2016 International Conference on Circuit, Power and Computing Technologies [ICCPCT]

A Practical Approach to E-mail Spam Filters to Protect Data from Advanced Persistent Threat

J. Vijaya Chandra,
IEEE Member, Research Scholar,
Department of CSE,
K.L. University, Guntur,
Andhra Pradesh., India.
vijayachandra.phd@gmail.com

Dr. Narasimham Challa,
Professor, Dept of CSE,
Vignan's Institute of Information Technology, Duvvada,
Vishakapatnam, A.P.,India.
narasimham_c@yahoo.com

Dr. Sai Kiran Pasupuleti,
Professor,
Department of CSE,
K.L. University, Guntur,
Andhra Pradesh., India.
psaikiran@kluniversity.in

Abstract— **Time based Self-destructing email mainly aims at protecting data privacy. In this paper we discussed the spear phishing process as a part of advanced persistent threat attack which gathers information and targets an individual or organization. It implements of social engineering techniques to gather data regarding recipient. Malicious emails are sent by combining the psychological and technical tricks, where phishing emails contains web-links that provoke the recipient to click on them, these links contains websites that are infected with malware. We also concentrated on Spam Emails and Targeted Malicious E-mails. In this paper we discussed recipient side detection techniques, such as spam or Junk mail filters using mathematical concept of Bayesian spam filtering. We contribute a clear indication of behavioral structure of Advanced Persistent Threat and a self-destructive mechanism is adopted as Defense System to protect sensitive confidential data from intruders. A mathematical approach is given along with the computational practical analysis and experimental result.**

Index Terms **- Advanced Persistent Threat, Spear Phishing, Self destructive E-mails, Spam E-mails, Targeted Malicious E-Mails.**

I. INTRODUCTION

Spear Phishing emails are targeted exploratory attacks based on social engineering that targets the victim's sensitive confidential data. It is a criminal attempt based on human psychology; technical procedural actions are played on victims, so that they open attached files, clicks embedded links and reveals sensitive information, which are commonly classified as the advanced targeted attacks by mining social networks [1]. Phishing websites were linked with the spear phishing emails and are exact duplicates of the original websites, undefined by the victims. Spear Phishing uses the amalgamate of zero-day application exploits, dynamic uniform resource locators, email spoofing, back door exploits and drive by downloads to bypass the conventional defenses. Advanced Spear Phishing attacks influence zero-day vulnerabilities and plug-ins in web browsers. Attackers use multiple attack vectors such as Internet, Email, Physical (USB) and Deception. They also use Trust Exploitation using different Techniques such as Social Engineering, Botnet, Spear Phishing, and Drive-to-click Strategy. The attackers pursue their objective that is to steal information over an extended period of time using sophisticated technologies and methods. This is known as Advanced Persistent Threat [2].

Advanced Persistent Threat Attacks are mostly targeted against an individual or particular organization, group or industry. It is a stealthy continuous process and potential adversary that possesses sophisticated levels of expertise, extensive research, trusted exploitation and significant resources which allow it to create opportunities to achieve its objectives at multi stages by using multiple attack vectors. When the system has been infiltrated, the attackers elevate privileges and create backdoors for feature intrusions [3].

The world famous notable incidents or cases which are considered as the Advanced Persistent Attacks are

A. Operation Aurora

Highly Sophisticated Targeted Attack on Google from china, where Intellectual property of Google was found theft [4].

B. Stuxnet

It is an American Israel Joint Program to Sabotage Iran's nuclear program [5].

C. *Operation Shady* RAT

Operation Shady Remote Access tool is an ongoing Series of Attacks mostly on organizations related to USA, that are continued still now, since 2009 that was originated from china and found by McAfee [6].

D. GhostNET

More than 100 countries are targeted by GhostNet operation which is associated with Advanced Persistent Threat as attackers used phishing methods and remote control tools. These attacks originate from china but china government denies it [7].

E. Darkhotel

The Business Executives in the modern world move from one country to another and stays at hotels. The attackers targets the hotel internet with spying software, once the executives connect to the hotel internet they trip them to install software updates, which installs the bundles of Trojans and key loggers [8].

