Spear Phishing: Diagnosing Attack Paradigm

Deepali Pande[1], Dr. P. S. Voditel[2]

Department of Computer Application,
Shri Ramdeobaba College of Engineering and Management, 440013 Nagpur,
Maharashtra, India.

Abstract
Internet is a rich source of web media and social networking applications. A cluster of users interconnect using those forming mutligroups. But the usage of web resources imprudently is causing doors to phishing, pharming and targeted phishing attacks. Careless use of social networking applications like LinkedIn, pinterest, whatsapp, face book and twitter barely from smart phones have become extrinsic sources for phishing and pharming attacks. Hence, it is essential to understand the pinholes of these attacks and their relationship with variants of user-agents on distributed platform. In this paper, we direct our survey in finding extrinsic porches influential to nasty invasions as attack entry point analysis. Also, we incline our detection considering recursive NM cache poisoning as the source of spear-phish attack. We present a detail analysis to determine spear- phishing. We evaluate and compare the spear phish feature detection attributes with PhishTank, a benchmark dataset.

Keywords:
phishing; spear phishing; pattern mining; classification

1. Introduction

Phishing is a crime which involves ethical and technical artifices in stealing consumers' personal data for malevolent practices. The intruders misuse social engineering schemes ethically in disguising the users to practice phishing attacks. Phishing attack thrives on persuading the users to acquire information from them. It is usually implemented after learning various semantics from the users' interaction. Some of the most common phishing techniques include implanting crime-wares to steal credentials, intercepting through the websites, hijacking the consumers' computer to compensate for locked data and many more. The computer scientists have lavishly been working to provide solutions which can curb phishing. The taxonomy of solutions for phishing has been categorized into two branches namely client based solutions and server side solutions. IT sector have been providing server side solutions for the phishing problem through schemes and soft ware's like anti-spams, anti-malwares, anti-virus update patches for dual layer security and many more.

The client based solutions majorly include email analysis and are implemented on end-user side through methods like browser plug-ins. The method of email analysis deals with filtering the emails as spams after content analysis by training a Bayesian filter. Other approaches include black listing the unreliable URLs after identified as malicious. The blacklist is then queried by the browser at runtime whenever the page is loaded. If it is found doubtful then an alert message is generated to block it. Many recent methods used the technique of controlled information flow in which soft -wares are used to monitor URL obfuscations and fake domain names. This approach provides solution through an anti-phishing toolbar and browser plug-ins which monitors and generates alert messages in case of distinguished phishing activity. A most recent scenario as studied from APWG uses visual similarity analysis to distinguish a phishing website from the legitimate one. It has been very efficient to discriminate and control phishing attacks. But it succumbs to barn targeted attacks namely the spear phish attack. Spear phishing is a dogma in which a particular organization or a specific user is targeted. The victim is lavished with the difficulty in reasoning the authenticity of the URLs and emails and may thereby fall into a deceitful mesh.

The approaches discussed in the section above dealt with issues of phishing and practices to curb the same. But the problem of spear phishing being a complete non-generic ideology has proven difficult to hardly find a solution with. The most common reason as been surveyed lies in the fact that users tend to ignore warnings from anti-phishing tools. The spear phishing message tends to be tough to differentiate. It is well learned from the cases that the same message is sent in bulk to indicate it as a spam. Also, the contexts of these messages seem to be more convincing. If the phisher knows more information about the targeted user then the latter can be well differentiated into spear phishing attack. Such a message seems to be novel and unseen outside. Also, if this is the case, the message has less chances of being detected as a spam by any anti-spam detectors. The recent categories of attacks which are less in defence include targeted malware attacks and session hijack attacks. They induce through phishing mails. These may be fortified if the former spear phish is left undetected. Hence, in broader terms a spear phish attack can be seen as a kick off to the latter ones described shortly.

Manuscript received February 5, 2026
Manuscript revised February 20, 2026
https://doi.org/10.22937/IJCSNS.2026.26.2.11

2. Motivation

All internet using organizations have firewalls as a mandatory security measures utilized with. It is an essential component designed to inspect the security issues like packet inspection, data authentication, integrity and confidentiality of the internal network. But firewalls suffer from major disadvantages like it cannot prevent users or attackers dialling into or out of the intranet work. It cannot insist of password protection or restrict for misuse of passwords. In short firewalls are poor tools which can screw spear phishing when the attacker resides inside the organization and thus cannot monitor poor decisions rooting the victim to such an attack.

