Spam Detection and Fake User Identification in Twitter: An Analysis of Machine Learning Models

Radhika S[1], Phani Sri Siva Charan K[2], Sai Khush Kumar M[3],
Ramanjaneya Reddy K[4], Manish Prasad[5]
Associate Professor, Department of Computer Science and Engineering[1]
Students, Department of Computer Science and Engineering[2,3,4,5]
Raghu Institute of Technology, Visakhapatnam, AP, India

Abstract: *Spammer detection and fake user identification are significant issues in the realm of social media, with Twitter being no exception. The detection of spammers and fake users on Twitter is essential for preserving the platform's integrity and protecting users from online scams and fraud. This project paper aims to conduct a comprehensive study of the different techniques and algorithms used for spammer detection and fake user identification on Twitter. We will evaluate the effectiveness of traditional techniques and machine learning-based methods and propose a novel approach combining both methods for spammer detection and fake user identification on Twitterpils.*

Keywords: Social Networks, Spammer Detection, Fake User Identification, Machine Learning, Decision Trees, Random Forests.

I. INTRODUCTION

Using the internet to get all kinds of information from anywhere in the world has become an incredible thing. The increasing demand of social networking sites allows users to collect more data and information about users. The large amount of information available on these sites has also attracted fraudulent users [1]. Twitter has quickly become an online resource for information about its users. Twitter is an Online Social Network (OSN) where users can share anything from news, thoughts and even their own thoughts.

There can be many discussions on different topics such as politics, current affairs and important events. When a user tweets, this is immediately announced to his followers and they are provided to publish the information they receive on a wider level [2]. With the development of OSN, the need to study and analyze user behavior on online social platforms has become stronger. Many people who don't know much about OSNs are easily fooled by scammers. Also, there is a need to attack and control OSN users just for advertising and thus spamming other people's accounts.

Recently, the discovery of spam on social networking sites has attracted researchers. Spam detection is a dangerous task in social network management. Spam on the OSN site should be identified to defend end users from various malicious activities and protect their security and privacy. These dangerous tactics used by spammers can do serious harm to real communities. Twitter spammers have many purposes, including spreading misinformation, fake news, rumors, and false statements.

Spammers reach their nasty targets through advertising and many other methods, promoting different names and randomly spreading their interests after spamming. These activities are vulnerable to legacy clients who are not known to be spammers. It also lowers the reputation of the OSN platform. Therefore, it is necessary to establish a strategy to encounter hackers so that action can be taken against their crimes [3]. A lot of research has been done to investigate Twitter spam.

Recently, the discovery of spam on social networking sites has attracted researchers. Spam detection is a dangerous task in social network management. Spam should be identified on the OSN site to protect users from various malicious activities and protect their security and privacy. These dangerous tactics used by spammers can do serious harm to real communities. Twitter spammers have many purposes, including spreading misinformation, fake news, rumors, and false statements.

DOI: 10.48175/IJARSCT-9178

There were also some calls from Twitter for fake customers to close the status of the current image. Ting Min et al. [4] explores new methods and techniques for detecting Twitter spam. The research described above provides a comparative study of the current system. On the other hand, the authors in [5] explored the different behaviors presented by spammers on the Twitter social network.

Spammers achieve their malicious goals by advertising and in many ways promote different domains and then randomly send spam to expand their interests. These activities are vulnerable to legacy clients who are not known to be spammers. It also lowers the reputation of the OSN platform. Therefore, it is necessary to establish a strategy to detect spammers so that action can be taken against their crimes [3]. A lot of research has been done to investigate twitter spam.

This study also provides data analysis that confirms the presence of spammers on the Twitter social network. Despite all the research, there are still gaps in the available literature. Therefore, we review technical methods for detecting Twitter spammers and identifying fake users to close the gap. In addition, this survey presents the taxonomy of Twitter spam detection methods and tries to explain in detail the latest developments in this field.

To cover the latest technology available, some research has also been done on fake users from Twitter. Ting Min et al. [4] explores new methods and techniques for detecting Twitter spam. The research described above provides a comparative study of the current system. On the other hand, the authors in [5] explored the different behaviours presented by spammers on the Twitter social network.

