IEEE Access
Multidisciplinary ⋮ Rapid Review ⋮ Open Access Journal

Received December 25, 2016, accepted January 3, 2017, date of publication January 20, 2017, date of current version March 13, 2017.

Digital Object Identifier 10.1109/ACCESS.2017.2656635

Sybil Defense Techniques in Online Social Networks: A Survey

MUHAMMAD AL-QURISHI[1], (Student Member, IEEE),
MABROOK AL-RAKHAMI[1], (Student Member, IEEE), ATIF ALAMRI[1], (Member, IEEE),
MAJED ALRUBAIAN[1], (Student Member, IEEE), SK MD MIZANUR RAHMAN[1,3], (Member, IEEE),
AND M. SHAMIM HOSSAIN[2], (Senior Member, IEEE)
[1]Research Chair of Pervasive and Mobile Computing, King Saud University, Riyadh 11543, Saudi Arabia
[2]Department of Software Engineering, College of Computer and Information Sciences, King Saud University, Riyadh 11543, Saudi Arabia
[3]Department of Information Systems, College of Computer and Information Sciences, King Saud University, Riyadh 11543, Saudi Arabia

Corresponding author: M. S. Hossain (mshossain@ksu.edu.sa)

This work was supported by King Saud University, Deanship of Scientific Research, Research Chair of Pervasive and Mobile Computing.

ABSTRACT The problem of malicious activities in online social networks, such as Sybil attacks and malevolent use of fake identities, can severely affect the social activities in which users engage while online. For example, this problem can affect content publishing, creation of friendships, messaging, profile browsing, and commenting. Moreover, fake identities are often created to disseminate spam, use the private information of other users, commit fraud, and so on. A malicious person can generate numerous fake accounts for these purposes to reach a large number of trustworthy users. Thus, these types of malicious accounts must be detected and deactivated as quickly as possible. However, this objective is challenging, because a fake account can exhibit trustworthy behaviors and have a type of name that will prevent it from being detected by the security system. In this paper, we provide a comprehensive survey of literature from 2006 to 2016 on Sybil attacks in online social networks and use of social networks as a tool to analyze and prevent these attack types. We first review existing Sybil attack definitions, including those in the context of online social networks. We then discuss a new taxonomy of Sybil attack defense schemes and methodologies. Finally, we compare the literature and identify areas for further research in Sybil attacks in online social networks.

INDEX TERMS Online social networks, Sybil attack and defense, Twitter, Sybil impact.

I. INTRODUCTION

The growth of online social networks (OSNs), such as Facebook, Twitter, and YouTube, has revolutionized the ways in which people interact, think and conduct business [1], [112]. Recently, users have shown tremendous activity on social media sites, resulting in an unprecedented amount of user-generated content at a continual pace [2], [3]. For example, the average number of monthly Google searches is more than 100 billion, and more than 50 billion web pages are indexed on Google. Meanwhile, YouTube has more than 1 billion users, which is almost one-third of all users on the Internet. On average, these users upload 100 hours of video every minute [4]. One of the largest OSN is Facebook. It is used by normal users, as well as celebrities, politicians, and other public interest figures, to share content with other users. The site has 1.3 billion users who spend 640 million minutes each month on its 54 million pages. Similarly, Twitter has 1.3 billion registered users, and its 320 million active users spend 170 million minutes on the site each month [5]–[7].

The immense number of users on OSNs has made them the de facto source of reliable identity services for the web. Using application programming interface (API) services, such as those for Twitter, YouTube, and Facebook, users can register for and log into third-party applications using their respective OSN account [8]. The simplicity of this process is attractive for users, while third-party apps benefit by gaining access to such vast OSN repositories of personal information. In 2010, 250 million Facebook users employed Facebook Connect on more than two million sites [9]. Facebook is now leveraging its "ecosystem" to expand into banking and payment processing with the intention of directly connecting its user base to online retailers.

As the popularity and influence of OSNs have increased, the incentives to attack these systems have also grown. A malicious user can create multiple false identities to gain access to sensitive private information to perform various types of several cybercrimes, such as compromising data integrity, "trolling" (making deliberately provocative or offensive online postings or criticism of opinions), rigging popularity, scamming, and breaking trust in online associations. These OSN attacks are types of Sybil attacks. The Sybil attack is named after the subject of the book Sybil, which describes a person diagnosed with dissociative identity disorder [10].

OSNs have become an integral part of contemporary life, with many people relying on them in the realm of work, social interactions, information sharing, and other aspects of daily living. Any negative impact on these areas due to Sybil attacks not only damages the user experience, but also the marketability and advertising potential of the given OSN. Outlined below are the key aspects that consider the above circumstances and motivate this work to survey defense schemes against Sybil attacks on OSNs.

- OSN openness
  OSNs are created as open platforms. To malicious users, the most attractive aspect of OSNs is their ready connection to many users at a minimal cost when compared to other internet channels. For example, in May 2009, many legitimate Twitter accounts were hijacked to spread advertisements [11], [12]. Moreover, in February 2010, the accounts of thousands of Twitter users, including the Press Complaints Commission and BBC correspondent Nick Higham, were hijacked after a viral phishing attack [13].
- Social connection blind trust
  Another major OSN loophole is that online users tend to trust their social connections and blindly value them. Thus, they may fall into a trap set by fake social connections perpetrated by cyber attackers. For instance, users are more likely to click a spamming link shared by a careless friend than the one they find on a random web page [14]–[17]. Spammers employ many techniques to send unwanted messages to users of OSNs, such as Facebook and Twitter. Such messages or tweets typically present either advertisements or fraudulent information, thereby enabling the perpetration of phishing attacks or the spread of malware through embedded URLs. For example, in August 2009, it was reported that nearly 11% of all Twitter posts was spam [3], [18].
- OSN recommender system limitations
  Recommender systems in OSNs are divided into four types according to the given OSN community: interest-, friend-, location-, and random-based communities [19]. The attackers strive to mimic the target profiles by building fake accounts that are consistent with one or more those of communities. Hence, the OSN recommender system directly chooses these fake accounts to recommend them to its users. Most OSN users prefer to have as many "friends" as possible to appear popular. Once the user adds the friends to her/his friendship list, the attacker can begin generating fake or otherwise malicious content [20]–[23].
- Reputation system vulnerability
  Similar to recommendation systems, reputation systems are likewise vulnerable to malicious users [23]. A reputation system calculates the numerical reputation of individual identities based on pairwise feedback between the identities [24]. A significant need exists to prevent fake identities from artificially boosting the attacker's reputation.
- Fake ratings and reviews
  Online reviews are helpful to online users and off- and online buyers. Thus, online reviews are very popular in e-commerce systems, such as Amazon, Yelp, and Google Play. Recently, however, online companies have had to curb the rising number of fake ratings and reviews that promote personal interests or bolster undeserving products, restaurant, apps, etc. In simple terms, these fake ratings and reviews generally subvert the reputation of a certain product or service.

The above aspects have made social networks attractive to Sybil attackers who manipulate large sets of malicious accounts to launch attacks [25]–[27]. Therefore, it is very important to detect Sybil attackers in an OSN to prevent their malicious activities. In particular, it is necessary to evaluate how liable an OSN user is in terms of his/her activities. Questions can be posed in this regard, such as "How does one identify users who are who they say they are, and whose activities can be considered genuine in a social network?" and "Does a specific method or strategy exist for detecting Sybil attackers in an OSN?"

In short, as the channels of global communication have rapidly grown, including OSNs, Sybil attacks have become increasingly prevalent. This survey covers recent literature on Sybil attacks in OSNs and techniques for protection against them. For many Sybil attackers, an attractive feature of OSNs is the large amount of personal information shared through them. This presents an opportunity for attackers to influence the integrity of data and communication in these networks. In other words, Sybil attacks present a major challenge to cyber security and improved defense mechanisms are required.

The remainder of this paper is organized as follows. Section II describes the background of this literature review on Sybil attacks in OSNs. Section III presents a Sybil attack defense solutions and schemes. Section IV discusses studies performed on detecting Sybil attacks in OSNs. Performance measurements, Benchmark datasets and matrices used in relevant studies are then addresses. Section V concludes the survey by summarizing the paper and outlining directions for future research.

II. BACKGROUND

In this section, we provide a general background and overview of OSN and some of its properties that help defending against Sybil attacks.

A. OSN OVERVIEW

An OSN can be described as a website that provides a venue for users to connect with other users, friends, and family members [28]. An OSN user is represented by an account profile. These three terms—user, account, and profile—can be used interchangeably. The profile describes the user's social attributes, such as name, gender, interests, and contact information. The relationship between accounts can be either two-sided, such as friendships in Facebook, or one-sided, such as a followership in Twitter [17]. Apart from creating reliable connections, an OSN can enable a user to share photos, music, videos, and other individualized information with certain friends or the general public [29], [30]. In addition to Facebook, Twitter, YouTube, some highly popular OSNs include LinkedIn, Google+, Instagram, and Myspace. These networks are a very suitable way for enabling users to remain in contact with other users around the world. Moreover, they are convenient because they enable users to create new connections with people who have similar affinities, such as a similar profession or similar interests [31]–[33].

In addition, an OSN is considered a social structure that can be represented as a social graph, as shown in Fig. 1, specifically G = (V, E), where V represents a set of users in the network as nodes. Relationship E between them is described as a set of edges [32]. The degree of a node indicates its centrality. Accordingly, the centrality of a node is considered to be a local measure that does not consider the global properties of the network [34], [35]. This type of social graph also features interdependencies or "ties." These ties can be manifold and assorted, and may include concepts, such as genealogy, age, ideas, gender, education, club membership, and political affiliations, among other characteristics that may influence relationships in the OSN, as described in [29], [32], [33], [35].

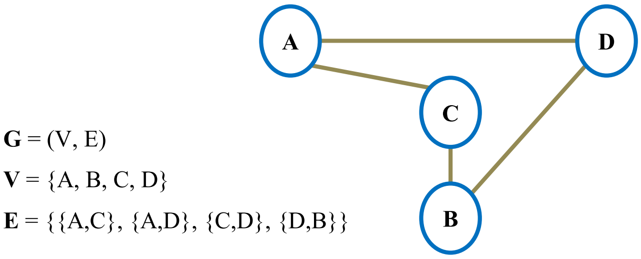


FIGURE 1. Social network structure.

OSN users participate in various social activities. A high likelihood exists that some malicious activities will occur in the midst of those activities. Social graphs play a critical role in developing a robust model of trustworthiness that is based on the different social activities [36]. Heterogeneous social graphs can be developed and used to capture the various social activities, such as posting new tweets, retweeting, and directly sending messages on Twitter; making new posts and sharing information on Facebook; and many other activities on other platforms. By using social graphs, it is possible to develop different measures of trustworthiness and adopt simple methods for detecting any form of malicious activity.

In OSNs, users, along with their social relationships and activities, are categorized as major entities. Thus, malicious activities can be differentiated from legitimate ones by measuring the extent to which each social activity is trustful and classifying the activity in the relevant category. When using social graphs, each entity is intuitively calculated to have a trustworthiness score. If the trustworthiness of an entity is deemed low, it is highly likely that the social activities with which it is associated will be unreliable. This is the reason why every suitable trust model that is used with social graphs must feature a propagation process to validate the information obtained [37]–[39].

Social graph properties encompass three major aspects: diameter and mean path length, centrality and nodal power, and degree distributions [35], [40].

- Diameter and mean path length:

The diameter of a social graph is defined as the greatest distance between any two nodes. It can be described as the maximum value of the shortest path length between a pair of nodes in the social graph. For instance, if $l(i, j)$ is considered to be the length between nodes i and j, then diameter d is considered the maximum $l(i, j)$ of all possible node pairs.

$$d = max_{x,y} l(i, j). \tag{1}$$

On the other hand, the mean path length, l_m, of a social graph is considered the mean distance between all the nodes that are located in the graph [35].

$$l_m = \frac{1}{n(n-1)} \sum l(i, j) \tag{2}$$

where l_m is bounded by the diameter; at times, it can be much shorter than d.

