Semiconductor Optoelectronics
ISSN: 1001-5868

FAKE PROFILE DETECTION ON SOCIAL MEDIA NETWORKS: A SYSTEMATIC REVIEW

Sumitra Menaria
Ph.D. Research Scholar, Computer Science and Engineering, Gujarat Technological University, Ahmedabad, Gujarat,India, Sumitra.menaria@gmail.com

Dr. Viral HBorisagar
Assistant Professor, Computer Engineering Department, VGEC, Chandkheda, Ahmedabad, Gujarat,India, viralborisagar@yahoo.com

Abstract:
People of all ages now spend the majority of their time on social media sites like Facebook, Instagram, blogs, and Twitter because they are a popular way to share information quickly and widely, which increases the number of users. The likelihood of disclosing inaccurate information and falling victim to fraudulent accounts is growing along with the sharp increase in daily visitors on such sites. A fake account is a common means to propagate false information, forward spam, forward URLs that cause phishing assaults, and steal contacts for one's own gain or the harm of rivals. Hence Identification of fake users and spammers on online social networks (OSNs) is a common study topic.. In this paper, we analyzed various affects done by fake profiles and recent techniques like machine learning algorithms such as Support Vector Machine (SVM), Logistic Regression (LR), Random Forest (RF) and K-Nearest Neighbors (KNN) and deep learning method used for detecting fake profiles. In order to prevent damage caused by fake profiles, comparison of various methods for re-identifying or cross-platform verification of profiles is also presented.

Keywords: Fake profile, profile cloning, cross platform identification, online social media network, profile re-identificati;on.

I INTRODUCTION:

Getting connected via social media has become more and more popular in recent years. With social networking, individuals will be able to check for friends with mutual interests or missing ties. As per (Global Social Media ResearchSummary August 2020), half of the world is online now, 4.57 billion people are using the internet now and out of that 3.81 billion people are using online social media in year 2020. Which is an increase of 9.2% year on year (Dean). As per the article by Esteban Ortiz-Ospina(Ortiz-Ospina)founder of our world in data, Facebook is the largest social media platform in the world. It has 2.4 billion users. Other social media platforms including Youtube and Whatsapp also have more than one billion users each[1].

Online social media are more susceptible to fake profiles as they see faster growth. Fake profiles are those that have been constructed with fictitious names and photos in order to get financial or personal advantages. Because so many individuals use social media sites online, businesses may reach a large audience there to sell their goods.

Many people utilise fictitious identities to obtain information for study and a variety of other purposes. Therefore, recognising fraudulent profiles is a crucial problem that needs to be solved for OSN's security and user privacy.

People creates fake profiles on OSNs for many reasons, some of the reasons are

1. **Link Farming:** when users attempt to increase their connections or followers. It can be used to persuade individuals about a certain issue or to target a sizable population for disseminating misleading information. This link is utilized by spammers in phishing attacks [2]–[5].
2. **Identity Theft:** When a fraudster poses as someone else and attempts to steal the victim's identity for their own or an organization's gain [4]–[7]
3. **Cyber bullying:** Unauthorized users have the ability to utilise false profiles to transmit offensive material or engage in cyberbullying. [8], [9]
4. **Stealing Personal Information:** Through intense online interactions with friends, fake users can target authenticated users and steal the victims' personal information [10], [11]
5. **Black Marketing:** Selling accounts with substantial follower numbers is a multimillion dollar endeavour. People build large fan bases, which they then sell to businesses. Businesses utilise them to promote their goods. [8], [11], [12]

There are different type of fake profiles, some of the examples are:

1. **Sybil Account:** According to Al-Qurishi et al. [1], Mateen et al. [8] and Mezhuyev et al. [10],a hacker will manually create many accounts in an effort to jeopardise security and privacy. This kind of attack is frequently used by businesses to boost their ratings on social media networks or e-commerce websites. Such profiles are typically utilised to obtain the most impact when engaging in criminal activity. The option to create and manage several accounts or identities, which appear to be truly unique identities, is present here on the same node.