For classification, we identified four methods for advertising spammers that help identify false users. Spammers can count on: (i) fake content, (ii) URL-based spam detection, (iii) trending topics spam detection, and (iv) fake client identification. Table 1 provides a comparison of the latest technology that can help users understand the importance and effectiveness of the plan, as well as provide a comparison of objectives and results. Table 2 compares the different features used to detect spam on Twitter. We hope this research will help readers find more information about the search spammer at some point.

This article is based on Part II, which presents the taxonomy of spammer search strategies on Twitter. Part III discusses a comparison of suggested methods for catching spammers on Twitter. Part IV presents the overall analysis and discussion, while Part V concludes the article and offers some suggestions for future work.

For classification, we identified four methods for advertising spammers that help identify false users. Spammers can count on: (i) fake content, (ii) URL-based spam detection, (iii) trending topics spam detection, and (iv) fake client identification. Table 1 provides a comparison of the latest technology that can help users understand the importance and effectiveness of the plan, as well as provide a comparison of objectives and results. Table 2 compares the different features used to detect spam on Twitter. We hope this research will help readers find more information about the search spammer at some point.

This article is based on Part II, which presents a taxonomy of spammer search strategies on Twitter. A comparison of suggested methods for detecting spammers on Twitter is discussed in Chapter III. Part IV presents the overall analysis and discussion, while Part V concludes the article and offers some suggestions for future work.

II. SPAMMER DETECTION ON TWITTER

In this article, we describe a classification of spam detection techniques. Figure 1 shows the taxonomy for identifying spammers on Twitter. The registration data is divided into four main categories:

(i) fake content, (ii) URL-based spam search, and (iii) fake user identification. Each authentication type is based on certain criteria, methods and search algorithms.

The first category (fake content) includes various techniques such as replication testing, malware alerts, and Lfun scheme methods. In the second category (URL-based spam detection), spammers are identified in URLs by different machine learning methods. The third category (spam on different topics) was analyzed by Naive Bayes classifiers and different languages. The last category (the user reporting a fake account) is based on the hybrid method of detecting fake customers. The procedures associated with each identified spammer group are discussed in the following columns.

2.1 Fake Content Based Spammer Detection

Gupta et al. [6] provides an in-depth explanation of the products affected by the increasing bad content. Many people with social media profiles were found to be responsible for the spread of fake news. To identify fake accounts, the authors selected accounts created immediately after the Boston bombing and subsequently banned by Twitter for violations and incidents. engine 7.

9 million unique tweets written by 3.7 million unique users. This document is known as the largest document on the Boston bombing. Authors categorize fake content by focusing on the time distribution of tweets is calculated based on the number of tweets posted per hour. 4,444 fake tweet accounts were analyzed based on the activity of spam-generated accounts.

It has been observed that most of the fake tweets are shared by people who follow them. The source of tweet analysis is determined by the medium that published the tweet. It was determined that most of the tweets containing information were produced from mobile devices, while the tweets that did not contain information were mostly produced from the relevant website. The role of user behavior in identifying fake content is calculated as follows:

(i) average number of spam or non-spam approved accounts, and (ii) user account follower count. Publishing fake content is defined by the following indicators: (i) social reputation, (ii) international cooperation, (iii) collaboration, (iv) affiliation, and (v) credibility.

After that, the authors used a regression model to ensure that all effects of ad content were false at that time and predicted future growth of false content.

Koni et al. [7] proposed a method to provide malicious alerts using a series of real-time tweets captured by the Twitter API. Then, a group of tweets thinking about the same topic is collected to create a notification. Strategic planning is used to evaluate Twitter tweets, analyze the progress of authoritative events, and report the events themselves.

Table 1: Comparison between proposed methods for spam detection in Twitter.

Ref.	Proposed Method	Goal	Data Set	Result
[15]	The Dirichlet distribution was used statistically to identify spammers on Twitter.	Separate spammers from non-spammers	Real Twitter data	Research using Twitter data demonstrates that both supervised and unsupervised algorithmic methods provide useful results.
[16]	An efficient unified waiting system for URL anomaly detection	detection of abnormal user interaction behavior	It uses a Twitter dataset with users' most recent 200 tweets.	The number of URL spammers that are used daily may be successfully analysed using anomalous detection.
[2]	Classification of users as spammers and non-spammers via manual inspection	Spammer discovered on Twitter	1.8 billion tweets and more than 1.9 billion links are part of the Twitter dataset.	Spammers are classified using a broad range of criteria, making the process substantially more resistant to them and their spamming tactics.
[17]	Three different types of cascade information—TSP, SS, and cascade filtering—that are produced using a spam detection algorithm have been deployed.	Using the characteristics of social networks in the specific social environment, spammers have been categorised.	Real Twitter dataset	Instead of looking at the entire network, the scalable techniques assess users' cantered 2-hop social networks.

