International Journal of Applied Information Systems (IJAIS) – ISSN : 2249-0868
Foundation of Computer Science FCS, New York, USA
Volume 12 – No. 9, December 2017 – www.ijais.org

Spam Detection Approaches and Strategies: A Phenomenon

Balogun Abiodun Kamoru
Dept. of Software Engr and Information Systems
Faculty of Computer Science and Information Technology
Universiti Putra Malaysia
Serdang 43300 Selangor Malaysia

Azmi Bin Jaafar
Dept. Software Engr. and Information Systems
Faculty of Computer Science and Information Technology
Universiti Putra Malaysia
Serdang 43300 Selangor Malaysia

Masrah Azrifah Azmi Murad
Dept. of Software Engr. and Information Systems
Faculty of Computer Science and Information Technology
Universiti Putra Malaysia
Serdang 43300 Selangor Malaysia

Ezema Onyeka Ernest
Dept.of Software Engr and Information Systems
Universiti Putra Malaysia Serdang 43300 Selangor Malaysia

Marzanah Binti A. Jabar
Dept. of Software Engr and Information Systems
Universiti Putra Malaysia
Serdang 43300 Selangor Malaysia

ABSTRACT
The massive increase of spam is posing a very dangerous and serious threat to our email and social networks. It is pertinent and imperative to step up the spam detection approach and strategies in email and social networks. In recent years online spam has become a major problem for the sustainability of the internet globally, Excessive amounts of spam are not only reducing the quality of information available on the email and social networks but also creating concern among the users of email and various social networks.

This paper aims to analyze existing research works in spam detection strategies and approaches, state of art, the phenomenon of spam detection, to explore the rudiment of spam detection, to proposed detection scheme and potential online mitigation schemes. The paper will surveys various anti spam strategies for email and social networking. In the literature we have studied that many anti spam strategies have been discovered and work on but they are still open challenges to these different approaches and techniques, while some of them are highlighted in this articles. It is very important to work on spam detection and reposition it for the better of the world..

General Terms
Spam Detection, Anti-Spam Approaches, Social Networks, Unsolicited Bulk E-mail, Unsolicited Commercial E-mail, Algorithms

Keywords
Spam Detection, Anti-Spam Strategies, Spam Mitigation, Spam Detection Scheme

1. INTRODUCTION
Spam is Unsolicited Commercial Email (UCE), is not a new problem causing complaints from many internet users globally. Spamming is the act of sending unsolicited commercial email, involves the sending of nearly identical emails to thousand or even millions of recipients without the recipients' prior consent or even violates recipients explicit refusal [1] [2] [3]. Internet is used on a daily basis to search for information and acquire knowledge [4]. Spam is increasingly being used to distribute virus, spyware, links to phishing websites, etc. The problem of spam is not only threat but also annoyance that has become a dangerous phenomenon to our existence. Unsolicited Bulk email (UBE) is another category of emails that can be considered spam.. As suggested in recent reports by Spamhaus and Symantec [7][8].

For instance, Symantec has detected 44% increase in phishing attempts from the first half of 2016 to the second half. Statistics from the Distributed Check sum Clearinghouse (DCC) Project [6] shows that 54% of the email messages checked by the DCC network in 2016 are likely to be from bulk email. According to MX logic [5] shows that an average 80.78% of the email messages delivered to their clients during week of March 24-30,2016 are considered to be spam email, with peaks more than 90%.

There are six main forms of spam, and they have different effects on End users globally like: (1) E-mail spam; (2) Comment spam; (3) Instant Messenger Spam;(4) Unsolicited text messages; (5) social networking spam; (6) Blogging and live stream spam[9].

