مجلة التربوي
Journal of Educational
ISSN: 2011- 421X
Arcif Q3

معامل التأثير العربي 1.63
العدد 22

مجلة التربوي

مجلة علمية محكمة تصدر عن

كلية التربية / الخمس

جامعة المرقب

العدد الثاني والعشرون

يناير 2023م

ضوابط النشر:

يشترط في البحوث العلمية المقدمة للنشر أن يراعى فيها ما يأتي :

-أصول البحث العلمي وقواعده .

-ألا تكون المادة العلمية قد سبق نشرها أو كانت جزءا من رسالة علمية .

-يرفق بالبحث تزكية لغوية وفق أنموذج معد .

-تعدل البحوث المقبولة وتصحح وفق ما يراه المحكمون .

-التزام الباحث بالضوابط التي وضعتها المجلة من عدد الصفحات ، ونوع الخط ورقمه ، والفترات الزمنية الممنوحة للتعديل ، وما يستجد من ضوابط تضعها المجلة مستقبلا .

تنبيهات :

-للمجلة الحق في تعديل البحث أو طلب تعديله أو رفضه .

-يخضع البحث في النشر لأولويات المجلة وسياستها .

-البحوث المنشورة تعبر عن وجهة نظر أصحابها ، ولا تعبر عن وجهة نظر المجلة .

Information for authors

1- Authors of the articles being accepted are required to respect the regulations and the rules of the scientific research.

2- The research articles or manuscripts should be original and have not been published previously. Materials that are currently being considered by another journal or are a part of scientific dissertation are requested not to be submitted.

3- The research articles should be approved by a linguistic reviewer.

4- All research articles in the journal undergo rigorous peer review based on initial editor screening.

5- All authors are requested to follow the regulations of publication in the template paper prepared by the editorial board of the journal.

Attention

1- The editor reserves the right to make any necessary changes in the papers, or request the author to do so, or reject the paper submitted.

2- The research articles undergo to the policy of the editorial board regarding the priority of publication.

3- The published articles represent only the authors' viewpoints.

مجلة التربوي
Journal of Educational
ISSN: 2011- 421X
Arcif Q3

معامل التأثير العربي 1.63
العدد 22

Performance Evaluation of Blacklist and Heuristic Methods in Phishing Emails Detection

Munayr Mohammed Amir[1], Melad Al-Daeef[2]

Faculty of Information Technology - Elmergib University[1,2]
mmameer@elmergib.edu.ly[1], mmaldaeef@elmergib.edu.ly[2]

Abstract- Phishing is a cybercrime in which, attackers try to fraudulently retrieve users' credentials by mimicking trusted communication channels. Phishing attacks are usually start with email means.Many methodsareimplemented to detect phishing emails. Blacklists and heuristics anti-phishing methods are commonly used to mitigate the impact of phishing crime.Previous studies revealed that each of these methods still has its drawbacks when implementedalone to detect phishing emails.In order to enhance the performance of these two methods, it is widely suggested by the researchers to combine these two methods to work as one anti-phishing system. Thus, if one method fails to detect the attack, the other method can catch it. In this study, blacklist and heuristics methods have produced an acceptable accuracy rate in phishing detection when they have cooperatively implemented, they have achieved up to 93% of TP accuracy rate.

Index Terms- phishing email, blacklist, heuristics, URL-based features

1. Introduction

The Internet nowdays has an obvious impact onhumans' life way. The internet has certainly brought a convenient lifestyle that has made people more dependent on it for a lot of work. This convenient lifestyle, however, has opened new avenues for cybercrime activities.Phishing is one of such crimes in which, the phishers try to steal users' credentials such as passwords and cridet card numbers.Phishing attacks are usually launched through simulated emails that falsely claim sent from trusted parties such as organizations or banks that the victim deals with. It is a useful countermeasure, therefore, to fight phishing attacks at the email level and kill phishing attacks in the cradle[1].Phishing is a cybercrime in which, attackers try to acquire sensitive information by impersonating a legitimate entity, through the use of electronic communications.Many reports show the increasing number in phishing attacks. In the first quarter of 2022, Anti-Phishing Working Group APWG[2] observed 1,025,968 total phishing attacks. In the second quarter of2022, APWG observed 1,097,811 total phishing attacks, a new record and the worst quarter for phishingthat APWG has ever observed. Another report by APWG [3] show that the number of recent phishing attacks has more than doubled since early 2020, while the APWG has observed a number between 68,000 and 94,000 attacks per month.It is a common scenario when phishing emails contain fake URLs to deliver the victims to phishing websites [4]. This study, therefore, evaluate the performance of two most commonlyimplemented anti-phishing methods, they are, blacklist and URL-based heuristics methods.

