Grenze International Journal of Engineering and Technology, June Issue

Phishing E-Mail Detection and Blocking it based on the Header Elements

Sulaiman Awadh Ali Obaid Maeli[1] and Ajay U Surwade[2]
[1-2] School of Computer Sciences, Kavayitri Bahinabai Chaudhari North Maharashtra University, Jalgaon, INIDA
Email: suliman032@gmail.com, ajaysurwade@gmail.com

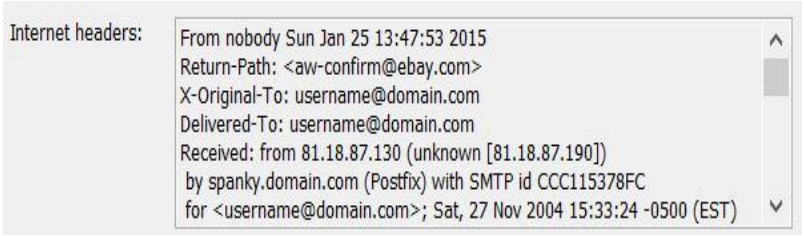
***Abstract*—E-mail is one of the most important modern official means of communication with high reliability, and this is the reason for its widespread and wide popularity. However, this does not make it safe from threats and attacks. The major threat to email is spamming and phishing, which cause a lot of financial losses to the victims. In this paper, we have described a filter based on analyzing the email header elements and its characteristics, extracting the most important features, and testing it on many rules and conditions that can detect and block phishing email messages. This filter is tested on five standard datasets containing spam and phishing emails using header information only and has achieved an overall average accuracy of about 96.31 percent.**

***Index Terms*— Email header, features extraction, Phishing emails, Black-lists, White-lists.**

I. INTRODUCTION

Email is the most important and best way to communicate between companies, institutions, and offices, despite the wide spread of modern means of communication such as social media and others, due to its ease of use, strength of protection, and reliability. But it is not without some problems and defects that threaten protection, and the most prominent of these threats is phishing attack, which is the fraudulent act of pretending to be a reliable entity in a communication in order to obtain confidential user information (such as usernames, passwords, bank account information, or credit card information) and other information [1]. Due to the increase in phishing attacks and the significant financial losses it causes to individuals and companies, much research has appeared to study and block email phishing attacks.

In this paper, we have analyzed the elements of the email header and its various properties to create rules and conditions that can classify an email message as phishing email or non-phishing email. The email, like normal postal mail, contains two parts header part and body part. Fig. 1 shows the header elements in the email message.



Grenze ID: 01.GIJET.9.2.31

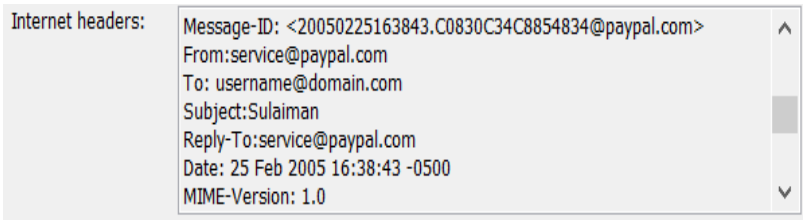


Figure 1. Message Headers in the E-mail

The header: represents an envelope of the E-mail containing information like the sender's and receiver's email addresses, email subject, message journey information across the various servers, cc, bcc, etc [2]. It consists of following fields such as:

i. **From**: contains Email sender information like name and email address of the sender.
ii. **To**: contains the email address(es) of the receiver(s) of the email, which may be delivered to a single recipient or several recipients. It's a mandatory field. The message has to have at least one recipient's address.
iii. **Subject**: contains information about message content.
iv. **Received**: contains information about the message journey, which involves details of mail transmission servers which it has travelled. It can be used to track the message's path.
v. **Reply-To**: includes an email address that is immediately inserted in the "To" field when the user replies to the email message.
vi. **Return-Path**: includes the sender's details, like their email and a link to reply to them. It is added by the server that delivers the message to the recipients.
vii. **Message ID**: It is a special identifier that is given to each message by the host when the message is created. divided into two parts, local and domain, which are separated by an at-sign and enclosed in angle brackets. "<" local-part "@" domain-part ">" [3].