- Power and centrality:

Power in this context is defined as a basic property of the social structures that have close relations with centrality [35], [41], [42]. Various techniques have been used to facilitate research on the graph property of power. However, the three main measures used to describe power or centrality is the degree, closeness, and betweenness. In this regard, the degree refers to the number of edges of a certain node. This property is often normalized by the sum of the edges that are available in a graph.

$$deg_i(g) = \sum\nolimits_{\boldsymbol{i=1}}^{\boldsymbol{n}} \boldsymbol{e(v_i)}. \tag{3}$$

Closeness is described as the typical number of "hops" from a certain node to the other nodes that are located in the social graph. Closeness can be calculated as

$$C_i\,(v) = \left[\sum_{j=1}^{n} d(i, j) \right]^{-1}$$

Lastly, betweenness refers to the possibility that any node will require a different node in order to reach an additional node using the shortest path [35], [40]. Betweenness is computed by using Eq. 4.

$$B_i\,(v) = \sum\nolimits_{i \neq \alpha \neq j \in V} \frac{\sigma_p(v)}{\sigma_p} \tag{4}$$

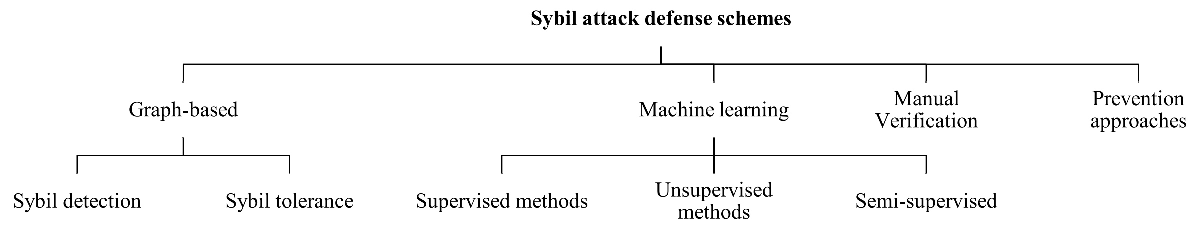


FIGURE 2. Proposed classification of Sybil schemes.

where σ_p is the total number of the shortest paths from i to j, and $\sigma_p(v)$ is the number of paths that cross the node.

- Degree distribution:

The degree distribution implies the distribution of degrees for all the nodes that are available in a graph. This property very closely follows a power law for all real social networks. A large percentage of nodes feature a minimal degree. However, a small number of nodes have a very large degree [43]. Power law is represented in Eq. 5.

$$P(k) \sim k^{-\gamma} \tag{5}$$

where γ is a constant.

In sum, social networks are important sources of information, and they can be modeled using graph theory. However, they can be difficult to analyze. The analysis may require considerable knowledge of a means to calculate the metrics and decipher appropriate meaning from the statistics obtained.

B. DEFINITION OF A SYBIL ATTACK

The name Sybil is from a nonfiction book about a woman named Sybil with dissociative identity disorder. John Douceur was the first to use this term in computer security [10], [44], [45], specifically with regard to an attacker who subverts a system by creating a large number of pseudonymous identities (i.e., user accounts) and employing them to significantly impact the system. In an ideal world, every user should have only one real identity; however, this is often not a reality in practical situations. Twitter, Facebook, and almost all other OSNs offer a lightweight process for creating an account with only email confirmation. Thus, it is very easy for a user to have more than one account.

In OSNs, a Sybil is a fake account with which a user attempts to create multiple identities to make as many friends as possible with legitimate accounts. A Sybil account can lead to many malicious activities in an OSN. For example, a Sybil can outvote legitimate users through Internet options [24]. In addition, it can control some accounts by using the fake identities to provide misleading information [46], [47]. Moreover, false reputation can be created based on Sybil accounts [48].

Generally, a Sybil attacker is a user that creates many fake identities to increase the power and influence of the attacker in the network and thereby generate engage in malevolent activities, such as disseminating social spam, distributing malware, distorting online ratings, executing phishing attacks, and so on [28], [49], [50].

III. SYBIL ATTACK DEFENSE: SOLUTIONS AND SCHEMES

Recently, interest and efforts for defending against malicious attacks in social networks using reliable solutions and schemes have increased [28], [44], [49]–[52]. Several proposed schemes strive to minimize Sybil attacks in an OSN by utilizing the properties of the OSN's social structure. Unlike the traditional solutions discussed in the previous section, these existing solutions and schemes do not require central trusted identities. Instead, they solely rely on the trust personified in the existing social relationships that occur between the users of the OSN. Most literature on existing Sybil defense mechanisms show that these mechanisms are in early stages of development. Therefore, most of these researchers describe new algorithms in their papers; however, they do not present a method by which all the proposed schemes can detect the occurrence of Sybil attacks [51].

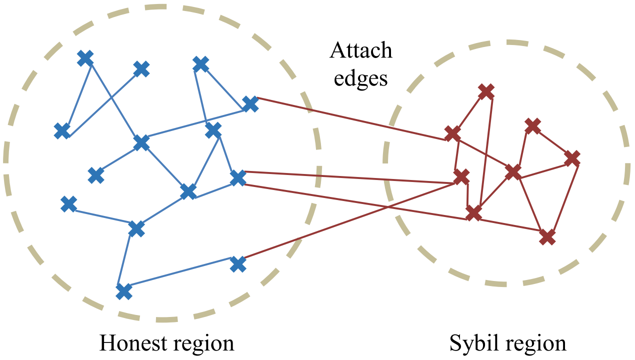


FIGURE 3. Legitimate and Sybil regions.

In this section, we provide an overview of different Sybil schemes (as shown in Fig. 2) that have been used by researchers to execute various Sybil detection and prevention algorithms and tools. In general, the Sybil schemes are divided into four main categories: Graph-based schemes, machine-learning-based schemes, manual verification and Prevention approaches. Graph-based (also called network-based [51]) schemes are further divided into subcategories: Sybil detection and Sybil tolerance. Meanwhile, the machine learning schemes are divided into respective supervised, unsupervised, and semi-supervised methods.

The classification approach is based on our perspectives and understanding of the given problem.

A. GRAPH-BASED METHODS

Graph-based methods use social network information in the case of Sybil to represent interdependencies between objects using edges or links. These methods strongly rely on social graph properties to distinguish Sybil users from legitimate users. They can be classified into Sybil detection and Sybil tolerance schemes, respectively. Most of the graph-based solutions rely on the assumptions outlined below.

In the social graph, legitimate and Sybil regions exist, and the connections between them are loose, as shown in Fig. 3. There are fewer edges between legitimate and Sybil accounts than between various Sybil accounts.

The legitimate region in the social graph is fast mixing, which means that the number of random walk movements from the initial state to the stationary state is small, as depicted in Fig. 4.

The Sybil attacker may infiltrate the social graph but cannot produce meaningful interactions; thus, the number of escape tails to the attacker is always zero.

1) SYBIL DETECTION SCHEMES

Most of the Sybil detection schemes are based on the concepts of the graph random walk and mixing time. The random walk is given by the following probability matrix:

$$P(i,j) = \begin{cases} \dfrac{1}{\deg(i)} & \text{if } ij \in E \\ 0 & \text{otherwise} \end{cases} \qquad (6)$$

The mixing time, $T(\epsilon)$, of a random walk is given by:

$$T(\epsilon) = max_{u\in V} min_t \{t : \Delta_u(t) < \epsilon\} \quad \text{for any } \epsilon < 0. \qquad (7)$$

Fast mixing is a crucial assumption on which most Sybil defense approaches depend [53]. For instance, given that n is the number of vertices, the mixing time is given by:

$$T(\epsilon) = \min\left(\log(n), \log(\frac{1}{\epsilon})\right) \quad \text{for } \epsilon = \theta(\frac{1}{n}). \qquad (8)$$

Eventually, the mixing time is $T(\epsilon) = O(\log(n))$.

Some of the most common approaches are social-graph-based techniques, such as the first ones proposed for this purpose—SybilGuard [26] and SybilLimit [55]—among other Sybil-defense approaches. Other approaches have included Sybil-resilient systems that use application knowledge to control Sybils. Below all of these schemes and methods are detailed.

a: SybilGuard [26]

This scheme identifies Sybil nodes by employing random routes. With respect to social networks, random routes are described as a special type of random walk whereby each node employs random variations that have already been computed [35], [56]. Nodes use these variations based on a one-to-one mapping technique from the incoming to the outgoing edges. One major shortcoming of SybilGuard is that the scheme suffers from a very high number of false-negatives [54], [57]. This occurs because every attack edge can easily introduce Sybil nodes without being easily noticed.

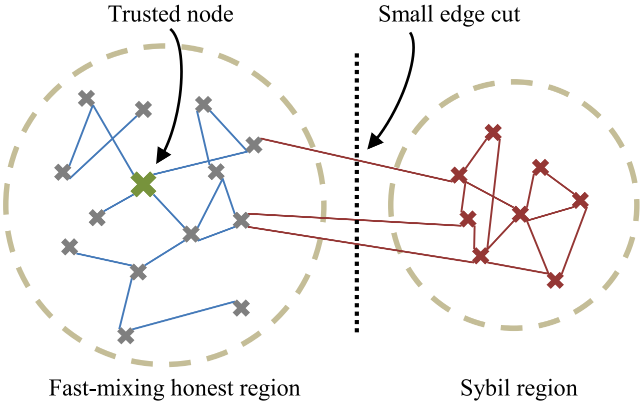


FIGURE 4. General assumptions.

The Sybil nodes can be denoted as $O(\sqrt{n}\mathrm{log}n)$. Although the improved version of the SybilGuard scheme can reduce the $O(\sqrt{n}\log n)$ value to $O(\mathrm{log}n)$, this value is still large compared to the proven lower bound by a factor of log n. Furthermore, it remains necessary to test all the suspect nodes in the social graph in order to detect the Sybil region in an OSN. The ranking ability of SybilGuard is determined based on varying random walk lengths. The cut-off can still be determined regardless of whether a walk intersection occurs.

b: SybilLimit [25]

SybilLimit executes tasks by performing independent random walks characterized as $O\left(\sqrt{m}\right)$. The random walks have lengths denoted as $O(\log n)$ based on each node. For SybilLimit to mark a suspect as non-Sybil, two conditions must be fulfilled. The first condition is referred to as the intersection condition; the second is the balance condition [51], [58]. The intersection condition necessitates the crossing of the last edge of one of the random walks presented on the trusted node together with the suspect. On the other hand, the balance condition reduces the number of non-Sybils that are experienced in each attack edge. When using SybilLimit, each tail of a random walk is given a ''load'' whose threshold is limited to a specific value. Accordingly, the load is incremented each time the trusted node considers another suspect to be non-Sybil. This scheme utilizes two basic parameters to control the quantity of nodes that are marked as non-Sybil in an OSN. These two parameters include the number of random walks from every node as well as the lengths of these walks. Whenever these two parameters increase, an increasing number of nodes are marked as non-Sybil.