Most systems have security holes for patch updating yet spear phishing being a semantic attack cannot be restricted. The principal information most likely the username and passwords have proven to be back doors for creeping into security holes when these have been learnt by the spearphishers. A spear phishing attack can be as good as a knock towards intrusion in the system. Moreover, it has been learnt from the survey that phishing attacks can be controlled through security soft ware's, transport layer security updates, anti-viruses and patches for boosting security in operating systems but spear phishing which is a kind of known party attack cannot be controlled using them. A wide group of systems implant a honey spot for curbing phishing attempts but intra-network spear phishing has no solution in such measures. Hence, after learning that spear phishing is a non-generic ideology, merely depending on goal-specific honey spot to curb it can never be a good solution.

3. Literature Survey

The APWG keeps a record on various strategies targeted for phishing the sites. A recent report highlights website intrinsic phishing activities. It states that a website may have thousands of URLs targeting a brand per domain. An email is sent to multiple users which direct them to a specific phishing website. This website may have plenty of URLs all directed to a common attack destination. A negotiable remedy to such strategy can be complete browser blocking or sometimes email blocking, both being unsuccessful. The difficulty complying it is requirement of full URL to block. Though plenty of researches to thwart phishing had been changing the scenario for phishers, still it has not yet fully put to an end. APWG reported a rogue steering in the on cast of phishing activity strengthening brand/domain pairs. URLs in big amount are being hosted to target a specific brand reporting to one more type in the taxonomy of phishing pronounced as spear phishing scoring a success count of 91% round the web. The phishers primarily target the social networking sites addressing to the personal information available at such places can be trivially utilized for identity based thefts. In some cases the evidences with APWG depict the deployment of worms for spear phishing to sneak-in the login information. Another approach to gain control of personal information includes data sharing which usually seems to be legitimated and trustworthy yet risky for the victim. The phishers obfuscate the victim through usage of altered links to direct him to fake page for stealing details. One of the examples in APWG showed that attackers broke into TD Ameritrade's database constituting of about 6.3 million customers' personal details and launched a spear phishing attack for gaining usernames and passwords [1-5]. Many of the known spear phishing attacks used fake phishing websites to gain control of web pages for directing to phishing page. One more strategy in spear phishing includes targeting business websites using zip archives and tar files. These seem to contain the user correspondence but after accessing them, it scrambles and locks files on the computer and targets the general public using a link in the email. This link then compels the owner to make payments for unlocking and decrypting the former files. This report is notified in the phishing activity trends report, second quarter 2014.

3.1. Knowing Spear phishing

It is a type of phishing attack wherein the intruder targets a single organization or company for stealing personal information. It is usually an email generated from individuals who are familiar with the victim. Unlike phishing attacks which broadcast the phished pages, spear phishing hones on specific organization. A plain phishing attack can be defined as an artificial trap set by the cyber criminals who use legitimate mails to obfuscate the victim to divert to fake web tools for breaching personal information. Spear phishing differs in the case that the intruder seems to be the one from the victim's known people list. He may be the one from internal organization relationships or may be from external links. To launch a spear phishing attack, the intruder uses personal characteristics of the victim to steal data. Sometimes, spear phisher targets a cluster of organization/people with the same characteristics previously analysed to target them. In such cases, the spear phishing mails may be apparently sent from these phishers within the organizations itself making them more deceptive.

3.2 How is Spear phishing Targeted?

The cybercriminals try to launch crime wares which spawn through the victims' personal web links in order to steal users' credentials. The crime wares configure key loggers which have in-built patches for information and pattern tracking. These key loggers monitor specific actions to capture certain information. The correspondence component, most likely the email seems to arrive from the most trustworthy source which the victim is frequently in contact with. The email depicts urgency and legally requests for personal information. It may also inject some worm wares which hijack the mandatory control points of the users' computer by the time the victim is busy in interacting with the spear phished page. The leniency of being unmindful because the email appears from someone more urging makes the victim implicated into spear phishing attack. The victims in such cases are less vigilant [4-7].