1.1 Types of Phishing Attacks

Phishing websites are usually designed to be identical to the original ones; they falsely claim being legitimate with the aim of deceiving both of the search engines and Internet users. This type of websites includes; spam, concocted, and spoof sites. *Spam sites* are designed to deceive search engines to increase their rank scores. *Concocted sites* are

deceptive sites that appear as legitimate commercial ones with the aim of failure-to-ship fraud; they just disappear after collecting customers' money without providing the agreed-upon goods or services[5]. They commonly presented to the victims as real escrow, financial, delivery, retail, and payment services. *Spoof sites* are copying of the real commercial websites that designed to deceivethe users disclose their credentials such as passwords, credit card numbers and so on. eBay, PayPal, and various banking are common examples of spoofed websites[5].

Phishing attacks are generally fall in two categories; social engineering and malware-based attacks[6]. Attackers in the social engineering phishing base usually try to control the victims' accounts by sending them simulated emails with fake URLs that deliver to phishing websites. Social engineering-base attacks, also known deceptive phishing, is further classified into email-based and website-based phishing. Malware-based phishing on the other side uses a variety of malicious programs that run on the victims' machines. This type of phishing is further classified as; keyloggers/screen loggers, session hijacking, host file poisoning, DNS phishing and content injection[7].

1.2 Blacklist and Heuristics Anti-Phishing Countermeasures

One of popular methods to combat phishing attacks include blacklists[5],in which, the suspicious URLs, phishing email senders, IP addresses or keywords are recorded.The content of blacklists is periodically updated, and then,is utilized by phishing email and website detection systems to block the threat sources.Phishing blacklists are usually obtained either by user feedback or from the reports by the third parties who perform phishing URL detection. Although their accurate results in detecting phishing instances, blacklists, however, cannot detect the fresh or zero day phishing instances because of the update time lag of the lists' content[8].

Another method which is widely used to detect phishing attacks is the heuristics method. This method is used to check emails' or websites' characteristics that include, URLs, HTML code, or page content to determine whether they pose a threat or not[9]. The heuristics based Anti-phishing systems more efficient than list-based systems in detecting fresh phishing instances [5]. In this research, the characteristics of URLs that extracted from email content are used to examine the email, and therefore, identify it as either phishing or legitimate one.

1.3 URL-Based Phishing Detection

Phishing detection by analyzing email's content is a useful countermeasure since simulated emails are usually used to launch most of the phishing attacks by hiding the fake URLs in the contents of such emails. By using emails, phishers can easily reach a huge number of victims and show them fake URLs that take to phishing websites[4][10]. Email filters -heuristics based methods- are widely used to prevent phishing emails from reaching users' inboxes. Numerous of such filters have been implemented to classify emails based on the natural language cues and the keywords in their contents. Researchers in this field have tried to improve the performance of phishing email filters by the analysisof URLs in email's contents to detect URL-based phishing indications. Such indications include, but not limited to; the number of dots in a given URL, the number of special characters, the presence of hexadecimal characters or IP addresses instead of the domain name, and URL length [11][12].