978-1-5090-1277-0/16/$31.00 ©2016 IEEE

II. Related Work

A. Spam Mails or Junk Mails

Spam is an un-solicited Junk mail by the companies to sell their product or services or exploit the organization intent to do fraud, spammer takes a database of list and sent emails built in thousands as bulk mails, few will in trap of spammer. Spam Mail wastes a great amount of bandwidth and space of both the ISP of senders and receivers. The most common precautions are Ignore Emails from unknown senders, be careful with delivery failure emails, don't give your primary email address to marketing agencies or in websites for any type of promotions, use email filtering systems available with email packages.

B. Targeted Malicious E-mails

They are particular concern designed to capture sensitive confidential data from victims. These attacks are not only misleading inexperienced end users but also technically aware end users, mainly targeted email service providers employers. The mails does not ask directly the email addresses in the servers but asks or provoke or insists to click on provided links in the emails that contains malware which infect not only single employee system but maps to spread all the systems of the company to elevate privileges and company resources. They even use social engineering and psychological technologies such as humans will not read the every letter in a word, they read sequentially and randomly, for example a victim may not find difference in between support@microsoft.com and support@mircosoft.com where the only two letters r and c are swapped. Thus the attackers may benefit without the knowledge of the victims [9].

C. E-Mails Delivery Architecture

Sender sends bulk or a group of mails regarding business promotion or malicious mails intentionally sent to attack through Email Service Provider Servers. Email Firewalls filters the malicious mails based on previous feedback which is stored at spam databases and block-list based on the feedbacks at DNS web Servers. Emails are again filtered at the Email Servers and if found as spam they will send to the spam folder.

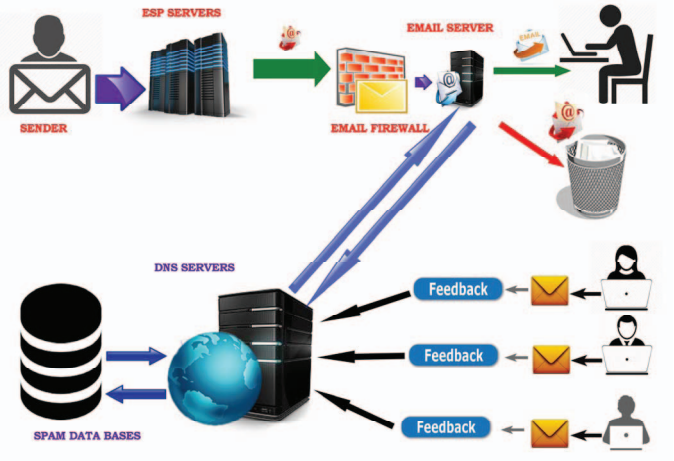

Figure 1: Block Diagram of E-mail Delivery Architecture

D. Extended Simple Text Transmission Protocol

Extended SMTP (ESMTP) has been brought into existence as the number of email users grew and additional functionality was sought in mail clients and SMTP servers. ESMPT is a connection-oriented, authenticated, text/graphic based protocol communicates between sender and receiver. It supports Audio /Video files and text in different languages. It works as inter server transport protocol and mail submission protocol.

E. Email Security Issues

The need for E-Mail security is well connected with the everyday emerging risks, threats and attacks done by cybercriminals using social engineering techniques to theft Identity of victims. Email security Lacks due to the following vulnerabilities [10]

- E-Mail communication can be exploited by intruders to gather the confidential and sensitive information from an organization.
- E-Mail information can be eavesdropped, intercepted and modified.
- E-Mail Infrastructure loop hole such as E-Mail server may be compromised and the entire organization may be taken into control including the users.
- Advanced Persistent Threat attacks on the E-Mail server makes the attacker to retrieve user names and user passwords.
- Distributed Denial of Service (DDoS) attacks may be performed on the mail server makes denying service to the legitimate and valid users.
- Malicious entities may use the prevailing security flaw to send inappropriate, sensitive and advertising information through mail by using the legitimate identity via mail server.
- SQL Injections and other hacking techniques crack the passwords.
- Spam and Phishing Mails breach the Email security.