3.3 Hazards of Spear phishing

Being a victim of spear phishing, the original identity of the victim is masqueraded to lead another spear phishing attack on the victims' group eventually breaking the complete chain of trust. The victims' computer may be hijacked creating a bottleneck for denial of service attacks if the computer is of special concern. Another kind of danger is the victim may be asked to pay in order to rescue the seized data. The most hazardous part recently inspected is that a spear phishing attack can configure a sniffer application exploiting the administrators' rights and can read or monitor data exchanges. Such activities have proven to hijacking range of computers in the victims' organization [6-10].

3.4 Diagnosing Spear phishing Attack

Spear phishing exhibits some characteristics because of which it becomes tractable. To ease this task, the spear phishing attack has been grouped into two categories namely internal attacks and external attacks. The act of phishing when the phisher belongs to the attacked organization is called internal attack. It may also be incurred through a botnet inside the organization. A botnet can capture the targeted computer very easily. It could inject malware to have a handshake on a set of operations involved in fooling the victim to capture its personal data. Example of this is the famous Anti-virus update message asking the user to provide details for making update in progress. Since initiated by a botnet, it will induce all the computers with such an update thus making a door to hijack services. It may then launch a demand in lieu of releasing the services. In such cases the botnet based spear phish attack is hazardous to curb. The internal spear phishing attack is easy to diagnose in such a case wherein the characteristic of an infected mail would be 'mail to one' because in big organizations email broadcasting is one of the means of many-to-many communication. In some cases it may be one-to-many communication. But in no sense, it can be a many-to-one/one-to-one communication. Hence learning this characteristic, it is treated as a symptom to detect internal spear-phishing attacks.

External spear-phishing is incurred from exposed computers or e-wares stepped-in through internet. Here, the firewalls can also be induced through masquerading nodes present in the outskirts of the firewalls. The external spear phishing is usually accomplished by inducing a zombie process through emails and controlling this zombie computer externally through botnet. The spear-phisher is indirectly involved in implementing such an attack. The external spear-phish attacks may also be auto-driven. After addressing these issues, the characteristic of external spear-phish attacks can be judged that these are 'compulsive in behaviour'. It may also reflect good frequency count in the mailbox until and unless it is processed. Hence, external spear-phishing attack can be diagnosed through symptoms like 'frequency' and 'impulsiveness to accesses'.

3.5 Analyzing Spear Phishing Attack

The time duration between the sender and the receiver is considered the major factor in deciding vulnerabilities of the spear phishing attack. The compromise can captured when the payload of the email is delivered and is executed. This proves the indication of chances in being spear phished. Also, the ratio of emails left into the inbox undetected as spear phish from past history gives a likelihood in detection of spear phishing. Analyzing in this direction lessens the problem of reporting benign emails being flagged as malicious. Picture1 illustrates a pinpointing spear phishing attack targeted after triggering malicious emails.

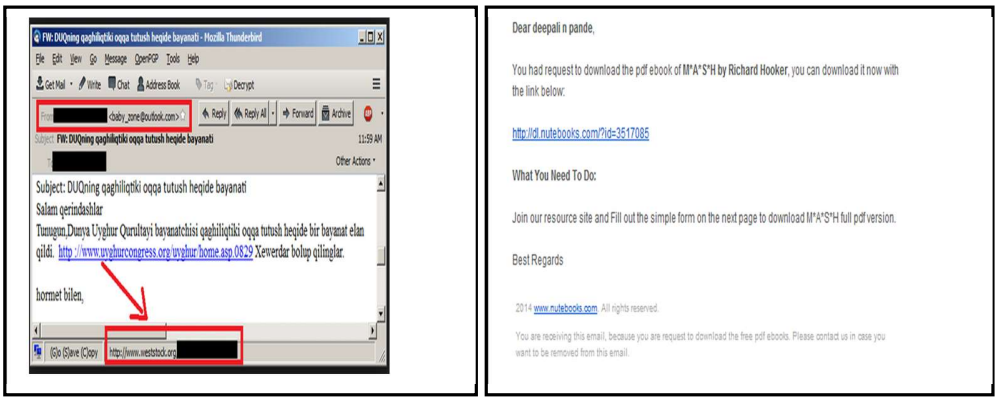

Figure 1 Snapshot of spear phish attempt redirecting to authentication demand link. The hyperlink traversed in through a book purchase email.

