International Journal of Recent Technology and Engineering (IJRTE)
ISSN: 2277-3878 (Online), Volume-8 Issue-6, March 2020

Social Media Fake Profile Detection Technique Based on Attribute Estimation and Content Analysis Method

Nitika Kadam, Harish Patidar

Abstract: ***Social media is a virtual place where every age group members are available. Members are using this platform to share knowledge and information. But some of them are misuse and abuse the services of social media. In order to perform malicious activities the users and BOTs are creating fake profile, these profiles are used for spreading the unsolicited and malicious contents. Therefore the proposed work is motivated to study about the social media platform and it's security and privacy challenges. In this context, the recent year's progress in the domain of social media is explored, and the literature is collected and categorized according to the area of work. Among them concepts of profile attribute evaluation and the content analysis based techniques are most promising. Therefore by using the concept of both the kinds of technique a new model is tried to design and their functional aspects are discussed. In near future the proposed working model is implemented and their performance is demonstrated.***

Keywords: social media analysis, security and privacy, fake profile detection, data mining and techniques, survey.

I. INTRODUCTION

Data mining and its techniques are improving day by day and its relevant areas of applications are also expanding. These techniques are helpful on decision making, pattern recognition, predictions and others. Due to it's ability of automated and efficient data analysis a number of domains are accepting this technology such as engineering, medical, business, banking, and more. On the other hand as different communication channels and the digitization of record are increasing the new challenges in data mining is involved. Additionally the amount of data in different data repositories and data bases rapidly increases. In this context the data mining techniques are offering the methods and tools to analyze the data and recover the required application objectives. Therefore these techniques are frequently adopted in different areas [1].

In this presented work the application of data mining technique in security and data management process is explained. Therefore the social media data analysis and fake profile detection is the main area of study.

Manuscript received on February 10, 2020.
Revised Manuscript received on February 20, 2020.
Manuscript published on March 30, 2020.
* Correspondence Author
Nitika Kadam*, Bachelor of Engineering, Computer Science from LNCT Indore.
Harish Patidar, Bachelor of Engineering, Computer Science and Engineering discipline, MIT Ujjain.

Social media is become most popular platform and every age groups are keenly interested on the services of social media. It is virtual world where anybody can join and meet new and old
members. People spend their important time on Social media platforms like twitter, facebook etc. [2]. The users in these platforms are very active. But there are two kinds of users first who are technically sound and know the social media usages limit and some of them are new. Sometimes these new users are not much having the understanding about the technology and their usages [3].

Moreover it, there are one more user's category present in social media who are using this platform for other purposes i.e. promotions, activity, events, political views and advertisements. Among some of them legitimately working on social media but some of them are abuse the innocent users by targeting them, and promoting and distributing the hate, spamming, phishing and other kinds of serious attacks [4]. Therefore identification and differentiation between fake and legitimate profiles in social media required for developing friendly and secure social media.

II. ONLINE SOCIAL NETWORK

In this age of technology every small and large requirements of user is obtainable through the internet. Internet provides the services at the user's door steps through different kinds of applications such as internet banking, search engine, e-commerce and etc. Among them a web application named as social networking becomes more and more popular in recent years. These applications are termed as "online social networking (OSN)" [5]. It is a big platform for any web user to search people, share data, and express their emotions and others [6]. The purpose of OSN is to join some users, create a profile, and use different applications hosted. Additionally it is an easy way to share information with their contacts or for public use. A user wants to actively participate in OSN required to make visible the profiles and connects. It is core functionality of social network which makes a person to be searchable [7]. On the other hand some privacy issues related to information leakage, identity based frauds, discloser of private and sensitive information invites malicious attacks. Using these information reputation slanders, personalized spamming, and phishing kinds of attacks are deployed [8]. There are two kinds of architecture is available for OSN, namely client server and peer to peer architecture. The client server based model is a centralized implementation of OSN and P2P is a distributed architecture.

