Journal of Language, Technology & Entrepreneurship in Africa (JOLTE)
Vol. 12 No. 2 2021

Fake Profile Identification on Online Social Networks

Catherine Njoki Muraya
cnmuraya@gmail.com

Fredrick Mzee Awuor
School of Information Science and Technology
Kisii University, Kenya.
fawuor@kisiiuniversity.ac.ke

Benard Magara Maake
School of Information Science and Technology
Kisii University, Kenya.
bmaake@kisiiuniversity.ac.ke

Abstract

Online social networks are web-based applications that allow user to communicate and share knowledge and information. The number of users who make use of these platforms are experiencing rapid growth both in profile creation and social interaction. However, intruders and malicious attackers have found their way into the networks, using fake profiles, thus exposing user to serious security and privacy problem. Every user in the online social network should verify and authenticate their identities, with the other users as they interact. However, currently verification of user's profiles and identities is faced with challenges, to the extent that a user may represent their identity with many profiles without any effective method of identity verification. As a result of this vulnerability, attackers create fake profiles which they use in attacking the online social system. In addition, online social networks use a logically centered architecture, where their control and management are under a service; provider, who must be entrusted with the security of data and communication traces; this further increases the vulnerability to attacks and online threats. In this paper, we demonstrate the causes and effects of fake profiles on online social networks, and then provide a review of the state-of-the-art mechanism for identifying and mitigating fake profiles on online social networks.

Keywords: *online social networks, fake profiles, sybil attack, fake accounts*

1. Introduction

Online Social Network (OSN) is a graphical structure with nodes and edges that represents users and their interactive activities respectively Depending on the type of the network being used, edges and nodes are either labelled or unlabeled (Khan and Lee, 2018). The links are the various forms of ties or interdependencies like friendship or similar interest, developed between nodes

that are connected through the Internet. This enables users to strengthen their social ties with other users who have similar interests, activities and real-life inter connections (Coletto, Garimella, Gionis & Lucchese, 2017). They also create personal profiles which they use to connect, exchange updates, and meet new individuals with similar interests (Reddy, Dey & Sinha, 2019). Previous research, adds that, within the social networks, users are therefore identified by a profile, which may contain various features like a picture, name, address and birth date (Aghasian, Garg, Gao, Yu, & Montgomery, 2017). However, research work by Romanov, Semenov & Veijalainen, (2017), indicate that there is no suitable method in place to thoroughly verify that the person whose identity is referred to in the profile, truly created and manages it, implying that someone could be exploiting someone else's identity (Romanov et al., 2017).

Through research work by Jia and Xu (2016), OSN, has evolved from a special phenomenon to a widespread recognition, where users interact and share information, as result, transforming how individuals communicate with one another. The platforms have given their users unimaginable privileges, including the ability to post events in real time, even before they are published, and to access never-ending uncensored material. (Adedoyin et al., 2014)

From the above-mentioned aspects, it is clear that, many people and organizations are highly depending on online social networks, not only for information exchange but also decision making. However, this popularity combined with user's minimal interaction control including flexibility to distribute content with no supervision has attracted the interest of malicious attackers, who join the networks to carry out malicious activities (Aghasian et al., 2017). Therefore, Meligy et al., (2017) conclude that, the platform becomes favorable grounds for scammers who impersonate user's identities by creating fake profiles which they use to carry out online attacks and activities that harm users' reputation and privacy. Fake profiles are used to spread false information and steal follower's identities, (Chen et al., 2018). The increased attacks on OSNs have implications on the privacy and security of individual user together with their data. Therefore, it has become necessary to design and implements appropriate methods that will mitigate the creation of fake profiles on online social networks (Hardjono and Smith, 2019). Once this is successfully done it will minimize online threats and users will be able to approach social networks more confidently without fear of breach of their privacy and data security.

