RESEARCH ARTICLE OPEN ACCESS

Phishing Page and Malicious URL Detection via Support Vector Machine using Page Layout Feature

Dr. E.K. Vellingiriraj ME., Ph.D.,[1] , G.Savitha[2]

[1]Head of Department of Computer Applications, Nandha Engineering College (Autonomous), Erode, Tamilnadu, India.

[2]Final MCA, Department of Computer Applications, Nandha Engineering College, (Autonomous),Erode, Tamilnadu, India.

Email: [1]giri@nandhaengg.com , [2]savithagnana@gmail.com

Abstract— The web technology has come the corner gravestone of a wide range of platforms, similar as mobile services and smart Internet-of- effects (IoT) systems. In several surroundings, stoner data is aggregated for a pall grounded platform, where web operations are used as a key interface to pierce and configure stoner data. Securing the web interface requires results to deal with pitfalls from both specialized vulnerabilities and social factors. The bushwhackers use web runners visually mimicking licit websites, similar as banking and government services, to collect druggies' sensitive information. Being phishing defense mechanisms grounded on URLs or runner contents are frequently finessed by bushwhackers. The World Wide Web has come the most essential criterion for information communication and knowledge dispersion. It helps to distribute information timely, fleetly and fluently. Identity theft and identity fraud are appertained as two sides of cyber crime in which hackers and vicious stoner s gain the particular data of being licit druggies to attempt fraud or deception provocation for fiscal gain. E-Mails are used as phishing tools in which licit looking emails are transferred making the genuine druggies identity with licit content with vicious URLs. SpamE-Mails emerges or transforms as Phishing matters. Spoofed Matters plays a vital part in which the hackers pretends to be a licit sender posing to be from a licit association which divulges the stoner to give his particular credentials. The content may escape from Content grounded pollutants or the dispatch may be without any body of the communication except vicious URL in it. This paper identifies vicious URLs in dispatch through reduced point set system. In addition, phishing runners are plant out grounded on CSS attributes values.

Keywords— ***Data Mining, Phishing Mails, Anti-SPAM Filtering,*** *Phishing Classification*

I. INTRODUCTION

Data mining is the process of rooting patterns from data. Data mining is seen as an decreasingly important tool by ultramodern business to transfigure data into an instructional advantage. It's presently used in a wide range of profiling practices, similar as marketing, surveillance, fraud discovery, and scientific discovery.

The affiliated terms data dredging, data fishing and data poking relate to the use of data mining ways to test portions of the larger population data set that are (or may be) too small for dependable statistical consequences to be made about the validity of any patterns discovered (see also data-poking bias). These ways can still, be used in the creation of new hypothesises to test against the larger data populations.

The homemade birth of patterns from data has passed for centuries. Beforehand styles of relating patterns in data include Bayes'theorem (1700s) and retrogression analysis (1800s). The proliferation, ubiquity and adding power of computer technology has increased data collection and storehouse. As data sets have grown in size and complexity, direct hands-on data analysis has decreasingly been stoked with circular, automatic data processing. This has been backed by other discoveries in computer wisdom, similar as neural networks, clustering, inheritable algorithms (1950s), decision trees (1960s) and support vector machines (1980s).

A primary reason for using data mining is to help in the analysis of collections of compliances of geste. Similar data are vulnerable to collinearity because of unknown interrelations. An necessary fact of data mining is that the (sub-) set (s) of data being analysed may not be representative of the whole sphere, and thus may not contain exemplifications of certain critical connections and behaviours that live across other corridor of the sphere. To address this kind of issue, the analysis may be stoked using trial- grounded and other approaches, similar as Choice Modelling for mortal-generated data. In these situations, essential correlations can be moreover controlled for, or removed altogether, during the construction of the experimental design.

There have been some sweats to define norms for data mining, for illustration the 1999 European Cross Industry Standard Process for Data Mining (CRISP-DM1.0) and the 2004 Java Data Mining standard (JDM1.0). These are evolving norms; latterly performances of these norms are under development. Independent of these standardization sweats, freely available open- source software systems like the R Project, Weka, KNIME, RapidMiner and others have come an informal standard for defining data-mining processes. Specially, all these systems are suitable to import and export models in PMML (Prophetic Model Markup Language) which provides a standard way to represent data mining models so that these can be participated between different statistical operations. PMML is an XML- grounded language developed by the Data Mining Group (DMG), an independent group composed of numerous data mining companies. PMML interpretation4.0 was released in June 2009.

