ISSN(Online): 2320-9801
ISSN (Print) : 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(A High Impact Factor, Monthly, Peer Reviewed Journal)

Website: www.ijircce.com

Vol. 6, Issue 8, August 2018

How Bigdata Analytics and Machine Learning can help in Social Engineering (Phishing)

Govind Singh Mahara[1]

Ph.D. Researcher, Department of Computer Engineering, RKDF University, Gandhi Nagar, Bhopal, India[1]

ABSTRACT: The social engineering attacks are the critical area for cybersecuritynowadays, this is most common and easy ways to enter personal account and get things, social engineering is a way of using technologies by a human on the human. WIFI is once is the most Insecure system for this purpose because your devices get connected and another activity of syncing happens automatically. In this tries to explain how the social engineering attacks are serious to organizations.

Social engineering is done by adds and emails, familiarity exploit. Gathering & using information, pretexting. baiting. Tailgating.

KEYWORDS: Social Engineering, Phishing, Machine Learning, Cyber threats, Bigdata, Fraud, Hacking

I. Introduction

This paper guides about social engineering attack in the cyber world. The take higher growth in business more and more business are adopting technology so that they can make more business opportunities. This adaptation and growth in technology have given great value to Internet uses, but another hand it is also used in a tremendous way to get criminal benefits, Email, SMS are a good way to get into phishing which is the biggest threat to any organization or business. Phishing is one type of social engineering in which criminal mind also symbolize as a phisher, attempts to fraudulently

get user credentials by mimicking communication from a trusted source or public organization in an automated action.

Phishing emails also have some link where victims must click, and it will guide to other website/web page where personal information/credentials are requested. According to Ant phishing working people, 25000 phishing campaigns are launched per month. The purpose of phishing emails is to get personal details as much as possible (Personal identity, credit card number authentication other information like username, password). Phisher target is always high-class profiles so that they can steal proprietary information that includes social security number. Phishing is a jeopardize all Internet Business community. It damages more then monetary trust that builds with their elements are deteriorating. Which cause economical loss with resource and time. However, report of phishing events is rarely reported by an organization or very small part of the actual event get reported because the release of negative things can cause more damage to organization image in the market that will hit investors. The statistical data projects high risk over Small size business or middle -size business according to (DBIR Verizon,2108) 4% of people click on any given phishing email, professional 21 %, the education sector has 41% and public sector 32 %, small business companies do not have an effective gateway solution, security policy and knowledge and resources for controlling these attacks. Due to these small businesses are easy targeted in email attacks.

People understanding phishing process and their characteristics can effectively encounter and also improve employee ability to make a decision over these emails. This purpose the study will open few points to broaden knowledge area which can help to make better decisions, protection making in person or organizational.

II. RELATED WORK

According to research survey 76% of target attacks being with a spear-phishing email containing a malicious attack or link using techniques which aredifficult to detect standard email of endpoint security (Trend Micro). Social engineering attacks can simply understand as it's all about finding gaps and use those against human, hack the human by the humanis more suitable in this case. Mostly this is a category of hacking personal information and use that again them way of using technology by human Wi-Fi is most hackable system as it connects automatically by that gabs can be identified easily secondly phishing/email is the most useableattach, spear phishing an email that appears to be from an individual or business that is known to you. Social Engineering is costly especially for larger organizations 48% of large companies and 32% of companies of all sizes have experienced 25 or more social engineering,48% of all participants cite an average per-incident cost of over $25,000,30% of large companies cite a per-incident cost of over $100,000 97% of security professionals and 86% of all IT professionals are aware or highly aware of this potential