2.1 Client-Server Architecture — Almost all OSNs are web based system. Such kind of network can be termed centralized OSN. Twitter and Facebook are the central providers which provide the processes like access, storage and maintenance. The key advantage of this model, it is easy of implement and management.

2.2 But a single point of failure may affect the network performance and can stop all the communication [9].

2.3 P2p Architecture — OSN can also be designed using P2P architecture. The implementation of P2P OSN is performed in decentralized manner. Here users' personal spaces are used for storage. This OSN support direct data exchange and support local services even when Internet is not available [10]. P2P architecture supports communication using local connectivity. This architecture is much supportive for the resource provisioning and data sharing and communication.

III. LITERATURE SURVEY

The fake profiles in a social media are growing for distributing the malicious and spam contents. Sometimes it becomes very harmful for society in these directions some essential contributions are provided in this section. According to **Cao Xiao et al [11]** fake accounts are kinds of malicious users who use a social media account to send spam, abuse the system or commit fraud. Authors report an approach for identifying fake accounts. These technique usages supervised learning for set of account classification. The model achieved AUC 0.95 - 0.98. Similarly ***Zaher Yamak et al [12]*** are analyzing the behavior of multiple fake accounts. The methodology provides a set of features to use in data analysis. The results show that is enables to detect 99% of fake accounts. ***Myo Myo Swe et al [13]***, also works to detect fake accounts. The topic modeling and keyword extraction techniques were used to create a blacklist. The technique produces 95.4% accuracy and 0.95 true positive rates. ***M. Mohammadrezaei et al [14]*** determine fake accounts by measuring features i.e. common friends. Thus cosine, L1-measure, Jaccard, and weight similarity were calculated and prepared a graph of the social network correspondence. The Medium Gaussian SVM algorithm has been used to predicts fake accounts. The results shows the curve=1 and false positive rate=0.02. The multimedia based social network such as YouTube is investigated by **Yixuan Li et al [15]**. And propose a method, to demonstrate Scaling, and also the fake activities on target platform to track over time with an accuracy of 98%.

Now in these days not only the human create fake accounts the machine learning algorithms are also used for deploying the fake account. These algorithms are known as social bots. **Onur Varol et al [16]** presented a system to identify the bots on social media. In this context thousands of features extracted from public data with the associated meta-data. The Twitter bots dataset were used to benchmark the classification system. The system demonstrates 9% - 15% of Twitter accounts are bots. **E. V. D. WALT and J. ELOFF [17]** apply a set of features to classify fake accounts. and ***S.* Cresci et al [18]** study about this phenomenon on Twitter. Additionally they proof quantitatively about spam-bot design. The finding shows not a single kind of method available to identify social spam-bots. **E. Ferrara et al [19]** design an algorithm to exhibit human-like behavior. To create a healthy social media ecosystem, so it is essential to eliminate fake Likes. **S. Gurajala et al [20]** study 62 million Twitter user profiles and presents a strategy to identify fake profiles creation.

Fake news or rumor is an another key issue in social media **Sebastian Tschiatschek et al [21]** were works to recognize fake news over social media. The aim is to select a subset of news, to take opinion from an expert, and stop fake news distribution. Using this approach we can prevent or minimize the misinformation. ***K. Shu et al [22]*** present a review for identifying fake news, including characterizations to provide evaluation of model. Finally authors have been discussed about research areas, issues, and future directions. **E. Ferrara et al [23]** study trending memes that attract attention and designed a ML system to recognize campaigns. With hashtags a millions of posts were used to prove accurate recognition is possible up to 95%. **J. Song et al [24]** propose a method called CrowdTarget. The aim is to recognize, post, page, and URL not accounts. According to results the CrowdTarget accurately distinguish tweets with low error rate of 0.01, and accuracy up to 0.98. **V. L. Rubin et al [25]** report three types of fake news with advantages and limitations.