LinkGuard is a URL-based phishing email detection approach was proposed by Chen and Guo in 2006. The URLs based on this approach, are extracted from emails' contents and classified into five categories to check the following criteria; *a)*If there any mismatch between the destination DNS name in the visible link and that in the actual link. *b)*If the dotted decimal IP address was used instead of the DMN name.*c)*If the URL's alphabets are encoded and formed into their corresponding ASCII codes and/or any special characters such as @ character. *d)*If the destination information in the anchor does not contain a hyperlink. *e)*If the URL redirects to phishing website by utilizing the hosting domain vulnerabilities. In 2007,researchers in [13] have identified some fine-grained heuristics to distinguish between legitimate and phishing URLs,theyhave achieved a 97.3% accuracy rate. In 2007, the PILFER algorithm was developed in [14]. to identify phishing emails by implementing a number of 10 classification features. PILFER algorithm employed the following features; IP-based URLs, the age of the domain, non-matching ULRs between the hyperlink and anchor tag (the visual and actual links), HTML and JavaScript presence, the number of links and domains, and periods number in examined URL.A number of 7,810 emails were used to evaluate the PILFER algorithm, results show 96% of identified phishing emails.In 2009, researchers in [15] have proposed the PhishCatch anti-phishing tool. In which, they have used a set of filters and weighted rules that include; the length of hyperlinks, the differences in the Received From, and From fields of the email, and the differences between hyperlinks and anchor tags. PhishCatch results show 80% of detection. To classify emails, researchers in [16], 2010, have used the confidence-weighted model that trained on features derived exclusively from the URLs. Their approach had achieved higher than 97% of detection accuracy on new phishing URLs when the model has continuously trained.The lexical URL analysis approach to identify phishing emails was used in [17], 2011,their approach is based on the fact that most of phishing emails contain fake URLs, thus, the lexical analyzing of these URLs can achieve highdetection accuracy. Another study in 2012 by the same authors in [17]show the advantages of using the lexical URL analysis technique in phishing detection process.

Therefore, due to their promising reliability, URL-based features are usedin this study to evaluate the performance of heuristic-based method in detecting phishing emails. A new URL-based feature namely the FldrNameLengthwhich proposed in our previous study is used here also.

One limitation of email filtering technique, however, is the social engineering approaches [18][19]. In spear phishing, for example, phishing emails usually contain recipients' personal information that mined from the web. If the victims see their names and affiliations in the received emails, they well just trust them and may be easily tricked [20]. Email filtering approach, also, cannot be applied to detect phishing websites that have not advertised through emails.

Because of the phishing detection limitations that associated with blacklist method and heuristic-based methods, and to enhance the detection accuracy rate, in this research, the two methods are cooperatively implemented to detect phishing emails, and their detection accuracy is evaluated on the cooperation base. Two types of datasets are used in the performance evaluationexperiment, they are; legitimate emails datasetand phishing emails dataset. The legitimate emails dataset comprises of 10000 emails collected by Shetty & Adibi[21], Cormack & Lynam[22] and Spamassassin public corpus[23].The phishing emailsdataset comprises of 2916 phishing emails collected by Nazario[24].

2. Blacklist Anti-Phishing Method Implementation

To determine whether the evaluated email is phishing or legitimate instance, all URLs in the checked email are examined against the blacklist of previously known phishing URLs that obtained from PhishTank database. A given URL is considered as phishing if it is matching any of blacklisted URLs. The email from which this URL was extracted is, therefore, identified as a phishing instance. Figure1 shows the operations' flowchart of the blacklist checking process. Initially,the received email is considered as a legitimate one, thus, itsphishing email status=False.The system calls the blacklist method to check all URLs that extracted from this email's contentagainst the content of PhishTank database.