Various legal means of anti-spam attempts have been work on by previous research[10] [11]. Legislation specifically targeted at email spam as well as unwanted messages in general have been introduced in some countries, such as the United State of America. Before targeted legislations are introduced, some existing laws are sought for fighting spam. Possible approaches and techniques are based on laws and statutes that combat fraud antiracketeering and anti harassment. These approaches are considered ineffective as they require considerable costs and efforts for the prosecutor to prove the relevance between the spam messages and the law. Another challenging problem to the legal approach is the limited jurisdiction of the law concerned. Also, many legislator are forced to leave loopholes in the legislations to avoid infringing the freedom of speech [10]. These often allow spammers to slip through and the restriction merely becomes a burden to legitimate senders.To reduce or mitigate spams, various anti spam methods have been proposed in state-of-the-art research[2] Heymann et al, classified anti-spam strategies into three categories: (i) Prevention Based; (ii)

Detection Based; (iii) Demotion Based. There are various anti-spam strategies as content based, link based, graph analysis, clocking scheme but in spite of having various anti-spam strategies there are various open challenges to these anti-spam strategies and approaches, phenomenon which need to be addressed.

To prevent users from being overwhelmed by spam, many internet service providers (ISP) and organizations deploy spam filters at the email server level. The family of Naive Bayes (NB) classifiers [13] is probably one of the most commonly implemented, which is also embedded in many popular email and social networks clients. They extract keywords and other indicators from email messages and determine whether the messages are spam using some statistical or heuristics scheme. However, spam senders (spammers) nowadays are using increasingly sophisticated techniques and approaches to trick content based filters by clever manipulation of the spam content [14]. for analysis. Also , words with scrambled character order can render vocabulary-based detection techniques ineffective, yet human can still understand the scrambled words. As a consequence, content-based filters are becoming less effective and hence other approaches are being explored to complement them,

One popular approach is based on blacklists and whitelists. A blacklist is a list of senders whose emails are blocked from getting through to the recipients. A whitelist is just the exact opposite. While a blacklist specifies who is to be kept out allowing all others to pass, a whitelist only allows those who are already on the list to get through . Since Spammers almost always spoof the " From" field of spam messages, blacklists usually keep IP addresses rather than email addresses. For incoming messages from senders not on the lists, content-based filters may be applied so that the approaches can complement each other.

In this paper, we propose to analyze existing research works in spam detection strategies and approaches, state of art, the phenomenon of spam detection, to explore the rudiment of spam detection. The paper will surveys various anti spam strategies for email and social networking. In the literature we have studied that many anti spam strategies have been discovered and work on but they are still open challenges to these different approaches and techniques, to look into the online mitigation methods .

The rest of this paper is organized as follows; section 2 review various form of related work on spam detection method in email and social networks, section 3 describes the anti -spam strategies and need for spam detection addressed by this paper. Section 4 details the spam detection techniques and section 5 explores the possible challenges and spam mitigation strategies and to proposed detection scheme. Section 6 present the conclusion and final part of this paper is section 7 which is our reference.have to improvise.

2. RELATED WORK ON SPAM DETECTION IN EMAIL AND SOCIAL NETWORKS

2.1 Related Work

A Spam detection method tries to determine whether a sender is a spammer or legitimate sender. Balogun et al,2017[14] Spam detection on social networks and email mainly focuses on the following : (1) Anomaly detection;(2) Fault detection; (3) Malware detection;(4) Intrusion detection. If a considerable effort is not made to find a technological solution to the menace of spam. The internet email and social email is in danger as an important medium of communication.

Ahmed and Abulaish,2013 [15] present that spammers are trying to a new approach to gain access through social media and email. While most of the previous work on social spam and email spam has focused on spam prevention on a single email or social spam. Social spam and email spam is relatively new research area and the literature is still sparse [4][5][7]. A large number of social spam and email spam classifiers have been used in spam detection but choosing the right classifier and the most efficient combination of them is still problem with previous scholar work. Although there are still limited studies on spam detection..

Taylor[16] discussed the domain reputation system deployed in Google's Gmail System, the reputation maintains the reputation for each domain that sends to email to Gmail. This reputations are calculated based on previous results from statistical filters and user feedback. If the reputation of a domain is good, the domain will be white listed and the reverse will be blacklisted. The emails from senders in neither lists are further processed with statistical anti-spam filters for making the final decision. Email classification results are logged as auto spam or auto non-spam events. Users can send feedback to the system by clicking on a button in the webmail interface for reporting misclassification. These events are also logged and used during the next update reputations.

Taylor also discussed the problem of spoofed source addresses which can affect sender-based detection systems. The sender policy framework (SPF) and Domain-based email authentication (Domain Keys) mechanisms are used to authenticate whether an email is really sent from the domain that it claims to be from. Besides reputation systems, heuristics-based approaches have also been explored.