4. Methodology

To study the source of phishing, we laid an experimental set-up for detail examination of phishing attacks henceforth identifying spear-phishing from them. In doing so, we collected bag-of-features of DNS spoofing activity records. A dataset tryfreedo.myd was featured after pattern mining using existing machine learning algorithms. The process of pattern mining is described in Figure 1.We take into assessment, the behaviour patterns of various hyperlinks from malicious email sources. These patterns are formed after categorization of benign and malicious emails. We extract features from both, the database of malicious as well as benign records. To distinguish the benign records from malicious ones, we compute signatures from the threshold parameter as evaluated from the n-grams processing of pattern categorization phase. This is helpful in reducing the chances of true positives being unnecessarily identified as malicious one in the discrimination of spear phishing. A set of patterns based on behaviour analysis of malicious email detection is used in the judgement of spear phishing attack. The information gain based feature selection on behaviour patterns is used to design the classifier using the concept of naive bayes classification. N –grams processing yields information gain in one class and its absence in other class. The features are selected and byte sequences of n-grams are generated in differentiation model for benign and malicious mails. We take into consideration the sequences extracted from repetitive words, domain registry data, and system call records from program codes. These n-gram information gain nuggets are then used to generate signatures for benign records.

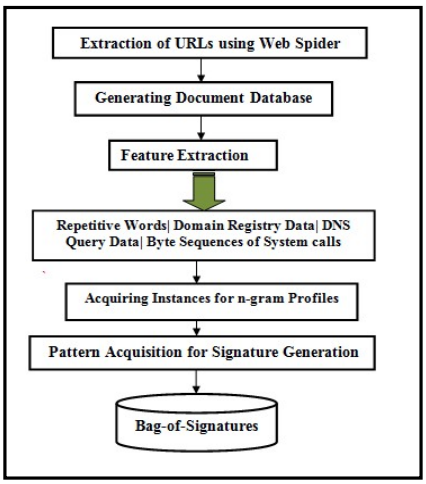

Figure 2 Pattern Mining for Feature Selection using machine learning methods

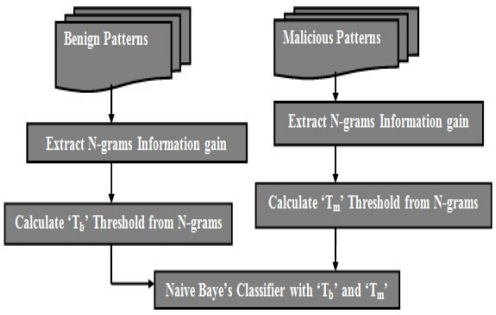

Figure 3 Spear phishing profiles Detection using Information Gain

5. Feature Mining

A spear-phish attack to target a dummy domain website, www.tryfreedo.com was designed for attack behaviour modelling. In the experimental set-up, studying various characteristics of phishing were the objectives to exemplify. During pre-attack scenario, the query section and the identifier of responses have a corresponding one-to one homomorphic match. Also, in the post attack scenario, during attack monitoring, we analyzed that the spearphishers typically originating from recursive Name Server exploitability, has a premature delivery response. We also experimented that the spear phish attack is easy to launch during the name resolution process because once and all the specification details are handed over to the DNS client, then to capture the dll files really becomes a bot-attack. Also, the spear phishing attack launching requires a malicious Name resolver which silently takes over the target system by modifying the 'resolve.conf' during DNS name resolution. We therefore, target the spear phish attack scenario as redirection of name server record to other domain to steal the authentication details, which is commonly known as DNS redirection.

6. Problem Statement

Hypothesis: We aim our attack detection hypothesis that the spy-bot installers through cross-site scripting cause intrusions for phishing. Hence lay an experimental set-up to detect the extrinsic porches causal for spear phishing.

We begin the spear phish attack analysis by first identifying whether query block is present in the transmission packet. If it is present, our target is to diagnose the type of query namely recursive query, iterative query or inverse query. If the type is 'recursive query' then we treat it as a source of spear phish attack and judge it using attack-parameter analysis. We describe various parameters for analyzing spear phish as a part of features of spear phish attack in the section described above.

