IJITCE

ISSN 2347- 3657

International Journal of Information Technology & Computer Engineering

www.ijitce.com

Email : ijitce.editor@gmail.com or editor@ijitce.com

International Journal of
Information Technology & Computer Engineering

ISSN 2347–3657

Volume 12, Issue 4, 2024

https://doi.org/10.62646/ijitce.2024.v12.i4.pp145-154

ADVANCED TECHNIQUES FOR SPAMMER DETECTION AND FAKE USER IDENTIFICATION ON SOCIAL NETWORKS

Bayi Lavanya[1], Dr. N. Baskar[2]

[1]PG Scholars, Department of CSE**, Malla Reddy Engineering College For Women -MRECW (**Autonomous Institution, UGC, Govt. of India.**)**

[2] Associate Professor, Department of CSE, **Malla Reddy Engineering College For Women -MRECW (**Autonomous Institution, UGC, Govt. of India.**)**

Email id: lavanyanaidu712@gmail.com[1]**, baskarsrkv@gmail.com**[2]

Abstract: Social networks are used by millions of drug addicts around the world. Transactions on social sites such as Twitter and Facebook have a huge impact on daily life and sometimes lead to unwanted consequences. Popular social networks have become targeted platforms for spammers who spread a huge amount of useless and dangerous information. For example, Twitter has become one of the most used platforms of all time, which can send a disproportionate amount of spam. Fake drug addicts send unwanted tweets to drug addicts to promote services and websites that not only influence real drug addicts but also disrupt their resource consumption. It also increases the chances that junkies will use fake identities to obtain invalid information, leading to the spread of malicious content. Recently, detecting spammers and associating fake junkies on Twitter has become a popular research area in modern Internet Social Networks (OSNs). In this article, we discuss methods to identify spammers on Twitter. We also present a number of methods for detecting spam on Twitter and evaluate the methods according to their ability to detect (i) fake content, (ii) URL-based spam, (iii) trending content spam, and (iv) fake junkies. The presented methods are compared based on certain features such as stoner features, happiness features, graph features, structural features, and temporal features. We hope that the presented work will be a useful resource for experimenters to find the most important recent developments in Twitter spam detection on a single platform.

KEYWORDS: *Classification, fake user detection, online social network, spammer's identification, Spam Detection, Social media.*

https://doi.org/10.62646/ijitce.2024.v12.i4.pp145-154

INTRODUCTION:

Thanks to the Internet, it has come veritably easy to get any kind of information from any source around the world. The growing demand for social networks has allowed medicine addicts to collect a huge quantum of information and data about medicine addicts. The huge quantum of data available on these networks has also attracted the attention of fake medicine addicts. Twitter has snappily come an online source of real-time information about medicine addicts. Twitter is an online social network(OSN) where medicine addicts can partake anything news, opinions, moods, etc. You can have lots of exchanges on a variety of motifs, including politics, current affairs, and important events. When a stoner tweets commodity, the content is transferred directly to their followers, helping to spread the information they entered to further followers. With the development of OSN, there's a growing need to study and dissect the relations of medicine addicts on online social platforms. numerous people who don't have important information on OSN can fluently be deceived by scammers. It's also necessary to crack down and control those who use OSN only for advertising purposes and shoot large quantities of spam to other people's accounts for that purpose. lately, detecting spam on social networks has attracted the attention of experimenters. In maintaining the security of social networks, detecting spam is a tedious task. Reducing spam on social media is pivotal to cover medicine addicts from colourful kinds of vicious attacks and ensure their safety and insulation. similar dangerous acts by spammers lead to wide destruction of real communities. Spammers achieve their vicious objects through advertising, through colourful means similar as transferring spam dispatches to support colourful mailing lists and eventually promoting their interests. This trouble is confusing former medicine addicts who are known as non-spammers. It also reduces the quality of the OSN platform. Thus, it's important to develop a system to describe spammers and take corrective measures to stop their vicious conditioning. Several exploration systems have been conducted regarding Twitter spam. To keep up with the rearmost technology, multitudinous tests have also been conducted to identify fake stoner senders on Twitter. It provides a test of new styles and styles to describe

Twitter spam. The following tests represent a relative study of current approaches. Meanwhile, the authors conducted tests of different actions of spammers on the social network Twitter. The study also provides a literature review that acknowledges the frequency of spammers on the social network Twitter. Despite all the exploration, there are still gaps in the literature. To fill this gap, the current state of the art of detecting spammers and relating fake addicts on Twitter was surveyed. This report also presents a number of spam discovery styles on Twitter and discusses recent developments in this field. The ideal of this paper is to identify the different styles for detecting spam on Twitter and present the different styles by classifying these styles into several orders. To classify, we've linked four spammer reporting styles that help in assigning the identity of fake addicts. Spammers can be linked using(i) fake content,(ii) URL grounded spam discovery,(iii) trend spam discovery, and(iv) fake stoner impersonation.