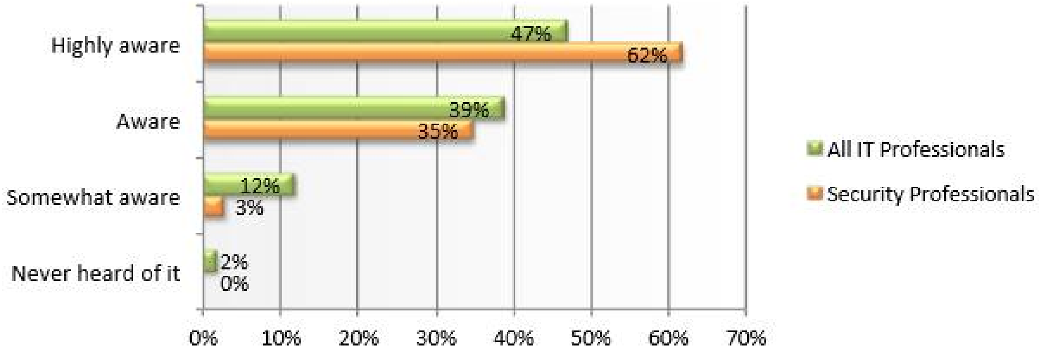


Figure 1: Awareness of Social engineering threats

Social engineer attack has different forms which impact personal information of human.
Phishing: Scams might be the most common types of social engineering attacks used today.
FamiliarityExploit: This is one of the best and it's a cornerstone of social engineering,In a nutshell, you are trying to make it appear perfectly normal to everyone that you should be there making yourself familiar to those that you want to exploit helps to lower their guard.
Gathering&Usinginformation: When it comes right down to it the key to being a successful social engineering in information gathering the more information you have about your mark the more likely you are to get what you want from him or her obviously.
Pretexting: is another form of social engineering where attackers focus on creating a good pretext or a fabricated scenario. That they can use to try and steal their victim's personal information.
Baiting: is in many ways like phishing attacks. However, what distinguishes them from other types of social engineering is the promise of an item or good hackers use to entice victims.
Tailgating: another social engineering all bock type is known as tailgating or "piggybacking" these types of attacks involve someone who lacks the proper authentication following an employee into a restricted area.

Phishing is a leading social engineering attack, it has no limits this fraud method has been growing rapidly approximately 8 million daily phishing attempts worldwide, Chinese phishers are more responsible for the full attack of phishing. There is a shared virtual server hacking attack, in this attack phisher hacks into webserver that hosts a larger number of domain websites /web application, Phisher places fake website pages to divert all information this way phisher can have thousands of websites live in few mints.

The APWG identified 263,538 attacks that used this strategy in 1Q 2018 this is 46% more by the Q4 2017.

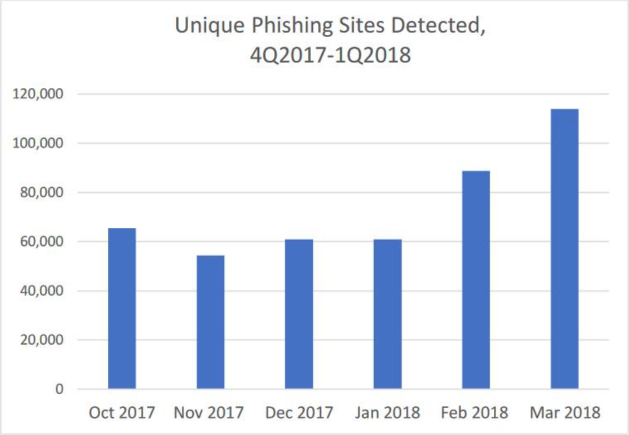
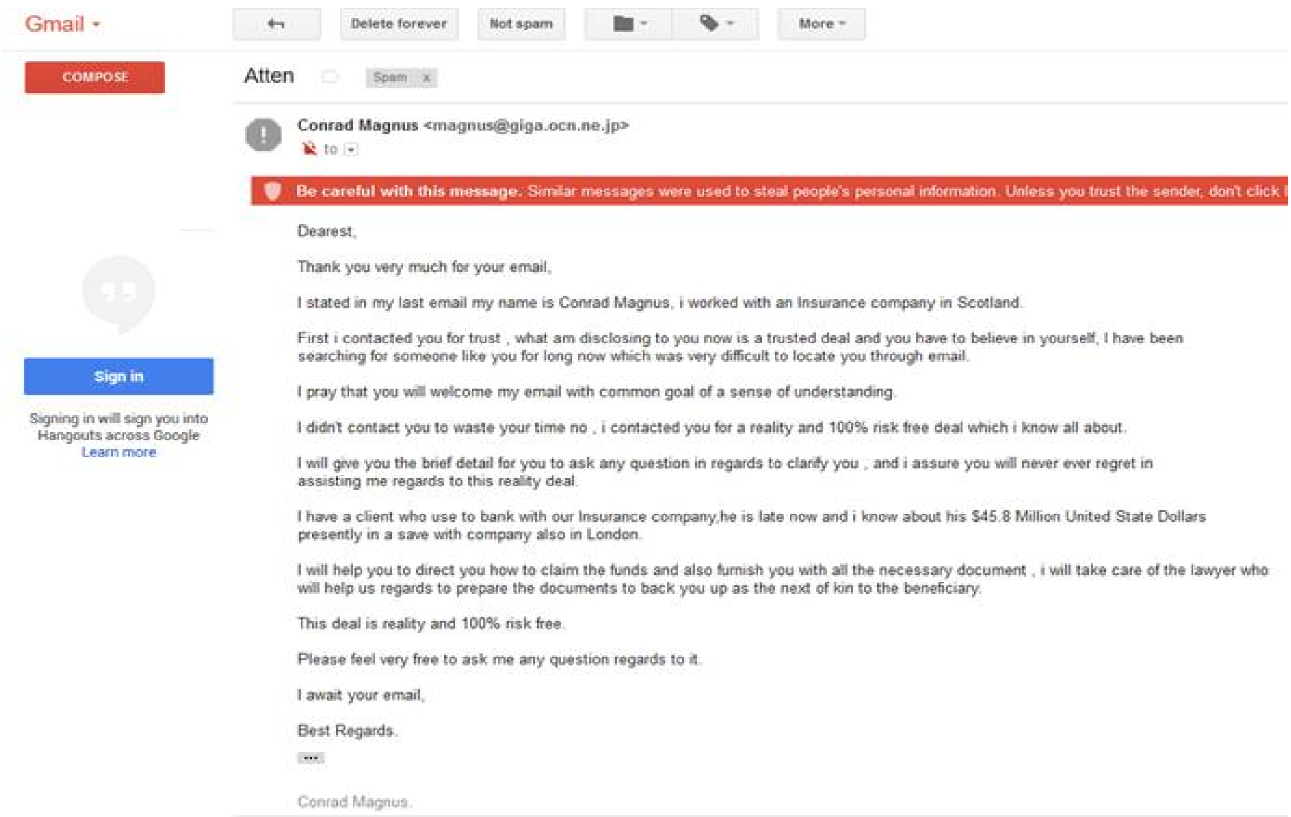