2. Problem Definition

The amount of data created, stored and shared within OSN is rapidly increasing, leading to a tremendous growth of these platforms. Today the efficient use of OSN data play a key role in sustainable economic development. As a valuable asset, data is providing significant advantages to companies in optimized decision-making process and predicting future trends which can inform people's economic preferences. However, with all these benefits, there is a growing concern about user security and privacy within the platforms. Intruders exploit shared information to create fake profiles for carrying out different malicious activities. Using the fake profiles, they solicit connection to the targeted victim and if the victim accepts the requested connection, the attacker gains access to the targets profile, which would normally be accessible to trusted contacts only. Furthermore, each victim's link to a fake profile, increases the profile's

seeming credibility, that can attract more connections, thus increasing vulnerability. The behavior of users on fake profiles differs from that of genuine people and even though on the surface, their features look identical, it's difficult to distinguish them without a reliable technique. The intruders exploit trust relationship, exposure and weakness of OSN platform to collect user's data without their knowledge and control. In addition, the centralized nature of OSNs make it favorable to amass large quantities of users personal and sensitive data, since they have little or no control over their stored data including, how it is used. Therefore, identification of fake profiles becomes a challenge, moreover, the rate at which social network data is currently growing and evolving further complicates it. To this end, this paper seeks to answer the following sub-problems:

I. What are the causes and effects of fake profiles on online social networks?

II. What is the state of the art mechanism for identifying and mitigating fake profiles on online social networks?

3. Online Social Networks: Definition, Structure and Architecture

Online social network (OSN) is a social structure where user interact, is built on a common thread of interest, either directly or indirectly. (BalaAnand, 2019). According to Penni (2017) OSN is a collection of web-based tools that are purposely built and dedicated to social interactions and facilitate communications for online-based technologies like Facebook. According to Kaplan and Haenlein (2010), social networks are a collection of web-based application that enable users to connect with others who share their interests and exchange user-generated content., According to previous research, the platform is an interactive computer-mediated technology that allows people to create and share information, ideas, career interests, and other kinds of expression through virtual communities and networks, therefore people perform various activities within the platform. Other study (Huang et al., 2014) indicate that, as a social structure the platform is made up of nodes, which are connected by one or more specific types of interdependencies such as, common interest and dislikes. OSN allow people to maintain their relationships, engage with acquaintances, and create new relationships with others based on shared features such as community, interests and interests.

Other study refers to OSN as a service, for instance, according to Boyd and Ellison (2007) online social networks, are web-based service that allows users to construct a public or semi-public profile inside a constrained environment. Other users then join the system, which links them to other users in the system. Adamic and Adar (2003) further adds that, OSNs collect information from the social connections of the individual user, create a vast social network that is interconnected and show users how they are linked to other individuals in the network.

According to a different study, OSN is an online community comprised of people who have similar activities. On the same note, (Schneider et al. 2009), adds that as an online community people, share common interests, activities, backgrounds, and/or friendships. The same meaning is shared by (Nair and Anbalagan 2021) who notes that, OSNs are systems which consist sets of individuals, nodes and the relationships between them, that is, the ties between individuals, over

a common platform on the Internet (Nair & Anbalagan, 2021). These relationships between entities or nodes are based on similar interests exhibited in the network or in real life.

The structure of online social networks can be recognized as a complex system, which are characterized by patterns of interactions and connectivity. Other authors, classified OSN into multiple categories based on the functions they offer to the user, which include applications that help individuals develop and maintain social relationships, application that make it easier to exchange media, and forums that allow people to share their expertise, news, and opinions (Aichner and Jacob, 2015). This includes systems that provides private social networks, business social networks, academic social networks, video sharing social networks, Instant messaging social networks (Adewole 2017).

OSNs provide diverse functionalities to its members, users can utilize a variety of applications (both paid and free) to market their businesses, connect with potential customers, and sell product online (Wani and Jabin, 2018) For instance, large commercial organizations are utilizing Twitter as a tool to extract business value because it allows customers to market their businesses and brands and buy directly through tweets (Aichner & Jacob, 2015) Politicians utilize OSN to boost their political efforts, while gamers, particularly young people, play multiplayer games or single alone games on the Facebook page. They are used by the general people for entertainment, reading news, sharing information, uploading media, and tagging their friends (Quattrone et al 2015)

Carlsen and Brincker (2020) argue that social networks provide robust and adaptive assistance, particularly during emergencies like the COVID pandemic, where people' needs are unpredictable and necessitate constant monitoring and adaptation. The platforms aids in tracking the epidemic by shedding light on both individual and social behaviors of people through big data analysis via Online and Mobile Social Platforms that assist in the tracking of people's physical contacts (Carlsen et al., 2020). As a user-generated content driver, OSN applications, are highly influential in a variety of situations, including purchasing/selling behaviors, entrepreneurship, political concerns, and venture capitalism. (Greenwood and Gopal 2015).