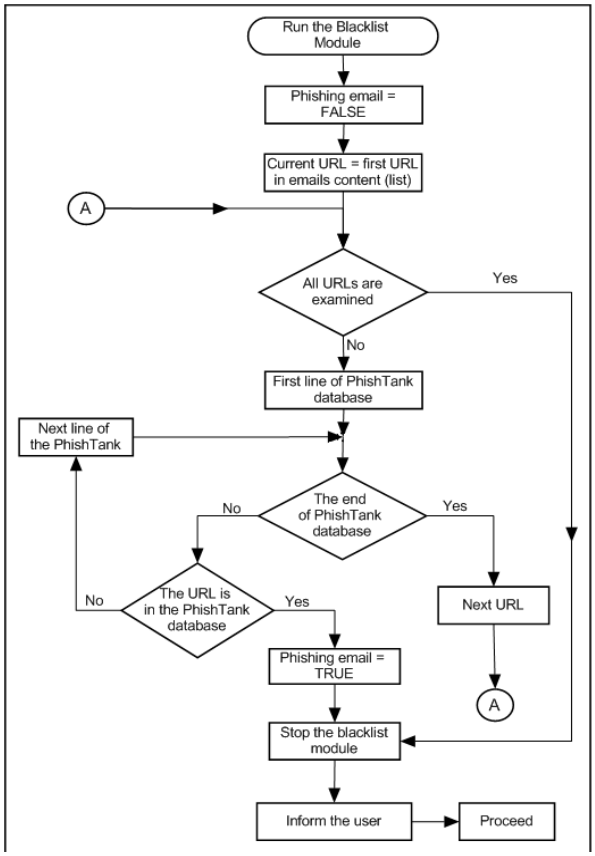

Figure 1: Flowchart Operations of the Blacklist Method

If anychecked URL found matching any of blacklisted URLs, the email from which this URL was extracted is marked as a phishing instance. All URLs in email are extracted using the Regexps technique.Table 1presents some examples of applied Regexps that have testedbefore implement them in our experiment. The Agent Ransack from Mythicsoft and the EditPad Pro Pad 7 tools that support the use of Regexpsareutilized to test all of implemented Regexps. These free tools can quickly and efficiently search files' contents and extract required information based on predefined searching patterns.

All emails in the datasets are one by one examined and overall result are used to evaluate the performance of the blacklist method in labeling the emails as either phishing or legitimate. If an email has not labeled -URLs in its content not found in the blacklist content-, the email is then rechecked using the heuristics anti-phishing method.

Table 1: Examples of Applied Regexps to Extract URLs from Emails' Contents

Regexp Pattern	Description
(http://|https://|www).*(|[a-z]{2,5})	To extract URLs in their general form
(http://|https://|www).+?[?]	To extract URLs ends with ? (end of the path)
(http://|https://|www).+?("|\s)	To extract remaining URLs that do not contain the ? mark

3. Heuristics Anti-Phishing Method Implementation

As in the previous method, all emailsareexamined to determine whether theyare phishing or legitimate instances. Thismethod is implementedto examinea given email by utilizing 12 URL-based features that extracted from email's content.A given URL is identified as phishing if its characteristics positively meet one of the employed 12 URL-based features that presented in Table 2.If a given URL identified asa phishing, the email from which this URL was extractedis, therefore, labeled asa phishing instance.

Table 2:URL-Based Features that used by the Heuristic Anti-Phishing Method

No.	URL-based Feature	Description
1	FldrNameLength	URL with a folder or sub-folder longerthan 30 characters is labelled as a phishing instance.
2	KeyWordURL	URL includesa suspicious Keyword is labelled as a phishing instance.
3	IP Address	URL contains an IP address is labelled as a phishing instance.
4	DMNDashes&Dots	DMN includes more than 4 dots and/or dashesis labelled as a phishing instance.
5	ImgHttps	URL contains anphishingimg src=https:// addressis labelled as a phishing instance.
6	DMN Semantics (DMN Naming)	URL contains unwanted character is labelled as a phishing instance.
7	Non-Standard Port Number	URL contains non-standard port number is labelled as a phishing instance.
8	MoreThanOneDomainURL	URL contains morethan one DMN is labelled as a phishing instance.
9	onMouseOver	URL uses onMouseOver property is labelled as a phishing instance.
10	URL-HEXCoding	URL's TLD usesURL-HEXCodingis labelled as a phishing instance.
11	Form Tag	URL contains input form is labelled as a phishing instance.
12	using of "@" character	URL has @ symbolis labelled as a phishing instance.

As the Figure2 shows, the suspicion status of the checked email is initially set to False since this email was not identified as a phishing instance by the blacklist checking method. URLs' examination process starts with the first URL extracted from email's content. Operations of the heuristic-based methodare ended when all URLs are examined. The currently checked URL is identified asa suspiciousor phishing instance if it meets one or more of the 12 features that listed in table 2.

