Proposed framework for Spam recognition in big data for Social Media Networks in smart environment

1st Minimol Anil Job
Dept. of Computing
Arab Open University
Bahrain
m.aniljob@aou.org.bh

2nd Jitendra Pandey
Dept. of Computing,
Middle East College, Muscat, Oman
jitendra@mec.edu.om

***Abstract*—Social Media Networks (SMNs) are becoming more and more popular across the globe in the past few years. Netizens share all their personal information regarding day to day activities, views, and opinions across various SMNs. Simultaneously, it can be observed that throughout the most popular SMNs face frequent social spam problems in various formats. Big Data theory is gaining much more attention and it is expected that SMNs will have more interactions with each other shortly. This would enable a spam link, content or profile attack to easily move from one social network like Twitter to other social networks like Facebook. Consequently, effective discovery of spam has turn out to be a noteworthy and prevalent issue. Proposed research highlights spam discovery across several SMNs by leveraging the data of sensing analogous spam inside an OSN (Open Source Network) and using it across various SMNs. Authors have selected Twitter and Facebook as the research marks, In this research paper, the authors proposed and presented a spam detection framework to find out spam on more than one social network those are most common features in terms of contents, behavior, posts and user involvement etc. Spam detection techniques can significantly facilitate in various social network to measure the vulnerability. Based on this fact the researchers have proposed and presented a spam detection framework to find out spam on more than one social network.**

Keywords—Social Media Networks, Big Data, Spam messages, Spam detection, framework

I. INTRODUCTION

In this digital era, the information on the web is accessible by users 24/7 globally. To obtain useful information users habitually identify useful web pages by querying search engines. When the users submit their search query, the search engine recognizes related pages which include the search criteria on the web. Once the match is identified search engines display the users with the links to resulting pages. Spammers usually use methods like Search Engine Optimization (SEO) or SQL injection to increase the page rank of the focus web page in search results. Social networks are growing extremely fast and there are many options with the users now specially with the young population. Twitter, Instagram, snapchat and Facebook being more common with the population. Over millions of audiences are attracted to these social networks in a month. Due to the increase in the usage of smart devices, society has taken these social media applications as an important medium of communication. At the beginning spams are introduced only in mails. Now it can be observed that spams have been extended to Social networks severely. Every social networking application has their own mechanism to filter and detect the spam content they witness. However, the challenge is that keeping the huge amount of data in picture, the kind of infrastructure is required to manage this is huge. And as soon as they get familiar with one, there is another one waiting for them. Every social network has its own unique characteristics with respect to the type of data they have and the king of secure environment they need. The first characteristic is 'data retrieval its usage and analysis is totally reliable on trust'. The second characteristic is 'end user knowledge about the possible intrusion at various stages'. To access the network of trust of other users is more complicated as there is no reliable and trustworthy authentication mechanism provided by the so called most popular social media platforms. As we are aware, presently Facebook has become highly popular among users and is used most number of users across the world over the internet.

Facebook administrators claim on their website that as of the third quarter of 2018, above 2.27 billion active users per month (Facebook MAUs). As per the source Facebook 10/20/2018, they experience almost ten percentage increase every year. According to the reports, the number of active users in the Facebook has increased to one billion in the third quarter of 2012. Thus Facebook has become the first social network with huge number of active users. Facebook defines active users as the users who are logged in to Facebook during the last 30 days.

From a Source published by Facebook on 2/01/17, there are 1.15 billion mobile users active daily until December 2016. It shows an increase of 23 percent year-over-year. Another claim by Facebook is that 4.75 billion pieces of content shared daily as per the report of May 2013. It shows a 94 percent increase from August 2012. According to 'zephoria.com 2018', Facebook estimated that in each month almost 2.6 billion people use Facebook, WhatsApp, Instagram, or Messenger. And also they claim that every day on average of 2 billion people or even more use at least one of the Facebook family of services. (zephoria.com 2018). Generally all the social media platforms have an authentication method before they activate the user profile. Once account is activated, depending on the platform they can operate in various ways.