In addition, similar to SybilGuard, SybilLimit relies on an estimation procedure when determining the length of the random walk as well as the quantity of random walks that are required. In general, these two parameters are crucial for determining the occurrence of a cutoff.

c: SybilInfer [62]

This scheme features a centralized Sybil defense algorithm that leverages a Bayesian inference approach. When using this approach, the scheme assigns a Sybil probability that indicates the degree of certainty to each node that occurs within the OSN. This scheme can achieve a low false-negative rate at the cost of high computation overhead. The general time complexity of this scheme can be summarized using $O(|V|2\log|V|)$, where V represents the set of vertices that occur within the social graph. An evaluation of SybilInfer shows that the scheme can handle a maximum of 30,000 nodes. This number of nodes is considered much smaller than that of a regular-size OSN [51]. In determining the cutoff, SybilInfer divides the nodes based on a threshold value with consideration of the possibility that a node will be considered non-Sybil.

d: GateKeeper [59]

GateKeeper is another decentralized Sybil defense solution that strongly relies on the assumption that OSNs are considered random expanders. This scheme is considered an improvement over the SybilLimit scheme by a factor of $O(\log n)$ on the random expander graphs. Generally, the enhancement was shown to be useful for the GateKeeper scheme, especially whenever the attacker dominates only the $O(1)$ attack edges. This scheme can admit an immense number of legitimate nodes while limiting the number of Sybils that are simultaneously admitted in each attack [60]. The ability of GateKeeper to achieve this objective is represented using $O(\log k)$, where k represents the quantity of attack edges. However, the assumption made by GateKeeper is too strong and has yet to be effectively validated. Generally, GateKeeper also suffers from a high false-positive rate, as well as negative rates; moreover, its ability to identify Sybil nodes on real-world asymmetric topologies is insufficient.

e: SybilShield [61]

SybilShield is another OSN-based protocol that assumes that a network consists of a number of small, medium, and large communities. SybilLimit and SybilGuard, on the other hand, assume that a network consists of a respective legitimate and Sybil region. SybilShield's other assumption is that two given networks are fast mixing, and a malicious attacker can create numerous identities, but few trusted relations exist between a legitimate node and a Sybil node. The protocol defines an edge between different communities as a foreign edge, meaning that edges formed between legitimate and illegitimate communities are fewer than the number of edges formed between legitimate communities.

SybilShield is intended to address the limitations of SybilGuard, whereby a legitimate suspect node can be treated as a Sybil node. It achieves this objective through modified random walks, known as the random route approach, whereby a suspect node is deemed legitimate if the random routes of a suspect node and verifier intersect. The main concept is to provide an opportunity to suspect nodes through an agent walk approach to mitigate the problem. Additionally, agents are selected from communities other than the verifier's community.

f: SybilDefender [63]

SybilDefender is a relatively different scheme that leverages network topologies to defend against Sybil attacks. This scheme can correctly identify Sybil nodes, even in situations in which the number of Sybil nodes introduced by each attack edge approaches the detectable lower bound theorized by most related studies. Furthermore, this scheme can easily and effectively detect the Sybil community surrounding a node based on different aspects of structure and size. It leverages random walks to large OSNs. It additionally outperforms the existing state-of-the-art approaches based on a single or double order of magnitude in running time and accuracy.

The general design of a SybilDefender scheme includes a Sybil identification algorithm that identifies the Sybil nodes, a detection algorithm for the Sybil community, and two approaches that limit the quantity of attack edges experienced in the OSN. A major drawback of the SybilDefender scheme is that it only relies on performing a limited number of random walks within the social graph.

g: SybilRank [66]

SybilRank is a new tool for OSN operators. It relies on social graph properties to rank users based on their perceived likelihood of being Sybils. This scheme is considered efficient in a computational manner and can scale graphs using several nodes. Even though SybilRank achieves an equivalent or even higher accuracy when compared to the other schemes, it is has a cost of $O(n \log n)$ compared to the number of seeds. SybilRank additionally addresses the most important limitations of the current defenses based on social graphs.

First, it influences its support for various trust seeds to reduce the number of false-positives that result from the existence of multiple communities considered non-Sybil. Secondly, the scheme enables a very flexible seed selection process that makes it much more difficult for attackers to target these seeds. Furthermore, its effectiveness is only moderately decreased when the distance of the Sybil decreases in relation to that of the trust seeds. This scheme is thus suitable for large-scale attacks in which fake accounts can be developed and maintained at an extremely low cost. In addition, SybilRank can be deployed on a social graph that features strong relationship edges.

h: SybilFence [67]

The author of [67] strived to improve the use of social-graph-based Sybil defenses through integration of user feedback in OSNs to limit the social links of users who receive negative feedback. The author provided an analysis of the concept of negative feedback in OSN by conducting a study on Facebook accounts in the black market considering that they can serve as excellent examples of perfectly maintained fake accounts.

In [68], the authors show that these accounts actually appear real even though they are fabricated by the vendors who sell them. Despite their appearance, they actually receive considerable negative feedback in the form of rejections. In Facebook, a pending request is usually perceived as a rejection. The investigated accounts had many pending requests but, interestingly, a more than 50% acceptance rate. This finding was attributed to the use of triadic closure principles, whereby users send requests to friends of users they have already befriended. The designed system incorporated user negative feedback into social-graph-based defenses. The objective was to model an OSN using two graphs, i.e., modeling the underlying social graph as an undirected graph, and modeling the user feedback graph as a directed graph.

Considering that Sybil attackers may create multiple fake accounts to execute attacks, the user set is divided into two subsets consisting of non-Sybil users and Sybils. The aim is to build a weighted defense graph in which users who have received negative feedback are subjected to reduced weights on social edges to mitigate the impact of the entrance of the Sybil attack edges. In the initial design, SybilRank is a proof of concept.

The SybilFence evaluation process showed better insights on detection accuracy of the initial design compared to SybilRank. Sybil attacks were simulated in four different social graphs—Facebook, Synthetic, ca-astroph, and ca-Hepth. The Synthetic graph was generated using the scale-free model, while the Facebook graph was sampled using the forest fire sampling technique. The simulation results showed that SybilFence actually improved SybilRank accuracy in terms of ranking normal users higher than Sybil users. It also showed greater immunity to simulated attacks in which Sybil users solicited additional social connections. However, this scheme should be applied on a real OSN graph, rather than a synthetic graph.

i: Integro [69]

Integro is an extension of SybilRank and likewise studies the posted content. The authors of [69] proposed Integro as a better solution for detection of Sybils. They stated that this method achieved 95% precision in Sybil detection, whereas SybilRank achieved 43%. The method was tested on Facebook, RenRen, and Tuenty, and it proved to be effective. Integro uses various node features to detect Sybils and identify potential victims in a non-adversarial setting. The developers of the method employed only known legitimate profiles, which accepted or rejected friendship requests sent by known Sybil profiles. The authors did not accept the assumption that Sybil nodes can have a limited quantity of friends and therefore attack edges.

The authors first proposed detection of victim accounts based on user-level activities. They acknowledged that many users are not cautious in OSNs and accept friend requests from users whom they do not know, especially if they have common friends. They model social networks as an undirected graph with the nodes representing user accounts and the edges representing a bilateral social relationship among nodes. The nodes have a degree equal to the sum of the weights on their edges. The set of nodes is divided into two parts: real and fake accounts. Then, a Sybil region is detected. Integro sets a feature vector for each node to predict the probability that the user will be a victim. It also counts a vulnerability score, which represents the level of the probability that a user will become a victim. Using the power iteration method, Integro computes trust values.

It is worth mentioning that Integro is limited to undirected social graphs, and it delays the consideration of new accounts. Thus, it has a major limitation in terms of the accuracy of Sybil detection among new users. It is furthermore limited in its method of correctly detecting Sybils in cases in which a large number of fake accounts can collect more trust than a low number of real ones.

2) SYBIL TOLERANCE

Sybil tolerance methods limit the effects of Sybils that are present in an OSN. These defense approaches utilize many techniques for providing a particular credit for each edge in the graph. They ultimately limit the impact of attack edges. The networks that apply this concept are called credit networks (Fig. 5), wherein a node trusts another node by giving a pairwise credit to its link up to a certain limit.

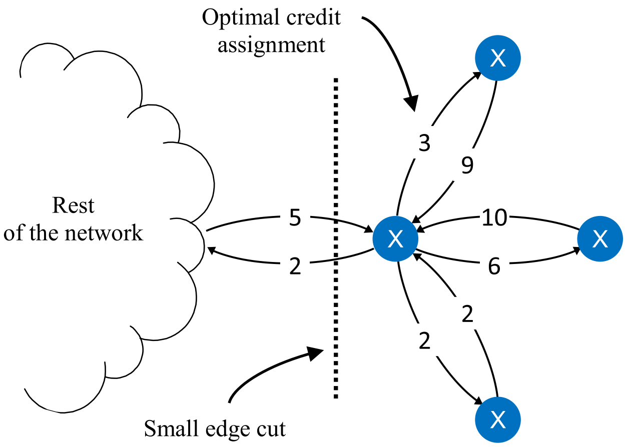


FIGURE 5. Credit network.

The three renowned approaches in this class—Ostra, SumUp, and TruTop—are detailed below.

a: Ostra [70]

Ostra is a monitoring system that reduces unwanted communication in online exchanges. The authors of [70] designed Ostra in three phases: authorization, transmission, and classification. Ostra uses a credit balance to decide whether to issue a token to enable the sender to transmit data to the receiver. Each user is assigned credit balance B within ranges L and U. B_U^L is the current credit of a sender and receiver, respectively. When the sender transmits a message to the receiver, sender L is increased by one and receiver U is decreased by one. If the message is unwanted, then the sender transmits one credit to the receiver.

Ostra requires that recipients classify incoming communication as either wanted or unwanted. Providing explicit feedback is a slight burden on the user. Ostra can be used only in conjunction with the ''invitation-only'' social network. The authors were unable to obtain a communication trace of the same scale as the OSN that they considered. Consequently, they made assumptions about the likely communication patterns in the OSN. They concluded that users communicate with nearby users much more often than with users who are far away in the OSN.

b: SumUp [60]

SumUp is a centralized scheme that addresses the problem of vote aggregation in a network. When using this scheme, three conditions must be met. First, all the votes from the legitimate nodes must be accepted. Second, if e_A is considered the quantity of attacking edges from an attacker, then the fake voting should be confined to e_A only. Third, if Sybil sends a fake vote in a continuous manner, then it should be rejected in the future. SumUp features an adaptive vote flow aggregation technique that limits the quantity of fake votes that are received from an attacker. The scheme also sums up the opinions of a legitimate source by carrying out a calculation of a set of max-flow links on the trust graph. The number of forged votes is often constrained by the quantity of attack edges among Sybil nodes. This is because only opinions with non-zero flows are often considered.

This scheme uses C_{max} in deciding the maximum quantity of votes that should be accepted by the system. SumUp can prohibit an attacker who progressively misbehaves. This scheme additionally assumes that a mini-cut exists between legitimate nodes and a cote collector occurs at the collectors. However, between the Sybils and legitimate nodes, the mini-cut occurs at the attack edges. Nevertheless, SumUp accepts $O(\log n)$ Sybils or attack edges, and it requires knowledge of the overall system. Major disadvantages of SumUp are its high respective computation and run-time requirements.

c: Canal [71]

a Sybil tolerance system that works in large networks. This system substitutes accuracy with speed as the authors in [71] demonstrate that its approximation rarely affects the users plus it does impact on the Sybil tolerance properties of the application. Authors claim that Canal can be directly fixed into existing Sybil tolerance schemes, which reduces credit payment latency from a few seconds to microseconds. Canal takes advantage of a novel landmark routing based algorithm which routes credit payments through landmark nodes. Canal consists of two components:

1) Universe Creator Processes-It Continuously selects new landmarks

2) Path Stitcher Processes-It continuously processes incoming credit payment requests. Canal uses these components to continually calculate new landmarks in parallel with doing flow calculations.

d: TrueTop [72]

TrueTop identifies the top influential users of a specific group of users on Twitter and other OSNs. A synthetic simulation was viable because real large-scale experiments were carried out on Twitter and the terms of service were consistently violated. Four datasets were used in the methodology and consistent results were obtained. The methodology models the strength of Sybil attacks on Twitter based on the α parameter. The α parameter refers to the ratio of the total weight of the edge in the non-Sybil region over the one from the Sybil region. The model assumes that there is often zero interaction between the Sybil and the non-Sybil regions. However, in real practice, users in the Sybil region always intend to initiate interactions with the other users in the non-Sybil region. This often occurs for obvious objectives, such as for spamming and phishing [11], [68], [73]–[75], to gain high scores of influence. Thus, as expected, a smaller number of credits remained in the Sybil region compared to when they were exposed to the TrueTop attacker model.