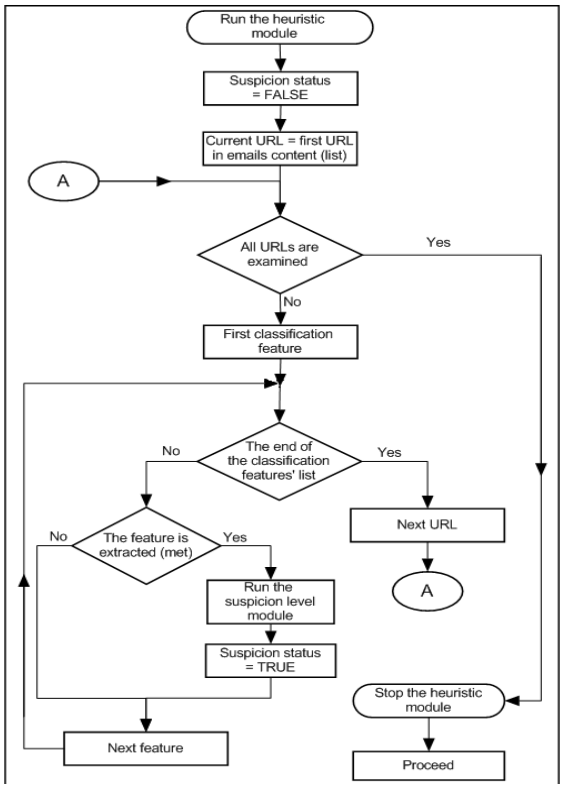

Figure2: Flowchart Operations of the Heuristic Method

4. Experiment of Blacklist and Heuristic Based Methods

This section presents the performance of blacklist and heuristic-basedmethods that evaluated based on the produced TP and FP results.Table 3 shows the implementation of confusion matrix method to calculate the performance metrics of these two methods. Before discussing the results, we identify what TP, FP, FN and TN stand for.

TP stands for True Positive (phishing instance is correctly identified as phishing).
TN stands for True Negative (legitimate instance is correctly identified as legitimate).
FP stands for False Positive (legitimate instance is incorrectly identified as phishing).
FN stands for False Negative (phishing instance is incorrectly identified as legitimate).

Table 3 shows that 2916 emails are considered as an actual positive (phishing) since they come from phishing email dataset. It also shows that 10000 emails are considered as an actual negative (not phishing) since they come from legitimate email dataset.The implementation experiment of using the blacklist and heuristic to label all emails in the two datasets has producedthe TP, FP, FN, and TN resultsthat presented in Table 3. Appendix.1shows a sample results of phishing email checking experiment by using the blacklist and heuristic methods.Produced results have shown a 93% of TP and 12% FP rates. This high FP result still acceptable sincea huge size of 10000 legitimate emails dataset were analyzed in the checking experiment. We must consider the fact that, all datasets contain noisy data which undoubtedly will affect the detection accuracy rate. This was the case with such big number of analyzed emails.

Table3: Confusion Matrix of the Heuristic and Blacklist Methods Evaluation

Total = 12916	**Identified Positive (P)**	**Identified Negative (N)**
Actual Positive (P) = 2916	(TP) 2716	(FN) 200
Actual Negative (N) = 10000	(FP) 1238	(TN) 8762

Results in the above confusion matrix show that:
The total number of checked emails = 12916.
The number of actual positive emails = 2916.
The number of actual negative emails = 10000.
The number of emails that identified as positive = TP + FP = 2716 + 1238 = 3954.
The number of emails that identified as negative = TN + FN = 8762 + 200 = 8962.
Based on the checking process, the heuristic and blacklist-basedmethods have together produced the following results:

$$\text{TP Rate} = \frac{TP}{P} = \frac{TP}{TP+FN} = \frac{2716}{2716+200} = \frac{2716}{2916} = 0.93.$$

$$\text{FP Rate} = \frac{FP}{N} = \frac{FP}{FP+TN} = \frac{1238}{1238+8762} = \frac{1238}{10000} = 0.12.$$

$$\text{TN Rate} = \frac{TN}{TN+FP} = \frac{8762}{8762+1238} = \frac{8762}{10000} = 0.88.$$

$$\text{FN Rate} = \frac{FN}{FN+TP} = \frac{200}{200+2716} = \frac{200}{2916} = 0.07.$$

