International Journal of Science and Research (IJSR)
ISSN (Online): 2319-7064
Index Copernicus Value (2013): 6.14 | Impact Factor (2013): 4.438

Survey Paper on Phishing Detection: Identification of Malicious URL Using Bayesian Classification on Social Network Sites

Vaibhav V. Satane[1], Arindam Dasgupta[2]

[1]Department of Information Technology, AVCOE Sangmner, A. Nagar, Maharashtra, India

[2]Department of Information Technology, AVCOE Sangmner, A. Nagar, Maharashtra, India

Abstract: ***Nowadays people uses online social networking sites for communications with others. People post their message on social network sites. The messages are of various types such as text, images, audio and video. Sometime for these messages the people post their URL and request their friend to visit that site to show the messages. The fraud user uses malicious URL and post on social networking sites. These malicious URL contains viruses which harms user system. Malware is a one of the attack in which fraud URL creates replica of their own when user clicks on these URL and acquire resource. Sometime malicious URL directed towards a website which is a fraud website. These fraud websites are used to steal user's confidential information. By using Bayesian classification identify the fraud URL on social networking sites and improve the security of social networking sites.***

Keywords: Bayesian Classification, Suspicious URL, Social Networking Sites (SNS), Phishing Websites, Social spam guard, CANTINA.

1. Introduction

Phishing is a type of attack that used to induce online user to get their personal information such as credit card number, bank details which causes financial loss of user. In last few years the use of online social network becomes more popular such as You Tube, Facebook and Twitter. All over world people use SNS to communicate with other people they share their data with them. The shared data may be confidential. AS a use of SNS increases parallel the creation of computer malware, such as viruses and phishing attacks, has also continued to rise. One of the most famous examples of this new attack is the Koobface virus. The Koobface virus spreads through hyperlinks that appear to come from one of your friends, usually advertising a funny video. When the victim clicks the link to watch the video. They are met with a pop-up message stating that they need to update their Adobe Flash player. When the user clicks to download the "update", they are actually downloading a Trojan horse which installs both a web proxy and a backdoor on the victim's system. Several projects have been implemented successful in the area of malware and fraud detection, including examples such as CANTINA, Net craft, and Spoof Guard; none to date have utilized the tremendous wealth of information found in social networks.

According to RSA report in year 2014 nearly RSA identified 33, 145 phishing attacks in August, RSA estimates phishing cost global organizations $282 million in losses. The U.S. remained the most targeted country in August with 61% of phishing volume. China, Netherlands, United Kingdom and Canada were collectively targeted by 20% of total attacks. Hong Kong remained the second top hosting country for phishing in August with 13% of total attacks. The volume of attacks hosted in Italy doubled from July from 3% to 6% [1].

According to APWG global phishing report there were at least 123, 741 unique phishing attacks worldwide. Most of the growth in attacks came from increases in attacks against vulnerable hosting and use of maliciously registered domains and subdomains. An attackis defined as a phishing site that targets a specific brand. The number of domain names in the world grew from 271.5 million in November 2013 to 279.5 million in April 2014.It identifies 22, 679 domain names registered maliciously by phishers and most of these registrations were made by Chinese phishers, especially using free domain name registrations. The targets included more large and small banks in Latin America, India, and the Middle East. The phishing attack using social networking site and email is around 23.1%, 25.7%, 32.4% and 12.8% for bank, ecommerce industry and money transaction respectively over total phishing attack [2].

2. Related Work

Nowadays most people uses internet for various purposes such as online shopping like purchasing or selling products, chat with friends, sending mail. Internet users now spend more time on social networking sites Information can spread very fast and easily within the social media networks. Social media systems depend on users for content contribution and sharing. Facebook had over 1.3 billion active users as of June 2014. there are over 1.3 billion (the number is keep growing) pages from various categories, such as company, product/service, musician/band, local business, politician, government, actor/director, artist, athlete, author, book, health, beauty, movie, cars, clothing, community. Fans not only can see information submitted by the page, but also can post comments, photos and videos to the page.

Justin Ma et. al proposed Identifying Suspicious URLs: An Application of Large-Scale Online Learning in which

explores online learning approaches for detecting malicious web sites using lexical and host-based features of the associated URLs. Use various lexical and host-based features of the URL for classification, but exclude web page content. If properly automated, this technique can afford the low classification overhead of blacklisting while offering far greater accuracy. in this proposed system built a URL classification system that uses a live feed of labeled URLs from a large web mail provider, and that collects features for the URLs in real-time [3].