Spamming is another objective of fake profile creation. The spam is distributed for phishing and trap purpose also. **F. Masood et al [26]** provided a survey of spammer's detection techniques. Taxonomy of Twitter spam detection approaches is also presented: (i) fake content, (ii) spam based on URL,

(iii) spam in trending topics, and (iv) fake users. **I. Sen et al [27]**discovers some essential feature for recognizing genuine like on Instagram. Using analysis of user behaviours to like a post, they successfully developed a method to detect fake likes with a precision of 83.5%. **P. Ratna et al [28]** explore the complexities of "fake likers" detection. To uncover this: (1) a number of profiles fake and legitimate were collected,

(2) analyze characteristics of both, (3) identify effective features, and (4) evaluate their performances. Results show methods provide accuracy = 0.871, false positive rate = 0.1, and false negative rate = 0.14. **A. M. A. Zoubi et al [29]** investigate the nature of spam profiles to improve a spam detection system. Dataset of 82 Twitter's profiles are collected and analyzed. Moreover, a feature selection process is used to identify the most influencing features.

On the other hand some of the authors assumed the social media issue as the security attack. Therefore, **J. Jia et al [30]** discussed Sybil detection in OSN. Random walk based technique were used to structure an OSN. In order to decide the place of user a reputation scores were used. Due to limitations of existing methods: 1) class labels benign or Sybils, not both, 2) low classification rate, and 3) not effective to manage the noise. **S. Rathore et al [31]** survey about security and privacy concern based on various threats during multimedia content sharing. They discuss various existing solutions to preserve user.

R. Kaur et al [32] discusses different types of anomalies and characteristics by review of number of techniques. It also includes review of data mining approaches. **O. Goga et al [33]** introduced a method to develop a dataset for impersonation attacks. The 16,572 cases are taken from Twitter and classified Identity Impersonation attacks on target celebrities.Findings show (i)
IIA are much broader than believed and (ii) Attackers are creating real type of accounts.
B. Viswanath et al [34] propose Stamper, for detecting tampered crowd computations. The design is based on two factors: (1) Sybil attack detection. (2) Sybil identities cannot forge the timestamps of activities. Additionally **N. Laleh et *al***offered a risk analysis technique based on user behavior. These behaviors can be 'normal behavior' if assessment score beyond then a threshold.The above discussed literature is summarized in this section, the table 1 shows the working domain of different authors in social media security and improvements.

Table 1 Contributions in Social Media

Authors	Domain of work
Cao Xiao et al [11]	Profile based account
Yixuan Li et al [12]	Profile based account
Zaher Yamak et al [15]	Profile based account
Myo Myo Swe et al [15]	Profile based account
M. Mohammadrezaei et al [18]	Profile based account
Onur Varol et al [14]	Bots
E. V. D. WALT and J. ELOFF [17]	Bots
S. Cresci et al [22]	Bots
E. Ferrara et al [26]	Bots
S. Gurajala et al [31]	Bots
Sebastian Tschiatschek et al [13]	Fake news
K. Shu et al [21]	Fake news
E. Ferrara et al [30]	Fake news
J. Song et al [31]	Fake news
V. L. Rubin et al [32]	Fake news
F. Masood et al [18]	Spamming
I. Sen et al [20]	Spamming
P. Ratna et al [27]	Spamming
A. M. A. Zoubi et al [23]	Spamming
J. Jia et al [24]	Attack and security
S. Rathore et al [25]	Attack and security
R. Kaur et al [28]	Attack and security
O. Goga et al [34]	Attack and security
B. Viswanath et al [35]	Attack and security
N. Laleh et al [29]	Risk assessment

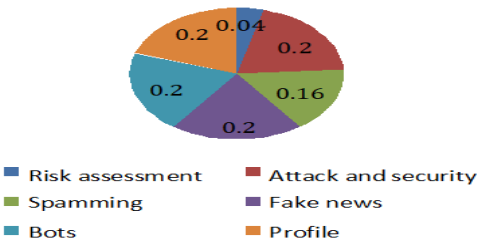

Figure 1 Social Media Security Contributions

IV. PROPOSED WORK

The recent contributions in the area of social spam or malicious profile detection is studied, the entire domain contains three major part of working systems.