A recursive query is mostly used for information collection during spear-phishing since the DNS server which receives this query has to fetch the answer and return back the needful information. During a typical browsing session, the 'resolve.conf' puts complaint DNS servers' addresses in it, which diverts all queries to these addresses. Only some of the DNS servers from the list in 'resolve.conf' may be recursive query based. The plausible attack paradigms used for spear phishing include session hijacking and DNS address replacement and is usually mounted by conquering the 'resolve.conf'. The victim node is fully unaware of these paradigms. A weaker point in spear phish attack mounting is that the root servers fully work on iterative query system. A major significance of iterative query system is that the DNS server sends a referral to another DNS server which may have an answer for the query. Exploiting this property, the early identification of spear phish attack launch is possible. The querying property of DNS name servers provided in the 'resolve.conf' can be monitored and shielded. In the experimental phase, we examined that if all the DNS servers use iterative query system then the chain of dependency in spear phishing is destroyed and henceforth causes early identification of attackers in the chain.

7. Experimental Setup for Attack Detection

We laid an experimental set-up of zabbix tool to find the causal vulnerabilities for spear phish attack. We configured zabbix as a network server to keep track of multiple instances and one as a client. Our target was to monitor phishing activities on agent-based and agent-less platforms. A website sniffer tool version 1.5 was used to monitor www.tryfreedo.com for an examination period of 99 days. The table I illustrates the calculation of window of opportunity for an intruder to launch spyware. We used DNS function exploit kits freely available for estimating the probability of a phishing attack by the intruder targeting externals ports as the source for attacks to enter.

Table I. DNS Activity monitoring using DNS Function Exploit kit.

Existing Tools	$N=\frac{\log(1-Q)}{\log(1-P)}$ Window of Opportunity	Probability of Success 'Q'	Number of Fake Responses 'F' (in thousands)	Probability of Guessing ID
Remote DNS Cache Poisoning Exploit BIND 9.5.0	47-52 ms	91%	40-50	43.44% - 15 minutes
DNS Cache Poisoning Exploit 9.0 BETA	24 – 57 ms	87%	18	67% - 30 minutes
Remote DNS Cache Poisoning Flaw Exploit	12-35ms	98%	37	76 %- 50 minutes

The workstation holding the tryfreedo.com was configured under SNMP as managed server agent to record the activity under MySQL database tryfreedo.myd. The records were analyzed using pattern mining described in figure 2. The PhishTank database and tryfreedo, where used to generate patterns for stylometric feature vectors on spams. As a part of pre-processing, the hyperlinks were parsed and URLs for images and links to other pages were retrieved from the tryfreedo.myd. The DNS query sniffer generated a record consisting of descriptions like hostname, port number, response time, duration and content into a dnsquery.csv file. The hostnames and port numbers were used as a testament for similarity assessment of the records obtained in the tryfreedo.myd. We used the formula described below to identify the scores of legitimate versus phished URLs on well known machine learning algorithms like bayesian classifier and nearest neighbour classifier.

$$\textbf{Scorematch} = \frac{\text{Number of Patterns matched}}{\text{Total Patterns Extracted}}$$

Doing this, a list of ports proving frequency of intrusions were recorded as a memoir from the DNS activity monitoring kits used in the experiment, as shown in the table II.

Table 2 List of memoirs collected in the experimental phase.

Service	Renderers
File Transfer Protocol-Port 21	Like HTTP Web Server Microsoft SQL Server Windows File Sharing Probe Gnutella peer to peer file sharing tool
Simple Mail Transfer Protocol Port 25	Like HTTP Web Server Microsoft SQL Server Windows File Sharing Probe Gnutella peer to peer file sharing tool
Domain Name System Service Port 53	Like HTTP Web Server Microsoft SQL Server Windows File Sharing Probe Gnutella peer to peer file sharing tool
Hypertext Transfer Protocol Port 80	Like HTTP Web Server Microsoft SQL Server Windows File Sharing Probe Gnutella peer to peer file sharing tool

8. Evaluation and Analysis

As, a part of experiment, the default internet explorer browser was installed. The spam filter properties were disabled for period of 99 days in tuning with the zabbix activity monitor. During the activity, we collected 570 phishing URLs in which mostly are diversion links, stegano-images, dll installers. The machine learning algorithms developed using java were tested using tryfreedo.myd and phishtank , available under APWG site. Table III, shows illustrations of phishing attachments and diversal links and executables found on a prominent frequency.

Table 3 List of top -7 most frequently occuring attachments. The attachment names appear as genuine names.