Xin Jin et. al proposed SocialSpamGuard as spam detection System for Social Media Networks security. Due to the huge amount of posts (over billions) on social media, manually checking every post to pick up the spams is impossible. scalable active learning approach proposed to manually verify as many spams as possible This system has several benefits automatically harvesting spam activities in social network, Introducing both image and text content features and social network features to indicate spam activities Integrating with our GAD clustering algorithm to handle large scale data and Introducing a scalable active learning approach to identify existing spams with limited human efforts, and perform online active learning to detect spams in real-time[4].

Weibo Chu et. al study the study the effectiveness of machine learning based phishing detection using only lexical and domain features, which are available even when the phishing Webpages are inaccessible. many phishing Webpages are short-lived, typically less than 20 hours, and URLs may change frequently (fast-flux).in this research work only lexical features and domain features are used in our phishing detection. These features are readily available without accessing the Webpage, and thus can be used even if the phishing URLs are no longer accessible. Accessing a suspicious Webpage may bring additional risks since today's phishing Webpages may contain malicious code such as Secure bank Phishing Trojan. In his research work indicate that phishing detector is highly effective even with the reduced types of discriminative features, with detection rates better than 98% with false positive rates at 0.64% or less [5].

Kurt Thomas et. al proposed a systematic approach for detecting large-scale attacks on Twitter that we leverage to identify victims of compromise, track how compromise spreads within the social network and evaluate how criminals ultimately realize a profit from hijacked credentials. Criminals succeed in hijacking accounts from users around the globe, irrespective of user understandings. Promising, casual, and core users with hundreds to thousands of followers all fall victim to attacks. At the duration of 1 day in assumed dataset correlate with 21% of victims never returning to twitter after the service wrests control of a victim's account from criminals. Furthermore, 57% of victims lose friends post-compromise in response to spam the victim's account send[6].

Yue Zhang et. al proposed CANTINA Approach for detecting phishing website. Based on the TF-IDF information retrieval algorithm. CANTINA examines the content of a web page to determine whether it is legitimate or not. The term frequency (TF) is simply the number of times a given term appears in a specific document. The inverse document frequency (IDF) is a measure of the general importance of the term. Roughly speaking, the IDF measures how common a term is across an entire collection of documents [7].

Mahmoud Khonji et al. conduct surveys the literature on the detection of phishing attacks. Phishing attacks target vulnerabilities that exist in systems due to the human factor. Many cyber-attacks are spread via mechanisms that exploit weaknesses found in end users, which makes users the weakest element in the security chain [8].

Chia-Mei Chen et.al proposed a suspicious URL identification system for use in social network environments based on Bayesian classification attackers may use SNSs as vehicles, using compromised user accounts to post messages that contain malicious URLs. Social network users typically trust the information that their friends submit in posts and feeds; thus, they become the victims of social engineering attacks. Malicious URLs used in social networks may make use of the trust and social relationships which resembles phishing websites in spam; thus, multiple sets of features are proposed for detecting spam or malicious URLs. In the first module, data collection, posts are collected including time and content. Posts that lack URL information are considered benign. In the second module, feature extraction, the proposed features are retrieved and a feature vector is constructed for classification. In the third module, the Bayesian classification model, posts are classified based on affected classification model [9].

Neda Abdelhamid et. al. proposed Multi-label Classifier based Associative Classification method for website phishing. Associate classification is effectively detecting phishing websites with high accuracy and MCAC generates new hidden knowledge (rules) that other algorithms are unable to find and this has improved its classifiers predictive performance [10].

Wei Xu et al. developed techniques to detect spammers in social networks, and aggregated their messages in large spam campaigns. To collect the data about spamming activity, create a large and diverse set of "honey-profiles" on three large social networking sites, and noted the kind of contacts and messages that they received then analyzed the collected data and identified anomalous behavior of users who contacted that profiles[11].

Maher Aburrous et al. proposed model based on fuzzy data mining techniques which is effective tool in assessing and identifying phishing websites for e-banking. The proposed model is based on fuzzy logic combined with data mining algorithms to characterize the e-banking phishing website factors and to investigate its techniques by classifying the phishing types and defining six e-banking phishing website attack criteria's with a layer structure.[12]

In the research work of Ram Basnet proposed the Detection of Phishing Attacks: A machine learning

Approach use Support Vector Machine algorithm also comparable study of Biased Support Vector Machine (BSVM) and Artificial Neural Networks both gave the same accuracy of 97.99%. The use of machine learning from a given training set is to learn labels of instances phishing or legitimate emails [13].

Justin MA et al. proposed a real-time system for gathering URL features and pair it with a real-time feed of labeled URLs from a large Web mail provider. From these features and labels, we are able to train an online classifier that detects malicious Web sites with 99% accuracy over a balanced dataset [14].