In real life the practice is that usually before making a new friend we look for so many things, on the other hand on social media platform we tend to accept requests from unknown profiles as well. There are almost 83 million fake profiles according to the report by CNN. Whether it is fake or real the users are considered as potential customers. The

reasons for creating fake profiles are many. Sometimes it happens from professionals while they are doing research and testing. Therefore spam detection on social networking should be considered with highest priority and appropriate techniques need to be identified. In this study, the researchers suggest and present a spam recognition framework for Social Media Networks.

II. Literature Review

Many researches have conducted since 2008 related to social media spam detection. Researchers have studied various methods to identify spamming behavior in internet. We can observe that majority of these techniques did not produce 100% accuracy in spam detection.

According to various research reports there are no specific way in assessing discrete user spam reports to classify spam messages in SMNs. Use of Machine learning algorithms by some researchers to categorize users or content as resources is a common practice. [1] [2] [3] [4] [5]. Additionally the similar approach is used for categorizing messages based on features which are associated with the contents or features of spammer networks.

DeBarr and Wechsler [2] in their paper 'Using Social Network Analysis for Spam Detection' describes about the use of centrality in the social graph of a social networking site to predict spam detection such as the probability of a user is likely to post spam in a social network. In another research about Twitter, Wang [4] mentioned about another technique which is the use of graph based metrics to improve spam classification on a microblogging platform.

In a study by Mehta et al. [5] presents that simple unsupervised algorithm can be used in spam detection. This algorithm uses statistical properties of effective spam profiles. It says that these properties help to deliver extremely accurate and speedy algorithm for detecting spam. Studies shows that due to the advancement of technology, social networks such as Face book, MySpace, LinkedIn, Friendster [6] and Tickle have large number of members, almost millions, who use them as both social networking as well as business networking. Latest studies are carried out to influence social network into email spam discovery according to the Bayesian likelihood algorithm [9]. The concept this algorithm is to use social relationship between sender and recipient to decide proximity and trust value. The next step is augment or decrease Bayesian probability according to these obtained values.

One of the fact growing social network media, Face book, recommend a spam detection method, which is the Edge Rank algorithm [10]. This algorithm assign each Facebook post with a score produce from few features. These features can be number of likes, number of commentary or number of reposts and so on. Since it is the social media with millions of users accessing it 24/7, Facebook Research team is continuously contributing to the research field by finding solutions to the problems resulting from their users [7].

The Facebook research team also shares to its users the software, platforms, and codes to be downloaded. [8]. Some of the top research fields by Facebook research team are "Applied Machine Learning, Computer Vision, Connectivity, Data Science, Economics & Computation, Facebook AI Research (FAIR), Human Computer Interaction & UX, Natural Language Processing & Speech, Security & Privacy, Systems & Networking and Virtual Reality".

Focus on a content and network information X. Hu has proposed a framework for social spammer detection in his study [13]. In another study to process the challenges about real-time detection of spam and the scalability, X. Jin proposed a General Activity Detection (GAD) clustering algorithm [14]. B. Markines, C. Cattuto, and F. Menczer in their study mainly focused on six features at the social media. These features are mainly in three levels to specify the spam. They are in the levels of post, in resource level, and in user level [15].

While in another study about spam detection, H. Gao analyzed number spam accounts existing in social networks. This is done by identifying the percentage of malicious wall posts. Also using conceded accounts as well as creating accounts for the purpose of spamming [16]. On another study by C. Grier [17] tests the usefulness of URL blacklists. This is done for intercepting the scattering of Twitter spam via the link feature. While in a study of spam detection, M. Bosma proposed a framework combined with user features and spam reports to detect spam [21. K. Thomas et al [17] have analyzed different features and behaviors through the largest spam campaigns in his study on Twitter accounts. While in another study by K. Thomas etc. found that in social media networking, by preventing the spread of compromise in 24 hours it could spare almost 70% of victims [17].