DOI: 10.48175/IJARSCT-9178

[18]	Design of 18 resilient characteristics with explicit and implicit time properties	Just respond to the query of how to identify spammer	Dataset that was crawled and manually annotated	The retrieved traits can distinguish between legitimate users and spammers with an accuracy of up to 93%.
[7]	For the purpose of detecting Twitter spammers, an inductive e-learning methodology has been applied.	To determine the best feature to identify Twitter spammers, user behaviour and tweet content have been examined.	Crawlers have been used to identify spammers using a collection of 62 features	With a detection accuracy that surpasses results reported in the existing literature, the random-forest technique offers adequate results in the detection of spam users.
[19]	Four separate feature sets were used in a text pre-processing technique to test the spam and non-spammer classifiers.	The study's goal is to identify spam tweets, which increase the amount of data that must be gathered by relying just on a tweet's inherent features.	2 significant labeled datasets of spam tweets.	The limited feature set available in tweets, which is superior to the current spammer identification methods, was used to obtain an impressive result.
[21]	Edge weighting and centrality weighting, two trials, were run.	To recognize the weight that might enable a more accurate opinion based on assessment algorithms and to recognize the significance of each well-defined edge in order to identify the opinion leader.		Thus, compared to other evaluation algorithms to identify the opinion leader, the low in-degree weight, high betweenness weight, and low or no PageRank weight may provide 100% agreement.
[9]	For the objective of determining the performance of detection and strength based on enormous volumes of truth data, a wide variety of traditional machine learning methods are used.	The study aims to develop real-time Twitter spam detection tools.	The ground truth data set was produced using a random sample of about 30 million labelled tweets.	In a real-world setting, the Lfun technique can considerably improve spam detection accuracy.
[1]	On numerical features, the entropy minimization discretization (EMD) technique was employed.	Using Twitter's Naive Bayes algorithm as the foundation, a classification method effect of discretization is proposed in order to identify fraudulent	Since there is no publicly accessible dataset, we developed our own dataset using the Twitter API.	Compared to continuous values, naive Bayes can perform well with discrete values.

		accounts.		
[13]	The combination of user-based, content-based, and graph-based elements together has led to the development of a hybrid technique for the detection of spammers on Twitter.	Combining user-based and graph-based features will increase the accuracy of spam profile detection.	400k tweets and 11k users from a Twitter data set were used.	In comparison to any existing technique, the study's detection rate is higher and more precise.
[6]	In order to demonstrate the influence of people who circulate fraudulent content, regression prediction models have been utilised.	To categorise and make recommendations for how to combat distinct kinds of spam occurrences that occur on Twitter during events like Boston Blast.	Twitter API was used to extract almost 7.8 million tweets about the Boston Marathon bombing.	During the Boston bombing incident, 29% of the Twitter information that went viral was phoney. The remaining 49% contained accurate information, while 51% were generic opinions and remarks.