Spear-Phishing Attachment Source	Spear-Phishing Attachment Name
page-jjjj9.war	109-WX-AZ-PA4-2certified.exe
apr-123.jar	/var/directory0/attach_dominician.exe
urns.dll	111_dx_oooo_fpqw.html
img049897.scr	/attach/ui40064_2013_Article_296.exe
20uiu.rar	/ybhjournal.pone.0130968.t001.exe
scriptforyou.au3	/var/cdr/fvmgHZXfRz5fhytx-croppedCjBlH.exe
qrtyui.rar	/nygvxspl_1k9_stylesheet.exe

During the analysis, it was found that spear-phishing can be promisingly identified from link diversion attacks. If a link diversion and a self installer is posted at the victim's system, then authentication details can be easily conquered through cross-site scripting or code injection as an aftermath. We used a computer system with minimalist configuration core i3, 2.4 GHz processor versus a smartphone with iOS and mutli-client web-applictions for reverse analysis. The table IV, shows the results of computation of legitimate versus phished websites detection on well known classifiers namely the bayesian classifier and the nearest neighbour classifier

.

Table 4. Analysis of legitimate/phishing links

Datasets	Machine Learning Algorithms			
	Bayesian Classifier		Nearest Neighbour Classifier	
	% Match		% Match	
	Legitimate	Phished	Legitimate	Phished
Tryfreedo	82 TPR/8 FPR	92 TPR / 8 FPR	88 TPR / 5 FPR	93 TPR/ 4 FPR
Phishtank	84 TPR/ 4 FPR	96 TPR/7 FPR	81 TPR / 7 FPR	85 TPR / 6 FPR

c

The distribution of classification parameters was not uniform in tryfreedo. The discrimative attributes used for comparative analysis in similarity assessment of tryfreedo and phistank had varied values. The precision-recall exhibit informative results under variant scenarios.

.

Table 4. Accuracy of Feature Retrieval using Precision and Recall. The recall percentage shows significant retrieval as compared to benchmark PhishTank. Result show balanced AUC relationship.

Feature Attributes	URL length > 60	Redirection	Sub-Domain	Long domain >25	Requesting URLs	URL Anchor	Tag Link Abnormality	Age of Domain	DNS Record	Page Rank	Traffic	Multiple Links
PhishTank												
Precision (%)	97.66	94.12	97.14	88.22	87.89	91.17	89.23	81.15	91.10	93.21	95.37	95.16
Recall (%)	97.45	94.36	97.16	82.17	86.71	90.13	88.24	80.02	92.18	94.43	96.12	96.89
tryfreedo												
Precision (%)	94.01	90.55	99.01	89.76	82.45	99.89	78.90	77.98	90.17	74.16	91.11	92.22
Recall (%)	91.99	93.47	98.19	87.61	80.90	91.22	81.23	80.01	92.18	62.23	80.18	98.15

We installed web scanner software to analyze whether the attacks focus spear phishing or phishing attacks. Some of existing soft-wares like arachni web scanner was used to acquire features from the user-agent used at all the terminals in the complete experiment process. The arachni is complaint to distributed architecture, in that multiple clients can be remotely scanned.

A set of two smart phones with iOS and two Windows based 64 bit clients were treated with arachni webscan. The user agents iCab, dolphin, internet explorer and mozilla firefox were dynamically scanned and informative features like IP address, hostname, Proxy IP, referrer page, and whether java & dot net enabled were monitored and stored as binaries. Table VI illustrates the values retrieved. The properties Java-enabled and dot net enabled revealed open-porches for phishing.

Table 4 The user agents with properties like Java enabled & Dot Net enabled as true, revealed porches set open for phishing.

Features Selected from User-Agents	
IP Address	192.68.X.X ::192.68.X.X
Hostname	192.68.X.X ::192.68.X.X
Proxy IP	192.68.X.X ::192.68.X.X
User-Agent	iCab, Dolphin, Internet Explorer, Mozilla
Referrer Page	www.catchfreedo.com
Java enabled	TRUE
Dot Net enabled	TRUE
Number of Plug-ins	2

The runtime activity of the top-10 attachments were scrutinized using compliant application servers and web containers. We used eclipse-neon 64-bit level compatibility to assess the targeted activity of the pin-pointing attachments. The contents revealed usage of JOnAS with EJB container for typical attachments of web codes and extension as '.war', also Jetty with java servlet container revealed hidden activity. The executables exhibit deception of social-net links.