Harris proposed a heuristic method called Greylisting [13] to avoid receiving spam at the recipient's email transfer agent (MTA). when a recipient MTA that uses Grey listing receives a delivery attempt, the MTA will respond with an SMTP temporary error messages. The recipient MTA will record the identity of the recent attempts of delivery so that the next attempt will be accepted. Legitimate senders that conform to the standard will have their message delivered. Whereas spammers, who concern more about coding simplicity and speed of the spamming engine, ignore any error message and move on to the next recipient in the list instead of retrying. Thus, Spam can be avoided.

Structural features in email social networks may also be exploited for sender-based spam detection. Gomes et al presented a graphic-theoretic analysis of email traffic and proposed several features that can serve to distinguish between spam and legitimate email. Although they did not present any spam detection study in this paper, the features proposed can be used for spam detection. In particular.

Boykin and Roychowdhury [3] proposed an automated anti-spam tool that exploits the properties of a social network, the nodes in the network are clustered to form spam and non-spam component automatically.

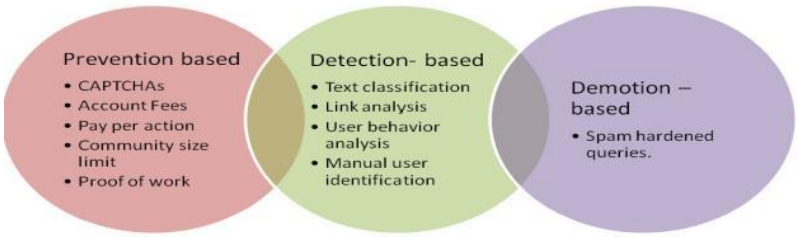

Fig 1: Anti-Spam Strategies

3. ANTI-SPAM STRATEGIES AND NEED FOR SPAM DETECTION

3.1 Anti-Spam Strategies

Anti-spam strategies are part of the strategies to mitigate spam in email and social spam, the following are anti-spam:(i) Prevention Based (ii) Detection Based (iii) Demotion Based.

- **Prevention Based**: This approach aims at making it difficult for spam content to contribute to social tagging system by restricting certain access through interfaces (Such as CAPTCHA which stands for "completely automated public turing test to tell computers and human parts") or through usage limits (such as tagging quot e.g Flickr introduced a limit of 75 tags per photo.

- **Detection Based**: Detection based approaches identify likely spams either manually or automatically by making use of machine learning (such as text classification) or statistical analysis (such as link analysis) and then deleting the spam content or visibly marking hidden to the user. For these methods, we can treat the corpus as set of objects with associated attributes. In e-mail spam, the messages ar objects and the headers are attributes. In web spam, the web pages are objects and attributes might be inlinks, outlinks, page content and various meta data.

- **Demotion Based**: The approach reduces the prominence of content likely to be spam. For instance rank based methods produce ordering of a system's , tags or users based on trust score. the figure below

3.2 Need For Spam Detection

The Spam is a threat to the users of internet globally. Due to the challenges for the service providers because of the following negative effects are [2]:

- Spam deteriorates the quality of search results and deprive legitimate websites of revenue.
- Spam have economic impact since highest ranking provides large free advertising and so an increase in web traffic volume.
- User trust is weaken due to the search engine provider which is especially tangible issue to zero cost of switching from one search provider to another.
- Spam websites are means of malware and adult content dissemination and phishing attack.
- Identification of the most appropriate tags for the given content and to eliminate the spam tag.