Research done by Chowdhury et al., (2014), highlights that, OSNs utilize client – server architecture, which, is a centralized implementation of the networks, for example, service providers like Myspace, Instagram, Facebook and Twitter, are the central providers for all services including access, storage and maintenance to those specific networks. The client–server-based model is controlled by a single administrative authority and all the users details together with their generated content are stored and owned by single service providers and with total control over all of the user's information (Romanov et al., 2017).

Different literature has presented other key aspects of online social network designs, for instance Pallis et al., (2011) indicated that the entire online social system is formed by three key layers the data storage, content management and application layer. In addition, the platform is made up of several application servers that offer a variety of services and APIs. A user sends a request to the OSN platform, which then sends it to the correct application server. User connection relationships are tracked and managed by graph servers, in terms of content distribution, the authors noted that, cache servers reduce the strain on application servers, allowing dynamic web applications to run faster. In addition, they are user relationships and trust that evolve over time.

OSNs consist of a large number of small files that must be read and updated often by a large number of users, as well as propagation of file updates to ensure data coherency (Pallis et al., 2011).

4. Fake Profiles on Online Social Networks

Fake profiles are a types of identity theft in which a real user is impersonated for a variety of malicious motives (Fire et al., 2014), that are meant to be used for purposes that violate OSN's terms of service, such as spamming (Gupta and Kaushal, 2017). According to other study, fake accounts are also employed in adenosine triphosphate situations to gather intelligence, develop trust, and distribute malware or a link to it (Ojo 2019). Other forms of malicious actions also make use of such fake profiles (Romanov et al., 2017).

Research by Josh et al (2020) concludes that as the number of people utilizing OSNs grows, so does the number of fake social media profiles and accounts, developed for cyber extortion or commit cybercrime in an anonymous or untraceable manner. Fake accounts owners may try to take advantage of people's goodwill through false announcements or disseminating misleading news in order to defraud them of their money. Furthermore, users create many accounts that are not owned by anyone, with intentions of increasing the number of votes in online voting systems and online gaming in order to earn referral bonuses (Joshi et al., 2020).

As a result of increased usage of OSNs, many users become exposed to privacy and security threats unknowingly, in addition a big number, are oblivious to the security implications of these threats, which include privacy violations, identity theft, and sexual harassment. (Fire et al., 2014). According to research done, Users are willing to reveal personal and confidential information about themselves, which can be used to damage individuals in both the virtual and physical worlds.

Many people use social networking sites to share their photos and those of their friends. Every day, tens of millions of images are uploaded to Facebook (Fire et al., 2014). In addition, many Facebook user profile pictures are viewable and downloadable by the public. For example, the Faces of Facebook website (Chu et al., 2012) allows, people to examine the profile images of almost 1.2 billion Facebook users. These images can be used to build a biometric database that can be used to track down OSN users without their permission.

Fake profiles can be used to launch Sybil attacks, send spam, and even tamper with OSN data (Chu, Z et al., 2012). According to a recent study, the market for buying false followers and retweets is already a multimillion-dollar industry. However, Boyd and Ellison, (2010) notes that extensive damage can be caused by a few but well-manipulated fake profile, the authors illustrated this with the fictional profile of Robin Sage that connected hundreds of users from different OSNs (Boyd & Ellison, 2010)

Research by (Josh, et al 2020) indicate, that fake profiles are formed in order to carry out cyber extortion or perform cybercrime in an anonymous or untraceable manner. They sometimes take advantage of people’s kindness by spreading false announcements or by disseminating fake news through these accounts to usurp money from innocent people. Moreover, the scammers create multiple fake profiles to get a hike of votes in online voting systems and also in online gaming in

the greed of getting referral incentives (Joshi et al., 2020). The victim gains a false advantage against the other members.

On a larger scale, social networks such as Facebook impacts of fake profile can span millions, and even billions, of users, and suffer from the threats created to spam honest users, forge identities or any other range of uses. (Gupta et al., 2015). Furthermore, these attacks can carry harmful consequences, for example applications such as peer to peer file sharing with no trusted authority run the risk of data loss if an adversary can compromise enough of the system to outvote honest users. Jia et al., (2016) adds that an attacker may pose as a trustworthy user and diverts the honest user to other pages in order to propagate malware, which results in a phishing attack. (Jia et al., 2016). An adversary establishes many fake profiles in order to obtain sensitive information from users such as complete names and bank account numbers in order to commit a range of cyber-crimes. (Sahoo and Gupta, 2019).