Figure 7 Information Gain Index of Top-9 Distinguished Features pinpointing the attack as spear phishing.

Pinpointing Features	Information Gain Ratio
has_redirection	0.041
age_of_domain	0.012
has_attachment	0.017
has_miscellaneous_plugins	0.087
has_posting	0.019
type_posting_is_dll	0.871
type_posting_is_war	0.912
type_posting_is_jar	0.964
type_posting_is_rar	0.926

setting feature learning association rules using pinpointing features describing type of attack as spearphishing. The top-9 informative features directing towards spearphishing detection with a hitting information gain index is illustrated in Table VII. We discriminate the spearphishing from phishing using association rule minning on the pinpointing attributes as described below.

Association Rule for SpearPhishing Detection

If *Redirection* = = **TRUE &&**
Age_of_Domain **> 75 &&**
has_attachment = = **TRUE &&**
has_miscellaneous_plugins **&&**
has_post **== TRUE &&** *type* = = *(is_posting_dll|is_posting_war|is_posting_rar)*
then *attack_type* = *Spearphish*
else if *Redirection* == **TRUE &&**
Age_of_Domain **< 75 &&**
has_attachment = = **TRUE &&**
has_post = = **FALSE**
then *attack_type* = *Phishing*

9. The Tryfreedo Dataset

The dataset acquired in manner described above, was focused to record total 11 attributes which converge towards phishing attacks. It has total 912 records from which 570 regarded cleanly as phishing. The tryfreedo.myd was compared for similarity with benchmark dataset phishtank freely available under UCI machine learning repository. The percentage of recall over precision show significant retrieval rate. The records evaluated as phishing were further used for discrimination of spearphishing. We computed information gain index on 570 dataframes after

The figure 3 highlights the density of information captured and refined for spearphish analysis. The machine learning tasks was evaluated using Java on eclipse neon purely using algorithmic approach. We also used Ri386 for dataframe calibration and various data preprocessing tasks mining inbuilt packages. We illustrate our pseudocode used for categorization of spear phishing attacks from pharming attacks and spamming attacks in snippet 1. The hypothesis that a spear phish can be launched only through intrusion, is proven and analysed in experiment. We demonstrate the clarity of the proposed hypothesis in table VII showing inclination on features attributes pinpointing the targetted attack.

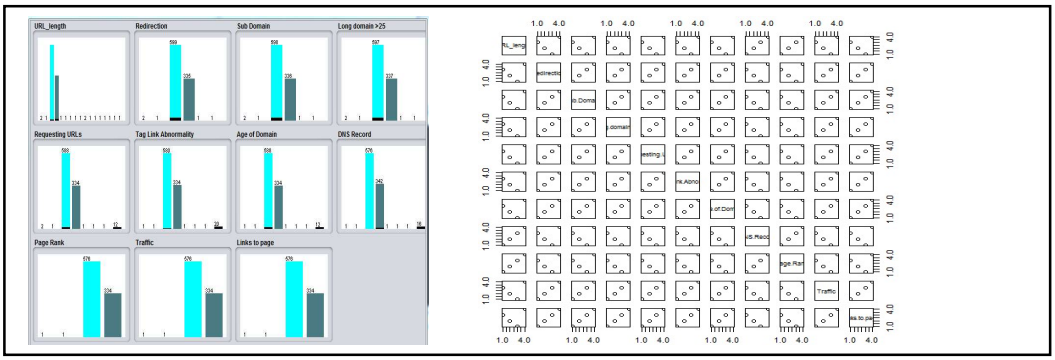

Figure 3 Density Graph of Tryfreedo Dataset

We used different combinations of cross-fold validation on tryfreedo under same testing conditions. On a generic score of 10 fold cross validation, different variants of classifer algorithms bayesian classifier, bayesNet, OneR, ZeroR, RandomForest, j48 were used to compare the results of classification. The results of classification on tryfreedo in comparison of PhishTank show an initial score of ROC with 0.9881 and 0.996. The ROC curve plots, shown in figure 4, specify accuracy of detection.