Finally, Security plays a major role in communication. E-Mail attacks acknowledged as intentionally misguiding, fraudulent, or known to result in fraudulent activity on the part of the sender for example Pyramid schemes and chain letters. E-Mail attacks that typically alert an end user psychologically that they are entitled lottery of huge amount in dollars, As Service Charges asking for little amount, un employee regarding a new job. These attacks are mostly known as Nigerian frauds and are also referred to as advance fee fraud. E-Mail attacks sends mail that appear to be from a Multi-National Company, but are not. It is also known as "brand spoofing" or "phishing," as these messages are often used to trick users into reveals the personal information such as email address, financial and private information such as usernames and passwords. Examples: account notification, credit card verification and billing update [11].

III. MATHEMATICAL AND COMPUTATIONAL ANALYSIS

A. Conditional Probability

It is used to calculate the probability that an event A occurs when it is known that an event B has occurred, where B has positive probability. The symbol for this probability is P [A|B] and reads "the conditional probability of A, given B".

In general, to calculate the probability that A occurs, given that B has occurred, means reevaluating the probability of A in the light of the information that B has occurred. Thus, B becomes our new sample space and we interested only in the part of A that occurs with B, that is A ∩ B. Thus we must have the formula

$$P[A \mid B] = \frac{P[A \cap B]}{P[B]},$$

If P[B]>0. The Conditional Probability of A given B is not defined if P[B]=0. In P[A ∩ B] was divided by P[B] so that P[B][B] = 1, making P[.|B] a probability measure. The event B in above equation is often called as the ***conditioning event***.

B. Multiplication Rule

Implementation of Multiplication Rule for events A and B
P[A ∩ B] = P[A]P[B|A],
If P[A]≠0, and
P[A ∩ B]=P[B]P[A|B]
If P[B]≠0 (*if either* P[A]=0 or P[B]=0) *then* P[A ∩ B]=0
The General Multiplication Rule for events $A_1, A_2 \ldots, A_n$,
$P[A_1 \cap A_2 \cap \ldots \cap A_n] = P[A_1]P[A_2|A_1]P[A_3|A_1 \cap A_2] \ldots$
$X\ P[A_n|A_1 \cap \ldots \cap A_{n-1}]$
Provided all the probabilities on the right are defined. A sufficient condition for this is that
$P[A_1 \cap A_2 \cap \ldots \cap A_n]>0$, since $P[A1] \geq P[A1 \cap A_2] \geq \ldots \geq P[A1 \cap A_2 \cap \ldots \cap A_{n-1}]$.

C. Bayes Theorem

Suppose the events A1,A2,….,An form a partition of Ω then for any event A with P[A]>0,

$$P[A_i][A] = \frac{P[A_i]P[A \mid A_i]}{P[A_1]P[A \mid A_1] + P[A_2]P[A \mid A_2] + \ldots + P[A_n]P[A \mid A_n]}$$

Where i = 1, 2, 3,……., n.
For each i,

$$P[A_i][A] = \frac{P[A_i \cap A]}{P[A]} = \frac{P[A_i]P[A \mid A_i]}{P[A]}$$

To Calculate the P[A], Apply the law of total probability
The $P[A_i]$, i = 1,2,3,…….n, are called prior or priori probabilities and $P[A_i|A]$, i=1,2,3,…..n, are called posterior or posteriori probabilities. To calculate the posterior probabilities using Bayes theorem, we must know both the prior probabilities $P[A_1]$, $P[A_2]$, $P[A_3]$,…., $P[A_n]$ and the conditional probabilities $P[A|A_1]$,……,$P[A \mid A_n]$.

D. Naïves Bayer's theorem for Spam Filtering

This is self-adopting continuous approach of learning from the new spams and considers the whole message into account based on the strings, the Identification of mail that is legitimate or spam takes place and classified separately based on tokens. Where tokens are considered as group of words, a group of characters can be stored in character array vector called as string. Based on the strings the learning process analyzes a mail to calculate its probability of being spam [12].

From Bayes theorem and the conditional probability and multiplicative rule, we classify the legitimate mails and spam mails using probability. The probability of a mail **m** with vector x = {x_1, x_2, …. , x_n} belong to a category C is

$$P(C = c \mid \vec{X} = \vec{x}) = \frac{P(C = c).P(\vec{X} = \vec{x} \mid C = c)}{\sum_{k \in \{spam, legit\}} P(C = k).P(\vec{X} = \vec{x} \mid C = k)}$$

Computing the probability of the email message containing block words is identified as spam. Suppose the suspected message contains the words like "click here", "free", "Viagra", "replica" etc., most people received mails with such words are spam as per the analysis [13].

