead to potential data loss. Phishing has become one of the deadly attacks. Phishing attacks' main incentive is through the use of e-mail communication and such attacks may result in data loss, compromised accounts, and remote code execution techniques.

Since phishing attacks aim at exploiting weaknesses found in humans, it is difficult to mitigate them. Several approaches have been introduced to thwart phishing attack. The use of vendor-based solution is a kind of traditional method to deter phishing email, but they do not prevent an end user from clicking on a malicious link. To deter such issues, a variety of Intrusion Detection Systems (IDS) and Intrusion Prevention Systems (IPS) approaches is proposed to identify and to deter phishing e-mails, but they lack feasibility or cannot be used when e-mail communication becomes encrypted. The challenges faced by those traditional methods include mitigation time and network performance.

1.1 Types of phishing attack

a. Deceptive phishing: The attacker tries to attain private information from the victims. Attackers use such information to pilfer money or to introduce other kind of attacks.

b. Spear phishing: Aims at specific individual instead of a group of people. Attackers look for their target on social media and other sites. This kind of attack is the first step to enter into company's defenses and carry out a targeted attack.

c. Whaling: When attackers aims at people like CEO, it is called whaling. This type of attack is concerned with high level executives because they are able to access a great deal of company information.

d. Pharming: Leads users to a fraudulent website that appears to be legitimate. Attackers can affect either the user's computer or the server and redirect the user to an illegitimate site even if the correct URL is typed in.

1.2 Causes of phishing

- Lack of user awareness
- Sensitivity of browsers
- Restricted use of digital signatures
- Ambiguous e-mails
- Incorrect source address
- Unavailability of secure desktop tools

1.3 Phishing Techniques

a. Link manipulation: this technique makes use of links in an e-mail appear to belong to spoofed organization. The common trick used by the phishers are misspelled URL's or the use of sub domain.

b. Filter evasion: In order to make the detection of text commonly used in phishing email harder for the anti-phishing filters, instead of text phishers make use of images

c. Website Forgery organization: There are certain cases in which some phishing scams use Java script commands in order to change the address bar. This can be done by either setting a picture of genuine URL over the address bar or by closing the original address bar and opening a new one with genuine URL.

d. Phone Phishing: This kind of phishing involves messages that are stated to come from a bank told users to dial a phone number regarding issues with their bank accounts. Once the phone number is dialed, it prompts told users to enter their account number and pin.

2. RELATED WORKS

Researchers have devoted a variety of techniques for preventing phishing attack. The phishing detection algorithm infeasibility is due to two major issues:(1) Due to the dynamics of phishing attack in real time, their detection rules cannot be changed dynamically as those rules are static. (2) Detection and mitigation time for phishing attack are too long to be determined.

Blacklisting is a common method to protect against Phishing attack. Pawan Prakash et al [1] proposed PhishNet to address the problems associated with blacklisting. The major problem with blacklisting is incompleteness, even though blacklist provides design simplicity and ease of implementation. PhishNet consists of two major componenets:1) URL Prediction component, inspects the current blacklist and methodologically generates new URL by employing various heuristics. Then with the help of DNS queries and content matching techniques, it tests whether this newly generated URL is malicious.2) approximate URL matching component, with the existing blacklist it performs an appropriate match of a new URL.

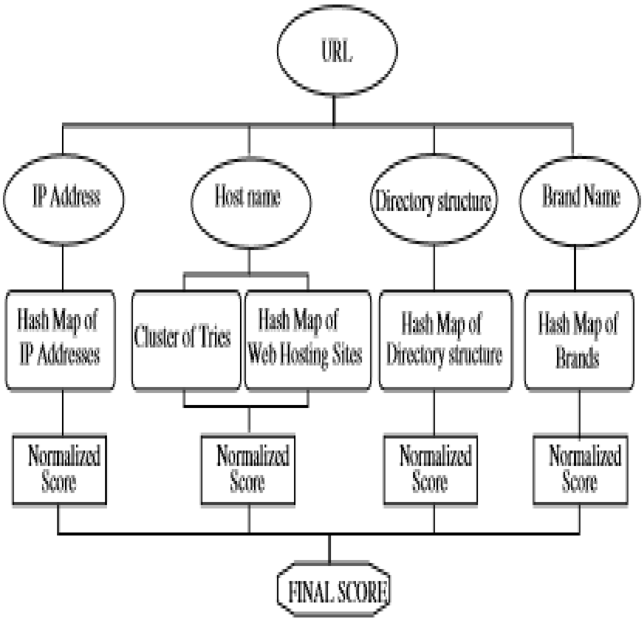

Fig 2.1: Computing the score of a new URL in PhishNet. In order to synthesis the new URL from the existing blacklist entities, five different heuristics are used. Once the new URL's are created, it is subjected to a validation process. In approximate matching, first the input URL is break down into four different entities-IP address, hostname, directory structure and brand name. Then the individual entities are scored by matching them with the corresponding fragments of the original entries. This is done to generate one final score. If this score calculated is greater than the pre-defined threshold, it indicates the URL as a phishing site.