One of the most common and practical attacks assaults against OSN platforms is the Sybil attack (Meligy et al., 2017)). In this attack, the adversary attempts to impersonate real users across OSN in order to obtain unfair trust to a certain user or community. Yang et al., (2014) shows that since users' profiles and other online threats on OSN platforms are not protected from sybil attacks due to a lack of effective authentication procedures, the attacks are done successfully. To execute spy and eavesdropping activities, OSN criminals can build false profiles by duplicating the existence of a certain account or by creating fake accounts. User's profile has become one of the targeted resources by spammers, who leverage the trust relationship among users to acquire more victims on OSNs.

5. Identification and Mitigation Fake Profiles on Online Social Networks

Researchers have identified various methods and techniques for detecting fake profiles and their associated behavior on online social networks. The proposed methods rely on several categories of algorithms and profile features (user-based, graph-based, content-based, and time-based). (Joshi et al., 2020). The techniques, according to Adewole et al (2016), can be divided into three main categories: crowdsourcing, graph based, and machine learning.

5.1 Crowdsourcing Approach

To identify fake accounts, the crowdsourcing method employs a human detection approach. To discover suspicious actions, the techniques entail the use of large and distributed set of workers known as crowd workers. They analyze the social network accounts and determine whether they are Sybil or authentic by looking at the information on their profiles.

Algorithms in this category include:

According to Wang et al. (2010), crowdsourcing algorithms, are used to identify fake profiles on social network. The authors emphasized that these techniques use human detection methods to detect fakes, by examining clickstream data in order to detect the presence of suspicious accounts. The authors designed a multi-tier Sybil detection technique based on crowdsourcing.

This study was one of the first steps in the creation of scalable and accurate crowdsourced Sybil detection system (Wang et al., 2010).

The evaluation of how crowds' wisdom can be used to assess security in the websites, including phishing was done by (Moore and Clayton 2008) Their study indicated that, users' involvement rates follow a power-law distribution, the accuracy of users' reports varies, and users with more expertise tend to have higher accuracy, thus becoming easier to detect fake profiles.

Wang et al. (2013) worked on crowdsourcing research on Amazon's Mechanical Turk for the task of detecting Sybil on online social networks. The authors, highlight that there is a lot of variation in terms of crowd users in their reporting accuracies, which needs to be taken into considerations for building a practical system. The same is shared by Freeman (2017), who identifies the disadvantages of using user's feedback for Sybil detection in social network sites. The authors used LinkedIn datasets to inform the research, which indicated that only a small number of skilled users (who have good accuracy that persists over time) for detecting fake accounts.

5.2 Graph Based Techniques

SybilGuard algorithms were proposed by Yu et al., (2008) the methods relied on based on fast-mixing assumption and random walks. A decentralized protocol that reduces the corruptive influence of Sybil attacks, including those that use IP harvesting and those launched from botnets outside the system. The design is based on a unique understanding of social networks, where, identities are nodes in the graph and (undirected) edges are trust relationships built by humans (e.g., friend relations). Attack edges are the connections between the honest area (which contains all the honest nodes) and the Sybil region (which contains all the Sybil identities created by malicious users). The procedure ensures that the number is not exceeded of attack edges is independent of the number of sybil identities, and is limited by the number of trust relation pairs between malicious users and honest users (Yu et al. 2008).

Yang et al., (2016) developed VoteTrust algorithm to detect fake accounts using signed graph. VoteTrust, is a system that uses user's activities such as initiating and accepting links to guard against Sybil attacks, it provides security assurances by demonstrating that Sybil can only send a limited number of requests to actual users. VoteTrust beats traditional ranking systems in real-world testing, by detecting real Sybils with great precision.

Viswanath et al. (2010) investigated Sybil defense strategies and devised a community-based Sybil detection method. Local communities, where the groupings of nodes are more tightly knit than the rest of the graph, around a trusted node are detected by Sybil defense techniques. The authors show that well-defined community structures are naturally more vulnerable to Sybil attacks, and that Sybils can deliberately target their ties in such networks to increase the effectiveness of their attacks. Sybil defense systems rely on detecting social network communities.

On the same note, Liu et al., (2015) suggested a two-step strategy for a community-based solution. The first phase groups accounts into communities, and the second assigns a label to each account in the community based on the accounts' and community's shared characteristics.