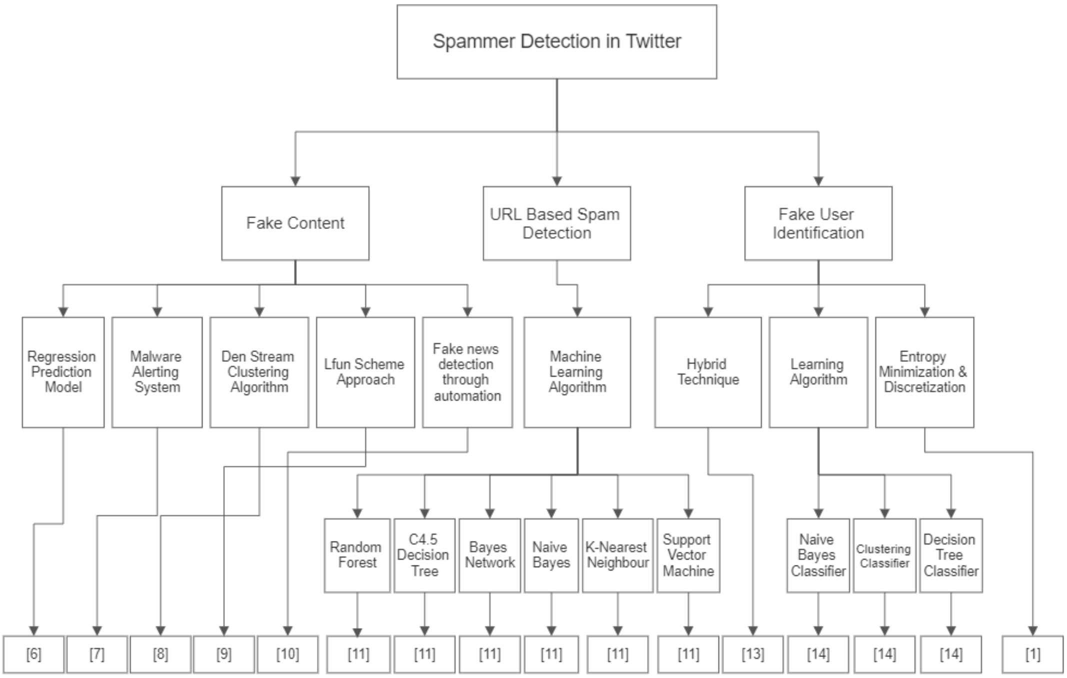

FIGURE 1. Taxonomy of spammer detection/fake user identification on Twitter.

The plan uses information contained in tweets when the user identifies spam or malware, or when a security warning is issued by the certificate. The warning message includes: (i) real-time data extraction from tweets and users, (ii) pre-planning and filtering based on Naive Bayesian algorithm to remove false information tweets, (iii) sigmoid of spammers data analysis detection window (iv) notification subsystem used when an event is generated, the system sends tweets about the same topic, among which tweets are different from places in the group, tweets closest to the group represent the whole system and (v) comments. The process is guaranteed to be efficient and effective to detect

DOI: 10.48175/IJARSCT-9178

some influences and compliments in circulation. In addition, Eshraqi et al. [8] considers different characteristics to identify spam and then identifies spam tweets with the help of stream-based clustering algorithm.

Several user accounts are selected from a large database and random tweets are selected from these accounts. These tweets are then classified as spam, not spam. The authors claim that the algorithm can classify information as spam and non-spam with high accuracy and identify fake tweets with accuracy. Many factors can be used to identify spam. For example, graph-based features are cutting-edge social networks that are graphically similar to Twitter.

If the number of followers is less than the number of followers, the reliability of the account decreases and the probability of spam accounts is high. Similarly, context-based content includes tweet reputation, HTTP links, mentions and replies. In terms of physical characteristics, if the user's account sends a lot of tweets in a short time, it is a spam account. The data of this study consisted of 50,000 users. The system can detect spammers and fake tweets with a high degree of accuracy.

Chen et al. Twitter has proposed a Lfun (Learn from Untagged Tweets) strategy to solve various problems in spam detection. [9]. Their framework consists of two parts, learning from perceived tweets (LDT) and learning from human tags (LHL).

These two elements are used to retrieve spam tweets from a series of anonymous tweets easily collected by the Twitter network. When spam tweets are received, they are sorted together using a random forest algorithm. The effectiveness of the system is measured by the detection of spam tweets. Tests were run on real-world data for ten consecutive days, with 100,000 tweets per day for both spam and non-spam. F-test and test value were used to evaluate the effectiveness of the proposed method.

The results of the proposed method show that the method is effective in detecting spam in the real world. Also, Buntain et al. [10] proposed a method for detecting fake news on Twitter by estimating the accuracy rate in two trust-centered datasets. The technique was applied to Twitter's fake news data, and the model was trained against a large number of users based on journalistic analysis. Two Twitter datasets are used to examine the integrity of OSNs.

The first database, CREDBANK, is a crowd sourced database for analyzing the authenticity of events on Twitter, while the second database, called PHEME, is a journalistic database of rumors on Twitter and tracking news. A total of 45 attributes were disclosed, divided into four categories: attribute attributes, user attributes, content attributes, and physical attributes. Alignment tags in PHEME and BUZZFEED have classes that describe whether the story is true or false. The results of the analysis help probe social media data to see if stories support similar patterns.