Several design choices for TrueTop represent the incomplete evaluations involved in the research methodology. Thus, the likelihood was shown that various possible strategies can be used by attackers to gain control of OSNs. In addition, the findings provided only a few relevant results and showed that the other design choices and attack strategies would yield similar results

B. MACHINE-LEARNING METHODS

In this section, we provide an overview of the different supervised approaches that have been used by researchers to detect Sybil attacks in OSNs. The advantages and drawbacks of each method are additionally highlighted. Machine learning is a technique for autonomously acquiring and integrating knowledge obtained from experience, analytical observations, etc. [75]. It is usually divided into two main types:

- Supervised learning, such as regression models, naive Bayes, support vector machine (SVM), and decision tree models.
- Unsupervised learning, such as clustering algorithms (K-means, fuzzy C-means, hidden Markov models, etc.).

Machine learning techniques are designed to solve problems involving massive amounts of data with many variables. These techniques are commonly used in areas such as pattern recognition (speech and image processing) and financial algorithms (credit scoring and algorithmic trading) [77].

1) SUPERVISED METHODS

Currently, some of the most commonly used supervised detection methods are the naive Bayes, SVM, nearest neighbor (e.g., k-NN), linear/logistic regression, and least squares models.

In sum, the success of supervised methods is highly dependent on the use of domain knowledge of data to construct features. This concept is also known as feature engineering [78]. This approach can be time consuming compared to feature selection, which only involves returning a subset of important features. It can be effectively illustrated by the logistic regression models in OpinSpam, which had to be fed 36 features in the purview of content and behavior information. Experts with extensive knowledge of Amazon and its corresponding review dataset were therefore required. Few studies have been conducted to detect Sybil attacks using a supervised technique. First, OSN providers do not give access to their databases; access is given only through public APIs. Thus, building the ground truth is the default. Second, training data on a large scale is difficult to achieve without full access. Despite those obstacles, some researcher have obtained access to RenRen [79], the largest Facebook-like OSN in China. They performed a supervised technique to distinguish between Sybil and non-Sybil accounts. That researched is outlined below.

a: UNCOVERING Sybil [54]

The main aim of this research was to make two key contributions to Sybil detection in OSNs. The first was to use available data on the characteristics of Sybil in the wild to improve a feature-based Sybil detector that was adopted in RenRen. The author first showed how a properly crafted detector could find 99% of Sybils with very low false-negative and positive rates. The second contribution was to characterize the topology of Sybil graphs in OSNs, marking the first topology characterization of its kind. The author showed that Sybils do not abide by the assumptions previously used in detectors behind community-based Sybil detection algorithms, e.g., SybilGuard, SybilLimit, SumUp, SybilInfer, etc. Thus, the author closely examined if the Sybil accounts had some form of relationship in the background. The results showed that Sybil accounts actually colluded to execute attacks.

The author then studied new forms of attack models that could be used to bypass the detector. It was shown that closely interwoven communities between Sybil users could succeed in this task. Using this information, the author demonstrated that the addition of a parameter, known as the external acceptance ratio, could improve the efficiency in Sybil detection as new attack models emerged.

In addition, the study compared the performances of two Sybil detection algorithms: a simple threshold detector and a complex learning algorithm detector, SVM. The performance of the tuned threshold detector was virtually equal to that of the rather expensive SVM. The study also showed that use of the detector could be inaccurate when Sybil attacks employ new model-forming tightly knit communities. A better detector was determined to be one that combines a feature-based detector and an enhanced community detection approach.

b: SOCIAL TURING TESTS [80]

The large study presented in [80] explored the feasibility of humans detecting Sybils in OSNs, while analyzing the detection accuracy of turkers. Under a certain set of conditions, the experts found that turkers delivered optimal results, while humans exhibited no uniform detection accuracies. The author analyzed this data and used it to drive the design of a multi-tier crowdsourcing system. The study of the system showed that it could be quite efficient as a standalone detection system or as a complimentary system to other detection techniques. The study of the human crowdsourcing capability was performed with the aid of a corpus of ground truth data Sybil accounts accrued from RenRen and Facebook. The two datasets had three sets of profiles: confirmed Sybils, confirmed legitimate users, and suspected profiles. Using crowdsourcing was not appropriate, however, because most members of the crowd lacked relevant expertise, especially in terms of being able to verify turkers. As a consequence, crowdsourcing was found to be inherently inconsistent in its assessments [81].

2) UNSUPERVISED METHODS

Some researchers have focused on detecting Sybil attacks in social media [111] using clustering and latent variables techniques [82]. Clustering is the process by which objects with similar characteristics are grouped in a set. Scavenger, for example, looks for URLs in Facebook wall posts; the URLS are then used to build a wall-post similarity graph. The graph is then used to select clusters that are most likely to result from spam campaigns. A significant feature of clustering is that it can uncover even hidden structures in unlabeled data while also having the potential to summarize key features [83]–[85].

Latent variable models are other clustering models that employ latent variables as a representation of hidden variables. Examples include spamicity of reviewers in Amazon. Spamicity, in this case, is the likelihood of something being spam. These concepts are well incorporated in an unsupervised Bayesian inference framework known as the author spamicity model (ASM). The framework creates a hypothesis that differentiates fake users and from real ones in a behavioral perspective. This task is performed by representing these two clusters in the latent space. Evaluation results showed that unsupervised spamicity models are actually very effective [86].

a: Clickstream [87]

The authors presented a method that leverages clickstream models on the basis of experimental approaches. They analyzed clickstream activity of click patterns of normal and Sybil users by employing logs from different social networks. Clickstream models were then proposed to characterize the click behaviors of users. These models were used to create a Sybil detection system independent of input from a service provider. The unsupervised Sybil detector created by the author of [87] exhibited excellent performance metrics, especially in terms of accuracy in real-world OSNs, namely RenRen and LinkedIn. Of one million random user accounts fed into the detector, it detected 22,000 Sybil users. In LinkedIn, out of the 40,000 analyzed users, the detector identified approximately 1,700 as Sybils out of 4,000 that had already been identified as Sybils.

b: SybilExposer [88]

A Sybil Community detection algorithm that depends on the properties of social graph communities to rank communities based on their apparent likelihood of being fake. Their motivation to conduct a study on SybilExposer is influenced by the various schemes that have been proposed to provide defense against Sybil attacks. Most of the schemes that have been proposed previously are either dependent on the right choice of known trusted nodes, or on a high running time cost. To investigate this problem, the authors carry out several experiments on real-world online social network graphs to identify the true positive rate, the false positive rate, and the running time complexity compared to the state of the art.

Authors proposes a different algorithm, SybilExposer, which deals with the weaknesses of manual inspection, high computation costs, and the need to consider various Sybil communities. This algorithm also enhances identification, and has over 20% capability to detect Sybil within a community compared to SybilRank. This algorithm can be used as a scalable first line of identification of various Sybil communities in very large networks. The simulation results indicates that the proposed algorithm performs much better in terms of computational costs and effectiveness, compared to other algorithms. However, future research directions should consider extending the scheme of the algorithm to deal with non-community Sybil nodes as well as edges that occur between Sybil communities.

3) SEMI-SUPERVISED METHODS

The semi-supervised learning is a technique that uses a set of labeled and unlabeled data. It is thus between supervised learning techniques that use labeled data and unsupervised learning techniques which use only non-labeled data. It has been shown that the use of unlabeled data, in combination with labeled data, significantly improves the quality of learning. Many researchers employ this method for their works.

a: SybilBelief [64]

SybilBelief is a semi-supervised learning framework that relies on loopy belief propagation and Markov random fields. The inputs of this scheme include the social network of the nodes in the system, a small number of known Sybils, and a small number of known benign nodes. In its operation, the scheme operates by propagating the label information from the benign node, which is known to the remaining nodes that are unknown in the system. In addition, SybilBelief can accurately identify the Sybil nodes with few cases of false-positive and false-negative rates. Moreover, it is resilient to noise and can perform orders of a magnitude better than other Sybil classification mechanisms. Experimental results obtained in [64] showed that SybilBelief performed an order of magnitude better compared to SybilInfer and SybilLimit. A major limitation with that study, however, is that the authors did not evaluate their approach using real Sybil users. In addition, it showed a lower accuracy when the number of attack edges increased, and it failed to work on weak trusted networks [65].

b: SybilFrame [65]

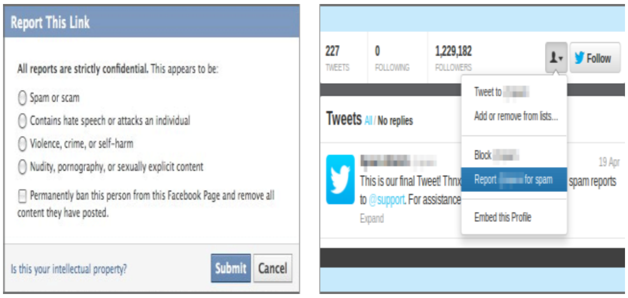
SybilFrame is used to solve the problem of Sybil attacks in OSNs with weak trust. It enables the defending of this kind of network, while operating in conditions where the number of attack edges is large. It uses a multi-stage classification mechanism that analyzes heterogeneous sources and types of information on the profiles in the OSN. The authors of this method used data on suspended accounts to serve as the basic truth for Sybil attacks. This method can additionally be used to rank accounts in terms of trustworthiness. SybilFrame gathers all the information on the users to define if they are Sybil or not. During this stage, the computing of the node prior and edge prior information takes place.

To conclude on edge priors, the authors assigned lower scores to attack edges and higher ones to the edges between the trustworthy nodes. Secondly, it correlates the information from local classifiers with the global properties of the network. In this stage, pairwise Markov random field and loopy belief propagation methods are used to make the conclusions on the posterior information. In the method framework, local classifiers are considered more effective when combined with the global information. These authors also studied the structural differences of Sybil and non-Sybil nodes. They determined that most Sybil accounts are isolated, which makes it difficult to detect them, and new effective methods should be created.

C. MANUAL VERIFICATION AND USER FEEDBACK

OSNs are targets for disseminating false and rumor-related content. This content proliferates every day and misleads thousands of OSN users. The content may include false news, doctored images, fake accounts, or any type of content that can be circulated among OSNs. Content can be analyzed in different ways to determine whether it is true or false, and it is the responsibility of the user to analyze content before commenting on or sharing it. Many social networks, such as Facebook and Twitter, enable users to report content that violates their rules or terms of privacy. This content could be spam, harmful or malware links, multiple accounts, unrelated content, or aggressive or abusive content. OSNs also encourage users to recommend any additional action to improve their security and privacy.

Many OSNs have authenticated public figures and key brand identities and provided them with verified accounts. These accounts are distinguished in Twitter and Facebook by a blue badge (Fig. 6), which can be a sign to differentiate fake or normal accounts from the verified one. Typically, verified accounts are provided only to celebrities, politicians, and major brands; however, it is not impossible for ordinary people or small brands to obtain a verified account. Posts and contents of these verified accounts are most likely to be true.