1. Spam or unsolicited post identification
2. Detection of bots
3. Detection of profile

Additionally most of the authors are works either with the profile attributes or the content which are distributed. There are not single methods available that utilizes the both content and profile attribute to ensure the profile is fake or legitimate. Therefore in this work we are utilizing both the features (i.e. content and profile attributes), additionally some of the features are developed by own to estimate and rank the risk level of a target profile.The proposed methodology is demonstrated as per figure 2. The system contains the components and processing aspects, it also helps to keep a track record of the work involved.

4.1 Social Media Platform: A data mining techniques requires some initial samples to train the learning algorithms and then we can apply it with the real world applications or issues for recognizing the patterns, classification prediction and etc. Therefore to prepare the synthetic dataset, proposed methodology identifies the social media platform which can disclose the data and profile information automatically. In this context [36] works on the twitter dataset for fake profile detection. Additionally in [37] the authors are works for fake profile detection using machine learning and NLP (Natural Language Processing), where they concentrated on Facebook dataset. In this article dataset of 1337 false users and 1481 true users has been used which is publically available. The authors of [39] are usages the twitter dataset for finding the fake profiles by analysis of profile characteristics and activity-based pattern detection. In the article [40] the user's fake Profiles are identified. As per the publically available article, Random Forest & Deep Convolutional Neural Network (CNN) has been used with two machine learning algorithms. To identify social bot colleagues in academia collected various bot dataset which is presented in their training data. The paper [42] detects the spam as well as the spammer and fake profile in social media. In this context they consume the twitter social media data.
The paper [43] contributes on analyzing data on LinkedIn social media. In
order to do this they are usages the technique ofclustering.In order to find the effective dataset for experimentation and system design we explored almost 8 papers. In theseresearch articles different social media based datasets are used. The different datasets are demonstrated in table 2.

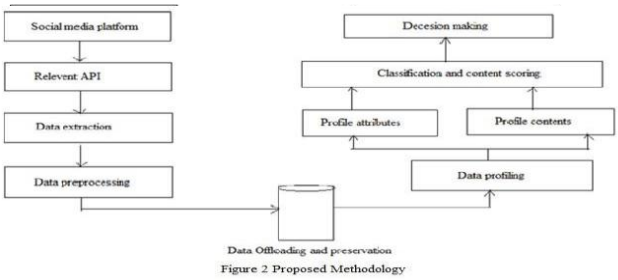

Figure 2 Proposed Methodology

References	Dataset used
[36]	Twitter
[37]	Facebook
[38]	Public spam user dataset
[39]	Twitter
[40]	Twitter
[41]	Academia
[42]	Twitter
[43]	LinkedIn

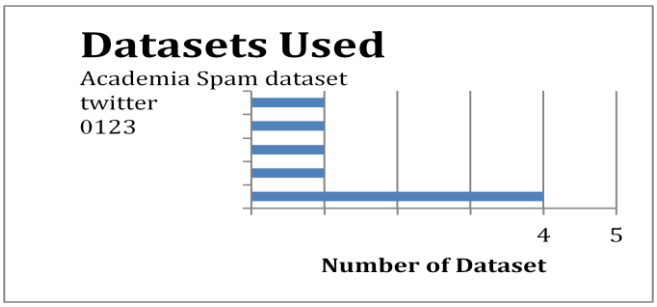

Figure 2 Dataset Used

According to the survey of different research articles for finding effective datasets for social media fake profile detection different social media data is used. Among them the twitter dataset is frequently used for machine learning based approaches.

4.2 relevant api*:* Now in these days almost every online platform publishes their data and feeds using some specific API (application programming interface), therefore different platform usages the different APIs for the data extraction. As previously made conclusion we are intended to work with the twitter datasets. Therefore the twitter fake profile dataset is found in [44] and [45] which can directly downloadable and useable with the learning algorithms. In addition of we also find the tutorial for discovering the twitter data feeds and profile information using [46]. Here we can use both the kinds of datasets like [44] and [45], to demonstrate the methodology. Additionally by using the available techniquesof [46] for data extraction we justify the work against the benchmark datasets.