The greater the number of communities to which two accounts belong, the more similar they are. The lack of scalability is one of the algorithm's most noticeable challenge. Yu et al. (2010) suggested a clustering approach for separating spam campaigns from wall messages.

Devineni et al., (2015) introduced the PowerWall algorithm, which is based on a social graph's modified power law property. Users of Facebook wall activities were analyzed using PowerWall, a power law distribution extension, in an attempt to catch problematic accounts. The writers looked at users' wall posts and found odd patterns, such as accounts that post the same number of messages every two days and another that posts every night without doing any other activity.

5.3 Machine Learning algorithms

Machine learning algorithms were employed to detect fake profiles. The techniques use a number of methodologies and approaches, including supervised, unsupervised and semi supervised methods. A supervised machine learning technique takes a labeled dataset and creates a model that can predict the class label for new data as an output. Unsupervised learning differs from supervised learning in that during the training stage, no labeled data is present, and the system learns from the data itself by recognizing relationships or similarities among the occurrences in the dataset. To create a model, a semi-supervised method uses a small amount of labeled data as well as a big amount of unlabeled data.

I. Supervised Learning Methods

. Using Machine Learning models, Alsaleh et al. (2014) categorized Twitter accounts as human, bot, or Sybil. They compared current Sybil account identification algorithms and investigated various types of Twitter Sybil account detection attributes in order to develop an effective and practical classifier. On Twitter, GalánGarca et al. (2014) discovered spammers accounts used for cyberbullying. By examining the content of the comments created by both profiles, the author provides a methodology for detecting and associating false profiles on the Twitter social network that are used for defamatory actions to a real profile inside the same network. They give a successful real-world use case in which this methodology was used to detect and stop a cyberbullying scenario at an elementary school.

Yang et al. (2014) used ground-truth data regarding Sybil behavior in the field to design a measurement-based, real-time Sybil detector on Renren. The authors show that Sybils on Renren do not follow earlier work on decentralized Sybil detectors since they do not obey behavioral assumptions. Eighty percent of Sybils do not interact socially with other Sybils, preferring to form friendships with regular users.

II. Unsupervised Learning Methods

On the Renren network, Jiang et al. (2012) proposed the Sybil group detector, the first attempt to detect and evaluate sybil groups. The authors created a sybil group detector using numerous attributes, apply it to Renren, and discovered 2,653 sybil groups and 989,764 sybil users. By

creating automatic validation technique for sybil groups, by assessing the action time similarity of users in a group,

For fake account detection, Gani et al. (2012) suggested a methodology that depends on a machine learning model, social interaction, and authorship analysis. The study solves the difficult challenge of detecting numerous identities and is based on the concept that each person is distinct and identifiable, whether in writing style or social behavior. Through three levels of representation space extraction, learning layer, and validation, the authors proposed a system for grouping related identities and allowing the finding of many identities belonging to one author. The authors looked at features that are common in social networks (activity and topology). The limitation has been set to 20 since it has been demonstrated that a user cannot manage an endless number of identities. This has to be verified by domain specialists, but it's already a useful reduction in author search space.

To detect fake profiles, Lee and Kim (2014) developed a model based on name-based attributes. They offer a new detection scheme in their paper to filter possibly fraudulent account groupings around the time of their establishment. To identify fake accounts generated using similar methods, the scheme uses the distinctions between algorithmically generated account names and human-made account names. For accounts created in a short period of time and apply a clustering method to group accounts with comparable name-based attributes and a classification algorithm to categorize malicious account clusters. The authors, analyzed 4.7 million Twitter accounts. The technique achieves reasonable accuracy while relying just on account names and their creation times. As a result, it can be used as a quick filter against fake account groups to conduct an in-depth analysis selectively.

Kiruthiga et al. (2014) developed an expanded clone spotter approach based on classification and clustering techniques. One of the most insidious attacks on Facebook is the cloning attack. The photographs and personal information about a person are generally stolen by attackers and used to construct fake profile pages, which began sending friend requests using the cloned profile. If a real user's account was suspended, they would issue a fresh friend request to their friends. Simultaneously, the cloned one sends the request to the person. It was difficult for users to distinguish the genuine one at the time. The clone attack is detected in the proposed system based on user activity time periods and click patterns to determine the resemblance between the cloned profile and the real one on Facebook. The performance of the similarity between users is improved by using Cosine similarity and the Jaccard index.