Martin Roesch et al [2] introduced SNORT which is a lip-cap based packet sniffer and logger. This can be used as a light weight Network Intrusion detection System. SNORT detects a variety of attacks and probs and perform content matching. Command line switches and optional Berkeley Packet Filter commands can be used to configure SNORT.

The major advantage of this method is that the decoded output of SNORT is more user friendly than tcpdump's output. tcpdump is a packet analyzer that works on command line. The main function is that it enables the user to display TCP/IP and other packets being transmitted or received over the network to which the computer is connected. SNORT mainly aims at gathering packets and process them in a detection engine. This comprises of three subsystems: the packet decoder, the detection engine, and the logging and alerting subsystem. The main function of the packet decoder is to set pointers into the packet data for latter analysis by the detection engine. The decode engine is implemented with the help of protocol stack and each of the subroutine imposes order on the packet data. These decoding routines are called through the protocol stack in an in-order form and ends up at the application layer.

A two-dimensional linked list is used for maintaining the detection rule in the SNORT. For each packet in both the directions, the rule chains are searched recursively in the detection engine. The rule that first matches the decoded packet in the detection engine triggers the action specified in the rule definition and returns. The alerting/logging subsystem is selected at run-time. This consists of three logging and five alerting options. Alerts may either be sent to syslog or can be sent as WinPopup messages. The syslog alerts are sent as security messages whereas WinPopup messages are sent as event notification.

SNORT rules are rules that helps to detect a variety of unfavorable or merely suspicious network traffic and are simple to write. Pass, log or alert are the three action derivatives that are used by SNORT when a packet matches with the specified with the rule pattern. Pass rule drop the packet. Log rule writes the user selected full packet to the logging system. The function of the alert rule is to generate an event notification and then log the full packet using the selected logging mechanism to enable later analysis.

Malicious web page also compromises host on the internet. A third party can embed a malicious script into a web page and when a user visits this page, the script is first executed and
thereby compromises the browser. Davide Canali et al [3] designed a filter called Prophiler that can quickly discard pages that are benign, forwarding only the pages that are likely to contain malicious code. This filter uses static analysis techniques to quickly examine a web page for malicious content. This analysis makes use of the features derived from HTML content of the page, from corresponding URL and from the associated Javascript code.

The aim of Prophiler is to differentiate web pages that are collected by the web crawler. The features extracted from the web pages helps in this classification and these are collected from the page's content and page's URL. HTML based features include statistical information about the raw content of the page and structural information that is obtained from parsing the HTML code. JavaScript features result from the static analysis of the JavaScript file. In order to make those feature extraction difficult, malicious JavaScript scripts are obfuscated and packed. By examining the URL of certain web page, it is possible to predict whether that web page is malicious. Some information contained in the URL and associated with the referenced host can be used to help in the detection of malicious web pages. But the problem with this method is that feature collection time is very slow.

CANTINA is a novel based approach by Yue Zhang et al [4] that aims to detect phishing web sites. The concept behind this method is TF-IDF information retrieval algorithm. In order to determine whether a web page s legitimate or not, CANTINA examines the content of a web page. The term frequency (TF) is defined as the number of times a given term appears in a particular document. The inverse document frequency (IDF) is a measure of the general relevance of the term. To measure how important a given word is to a document in a corpus, the TF-IDF yield a weight.

After the calculation of TF-IDF score for a given web page, CANTINA generate a lexical signature and is feed into the search engine. If the domain name of the current web page matches the domain name of the N top search results, then it is considered as a legitimate web site. Otherwise, it is a phishing site. The problem with this method is that it does not have any dictionary for any languages other than English. Also it have some performance problem due to the time lag involved in querying Google.

Ningxia Zhang et al [5] applied a multilayer feed forward neural network to detect phishing emails. Machine learning for phishing classification involves the use of feature set extracted from an email and apply a learning algorithm to classify an email to phishing or legitimate based on the selected features.