Figure2: Unique Phishing Sites Detected published by APWG

Spammers taking advantage of occasion /event and holidays offerings, phisher create the spammers for these events, Therewherea larger number of phishing campaign run over on Japanese earthquake about Olympic and Xmas.
It can be any big event. The way criminal /Phisher executes operation there is much great opportunity to get success.
Also popular scam of economical fears these scams include phishing emails those are coming from some financial institution or some economical upcoming are there.



Atten

Conrad Magnus <magnus@giga.ocn.ne.jp>

Be careful with this message. Similar messages were used to steal people's personal information. Unless you trust the sender, don't click

Dearest,

Thank you very much for your email,

I stated in my last email my name is Conrad Magnus, i worked with an Insurance company in Scotland.

First i contacted you for trust , what am disclosing to you now is a trusted deal and you have to believe in yourself, I have been searching for someone like you for long now which was very difficult to locate you through email.

I pray that you will welcome my email with common goal of a sense of understanding.

I didn't contact you to waste your time no , i contacted you for a reality and 100% risk free deal which i know all about.

I will give you the brief detail for you to ask any question in regards to clarify you , and i assure you will never ever regret in assisting me regards to this reality deal.

I have a client who use to bank with our Insurance company,he is late now and i know about his $45.8 Million United State Dollars presently in a save with company also in London.

I will help you to direct you how to claim the funds and also furnish you with all the necessary document , i will take care of the lawyer who will help us regards to prepare the documents to back you up as the next of kin to the beneficiary.

This deal is reality and 100% risk free.

Please feel very free to ask me any question regards to it.

I await your email,

Best Regards.

Conrad Magnus.

III. PROPOSED ALGORITHM

There is a need foran advance email filter technique to detect social engineering attacks. Mail server message log looks as follows

Apr 01 06:43:39 zimbra postfix/smtpd[31272]: NOQUEUE: filter: RCPT from unknown[115.159.87.234]: <anna@1g77.com>: Sender address triggers FILTER smtp-amavis:[127.0.0.1]:10024; from=<anna@1g77.com> to=<removed> proto=ESMTP helo=<anna.1g77.com>

It perfectlyshows this domain does not exist, need to identify these type of domain has block those asap, by using bigdata analytics we can see how many domainsis newly born and what are those.

