Detection of phishing attacks

Muhammet Baykara
Department of Software Engineering,
Faculty of Technology
Firat University, Elazig, Turkey
mbaykara@firat.edu.tr

Zahit Ziya Gürel
Department of Software Engineering,
Faculty of Technology
Firat University, Elazig, Turkey
ziyagurel55@hotmail.com

***Abstract*— Phishing is a form of cybercrime where an attacker imitates a real person / institution by promoting them as an official person or entity through e-mail or other communication mediums. In this type of cyber attack, the attacker sends malicious links or attachments through phishing e-mails that can perform various functions, including capturing the login credentials or account information of the victim. These e-mails harm victims because of money loss and identity theft. In this study, a software called "Anti Phishing Simulator" was developed, giving information about the detection problem of phishing and how to detect phishing emails. With this software, phishing and spam mails are detected by examining mail contents. Classification of spam words added to the database by Bayesian algorithm is provided.**

Keywords—information security; intrusion detection; phishing attacks; intrusion detection systems

I. INTRODUCTION

Phishing is defined as the fraudulent acquisition of confidential data by the intended recipients and the misuse of such data. The phishing attack is often done by email. An example of Phishing; as if e-mail appear to be from known web sites, from a user's bank, credit card company, e-mail, or Internet service provider. Generally, personal information such as credit card number or password is asked to update accounts. These emails contain a URL link that directs users to another website. This site is actually a fake or modified website. When users go to this site, they are asked to enter personal information to be forwarded to the phishing attacker [1, 2].

Phishing is often used to learn someone's password or credit card information. With the help of e-mail prepared as if coming from a bank or official institution, computer users are directed to fake sites. In general, the information that is stolen by a phishing attack is as follows:

- User account number
- User passwords and user name
- Credit card information
- Internet banking information

The Anti Phishing Simulator, which is designed to prevent serious threats like this, catches malicious e-mails arriving at e-mail addresses integrated into the system. This system also provides URL based control. The system evaluates the keywords included in the existing database and thus determines the contents of the mail [3, 4].

II. PHISHING ATTACKS

Phishing sends a fraudulent transmitter that appears to come from a real source. It is usually done by e-mail. The aim is to steal sensitive data such as credit card and login information or to install malicious software on the victim's machine. Phishing is a common type of cyber attack that everyone must learn to protect themselves. Phishing is start with a fake e-mail or other type of transmission designed to attract a victim. In this type of attack, the message appears to come from a trusted source. If the attacker is deceiving the victim, it is mostly encouraged to provide confidential information in a fraud web site. Sometimes malware is downloaded to the target computer. Attackers provide financial gain by having their victim's credit card information or other personal data. Sometimes, phishing e-mails are sent to retrieve login details or other details of employees to use for an advanced attack against a specific company. Cyber attacks, such as Advanced Permanent Threats (APT) and ransom software, usually start with phishing.

In a phishing attack, attackers can use social engineering and other public information resources, including social networks like LinkedIn, Facebook and Twitter, to gather background information about the victim's personal and work history, interests and activities. With this pre-discovery, attackers can identify potential victims' names, job titles and e-mail addresses, information about the names of key employees in their colleagues and organizations. This information can then be used to prepare a reliable e-mail. These attacks, including attacks by advanced persistent threat groups, usually start with an e-mail containing a malicious link or attachment. In this type of attack, the most popular vulnerability or clickable phishing environments have been identified as the most popular Facebook feeds. When phishing attacks are created, they are often used for unrealistic news, such as those created around major events, holidays and anniversaries. Usually a victim receives a message that appears to have been sent by a known person or organization. The attack is carried out via a malicious file injection that includes phishing software or through links to malicious websites. In either case, the goal is to direct the user to a malicious website to install malicious software on the device or to trick them into disclosing personal and financial information, such as victims, passwords, account IDs, or credit card details. A successful phishing message is usually shown from a well-known company; it is difficult to tell from the original messages: in

978-1-5386-3449-3/18/$31.00 ©2018 IEEE

phishing e-mails, company logos and other descriptive graphics and data collected from the company. As with other link manipulation techniques, the use of subdomains and misspelled URLs (often spelling mistakes) is common. Phishing attackers use JavaScript to place a legitimate URL of the URL onto the browser's address bar. The URL generated by navigating through an embedded link can also be modified using JavaScript. Defense against phishing attacks should begin with training and informing users to identify phishing messages; but there are other strategies that can reduce successful attacks. For example; a network gateway e-mail filter can capture many targeted phishing e-mails and reduce the number of phishing e-mails reaching users' inboxes [3]. Figure 1 illustrates the process of performing phishing attacks.