The first step in this approach is to select and define a set a feature that capture the characteristics of the phishing emails. Structural features, link features, element features and word list features are the kind of features that are extracted from the phishing email. For the purpose of phishing detection, multilayer feed forward neural network is used. In this case the connections between neurons does not form a directed cycle. The number of input and output depends on the number of computational units in the input and output layer.

A serious problem that is found in large network with large number of parameters is over fitting. Large networks are very slow to use and thus it is difficult to deal with over fitting. The technique that is introduced to solve this problem is dropout. Nitish Srivastava et al [6] mention about the problems caused by overfitting and how it can be over come with the help of dropout. Dropout can be defined as the dropping out of units in a neural network i.e temporarily remove the units from the network along with all its incoming and outgoing connections. For each of the input layer, a vector of independent Bernoulli random variables is obtained, sampled and multiplied element-wise with the outputs of that layer to get the thinned output which are used as input to the next layer. This process is applied at each layer and it helps to reduce to a sub network from a large network. This resulting network is used without dropout which can be used for classification problem. The main drawback of dropout is the increase in training time.

The major issue that is faced by SDN (Software Defined Network) is data to control plane saturation attack. Haopei Wang et al [7] addresses this issue. In order to consume resource in both data and control plane, an attacker can produce a large amount of table-miss packet_in messages. To mitigate this security threat, an efficient, lightweight and protocol-independent defense framework called FLOODGUARD for SDN networks is introduced. Here table-miss refers to the new flow for which there is no matching flow rules installed in the flow table.

FLOODGUARD consists of two modules: proactive flow rule analyzer and packet migration. The proactive flow rule analyzer generates proactive flow rules once they are activated and are directly installed into the data plane switches. These flow rules are updated dynamically by the flow rule analyzer. The packet migration modules comprise of two components: migration agent and data plane cache. The main function of the migration agent is to detect saturation attack. Along with this, it also helps to migrate table-miss packets to the data plane. Data plane cache temporarily caches table-miss packets during saturation attack. The data plane cache will slowly send the table-miss packets as packet_in messages to the controller by using rate limiting and round-robin scheduling algorithm.

David Hay et al [8] describes Deep Packet Inspection (DPI) as a service to the middle box. Middle box is a device found in the computer network that converts, examines, filter and manipulates network traffic such as firewall. Deep Packet Inspection (DPI) is a data processing method that examines in detail the data being sent over a computer network. Here the payload of the packet is inspected against a set of patterns. DPI is a common task in many middleboxes. The entity that maintain the DPI process across the network is DPI controller and is responsible to communicate with SDN to realize the appropriate data.

Registration and pattern set management are the two kind of procedure that takes place between middleboxes and DPI controller. DPI controller's task is to register middleboxes that use its service. A middlebox registers itself to the DPI service using a registration message. It may inherit the pattern set of an already registered middlebox. The DPI service responsibility is only to indicate appearances of patterns and these are added to and removed from the DPI controller using dedicated messages from middleboxes to the controller. In order to maintain multiple pattern sets, a virtual DPI algorithm is presented. Finally, the DPI controller initializes DPI service instances.

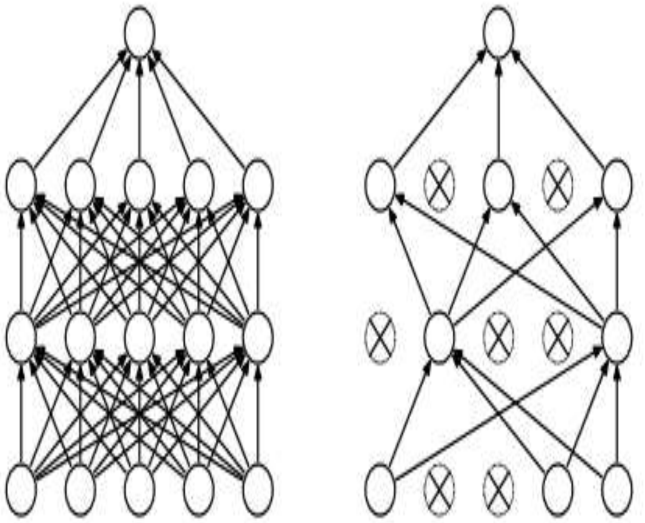

Fig 2.2: a) Standard Neural Net b) After applying dropout

Ben Pfaff et al [9] describes the design and implementation of Open vSwitch. It is an open source and multi-platform. It is known for its flexibility and general purpose usage. Open vSwitch is mainly used by people who select an operating system distribution and hypervisor. This made the Open vSwitch design to be quite modular and portable. The two important components that direct packet forwarding in Open vSwitch is ovs-vswitchd and datapath kernel module. From a physical NIC or a VM's virtual NIC, the datapath module in the kernel receives the packets. In user space, it is the ovs-vswitchd determines how the packet should be handled, it then passes the packet back to the datapath with the desired handling. Open vSwitch is commonly used as an SDN switch, and OpenFlow is the main way to control forwarding. The main function is that it allows a controller to add, remove, update, monitor and obtain statistics on flow tables and their flow.