The email has another part called Body which contains the content to be read by the recipient; it could be text, an image, an attachment, or both [4]. The phishers or spammers used to modify these headers information by doing forgery. Thus, changes in these fields are important features.

II. Related Work

Tianda Yang, Kai Qian, Dan Chia-Tien Lo, K. Al Nasr and Ying Qian (2015) Combines different filtering techniques. In hard cases, email header meta features can be used to correctly classify spam. In his experiments, a Naive Bayes filter incorporates these meta features [5].

Ankit Kumar Jain & B. B. Gupta (2016) proposed approach for phishing web pages detection by checking the legitimacy of a webpage not in a whitelist using hyperlink features. Also, this approach detects various types of phishing attacks, like DNS poisoning, zero-hour attacks, etc [6].

D. Kaur and S. Kalra (2016) developed hybrid methods to detect phishing attacks by using whitelists with another technique to check and classify URLs as phishing or non-phishing [7].

Ghogare, Pramod, Surwade, Ajay and Patil, Manoj (2018) devised an approach for spam classification using feature selection and extracting sender email from the message header and used it for classification [8].

Omar Abahussain & Yousef Harrath (2019) have checked incoming mail and examined the email name and email address in the From field, as well as searching for URLs in the email content, extracting the domain name, and comparing it with known domain addresses on the blacklist [9].

T. Krause, R. Uetz and T. Kretschmann (2019) presented a new approach based on meta data for spam classification using a static set of engineered features with automatically extracted features from header data only, without analyzing an email’s body [10].

Thashina Sultana, K. A. Sapnaz, Fathima Sana, and Mrs. Jamedar Najath (2020) proposed a model for detecting spam email and adding the IP address of the spam sender to the blacklist [11].

Anchit Bijalwan (2020) suggested the blacklists method for network traffic to filter infected packets. This technique is used to detect malware and botnets. He has described packet filtering procedure; all suspicious IP addresses are classified as blacklisted [12].

Kulkarni, Priti, Saini, Jatinderkumar, and Acharya, Haridas (2020) Investigated the header attributes of emails using five different feature selection techniques and five different machine learning classifiers [13].

Ajay U. Surwade (2020) developed Origin based Filter which blocks phishing e-mail by extracting header part information of e-mails using Blacklist approach [14].
Youness Mourtaji, Mohammed Bouhorma, Daniyal Alghazzawi, Ghadah Aldabbagh, and Abdullah Alghamdi (2021) developed a solution based on a hybrid rule-based approach that extracts features from six different methods, including the blacklisting method, which checks the domain name against two antiviruses blacklists that consider this domain blacklisted [15].

III. METHODOLOGY

This work is based on analyzing the important header element to extract important features to design an origin-based filter that has some rules and conditions, as mentioned in Section-IV, to classify the email as phishing or non-phishing. The methodology adopted for this origin-based filter (OBF) as shown in Figure-2.