II. RELATED WORKS

In the paper (1) the authors stated that for times, senders have abused SPF- authorized and DKIM- inked dispatches to achieve sphere- position dispatch authentication. Grounded on that authentication, colorful correspondence receivers have tried to cover senders by using (DKIM) and/ or (SPF) results to descry and block unauthorized dispatch. (A detailed discussion of the pitfalls these systems attempt to address can be plant in (DKIM-THREATS).) Still, there has been no single extensively accepted or intimately available medium to communicate sphere-specific communication authentication programs, or to request reporting of authentication and disposition of entered correspondence.

In the paper (2) the authors stated that in recent times, there has been a dramatic shift from bulk spam emails to targeted dispatch phishing juggernauts. Similar attacks have started to beget huge brand, finan-cial and functional damage to organisations encyclopedically. Phishing attacks involve simple, straightforward, masquerading methodology. The end is to bait and trick an unknowing victim in order to evoke as important information as possible, using SMS, dispatch, WhatsApp and other messaging services, or phone calls that have been caricatured to appear is if they're from known, dependable musketeers or associates.

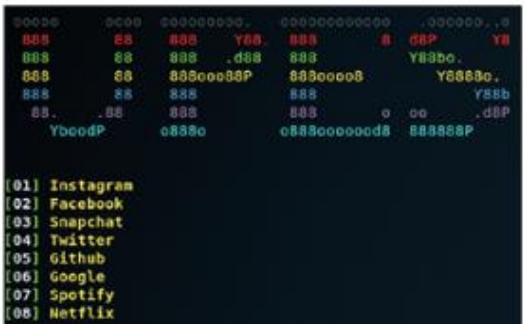
The intent is to get the victims to click and log into reproduced web doors similar as company intranet or bank spots and social networking spots similar as Facebook, Instagram, Twitter or indeed Yahoo and Gmail spots. Once the unknowing victims click on the URL transferred by the bushwhacker, rather of the original point they're directed to the bushwhacker’s fake point. On trying to log in or submitting infor-mation on that website, victims give the bushwhacker with sensitive information.

This can include stoner ID, dispatch, pass- word, address, mobile number, date of birth and payment card details, among other effects. Cyber bushwhackers have enhanced their methodologies to include personalised attacks. Targeting elderly- position, high- value help similar as the head of HR, C- position directors similar as CISO, CTO, FEATURE Computer Fraud & Security September 202016CFO or board members is an advanced form of phishing attack against individualities, known as whaling.

Spear-phishing attacks, on the other hand, target specific individualities within the organisation, and are largely personalised. Similar individualities include finance platoon members, IT security platoon members or indeed new hires. Piecemeal from using reproduced web doors, bushwhackers also target two- factor authentication by copying one- time watchwords (OTPs) as well as creating fake QR canons which, if scrutinized by mobile phones, respond by offering huge abatements at caffs, grocery stores or ménage service stores in return for online payment, which obviously goes to the bushwhacker’s account.

PHISHING TAXONOMY

The authors classified phishing attacks based on new and upcoming tactics adopted by cyber attackers while luring victims and performing fraudulent activities to obtain personal and sensitive information. Tactical and social engineering techniques are detailed in Figure 2.1.

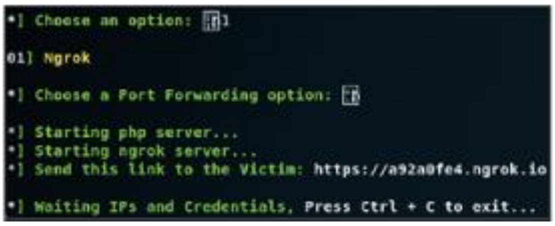


```
888          88    888     Y88.  888        8  d8P      Y8
888          88    888    .d88   888           Y88bo.
888          88    888ooo88P     888ooooo8         Y8888o.
888          88    888           888                  Y88b
 88.       .88     888           888        o  oo      .d8P
   YboodP          o888o         o888ooooood8  888888P

[01] Instagram
[02] Facebook
[03] Snapchat
[04] Twitter
[05] Github
[06] Google
[07] Spotify
[08] Netflix
```

FIGURE 2.1 PROPOSED UNIQUE PHISHING TAXONOMY

Cyber culprits perform phishing conditioning for plutocrat, and to insure that their scheme is effective and evades discovery, they don't make rational or ethical opinions. To combat phishing, this exploration presents the phisher’s mindset and methodology of attack. As shown in Figure2.2, the bushwhacker’s toolkit has options to choose from, including using reproduced social media spots, gathering two- factor authentication OTP law or using a QR law in the form ofpre-designed templates.