Report This Link

All reports are strictly confidential. This appears to be:

○ Spam or scam
○ Contains hate speech or attacks an individual
○ Violence, crime, or self-harm
○ Nudity, pornography, or sexually explicit content

□ Permanently ban this person from this Facebook Page and remove all content they have posted.

Is this your intellectual property?

[redacted] 19 Apr
This is our final Tweet! Thnx ... spam reports to @support. For assistance ...
Expand

FIGURE 6. **Reporting harmful accounts and content in Facebook and Twitter.**

D. PREVENTION APPROACHES

Traditional approaches to defeating Sybil attacks are based either on trusting central authorities or connecting identities to resources that cannot be easily obtained by the attacker. This prevents the attacker from creating Sybil identities at the outset. These traditional approaches are called Sybil prevention. Many OSNs adapt these approaches, such as Cyworld,[1] to prevent Sybil attacks. In these approaches, the users must provide verified identities when creating new accounts, such as a social security or passport number. A simpler approach is to enforce users to solve a challenge-response test, such as CAPTCHA or crypto-puzzles, as a requirement to obtain access to system services [91], [92]. A more recent approach is to verify the user identity by sending a verification SMS message to the user's phone. Although this approach is widely used, attackers can bypass it by using virtual mobile number services, or by obtaining many accounts using disposable phones.

E. OTHER SUPPLEMENTARY WORK

Other researchers have moved to work on improvements on certain dimensions. For example, in [94] authors are trying to solve the problem of measuring Sybil groups that exist in large-scale online social networks based on different levels. These levels include malicious activities, individual information, and social relationships among other related aspects. The authors try to leverage the characteristics of Sybil groups in order to improve the security mechanisms of OSNs to provide defense against Sybil attacks.

The work in [95] try to solve the problem of retroactively identifying fake profiles by analyzing 62 million Twitter user profiles that are available in the public domain. The authors identify a sub-set of fake user accounts that are considered to be highly reliable after using a pattern-matching algorithm on the screen-names, as well as an elaborate analysis of the tweet update times. The authors try to reveal the distinct behavior of the user accounts that are labelled as fake, based on a ground truth data set, after analyzing the statistics of profile creation as well as the URLs of the accounts considered to be fake.

Similarly, authors in [96] trying to describe an accessible approach that can facilitate the finding of groups that consist of fake accounts that are registered by a similar actor. The major technique that they employ in their study is the use of a supervised machine learning pipeline that is useful in classifying a whole cluster of accounts as either legitimate or malicious. In their research study, they make use of major features like name, company, email address, or university, and these details feature some essence of patterns within a cluster. The authors apply their framework during the study and analyze the account data collected on LinkedIn, as grouped by two factors: registration IP address, and the registration date.

In [97], authors propose a solution to solve the problem of graph-based Sybil defense mechanisms experiencing high false detection rates. The authors focus on enhancing these mechanisms, through the consideration of extra graph structural information that makes up the building blocks of social networks. Their solution is based on proposing an effective graph pruning technique that seeks to enhance the local structural similarity between the neighboring nodes that are available in a social network.

FakeBook [98] try to explore the problem of malicious entities that are trying to exploit the vulnerabilities as well as the weaknesses of online social networks. The authors carry out an investigation on the possible approach that can be used to deal with the problem of malicious entities in online social networks. The authors try to explore this problem originally by analyzing social network graphs based on a dynamic point of view, putting more consideration on the context of threats to privacy.

Recent work by [99] looked to extend the work against Sybil attacks in distributed systems by creating a framework that accounts for the adversary's ability to create periphery attacks. This work was inspired by the shortcomings of graph-based Sybil detection models designed in prior literature whereby it is assumed that the number of links that an adversary can make between Sybil and honest nodes is restricted. While models leveraging this edge limiting assumption have worked against certain types of Sybil attacks, it has also been shown that they fail when Sybil attacks are characterized by isolated Sybil's connected to many edges to honest nodes. This sort of attack is known as periphery attacks and they violate the edge limiting assumption thus rendering conductance-based Sybil defense strategies futile.

IV. DISCUSSION

In this survey, we discuss studies performed to date on detecting Sybil attacks in OSNs. Many researchers have strived to show the devastating effects of Sybil attacks in OSNs, such as those of spamming or other malicious activities. These research interests have led to the development of

[1] http://www.cyworld.com/

methods to mitigate Sybil groups in OSNs, most notably community-based approaches that use social graphs to detect them.

A. IMPLICATIONS

Generally, Sybil attack defense techniques are divided into four categories: graph-based, machine learning, manual verification, and prevention approaches. The graph-based approaches consist of Sybil detection and Sybil tolerance techniques. Sybil detection techniques are used to identify and discover identities that tend to be Sybil by using some features of the OSN's structure.

Unlike traditional approaches, these schemes do not require central trusted identities. Instead, they depend on the trust that is incorporated in existing connections between these identities. Almost all Sybil detection techniques depend solely on community-detection algorithms to identify weak connections between different regions. They then apply the general assumptions of Sybil detection techniques. These assumptions, which are discussed in Section III-A are outlined below.

First, Sybil detection techniques assume that, in the social graph, two main regions exist—a legitimate region and a Sybil region—and the attacked edges between these regions are fewer than the edges between Sybil accounts. This assumption is based on the synthetic social graph; however, in a realistic graph, this assumption would not hold for an OSN.

Second, they assume that a legitimate region in a social graph is always fast mixing; however, this assumption cannot be confirmed as being true [50], [51], [53], [54] because real social graphs are randomly formed and dynamically grow. Moreover, Mohaisen et al. [50], [53] mentioned that the calculation of the property of mixing time in the social graph was used without careful measurements in the respective studies. Third, they assumed that attackers can infiltrate the social graph but cannot produce meaningful interactions, and thus the number of escaping tails to the attacker is always zero. Additionally, Sybil detection techniques rely on relationship information, which is not sufficient for distinguishing a Sybil region from a non-Sybil region.

Sybil tolerance techniques employ different algorithms by utilizing the graph structure and transaction history to limit the arbitrary impact of gaining multiple identities by the attacker. Unlike Sybil detection approaches, Sybil tolerance does not rely on community-detection algorithms, and it does not assume that a legitimate region is fast mixing. Rather, it relies on the assumption that a limited number of attack edges exist between the two regions. Sybil tolerance techniques furthermore leverage the transactions between different users and the credit network. As a result, existing systems are application-specific solutions. Unfortunately, Sybil detection and tolerance techniques are not applicable to real-world cases because they require previous knowledge about the whole graph and do not scale to large social networks on account of their complexity.

In addition to graph-based solutions, machine-learning approaches were developed to tackle the above problem. They are presented in [54], [80], [87], [82]. The approaches presented in each of these papers work well with respect to their own assumptions about differentiate between Sybil and non-Sybil nodes/users. However, they lack a clear conceptual unified framework to evaluate and compare those methods, especially when confronting different attack strategies and social structures. Machine-learning approaches are based on specific features, which make them vulnerable to Sybil attacks that are executed using different strategies. These approaches therefore require a large amount of ground truth data for training, which would not be feasible in current OSNs.

Manual verification approaches are often provided as tools from OSN operators; however, they solely depend on user feedback about the credibility of a certain profile or content. Additionally, many OSN operators impose strong identity instructions by requiring additional verification methods to allow a user to obtain an account in the network. Such an approach is classified as a Sybil prevention approach, whereby the user must send verified documents to prove his/her identity. This tactic is intended to prevent a single user from creating multiple accounts under different identities.

The core principles of the Sybil defense approach, specifically rankings and community detection, were addressed by Viswanath et al. [51]. They studied how the schemes would respond to changes in the OSN or behavior of the attacker. They acknowledged that, with the increase or decrease of a Sybil partition, an ordering could be imposed on the nodes by which they are added or removed. They concluded that all schemes use the ranking of nodes to define Sybil and the parameter settings define the cutoff point. The performances of Sybil defender schemes were evaluated on OSNs with different structures. It was determined that these schemes are more efficient in networks with minimal community structures. Table 1 summarizes all related schemes and methods based on the algorithms/ methods, datasets, and measurements used.

B. PERFORMANCE MEASUREMENTS

Different performance measurements were used to investigate the validity of Sybil defense techniques in OSNs, as we can see in Table 2. A common approach to measuring performance is the Receiver Operating Characteristic (ROC) curve, which presents the false-positive rate versus the false-negative rate for different values of a threshold, as shown in Fig. 7. Additionally, many researchers use modularity, centrality, precision, recall, sensitivity, and specificity, which are likewise common measures for different analysis tasks.

The developed techniques were all experimentally proven effective in the defense against Sybil attacks. SybilGuard [26] was shown to reduce the number of Sybil groups for 99.8% of legitimate users. That is, the scheme allows for 99.8% of legitimate users to accept other legitimate nodes.

TABLE 1. All related schemes and methods based on the methods, datasets, and measurements used.

Paper Author and Year	Scheme name	Defense Type	Algorithms/ methods	Assup-tion	Datasets	Measurements
Graph-based: Sybil detection schemes						
Yu et al. (2008) [26]	SybilGuard	Distributed	Incremental maintenance algorithm	(2)	Kleinberg's synthetic social network model	Random walk performed by each node
Yu et al. (2010) [25]	SybilLimit	Distributed	Approximation algorithm and sampling algorithm	(2)	Friendster, LiveJournal, DB LP, & Kleinberg	Random walks
Danezis & Mittal (2009) [62]	SybilInfer	Centralized	Bayesian inference	(2)	Power-law network and LiveJournal	Random walks
Tran et al. (2011) [59]	Gatekeeper	Distributed	Breadth-first search and random walk	(1,2)	YouTube, Digg	FPR and FNR rates
Shi et al. (2013) [61]	SybilShield	Distributed	Community detection algorithm with agent nodes	(1,2)	MySpace	Modified random walks
Wei et al. (2013) [63]	SybilDefender	Centralized	Limited (partial) random walk performed by node	(2)	Facebook and Orkut	Varying random walk length
Cao et al. (2012) [66]	SybilRank	Distributed	Community detection but not Sybil resilient algorithm	(1)	Facebook	ROC curve, FPR rates, & FNR rates
Cao & Yang (2013) [67]	Sybilfence	Distributed	Social-graph-based Sybil defenses	(2)	Facebook graph	ROC
Boshmaf et al. (2015) [69]	Íntegro	Distributed	Graph partitioning algorithms, random walk	(2,3)	Facebook, Tuenti, LinkedIn	ROC
GB: Sybil tolerance						
Mislove et al. (2008) [70]	Ostra	Distributed	"Remove" certain edges based on user feedback	(2)	YouTube	Proportion of attackers, maximum credit imbalance per link and false classification probability
Tran et al. (2009) [60]	SumUp	Centralized/ Decentralized	Voting envelope, approximation algorithm	(2)	YouTube, Flickr, Digg	Number of voters, time step and number of attack edges
Zhang et al. (2015) [72]	TrueTop		Modified PageRank algorithm		Twitter	Counting the final credits at every vertex
Viswanath et al. (2012) [99]	Canal	N/A	Landmark routing-based techniques	(2)	N/A	N/A
ML: Supervised methods						
Z. Yang et al. (2011) [54]	Uncovering social network Sybils in the wild	Centralized/ decentralized	Threshold-based detection scheme, SVM classifier	N/A	RenRen	Invitation Frequency, incoming and outgoing requests accepted, clustering coefficient
G. Wang et al. (2012) [80]	Social Turing tests		Crowdsource	N/A	Facebook, RenRen	Crowdsource, CDF, false negative Rate
ML: Unsupervised methods						
Wang et al. (2013) [86]	Clickstream	Centralized/ decentralized	K-nearest neighbor, nearest cluster, nearest cluster center	N/A	RenRen, LinkedIn	Normal cluster coverage, error rate, false positive rate, false negative rate
Cai & Jermaine (2012) [79]	Latent community model		Latent community model	N/A	Digg	Sensitivity, false positive/false negative rates
ML: Semi-supervised methods						
Gong et al. (2014) [64]	SybilBelief	Distributed	Loopy belief propagation	(2)	Facebook, Slashdot, Email	AUC ROC curve
Peng et al. (2015) [65]	SybilFrame	Distributed	Pairwise Markov random field model, loopy belief propagation	(2)	Twitter and Facebook	AUC, FNR, FPR
Analytical Studies						
Bimal et al. (2010) [51]	Comparative study	N/A	Off-the-shelf community detection algorithms	N/A	YouTube, Astrophysicists, Advogato, Facebook.	ROC, modularity, centrality measure
Pise & Kumar (2014) [93]	Comparative Study	N/A	Sybil identification algorithm, & Sybil community detection algorithm	N/A	Myspace data set	Theoretical probability analysis & experiments
Mohaisen et al. (2010) [53]	N/A	Distributed	Breadth first search (BFS) algorithm	N/A	Facebook, LiveJournal, Youtube, DBLP, Enron, Wiki-vote	Mixing time
Mohaisen et al. (2013) [50]	N/A	Distributed	Breadth-first search algorithm	(3)	DBLP, Facebook, Livejournal, Youtube, Wiki-vote, Enron, Epinion, DBLP	Interaction and mixing time