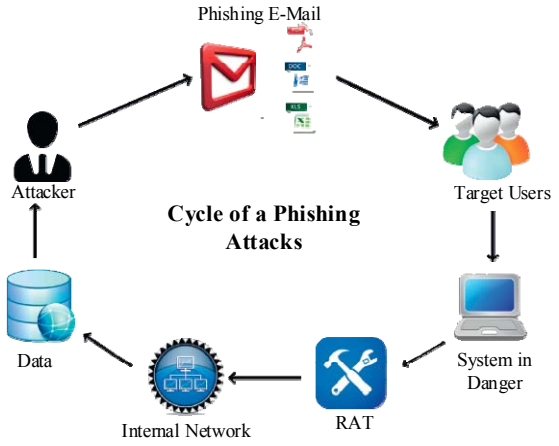

Fig 1. The processing cycle of phishing attacks

III. RELATED WORKS

Liu P et. al. have tried to find an effective solution for filtering spam e-mails in their work. The recommended approach in the study is to use the text of the e-mail as a keyword only to perform complex word processing. In their study conducted, 4327 emails in the CSDMC2010 SPAM training data set were evaluated. Experimental results at the end of the study showed that the proposed algorithm had an accuracy of 92.8% [1].

Agrawal N et. al. have studied content-based filtering techniques in their work. Regardless of the content of incoming emails, they suggest an original spam filtering approach based on the header information of the post. The aim of suggested approach is to optimize network and server performance [2].

Chandra J. V. et. al. have focused on the punctuality process in the context of an improved continuous threat attack that collects information in their work and targets a person or an organization. They also focused on spam emails and targeted malicious emails, mathematical use of Bayesian spam filtering, and buyer-side detection techniques such as spam or junk mail filtering [3].

A comparative study of various spam filtering techniques based on various features has been presented to derive the best results based on the requirements of Sharma A. K. et. al. to maximize the efficiency of spam control algorithms or various existing data mining algorithms in their work [4].

Vyas T. et. al. In their work dealt with different classification techniques that use WEKA to filter spam mail. As a result, they show that Naive Bayes technique provides good accuracy (close to the highest level) among other techniques and is the fastest algorithm. In addition, a comparative study of each technique in terms of accuracy and time has been provided [5].

Thomas J. et. al. have described property selection techniques used in the generic text classification for spam filtering in their work. In addition, classification and estimation; header was performed using different items, such as e-mail and the subject body. Different feature selection methods are presented comparatively. As a result of extensive experiments, the selection of Weighted Mutual Information feature by e-mail header and body has been shown to be effective in email classification [6].

The goal of AlRashid H. et al.'s work is to provide a solution to reduce the rate of false positive emails. The current e-mail addressed the problem by examining the behavior of spam filters and highlighting the different reasons behind the failure of the email. On the basis of this problem, they have developed an algorithm that helps users of email to send securely in safe transmission. The proposed algorithm is based on reversing the mechanism of spam filters on the client side [7].

Dhanaraj S. et al. have analyzed in detail the software and methods developed to prevent spam mails in their work. They have reviewed the systems on spam filtering and made suggestions [8].

The performance measures of certain controlled machine learning techniques categories such as Bayes algorithms, Bayesian algorithms, tree algorithms, artificial neural network, and support vector machines for classification of a spam e-mail collection held by the UCI Machine Learning Storage have been compared in the study of Panigrahi P. K. The purpose of the work done is to examine the contents of the emails, to learn the limited data set available, and to develop a classification model that can predict whether an email is spam [9].

In Toit T. et al.'s study, the performance of the Artificial Neural Network on a public e-mail server is investigated in the context of statistical spam filtering. It is also compared with a neural network, Naive Bayesian classifier and Memory based technique [10].