4. SPAM DETECTION TECHNIQUES AND APPROACHES

4.1 Spam Detection Techniques

The spam detection techniques can be categorized into 4:

1. Content Based
2. Link Based
3. Algorithms that exploits click stream
4. Semantic based spam detection

4.1.1 Content Based

Spam detection techniques which analyze content features such as word count or language models and content duplication. Fetterly et al proposed that web spam pages exhibit some anomalous properties as : (1) URL of spam pages have exceptional number of dots, dashes, digits and length, (2) Most spam pages that resides on the same host have very low word count variance,(3) Content of spam pages changes very rapidly.[5] Features based on HTML page structure to detect script generated spam pages. In this preprocessing is made by removing all the content and considering only layout of the page. They applied finger printing technique with subsequent clustering to find groups of structurally near spam pages [5]. Mishne et al proposed a line of work on language modelling for spam detection. They proposed an approach of spam detection in blogs by comparing the language models for blog comments and page linked with these comments. They use KL divergence as a measure of discrepancy [7]. in other work by Sydow lingusitic features were analyzed for web spam detection by considering lexical validity and content diversity, syntactical diversity and entropy, usage of active and passive voices and various other NLP feature [8].

4.1.2 Link Based

The link based approach analyze link based information such as neighbor graph connectivity. Based on identification of suspicious nodes and links and their subsequent down weighting. Extracting link based features for each node and use various machine learning algorithm to detect spam. Graph regularization technique for spam detection. In this link information is used to compute global importance scores for all pages pi to page pj. Algorithm follows repeated improvement principle i.e the true score is computed as convergence point of an iterative updating process [5]. Algorithms belonging to this category represent pages as feature vectors and perform standard classification or clustering analysis. Studies link-based feature to perform website categorization based on their functionality, their assumption is that sites sharing similar structural pattern, such as average page level or number of out links per leaf page,

share similar roles on web. For e.g web directories mostly consists of pages while spam site have specific topology aimed to optimize page Rank boost and demonstrate high content duplication. Overall, eacg website is represented as a vector of 16 connectivity and a clustering is performed using cosine as a similarity measure [5]

4.1.3 Algorithms that exploits click stream

Data and user behaviour of data , query popularity data or information and HTTP session information, since click spam aims to push " Malicious noise" into a query log with the intention to corrupt data, used for the ranking function construction, most of the counter methods study the ways to make learning algorithm robust to this noise. Other anti-click-fraud methods are driven by the analysis of the economic factors underlying the spammers ecosystem. Interesting idea to prevent click spam is proposed.. The author suggests using personalized ranking functions, as being more robust, to prevent click fraud manipulation[3].

4.1.4 Semantic Based Spam Detection

To overcome the drawback of content based spam detection, semantic based detection method is used where instead of content semantic of the web site is analyzed.

5. POSSIBLE CHALLENGES, SPAM MITIGATION STRATEGIES AND PROPOSED SPAM DETECTION SCHEME

5.1 Possible Challenges in Spam Detection Techniques

In literature review we have studied that many anti-spam strategies have been discovered but still there are some possible challenges to these techniques. Some of them highlighted below:

It is observed that interaction across social network become popular. For e.g. users can use their Facebook account or email account to log in some other social network services. Thus future challenges is to investigate how trust model across domains can be effectively connected and shared.

However social network services are used by people from various countries , so various languages simultaneously appears in tags and comment. In such cases some text information may be regarded as wrong or considered as spam due to language spam. Therefore incorporating multilingual in trust modelling would solve this problem.

Most of the existing approaches based on text information assuming monolingual environment.

Trust modeling most of the current techniques for noise and spam reduction focus only on textual tag processing and user profile analysis while audio and visual content features of multimedia content can also provide useful information about the relevance of the content tag relation.

In trust modeling system user's trust tends to vary over time according to the users' experience and involvement of social networks. Only a few approaches deals with the dynamics of trust by distinguishing between recent and old tags. Future work considering dynamics of trust would lead to better modeling in real world application.

5.2 Spam Mitigation Strategies Scheme

A detection scheme needs a mitigation strategy to react to spam in e-mail and social networks. There are more than one way to use the legitimacy avenue provided by the social network based detection strategy scheme to mitigate spam. One of the more straight forward ways is to apply a threshold to the score which email from the sender will be filtered. While this approach is simple, we observe that it is unlikely that social network based detection alone is accurate enough for the purpose. Also, existing content-based schemes and rule-based schemes are still performing reasonably well. We prefer to use the social network based detection strategy scheme to complement rather than replace content-based filtering techniques.