5. Discussion

Produced results show the reliability ofthe blacklist and URL-based features -heuristic-methodsto identify legitimate and phishing emails. Some of anti-phishing tools that use the blacklist and heuristic methods such as the Microsoft IE Phishing Filter is basically utilize the blacklist method, it also utilizesthe heuristics method when encountering a site that not on its blacklist, its detection accuracy rate is up to 92%. The Hybrid client-side tool in [25] employs the URL-based features,itachieved a 87.5% of accuracy rate. PhishNet by Prakash [26]is another example of anti-phishing toolbars; itgenerates many of child URLs based on a given blacklisted URL, it achieved a 92% of detection accuracy rate. Many of the generated URLs, however, might be either innocent or not exist. Compared with the results from other tools, the experimen in this study has achieved up to 93% of accuracy rate.

6. Conclusion and Future Work

The blacklist is a widely used method to combat against phishing attacks. One of most known disadvantages of this method is that, its ineffeciency in detecting fresh phishing instances because of time lag between launching time of phishing URL and the time of this new phishing URL being blacklisted. Because of that, it is a common case where the performance of this methods is improved is to implement another anti-phishing method besides the blacklist method for the purpose of further verification if the blacklist method does not detect the phishing attack. The performance of the heuristics method was evaluated when it cooperatively implemented with the blacklist method. 12 URL-based heuristics were utilized in this studyfor the purpose performance evaluation.Experimental results show that the accuracy of phishing detection is within an acceptable scope.

For a future work, different datasets of phishing and legitimate can be used in the performance evaluation process for these two anti-phishing methods. Another evaluation attempt carried out by maintaing a blacklist athor than the blacklist from PhishTank database which was utilized in this study. Different URL-based heuristics set rather than the 12 URLs that implemented in this study can be utilized in further studies to evaluate the performance of the heuristic-based anti-phishing method.

References

[1] M. M. Al-Daeef, N. Basir, and M. M. Saudi, "Hybrid model of phishing email detection: A combination of technical and non-technical anti-phishing approaches," in *Lecture Notes in Engineering and Computer Science*, 2017, vol. 1.

[2] P. E. Reports, P. S. Trends, B. P. Measurement, E. P. Attacks, M. Targeted, and I. Sectors, "2 Quarter," no. September, 2022.

[3] Anti-Phishing Working Group, "Phishing Activity Trends Report 3rd Quarter," no. November, pp. 1–9, 2021.

[4] W. D. Yu, S. Nargundkar, and N. Tiruthani, "A phishing vulnerability analysis of web based systems," *Proc. - IEEE Symp. Comput. Commun.*, pp. 326–331, 2008, doi: 10.1109/ISCC.2008.4625681.

[5] A. Abbasi and C. Hsinchun, "A Comparison of Tools for Detecting Fake Websites," *IEEE*, pp. 47–54, 2009.

[6] M. Jakobsson and S. Myers, "Phishing and Countermeasures: Understanding the Increasing Problem of Electronic Identity Theft," *John Wiley Sons*, 2006.

[7] B. B. Gupta, A. Tewari, A. K. Jain, and D. P. Agrawal, "Fighting against phishing attacks: state of the art and future challenges," *Neural Comput. Appl.*, vol. 28, no. 12, pp. 3629–3654, 2017, doi: 10.1007/s00521-016-2275-y.

[8] H. Z. Zeydan, A. Selamat, and M. Salleh, "Current state of anti-phishing approaches and revealing competencies," *J. Theor. Appl. Inf. Technol.*, vol. 70, no. 3, pp. 507–515, 2014.

[9] S. Sarika and V. Paul, "Distributed Software agents for antiphishing," *Int. J. Comput. Sci. Issues*, vol. 10, no. 3, pp. 125–130, 2013.

[10] H. S. Rao and S. K. A. Nabi, "a Novel Approach for Predicting Phishing Websites Using the Mapreduce Framework," vol. 3, no. 10, pp. 505–510, 2014.