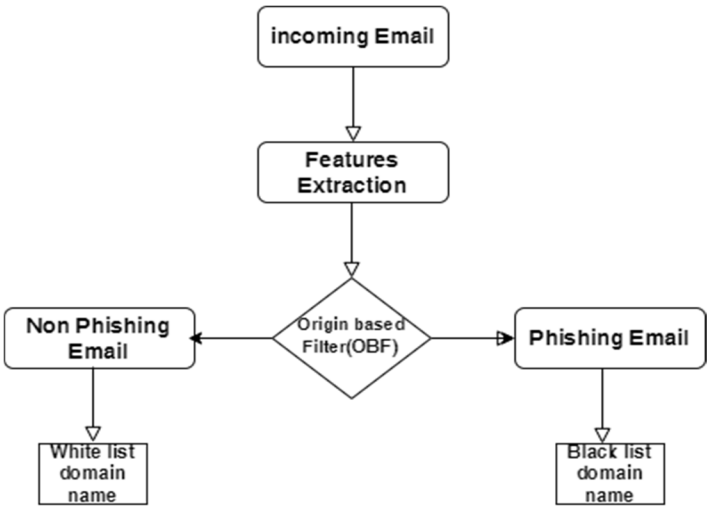


Figure 2. Structure of the OBF Filter

The features such as 'From', 'Reply-To', 'Return-Path' and 'Message-ID' as shown in Table-1, are extracted from the standard datasets such as Enron [16], Public Phishing Corpus [17], SPAM Archive [18], CSDMC2010_SPAM [19] and Spam Assassin [20].
According to the RFC822 protocol, the message header has been extracted and four standard fields selected for features extraction as shown in table I:

TABLE I. STANDARD HEADER FIELDS WITH EXTRACTED FEATURES

Field	Extracted Features	Probability Value
From	invalid address (wrong or fake mail address)	0 or 1
Reply-To	mismatching domain names between "ReplyTo" and "From" addresses.	0 or 1
Return-Path	mismatching domain names between "Return-Path" and "From" addresses.	0 or 1
Message-ID	invalid address or mismatching domain names between "Message-ID" and "From" addresses.	0 or 1

The Python code is developed using the rules mentioned in Section-IV. This code is tested with the standard datasets as mentioned above. This architecture is classifying emails as Phishing or Non-Phishing emails. The Accuracy of classification has been calculated and reported in Table-II to VI. The IP addresses or domain names of classified Phishing emails are extracted and are stored as the Blacklists similarly, the IP addresses or domain names of classified Non-Phishing emails are extracted and are stored in the Whitelists.

IV. IMPLEMENTATION DETAIL

The following rules and conditions for classification have been set after carrying out the detail analysis of the standard datasets which are described with necessary description:

a) From field is invalid mail (wrong or fake mail address).

A regular expression is used to validate the From field with various email address formats.

b) From field and Reply-To field has different domain name.
FROM != REPLY-TO

c) From field and Return-Path fields has different domain name.
FROM !=RETURN-PATH

d) Reply-To or Return-Path or both is empty.
len(REPLY-TO) == 0 OR len(RETURN-PATH) == 0
OR (len(REPLY-TO) == 0 AND RETURN-PATH) == 0)

e) From field and message id has different domain name.
from_domain!= message_id

There is one possible probability value for each condition mentioned above. "0" means "false" or "1" means "true.". The value "1" (ie, 'true') classify email as phishing email while, value "0" (ie, 'false') classify email as non-phishing email. The decision-making condition is as given below:

If ((FROM != RETURN-PATH) or (FROM != REPLY-TO) or (len(REPLY-TO) == 0) or len(RETURN-PATH) == 0 or (len(REPLY-TO) == 0) AND (RETURN-PATH) == 0) or (from_domain != message_id))

Then classify email as Phishing Email and Extract it's IP addresses or Domain names and store it as Blacklists.

Else, classify email as Non-Phishing Email and Extract it's IP addresses or Domain names and store it as Whitelists.

The results collected during experiments are reported in next sections.

V. Result and Discussion

The Tables-II to VI represents the results collected after the classifications are reported with the help of confusion matrix.