4.3 Data extraction*:* A number of APIs supporting the query based data extraction from the social media and online platforms. In this context, application prepares a query string and passes it to the target API, API finds the important data from Social media and returns it. In the similar manner the API described in [46] is going to be used for extracting the user data. In addition of the described dataset in [44] and[45] is also used for experimentation and system design. In addition of that a fake profile dataset is also available for finding fake profile in twitter [47]. This dataset is hosted on Github and contains 22 attributes additionally their detailed description is also available in [47]. Moreover it one more contribution in the same place we have found as described in [48]. That dataset is available in CSV (Comma-separated values) format. The total of 33 attributes is available in this dataset. Two separate files contains the entire profile information in which instance of data and file name of first file 1338 is fake. Therefore the file name is treated here as the class labels for the given set of attributes. Similarly, the second file is containing 1482 instance of data. That represents the legitimate profiles thus class label is termed here as legitimate. After combining both the files in a common CSV file the total of 2820 instance of data and two class labels we found. The following set of attributes is available in this dataset as given in table 3.

Table 3 Dataset Attributes

S. No.	Name	Type
1	Id	Numeric
2	Name	Text
3	screen_name	Text
4	statuses_count	Numeric
5	followers_count	Numeric
6	friends_count	Numeric
7	favourites_count	Numeric
8	listed_count	Numeric
9	created_at	Date & time
10	url	Text
11	Lang	Text
12	time_zone	Text
13	Location	Text
14	default_profile	Text
15	default_profile_image	Text
16	geo_enabled	Numeric
17	profile_image_url	Text
18	profile_banner_url	Text
19	profile_use_background_image	Text
20	profile_background_image_url_h ttps	Text
21	profile_text_color	Text
22	profile_image_url_https	Text
23	profile_sidebar_border_color	Text
24	profile_background_tile	Text

25	profile_sidebar_fill_color	Text
26	profile_background_image_url	Text
27	profile_background_color	Text
28	profile_link_color	Text
29	utc_offset	Text
30	Protected	Text
31	Verified	Text
32	Description	Text
33	Updated	Date and time
34	Dataset	Text