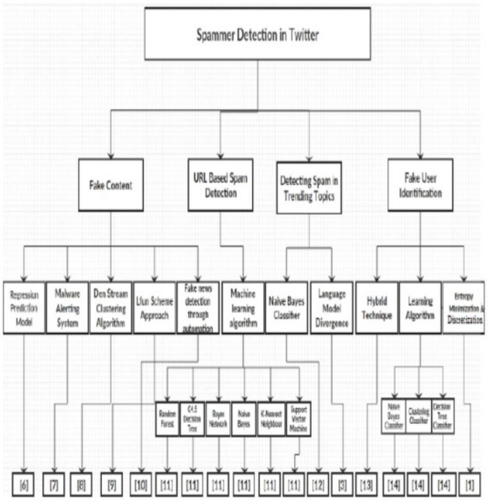
RELATED WORK

There are various schemes for multi-keyword search, targeting different functionalities of searches such as conjunctive keyword search, disjunctive keyword search, and subset search. proposed two different schemes for conjunctive keyword search based on Shamir's secret division and bilinear pairing, respectively, with the advantage that these schemes return files containing all the keywords searched for and have been proven to be secure in the standard model. A disjunctive keyword search scheme was presented in the next year that allowed retrieval of files containing any subset of the query keywords given. Predicate encryption for both conjunctive and disjunctive keyword searches has used another approach. Although these protocols correctly yield answers, such solutions are not without their limitations, as long as such solutions are concerned. The two major issues with existing approaches are that first, the indexing data structures must be rebuilt by the data owners, which is a very expensive process, and second, traditional solutions incur a high cost of computation during updates. As a result, these systems are not very efficient when dynamic operations, such as insertions or deletions of documents in large data sets, need to be conducted.

S.NO	YEAR	TITLE	AUTHORS	JOURNAL/CONFERENCE NAME	PROPOSED METHODS	LIMITATIONS
1	2017	Twitter fake account detection	B. Erçahin, Ö. Akta³, D. Kilinç, C. Akyol	Int. Conf. Comput. Sci. Eng. (UBMK)	Fake account detection on Twitter using various techniques	Limited to detecting only certain types of fake accounts
2	2017	An integrated approach for malicious tweets detection using NLP	S. Gharge, M. Chavan	Int. Conf. Inventive Commun. Comput. Technol. (ICICCT)	NLP-based approach for detecting malicious tweets	Approach might miss sophisticated attacks that bypass NLP
3	2018	Twitter spam detection: Survey of new approaches and comparative study	T. Wu, S. Wen, Y. Xiang, W. Zhou	Comput. Secur.	Survey of various Twitter spam detection techniques	Mainly focuses on spam, not broader fake content detection
4	2016	A survey on behaviors exhibited by spammers in popular social media networks	S. J. Soman	Int. Conf. Circuit, Power Comput. Technol. (ICCPCT)	Survey of spamming behaviors across social media platforms	Focuses on general behaviors, lacks specific detection methods
5	2017	Twitter analysis for real-time malware discovery	F.Concone,A.De Paola, G. Lo Re,Morana	AEIT Int. Annu. Conf.	Real-time analysis of Twitter for malware detection	Limited to malware-related analysis, no general spam/fake detection

https://doi.org/10.62646/ijitce.2024.v12.i4.pp145-154

SYSTEM ARCHITECTURE:



PROBLEM STATEMENT:

Multi-key word search schemes now currently experience several issues relating to efficiency and the computational cost of operations such as dynamic document insertion and deletion. Jaccard keyword search schemes based on Shamir's secret division and bilinear pairing guarantee Multi-key word search schemes now currently experience several issues relating to efficiency and the computational cost of operations such as dynamic document insertion and deletion. Jaccard keyword search schemes based on Shamir's secret division and linear pairing guarantee security and accuracy but require data owners to rebuild the index search structure, which is quite time-consuming and resource-intensive. Also, a more efficient, scalable solution that supports dynamic operations without heavy overhead is so much in demand and will provide much better performance as well as user-friendliness. security and accuracy but require data owners to rebuild the index search structure, which is quite time-consuming and resource-intensive. Also, a more efficient, scalable solution that supports dynamic operations without heavy overhead is so much in demand and will provide better performance as well as user-friendliness.