Trail show how we can minimize the risk of social engineering "Phishing" in the organization,

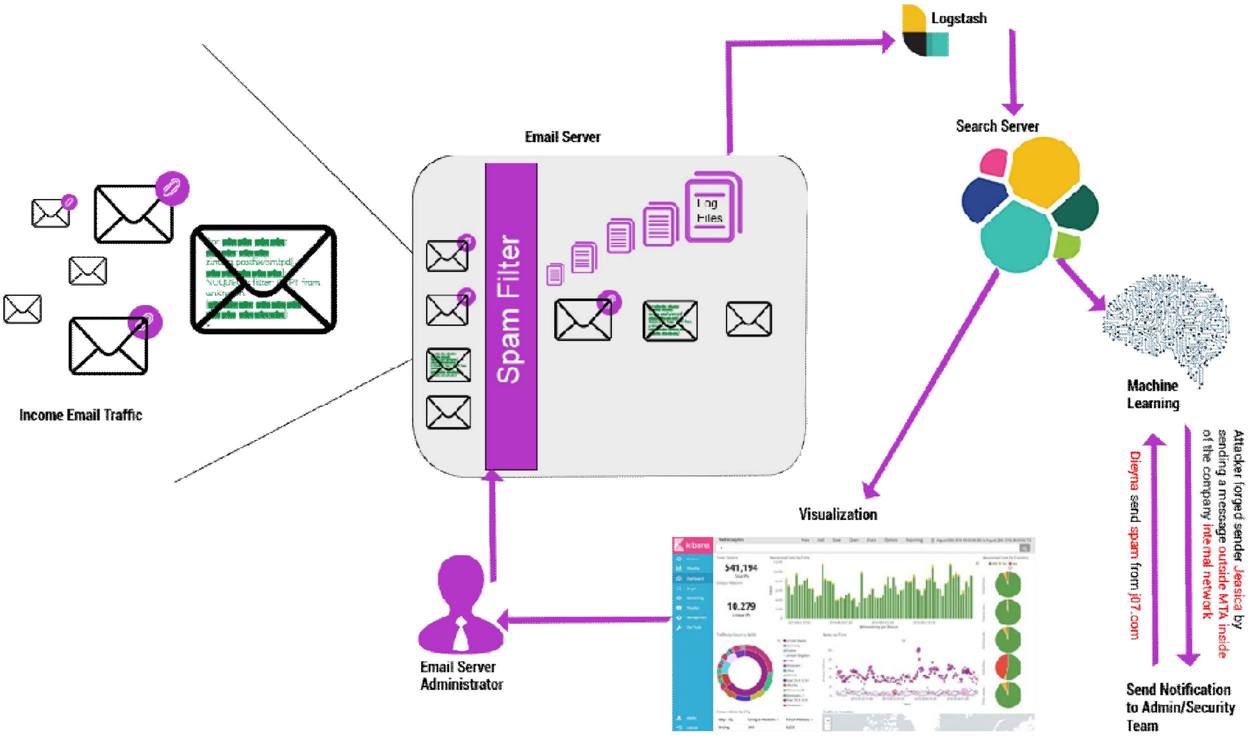

Logstash
Search Server
Email Server
Log Files
Spam Filter
Income Email Traffic
Machine Learning
Attacker forged sender Jeasica by sending a message outside MTA inside of the company internal network
Dieyna send spam from j07.com
Send Notification to Admin/Security Team
Visualization
Email Server Administrator

Figure 3: Proposed 5 Component Structure to Stop Social Engineering Attack.

Social Engineering email attack protection using 5 components. Email server spam filter which will be first component interacting with income email traffic it stops unwanted emails also logs are written.The second component read logfile and send to the search server, It will be continuously working with search server to send updated information to the server. One information is inside the search server, four and five component will start working on there given task four component is machine learning component will be taking data and training machine model by working on following questions

Who? Where? What? When? How?
Who: Sends the emails ? to whom?
Where? Does the email come from?
What? Is the intention of this email?
When? Is the email begin sent?
How? Is the email being relayed?

Once the model is trained and it will detect spam and scam emails e.g. “*Dieyna send spam from j07.com*” or “*Attacker forged sender Jessica by sending a message outside MTA inside of the company internal network*”, notification is sent to organization cybersecurity team as well as email server Admin.
Next component is visualization system which will get data to form search server and shows statistics when, where and who send emails, it will list down all new domains as well as a count of old domains which will give a statistical view of spam emails.
Social Engineering AttackProtection detects targeted attack emails and prevents them from reaching endpoints.

IV. Simulation Results

The system for protecting phishing emails/ emails attack have detected spam and provided good protection this research includes bigdata analytics with machine learning to protect email attacks. Trail figure shows how much traffic is coming from which domain all the others domain are declared spam emails which are notified to cybersecurity cell.