2.2 URL based Spam Detection

Chen et al. [11] evaluated machine learning algorithms to detect spam tweets. The authors examined the impact of various spam detection features, such as (i) spam to non-spam ratio, (ii) report size, (iii) time dependent data, (iv) factor analysis. and (v) data sampling. Nearly 600 million public tweets were collected for the first time to test the findings, and the authors used Trend Micro's website reputation to identify as many spam tweets as possible. Through this analysis, a total of 12 light weights are also classified to distinguish non-spam tweets from spam tweets.

The properties of the analyzed features are represented by the cdf plot.

These features are well known for machine learning-based spam classification and are then used in experiments to test spam detection. Four datasets were used to replicate the scenarios. Since there is no publicly available information for this project, the information is rarely used in previous studies. After analyzing the spam tweets, 12 features were collected.

These features fall into two categories, user-based features and tweet-based features. User-based features are defined by various factors such as account age and number of user likes, lists, and tweets. User-defined attributes are defined in JSON format.

On the other hand, tweet-based features include (i) retweet, (ii) hashtag, (iii) mention and (iv) URL count. The test results show that the variable feature distribution reduces the performance, while there is no difference in the training data set distribution.

2.3 Fake User Identification

A classification method was proposed by Erşahin et al. [1] Search for spam accounts on Twitter. The data used in this study were collected manually. Classification is done by analyzing usernames, profile and background pictures, friends and followers, tweet content, account descriptions and tweet counts. The database contains 501 fake numbers and 499 real numbers, with 16 attributes identified from data from the Twitter API.

Two tests are performed to detect counterfeit money. The first experiment uses Naive Bayes learning on the Twitter dataset including all discretized items, while the second experiment uses Naive Bayes learning on the discretized Twitter data.

Martin et al. [13] proposed a hybrid method that uses user-based content and graphics to identify spammer profiles. A model using three features is proposed to distinguish between non-spam and spam profiles.

The proposed system was analyzed using a Twitter dataset containing 11,000 users and approximately 400,000 tweets. The goal is to combine all these features to achieve better performance and accuracy. User-based features are created based on social and user account features. User-based features should be added to spam detection models. Because these activities involve user accounts, all behaviors associated with user accounts are defined.

These attributes include followers and followers, age, FF rate, and reputation. Instead, content attributes are linked to tweets that users post as spam bots, with more duplicate content tweets than non-spam posting duplicate tweets.

This function depends on user typed words or content. Spammers spread fake news by posting content with malicious URLs to promote their products. Content-based attributes include: (i) total tweet count, (ii) hashtag rate, (iii) URL rate, (iv) mention rate, and (v) tweet frequency.

Graph-based features are used to control the avoidance tactics that spammers do. Spammers use different techniques to stay undetected. They can buy fake followers from different third-party websites and transfer their followers to other users to become legitimate users. Graph-based features include input/output rating and averaging. Since no data is published due to Twitter's policy, the evaluation of the method was made using the data obtained from the previous drawing.

Results were evaluated with a combination of three methods, refined Naive Bayes and J48. The experimental results show that the detection of this method is higher than the current technology and its accuracy is high. Gupta et al. [14] proposed a strategy for catching spammers on Twitter and using a popular technique, namely, Naive Bayes, clustering and decision trees. The algorithm classifies accounts as spam or not spam. The database contains 1064 Twitter users and contains 62 user-specific or tweet-specific attributes. Spammer accounts for approximately 36% of data usage. Because spammers behave differently from non-spammers, certain behaviors or characteristics are defined where the two groups differ from each other.

Specific recognition depends on the number and tweet level of the user, such as followers or unfollowers, spam keywords, replies, hashtags and URLs [30] , [32]. After the feature has been implemented, the first step is to change everything else to the random feature. Later, the authors developed methods using clustering, decision trees, and Naive Bayesian algorithms. With Naive Bayes, accounts are identified by predicting whether certain accounts are spamming. In a group-based algorithm, the entire group of accounts is divided into spam and non-spam groups.

In the decision tree algorithm, the structure of the tree is created and decisions are made at each level of the tree. The results of the proposed method show that the combined algorithm is more effective in identifying non-spam accounts than finding spam accounts. The results of these integrated algorithms demonstrate the complete accuracy and effective detection of spam-free individuals.