In a study conducted in YouTube to stem social spam, S. Long designed a new methodology combines with three features which are word features, topic features and user-based features [23]. For spam detection in MySpace and Twitter, K. Lee proposed a honeypot-based approach [22]. To test spam in MySpace, J. Caverlee, L. Liu, and S. Webb proposed a reputation-based trust aggregation framework in their research paper. [22]. While in a study in the area of use of social networks academic sector, Yardi et al. [24] identified the ways of learning the behavior of a diminutive part of spammers in Twitter. Also to find that the behavior of spammers is dissimilar from legitimate users in the field of posting tweets, supporters, following associates.

In a research paper by Stringing et al. [25] examine spammer feature via creating a number of honey-profiles in three large social set of connections sites. The social network sites used for the study are Facebook, Twitter and MySpace. This is done with five common features used as potential for spammer detection such as 'follower-to-follower, URL ratio, communication similarity, communication sent and friend digit'. In this research work, even though convincible framework for spammer detection

has been determined but they could not produce detailed approaches of specification and prototype evaluation.
Due to the large number of daily users, social networking data fall into the category of big data and hence several methods could be applied to detect spam patterns using data mining [26]. Following five data mining techniques can be used in this area.
Anomaly detection: This is a method used to detect an 'abnormal behavior' among all typical case data. Associative learning: In this method, typical users those already have a behavior may perform some other more behaviors.
Cluster detection: In this method, data will be clustered in a group using either by similarity base or criteria. Classification: By knowing classification in advance, the given data can be categorize into classes. Regression: This is a prediction model used in data mining. By applying appropriate model in the available historical data, future behaviors can be easily predicted.

In another study by M. Bosma, E. Meij and W. Weerkamp, a framework for spam detection is presented [21]. The framework is modelled based on the HITS web link method and the bipartite graphs. In another study about spam detection by B. Markines, C. Cattuto and F. Menczer, social spam detection is proposed with six features such as 'plagiarism, valid link, number of advertisements, unrelated tags, tag spams and contents of sources'[26].

Researches show that typical data mining technique for spam detections are keyword search and linked-based search [27]. Keyword search mainly considers the documents that match the query best and Term Frequency–Inverse Document Frequency (TFIDF) is calculated. Linked-based ranking approach is for ranking the links and the popular algorithm is using the Page rank or number of HITS [17]. Researches show that spam detection on social networks mainly focuses on 'anomaly detection, fault detection, malware detection and intrusion detection'. If significant effort is not taken to find a proper technological solution to the threat of spam, the communication media such as email as well as social media applications will be in danger as an important medium of communication [18]. Spam detection on social networking has become a serious problem globally due to the global reach of applications and advancement of social media usage.

According to the research findings show that the current state of spam is increasing and more effort that is rigorous are required to control and stop them in an effective and efficient manner. It has been observed and reported that 75.9% of email messages are spam, and social networks are the most vulnerable attacks [19]. By reviewing the literature research, we can observe that a large number of classifiers have been used in spam detection. However, choosing the right classifier and the most efficient combination of them is still problem. In this paper, the researchers identified a classifier to detect spam in social media and present a framework using the identified classifier.

III. Summary of Related Researches

After the literature research, the researchers have concluded that there are many works have been conducted about social spam detection; however, most previous work on social spam has concentrated on the methods and techniques for spam detection and prevention on a single social network. It has been identified that these works are done either for Facebook [11, 22] or MySpace [16] or Twitter [4]. To understand the approach it is noted that various methods have been discussed in these research such as collaborative filtering, friend graph analysis, classification, behavioral analysis etc. Authors have taken into consideration the key findings from the previous researches while proposing the new framework. Different classifiers proposed by various researchers have been tested in spam detection earlier and it is found that it is a big challenge to choose the right one for the same purpose. Previous work by Byungki et al. [5] proposes a Bayesian framework. The framework uses a method of investigating the integration of text and image classifiers. This method has considered theoretically efficient and practically reasonable method of combination

In cross-domain text classification, several novel classification approaches are proposed and implemented by various researchers. In a paper, Pu Wang et al. [17] presented 'semantics based algorithm' for cross-domain text classification using Wikipedia based on clustering classification algorithm. Elisabeth Lex et al. [20] described a novel and efficient 'centroid based algorithm', which is known as Class Feature Centroid Classifier (CFC) for cross-domain classification of weblogs. The research also discussed the tradeoff between complexity and accuracy in applying the method. Kurt Thomas et al. [15] in a study proposed a URL spam filtering technique to address diff erent web services similar to social network services. This study presented a real time URL spam-filtering system named 'Monarch' and it demonstrated a way of deployment of web services on cloud infrastructure.