To induce the reproduced forged Twitter link, the authors set up a rear lair using an Ngrok deputy on the bushwhacker’s command and control (C&C) garçon. This deputy operation launches multiple virtual coverts as original network services. These capture the network business for detailed examination.



```
*] Choose an option: 1

01] Ngrok

*] Choose a Port Forwarding option:

*] Starting php server...
*] Starting ngrok server...
*] Send this link to the Victim: https://a92a0fe4.ngrok.io

*] Waiting IPs and Credentials, Press Ctrl + C to exit...
```

FIGURE 2.2 THE INITIAL UNIQUE PHISHING TOOLKIT

This helps gather the victim’s sensitive details, using one unique phishing system from three options to maximize success. However, this is followed with options to choose between reproduced Instagram, Facebook.

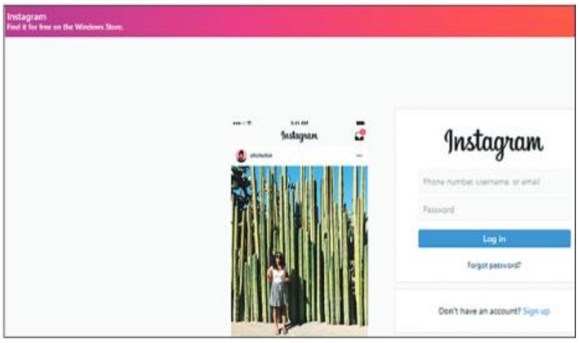
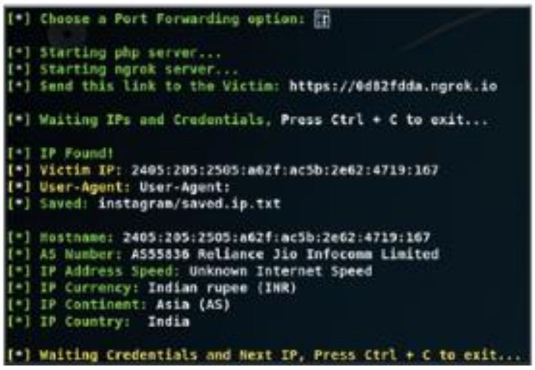

Instagram
Find it for free on the Windows Store.
Instagram
Phone number, username, or email
Password
Log in
Forgot password?
Don't have an account? Sign up

FIGURE 2.3 THE SOCIAL MEDIA SITE OPTIONS

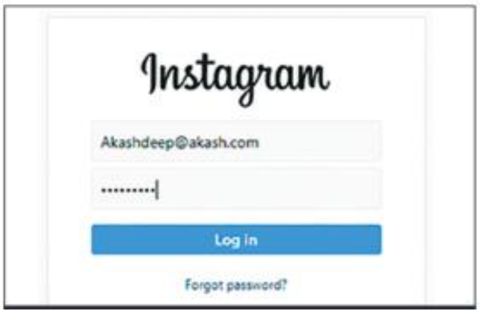
If the attacker starts with the social media spots option. Assuming the attacker chooses the Twitter social media template (option 4), a hinder deputy garcon is started on the attacker's system. In our attack, the forged link is https//e89e09404a68.ngrok.io as shown in Figure2.4.



```
[*] Choose a Port Forwarding option: 

[*] Starting php server...
[*] Starting ngrok server...
[*] Send this link to the Victim: https://0d82fdda.ngrok.io

[*] Waiting IPs and Credentials, Press Ctrl + C to exit...

[*] IP Found!
[*] Victim IP: 2405:205:2505:a62f:ac5b:2e62:4719:167
[*] User-Agent: User-Agent:
[*] Saved: instagram/saved.ip.txt

[*] Hostname: 2405:205:2505:a62f:ac5b:2e62:4719:167
[*] AS Number: AS55836 Reliance Jio Infocomm Limited
[*] IP Address Speed: Unknown Internet Speed
[*] IP Currency: Indian rupee (INR)
[*] IP Continent: Asia (AS)
[*] IP Country:  India

[*] Waiting Credentials and Next IP, Press Ctrl + C to exit...
```