III. MACHINE LEARNING ALGORITHMS USED

In the project of detecting spammers on Twitter, machine learning algorithms play a crucial role in identifying patterns and features of spammers. Here are some of the machine learning algorithms that can be used for Twitter spam detection:

1. **Support Vector Machine (SVM):** SVM is a supervised machine learning algorithm that can be used for binary classification. SVMs are effective in high-dimensional spaces and can efficiently handle large datasets. In the context of Twitter spam detection, SVM can be used to classify a user as a spammer or a legitimate user based on their profile information, tweets, and network features.

DOI: 10.48175/IJARSCT-9178

2. **Random Forest:** Random Forest is a supervised machine learning algorithm that uses an ensemble of decision trees to perform classification. Random Forest is known for its high accuracy and ability to handle noisy datasets. In Twitter spam detection, Random Forest can be used to identify spammers based on their behavior, such as the frequency of tweets, the time of posting, and the content of the tweets.
3. **Naive Bayes:** Naive Bayes is a probabilistic machine learning algorithm that can be used for binary classification. Naive Bayes is simple and fast, making it ideal for real-time applications such as Twitter spam detection. Naive Bayes can be used to classify users based on their profile information, tweets, and network features.
4. **Decision Tree:** Decision Tree is a supervised machine learning algorithm that can be used for classification and regression tasks. Decision Tree is easy to understand and interpret, making it useful for identifying features that are important for spam detection. In the context of Twitter spam detection, Decision Tree can be used to classify a user as a spammer or a legitimate user based on their tweet content and network features.

IV. LITERATURE REVIEW

Several studies have been conducted to identify spammers and fake users on Twitter. One common approach is to analyze user activity patterns, such as the frequency of tweets, the number of followers, and the engagement rate. These features help distinguish genuine users from fake users as spammers and fake users tend to have high activity levels and low engagement rates.

Another approach is to analyze the social network structure, including the user's connections and interactions with other users. For example, a user with a large number of followers but few interactions with other users may be classified as a fake user. Additionally, the type of content shared by users can also be analyzed to identify spammers and fake users. For instance, spammers often post repetitive content with a high level of promotion, whereas genuine users tend to post diverse content that reflects their interests and personality.

Machine learning algorithms have also been employed for spammer detection and fake user identification on Twitter. Decision trees, random forests, and neural networks are commonly used for classification as they can handle large datasets and generate accurate results. Decision trees are particularly useful for identifying the most relevant features for classification. Random forests can handle missing data and noisy data, and neural networks can learn complex patterns and relationships.

V. CONCLUSION

In this article, we examine the techniques used to detect spammers on Twitter. Additionally, we propose a taxonomy of Twitter spam detection methods and group them into fraudulent content detection, spam detection on different topics, and fraudulent recruitment strategies. We also compare planning strategies based on various attributes such as user attributes, content attributes, display attributes, physical attributes, and physical attributes. In addition, these strategies are compared according to the stated objectives and the data used. The proposed review should help researchers find information on the most advanced Twitter spam detection techniques on the map.

Despite the development of efficient and effective methods for spam detection and user identification on Twitter [34], there are some open areas that require further attention by researchers. These problems are briefly explained as follows: Detection of fake news on social media is a problem that needs to be investigated because these news stories can have a significant impact on both individuals and groups [25]. Another important issue worth investigating is identifying the sources of gossip on social media. Although some studies based on statistical methods have been done to investigate the source of the rumors, more polished approaches, ex: social media , as a social network, can be proven useful.

REFERENCES

[1]. B. Erçahin, Ö. Aktaş, D. Kilinç, and C. Akyol, ‘‘Twitter fake account detection,’’ in Proc. Int. Conf. Comput. Sci. Eng. (UBMK), Oct. 2017, pp. 388–392.

[2]. F. Benevenuto, G. Magno, T. Rodrigues, and V. Almeida, ‘‘Detecting spammers on Twitter,’’ in Proc. Collaboration, Electron. Messaging, Anti- Abuse Spam Conf. (CEAS), vol. 6, Jul. 2010, p. 12.

[3]. S. Gharge, and M. Chavan, ''An integrated approach for malicious tweets detection using NLP,'' in Proc. Int. Conf. Inventive Commun. Comput. Technol. (ICICCT), Mar. 2017, pp. 435–438.