III. Semi-Supervised Methods

The usefulness of this learning strategy has been proved by (Li et al., 2019). The Transductive support vector machines (TSVM) algorithm was trained using both image and document object model attributes. To solve the TSVM's local convergence problem, the authors developed a quantum-inspired evolutionary method. This method appears to be quite promising, and it could be particularly beneficial in social networks when there is a scarcity of publicly labeled data to identify fake accounts.

6. Conclusion

The core challenge faced by these algorithms is how to mitigate fake profiles in a way that is both scalable and effective, this highly affect their optimal performance. In this case, fake profile identification requires mechanism that can operate on billions of users together with their daily actions to identify dozens of different fake profiles. Therefore, scalability and data storage issues should initially be addressed while operating on big data that is available in OSNs.

Resolving the increased cases of privacy and security breaches currently found in OSNs, is a significant area for future research. This will not only make the user to be aware of how their personal and sensitive data is being collected and used but will also assist in preventing online attacks and misuse of data. Moreover, organizations can utilize OSN data without much concerns of compromising its security since users will be in control of the data shared on the platforms.

References

Adamic, L. A., & Adar, E. (2003). Friends and neighbors on the web. *Social Networks*, *25*(3), 211-230. https://doi.org/10.1016/s0378-8733(03)00009-1

Adedoyin-Olowe, M., Gaber, M. M., & Stahl, F. (2014). A survey of data mining techniques for social media analysis. *Journal of Data Mining & Digital Humanities*, *2014*. https://doi.org/10.46298/jdmdh.5

Adewole, K. S., Anuar, N. B., Kamsin, A., Varathan, K. D., & Razak, S. A. (2017). Malicious accounts: Dark of the social networks. *Journal of Network and Computer Applications*, *79*, 41-67. https://doi.org/10.1016/j.jnca.2016.11.030

Aghasian, E., Garg, S., Gao, L., Yu, S., & Montgomery, J. (2017). Scoring users' privacy disclosure across multiple online social networks. *IEEE Access*, *5*, 13118-13130. https://doi.org/10.1109/access.2017.2720187

Aichner, T., & Jacob, F. (2015). Measuring the degree of corporate social media use. *International Journal of Market Research*, *57*(2), 257-276. https://doi.org/10.2501/ijmr-2015-018

Alsaleh, M., Alarifi, A., Al-Salman, A. M., Alfayez, M., & Almuhaysin, A. (2014). TSD: Detecting Sybil accounts in Twitter. *2014 13th International Conference on Machine Learning and Applications*. https://doi.org/10.1109/icmla.2014.81

BalaAnand, M., Karthikeyan, N., Karthik, S., Varatharajan, R., Manogaran, G., & Sivaparthipan, C. B. (2019). An enhanced graph-based semi-supervised learning algorithm to detect fake users on Twitter. *The Journal of Supercomputing*, *75*(9), 6085-6105. https://doi.org/10.1007/s11227-019-02948-w

Boyd, D. M., & Ellison, N. B. (2007). Social network sites: Definition, history, and scholarship. *Journal of Computer-Mediated Communication*, *13*(1), 210-230. https://doi.org/10.1111/j.1083-6101.2007.00393.x

Carlsen, H. B., Toubøl, J., & Brincker, B. (2020). On solidarity and volunteering during the COVID-19 crisis in Denmark: The impact of social networks and social media groups on the distribution of support. *European Societies*, *23*(sup1), S122-S140. https://doi.org/10.1080/14616696.2020.1818270

Chowdhury, S. R., Roy, A. R., Shaikh, M., & Daudjee, K. (2014). A taxonomy of decentralized online social networks. *Peer-to-Peer Networking and Applications, 8*(3), 367-383. https://doi.org/10.1007/s12083-014-0258-2

Chu, Z., Widjaja, I., & Wang, H. (2012). Detecting social spam campaigns on Twitter. *Applied Cryptography and Network Security*, 455-472. https://doi.org/10.1007/978-3-642-31284-7_27

Coletto, M., Garimella, K., Gionis, A., & Lucchese, C. (2017). Automatic controversy detection in social media: A content-independent motif-based approach. *Online Social Networks and Media, 3-4*, 22-31. https://doi.org/10.1016/j.osnem.2017.10.001