[11] G. Xiang, J. Hong, C. P. Rose, and L. Cranor, "Cantina+: A feature-rich machine learning framework for detecting phishing web sites," *ACM Trans. Inf. Syst. Secur.*, vol. 14, no. 2, p. 21, 2011.

[12] X. Gu, H. Wang, and T. Ni, "An Efficient Approach to Detecting Phishing Web ⋆," vol. 14, no. 61070121, pp. 5553–5560, 2013, doi: 10.12733/jcis6350.

[13] S. Garera, N. Provos, M. Chew, and A. D. Rubin, "A framework for detection and measurement of phishing attacks," *Proc. 2007 ACM Work. Recurr. malcode - WORM '07*, p. 1, 2007, doi: 10.1145/1314389.1314391.

[14] I. Fette, N. Sadeh, and A. Tomasic, "Learning to detect phishing emails," *Proc. 16th Int. Conf. World Wide Web - WWW '07*, p. 649, 2007, doi: 10.1145/1242572.1242660.

[15] W. D. Yu, S. Nargundkar, and N. Tiruthani, "PhishCatch-A phishing detection tool," *Proc. - Int. Comput. Softw. Appl. Conf.*, vol. 2, pp. 451–456, 2009, doi: 10.1109/COMPSAC.2009.175.

[16] A. Blum, B. Wardman, T. Solorio, and G. Warner, "Lexical feature based phishing URL detection using online learning," *Proc. 3rd ACM Work. Artif. Intell. Secur. -*

AISec '10, p. 54, 2010, doi: 10.1145/1866423.1866434.

[17] M. Khonji, Y. Iraqi, and A. Jones, "Lexical URL analysis for discriminating phishing and legitimate e-mail messages," *Internet Technol. Secur. …*, no. December, pp. 11–14, 2011, [Online]. Available: http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=6148476

[18] B. T. N. Jagatic, N. A. Johnson, and M. Jakobsson, "Social Phishing," vol. 50, no. 10, 2007.

[19] T. Ronda, S. Saroiu, and A. Wolman, "iTrustPage: A user-assisted anti-phishing tool," *ACM SIGOPS Oper. Syst. Rev.*, vol. 42, no. 4, pp. 261–272, 2008, doi: 10.1145/1352592.1352620.

[20] Trend Micro, "Spear-phishing email : most favored APT attack bait, Trend Mciro incorporated research paper," *Ressearch Pap.*, pp. 1–8, 2012, [Online]. Available: http://www.trendmicro.co.uk/cloud-content/us/pdfs/security-intelligence/white-papers/wp-spear-phishing-email-most-favored-apt-attack-bait.pdf

[21] J. Shetty and J. Adibi, "The Enron email dataset database schema and brief statistical report," *Distribution*. 2004. [Online]. Available: http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.91.3560&rep=rep1&type=pdf

[22] G. Cormack and T. Lynam, "TREC 2005 Spam Track Overview," *Sixteenth Text Retrieval Conference (TREC 2007)*, no. Trec. pp. 1–9, 2005. [Online]. Available: http://trec.nist.gov//pubs/trec14/papers/SPAM.OVERVIEW.pdf

[23] Spamassassin, "public corpus, http://spamassassin.apache.org/publiccorpus." 2006.

[24] J. Nazario, "Phishing Corpus." p. 0, 2007. [Online]. Available: http%5Cn//monkey.org/~jose/wiki/doku.php?id=phishingcorpus

[25] F. Kausar, B. Al-Otaibi, A. Al-Qadi, and N. Al-Dossari, "Hybrid Client Side Phishing Websites Detection Approach," *IJACSA) Int. J. Adv. Comput. Sci. Appl.*, vol. 5, no. 7, pp. 132–140, 2014, [Online]. Available: www.ijacsa.thesai.org

[26] P. Prakash, M. Kumar, R. Rao Kompella, and M. Gupta, "PhishNet: Predictive blacklisting to detect phishing attacks," *Proc. - IEEE INFOCOM*, 2010, doi: 10.1109/INFCOM.2010.5462216.