PROPOSED MODEL:

In this direction, the proposed system develops a more effective and secure solution through Bloom filters-based index tree structures that reduce rebuilding the search index structure. Consequently, time consumption and resources decrease, contributing to a significant improvement in search performance. A probabilistic data structure called a Bloom filter uses little memory and can produce a small number of false positives while rapidly determining if an element is a part of a set. Bloom filters, when incorporated into index tree structures, assist in monitoring changes or new data entries, eliminating low-probability or irrelevant candidates at the beginning of the search process.

As a result, rebuilding search indexes less frequently requires computational overhead, allowing spam detection systems to get updates in real time. The hybrid method works well for large-scale spam filtering because it guarantees quicker searches, improved detection accuracy, and less storage usage.

The system efficiently supports dynamic operations such as insertions and deletions of documents, making it very salable and adaptable to huge data sets. The core of the system is the usage of the vector space model that creates an index vector for every file in the outsourced dataset. This vector space model, when hybridized with the frequently used Term Frequency-Inverse Document Frequency (TF-IDF) weighting technique, gives enhanced search accuracy by prioritizing more relevant documents. Furthermore, it also computes the relevance score between the search query and indexed files using cosine similarity for the proper ranking of the results. The integration of these methods in the proposed system reduces computational costs both during the search and update processes. This makes it a powerful and applicable solution for environments that require frequent updates as well as powerful multi-keyword search, which addresses previous inefficient systems.

ANALYZE VSM AND TF-IDF :

The accuracy of spam detection is much improved when the vector space model (VSM) is hybridized with the Term Frequency-Inverse Document Frequency (TF-IDF) weighting technique. When it comes to spam filtering, emails and messages are handled like papers, and keywords and patterns that are deemed suspect are given special weight in the classification process. Since a pure VSM representation applies weights based just on raw term frequency, it may not be sufficient to distinguish between authentic and spam emails. This may lead to a preponderance of frequent or unimportant terms in the vector representation, which could result in incorrect categorization or false positives. TF-IDF integration, on the other hand, improves the model's ability to filter spam by highlighting terms like "winner," "lottery," or "click here" that are common in spam messages but uncommon in emails that are legal.

By giving higher weights to phrases that are specifically linked to spam and lower weights to common but uninformative words (like "hello" and "regards"), this hybridization enhances spam detection. Consequently, the system gains proficiency in differentiating between non-spam and spam information by detecting minute variations in word usage patterns. Greater precision and recall are achieved by the filtering method with TF-IDF-enhanced VSM, which lowers false positives and false negatives. Consequently, spam detection becomes more dependable.

Implementation

Tweet Admin

In this module, the Admin has to login by using a valid user name and password. After login successful he can perform some operations such as View Users and Authorize(Give link on user to view Profile),View all Uses Friend Request and Response,Add Spam Filter name,View All spamming accounts with profile details and Block,View All UnBlock request users details using decision tree format and Unblock by clicking user name ,View all User's Tweet Topic with Interactions and scores,View All Spam Account(Based on Virus,Malware) And Normal Account with Reasons based on Random Forest Tree,View All Spamming and Normal Behaviors based on Interactions by Filter Name and give link to show Number of both users in chart,View All Spamming and Normal Behaviors based on Tweet Meta Data by Filter Name and give link to show Number of both users in chart,View Number of Spamming Account and Normal Account in Chart

Friend Request Response

In this module, the admin can view all the friend requests and responses. Here all the requests and responses will be displayed with their tags such as Id, requested user photo, requested user name, username request to, status and time & date. If the user accepts the request then the status will be changed to accepted or else the status will remain as waiting.

User

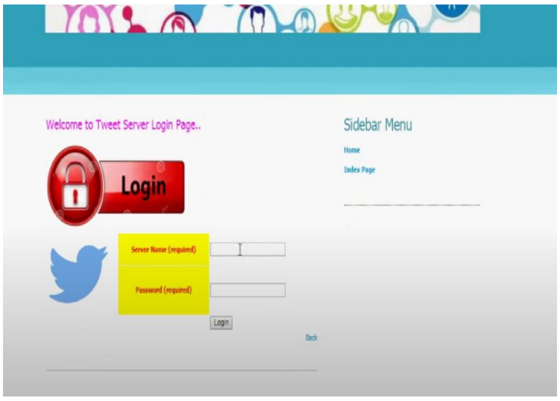
In this module, there are n numbers of users present. Users should register before performing any operations. Once a user registers, their details will be stored in the database. After registration is successful, he has to login by using authorized user name and password. Once Login is successful user can perform some operations like View Your Profile with community, Search Friends based on community, View Friend Request and Response,View My Friends based on community, Create Tweet Topic with tweet_postname, TAbout, TUses,content desc, Browse MetaData_desc, TweetURL, TDate and Time, TOwner, add TImage, Search