The performance of SybilInfer [62] was measured and compared to those of other state-of-the-art algorithms, i.e., SybilGuard [24] and SybilLimit. The performance of SybilInfer [62] was shown to be relatively good. SybilLimit [55] coped with up to f −0.02 compromised nodes as long as the node degree of the attacker was low. Very few compromised nodes were evident when SybilInfer [62] was used (f −0.01) within the interval in which SybilLimit [55] was applicable. SybiliInfer [62] enabled the nodes to control only 5% or fewer of the entities in the system. This was quite impressive when compared to the 30% control that SybilLimit[55] was shown to afford the nodes. In the limit of the SybilLimit [55] applicability range, i.e., when f −0.02, SybilInfer [62] limited the number of compromised entities to less than 8% while SybilLimit [55] could only manage 50% illegitimate entities.

The performance of SybilDefender [63] in the simulation evaluation showed a better performance than SybiLimit in terms of the false-positive and false-negative (FN) rates. In terms of computation, SybilDefender [63] likewise surpassed SybilLimit [55] in the time required to distinguish a legitimate node and a Sybil node. SybilDefender [63] required 0.87 s to test a Sybil node, while SybilLimit [55] required 11.56 s for the same network topology. For a legitimate node, SybilDefender [63] required 7.11 s, whereas SybilLimit required 83.55 s [55].

SybilRank [66] was shown to perform better than other Sybil defense techniques in the experimental evaluation. SybilRank [66] achieved at least a 70% accuracy rate, even when a 5,000 Sybil cluster had 1,500 attack edges. Despite the distribution of trust seeds among communities, the locations of trustworthy seeds did not influence the accuracy of the detection [58], [89].

The experiments performed on SybilFence [67] showed that it performed much better when the offset factor was large. The offset factor is the penalty assigned to nodes that receive rejections in a network. SybilFence performed better than SybilRank [66] under a Sybil flooding request.

The evaluation of VoteTrust compared it to other PageRank-like algorithms, specifically BadRank and TrustRank. It was shown that VoteTrust outperformed all of them. The comparison was conducted using a precision-recall curve. VoteTrust had higher precision with the same recall. TrustRank uses the same heuristics as those of VoteTrust; however, it may mix fewer popular users with Sybils. Thus, VoteTrust significantly improved the precision of TrustRank in both the manually checked dataset and the banned account dataset. SybilRank outperformed TrustRank; however, its performance was still insufficient because it had a higher false-positive rate. BadRank, on the other hand, did not efficiently detect Sybils in the human-checked dataset because Sybils were not present in the seed community.

The evaluation performed by Alvisi et al. in the study of Sybil defense evolution. The study adapted the variable length random walk algorithm of Andersen, Chung and Lang (ACL). The authors established two key results: ACL does need require preprocessing, and ACL tolerates denser attacks than SybilLimit.

Wang et al. conducted a study on the use of clickstream behavior for Sybil defense using a proposed unsupervised Sybil detector with excellent performance metrics. The detector was operated on real-world OSNs, including LinkedIn, RenRen, and Facebook, to test the accuracy. In LinkedIn, it could detect 1,700 users as Sybils of the 4,000 that had already been established as Sybils.

The performance of Sychrotrap was analyzed in terms of the results of the clustering pipeline under deployment in Facebook. The clustering pipeline was scrutinized under a variety of similarity thresholds. Under the set threshold values, the clustering jobs required less than 100 min thanks to the use of Giraph in incremental processing.

In a study that presented a tuned Sybil detector, two detection algorithms were compared for performance: a simple threshold detector and a complex learning algorithm detector, SVM. The detection accuracy performance of the tuned threshold detector was virtually equal to that of the rather expensive SVM. A better detector was determined to be one that combined feature-based detection and community detection.

In assessing the feasibility and performance of human crowdsourcing in Sybil detection, Wang compared the system's cost of using turkers to that of using moderate data, such as images in Facebook and Tuenti, at the time of the study. The cost of the turker workforce required for a human crowding system was determined to be much less expensive than the other workforces [90].

In the assessment of SybilControl [91], it was found that Regular Chord performed much better than SybilControl when no replications were used. The lookup process achieved a 1% success rate for the former. For two replications, the lookup processes in SybilControl were nearly 100%. The difference between the two was very small, showing that SybilControl had a very small impact on the lookup process. Similar results were demonstrated when the lookup hops were plotted versus the frequency of each protocol execution. The effects of increasing the frequency were drastic for no-replication Chord and no-replication SybilControl. Increasing the number of replications for both protocols reduced the effects of increasing the frequency; however, the number of lookup hop counts was still affected.

Three metrics were used in a comparison of SybilRank [66] and Sybil defense strategies: the ROC curve, false-positive rates, and false-negative rates. Cao and Yang's work on SybilFence was evaluated in a comparison to SybilRank. The evaluation entailed a simulation of user friend requests in social graphs, while friend request rejections were deemed negative feedback. Four graphs were simulated for this evaluation. The Facebook graph was sampled using the Forest Fire method. A Sybil region with 5,000 Sybil nodes was introduced to each graph. The author additionally employed three metrics to compare the Sybil defense strategies on these graphs: the ROC curve, false-positive rates, and false-negative rates.

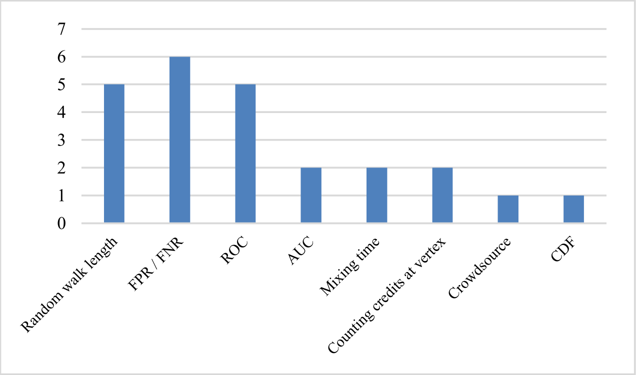


FIGURE 7. Performance measures used in the literature.

C. DATASETS AND SOCIAL SYSTEMS

The datasets used in analyzing various algorithms of Sybil defense schemes vary across studies, as shown in Table 3. The Stanford Network Analysis Platform (SNAP)[2] is one of the most powerful resources for analyzing social network data. SNAP provides free libraries and large open dataset collections that can be used for network analysis and graph mining. SNAP experimentally is scaled to massive networks that consists of hundreds of millions of nodes and billions of edges.

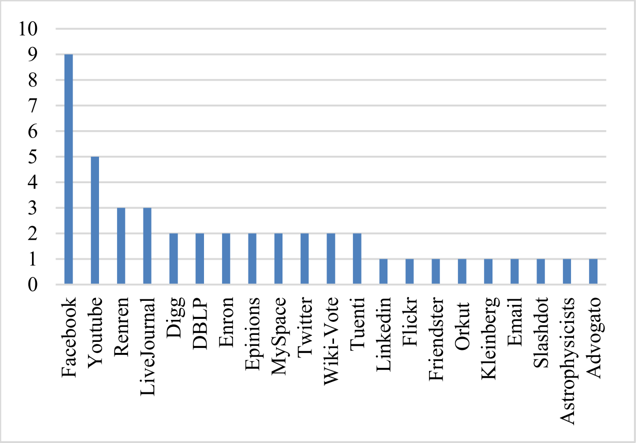


FIGURE 8. Number of studies per OSN dataset.

Figure 8 depicts a chart representing the number of studies associated with the targeted OSN. As shown in the chart, a significant number of studies applied their techniques on datasets collected from Facebook and YouTube. Facebook is an appropriate network on account of its explicit graph structure. Most of the studies focus on a structure level instead of a content level. Another reason could be the popularity of Facebook and YouTube, which makes them targets for social black markets to increase influence, views, followers, etc. Surprisingly, although it is popular, Twitter was not adequately studied in terms of Sybil attacks compared with Facebook and YouTube. This may be on account of its implicit graph structure. Generally, the studies that use Twitter datasets were targeting spam rather than Sybil attacks.

The guarantees of SybilGuard [26] were evaluated for experimental purposes using a simulation setup, which assumed that real-world social networks contain limited public information. The experiment employed Kleinberg's synthetic social network model to generate three graphs: a million-node graph with an average node degree of 24, a 10,000-node graph with an average degree of 24, and a 100-node graph with an average degree of 1.

The work on SybilInfer discussed two experiments that aimed to validate the protocol's ability to defend against Sybil attacks. The first experiment used data accrued from an application of SybilInfer in synthetic social network topologies. The second experiment used a sampled LiverJournal. A variant of Snowball sampling was used to collect the dataset that consisted of 100,000 nodes. The network was pre-processed, resulting in a social sub-graph with 33,000 nodes. SybilInfer was then applied to the topology without introducing an artificial attack.

The study by Wei et al. involved an experimental evaluation of SybilDefender using two datasets extracted from Orkut and Facebook. Construction of a Sybil region for the evaluation involved two popular network analysis models: the preferential attachment (PA) model and the Erdos–Renyi (ER) model. The experiment evaluated the Sybil identification algorithm and the Sybil community detection algorithm. Moreover, the experiment evaluated SybilDefender in a comparison with SybilGuard, SybilLimit, and other algorithms, such as GateKeeper and the community detection algorithm presented by Viswanath et al. SybilRank was evaluated by comparing it t other approaches: GateKeeper, Mislove CD, EigenTrust, SybilLimit, and SybilInfer. The author simulated social graphs to compromise Sybil regions. The Facebook graph used in the experiment was the result of Forest Fire sampling generated using a Barabasi scale-free model.

VoteTrust guarantees were evaluated using an experiment in RenRen. The experiment employed a PKU network with more than 200,000 users. The author added artificial Sybil accounts on the PKU network while determining real Sybils in the network. Adding artificial Sybils helped validate the theoretical bounds of the algorithm under attack strategies. Detection on a real network was conducted to compare VoteTrust to BadRank and TrustRank in real Sybil detection. The metrics used for the performance evaluation were precision and recall.

The study by Alvisi et al. on the evolution of Sybil defense strategies simulated four social graphs—Facebook, DBLP, Epinions, and WikiTalk—under attack models. Both ACL and SybilLimit were run on them to compare the performances of the algorithms. Wang et al., on the other hand, leveraged a dataset from RenRen to assess the possibility of employing user behavioral characteristics to study defense mechanisms against Sybil attacks. To validate the idea, the author analyzed the behavioral differences of Sybil and non-

[2] http://snap.stanford.edu/

TABLE 2. Performance measurements used in the literature.