$P(C = c \mid \vec{X} = \vec{x})$ is to identify the probability that a message is spam, knowing the block words.

$P(\vec{X} = \vec{x} \mid C = c)$ is the overall probability that any given mail is spam

$P(\vec{X} = \vec{x} \mid C = k)$ is the overall probability that any given mail is legitimate

The equation further simplified, where legit is the legitimate email and spam is the spam email and defined as

$$P(C = c \mid \vec{X} = \vec{x}) = \frac{P(C = c).\prod_{i=1}^{n} P(X_i = x_i \mid C = c)}{\sum_{k \in \{spam, legit\}} P(C = k).\prod_{i=1}^{n} P(X_i = x_i \mid C = k)}$$

Let legit → spam and spam → legit denote two error types. Invoking a decision-theoretic notion of cost, we assume that legit → spam is λ times more costly than spam →legit.
A message is classified as spam if the following criterion met:

$$\frac{P(C = spam \mid \vec{X} = \vec{x})}{P(C = legitimate \mid \vec{X} = \vec{x})} > \lambda$$

To the extent that the independence assumption holds and the probability estimates are accurate, a classifier based on this criterion achieves optimal results [14].

$$P(C = spam \mid \vec{X} = \vec{x}) = 1 - P(C = legitimate \mid \vec{X} = \vec{x})$$

Differentiates the spam and legitimate mails based on set t

$$P(C = spam \mid \vec{X} = \vec{x}) > t,$$

$$t = \frac{\lambda}{1+\lambda}, \lambda = \frac{t}{1-t}$$

IV. Statistical Analysis

Here λ establishes the concept of classification by accuracy, error rate and also allocating the penalty for improper classification leads to identify legitimate email as spam. If spam is shown as legitimate the user does not care much about it, it is negligible. Where t threshold is calculated based on the statistical analysis using the formula [15].

$$t = \frac{\lambda}{1+\lambda}$$

Here we observed three cases in case 1 where the value of λ is very high that is 999 which given a threshold value 0.999 means that the spam filter blocked the emails and they are discarded without further processing. In case 2 where the value of λ is medium that is 9 which given a threshold value 0.9 means that the blocking a legitimate mail will be considered seriously and penalized, where as passing a spam is not considered seriously. In case 3 the value of λ is 1 that is the threshold value is 0.5 mail user doesn't not respond much for being loss of a legitimate email.

The classification errors and accuracy of electronic mails is as legit → spam and spam → legit are considered as errors, the measure of errors occurred is known as error rate and the mails classified as legit →legit and spam → spam measurement is known as the accuracy rate. If $\mathbf{M_T}$ is the Total Mails then $\mathbf{M_L}$ is known as the number of Legitimate Mails and $\mathbf{M_S}$ is known as the number of Spam Mails. Acc_{wt} is the weight of Accuracy and Err_{wt} is the Error weight.

$$Acc_{wt}\% = \frac{(legit \rightarrow legit) + (spam \rightarrow spam)}{M_T} X100$$

$$Err_{wt}\% = \frac{(legit \rightarrow spam) + (spam \rightarrow legit)}{M_T} X100$$

TABLE 1: The Table Sample Data Set to Classify Spam and Legitimate Mails

Test Cases	Total Mails	Spam Mails	Legitimate Mails	legit →legit	spam→ spam	legit → spam	spam → legit	ACC_{wt} %	Err_{wt} %
1	5000	2500	2500	2000	2000	500	500	80	20
2	5000	1250	3750	3000	1200	750	50	84	16
3	5000	2500	2500	2200	2300	300	200	90	10
4	5000	3700	1300	1200	3500	100	200	94	06
Average								87	13

As the Naïves Bayer's classification theorem is a keyword based spam classifier, based on the feedback new keywords should be added to the spam database. After conducting a series of experiments the results are tabulated as above, where each test case consist of 5000 mails, the average of the results of accuracy weight 87% and error weight is 13% calculated. The Graphical Analysis is given based on the accuracy and error for classifying the spam and legitimate emails.