IV. The Suggested Framework

After considering all the researches and discussing various models, authors have tried to propose concise framework of the best proposed classifier, which explains specifications of every step and the tweaking target, additionally it provides the concluding fulfilled classifier. Overall framework for the proposed classifier is discussed in Fig. 1 along with discussing individual stage.

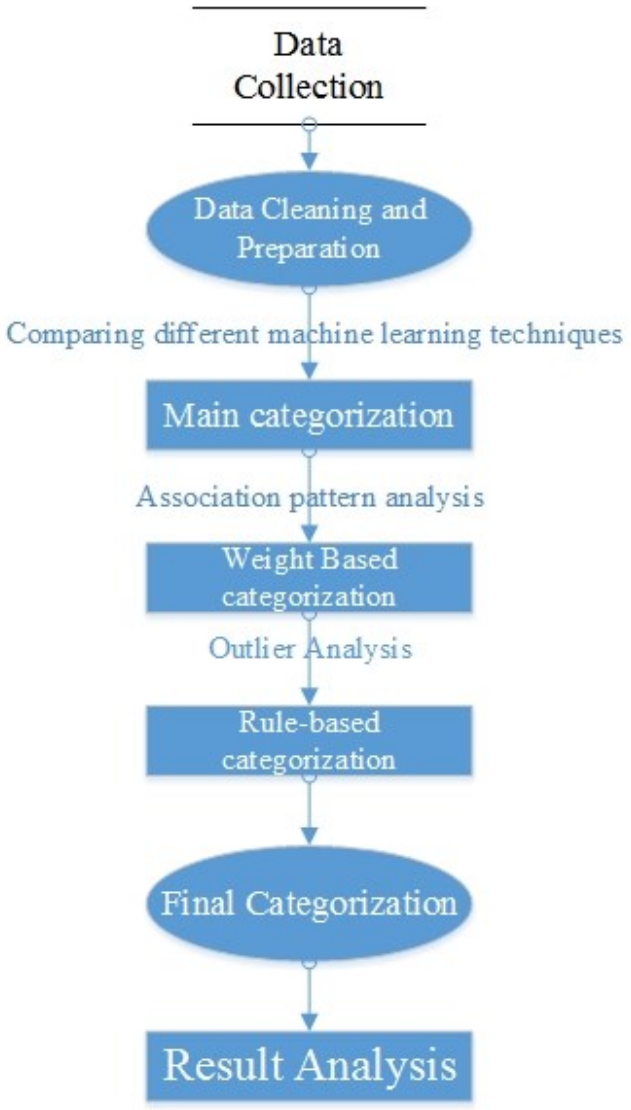

Fig. 1 Proposed Framework

After the acquisition of data using appropriate methods the next step is data Cleaning using suitable techniques and preparation. Various machine learning techniques such as classification or regression to be compared and select the appropriate algorithm to proceed. The next stage is to proceed with the categorization of the data. Starting with the main categorization and then using association pattern analysis the next stage can be reached which is weight based categorization. A Frequent pattern is defined as a pattern that happens recurrently in a data set. Association rule mining can be applied in such patterns to find all frequent patterns. Association patterns will produce strong rules from frequent patterns. Within a given dataset rules can be find that will predict the occurrence of an item based on the occurrences of other items in the data flow, in this case, the spam data flow in a social network. The next stage rule-based categorization is reached by using Outier Analysis. An Outlier can be defined as exceptional chance of occurrence within the given data set. Outier is usually happen possibly due to variability in the measurement and sometimes it indicates experimental error also. The last stage is the final categorization where the spam is detected and it is easy to identify mechanisms to remove the spams.