IV. DEVELOPED APPLICATION

Classification can be defined as an estimate of a particular outcome, based on specific qualifications, starting from the training data. To estimate the results, a particular classification algorithm works on a set of qualifications and a training set containing the relevant result, often referred to as the target or estimated quality. The algorithm tries to predict the results and investigate possible relationships between qualifications. Then, the algorithm is given an unseen data set, called the set

978-1-5386-3449-3/18/$31.00 ©2018 IEEE

of estimates, containing the same set of attributes, with the exception of an unknown set of estimates. The algorithm analyzes the input and generates an estimate. Prediction correctness indicates that the algorithm used is "good" [11].

After the preliminary stages of the data mining process, the choice of parameters and the choice of data set to be tested will affect the performance of the model that will be visible in the applications. For this reason, the result of the comparison will depend on the chosen classification algorithm. The data set (Table 1) used in this study is composed of the words that have the most used spam mail feature today. Each word has its own weight. These weights have been given a higher weight on the words that will cause the person to feel excited, fearful and hateful. In addition, these words are created with data mining, existing harmful words and harmful content created by the sites. With the Bayesian classifier-like algorithm, the weights of the words are calculated and a spam word count is made. At the same time certain rules are applied to prevent phishing attacks. Firstly, phishing attack links are detected on the Internet. In addition, spam mail and phishing attack possible to arrange links "add block" database was determined by selecting. URL control is provided on this page. In addition, site security checks and security password protocols (http-https) are checked.

TABLE 1. USED DATASET

id	kelime	deger	Kimlik	URL
1	kredi kartlarını kabul et	3	1	ofilmo.com
2	ücretsiz	4	2	dmp.gravity4
3	hepsi	1	3	www.whitebo
4	şimdi harekete geçin!	2	4	www.lesssec
5	kazanan sensin!	5	5	powertraf.com
6	ek gelir fırsatı	3	6	bahissirketler
7	bedava al	5	7	https://foruma
8	ücretsiz sermayeler	3	8	http://phbrin
9	tamamen doğal	3	9	thresholdofvi
10	hepsi yeni	2	10	memohaber.co
11	faizsiz kredi	4	11	buytoolbar.biz
12	şimdi uygula	2	12	full2hd2filmceh
13	çevrimiçi başvurun	2	13	baslattusu.co
14	dikkat	4	14	beehappyy.biz
15	deneme	1	15	filmcuks.com

Bayesian network classifiers are a special type of Bayesian networks designed for classification problems. Randomly obtained sets of variables and conditional dependencies are defined as a probabilistic model driven by graphs [12]. The Bayesian classifier, which is characterized as a graphical model that effectively encodes the common probability distribution for a large set of variables, provides an efficient representation of the multivariate probability distribution of a set of random variables and makes various calculations on that representation [13].

$$p(X) = \prod_{i=0}^{n-1} p(X_i \mid \prod_{X_i}) \quad \textbf{(1)}$$

BayesNet is a widely used class of models for representing probabilistic information. In BayesNet, border conditional dependencies represent non-bound nodes conditionally independent variables. In Equation 1, X =(X0… Xn-1) is treated as a vector of variables. ∏xi are the Parents of the Xi's in the network, ie the ancestral clusters. p(Xi/∏xi) Xi is a conditional probability. The distribution here can be used to create new examples from conditional and marginal probabilities. The benefits of Bayesian networks are as follows:

- They do not need to have prior knowledge of the problem.
- They have the ability to maintain a high level of interaction.
- They enable the architectural blocks to be efficiently assembled and fused in accordance with a certain order.
- They use data modeling to estimate the common distribution of solution that is positive in terms of outcome [14].

In this study, an application called "Anti Phishing Simulator" was developed to check the text content and determine whether the related message contained phishing elements. A simple flowchart of the implemented application is given in Figure 2. Today, an e-mail can be found in primitive ways whether it is a phishing message or not. For this are looked where this e-mail came from, whether a link with the message matches the actual website, whether the email or referrer web site is using some emotional or exciting words to get a response, whether it is spelling or grammar errors in the email or on the website. However, many people pay attention to this point unconsciously entering the links given to others accounts.