Appendix.1 Result sample of implementing the blacklist and heuristic methods

Results in this section means that, an email is classified as phishing based on either one or more classification feature, or blacklist checking result. The first column of the table below is dedicated to checked emails' paths and names, the second column is dedicated to the result from all URL-based classification features and blacklist checking method, and other columns are dedicated to all employed classification features. If the result was 1 in the second column, the email is classified as phishing, it classified as a legitimate otherwise.

Checked Email path and Name	Result by all URLs Based Features and Blacklist Check	MoreThanOneDomainURL	URL_IP	URL-HEXCoding	URLKeyWord	OnMouseOver	Imghttps	PortNumber	FormTag	@Character	DMNSemantics	DMNDashes&Dots	FldrNameLength	Blacklist Checking Result
E:\datasets\phishingemails\phishing (2610).txt	1				1			1				1		
E:\datasets\phishingemails\phishing (2611).txt	1		1										1	
E:\datasets\phishingemails\phishing (2612).txt	1				1			1				1		
E:\datasets\phishingemails\phishing (2613).txt	0													
E:\datasets\phishingemails\phishing (2614).txt	1		1											
E:\datasets\phishingemails\phishing (2615).txt	1				1		1					1		
E:\datasets\phishingemails\phishing (2616).txt	1											1		
E:\datasets\phishingemails\phishing (2617).txt	1		1		1			1						
E:\datasets\phishingemails\phishing (2618).txt	1												1	
E:\datasets\phishingemails\phishing (2619).txt	1		1		1								1	
E:\datasets\phishingemails\phishing (2620).txt	1				1		1							
E:\datasets\phishingemails\phishing (2621).txt	1		1		1									
E:\datasets\phishingemails\phishing (2622).txt	1											1		
E:\datasets\phishingemails\phishing (2623).txt	1				1									
E:\datasets\phishingemails\phishing (2624).txt	1		1		1									
E:\datasets\phishingemails\phishing (2625).txt	1				1									
E:\datasets\phishingemails\phish	1		1		1									

ing (2626).txt														
E:\datasets\phishingemails\phishing (2627).txt	0													
E:\datasets\phishingemails\phishing (2628).txt	1				1									
E:\datasets\phishingemails\phishing (2629).txt	1		1											
E:\datasets\phishingemails\phishing (2630).txt	1				1							1		
E:\datasets\phishingemails\phishing (2631).txt	1					1						1		
E:\datasets\phishingemails\phishing (2632).txt	1													1
E:\datasets\phishingemails\phishing (2633).txt	1											1		
E:\datasets\phishingemails\phishing (2634).txt	1		1		1									
E:\datasets\phishingemails\phishing (2635).txt	1	1			1		1							
E:\datasets\phishingemails\phishing (2636).txt	1													1
E:\datasets\phishingemails\phishing (2637).txt	1											1		
E:\datasets\phishingemails\phishing (2638).txt	1				1									
E:\datasets\phishingemails\phishing (2639).txt	1													1
E:\datasets\phishingemails\phishing (2640).txt	1		1											
E:\datasets\phishingemails\phishing (2641).txt	1		1		1		1							
E:\datasets\phishingemails\phishing (2642).txt	1		1		1								1	
E:\datasets\phishingemails\phishing (2643).txt	1				1									
E:\datasets\phishingemails\phishing (2644).txt	0													
E:\datasets\phishingemails\phishing (2645).txt	1				1		1							
E:\datasets\phishingemails\phishing (2646).txt	1				1									
E:\datasets\phishingemails\phishing (2647).txt	1													1
E:\datasets\phishingemails\phishing (2648).txt	1				1									
E:\datasets\phishingemails\phishing (2649).txt	1			1	1						1			
E:\datasets\phishingemails\phishing (2650).txt	1		1		1									
E:\datasets\phishingemails\phishing (2651).txt	1						1							
E:\datasets\phishingemails\phishing (2652).txt	0													

مجلة التربوي
Journal of Educational
ISSN: 2011- 421X
Arcif Q3

معامل التأثير العربي 1.63
العدد 22

الفهـــــــــرس