A. The difference between the existing system and the proposed Framework

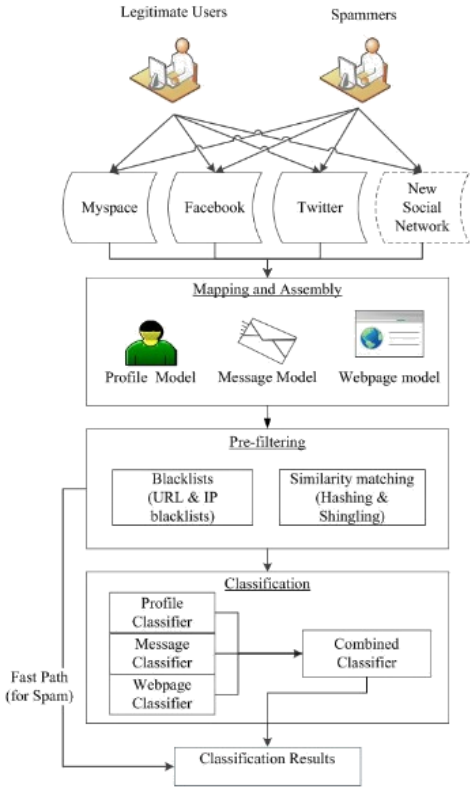

Fig. 2 Spam detection framework: Overview [21]

Some URLs were assessed manually and it was detected to be wrongly categorized as spam although they were genuine users. Testing was done again after enhancing the classifier and white-listing.

Fig. 3 Feedback mechanism [22]

The above framework also includes a response tool through which the authorization can be taken from the user, considering a case where the spam-labeled tweet is wrongly identified as spam, then the proposed model will rectify its label for the tweet in question and also all the similar tweets will be rectified. The proposed model architecture is summarized in Fig.3.

V. EXPECTED BENEFITS/ ADVANTAGES

The technological development with the new technique frequently necessitates an incessant re-search along with augmentation of techniques which is primarily involved in identifying the spammers, but interestingly you can't have a single trustful algorithm which is 100% correct in terms of handling of humanoid behavior.

Since twitter makes all its content available publically by default, it has become the commonly researched platform, in

comparison to the Facebook. Although Facebook is taking over in terms of number of users when compared to other SMN platform. And so few only have administered practices with it.

Facebook research team, as a part of their regular development, has achieved a tremendous amount improvement in the spam discovery perspective. This is indeed helping Facebook to retain the top spot amongst all SMN platforms, with increasing number of users every day.

VI. CONCLUSION

In this research paper, the authors proposed and presented a spam detection framework to find out spam on more than one social network. Authors have suggested that the proposed framework can be applied to multiple social networks to detect spam. Proper experiments can be done to demonstrate the efficiency of the framework. Further to add in the research authors plan to utilize live feeds from SMNs on testing and evaluating the proposed framework. In addition, as a future work, detecting spammers' behavior and integrating it to the framework will be considered. It is worthy to note that as the technology development is continuous process, so is the process of intriguing new spam techniques. No research can assure full guarantee over developing a fool proof system but it can be minimized. More research is required in this area though, with datasets from other SMNs as well, which can be a scope for further research.

VII. BIBLIOGRAPHY

[1] Benevenuto, F., Rodrigues, T., Almeida, V., Almeida, J., Zhang, C., Ross, K.: Identifying video spammers in online social networks. In: Proceedings of the 4th International Workshop on Adversarial Information Retrieval on the Web, pp. 45–52. ACM (2008)

[2] DeBarr, D., Wechsler, H.: Using Social Network Analysis for Spam Detection. In: Chai, S.-K., Salerno, J.J., Mabry, P.L. (eds.) SBP 2010. LNCS, vol. 6007, pp. 62–69. Springer, Heidelberg (2010)