10. Discussion and Conclusion

In the experimental evaluation we used DNS function exploiters on 3-way server-client; client-client; server-server configuration to identify and evaluate the source of spear-phish attacks. We developed a website www.tryfreedo.com to record the scrupulous activity as the agent. We used, two smatphones with webapps like facebook , linkedin, pinterest,instagram,twitter and shareit to send intrusion activities on tryfreedo by flaunting the website www.catchfreedo.com on bare internet explorer as a browser providing it only a minimal set of security,with spam filters disabled permanently. Also, in the trial period, we disabled the hotlink protection and leech protection features. It was examined that the attachments, fetched in the experimental phase are floated on the internet through webomedias. These attachments have dll files, war files, and self-installable plug-ins which target any users through sniffing activity. We have listed some of these malicious attachments in the table-III.The outcome exhibits, smartapps can be used for link reversals from the iOS smartphones. We calculated the true positives and the false negatives on the records obtained in the process. The results show, on an average a 83% correct identifcation of legimate websites whereas a 94 % correct identification of phishing URLs. As a part of the future work, we are targetting to reduce the false negative rate in identification of phished URLs. Also, we aim to identify cross-site script link-reversals as a stylometric feature in the feature collection phase of our proposed system of spear-phish detection.

Acknowledgement

The survey work is carried out under the guidance of Dr. P.S. Voditel. I am thankful to her for directing me towards the research goal. I extend my sincere gratitude to Shri Ramdeobaba College of Engineering and Management, Nagpur for supporting the research work. I am also thankful to the Liaison progam for launching free webowares like arachni. Arachni-web-scanner is a freeware with good coverage of multi –operating system support and proved very useful during the experiment work.

References

[1] A. Martino and X. Perramon, “Phishing Secrets: History, Effects, and Countermeasure,” in International Journal of Network Security, 12(1), pp. 37–45, January 2011.

[2] F. Aloul, “The Need for Effective Information Security Awareness,” in Journal of Advances in Information Technology (JAIT), 3(3), pp. 176-183, 2012.

[3] M. Cova, C. Kruegel and G. Vigna, “There is No Free Phish: An Analysis of “Free” and Live Phishing Kits,” in Proc. of the 2nd USENIX Workshop on Offensive Technologies, 2008.

[4] R. Dhamija, J. Tygar and M. Hearst. “Why Phishing Works,” in Proc.of the Conference on Human Factors in Computing Systems (CHI), p. 581-590, 2006.

[5] Y. Zhang, S. Egelman, L. Cranor and J. Hong, “Phinding Phish: Evaluating Anti-Phishing Tools.” in Proc. of the 14th Annual Network & Distributed System Security Symposium (NDSS), February 2007.

[6] M. Wu, R. Miller and S. Garfinkel, “Do Security Toolbars Actually Prevent Phishing Attacks?” in Proc. of the

Conference on HumanComputer Interaction (CHI), New York, pp. 601-610, 2006.

[7] S. Egelman, L. Cranor and J. Hong, "You've Been Warned: An Empirical Study of the Effectiveness of Web Browser Phishing Warnings," in Proc. of the Conference on Human-Computer Interaction (CHI), Florence, Italy, pp. 1065-74, 2008.

[8] G. Ollmann, "The Pharming Guide: Understanding and Preventing DNS-related Attacks by Phishers," NGS Secure, 2005. Available at: www.infosecwriters.com/text_resources/pdf/ThePharmingGuide.pdf

[9] L. Shujun and R. Schmitz, "A novel anti-phishing framework based on honeypots," in Proc. of the eCrime Researchers Summit, pp.1-13,September 2009.

[10] S. Sheng, M. Holbrook, P. Kumaraguru, L. Cranor, and J. Downs, "Who Falls for Phish? A Demographic Analysis of Phishing Susceptibility and Effectiveness of Interventions," in Proc. of the Conference on HumanComputer Interaction (CHI), Atlanta, Georgia, 2010.

[11] T. Jagatic, N. Johnson, M. Jakobsson and F. Menczer, "Social Phishing," in Proc. of the Communications of ACM, 50(10), pp. 94-100, October 2007.

[12] R. Dodge, E. Rovira, R. Zachary, and S. Joseph, "Phishing Awareness Exercise," in Proc. of the 15th colloquium for Information Systems Security Education, Fairborn, Ohio, June 13-15, 2011.

[13] http://www.arachni-scanner.com/download/, Arachni Web-Scanner For Windows Systems.

[14] https://archive.ics.uci.edu/ml/datasets/Phishing+Websites, UCI Machine Learning Repository.