2. **Sockpuppets:** According to Kacchi and Deorankar[13], Krishnan et al. [7], Mateen et al. [8], and Singh et al. [14], sockpuppet was made as a fictional online identity with the intention of deceiving people. Sockpuppets are frequently mass-produced by a single commanding person or group. They are typically employed for block avoidance, fabricating popular opinion, stacking ballots, and other related activities.

3. **Social Bots:** Boats are software programmes designed to perform certain tasks without the need for human intervention, according to Mateen et al. [8], Mezhuyev et al. [10], Narayanan et al. [15], and Xiao et al. [18]. They keep users occupied and act like humans. Bots frequently use synthetic personas to interact with people and create less recognisable social networks. In online social networks, it can also be used to influence people, publish comments, and send friend invites.

Various techniques for spotting the fake account have been looked into in the study that has been presented. In part II, related work to identify fake accounts has been discussed. Section III presents several features and databases that are available for false profile detection. Future actions are included in Section IV.

II RELATED WORK

Numerous thorough research have already been conducted to identify fake profiles on OSNs using a variety of techniques. In order to prevent fraudulent people from being added to our profiles, the first of three categories for fake profile identification is the analysis of fake friend requests[2]–[4], [6]. This sort of method is referred to as a pre analysis method. Using machine learning approaches, such as post analysis methods, the second way is to identify bogus profiles [5], [7], [8], [10], [12], [13]. The third method, sometimes referred to as profile re-identification, involves comparing a person's profiles across several platforms to determine whether they are authentic or not.

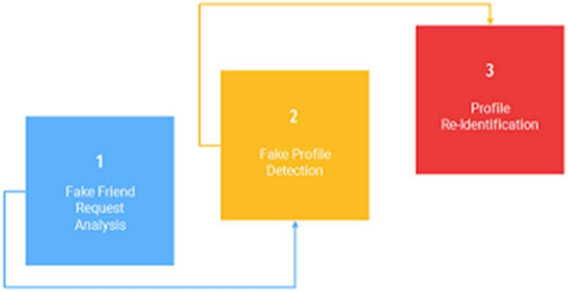

Fig 1. Types of analysis done on Social Media Profiles

A. Fake Friend Request analysis method:

As an illustration, imagine that A receives a friend request from B. First, manually count the number of friends that you have in common. A will trust B and accept B's friend request because A discovered that C and D have mutual friends. Although C doesn't actually know B, he accepted his friend request after spotting D, another common buddy of theirs. In this manner, a connection with false users is made using indirect trust. Therefore, [2]has attempted to establish trust amongst various nodes in order to identify authentic friend requests. The techniques listed below are used to identify bogus friend requests.

System for making decisions: A strategy for trustworthy decision-making over accepting friend requests on social media networks to determine "friend to be" has been put forth by Rahman et al. [12], Vitaliy, and Zakirul. The input data in this case is a graph whose edges are friend requests and whose vertices are user profiles. In this study, a three-stage strategy is adopted. The first step is a streamlined way in which the user profile of the person who submitted the friend request and its qualities are compared.

The second stage, which is an enhanced method, analyses the user's profile attributes as well as the profile of the person who submitted the friend request and the profiles of that person's friends. To determine the reliability of the buddy request, the third step compares the attributes obtained in the previous two steps.

A friend recommendation system based on semantics has been proposed by Wang et al. [16].

A similarity matrix was created using Freindbook as a data mining-based method based on the user's lifestyle. The similarity between users' living styles can be represented by a module by creating a friend matching graph. The major flaw in the proposed system, however, was data collecting.

The friend suggestion system created by Kacchi and Deorankar[13] is based on a variety of criteria, including rating and factors like shared interests, blood type, proximity to one another, and comparable blood types.

A trustworthy mechanism for making decisions has been developed by Mezhuyev et al. [10] to intervene in favour of the Request Acceptance.The author of this study analysed user and friend-to-be qualities as well as those of friend-to-be friends.For improved buddy recommendations, different characteristics like liking and disliking, mutual friends, behaviour, etc. were compared.