TABLE II. EMAIL CLASSIFICATION IN ENRON DATASET

Enron				
SPAM				
Folder Name	Phishing	Non Phishing	Total	Accuracy in %
BG	9402	598	10000	94.02
GP	13719	0	13719	100
SH	9256	13	9269	99.85

TABLE III. EMAIL CLASSIFICATION IN PUBLIC PHISHING CORPUS DATASET

Public Phishing Corpus				
Folder Name	Phishing	Non Phishing	Total	Accuracy in %
Phishing 0	398	16	414	96.13
20051114	411	27	438	93.83
Phishing 2	1398	25	1423	98.24
Phishing 3	2225	54	2279	97.63

TABLE IV. EMAIL CLASSIFICATION IN SPAM ARCHIVE DATASET

SPAM Archive				
Folder Name	Phishing	Non Phishing	Total	Accuracy in %
01/2020	2528	115	2643	95.64
02/2020	6780	715	7495	90.46

TABLE V. EMAIL CLASSIFICATION IN CSDMC2010_SPAM DATASET

CSDMC2010_SPAM				
Folder Name	Phishing	Non Phishing	Total	Accuracy in %
Spam	1315	63	1378	95.42

TABLE VI. EMAIL CLASSIFICATION IN SPAM ASSASSIN DATASET

Spam Assassin				
Folder Name	Phishing	Non Phishing	Total	Accuracy in %
20030228_spam	486	14	500	97.20
20030228_spam_2	1360	37	1397	97.30

The results for 'Spam Archive' dataset shown in Table-IV, the Folder name '02/2020' has minimum accuracy which is 90.46 %. This folder contains total 7495 spam emails which should have been classified as Phishing email but, only 6780 have been classified as Phishing and 715 are misclassified as non-Phishing.

The results for ‘Public Phishing Corpus’ dataset shown in Table-III, the Folder name ‘20051114’ contains total 438 phishing emails which should have been classified as Phishing emails but, only 411 have been classified as Phishing and 27 are misclassified as non-Phishing.
The results for Enron dataset shown in Table-II, the Folder name ‘BG’ contains total 10000 spam emails which should be classified as Phishing emails but only 9402 are classified as Phishing while, 598 have been classified as non-Phishing.
The results for ‘CSDMC2010_SPAM’ dataset shown in Table-V, the Folder name ‘Spam’ contains total 1378 spam emails which should be classified as Phishing emails but only 1315 are classified as Phishing while, 63 have been classified as non-Phishing.
The above mis-classification suggests that, the features extracted from these emails are not sufficient and need some more features for accurate classification. So, we need to investigate the other features along with existing features so that accuracy can be improved. In future we are planning to investigate other header features ‘Subject’ field of the email.

VI. Conclusions and Future Work

A phishing e-mail is a real threat to individuals and entities that must be detected and blocked before it reaches its target. In this paper, a filter has been designed based on the characteristics of some elements of the message header, extracting the features, and then classifying the message according to the conditions and rules that detect and block phishing mail. The experiment has been carried out on some standard datasets such as Enron, Public Phishing Corpus, SPAM Archive, CSDMC2010_SPAM, and Spam Assassin and has achieved 96.31% average accuracy of classification.
In future, we are planning to increase accuracy by selecting more elements from the header part and investigating those features which can improve the results in terms of accuracy of classification. We are planning to concentrate on the ‘Subject’ field of email header part which may help us to increase accuracy of phishing classification.