In Open vSwitch, the OpenFlow table from SDN controller receives ovs-vswitchd and matches the received packets from the datapath module against these OpenFlow tables.it then gathers the actions applied and the result is cached to the kernel datapath. For the packet classification, Open vSwitch uses tuple search classifier. Such flow table is implemented as a single hash table in tuple search classifier.

Tommy Chin et al [10] presents a new detection and mitigation approach called PhishLimiter where it proposes new technique for Deep Packet Inspection (DPI) and then leverage it with Software-Defined Networking (SDN) to identify phishing activities through e-mDPI approach uses two modes: Store and Forward (SF) and a Forward and Inspect (FI). In order to decide whether to choose SF or FI, for each of the incoming packet a score called PhishLimiter Score (PLS) is calculated. It is then compared with the predefined threshold value on the Open Vswitch called an OVS score. For each of the SDN flow, if the PLS value is greater than the OVS threshold value then the packet is placed in SF mode. In this mode, the packet is first placed in a buffer undergoes detection for malicious activities. With the help of ANN approach, if the packet is phishing then it is dropped otherwise it is forwarded. Then the value of PLS is updated and is compared again for the next packet. If the PLS value is less than the OVS threshold value, then the packet is placed in FI mode. Under this mode, packet is forwarded to the destination and copied for inspection. The copied packet is dropped once PhishLimiter determines if it has malicious intent.

ANN classifier algorithm has the features extracted from the URL and web-based code as an input value. The main type of features that are extracted includes URL based features, HTML based features and Domain based features. And the feature extractor algorithm outputs a vector that consist of thirty features. After the model is loaded, the input feature vector is attached to the output matrix as the first layer. In every loop of computation, weights matrix w is multiplied with previous layer result and then add the bias b. The resulting value is applied to a sigmoid function used as the next layers input. The sigmoid function has the value range from 0 to 1. By adjusting the size of the layer, it output only one probability defined as L. When L < 0.5 then the URL as non-phishing; otherwise, as phishing URL. ANN model provides the best average accuracy as compared to SVM, Logistic regression and Naive Bayes.