B. Fake profile detection:

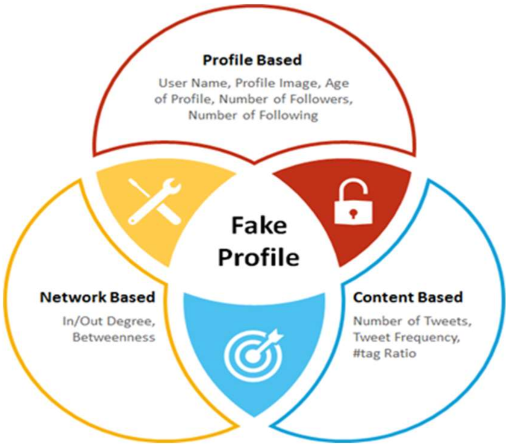
Profile Based
User Name, Profile Image, Age of Profile, Number of Followers, Number of Following
Fake Profile
Network Based
In/Out Degree, Betweenness
Content Based
Number of Tweets, Tweet Frequency, #tag Ratio

Figure 2. Methods for detecting fake profiles

A scalable method to locate a series of fraudulent logins created by the same person was developed by Xiao et al. [3]. A cluster of accounts was classified as malicious using supervised machine learning techniques. In order to condense user-generated data into a manageable area on which statistical features could be computed, the pattern encoding algorithm was created. In this study, support vector machines, logistic regression, and the random forest approach were applied. However, the system only supported English language alphabets and was unable to withstand hostile attacks.

For the purpose of identifying spammers on the Twitter platform, Mateen et al. [8] created a hybrid technique that makes use of both content- and graph-based features. In order to categorise profiles as spammer or legitimate accounts, J48, Decorate, and Nave Bayes classifiers are utilised. There were several features employed that may be removed utilising correlation without affecting the results.

Al-Qurishi et al. [1] proposed a system that had three modules: a data gathering module, a feature extraction method, and a deep regression module. Twitter users were divided into three groups based on their profiles, content, and networks by the data harvesting module. Additional categories for content-based data included temporal, topical, linguistic quality, and emotion-based categories. The only issue with the method is that sufficient data filtering is necessary despite its great accuracy.

A approach was suggested by Singh et al. [14] using a Support Vector Machine and a Neural Network. A machine learning technique was used to classify the extracted profile features as either bogus or real. For classification, a sizable amount of data from fake projects and the 2013 election was used.

IronSense is a machine learning-based browser plugin created by Narayanan et al. [15]. Their study concentrated mostly on separating important characteristics—such as the quantity of friends, followers, and status updates—between phoney and real users. SVM, logistic regression, and random forest models were employed as machine learning techniques. However, there is still room for development in terms of forecast accuracy and efficiency.

The false profile identifier presented by Krishnan et al. [7] is based on finite automata. Based on the individuals' birthplace, educational institutions, and place of employment, a regular expression is constructed. To check for similarities in login, users who sent friend requests were compared using regular expression. However, for the person who has friends in more communities, the regular expression that was generated was noticeably lengthier.

A deep learning-based system called DeepProfile with the WalkPool pooling function was created by Wanda and Jie[17].It uses a dynamic deep neural network and supervised learning to identify if a profile is legitimate or fraudulent based on an examination of its attributes. This strategy made advantage of three types of manufactured features: the account's identification, its relationships with other accounts, and its behaviour. Malicious accounts were identified by using designed traits and based on deviations in profile information.

C. Cross Platform re identification of the profile

Re-identifying people on online social media through cross platform is the area of research that has received the least attention.

It aids in determining a person's originality. Re-identification gets its name from the fact that it also offers confirmation.

For cross-platform user re-identification, researchers have focused on fact-based, relationship-based, and model-based solutions. Cross-platform identification can lower the false negative ratio. Supporting evidence can be gathered using the re-identification process to label any profile as fraudulent.

Hill and Nagle's [18] four-step approach for comparing user behaviour on two separate network entities across two different time periods was devised [5]. The random graph approximation technique used in this framework allows for estimation without comparison. Although less noise-tolerant than other methods, this one was resistant against missing links.

Since Zafarani et al. [19] demonstrated that solely link-based cross identification is insufficient, user re-identification was done using both link and content information. To find correlations between various profiles on various SMNs and crossed-over friends on various networks of the same base node, link information was used. For systematic analysis, a further behavioral modelling technique was taken into consideration. The success of the behavioral modelling approach for identification was due to the human beings' redundant behavioral patterns. Patterns of behavior were divided into three categories: "exogenous influences," "endogenous variables," and "pattern owing to human limitation."