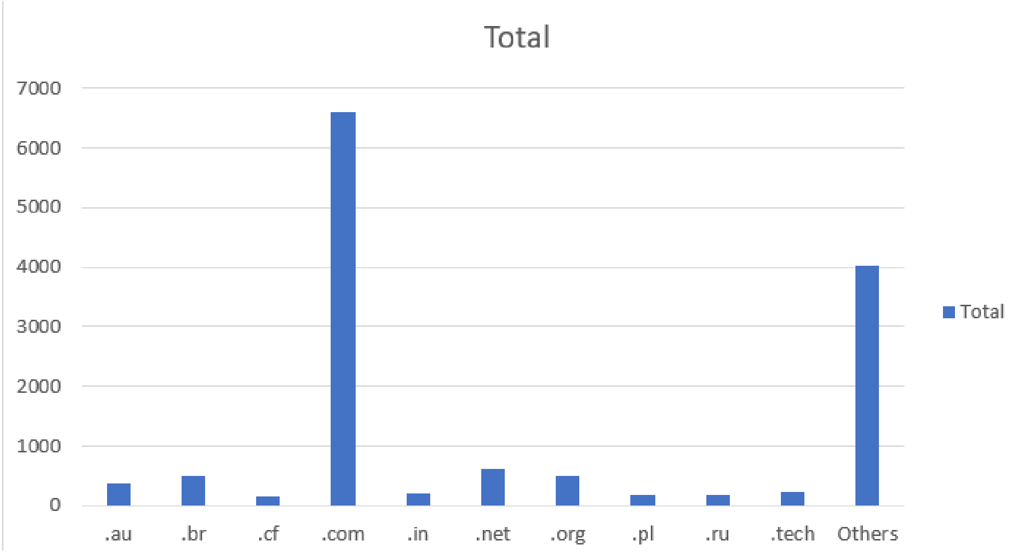


Figure 3: Proposed 5 Component Structure to Stop Social Engineering Attack.

V. Conclusion and Future Work

Also, there should be proper awareness program for employees so that they have knowledge about attacks Education is one of the best ways to building the trust and overcome phishing fears. Helping people to understand Which kind of emails are harmful, Phishing always evolves in different forms were taking advantage of human behaviorto mention its hack the human by the human.

Future work on this will be integrating auto modification of spam filters according to machine learning results. So that no human interaction is needed. Also, need to check possibilities of Robotic Process Automation and bigdata technologies contribution on automation.

REFERENCES

1. Ullah, F., and Babar, M.A (2018), "Architectural Tactics for Big Data Cybersecurity Analytics Systems: AReview"
2. Nicholas Carr. "The Limits of Social Engineering". https://www.technologyreview.com/s/526561/the-limits-of-social-engineering/
3. Neil DuPaul,"Hacking the Mind: How & Why Social Engineering Works"https://www.veracode.com/blog/2013/03/hacking-the-mind-how-why-social-engineering-works.
4. Davide Andreoletti, SUPSI and Enrico Frumento, CEFRIEL "https://www.dogana-project.eu/index.php/social-engineering-blog/11-social-engineering/92-cambridge-analytica "
5. David Kebo "Gartner: Social engineering, big data top security priorities for 2013" https://www.us-analytics.com/hyperionblog/qlikview/2012/11/gartner-social-engineering-big-data-top-security-priorities-for-2013"
6. Cárdenas, A.A., P.K. Manadhata, and S. Rajan, Big data analytics for security intelligence. *"https://downloads.cloudsecurityalliance.org/initiatives/bdwg/Big_Data_Analytics_for_Security_Intelligence.pdf"*.
7. https://www.us-analytics.com/hyperionblog/qlikview/2012/11/gartner-social-engineering-big-data-top-security-priorities-for-2013
8. Joseph A. Cazier, Christopher M. Botelho, "Social Engineering's Threat to Public Privacy" Appalachian State University.
9. Trend Micro, "https://www.trendmicro.tw/cloud-content/us/pdfs/business/datasheets/ds_social-engineering-attack-protection.pdf"*Trend Micro*.
10. Dimensional Research, THE RISK OF SOCIAL ENGINEERING ON INFORMATION SECURITY "https://www.stamx.net/files/The-Risk-of-Social-Engineering-on-Information-Security.pdf"
11. SymantecResearch, Fraud Alert:Phishing White Paper"http://www.symantec.com/content/en/us/enterprise/white_papers/b-fraud-alert-phishing-wp.pdf"
12. APWG, Phishing Activity Trends Report 1Q -2018 "https://docs.apwg.org/reports/apwg_trends_report_q1_2018.pdf"
13. DigiCert,inc,Phishing:A Primer on what phishing is and how it works"https://www.digicert.com/news/DigiCert_Phishing_White_Paper.pdf"