[4]. T. Wu, S. Wen, Y. Xiang, and W. Zhou, ''Twitter spam detection: Sur- vey of new approaches and comparative study,'' Comput. Secur., vol. 76, pp. 265–284, Jul. 2018.

[5]. S. J. Soman, ''A survey on behaviors exhibited by spammers in popular social media networks,'' in Proc. Int. Conf. Circuit, Power Comput. Tech- nol. (ICCPCT), Mar. 2016, pp. 1–6.

[6]. A. Gupta, H. Lamba, and P. Kumaraguru, ''1.00 per RT #BostonMarathon # prayforboston: Analyzing fake content on Twitter,'' in Proc. eCrime Researchers Summit (eCRS), 2013, pp. 1–12.

[7]. F. Concone, A. De Paola, G. Lo Re, and M. Morana, ''Twitter analysis for real-time malware discovery,'' in Proc. AEIT Int. Annu. Conf., Sep. 2017, pp. 1–6.

[8]. N. Eshraqi, M. Jalali, and M. H. Moattar, ''Detecting spam tweets in Twitter using a data stream clustering algorithm,'' in Proc. Int. Congr. Technol., Commun. Knowl. (ICTCK), Nov. 2015, pp. 347–351.

[9]. C. Chen, Y. Wang, J. Zhang, Y. Xiang, W. Zhou, and G. Min, ''Statistical features-based real-time detection of drifted Twitter spam,'' IEEE Trans. Inf. Forensics Security, vol. 12, no. 4, pp. 914–925, Apr. 2017.

[10]. C. Buntain and J. Golbeck, ''Automatically identifying fake news in popu- lar Twitter threads,'' in Proc. IEEE Int. Conf. Smart Cloud (SmartCloud), Nov. 2017, pp. 208–215.

[11]. C. Chen, J. Zhang, Y. Xie, Y. Xiang, W. Zhou, M. M. Hassan, A. AlElaiwi, and M. Alrubaian, ''A performance evaluation of machine learning-based streaming spam tweets detection,'' IEEE Trans. Comput. Social Syst., vol. 2, no. 3, pp. 65–76, Sep. 2015.

[12]. G. Stafford and L. L. Yu, ''An evaluation of the effect of spam on Twitter trending topics,'' in Proc. Int. Conf. Social Comput., Sep. 2013, pp. 373–378.

[13]. M. Mateen, M. A. Iqbal, M. Aleem, and M. A. Islam, ''A hybrid approach for spam detection for Twitter,'' in Proc. 14th Int. Bhurban Conf. Appl. Sci. Technol. (IBCAST), Jan. 2017, pp. 466–471.

[14]. A. Gupta and R. Kaushal, ''Improving spam detection in online social net- works,'' in Proc. Int. Conf. Cogn. Comput. Inf. Process. (CCIP), Mar. 2015, pp. 1–6.

[15]. F. Fathaliani and M. Bouguessa, ''A model-based approach for identifying spammers in social networks,'' in Proc. IEEE Int. Conf. Data Sci. Adv. Anal. (DSAA), Oct. 2015, pp. 1–9.

[16]. V. Chauhan, A. Pilaniya, V. Middha, A. Gupta, U. Bana, B. R. Prasad, and S. Agarwal, ''Anomalous behavior detection in social networking,'' in Proc. 8th Int. Conf. Comput., Commun. Netw. Technol. (ICCCNT), Jul. 2017, pp. 1–5.

[17]. S. Jeong, G. Noh, H. Oh, and C.-K. Kim, ''Follow spam detection based on cascaded social information,'' Inf. Sci., vol. 369, pp. 481–499, Nov. 2016.

[18]. M. Washha, A. Qaroush, and F. Sedes, ''Leveraging time for spammers detection on Twitter,'' in Proc. 8th Int. Conf. Manage. Digit. EcoSyst., Nov. 2016, pp. 109–116.

[19]. B. Wang, A. Zubiaga, M. Liakata, and R. Procter, ''Making the most of tweet-inherent features for social spam detection on Twitter,'' 2015, arXiv:1503.07405. [Online]. Available: https://arxiv.org/abs/1503.07405

[20]. M. Hussain, M. Ahmed, H. A. Khattak, M. Imran, A. Khan, S. Din,

[21]. A. Ahmad, G. Jeon, and A. G. Reddy, ''Towards ontology-based multilin- gual URL filtering: A big data problem,'' J. Supercomput., vol. 74, no. 10, pp. 5003–5021, Oct. 2018.