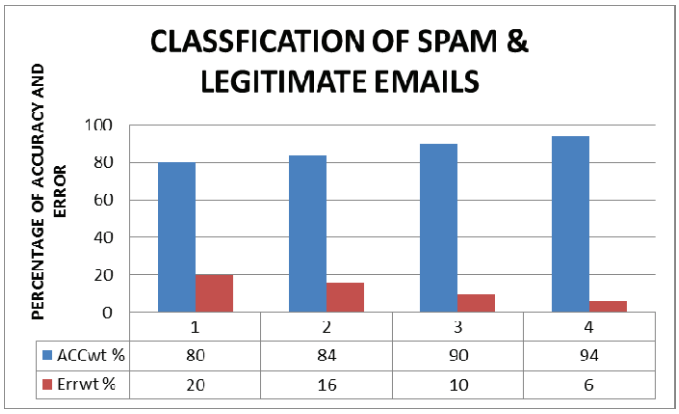

Figure 2: Graph of Accuracy & Error in E-mail Classification.

Spam mails are the phishing mails that are sent in number of thousands of mails of advertisements or unwanted materials or intentionally with fraud intensity. The major types of spams are adult, financial and games. Naïves Bayer's spam classifier is a content based and identification of new spam, will be based on feedback. Spear phishing and email spams are continuously increasing from last few years, the spams are broadly classified into two types they are identity based filtering or source based filtering ant based the second one is the content based filtering. We can support and identify legitimate mail by source based filtering such as the mails domain names gov.in, edu.in, ac.in.

V. A Practical Approach to E-Mails Filters

In Simulation Lab for the Experiment, we created a lab environment; to study Email Spam Filters based on Naïves baye's theorem for classification. Electronic mails are popular at business world, the e-commerce made emails as the best communication medium with the consumer, so many bulk messages are sent to the victims. Deleting the spams form the spam box is one of the possible ways but it is tedious and time consuming job. There we are in need of the spam filter where there will not be any human interaction [16].

For the Simulation process working with the spam filters a database consists of 60 words as samples with predefined probability is taken. These words are selected based on the block list strings that are identified by the international standards organization and also based on the observation that they are most likely present in spam messages. To calculate the legitimate and spam mails the following notations are used [17].

$$t = \frac{\lambda}{1+\lambda}, \lambda = \frac{t}{1-t}$$

The value of λ taken as 1.2 with a database of only 60 words a higher value of λ gives improper results. Emails from domain id '.edu' and '.ac.in' have been assumed as legitimate messages. In spite of words some phases are also considered as notations for spam such as **Billion Dollars, offers extra cash, for free,** etc.,

VI. Conclusion

Finally the Naïves baye's theorem for classification is implemented on the single string, multiple strings and string based on weight that means Save $ ≠ save dollars. To fool the spam filters new techniques are adopted by the spammers, as the classification is mainly based on the keywords such as porn, Viagra etc., spammers started sending the emails with replacing some alphabets in the keywords such as o is replaced by 0, i is replaced by !. Since as per the phycology, human read the letters in the words randomly even though they are having some spelling mistakes. Hence p0rn and v!agra can be easily read by victims. So based on the feedback and continuous monitoring spam keywords block list should be updated most frequently then only spam mails can be identified and defend.

The most common scam mails is the fraud job offer emails, most of them are using the logos of multinational companies and higher official names and signatures. The only way to identify the fraud mails and legitimate mails is that the email ids of multinational companies' newer use Gmail, Hotmail or Yahoo, they will have their official mail account.

The performance testing on the designed email spam filter is to calculate the accuracy, reliability and other factors. Continuous filtering System and Defense System is used protect sensitive confidential data from Advanced Persistent Threats. We leave the fully fledged implementation of the mechanism on commercial spam filter is for a future extension.

References

[1] Jingguo Wang; Herath, T.; Rui Chen; Vishwanath, A.; Rao, H.R., "Research Article Phishing Susceptibility: An Investigation into the Processing of a Targeted Spear Phishing Email", *IEEE Transactions on Professional Communication,* vol.55, no.4, pp.345-362, Dec. 2012.

[2] Vukalovic, J.; Delija, D., "Advanced Persistent Threats - detection and defense", *38th International Convention on Information and Communication Technology, Electronics and Microelectronics (MIPRO),* vol., no., pp.1324-1330, 25-29 May 2015.

[3] J. Vijaya Chandra, Dr. Narasimham Challa and Dr. Mohammed Ali Hussain, "Data and Information Storage Security from Advanced Persistent Attack in Cloud Computing", *International Journal of Applied Engineering Research*, Pp. 7755-7768, Volume 9, Number 20(2014).