A hybrid technique based on personal information was created by Ahmad and Ali [20]. Utilizing network and content features helped with re-identification. The author made use of a network feature, network followers, cross-posting, and one content feature. To determine whether two nodes on separate networks were comparable, the native approximation distance method was utilized. Crosslinks on distinct social media networks penalized the seed user, and crawl lists of users from both networks using cross link properties were also penalized. Levenshtein distance was calculated between attributes, and attributes with zero distance were regarded as the seed user. The seed user was then used as an input to gather relationships and followers in order to find user similarities across platforms.

Hashimoto et al. [5]Anonymized data are linked with different types of side data using machine learning algorithms.
The user's social media accounts and resumes were used as the target data, with side data being used for comparison. The author made an effort to address the issues of data availability and side data's inappropriateness. A machine learning technique was used to link side data with profile traits because resumes are readily available on any public platform, making data more easily accessible.

Using the Jaro-winkler similarity technique, Yadav S. et al. [11] created a method to detect identical user logins on Twitter. The three-layered strategy was only relevant for identifying duplicate logins on the same network and employed profile attributes to search users, content attributes to locate similarities in postings, and network attributes to find mutual friends and link connections.

Ali and Ahmad [21] Using information that was criticized by the user on Twitter about their other accounts on other social media, Tweeter was utilized to construct a data set for study. The unique trait to identify similar features was the user's screen name. To compute similarity between various accounts on social media, token-based and character-based distance measure algorithms were used.

III. DATASET AND FEATURE USED FOR FAKE PROFILE DETECTION

For the purpose of detecting false friend requests, fake profiles, and cross-platform profile re-identification, numerous researchers have created various techniques. For each of the three categories, features used by researchers are presented in this section. In table I, the specific profile-based features, content-based features, network characteristics, and temporal features employed by various studies are listed.

TABLE I. FEATURES USED FOR FAKE PRIFILE DETECTION

Method	Feature Type	Feature	Reference
Fake Friend Request Analysis	User Feature	Number of followers	[5], [8], [9], [12]
		Number of Following	
		Number of Replies	
		Number of Repost	
		Age of Account	
		Number of Replies	
	Content Features	Number of Retweets	
		Number of hash tags	
		Number of User Mentioned	
		Number of URL	
		Number of Retweets	
	Graph Features	In Degree	
		Out Degree	
		Betweenness	
	Temporal Feature	Time of tweet	
		Length of tweet	
		Tweet Frequency	
		Tweet sent in time interval	
		Ideal time in days	
Fake Profile	Profile Based features	Verified account (Y/N)	[1], [7], [9], [10], [13], [15]
		#Char of Screen name	
		#Digits of Screen name	
		Time Zone	
		Default profile picture (Y/N)	
		Default profile cover (Y/N)	
		Account age	
		Has profile Description (Y/N)	
		Profile description length	
		Bio has URL (Y/N)	
		#URL in Bio	
		Contains social networks contacts	
	Content-based Features	Temporal Features	
		Topic-based Features	
		Quality-based Features	
		Emotion-based Features	
	Network-based Features	# friends	
		# followers	
		# favorites	
		# tweets	
		# retweets	
		# mentions	
		# replies	
		# retweeted tweet	
		# friends distribution	
		# followers distribution	
		# favorites distribution	

		#replied by others	
		#retweeted by others	
		#mentioned by others	
		#favorite by others	
Profile Re-identification	Profile attributes	First name	[11], [19]–[21]
		Last name	
		Gender	
		Location	
		Education	
		Profession	
		Email	
		Language	
		Date of birth	
		Tag line	
		Profile URL	
		Location	
	Content attributes	Tweets	
		Video posts	
		Image posts	
		YouTube Links	
	Network attributes	Friendships	
		Group membership	
		Fan page participations	
		Connections	
		Followings	
		Followers	

Many different datasets are utilised in various methods for detecting false profiles. Details regarding the datasets are provided in this section. The Stanford Large Network Dataset, which Cao et al. [4] have worked on, is essentially a collection of online social networks where the edges indicate interactions between people. The dataset includes Facebook and Twitter data with 4039 nodes and 88,234 edges connecting them.