Different ways of combining filters have been explored in the literature. Segal et al.[9] proposed to form a pipeline of anti-spam filter components. An e-mail passes through each component in the pipeline one by one. Each component assigns a score to the email. An e-mail can be directly classified by an intermediate classifier and skip all subsequent components if the classifier determined the classification of the email with high confidence. Lynam and Cormack [14] explored different ways of combining anti-spam filters. Specifically, the voting of binary classifications from spam classifiers, the log-dds averaging of spam cores, the use of SVM on spam scores from different spam filters and the use of logistic regression to find the weights for computing the weighted average of spam scores for multiple filter strategies.

Since the main focus of this paper is to mitigate spam from e-mail and social networks, we intend to discuss only simplified views of three potential approaches in which the legitimacy sender scores may be used to complement existing score generating filters. In depth study on the benefits and effectiveness on advanced filter ensemble schemes are reserved for future work.

There are about three approaches to mitigate spam on email and social networks: (1) Parallel single thresholding approach, (2) Serial multiple thresholding approach, (3) Serial throtting and thresholding approach.

1. Parallel single thresholding approach

Many of the content-based spam detection schemes are able to generate a spam score, and so does the proposed social network based scheme. A natural way to combine the two is to run the two scheme in parallel so that each of them generate a score. The two scores are combined to give a decision.

An email is fed to both schemes, the content-based analyzer will analyze the content of the email and assign a score Sc to the email. The higher the score Sc is , the more confident that the analyzer thinks the email is spam. The proposed social network based scheme will identify the originator of the email concerned and query the score database is a legitimacy score, we may switch it to a spam score by a simple negation, I.e., Ss= (-1)Yi. This spam score can then be combined with other content-based and rule-based filters with, for example, a weighted sum, to generate a single spam score. Emails with a score higher than a certain threshold can be considered as spam.

2. Serial multiple thresholding approach:

To cope with the advanced techniques of spamming, content-based filters are getting more and more sophisticated. The sophistication also translates to heavier load on the spam

filtering module. On the contrary, the spam sender score is first determined offline. Only a lightweight query to the score database is needed during the online process. One may consider taking a serial approach by filtering spam with lightweight sender score approach first.

A serial multiple thresholding system, during the spam filtering process, the legitimacy score for email sender will first be fetched from the database, Two thresholds Ti >Ts on this score will be defined. Emails from senders with the legitimate score above Ti will be accepted directly to the inbox, skipping the content-based filter. Senders with a score lower than Ts will be considered spammers. Their emails can be rejected at this stage or flagged as spam directly. Email from senders with a score in between the two thresholds, i.e., Ts< Fi<Ti will be passed to content-based analyzer that will make the final decision. Spammy emails can be filtered or flagged accordingly.

This approach has several advantages. The sender based filtering scheme acts like an automatic whitelist and blacklist approach. As a result, the load on the content-based filter will be lowered. Additional resource intensive analysis on the email, such as Optical Character Recognition (OCR) on images, may now be enabled to improve the accuracy. Also, notice that some of the legitimate senders are allowed to skip the subsequence filters, an administrator may use a more aggressive threshold for those filters while maintaining the same false overall rejection rate.

3. Serial throtting and thresholding approach:
A variant of the serial approach is to throttle rather than to filter senders with the sender score. It is observed that spammers generally depend on a very high email delivery throughput to generate a revenue [7]. Li et al proposed to slow down the transmission rate of a suspected spam at the TCP level. Although they did not propose a way to identify suspicious senders, the authors showed that when enough recipients are using TCP damping against emails with a high spam likelihood, it is possible to lower the delivery throughput of spammers considerably . This may be used as a deterrent that drives spammer away from the server to avoid high delays.

Also, since emails are slowed down during the delivery but not entirely dropped, false rejection of legitimate email is much less expensive. An average user may not care about some minutes of delays in delivery. However, for spammers that require high delivery throughput, a slowdown in delivery hurts their profitability.

One potential problem is that the scheme requires online determination of the spam likelihood of an email. Content-based analysis may fall short in this aspect since limited information about the email is revealed during the earlier stage of the email delivery process, by the time the content of an email is being received and analyzed, it may already be near the end of the delivery. It may be too late for TCP damping to slow down the spammer enough to make a difference.