FIGURE 2.4 THE TWITTER LINK TO SEND TO THE VICTIM

Trained and alive workers can descry this link as an incorrect URL or forged link. The provocation for this disquisition is to increase phishing awareness and not factual phishing, so DNS spoofing is not applied also. That would easily give a legit link or a similar one to those used by Twitter.

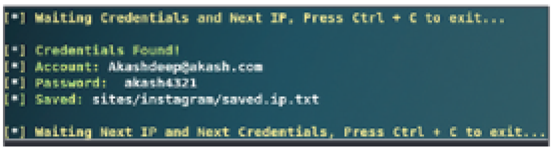


Akashdeep@akash.com

•••••••••

FIGURE 2.5 DISGUISING THE FORGED PHISHING LINK

The attacker has options to use enhanced phishing styles, a two- factor authentication attack to snare the victim's OTP or indeed bait the victim to shoot Paytm, Google or WhatsApp capitalist by surveying a QR Code, as illustrated in Figure 2.6



```
[*] Waiting Credentials and Next IP, Press Ctrl + C to exit...

[*] Credentials Found!
[*] Account: Akashdeep@akash.com
[*] Password:  akash4321
[*] Saved: sites/instagram/saved.ip.txt

[*] Waiting Next IP and Next Credentials, Press Ctrl + C to exit...
```

FIGURE 2.6 SPOOFED TWISTED LOGIN PAGE WITH THE ATTACKER'S C&C DASHBOARD

Features for phishing discovery

Creating phishing mindfulness generally involves having end druggies attend a course, read documents related to' good practices'and'Dos & Do n'ts'. Still, it's mortal nature to forget and concentrate on the micro task when going about daily chores. The authors reviewed over 300 phishing matters entered by Gmail and Yahoo. The top 15 unique tactics and phishing features espoused by cyber bushwhackers are presented in Table 1. The authors propose creatinganti-phishing rules by IT security brigades to check and validate these phishing features. These can fluently help descry and block phishing attacks.

The Internet of Effects

Bio-wearable, body detectors, Internet of Effects (IoT) bias or smart systems in our homes, services and structures have changed our lives for the better. These bias have come an natural part of our work and lives, plant everyplace in our services, houses, seminaries, vehicles, hospitals, manufacturing diligence and indeed on our bodies.

Malware stationed via phishing is able of controlling these bias, which may well beget further detriment than benefit. IoT bias and detectors generally collect data and are substantially connected to colorful networks and the Internet. This leads to an existent's tête-à-tête identifiable information (PII), position and voice being stored in these bias.

This PII can range from particular details similar as name, age, position, dispatch word credentials or indeed health data. Therefore for any cyber bushwhacker, there are easy, low- hanging means with value information, making any existent – not just high- value directors – implicit targets.
In the paper (3) the authors stated that to exclude spamemails, several textbookanti-spam systems were created to dissect the textual content ofe-mails and classify them. Owing to the good performance of these systems, spam dispatches began to do in images. This rendered textbookanti-spam systems useless, thereby fostering the development of imageanti-spam systems. Image processing is much further computationally precious than textbook processing, and the results of imageanti-spam systems have been inferior to those of textbook systems.
This paper proposed an imageanti-spam system that makes use of colorful styles of image point birth and an artificial neural model to classifye-mails. The birth styles are estimated both collectively and in combination. The neural

model is completely estimated using intimately available databases. The use of these databases is described in detail in order to grease reproducible results. Besides assaying the bracket capability of the proposed system, this study also evaluates its computational costs, including costs for rooting features and classifying images. The results are promising both in terms of rates of correct bracket and of false cons produced by theanti-spam system, as well as in terms of its computational cost.

III. METHODOLOGY

The exploration's purpose is to develop the URL analyzer system with the help of minimized phishing point set identifies the vicious URL in the emails.