On Facebook dataset, Zhang et al. [22] expanded their investigation. There are 63,731 people and 817,091 links in this dataset.Wang et al [9] .'s research utilised samples from the CREDBANK and PHEME Twitter datasets, which each contained information on 38,000 people and their connections. The data set offered by Mateen et al. [8] contains 467480 tweets and 10256 people, and is mostly useful for profile re-identification and content-based analysis. Facebook datasets are also available for research and can be pulled from Twitter using the Twitter API. The Github repository also offers a social media dataset with 2820 nodes, 1482 of which had authentic user information and 1338 false. Many studies lack information regarding the dataset they utilised due to concerns about confidentiality.

For detecting fake friend requests [9], [12], [14], [15], [17], the features listed in Table I are divided into three categories. Fake friend requests can be examined by four distinct types of groups, such as user profile-based features, which include the followers, activities, and friends of a certain account. The age of the account is another important factor to consider when comparing the amount of followers and followings over a specific period of time, which can help determine if an account is real or fraudulent. To determine if an account is false or real, another way is to evaluate content characteristics such the quantity of tweets, retweets, and tags.

A friend request can be identified as coming from a phoney account if it has sent more friend requests than it has received in comparison to other requests, according to a graph feature that includes studies of in and out degree and betweenness. Researchers may be able to determine whether a request is the result of a bot by analysing temporal features. It is possible to recognise tweets sent by bots by looking at the timing and space between tweets.

The second category is to identify fake profiles. Screen names and profile names are important factors in classifying a profile as fake or genuine. If the profile name is the same, the profile image is the next crucial factor in determining if an account is real or false.

The age of the profile is the following crucial factor. If a profile's name and photo match, we can compare the ages of the two profiles and declare the one with the younger age to be a fake. Features that are based on content aid in locating a certain person's social media activity. Identification of suspect activity is aided by the type of content shared and posted by profiles. The amount of posts made from a given account, the number of friend requests that come in and go out for that profile, the number of tweets or posts that are retweeted by friends in common, and other metrics were tracked by researchers using network based features to identify fraudulent profiles. All of the aforementioned aspects aid in increasing the system's ability to recognisefake profiles.

Profile re-identification is the third category of profile analysis [6], [19]–[22] on social media. Once we identify a suspicious account, it's crucial to cross-verify the account with other sources, such as other social media accounts linked to the suspected account. Profile-based attributes such profile name, date of birth, qualifications, and education can be examined for cross-verification.

Sharing the same post with similar-looking text, video, or image on different social media platforms allows users to compare many accounts and improves the detection system's accuracy. Similar buddy group on several social media networks is an example of a network property. A key factor in profile re-identification is the number of connections, followers, and following on multiple social media accounts.

IV CONCLUSION:

The many methods for spotting fraudulent profiles on social media networks have been explored in this study, together with the datasets and analytical features. Our research into the literature revealed that while various approaches have been developed to identify bogus accounts, this alone is insufficient. In order to prevent users from connecting with phoney profiles, techniques for warning them about incoming fake friend requests should be created.

Although there has been a lot of study on spotting fraudulent accounts, we discovered that much of it focuses on content-based methods, with relatively little work being done on feature-based, network-based, or graph-based methods.These two techniques can increase the effectiveness of finding fake accounts.

Since most of the work has only been done with offline data, methods that also function with online data should be created. Utilizing diverse social networks and big data analysis for the detection of false accounts in social networks, a new study in the area of collaboration and cross-platform analysis for re-identification of profiles can be conducted. By replacing predetermined labelled data with unstructured behavioral data and sentiment analysis of user social activity, we may further improve fake profile identification.

REFERANCES

[1] M. Al-qurishi, M. Alrubaian, S. M. Rahman, and A. Alamri, “A Prediction System of Sybil Attack in Social Network using Deep-Regression Model,” Futur. Gener. Comput. Syst., 2017, doi: 10.1016/j.future.2017.08.030.

[2] B. Dean, “Global social media growth rates.” .