Devineni, P., Koutra, D., Faloutsos, M., & Faloutsos, C. (2015). If walls could talk. *Proceedings of the 2015 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining 2015*. https://doi.org/10.1145/2808797.2808880

Fire, M., Kagan, D., Elyashar, A., & Elovici, Y. (2014). Friend or foe? Fake profile identification in online social networks. *Social Network Analysis and Mining, 4*(1). https://doi.org/10.1007/s13278-014-0194-4

Freeman, D. M. (2017). Can you spot the fakes? *Proceedings of the 26th International Conference on World Wide Web*. https://doi.org/10.1145/3038912.3052706

Galán-GarcÍa, P., Puerta, J. G., Gómez, C. L., Santos, I., & Bringas, P. G. (2015). Supervised machine learning for the detection of troll profiles in Twitter social network: Application to a real case of cyberbullying. *Logic Journal of IGPL*, jzv048. https://doi.org/10.1093/jigpal/jzv048

Gani, K., Hacid, H., & Skraba, R. (2012). Towards multiple identity detection in social networks. *Proceedings of the 21st international conference companion on World Wide Web - WWW '12 Companion*. https://doi.org/10.1145/2187980.2188098

Greenwood, B. N., & Gopal, A. (2015). Research note—Tigerblood: Newspapers, blogs, and the founding of information technology firms. *Information Systems Research, 26*(4), 812-828. https://doi.org/10.1287/isre.2015.0603

Gupta, A., & Kaushal, R. (2017). Towards detecting fake user accounts in Facebook. *2017 ISEA Asia Security and Privacy (ISEASP)*. https://doi.org/10.1109/iseasp.2017.7976996

Hardjono, T., Shrier, D. L., & Pentland, A. (2019). *Trusted data: A new framework for identity and data sharing*. MIT Connection Science & Engineering.

Huang, S., Lv, T., Zhang, X., Yang, Y., Zheng, W., & Wen, C. (2014). Identifying node role in social network based on multiple indicators. *PLoS ONE, 9*(8), e103733. https://doi.org/10.1371/journal.pone.0103733

Jia, H., & Xu, H. (2016). Measuring individuals' concerns over collective privacy on social networking sites. *Cyberpsychology: Journal of Psychosocial Research on Cyberspace, 10*(1). https://doi.org/10.5817/cp2016-1-4

Jiang, J., Shan, Z., Sha, W., Wang, X., & Dai, Y. (2012). Detecting and validating Sybil groups in the wild. *2012 32nd International Conference on Distributed Computing Systems Workshops*. https://doi.org/10.1109/icdcsw.2012.9

Joshi, S., Nagariya, H. G., Dhanotiya, N., & Jain, S. (2020). Identifying fake profile in online social network: An overview and survey. *Communications in Computer and Information Science*, 17-28. https://doi.org/10.1007/978-981-15-6315-7_2

K. Ojo, A. (2019). Improved model for detecting fake profiles in online social network: A case study of Twitter. *Journal of Advances in Mathematics and Computer Science*, 1-17. https://doi.org/10.9734/jamcs/2019/v33i430187

Kaplan, A. M., & Haenlein, M. (2010). Users of the world, unite! The challenges and opportunities of social media. *Business Horizons, 53*(1), 59-68. https://doi.org/10.1016/j.bushor.2009.09.003

Khan, J., & Lee, S. (2018). Online social networks (OSN) evolution model based on Homophily and preferential attachment. *Symmetry*, *10*(11), 654. https://doi.org/10.3390/sym10110654

Kiruthiga S., Kola Sujatha P., & Kannan A. (2014). Detecting cloning attack in social networks using classification and clustering techniques. *2014 International Conference on Recent Trends in Information Technology*. https://doi.org/10.1109/icrtit.2014.6996166

Lee, S., & Kim, J. (2014). Early filtering of ephemeral malicious accounts on Twitter. *Computer Communications*, *54*, 48-57. https://doi.org/10.1016/j.comcom.2014.08.006

Li, Y., & Liang, D. (2019). Safe semi-supervised learning: A brief introduction. *Frontiers of Computer Science, 13*(4), 669-676. https://doi.org/10.1007/s11704-019-8452-2

Liu, D., Mei, B., Chen, J., Lu, Z., & Du, X. (2015). Community based spammer detection in social networks. *Web-Age Information Management,* 554-558. https://doi.org/10.1007/978-3-319-21042-1_61