3.1 URL ANALYZER

Phishing URLs are anatomized grounded on verbal features and URL's host grounded features. The verbal point analyses the URL format. URLs contain both host name and the path. For illustration, consider‘www.annauniversity.edu/emmrc1/emmrctest.html’ , the host name iswww.annauniv.edu andemmrc1/emmrctest.html is the path. The proposed methodology analyses host grounded features similar as IP addresses given in the suspicious list, colorful verbal grounded features similar as URL encoding, presence of hexadecimal character, suspicious letters/ characters, or vicious IP addresses to hide and analyses the chances of words to check whether emails contain any suspicious links to avoid druggies falling by phishing attacks as illustrated in Fig3.1.

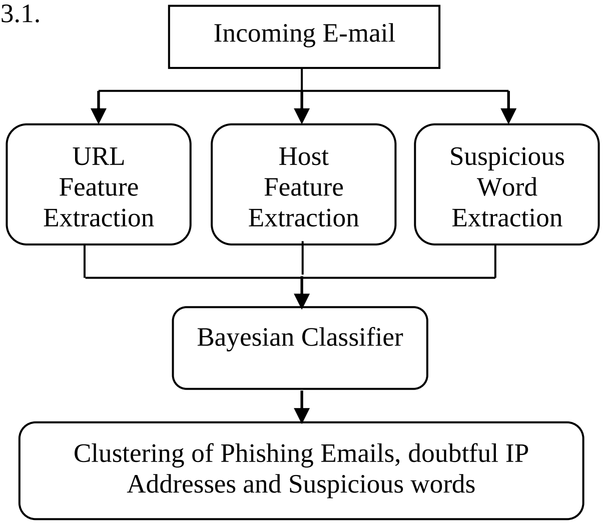

Incoming E-mail
URL Feature Extraction
Host Feature Extraction
Suspicious Word Extraction
Bayesian Classifier
Clustering of Phishing Emails, doubtful IP Addresses and Suspicious words

FIG 3.1 FINDING DOUBTFUL EMAIL, IP AND WORDS AND CLUSTER CLASSIFICATION

3.2 Verbal FEATURES

Verbal features analyses the format of the URL. It includes the length of the host name, length of the URL, the number of blotches, presence of suspicious characters similar@ symbol, hexadecimal characters and other special double characters similar as (‘.’,‘ = ’,‘$’, ‘’,etc.) either in the host or path name. IP addresses and hexadecimal characters are used to hide the factual URLs. For illustration consider URL http//www.bankingcompany.com/online/transaction/ website/phishing.html” which is docked using the IP address http//132.115.201.115 which looks like licit and not suspicious.

The URL can also be represented using hexadecimal base values with a ‘’ symbol. It may represent any special characters Spoofguard linked the‘@’and‘-’ symbol most prominent in phishing URLs.

A@ symbol in a URL will enable the URL at the left to discard which is licit URL and right to enter into the phishing point. Consider the URL http//www.citibank.com@phishingsite.com” will enter into “ phishingsite.com” and discards “www.citibank.com”. These kinds of ways use the factual phishing website to disguise and pose as licit spots.

3.3 HOST Grounded FEATURES

Host grounded features identify the position, proprietor and how vicious spots are hosted and managed. Age of the sphere is used to identify when vicious websites are hosted similar that they've lower age or fairly new to gain the stoner credentials. They will be lately registered transferring further matters and some disciplines may not be available indeed at the time of checking. It obtains the data in the number of months and some may be in times more lately.

The WHOIS lookups on the WHOIS garçon is used to recoup the sphere enrollment date, and if the sphere enrollment entry isn't plant on the WHOIS garçon, this point simply return-1, thinking it suspicious. The IP Addresses list are stored in a train and brought. All of them are checked in each of the correspondence contents for their presence.

3.4 SUSPICIOUS WORDS

Some words similar as includes Secure, Account, Update, Login, Verify, Signin, Banking, Notify, Click, Inconvenient, word etc and theirCo-Occurences in the phishing matters. So they're checked in all emails to classify them into phishing matters.

3.5 APPROACH- BAYES CLASSIFIER

Bayes classifier is acclimated in spam pollutants similar that individual features of URLs are distributed singly of the values of other features. Bayes theorem is used to calculate the probability of thesis for the event B, handed with the training data A,

. $P (B| A) = P (A| B) * P (B)/ P (A)$ (1)

It's frequently easier to calculate the chances, P (A| B), P (A), P (B) for the probability that's needed. Reasoning Baye's rule, assume that licit and phishing websites do equal in number and hence with equal probability, also the posterior probability that the point vector X belongs to a vicious URL. Then, P (A) = Probability of point F in phishing and licit dataset.