Tweet Topic by keyword and give Your Interactions (increase score while viewing) and view URL to see web page,View all your Tweets Topic with other Interactions and scores,View all your Friends Tweet Topic with other Interactions and scores and give your Interactions,View All Similar Friend's Tweets Topic,show all Spamming behaviors friends Topics with profile.

Searching Users to make friends

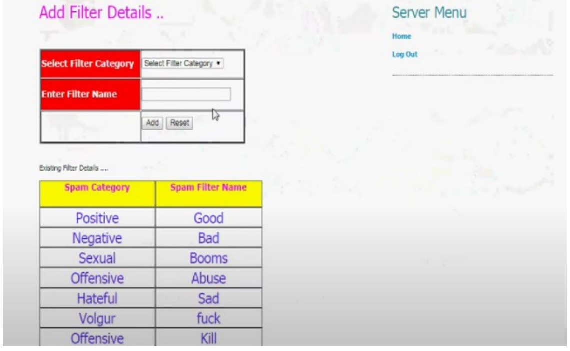
In this module, the user searches for users in Same Network and in the Networks and sends friend requests to them. The user can search for users in other Networks to make friends only if they have permission.

RESULTS:


Welcome to Tweet Server Login Page..

Server Name (required)	
Password (required)	

FIG:1-login

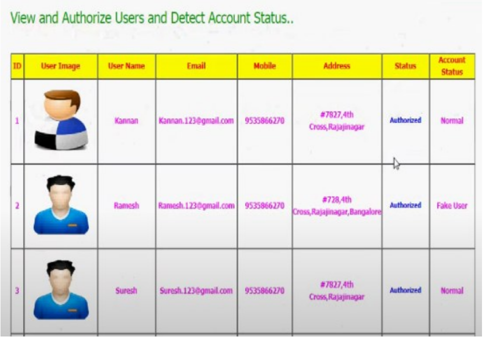

Add Filter Details ..

Select Filter Category	Select Filter Category ▾
Enter Filter Name	

Existing Filter Details ....

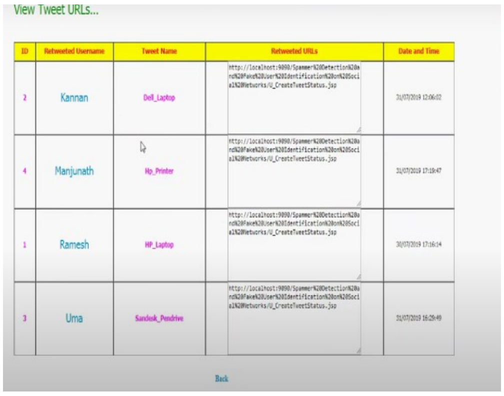
Spam Category	Spam Filter Name
Positive	Good
Negative	Bad
Sexual	Booms
Offensive	Abuse
Hateful	Sad
Volgur	fuck
Offensive	Kill

FIG:2-FILTER DETAILS


View and Authorize Users and Detect Account Status..

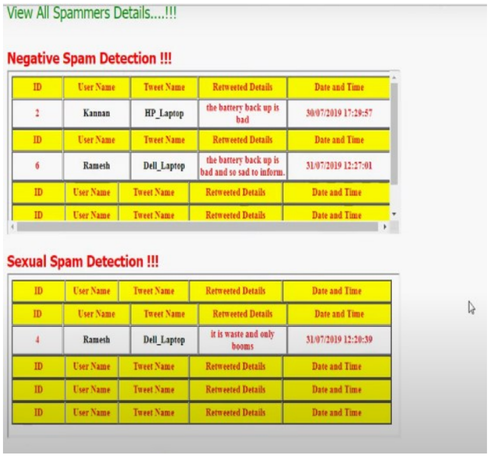
ID	User Image	User Name	Email	Mobile	Address	Status	Account Status
1		Kannan	Kannan.123@gmail.com	9535866270	#7827,4th Cross,Rajajinagar	Authorized	Normal
2		Ramesh	Ramesh.123@gmail.com	9535866270	#728,4th Cross,Rajajinagar,Bangalore	Authorized	Fake User
3		Suresh	Suresh.123@gmail.com	9535866270	#7827,4th Cross,Rajajinagar	Authorized	Normal

FIG:3- USER DETAILS


View Tweet URLs...