[3] Q. Cao, M. Sirivianos, X. Yang, and K. Munagala, “Combating Friend Spam Using Social Rejections,” 2015, doi: 10.1109/ICDCS.2015.32.

[4] Dave Chaffey, “Global social media researchsummary.” .

[5] M. Ichino and H. Yoshiura, “A Re-Identification Strategy Using Machine Learning that Exploits Better Side Data,” 2019 IEEE 10th Int. Conf. Aware. Sci. Technol., pp. 1–8.

[6] D. M. Freeman and T. Hwa, “Detecting Clusters of Fake Accounts in Online Social Networks Categories and Subject Descriptors,” pp. 91–101.

[7] P. Krishnan and D. J. Aravindhar, “Finite Automata for Fake Profile Identification in Online Social Networks,” no. Iciccs, pp. 1301–1305, 2020.

[8] M. Mateen and M. Aleem, “A Hybrid Approach for Spam Detection for Twitter,” pp. 466–471, 2017.

[9] A. Zhibo Wang, Jilong Liao, Qing Cao, Hairong Qi and Z. Wang, “Friendbook: A semantic-based friend recommendation system for social networks,” in IEEE Transactions on Mobile Computing, 2015, pp. 14(3):538–551, doi: doi: 10.1109/TMC.2014.2322373.

[10] V. Mezhuyev, Z. A. Bhuiyan, S. M. N. Sadat, S. Aishah, B. Zakaria, and N. Refat, “Reliable Decision Making of Accepting Friend Request on Online Social,” vol. 4, no. c, 2018, doi: 10.1109/ACCESS.2018.2807783.

[11] S. Yadav, A. Sinha, and P. Kumar, “Multi - attribute identity resolution for online social network,” SN Appl. Sci., vol. 1, no. 12, pp. 1–15, 2019, doi: 10.1007/s42452-019-1701-z.

[12] V. Mezhuyev, S. Member, S. M. N. Sadat, and A. T. Asyhari, “Evaluation of the Likelihood of Friend Request Acceptance in Online Social Networks,” IEEE Access, vol. 7, pp. 75318–75329, 2019, doi: 10.1109/ACCESS.2019.2921219.

[13] T. R. Kacchi and P. A. V Deorankar, “Friend Recommendation System based on Lifestyles of Users,” 2016.

[14] N. Singh, T. Sharma, A. Thakral, and T. Choudhury, “Detection of Fake Profile in Online Social Networks Using Machine Learning,” 2018 Int. Conf. Adv. Comput. Commun. Eng., pp. 231–234, 2018.

[15] A. Narayanan, “IronSense : Towards the Identification of Fake User-Profiles on Twitter Using Machine Learning Department of Computer Science.”

[16] J. Jia, B. Wang, and N. Z. Gong, “Random Walk based Fake Account Detection in Online Social Networks,” 2017, doi: 10.1109/DSN.2017.55.

[17] P. Wanda and H. Jin, “Journal of Information Security and Applications DeepProfile : Finding fake profile in online social network using dynamic CNN,” J. Inf. Secur. Appl., vol. 52, p. 102465, 2020, doi: 10.1016/j.jisa.2020.102465.

[18] S. Hill, “Social Network Signatures : A Framework for Re-Identification in Networked Data and Experimental Results,” 2009, doi: 10.1109/CASoN.2009.31.

[19] R. Zafarani, L. E. I. Tang, and H. Liu, “User Identification Across Social Media,” vol. 10, no. 2, 2015.

[20] W. Ahmad and R. Ali, “A Framework for Seed User Identification across Multiple Online Social Networks,” pp. 708–713, 2017.

[21] R. Ali, W. Ahmad, and R. Ali, “ScienceDirect ScienceDirect Social Account Matching in Online Social Media using Cross- Social Account Matching in Online Social Media using Cross- linked Posts linked Posts,” Procedia Comput. Sci., vol. 152, pp. 222–229, 2019, doi: 10.1016/j.procs.2019.05.046.

[22] Z. Zhang, S. Su, X. Liu, Y. Guo, and J. Zhang, “Efficient Multi-pair Active Friending in Online Social Networks,” 2018 IEEE Glob. Commun. Conf., pp. 1–6, 2018.