4.4 Data preprocessing: The data extracted from the social media may contain a significant amount of noise and unwanted data. In addition of sometimes we obtain abundant datasets which may contains the missing or null values, special characters and other kinds of unwanted data. Therefore minimization of noise and maximization ofinformative contains required to preprocess the data. Quality of data is improved with the help of processing techniques and same is used to learn the example patterns. In this context to preprocess the data attributes we investigated different research articles.First we explored about the Discrimination-Aware Classification Problem. To enhance the accuracy classifier has been used. **Faisal Kamiran et al [49]** concentrate on the case of binary sensitive attribute in two-class. And provide an algorithm that preprocesses the data to remove discrimination. They extend existing data preprocessing techniques, by changing class labels, and reweighing or re- sampling. **Sergio Ramírez-Gallego et al [50]** on data preprocessing they summarize categorize and analyze. Also it explored feature, instance selection and discretization. They conduct experiments using relevant contributions. Finally offered advices about existing data stream preprocessing algorithms. Data preprocessing from log files a time-consuming phase. **Michal Munk et al [51]** attempts to identify phases, necessary in preprocessing of educational data. The sequential patterns analysis is considered, and tries answering the question, the preprocessing phases has an impact on discovered knowledge. **Elaf Abu Amrieh et al**
[52] proposed a model with a new category of features, called behavioral features. This feature is related to learner interactivity. They use data mining techniques to evaluate impact of features. Wrapper and filter are the filter selection used in **Hua Yin et al [53]** with different dimensions. Feature selection is better, when dataset is less imbalance. Simple searching method is used in wrapper based feature selection.
By the different authors and researchers we found a number of techniques available for preprocessing of data. According to the study we observed that the techniques are used for preprocessing of data is mostly depends on the application and the nature of data. Therefore we include the techniques for handling the missing values and noisy attributes during the data preprocessing.
4.5 Data offloading and preservation: the preprocessed data in previous phase is used in further information extraction and decision making purpose. Therefore a structured manner of data organization and preservation required to implement. In this context in this phase a local database is prepared which can write the data on a given database for further use with the proposed data mining system.
4.6 Data Profiling**:** the offline data which is extracted from the social media contains the different user profiles and their contents. Here the content term is used to describe the data which is published or shared by the user. Thus the data profiling is required to separate the user based data, their consumption and activities. All the personalized user's data is required to investigate about the user behavior and activity conducted over the social media platform.
4.7 Profile Contents: the normal and fake both kinds of profiles are used for publishing the contents over the social media to target an audience, thus the "contents matter" for identifying the spammers and fake users of social media. Therefore the proposed work not only targeting the profile attributes it also considers the contents which are circulated using the target profile. Additionally the methodology is developed in such way by which both the property can be used for demonstrating the profile behavior more accurately.
4.8 Profile Attributes: most of the time the trolling and spreading contents are not done by the actual profile. In this purpose a new profile is created using the different unauthorized attributes. These attributes are available in the data which are listed over the social media such as date of profile creation, background image, location, mobile verification and other methodologies. Thus the available attributes in a profile is used for authenticating the user.
4.9 Classification and Content Scoring*:* The decision making task is not only can be made by using the profile attributes. It requires which kinds of data are spared using the profile therefore both kinds of aspects are involved in this work. First the attributes are classified according to the available literature methodology and also contents are also required to classify. Finally the scoring based on both the properties is performed for strongly providing the decisions about the profile is fake or legitimate.
4.10 Decision Making*:* That is the final outcome of the proposed fake profile detection system. Therefore by using the outcomes of the previous step of data analysis the conclusion is made about the target profile in terms of fake or legitimate. Thus using the profile attributes and by recovering the score by using sentiment analysis the decision are made more precisely. Therefore the proposed multi-factor social spam profile detection approach is accomplished.

V CONCLUSION & FUTURE WORK

This paper provides the survey and review over the existing and recently contributed research articles. These articles are working on various domain of security, the identified domain of research work are based human created fake accounts, BOTs based fake accounts, spamming and fake news distributions, risk assessment and attack oriented. Additionally, the understanding of social media network is also reported in this paper. Further by concluding the recent literature a new data model using the data mining and machine learning technique is proposed. The proposed working model is a combination of content based and profile attribute based method. The proposed method involves the advantage of both the techniques and accurately classifies the fake accounts.

Also to implement the proposed data model for recognizing the fake user profiles the data extraction techniques are concluded. In addition of that the different datasets available for this task is also reported and for further implementation of the techniques the dataset is finalized. Moreover for processing of data initially for improving the quality of learning and performance classifiers the preprocessing technique is also finalized. Additionally the basic overview of the proposed data model is given. In near future the proposed system is implemented and the performance is reported.