978-1-5386-3449-3/18/$31.00 ©2018 IEEE

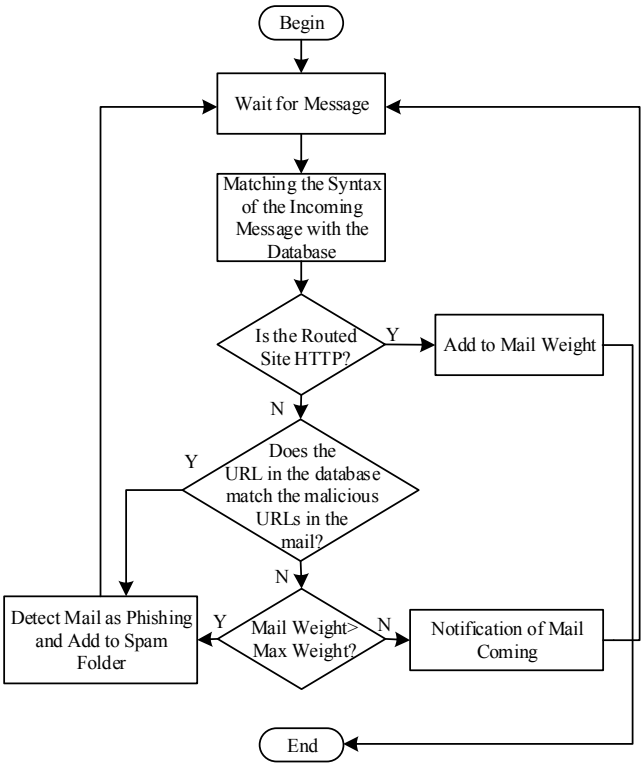

Fig 2. Flowchart of the application

In this work, "Anti Phishing Simulator" decides whether a message is phishing thanks to the Bayesian classification algorithm and the scores added to the database. It is instantly perceived as a spam message by the words that are exciting, phrases that increase the desire for shopping, and which contain unwanted content. In addition, these spam mails can be viewed from the spam section. At the same time, it is possible to add an unwanted site URL address or an unwanted word to the database with the "add spam" feature. The page's html code is displayed with the "URL control" feature for those who have mastered the computer programming language. In this way, it can be checked whether the links in the page are valid or not. Although the database is very large, there is "add spam" feature to manage it on demand. In this way, the user can filter out messages that are not really spam, but that he does not want to see. Figure 3 shows the logical operation of the life cycle of a phishing attack and the detection operation performed.