Given that our proposed scheme gives a sender score and an online query of the score is lightweight, one can afford to implement TCP damping with the legitimacy score. Senders with a high legitimacy score would be offered normal or preferential service while others are slowed down. One of the way to use an exponential decay function on the legitimate score to determine the extent of damping. The delay imposed by the server grows exponentially with the decrease of the legitimate score.

5.3 Proposed Spam Detection Scheme

figure below is an overview of the proposed solution for detecting spam senders. Email and Social networks are first constructed from email transaction logs. A social network can be represented by a direct graph where senders are represented as nodes and email transactions are represented as edges. After the feature extraction and pre-processing stages, a machine learning method, such as k-Nearest Neighbour (k-NN) classifier, can be used for the classification task. Some post processing on the classifier output may yield results that are more versatile. The proposed spam detection scheme will give a permanent solution to the possible challenges pose by the spam on email and social network.

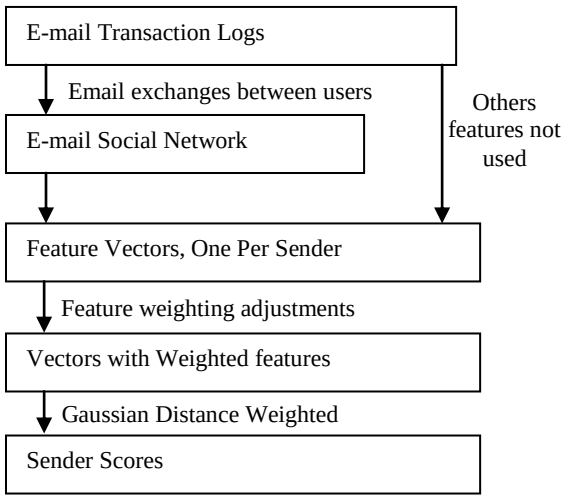

Fig2: Proposed Spam Detection Scheme

6. CONCLUSIONS

In this paper, we have explored the new direction for a proposed scheme of spam detection, we have been able to explore the possible challenges on spam detection, spam detection approaches and spam detection techniques. From the above study we have studied various spam detection approaches and techniques and explored the open challenges and issues which has to be addressed and left as open challenge for research. For future research it could be combine multi media content analysis with conventional tag processing and user profile analysis.

7. ACKNOWLEDGMENTS

We Thank Associate Professor Azmi Bin Jaafar for his ideas, Support and financial assistance through the help of Faculty of Computer Science and Information Technology, Univeristi Putra Malaysia. This work will not be possible without the help of Deputy Dean of the Faculty of Computer Science Universiti Putra Malaysia for her support during the cause of writing this paper. However, our gratitude goes to Associate Professor Dr.Marzanah A.Jabar for her comment, suggestion and positive contribution. We also thank all those researchers whose works are referenced. Finally, we wish to thank Ernest Onyeka Ernest a Master Research from Dept. of Information Security, Universiti Putra Malaysia for his support and contribution and lastly the anonymous reviewers for their valuable comments.

8. REFERENCES

[1] Bowman, M., Debray, S. K., and Peterson, L. L. 1993. Reasoning about naming systems. .

[2] Ding, W. and Marchionini, G. 1997 A Study on Video Browsing Strategies. Technical Report. University of Maryland at College Park.

[3] Fröhlich, B. and Plate, J. 2000. The cubic mouse: a new device for three-dimensional input. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems

[4] Tavel, P. 2007 Modeling and Simulation Design. AK Peters Ltd.

[5] Sannella, M. J. 1994 Constraint Satisfaction and Debugging for Interactive User Interfaces. Doctoral Thesis. UMI Order Number: UMI Order No. GAX95-09398., University of Washington.

[6] Forman, G. 2003. An extensive empirical study of feature selection metrics for text classification. J. Mach. Learn. Res. 3 (Mar. 2003), 1289-1305.

[7] Brown, L. D., Hua, H., and Gao, C. 2003. A widget framework for augmented interaction in SCAPE.

[8] Y.T. Yu, M.F. Lau, "A comparison of MC/DC, MUMCUT and several other coverage criteria for logical decisions", Journal of Systems and Software, 2005, in press.

[9] Spector, A. Z. 1989. Achieving application requirements. In Distributed Systems, S. Mullender.