In addition with CSS attributes checking, the proposed system checks the correspondence contents against the phishing correspondence disciplines like g00gle, micr0s0ft, etc which is the suspicious list. Also, the IP addresses are maintained in the suspicious list of which correspondence contents are checked. Also, the words like signin, corroborate, word, account, etc are also maintained in the suspicious list of which correspondence contents are checked.

All the phishing mails counts are also plant out. The IP addresses as well as suspicious words tentative probability are also plant out. In addition, arbitrary timber algorithm is used to prognosticate the model as it helps better in colorful ways. Random Forest is used then to develop a vaticination system in order to dissect and prognosticate the spammatters.

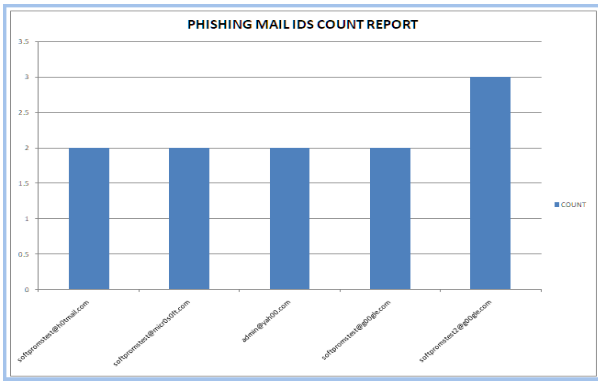


FIG 4.1 PHISHING MAIL IDS REPORT

IV. EXPERIMENT RESULTS AND FINDINGS

The following list of emails is fetched out with their count are extracted during phishing mail ids extraction.

During phishing mail ids extraction, the following list of emails is fetched out with their probability values.

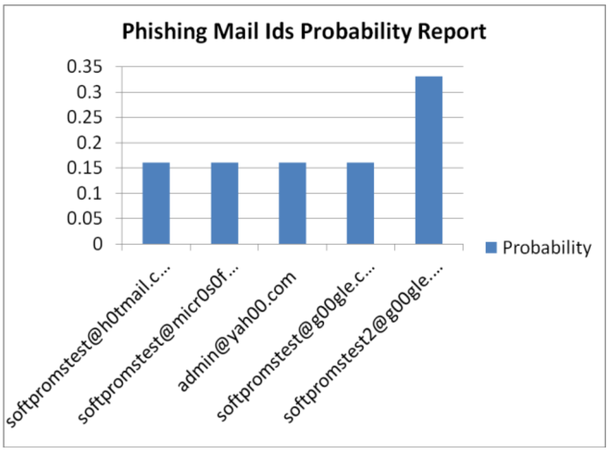


FIG 4.2 PHISHING MAIL IDS REPORT

During suspicious words extraction, the following list of words is fetched out with their count.

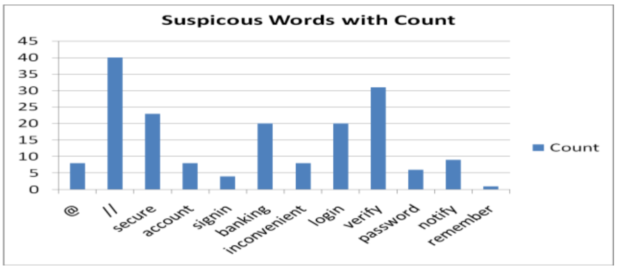


FIG 4.3 SUSPICIOUS WORDS REPORT

During suspicious words extraction, the following list of words is fetched out with their probability values.

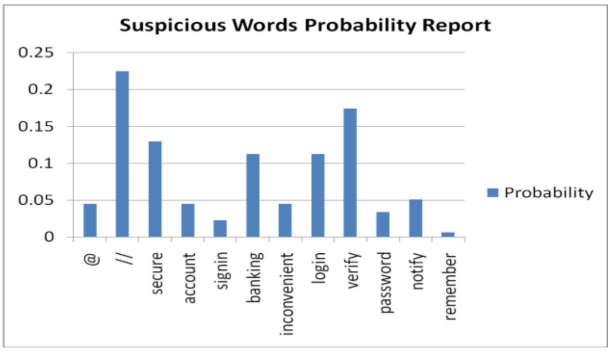


FIG 4.4 SUSPICIOUS WORDS REPORT

During suspicious IP addresses extraction, the following list of addresses is fetched out with their count.