Measurement technique	#Ref
Random walk length	[23][24][59][60][61]
FPR / FNR	[57][63][64][77][79][83]
ROC	[62][64][65][67][49]
AUC	[62][63]
Mixing time	[48][51]
Counting credits at vertex	[68][69]
Crowdsource	[77]
CDF	[77]

Sybil users by using ground truth data. The study was thus based on clickstreams for normal and Sybil users by employing data from RenRen as a ground truth data set. The study was based on the clickstream of 9,994 Sybils and 5,998 normal users on RenRen. In total, the data set had 1,008,031 and 5,856,941 clicks for Sybils and normal users, respectively.

The work by Zhi et al. presented a tuned threshold-based detector that was tested in RenRen in February 2011. Two sets of user accounts were used to build the detector (containing 1,000 Sybil and 100 non-Sybil accounts). The detector was implemented on the network for one year. Edge creation data from the experiment was then used to study the topological traits of the Sybil community. The feasibility of using human effort in Sybil detection was studied with a corpus of ground truth data from RenRen. This involved two datasets containing three different profiles: confirmed Sybils, confirmed legitimate users, and suspect profiles.

The performance overhead of implementing SybilControl in DHT was investigated using simulations in networks with 1,000 verifiable nodes. The nodes sent pings with challenges every 5 s with the puzzle time selected was 20 s. The experiment was intended to evaluate the lookup performance of Chord with and without SybilControl enabled. The success rate of attempted lookups was measured while the number of hops through nodes in the lookup were noted, including the failed nodes. The author hypothesized that the Chord lookup process would degrade when using SybilControl; thus, the use of multiple replications in SybilControl was proposed.

D. FAKE IDENTITIES LEVERAGING EMOTION IN SYBIL ATTACK

The question is -how do these fake identities fool normal users as they seek to form social connections with the and hence drive their agendas? It seems that it has more to do with human emotion, something that has been found to be a major factor behind activity on social networks. In fact, studies have established a correlation between social interaction in social networks and human emotion [102]–[110]. For instance, research done by the University of Glasgow's Institute of Neuroscience and Psychology identified vital emotions in social interactions, in this case: happy, sad, afraid/surprised and angry/disgusted. Such emotions are a primary reaction to stimuli and they clearly play a role in the choices made by users in their online communications [101], [102].

TABLE 3. Social network datasets used in the literature.

Social Network	Dataset range	#Ref
Facebook	#nodes: [182-3,097,165] #edges: [40,013 - 28,377,481]	[51][53] [63][64] [65][66] [67][69] [80]
Twitter	# nodes:[20,000,000 -1,768,065,468] # edges: [100,026,013 -256,000,000]	[65][72]
RenRen	# nodes:[200-10,000]	[80][86]
LiveJournal	# nodes: [100,000-2,000,000] # edges: [8,737,636-53,714,120]	[25][53] [62]
DBLP	# nodes: [106,002-614,981] # edges: [625,932-1,155,148]	[25][53]
YouTube	# nodes:[446,181-1,134,890] # edges: [1,728,938 -3,458,000]	[51][53] [59][60] [70]
Enron	# nodes: [10,000-33,696] # edges: [180,811 -105,343]	[53][66]
Wiki-vote	# nodes:7,066 # edges: 100,736	[53][66]
Epinion	# nodes:[10,000-75,879] # edges:[222,077-508,837]	[53][66]
Digg	# nodes:[539,242-594,426] # edges: [4,035,247-5,066,998]	[59][82]
Flicker	# nodes:1,530,000 # edges: 21,399,000	[60]
MySpace	# nodes: 100,000	[61][92]
Tuenti	# nodes: [60,000-11,291,486]	[66][70]
LinkedIn	# nodes: 40000	[86]
Friendster	# nodes: 932,512 # edges: 7,835,974	[25]
Orkut	# nodes:3,072,441 # edges: 117,185,083	[63]
Kleinberg	# nodes: 1,000,000 # edges: 10,935,294	[25]
Slashdot	# nodes: 82,168 # edges: 504,230	[64]
Astrophysicists	# nodes:14,845 # edges: 119,652	[51]
Advogato	# nodes:5,264 # edges: 43,027	[51]

Similarly, studies conducted by M. Guerini at Trento Rise in Italy and J. Staiano at Sorbonne Université in Paris found that emotion impacted social connections and activity when using social networks, albeit with different emotions sparking different types of discussions [103]. For example, emotions such as amusement, inspiration, and happiness drove users to broadcast via Facebook shares and tweets [104].

In an attempt to blend in well with normal users, social media spammers attach tailored images to their fake and compromised accounts. Such images have been found to cause human brains to release oxytocin, a 'feel good' chemical, meaning that the malicious user can easily create trust with the fake identity. For example, a malicious user targeting young male users can use an attractive photo of a woman in a bid to capture the attention of their target audience and they can even conjure up trust through emotional attachment [105]–[107].

These examples are just the tip of the iceberg, but the bottom line is that the strategies adopted by fake and compromised social media identities will keep evolving as social network operators create increasingly sophisticated methods to quash this behavior [108]–[110]. However, it is also expected that cognitive approaches will continue to prevail, seeing as they are one very effective way of forming smart social connections with innocent users.

E. IMPACTS OF FAKE AND SYBIL ACCOUNTS ON ONLINE SOCIAL NETWORKS

Research has shown that almost everyone who has spent some time on Twitter and Facebook has come in contact with fake profiles. As a matter of fact, 1 in 20 user accounts on Twitter are thought to be fake, while on Facebook, a little more than 1 in 100 active users are fake [108]. The fake accounts are usually run by low-paid humans or even by automated bots with the aim of inflating followers, pushing spam or malware and even influencing public discourse. The interesting bit is that fake accounts are quite easy to acquire online. A recent study done by Barracuda Labs found 20 sellers of fake accounts on the popular e-commerce platform, eBay. The research also evaluated the Google search results for the phrase "buy Twitter followers," where it was established that 58 out of the top 100 results sold fake followers, with 40% of the followers' accounts set up in the past 6 months. An average cost for acquiring 1000 followers was found to be very low, at around $18.

The use of fake accounts to create social support and to influence real users is escalating. For example, in the year 2010, a conservative group in Iowa State leveraged fake accounts to drive support for Republican Candidate Scott Brown in the Senate race. The campaign reached an audience of over 60,000 people, and some posit that this might have contributed to his victory in the election. The same scenario was witnessed in Mexico's 2012 election, wherein the Institutional Revolutionary Party used more than 10,000 fake accounts to skew online debates. Commercial campaigns have also been affected by automated accounts. A good example can be found in a 2014 study [109] on 12 million users of Weibo - a Chinese social network similar to Twitter. This particular study found that up to 4.7 million users participated in campaigns that looked to influence users on products and services. Most of them propagated messages and mentioned products and services from certain influential accounts - likely to be the brand behind the accounts. Sybil attack is bound to get much bigger in the coming years, as social networks become even more tightly coupled with personal spending. In other words, user may want to be very careful before accepting any new friend requests, as the black market in buying and selling fake accounts is bound to get even bigger.

V. CONCLUSION

In this paper, we surveyed state-of-the-art research relating to Sybil attack defense schemes and techniques in OSNs. We quantified these works from different perspectives, including the methodologies, algorithms, assumptions, and designed models. We further provided a discussion to summarize our observations based on the reviewed literature. This discussion may be useful for readers in gaining a better understanding of the problem of Sybil attacks and the methodologies that have been proposed to address them. Although various existing solutions are intended to resolve the problem of Sybil attacks in OSNs, they remain an immense problem. Further research is needed to stop fake users from causing negative effects, altering actions, and driving quantifiable outcomes while using online platforms.

Generally, a common evaluation to compare Sybil detection and Sybil tolerance schemes has not been undertaken for several reasons. First, obtaining the various datasets and design properties that each scheme uses is very challenging. Second, understanding the way in which the different techniques have been designed would enable them to be compared. However, this understanding has yet to be engendered. Moreover, the open problems of content analysis and content-based methods of node detection development remain. It is important to better analyze the "likes," posting of messages, spam keywords, etc. because this information may make the detection of Sybil nodes more effective. A superior detector is thus one that uses a mixture of feature-based detection and an enhanced social structure community detection approach.

Furthermore, the existing Sybil detection schemes require improvement. Researchers have the opportunity to address all of the above-mentioned drawbacks and continue to analyze the nature and speed of node mixing. In addition, researchers may address question if Sybil nodes may not always be malicious. For example, users might need several profiles in the network for various personal reasons. The developers of new Sybil detection methods must consider these concerns and define approaches to most effectively address them.

REFERENCES

[1] Y. Hu and M. Chen, "Information diffusion prediction in mobile social networks with hydrodynamic model," in *IEEE Int. Conf. Commun. (ICC)*, 2016, pp. 1–5.

[2] M. Chen, Y. Qian, S. Mao, W. Tang, and X. Yang, "Software-defined mobile networks security," *Mobile Netw. Appl.*, vol. 21, no. 5, pp. 729–743, Oct. 2016.

[3] M. AlRubaian, M. Al-Qurishi, M. Al-Rakhami, S. M. M. Rahman, and A. Alamri, "A multistage credibility analysis model for microblogs," in *Proc. IEEE/ACM Int. Conf. Adv. Social Netw. Anal. Mining*, Aug. 2015, pp. 1434–1440.

[4] M. Al-Qurishi, M. Al-Rakhami, M. AlRubaian, A. Alamri, and M. Al-Hougbany, "Online social network management systems: State of the art," *Proc. Comput. Sci.*, vol. 73, pp. 474–481, Jan. 2015.

[5] M. Al-Qurishi, R. Aldrees, M. AlRubaian, M. Al-Rakhami, S. M. M. Rahman, and A. Alamri, "A new model for classifying social media users according to their behaviors," in *Proc. 2nd World Symp. Web Appl. Netw. (WSWAN)*, Mar. 2015, pp. 1–5.

[6] Statista. (2016). *Statistics and Facts About Social Media Usage*. [Online]. Available: http://www.statista.com/topics/1164/social-networks/

[7] D. D. S. Gadgets. (Nov. 2016). *170 Amazing Twitter Statistics and Facts*. [Online]. Available: http://expandedramblings.com/index.php/march-2013-by-the-numbers-a-few-amazing-twitter-stats

[8] M. Al-Qurishi, M. Al-Rakhami, M. Alrubaian, A. Alarifi, S. M. M. Rahman, and A. Alamri, "Selecting the best open source tools for collecting and visualzing social media content," in *Proc. 2nd World Symp. Web Appl. Netw. (WSWAN)*, Mar. 2015, pp. 1–6.

[9] J. V. Grove. (2010). *Each Month 250 Million People Use Face-Book Connect on the Web*. [Online]. Available: http://mashable.com/2010/12/08/facebook-connect-stats/#F1vTxReW_kqx

[10] J. R. Douceur *et al.*, "The Sybil attack," in *Peer-to-Peer Systems*, P. Druschel, Ed. Heidelberg, Germany: Springer, 2002, pp. 251–260.

[11] H. Shen and X. Liu, "Detecting spammers on Twitter based on content and social interaction," in *Proc. Int. Conf. Netw. Inf. Syst. Comput. (ICNISC)*, Jan. 2015, pp. 413–417.