ID	Retweeted Username	Tweet Name	Retweeted URLs	Date and Time
2	Kannan	Dell_Laptop	http://localhost:9090/Spammer%20Detection%20and%20Fake%20User%20Identification%20on%20Social%20Networks/U_CreateTweetStatus.jsp	31/07/2019 12:06:02
4	Manjunath	Hp_Printer	http://localhost:9090/Spammer%20Detection%20and%20Fake%20User%20Identification%20on%20Social%20Networks/U_CreateTweetStatus.jsp	31/07/2019 17:19:47
1	Ramesh	HP_Laptop	http://localhost:9090/Spammer%20Detection%20and%20Fake%20User%20Identification%20on%20Social%20Networks/U_CreateTweetStatus.jsp	30/07/2019 17:16:14
3	Uma	Sandesk_Pendrive	http://localhost:9090/Spammer%20Detection%20and%20Fake%20User%20Identification%20on%20Social%20Networks/U_CreateTweetStatus.jsp	31/07/2019 16:29:49

FIG:5- TWEETS URLS


View All Spammers Details....!!!

Negative Spam Detection !!!

ID	User Name	Tweet Name	Retweeted Details	Date and Time
2	Kannan	HP_Laptop	the battery back up is bad	30/07/2019 17:29:57
ID	User Name	Tweet Name	Retweeted Details	Date and Time
6	Ramesh	Dell_Laptop	the battery back up is bad and so sad to inform.	31/07/2019 12:27:01
ID	User Name	Tweet Name	Retweeted Details	Date and Time
ID	User Name	Tweet Name	Retweeted Details	Date and Time

Sexual Spam Detection !!!

ID	User Name	Tweet Name	Retweeted Details	Date and Time
ID	User Name	Tweet Name	Retweeted Details	Date and Time
4	Ramesh	Dell_Laptop	it is waste and only booms	31/07/2019 12:20:39
ID	User Name	Tweet Name	Retweeted Details	Date and Time
ID	User Name	Tweet Name	Retweeted Details	Date and Time
ID	User Name	Tweet Name	Retweeted Details	Date and Time

FIG:6- SPAM DETECTION

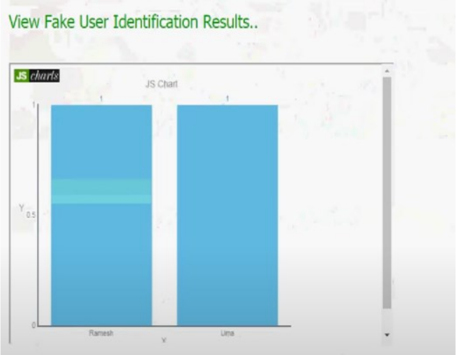


FIG:7- GRAPH

CONCLUSION:

We discussed several spammer detection techniques in the context of Twitter. In addition, we presented several spam detection methods over Twitter and classified them to be of types including "fake content detection", "URL-based spam detection", "trend content detection" and "fake user detection methods". Several comparisons are presented based on characteristics of user, satisfaction characteristics, graphic characteristics, structural characteristics, and temporal characteristics. The styles were also compared about the particular objects and data sets used. The developed overview aims at helping researchers find the latest information regarding spam detection styles on Twitter in an integrated way. Despite the effective and effective approaches to spam detection and identification of counterfeit drugs on Twitter, some undiscovered areas still attract the attention of experimenters. Therefore, to put this issue in a few words, the widely spread false news in social networks is an issue that should be questioned, for the same news influences both individuals and groups. There is also some related content that should be disseminated. Source of the rumor on the social network should be spread. As for more sophisticated approaches, there are those presented by B.B. Social network based approaches may also be applicable. Several studies have been conducted in the past to trace the sources of rumours, mainly based on statistical methods.

REFERENCES

[1] B. Erçahin, Ö. Akta³, 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.*

https://doi.org/10.62646/ijitce.2024.v12.i4.pp145-154

Technol. (ICICCT), Mar. 2017, pp. 435_438.

[4] T. Wu, S. Wen, Y. Xiang, and W. Zhou, ``Twitter spam detection: Survey 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. Technol. (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 popular Twitter threads," in *Proc. IEEE Int. Conf. Smart Cloud (SmartCloud)*, Nov. 2017, pp. 208_215.