Meligy, A., M. Ibrahim, H., & F. Torky, M. (2017). Identity verification mechanism for detecting fake profiles in online social networks. *International Journal of Computer Network and Information Security*, *9*(1), 31-39. https://doi.org/10.5815/ijcnis.2017.01.04

Moore, T., & Clayton, R. (2008). The consequence of non-cooperation in the fight against phishing. *2008 eCrime Researchers Summit*. https://doi.org/10.1109/ecrime.2008.4696968

Nair, R., & Anbalagan, B. (2021). A graph based cloned profile detection in online social networks. *International Journal of Scientific and Research Publications (IJSRP), 11*(6), 652-657. https://doi.org/10.29322/ijsrp.11.06.2021.p11484

Pallis, G., Zeinalipour-Yazti, D., & Dikaiakos, M. D. (2011). Online social networks: Status and trends. *Studies in Computational Intelligence*, 213-234. https://doi.org/10.1007/978-3-642-17551-0_8

Penni, J. (2017). The future of online social networks (OSN): A measurement analysis using social media tools and application. *Telematics and Informatics*, *34*(5), 498-517. https://doi.org/10.1016/j.tele.2016.10.009

Quattrone, A., Kulik, L., Tanin, E., Ramamohanarao, K., & Gu, T. (2015). undefined. *2015 10th International Conference on Information, Communications and Signal Processing (ICICS)*. https://doi.org/10.1109/icics.2015.7459926

Reddy, H., Dey, A., Dey, M., & Sinha, N. (2019). Detection of fake accounts in Instagram using machine learning. *International Journal of Computer Science and Information Technology*, *11*(5), 83-90. https://doi.org/10.5121/ijcsit.2019.11507

Romanov, A., Semenov, A., & Veijalainen, J. (2017). Revealing fake profiles in social networks by longitudinal data analysis. *Proceedings of the 13th International Conference on Web Information Systems and Technologies*. https://doi.org/10.5220/0006243900510058

Romanov, A., Semenov, A., Mazhelis, O., & Veijalainen, J. (2017). Detection of fake profiles in social media - Literature review. *Proceedings of the 13th International Conference on Web Information Systems and Technologies*. https://doi.org/10.5220/0006362103630369

Sahoo, S. R., & Gupta, B. (2019). Hybrid approach for detection of malicious profiles in Twitter. *Computers & Electrical Engineering*, *76*, 65-81. https://doi.org/10.1016/j.compeleceng.2019.03.003

Schneider, F., Feldmann, A., Krishnamurthy, B., & Willinger, W. (2009). Understanding online social network usage from a network perspective. *Proceedings of the 9th ACM SIGCOMM conference on Internet measurement conference - IMC '09*. https://doi.org/10.1145/1644893.1644899

Viswanath, B., Post, A., Gummadi, K. P., & Mislove, A. (2010). An analysis of social network-based Sybil defenses. *Proceedings of the ACM SIGCOMM 2010 conference on SIGCOMM - SIGCOMM '10*. https://doi.org/10.1145/1851182.1851226

Wang, A. H. (2010). DON'T FOLLOW ME - Spam detection in Twitter. *Proceedings of the International Conference on Security and Cryptography*. https://doi.org/10.5220/0002996201420151

Wani, M. A., & Jabin, S. (2018). Mutual clustering coefficient-based suspicious-link detection approach for online social networks. *Journal of King Saud University - Computer and Information Sciences*. https://doi.org/10.1016/j.jksuci.2018.10.014

Yang, Z., Wilson, C., Wang, X., Gao, T., Zhao, B. Y., & Dai, Y. (2011). Uncovering social network sybils in the wild. *Proceedings of the 2011 ACM SIGCOMM conference on Internet measurement conference - IMC '11*. https://doi.org/10.1145/2068816.2068841

Yang, Z., Xue, J., Yang, X., Wang, X., & Dai, Y. (2016). VoteTrust: Leveraging friend invitation graph to defend against social network sybils. *IEEE Transactions on Dependable and Secure Computing*, *13*(4), 488-501. https://doi.org/10.1109/tdsc.2015.2410792

Yu, H., Kaminsky, M., Gibbons, P. B., & Flaxman, A. (2006). SybilGuard. *ACM SIGCOMM Computer Communication Review*, *36*(4), 267-278. https://doi.org/10.1145/1151659.1159945