REFERENCES

1. E. E. Papalexakis, C. Faloutsos, N. D. Sidiropoulos, "Tensors for Data Mining and Data Fusion: Models, Applications, and Scalable Algorithms", ACM Transactions on Intelligent Systems and Technology, Vol. 8, No. 2, Article 16, Date: October 2016.
2. [Book] C. Wüest, "The Risks of Social Networking", Security Response, https://www.symantec.com/content/en/us/enterprise/media/security_response/whitepapers/the_risks_of_social_networking.pdf.
3. Prof. T. Ravichandran, "Enhancing Soft Skills And Personality", Indian Institute of Technology Kanpur, National Programme on Technology Enhanced Learning (NPTEL)
4. Romanov, A. Semenov and J. Veijalainen, "Revealing Fake Profiles in Social Networks by Longitudinal Data Analysis", DOI: 10.5220/0006243900510058, In Proceedings of the 13th International Conference on Web Information Systems and Technologies (WEBIST 2017), pages 51-58 ISBN: 978-989- 758-246-2
5. Guille, H. Hacid, C. Favre, D. A. Zighed, "Information Diffusion in Online Social Networks: A Survey", ACM, 2013,
6. VOL42 ISS2, June 2013
7. Whiting, D. Williams, "Why people use social media: a uses and gratifications approach", Qualitative Market Research: An International Journal Vol. 16 No. 4, 2013 pp. 362-369 q Emerald Group Publishing Limited 1352-2752
8. Maier, S. Laumer, A. Eckhardt and T. Weitzel, "Giving too much social support: social overload on social networking sites", European Journal of Information Systems (2014), 1–18 © 2014 Operational Research Society Ltd.
9. Gan, and L. R. Jenkins, "Social Networking Privacy—Who's Stalking You?", Future Internet 2015, 7, 67-93; doi:10.3390/fi7010067
10. Zhang and J. Sun, X. Zhu, Y. Fang, "Privacy and Security for Online Social Networks: Challenges and Opportunities", IEEE Network • July/August 2010 0890-8044/10/$25.00 © 2010 IEEE.
11. S. Buchegger, D. Schioberg, L. H. Vu, A. Datta, "PeerSoN: P2P Social Networking —Early Experiences and Insights", SNS'09,
12. March 31, 2009, Nuremberg, Germany, Copyright 2009 ACM
13. Xiao, D. M. Freeman, T. Hwa, "Detecting Clusters of Fake Accounts in Online Social Networks", AISec'15, October 16, 2015, Denver, Colorado, USA, © 2015 ACM ISBN 978-1-4503- 3826-4/15/10
14. Z. Yamak, J. Saunier, L. Vercouter, "Detection of Multiple Identity Manipulation in Collaborative Projects", WWW'16 Companion, April 11–15, 2016, Montréal, Québec, Canada,
15. ACM 978-1-4503-4144-8/16/04
16. M. M. Swe, N. N. Myo, "Blacklist Creation for Detecting Fake Accounts on Twitter", International Journal of Networked and Distributed Computing, Vol. 7(1); December (2018), pp. 43–50
17. M. Mohammadrezaei, M. E. Shiri, and A. M. Rahmani, "Identifying Fake Accounts on Social Networks Based on Graph Analysis and Classification Algorithms", Hindawi Security and Communication Networks Volume 2018, Article ID 5923156, 8 pages
18. Y. Li, O. Martinez, X. Chen, Y. Li, J. E. Hopcroft, "In a World That Counts: Clustering and Detecting Fake Social Engagement at Scale", WWW 2016, April 11–15, 2016, Montréal, Québec, Canada. ACM 978-1-4503-4143-1/16/04
19. O. Varol, E. Ferrara, C. A. Davis, F. Menczer, A. Flammini, "Online Human-Bot Interactions: Detection, Estimation, and Characterization", Proceedings of the Eleventh International AAAI Conference on Web and Social Media (ICWSM 2017)
20. E. V. D. WALT and J. ELOFF, "Using Machine Learning to Detect Fake Identities: Bots vs Humans", 2169-3536, 2018 IEEE, VOLUME 6, 2018
21. S. Cresci, R. D. Pietro, M. Petrocchi, A. Spognardi, M. Tesconi, "The paradigm-shift of social spambots", Web Science 2017 ACM, Perth, Australia, ISBN 123-4567-24-567/08/06
22. E. Ferrara, O. Varol, C. Davis, F. Menczer, and A. Flammini, "The Rise of Social Bots", Communications of The ACM, JULY 2016, VOL. 59, NO. 7