References

[1] Tak, Gaurav & Ojha, Gaurav. (2013). Multi-Level Parsing Based Approach Against Phishing Attacks with the Help of Knowledge Bases. International Journal of Network Security & Its Applications. 5. 10.5121/ijnsa.2013.5602.
[2] P. Mishra, E. S. Pilli and R. C. Joshi, "Forensic Analysis of E-mail Date and Time Spoofing," 2012 Third International Conference on Computer and Communication Technology, 2012, pp. 309-314, doi: 10.1109/ICCCT.2012.69.
[3] Hamid, I. R. A., Abawajy, J., & Kim, T. H. (2013). Using feature selection and classification scheme for automating phishing email detection. Studies in informatics and control, 22(1), 61-70.
[4] Beaman, C., & Isah, H. (2022). Anomaly Detection in Emails using Machine Learning and Header Information. arXiv. https://doi.org/10.48550/arXiv.2203.10408.
[5] Tianda Yang, Kai Qian, Dan Chia-Tien Lo, K. Al Nasr and Ying Qian, "Spam filtering using Association Rules and Naïve Bayes Classifier," 2015 IEEE International Conference on Progress in Informatics and Computing (PIC), 2015, pp. 638-642, doi: 10.1109/PIC.2015.7489926.
[6] Jain, A.K., Gupta, B.B. A novel approach to protect against phishing attacks at client side using auto-updated whitelist. EURASIP J. on Info. Security 2016, 9 (2016). https://doi.org/10.1186/s13635-016-0034-3 .
[7] Davneet Kaur and Sheetal Kalra, “Five-tier barrier anti-phishing scheme using hybrid approach”, Information Security Journal-A Global Perspective, 2016, DOI: 10.1080/19393555.2016.1215573.
[8] Ghogare, Pramod & Surwade, Ajay & Patil, Manoj. (2018). Effective E-mail Spam Filtering Using Origin Based Information. International Journal of Computer Sciences and Engineering. 6. 359-362. 10.26438/ijcse/v6i11.359362.
[9] O. Abahussain and Y. Harrath, "Detection of Malicious Emails through Regular Expressions and Databases," 2019 International Conference on Innovation and Intelligence for Informatics, Computing, and Technologies (3ICT), 2019, pp. 1-5, doi: 10.1109/3ICT.2019.8910291.
[10] T. Krause, R. Uetz and T. Kretschmann, "Recognizing Email Spam from Meta Data Only," 2019 IEEE Conference on Communications and Network Security (CNS), 2019, pp. 178-186, doi: 10.1109/CNS.2019.8802827.
[11] Thashina Sultana, K A Sapnaz, Fathima Sana, Jamedar Najath, 2020, Email based Spam Detection, INTERNATIONAL JOURNAL OF ENGINEERING RESEARCH & TECHNOLOGY (IJERT) Volume 09, Issue 06 (June 2020).
[12] Anchit Bijalwan, “Botnet Forensic Analysis Using Machine Learning”, Hindawi’s Journal of Security and Communication Networks, Volume2020, pp:1-9, February-2020. Article ID 9302318, https://doi.org/10.1155/2020/9302318.
[13] Kulkarni, Priti & Saini, Jatinderkumar & Acharya, Haridas. (2020). Effect of Header-based Features on Accuracy of Classifiers for Spam Email Classification. International Journal of Advanced Computer Science and Applications. 11. 10.14569/IJACSA.2020.0110350.

[14] A. U. Surwade, "Blocking Phishing e-mail by extracting header information of e-mails",(2020), International Conference on Smart Innovations in Design, Environment, Management, Planning and Computing (ICSIDEMPC), 2020, pp. 151-155, doi: 10.1109/ICSIDEMPC49020.2020.9299596.
[15] Mourtaji, Youness & Bouhorma, Mohammed & Alghazzawi, Daniyal & Aldabbagh, Ghadah & Alghamdi, Abdullah. (2021). Hybrid Rule-Based Solution for Phishing URL Detection Using Convolutional Neural Network. Wireless Communications and Mobile Computing. 2021. 1-24. 10.1155/2021/8241104.
[16] http://nlp.cs.aueb.gr/software_and_datasets/Enron-Spam/index.html. Last accessed on 10 September, 2022.
[17] https://academictorrents.com/details/a77cda9a9d89a60dbdfbe581adf6e2df9197995a Last accessed on 14 October 2022.
[18] http://untroubled.org/spam/ Last accessed on 28 September 2022.
[19] https://github.com/jdwilson4/Intro-to-MachineLearning/tree/master/Data/SPAMData Last accessed on 12 November 2021.
[20] https://spamassassin.apache.org/old/publiccorpus/ Last accessed on 2 October 2022.