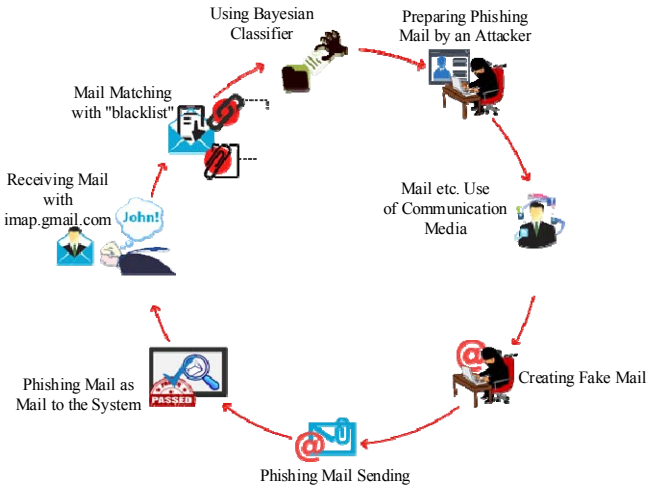

Fig 3. Logical operation cycle of the performed work

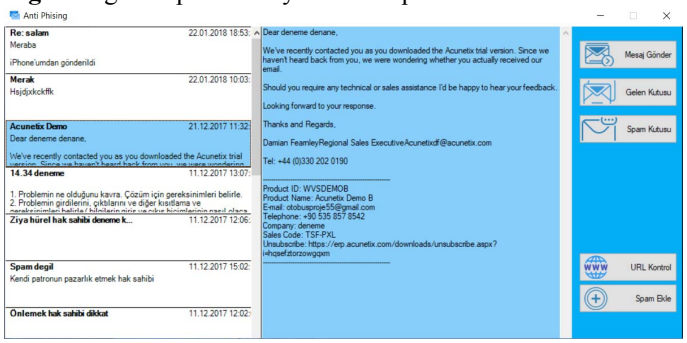

Fig 4. Mail system included in the application

Anti Phishing Simulator " aims to control the security of information and to prevent infringements, to check whether spam is available from the current database, to enable the user to create his own spam list, and to check whether the incoming mail has dangerous content. Figure 4 shows the inclusion of the mail account to be protected in the system. With this module, the user will also control and communicate without having to open the mail. Figure 5 shows a screenshot of the spam mails detected in the mail system. With this module, it is possible to determine the classification results of keywords and passages in the database by Bayes algorithm. Figure 6 shows the panel where the URLs in the mail are checked and the fake URL is detected.

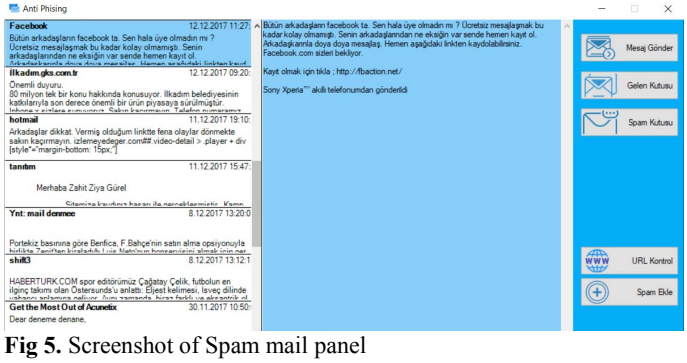

Fig 5. Screenshot of Spam mail panel

978-1-5386-3449-3/18/$31.00 ©2018 IEEE

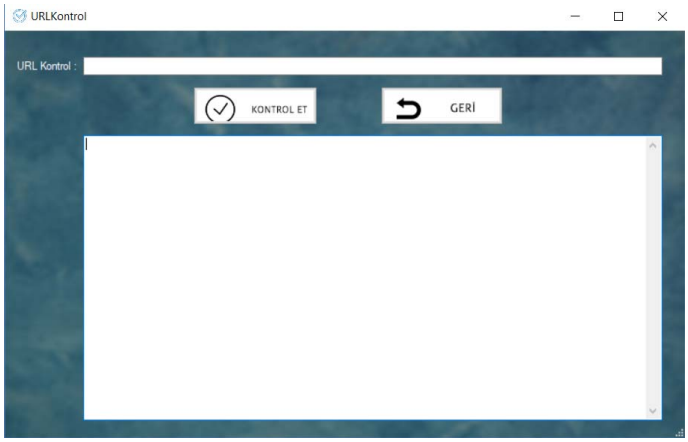

Fig 6. Screenshot of URL control module

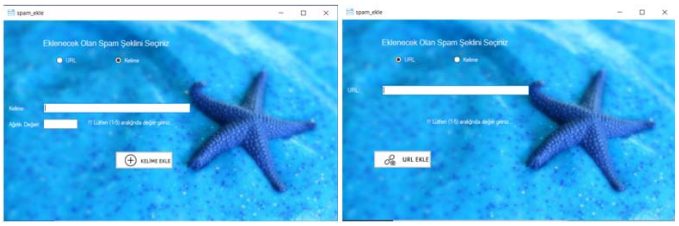

Fig 7. Adding spam words and URLs to the system

In addition to the words in the database, the study provides the opportunity to add spammy words and URLs that we want (Figure 7). These modules allow the user to mark spam as a spam that is not actually spam but does not want it to arrive.

V. RESULTS AND EVALUATION

E-mail is one of the most important communication methods. Increased spam e-mails cause traffic congestion, decreased productivity, phishing and this is a serious problem in terms of the world of information. The number of spam e-mails is increasing every year. For this reason, spam e-mail filtering is an important, meaningful and challenging issue. Due to the rapid spread of phishing attacks, different ways of protection have been developed. Real and fake web pages are sometimes very difficult to tell from the fact that fake pages are the same in terms of design.