3. Data Mining

Data mining is the process for extracting hidden knowledge and pattern from large database. Knowledge Discovery from Data is synonym for data mining. Machine learning approaches are commonly used to classify malicious URLs and anomalies. Machine learning uses different classification technique such as logistic regression, SVM, and Bayesian classification. Logistic regression is used to predict the outcome of a binary dependent variable based on various predictor variables. SVM represents training data as points in space and locates one hyperplane in order to classify the data into categories. The point representing training data are support vectors and the solid line represents the separating hyperplane used for classifying the test data.

Bayesian classification is statistical classifier. It is based on Bayes theorem

$$P\left(\frac{H}{X}\right)=\frac{P(X/H)P(H)}{p(X)}$$

X is a data tuple. In Bayesian terms, ***X*** is considered evidence.
H be some hypothesis, such as that the data tuple ***X*** belongs to a specified class *C*.
For classification problems, we want to
Determine *P*(*H* / *X*), the probability that the hypothesis *H* holds given the "evidence" or observed data tuple *X*.
P (*H*), which is independent of *X*.
P(*X*/*H*) is the posterior probability of *X* conditioned on *H*.
P (*H*) is the prior probability, or *a* priori probability, of *H*.
P(*X*) is the prior probability of *X*.

URL Features

Malicious Web sites use lexical and host-based features.

Lexical features:

These features allow to capture the property that malicious URLs tend to "look different" from benign URLs. These are Hostname, primary domain, path token, last path token and so. To implement these features, use a bag-of-words representation of tokens in the URL, where '/', '?', '.', '=', '-', and ' ' are delimiters [3].

Host-Based Features

These features describe properties of the Web site host as identified by the hostname portion of the URL. They allow us to approximate "where" malicious sites are hosted, "who" own them, and "how" they are managed. Such as WHOIS information, location, connection speed, Membership in blacklists and so [3].

4. Conclusion

Fraud user use social networks are often used to collect personal information as well as collect financial data of user. Extraction of feature set for finding mistrustful URL using Bayesian classification in social networking site. Two types of anomaly features were proposed: domain anomaly and social anomaly features. Domain anomaly features are used to identify possible malicious domains based on lexical and reputation factors, whereas social anomaly features represent anomalous user behaviors in social communications.

References

[1] RSA Online Fraud Report September 2014.
[2] APWG_Global_Phishing_Report_1H_2014.
[3] Justin Ma et.al "Identifying Suspicious URLs: An Application of Large-Scale Online Learning"*26thInternational Conference on Machine Learning*, Montreal, Canada, 2009.
[4] Xin Jin et.al "Social Spam Guard: A Data Mining Based Spam Detection System for Social Media Networks", *37thInternational Conference on Very Large Data Bases*, August 29th 2011, Washington.
[5] Weibo Chu et. al "Protect Sensitive Sites from Phishing Attacks Using Features Extractable from Inaccessible Phishing URLs" , *Microsoft Research Asia*, Beijing, China. 2011
[6] Kurt Thomas et. al "The Koobface Botnet and the Rise of Social Malware", *5th International Conference on Malicious and Unwanted Software, IEEE* 2010.
[7] Yue Zhang et. al "CANTINA: a content-based approach to detecting phishing web sites", in *Proceedings of the International World Wide Web Conference*, 2007.
[8] Mahmoud Khonji et al. "Phishing Detection: A Literature Survey", *IEEE Communication Surveys & Tutorials*, vol. 15, No. 4. 2013.
[9] Chia-Mei Chen et al. "Feature set identification for detecting suspicious URLs using Bayesian classification in social networks", *Elsevier Information Sciences* 289 (2014) 133–147
[10] Neda Abdelhamid et al. "Phishing detection based Associative Classification data mining", *Elsevier Expert Systems with Applications* 41 (2014) 5948–5959.
[11] Wei Xu, Fangfang Zhang, Sencun Zhu, "Worm detection in online social networks", in*: Proceeding ACSAC '10 Proceedings of the 26th Annual Computer Security Applications Conference*, 2010.
[12] Maher Aburrous et al. "Intelligent phishing detection system for e-banking using fuzzy data mining"

Elsevier Expert Systems with Applications, 37 (2010) 7913–7921.

[13] Ram Basnet, Srinivas Mukkamala, and Andrew H. Sung, "Detection of Phishing Attacks: Machine Learning

[14] Approach", in*: Springer Studfuzz 226*, pp. 373–383, Verlag Berlin Heidelberg 2008.

[15] Justin MA et al. "Learning to Detect Malicious URLs**",** *ACM Transactions on Intelligent Systems and Technology*, Vol. 2, No. 3, Article 30, Publication date: April 2011