[22]. B. Meda, E. Ragusa, C. Gianoglio, R. Zunino, A. Ottaviano, E. Scillia, and

[23]. R. Surlinelli, ''Spam detection of Twitter traffic: A framework based on random forests and non-uniform feature sampling,'' in Proc. IEEE/ACM Int. Conf. Adv. Social Netw. Anal. Mining (ASONAM), Aug. 2016, pp. 811–817.

[24]. S. Ghosh, G. Korlam, and N. Ganguly, ''Spammers' networks within online social networks: A case-study on Twitter,'' in Proc. 20th Int. Conf. Companion World Wide Web, Mar. 2011, pp. 41–42.

[25]. C. Chen, S. Wen, J. Zhang, Y. Xiang, J. Oliver, A. Alelaiwi, and M. M. Hassan, ''Investigating the deceptive information in Twitter spam,'' Future Gener. Comput. Syst., vol. 72, pp. 319–326, Jul. 2017.

[26]. David, O. S. Siordia, and D. Moctezuma, ''Features combination for the detection of malicious Twitter accounts,'' in Proc. IEEE Int. Autumn Meeting Power, Electron. Comput. (ROPEC), Nov. 2016, pp. 1–6.

[27]. M. Babcock, R. A. V. Cox, and S. Kumar, ‘‘Diffusion of pro- and anti-false information tweets: The black panther movie case,’’ Comput. Math. Org. Theory, vol. 25, no. 1, pp. 72–84, Mar. 2020.

[28]. S. Keretna, A. Hossny, and D. Creighton, ‘‘Recognising user identity in Twitter social networks via text mining,’’ in Proc. IEEE Int. Conf. Syst., Man, Cybern., Oct. 2013, pp. 3079–3082.

[29]. C. Meda, F. Bisio, P. Gastaldo, and R. Zunino, ‘‘A machine learning approach for Twitter spammers detection,’’ in Proc. Int. Carnahan Conf. Secur. Technol. (ICCST), Oct. 2014, pp. 1–6.

[30]. W. Chen, C. K. Yeo, C. T. Lau, and B. S. Lee, ‘‘Real-time Twitter content polluter detection based on direct features,’’ in Proc. 2nd Int. Conf. Inf. Sci. Secur. (ICISS), Dec. 2015, pp. 1–4.

[31]. H. Shen and X. Liu, ‘‘Detecting spammers on Twitter based on con- tent and social interaction,’’ in Proc. Int. Conf. Netw. Inf. Syst. Comput., pp. 413–417, Jan. 2015.

[32]. G. Jain, M. Sharma, and B. Agarwal, ‘‘Spam detection in social media using convolutional and long short term memory neural network,’’ Ann. Math. Artif. Intell., vol. 85, no. 1, pp. 21–44, Jan. 2020.

[33]. [31] M. Washha, A. Qaroush, M. Mezghani, and F. Sedes, ‘‘A topic-based hid- den Markov model for real-time spam tweets filtering,’’ Procedia Comput. Sci., vol. 112, pp. 833–843, Jan. 2017.

[34]. F. Pierri and S. Ceri, ‘‘False news on social media: A data-driven survey,’’ 2020, arXiv:1902.07539. [Online]. Available: https://arxiv. org/abs/1902.07539

[35]. S. Sadiq, Y. Yan, A. Taylor, M.-L. Shyu, S.-C. Chen, and D. Feaster, ‘‘AAFA: Associative affinity factor analysis for bot detection and stance classification in Twitter,’’ in Proc. IEEE Int. Conf. Inf. Reuse Integr. (IRI), Aug. 2017, pp. 356–365.

[36]. M. U. S. Khan, M. Ali, A. Abbas, S. U. Khan, and A. Y. Zomaya, ‘‘Segregating spammers and unsolicited bloggers from genuine experts on Twitter,’’ IEEE Trans. Dependable Secure Comput., vol. 15, no. 4, pp. 551–560, Jul./Aug. 2018.

DOI: 10.48175/IJARSCT-9178