The constant growth of e-mail users has resulted in unwanted e-mails becoming so widespread. Existing server and client-side anti-spam filters are being used to detect different features of spam e-mails. However, some effective tricks have been developed with the addition of spam senders' spam content as digital images, pdf and word; this extension has rendered it ineffective for current techniques based on analyzing digital text in the body areas of the e-mail. Most of the work strategy proposed in the study provides an anti-spam filtering approach based on data mining techniques that classifies spam and phishing e-mails. The effectiveness of these approaches is evaluated on the broad body of the simple text data set and the text embedded image data set.

"Anti Phishing Simulator" collects phishing and spam messages at a common point. In addition to getting spam messages in the spam box, it allows you to control the "spam box" whenever you want. Those who are technically qualified by the "URL Control" feature will be able to examine the link address in the mail in more detail. In the future, it is aimed to analyze mail content more thoroughly with basic text mining by increasing the spam keyword database much more. It is also aimed to obtain more accurate results and classification with artificial neural networks.

REFERENCES

[1] P. Liu and T. S. Moh, "Content Based Spam E-mail Filtering," 2016 International Conference on Collaboration Technologies and Systems (CTS), Orlando, FL, pp. 218-224, 2016.

[2] N. Agrawal and S. Singh, "Origin (dynamic blacklisting) based spammer detection and spam mail filtering approach," 2016 Third International Conference on Digital Information Processing, Data Mining, and Wireless Communications (DIPDMWC), Moscow, pp. 99-104, 2016.

[3] J. V. Chandra, N. Challa and S. K. Pasupuleti, "A practical approach to E-mail spam filters to protect data from advanced persistent threat," 2016 International Conference on Circuit, Power and Computing Technologies (ICCPCT), Nagercoil, pp. 1-5, 2016.

[4] A. K. Sharma and R. Yadav, "Spam Mails Filtering Using Different Classifiers with Feature Selection and Reduction Technique," 2015 Fifth International Conference on Communication Systems and Network Technologies, Gwalior, pp. 1089-1093, 2015.

[5] T. Vyas, P. Prajapati and S. Gadhwal, "A survey and evaluation of supervised machine learning techniques for spam e-mail filtering," 2015 IEEE International Conference on Electrical, Computer and Communication Technologies (ICECCT), Coimbatore, pp. 1-7, 2015.

[6] J. Thomas, N. S. Raj and P. Vinod, "Towards filtering spam mails using dimensionality reduction methods," 2014 5th International Conference - Confluence The Next Generation Information Technology Summit (Confluence), Noida, pp. 163-168, 2014.

[7] H. AlRashid, R. AlZahrani and E. ElQawasmeh, "Reverse of e-mail spam filtering algorithms to maintain e-mail deliverability," 2014 Fourth International Conference on Digital Information and Communication Technology and its Applications (DICTAP), Bangkok, pp. 297-300, 2014.

[8] S. Dhanaraj and V. Karthikeyani, "A study on e-mail image spam filtering techniques," 2013 International Conference on Pattern Recognition, Informatics and Mobile Engineering, Salem, pp. 49-55, 2013.

[9] P. K. Panigrahi, "A Comparative Study of Supervised Machine Learning Techniques for Spam E-mail Filtering," 2012 Fourth International Conference on Computational Intelligence and Communication Networks, Mathura, pp. 506-512, 2012.

[10] T. du Toit and H. Kruger, "Filtering spam e-mail with Generalized Additive Neural Networks," 2012 Information Security for South Africa, Johannesburg, Gauteng, pp. 1-8, 2012.

[11] Friedman, Nir, Dan Geiger, and Moises Goldszmidt. "Bayesian network classifiers." Machine learning 29.2-3, pp. 131-163, 1997

[12] Muralidharan, V., and V. Sugumaran. "A comparative study of Naïve Bayes classifier and Bayes net classifier for fault diagnosis of monoblock centrifugal pump using wavelet analysis." Applied Soft Computing 12.8, pp. 2023-2029, 2012.

[13] Bouckaert, Remco R. "Bayesian network classifiers in weka." Hamilton: Department of Computer Science, University of Waikato, 2007.

[14] E. Ardıl, "Esnek hesaplama yaklaşımı ile yazılım hata kestrimi." (2009).

[15] U Gürtürk, M Baykara, M Karabatak, "Identifying the Visitors with Data Mining Methods from Web Log Files", International Journal of Emerging Technologies in Engineering Research (IJETER), 5(3), 243-249, 2017.