[3] Irani, D.,Webb, S., Pu, C.: Study of static classification of social spam profiles in myspace. In: Proceedings of the 4th International Conference on Weblogs and Social Media (2010)

[4] Wang, A.H.: Don't follow me: Spam detection in twitter. In: Proceedings of the 2010 International Conference on Security and Cryptography (SECRYPT), pp. 1–10. IEEE (2010)

[5] Mehta, B., Hofmann, T., Fankhauser, P.: Lies and propaganda: detecting spam users in collaborative filtering. In: Proceedings of the 12th International Conference on Intelligent User Interfaces, pp. 14–21. ACM (2007)

[6] Danah, M.B., 2004, Friendster and publicly articulated social networking. Proceedings of Extended
Abstracts on Human Factors in Computing Systems, (CHI '04), Vienna, Austria, pp: 1279-1282. DOI:10.1145/985921.986043

[7] Hua Shen, Fenglong Ma, Xianchao Zhang, Linlin Zong, Xinyue Liu, and Wenxin Liang. Discovering social spammers from multiple views. Neurocomputing, 225:49–57, 2017.

[8] Mahdi Washha, Aziz Qaroush, Manel Mezghani, and Florence Sedes. A topic based hidden markov model for real-time spam tweets filtering. Procedia Computer Science, 112:833–843, 2017.

[9] Trust evaluation based content filtering in social interactive data, in: Proceed-ings of the 2013 International Conference on Cloud Computing and Big Data (CloudCom-Asia), IEEE, 2013, pp. 538–542.

[10] J. Kincaird, Edgerank: the secret sauce that makes Facebook's news feed tick, TechCrunch, 2010, http://techcrunch.com/2010/04/22/facebook-edgeran⟩.

[11] S. Yardi, D. Romero, G. Schoenebeck, Detecting spam in a Twitter network, First Monday 15 (1) (2009).

[12] G. Stringhini, C. Kruegel, G. Vigna, Detecting spammers on social networks, in: Proceedings of the 26th Annual Computer Security Applications Conference, ACM, 2010, pp. 1–9.

[13] X. Hu, J. Tang, and H. Liu. "Online social spammer detection," In 28th AAAI Conference on Artificial Intelligence, 2014.

[14] Brian, "Five main methods of detecting patterns in data mining," [Online]. Available: http://legallysociable.com/2012/04/05/five-main-methods-of-detecting-patterns-in-data-mining/. [Accessed 30 June 2012].

[15] M. Bosma, E. Meij and W. Weerkamp, "A Framework For Unsupervised Spam Detection In Social Networking Sites," 2012.

[16] B. Markines, C. Cattuto and F. Menczer, "Social Spam Detection," in AIRWeb, 2009.

[17] J. Pei, B. Zhou, Z. Tang and D. Huang, Data Mining Techniques for Spam Detection.

[18] Ismaila.I; Ali.S.,"Improved email spam detection model with negative selection algorithm and particles swarm optimization. In: Proceeding of Applied Soft Computing"; Applied Soft Computing 22 (2014) 11-27

[19] Ahmed.F., Abulaish,M., ." A generic statistical approach for spam detection in Online Social Networks". Computer Communications 36 (2013) 1120-1129. Science direct (Elsevier).

[19] Ahmed.F., Abulaish,M., ." A generic statistical approach for spam detection in Online Social Networks". Computer Communications 36 (2013) 1120-1129. Science direct (Elsevier).

[20] Byun.B;Lee.C;Webb.S;Irani.D; and Pu.C." An anti-spam filter combination framework for text-and-image emails through incremental learning" : In Proceedings of the Sixth Conference on Email and Anti-Spam (CEAS).2009.

[21] De Wang, Danesh Irani, and Calton Pu. A social-spam detection framework. In Proceedings of the 8th Annual Collaboration, Electronic Messaging, Anti-Abuse and Spam Conference, pages 46–54. ACM, 2011.

[22] Juan Martinez-Romo and Lourdes Araujo. Detecting malicious tweets in trending topics using a statistical analysis of language. Expert Systems with Applications, 40(8):2992–3000, 2013.