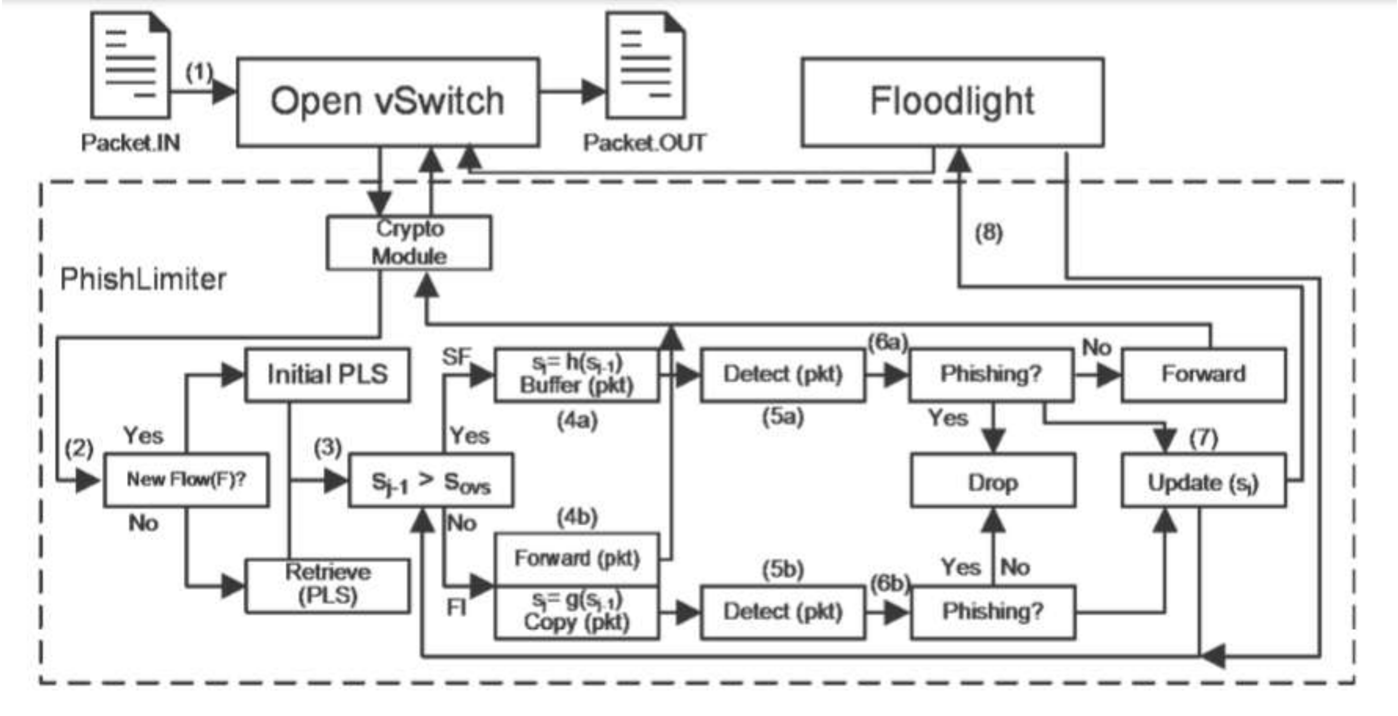

Figure 2.3: The detailed mechanism of PhishLimiter

3. CONCLUSION

Phishing attacks causes a major threat for computer system defenders, often forming the first step in a multi-stage attack. Traditional solutions described in the literature review focus on the use of inline inspection techniques such as an IPS or proxy service based on static string matching are not much feasible. PhishLimiter has the ability to handle network traffic dynamics for containing phishing attacks and can provide a better traffic management since it has a global view of networks due to SDN. An ANN model using a PLS system is developed for classifying phishing attack signature.

REFERENCES

[1] P. Prakash, M. Kumar, R. R. Kompella, and M. Gupta, “Phishnet: predictive blacklisting to detect phishing attacks,” in INFOCOM. IEEE, 2010, pp. 1–5.

[2] M. Roesch, “SNORT-lightweight intrusion detection for networks.” in USENIX LISA, 1999.

[3] D. Canali et al., “Prophiler: A fast filter for the large-scale detection of malicious web pages,” in Proceedings of the 20th International Conference on World Wide Web, 2011.

[4] Y. Zhang, J. I. Hong, and L. F. Cranor, “mboxCANTINA: A contentbased approach to detecting phishing web sites,” in Proceedings of the 16th International Conference on World Wide Web, ser. WWW ’07. New York, NY, USA: ACM, 2007, pp. 639–648.

[5] N. Zhang and Y. Yuan, “Phishing detection using neural network,” Department of Computer Science, Department of Statistics, Stanford University, Available at http://cs229. stanford. edu/proj2012/ZhangYuan Phishing Detection Using Neural Network. pdf (Accessed 23 April 2016), 2013.

[6] N. Srivastava, G. E. Hinton, A. Krizhevsky, I. Sutskever, and R. Salakhutdinov, “Dropout: a simple way to prevent neural networks from overfitting.” Journal of Machine Learning Research, vol. 15, no. 1, pp. 1929–1958, 2014.

[7] H. Wang et al., “FloodGuard: A DoS attack prevention extension in software-defined networks,” in DSN, 2015.

[8] A. Bremler-Barr, Y. Harchol, D. Hay, and Y. Koral, “Deep packet inspection as a service,” in Proceedings of the 10th ACM International on Conference on emerging Networking Experiments and Technologies. ACM, 2014, pp. 271–282.

[9] B.Pfaffetal.,“The design and implementation of openvSwitch,”in12th USENIX Symposium on Networked Systems Design and Implementation (NSDI), 2015, pp. 117–130.

[10] Tommy Chin, Member, IEEE, Kaiqi Xiong, Senior Member, IEEE, and Chengbin Hu, “PhishLimiter: A Phishing Detection and Mitigation Approach Using Software-Defined Networking.”

BIOGRAPHIES

Jibi Mariam Biju, she is currently pursuing M.tech in Computer Scinece and Engineering in Sree Buddha College of Engineering, Elavumthitta. Her research areas include the field of data mining, cryptography and security.

Anju J Prakash is working as Asst.Professor in Computer science and engineering in Sree Buddha College of Engineering, meanwhile pursuing her PhD in the field of image processing or data mining from Noorul Islam Centre for higher education.