[4] https://googleblog.blogspot.in/2010/01/new-approach-to-china.html.

[5] Langner, R., "Stuxnet: Dissecting a Cyberwarfare Weapon", *IEEE Security & Privacy*, vol.9, no.3, pp.49-51, May-June 2011.

[6] Dmitri Alperovitch, "Revealed: Operation Shady RAT", Threat Research, McAfee, 2011.

[7] Shishir Nagaraja, Ross Anderson, "The snooping dragon: social-malware surveillance of the Tibetan movement", Technical Report, Number 746,University of Cambridge, ISSN 1476-2986, p. 2. Retrieved March 31, 2009.

[8] "The Darkhotel APT: A Story of Unusual Hospitality". Kaspersky Labs. November 10, 2014.

[9] Khonji, M.; Iraqi, Y.; Jones, A., "Mitigation of spear phishing attacks: A Content-based Authorship Identification framework", International Conference for Internet Technology and Secured Transactions (ICITST), vol., no., pp.416-421, 11-14 Dec. 2011.

[10] J. Vijaya Chandra, Narasimham Challa, Sai Kiran P, Thirupathi RK, Krishna RV, " Numerical Formulation and Simulation of Social Networks using Graph Theory on Social Cloud Platform", Global Journal of Pure and Applied Mathematics. 2015; 11(2):1253–64.

[11] J. Vijaya Chandra, Narasimham Challa, Sai Kiran Pasupuleti , "Intelligence based Defense System to Protect from Advanced Persistent Threat by means of Social Engineering on Social Cloud Platform", Indian Journal of Science and Technology, Volume 8, Issue 28, October 2015.

[12] Jatana, N.; Sharma, K., "Bayesian spam classification: Time efficient radix encoded fragmented database approach," International Conference on Computing for Sustainable Global Development (INDIACom), vol., no., pp.939-942, 5-7 March 2014.

[13] Jian Zhong, Yilu Zhou and Wei Deng, "Filtering image-based spam using multifractal analysis and active learning feedback-driven semi-supervised support vector machine," IEEE Conference Anthology, pp. 1-5, China, 2013.

[14] Ze Li , Haiying Shen ,"SOAP: A Social network Aided Personalized and effective spam filter to clean your e-mail box", IEEE Proceedings in INFOCOM, vol., no., pp.1835-1843, 10-15 April 2011.

[15] "A decentralized and personalized spam filter based on social computing", International in Wireless Communications and Mobile Computing Conference (IWCMC), vol., no., pp.887-894, 4-8 Aug. 2014.

[16] Reddy, M.R., Yalla, P., J.Vijaya Chandra, "Design and implementation of integrated testing tool based on metrics and quality assurance",pp. 10463-10472,International Journal of Applied Engineering Research, Volume 9,Number 21(2014).

[17] Wanqing You, Kai Qian, Dan Lo, P. Bhattacharya, Minzhe Guo and Ying Qian, "Web Service-Enabled Spam Filtering with Naïve Bayes Classification", IEEE First International Conference onBig Data Computing Service and Applications (BigDataService), Redwood City, CA, 2015, pp. 99-104, March-April 2015.

About Authors

J.VijayaChandra is a Research Scholar at K L University; Interested Research areas are Cloud Security, Network Security, Intelligence Security and Data Security. Published 10 Research Papers for International Journals. He is Oracle Certified Associate and Member of IEEE.

Dr. Narasimham Challa, Ph.D., is Professor, Depart. of Computer Science and Engineering, Vignan's Institute of Technology and Science, Vishakapatnam, A.P., India. He has 20 Years of Teaching Experience, His research areas are Cryptography, Cloud Computing and Intelligent Security System. He Published About 50 Research Papers National and International Journals. Under his guidance 8 research Scholars are doing their Ph.D.in different Universities like JNTUK, K.L. University.

Dr. Sai Kiran Pasupuleti, Ph.D., is Professor, Dept. of Computer Science and Engineering, K L University, Green Fields, Vaddeswaram, Guntur District, A.P., INDIA. He is having rich teaching and Research Experience. His research areas are Mobile Computing, Cloud Computing and Computer Networks. He published about 25 Research Papers in International Journals.