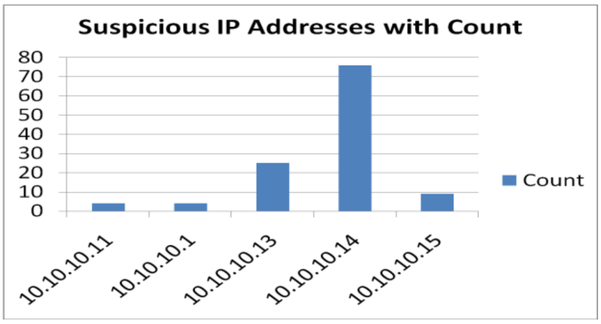


FIG 4.4 SUSPICIOUS IP ADDRESSES REPORT

During suspicious IP addresses extraction, the following list of addresses is fetched out with their probabilities.

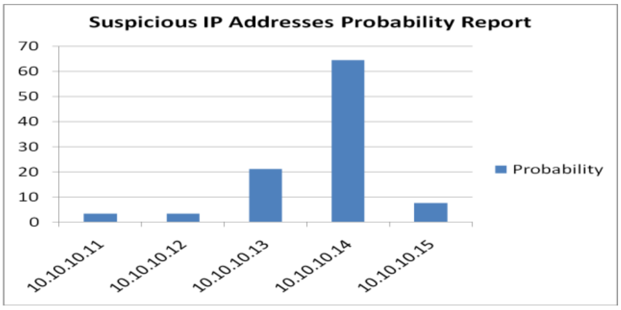


FIG 4.5 SUSPICIOUS IP ADDRESSES REPORT

During Sequential Pattern clustering threshold value is given as 0.5 and the following results are produced.

During Sequential Pattern clustering using cosine similarity the following groups of words are found out.

Output Clusters of Suspicious Words:	
1	secure account signin banking inconvenient @
2	login verify password //
3	@
4	//
5	Notify

TABLE 4.1 OUTPUT CLUSTERS OF SUSPICIOUS WORDS

During Sequential Pattern clustering using cosine similarity the following groups of IP Addresses are found out.

Output Clusters of Suspicious IP Addresses:	
1	10.11.11.11 10.11.11.11
2	192.160.1.1 192.160.1.1
3	20.20.20.20 20.20.20.20
4	10.10.10.11 10.10.10.11
5	10.10.10.12 10.10.10.12

TABLE 4.2 OUTPUT CLUSTERS OF SUSPICIOUS IP ADDRESSES

During Sequential Pattern clustering using cosine similarity the following groups of mail ids are found out.

Output Clusters of Phishing Mail Ids	
1	softpromstest@h0tmail.com
2	softpromstest1@micr0s0ft.com
3	admin@yah00.com softpromstest@g00gle.com softpromstest2@g00gle.com softpromstest2@g00gle.com

TABLE 4.3 OUTPUT CLUSTERS OF PHISHING MAIL IDS

V. CONCLUSION

Hackers Shirkanti-spam filtering ways using embedding vicious URL in the communication contents. So the URL analyzer system is used with the help of minimized phishing point set to identify the suspicious/ vicious URL in emails. Phishing Emails, Suspicious words and IP Addresses count are plant out. Phishing Emails, Suspicious words and IP Addresses tentative probability values are plant out. Affiliated Phishing Emails, Suspicious words and IP Addresses are grouped into clusters. Cosine similarity grounded successional pattern mining is used with threshold value to group the dispatch, words, IP address patterns in the dispatch data set. The results show that end druggies are ignorant of zero rather of'o'in the correspondence ids as well as one rather of'1'. So the developed operation is able of detecting similar correspondence ids as phishing matters. Then Random Forest is used to develop a vaticination system in order to dissect and prognosticate the spam matters.