23. S. Gurajala, J. S. White, B. Hudson, J. N. Matthews, "Fake Twitter accounts: Profile characteristics obtained using an activity-based pattern detection approach", SMSociety '15, July 27 - 29, 2015, Toronto, Canada © 2015 ACM. ISBN 978-1- 4503-3923-0/15/07
24. S. Tschiatschek, A. Singla, M. G. Rodriguez, A. Merchant, A. Krause, "Fake News Detection in Social Networks via Crowd Signals", WWW '18 Companion, April 23–27, 2018, Lyon, France © 2018 IW3C2, published under Creative Commons CC BY 4.0 License. ACM ISBN 978-1-4503-5640-4/18/04
25. K. Shu, A. Sliva, S. Wang, J. Tang, and H. Liu, "Fake News Detection on Social Media: A Data Mining Perspective", arXiv:1708.01967v3 [cs.SI] 3 Sep 2017
26. E. Ferrara, O. Varol, F. Menczer, A. Flammini, "Detection of Promoted Social Media Campaigns", Proceedings of the Tenth International AAAI Conference on Web and Social Media (ICWSM 2016)
27. J. Song, S. Lee, J. Kim, "CrowdTarget: Target-based Detection of Crowdturfing in Online Social Networks", Publication rights licensed to ACM, ACM 978-1-4503-3832-5/15/10
28. L. Rubin, Y. Chen and N. J. Conroy, "Deception Detection for News: Three Types of Fakes", ASIST 2015, November 6-10, 2015, St. Louis, MO, USA. Copyright © 2015
29. F. Masood, G. Ammad, A. Almogren, A. Abbas, H. A. Khattak,
30. U. Din, M. Guizani, and M. Zuair, "Spammer Detection and Fake User Identification on Social Networks", Special Section on Roadmap To 5G: Rising To The Challenge, Vol 7, 2019
31. Sen, A. Aggarwal, S. Mian, S. Singh, P. Kumaraguru, A. Datta, "Worth its Weight in Likes: Towards Detecting Fake Likes on Instagram", WebSci '18, May 27–30, 2018, Amsterdam, Netherlands © Association for Computing Machinery. ACM ISBN 978-1-4503-5563-6/18/05
32. P. Ratna, B. Satya, K. Lee, D. Lee, T. Tran, J. Zhang, "Uncovering Fake Likers in Online Social Networks", CIKM'16, October 24-28, 2016, Indianapolis, IN, USA c 2016 ACM ISBN 978-1-4503-4073-1/16/10
33. M. A. Zoubi, J. Alqatawna, H. Faris, "Spam Profile Detection in Social Networks Based on Public Features", 8th International Conference on Information and Communication Systems (ICICS), 978-1-5090-4243-2/17/$31.00 ©2017 IEEE
34. J. Jia, B. Wang, N. Z. Gong, "Random Walk based Fake Account Detection in Online Social Networks", 2017 47th Annual IEEE/IFIP International Conference on Dependable Systems and Networks (DSN), DOI:10.1109/DSN.2017.55
35. S. Rathore, P. K. Sharma, V. Loi, Y. S. Jeong, J. H. Park, "Social network security: Issues, challenges, threats, and solutions", Information Sciences 421 (2017) 43–69, © Elsevier Inc.
36. R. Kaur, S. Singh, "A survey of data mining and social network analysis based anomaly detection techniques", Egyptian Informatics Journal (2016) 17, 199–216
37. O. Goga, G. Venkatadri, K. P. Gummadi, "The Doppelgänger Bot Attack: Exploring Identity Impersonation in Online Social Networks", IMC'15, October 28–30, 2015, Tokyo, Japan, c 2015 ACM. ISBN 978-1-4503-3848-6/15/10
38. Viswanath, M. Ah. Bashir, M. B. Zafar, S. Bouget, S. Guha,
39. K. P. Gummadi, A. Kate, A. Mislove, "Strength in Numbers:

AUTHORS PROFILE

Nitika Kadam is Bachelor of Engineering in Computer Science from LNCT Indore in 2010. She holds a Master of Engineering degree in CSE from SDBCT Indore in 2015 and pursuing PHD from Oriental University Indore. With 7 years of teaching experience her research papers are published in international journals.

Harish Patidar holds a Bachelor of Engineering degree in Computer Science and Engineering discipline, from MIT Ujjain in 2005. He Completed his Master of Engineering degree from SVITS Indore in 2013 and SPSU Udaipur awarded him PHD in October 2017. 13 yearsof teaching experience is added in his profile. His 25 research papers are published in international journals, Thomson Reuters, Springer and Scopes Index journals are few of them. A patent is also published during his PHD work.