REFERENCES

[1] M. Kucherawy and E. Zwicky,Domain-Based Message Authentication,Reporting, Conformance (DMARC), Internet Engineering Task Force,document RFC 7489, Mar. 2015. [Online]. Available: https://www.rfc-editor.org/info/rfc748

[2] A. Bhardwaj, V. Sapra, A. Kumar, N. Kumar, and S. Arthi, ''Why is phish-ing still successful?''Comput. Fraud Secur., vol. 2020, no. 9, pp. 15–19,Sep. 2020.

[3] O. A. S. Carpinteiro, B. C. Sanches, and E. M. Moreira, ''Detectingimage spam with an artificial neural model,''Int. J. Comput. Sci. Inf.Secur., vol. 15, no. 1, pp. 296–314, Jan. 2017. [Online]. Available: https://www.academia.edu/36003643/ Journal_of_Computer_Science_IJCSIS_ January_2017_Full_Volume_pdf

[4] Justin Ma, Lawrence Saul, K., Stefan Savage and Geoffrey Voelker, M. "Identifying Suspicious URLs: An Application of Large-Scale Online Learning", In ICML '09: Proceedings of the 26th Annual International Conference on Machine Learning, pp. 681-688, 2009. 7

[5] Y. Niu, Y.-M. Wang, H. Chen, M. Ma, and F. Hsu. A Quantitative Study of Forum Spamming Using Context-based Analysis. In Proceedings of the Symposium on Network and Distributed System Security (NDSS), San Diego, CA, Mar. 2007.

[6] J. Ma, L. K. Saul, S. Savage, and G. M. Voelker. Identifying Suspicious URLs: An Application of Large-Scale Online Learning. In Proc. of the International Conference on Machine Learning (ICML), Montreal, Quebec, June 2009.

[7] Colin Whittaker, Brian Ryner and MarriaNazif, "Large-Scale Automatic Classification of Phishing Pages", In proceedings of NDSS, 2010.

[8] Neil Chou, Robert Ledesma, Yuka Teraguchi, Dan Boneh and John Mitchell, "Client-side defense against web-based identity theft", In 11th Annual Network and Distributed System Security Symposium (NDSS '04),San Diego, 2004.

[9] G. Casto, O. Fisher, R. Moll, M. Nazif, and D. Born. Client specification for the Google Safe Browsing v2.1 protocol. http://code.google. com/p/google-safe-browsing/wiki/ Protocolv2Spec, 2007.

[10] Blake Ross, Collin Jackson, Nicholas Miyake, Dan Boneh, and John Mitchell. A browser plug-in solution to the unique password problem. In Proceedings of 2005 USENIX Security Symposium, 2005.

[11] A. Attar, R. M. Rad, and R. E. Atani, "A survey of image spamming andfiltering techniques,"Artificial Intelligence Review, vol. 40, pp. 71–105,2013.

[12] G. Fumera, I. Pillai, and F. Roli, "Spam filtering based on the analysis oftext information embedded into images,"Journal of Machine LearningResearch, vol. 7, pp. 2699–2720, 2006.

[13] M. Dredze, R. Gevaryahu, and A. Elias-Bachrach, "Learning fast clas-sifiers for image spam," inProceedings of the Conference on Email andAnti-Spam (CEAS), 2007.

[14] B. Mehta, S. Nangia, M. Gupta, and W. Nejdl, "Detecting image spamusing visual features and near duplicate detection," inProceedings ofthe International Conference on World Wide Web (WWW), 2008, pp.497–506.

[15] T. Liu, W. Tsao, and C. Lee, "A high performance image-spam filteringsystem," inProceedings of the International Symposium on DistributedComputing and Applications to Business Engineering and Science(DCABES), 2010, pp. 445–449.

[16] H. Cheng, Z. Qin, C. Fu, and Y. Wang, "A novel spam image fil-tering framework with multi-label classification," inProceedings ofthe International Conference on Communications, Circuits and Systems(ICCCAS), 2010, pp. 282–285.

[17] C. Wang, F. Zhang, F. Li, and Q. Liu, "Image spam classificationbased on low-level image features," inProceedings of the InternationalConference on Communications, Circuits and Systems (ICCCAS), 2010,pp. 290–293.

[18] B. Al-Duwairi, I. Khater, and O. Al-Jarrah, "Texture analysis-basedimage spam filtering," inProceedings of the International Conferencefor Internet Technology and Secured Transactions (ICITST), 2011, pp.288–293.