[12] G. Cluley. (2009). *Acai Berry Spammers Hack Twitter Accounts To Spread Adverts*. [Online]. Available: https://nakedsecurity.sophos.com/2009/05/24/acai-berry-spammers-hack-twitter-accounts-spread-adverts/

[13] C. Arthur. (2010). *Twitter Phishing Hack Hits BBC, Guardian and Cabinet Minister*. [Online]. Available: https://www.theguardian.com/technology/2010/feb/26/twitter-hack-spread-phishing

[14] L. Jin, Y. Chen, T. Wang, P. Hui, and A. V. Vasilakos, "Understanding user behavior in online social networks: A survey," *IEEE Commun. Mag.*, vol. 51, no. 9, pp. 144–150, Sep. 2013.

[15] N. Abokhodair, D. Yoo, and D. W. McDonald, "Dissecting a social Botnet: Growth, content and influence in Twitter," in *Proc. 18th ACM Conf. Comput. Supported Cooperat. Work Social Comput.*, 2015, pp. 839–851.

[16] Y. Boshmaf, I. Muslukhov, K. Beznosov, and M. Ripeanu, "The socialbot network: When Bots socialize for fame and money," in *Proc. 27th Annu. Comput. Secur. Appl. Conf.*, 2011, pp. 93–102.

[17] M. Wani, M. A. Alrubaian, and M. Abulaish, "A user-centric feature identification and modeling approach to infer social ties in OSNs," in *Proc. Int. Conf. Inf. Integr. Web-Based Appl. Services*, 2013, p. 107.

[18] SciTechBlog. (2010). *A New Look at Spam by the Numbers*. [Online]. Available: http://scitech.blogs.cnn.com/2010/03/26/a-new-look-at-spam-by-the-numbers/

[19] X. Han *et al.*, "CSD: A multi-user similarity metric for community recommendation in online social networks," *Expert Syst. Appl.*, vol. 53, pp. 14–26, Jul. 2016.

[20] J. Tang, X. Hu, H. Gao, and H. Liu, "Exploiting local and global social context for recommendation," presented at the 23rd Int. Joint Conf. Artif. Intell., Beijing, China, 2013.

[21] Y. Zhang, M. Chen, D. Huang, and Y. Li, "iDoctor: Personalized and professionalized medical recommendations based on hybrid matrix factorization," *Future Generat. Comput. Syst.*, vol. 66, pp. 30–35, Jan. 2017.

[22] Y. Zhang, "Grorec: A group-centric intelligent recommender system integrating social, mobile and big data technologies," *IEEE Trans. Serv. Comput.*, vol. 9, no. 5, pp. 786–795, Sep./Oct. 2016.

[23] H. Yu, C. Shi, M. Kaminsky, P. B. Gibbons, and F. Xiao, "Dsybil: Optimal Sybil-resistance for recommendation systems," in *Proc. 30th IEEE Symp. Secur. Privacy*, May 2009, pp. 283–298.

[24] M. Sirivianos, K. Kim, and X. Yang, "FaceTrust: Assessing the credibility of online personas via social networks," presented at the 4th USENIX Conf. Hot Topics Secur., Montreal, Canada, 2009.

[25] H. Yu, P. B. Gibbons, M. Kaminsky, and F. Xiao, "SybilLimit: A near-optimal social network defense against Sybil attacks," *IEEE/ACM Trans. Netw.*, vol. 18, no. 3, pp. 885–898, Jun. 2010.

[26] H. Yu, M. Kaminsky, P. B. Gibbons, and A. D. Flaxman, "SybilGuard: Defending against Sybil attacks via social networks," *IEEE/ACM Trans. Netw.*, vol. 16, no. 3, pp. 576–589, Jun. 2008.

[27] A. ALghamidi, "General investigations shake hands with the media in a joint mission to enlighten," in *Alriyadh News Paper*, vol. 17048. Riyadh, Saudi Arabia: Alyamamh Press, 2015.

[28] R. Gunturu. (2015). "Survey of Sybil attacks in social networks," [Online]. Available: https://arxiv.org/abs/1504.05522

[29] S. Y. Bhat and M. Abulaish, "Analysis and mining of online social networks: Emerging trends and challenges," *Wiley Interdiscipl. Rev., Data Mining Knowl. Discovery*, vol. 3, no. 6, pp. 408–444, 2013.

[30] C. T. Butts, "Social network analysis: A methodological introduction," *Asian J. Social Psychol.*, vol. 11, no. 1, pp. 13–41, 2008.

[31] A. Chinchore, F. Jiang, and G. Xu, "Intelligent Sybil attack detection on abnormal connectivity behavior in mobile social networks," in *Knowledge Management in Organizations*, L. Uden, Ed., Heidelberg, Germany: Springer, 2015, pp. 602–617.

[32] Y. Hu and L. Shi, "Visualizing large graphs," *Wiley Interdiscipl. Rev., Comput. Statist.*, vol. 7, 2015, pp. 115–136.

[33] M. Oliveira and J. Gama, "An overview of social network analysis," *Wiley Interdiscipl. Rev., Data Mining Knowl. Discovery*, vol. 2, no. 2, pp. 99–115, 2012.

[34] J. Scott, *Social Network Analysis*. Thousand Oaks, CA, USA: Sage, 2012.

[35] M. O. Jackson, *Social and Economic Networks*, vol. 3. Princeton, NJ, USA: Princeton Univ. Press, 2008.

[36] J. Piersa and T. Schreiber, "Spectra of the spike-flow graphs in geometrically embedded neural networks," in *Proc. Artif. Intell. Soft Comput.*, 2012, pp. 143–151.

[37] P. A. Dow, L. A. Adamic, and A. Friggeri, "The anatomy of large facebook cascades," in *Proc. ICWSM*, 2013, pp. 145–154.

[38] J. Ugander, B. Karrer, L. Backstrom, and C. Marlow. (2011). "The anatomy of the Facebook social graph." [Online]. Available: https://arxiv.org/abs/1111.4503

[39] A. M. Greenberg, W. G. Kennedy, and N. D. Bos, *Social Computing, Behavioral-Cultural Modeling and Prediction: 6th International Conference, SBP 2013* (Lecture Notes in Computer Science). Washington, DC, USA, Apr. 2013.

[40] L. C. Freeman, S. P. Borgatti, and D. R. White, "Centrality in valued graphs: A measure of betweenness based on network flow," *Social Netw.*, vol. 13, no. 2, pp. 141–154, 1991.

[41] D. Gómez, E. González-Arangüena, C. Manuel, G. Owen, M. del Pozo, and J. Tejada, "Centrality and power in social networks: A game theoretic approach," *Math. Social Sci.*, vol. 46, no. 1, pp. 27–54, 2003.

[42] S. P. Borgatti and M. G. Everett, "A graph-theoretic perspective on centrality," *Social Netw.*, vol. 28, 2006, pp. 466–484.

[43] I. Stanton and A. Pinar, "Constructing and sampling graphs with a prescribed joint degree distribution," *J. Experim. Algorithmics*, vol. 17, pp. 3–5, Jul. 2012.

[44] Y. Boshmaf, "A Quick Survey of Social Network-Based Sybil Defenses," Univ. British Columbia, Vancouver, Canada, 2012. [Online]. Available: http://blogs.ubc.ca/computersecurity/files/2012/04/571B_Sybil_Defenses1.pdf

[45] M. AlRubaian, M. Al-Qurishi, S. M. M. Rahman, and A. Alamri, "A novel prevention mechanism for Sybil attack in online social network," in *Proc. 2nd World Symp. Web Appl. Netw. (WSWAN)*, Mar. 2015, pp. 1–6.

[46] M. Conti, R. Poovendran, and M. Secchiero, "Fakebook: Detecting fake profiles in on-line social networks," in *Proc. Int. Conf. Adv. Social Netw. Anal. Mining (ASONAM)*, 2012, pp. 1071–1078.

[47] M. Y. Kharaji and F. S. Rizi. (2014). "An IAC approach for detecting profile cloning in online social networks." [Online]. Available: https://arxiv.org/abs/1403.2006

[48] K. Hoffman, D. Zage, and C. Nita-Rotaru, "A survey of attack and defense techniques for reputation systems," *ACM Comput. Surv. (CSUR)*, vol. 42, no. 1, 2009, Art. no. 1.

[49] H. Yu, "Sybil defenses via social networks: A tutorial and survey," *ACM SIGACT News*, vol. 42, no. 3, pp. 80–101, Sep. 2011.

[50] A. Mohaisen and J. Kim. (2013). "The Sybil attacks and defenses: A survey." [Online]. Available: https://arxiv.org/abs/1312.6349

[51] B. Viswanath, A. Post, K. P. Gummadi, and A. Mislove, "An analysis of social network-based Sybil defenses," *ACM SIGCOMM Comput. Commun. Rev.*, vol. 41, no. 4, pp. 363–374, 2011.

[52] F. Li, B. Liu, Z. Xiao, and Y. Fu, "Detecting and defending against Sybil attacks in social networks: An overview," in *Proc. 9th Int. Conf. Broadband Wireless Comput., Commun. Appl. (BWCCA)*, Nov. 2014, pp. 104–112.

[53] A. Mohaisen, A. Yun, and Y. Kim, "Measuring the mixing time of social graphs," in *Proc. 10th ACM SIGCOMM Conf. Internet Meas.*, 2010, pp. 383–389.

[54] Z. Yang, C. Wilson, X. Wang, T. Gao, B. Y. Zhao, and Y. Dai, "Uncovering social network Sybils in the wild," *ACM Trans. Knowl. Discovery Data (TKDD)*, vol. 8, no. 1, 2014, Art. no. 2.

[55] H. Yu, P. B. Gibbons, M. Kaminsky, and F. Xiao, "Sybillimit: A near-optimal social network defense against Sybil attacks," in *Proc. IEEE Symp. Secur. Privacy*, May 2008, pp. 3–17.

[56] L. Alvisi, A. Clement, A. Epasto, S. Lattanzi, and A. Panconesi, "Communities, random walks, and social Sybil Defense," *Internet Math.*, vol. 10, no. 3–4, pp. 360–420, 2014.

[57] W. Wei, F. Xu, C. C. Tan, and Q. Li, "Sybildefender: Defend against Sybil attacks in large social networks," in *Proc. IEEE INFOCOM*, Mar. 2012, pp. 1951–1959.

[58] A. Mohaisen, N. Hopper, and Y. Kim, "Keep your friends close: Incorporating trust into social network-based Sybil defenses," in *Proc. IEEE INFOCOM*, Apr. 2011, pp. 1943–1951.

[59] N. Tran, J. Li, L. Subramanian, and S. S. Chow, "Optimal Sybil-resilient node admission control," in *Proc. IEEE INFOCOM*, Apr. 2011, pp. 3218–3226.

[60] D. N. Tran, B. Min, J. Li, and L. Subramanian, "Sybil-resilient online content voting," in *Proc. NSDI*, 2009, pp. 15–28.

[61] L. Shi, S. Yu, W. Lou, and Y. T. Hou, "Sybilshield: An agent-aided social network-based Sybil defense among multiple communities," in *Proc. IEEE INFOCOM*, Apr. 2013, pp. 1034–1042.

[62] G. Danezis and P. Mittal, "SybilInfer: Detecting Sybil nodes using social networks," in *Proc. NDSS*, 2009, pp. 1–15.

[63] W. Wei, F. Xu, C. C. Tan, and Q. Li, "SybilDefender: A defense mechanism for Sybil attacks in large social networks," *IEEE Trans. Parallel Distrib. Syst.*, vol. 24, no. 12, pp. 2492–2502, Dec. 2013.

[64] N. Z. Gong, M. Frank, and P. Mittal, "SybilBelief: A semi-supervised learning approach for structure-based Sybil detection," *IEEE Trans. Inf. Forensics Security*, vol. 9, no. 6, pp. 976–987, Jun. 2014.

[65] P. Gao, N. Z. Gong, S. Kulkarni, K. Thomas, and P. Mittal. (2015). "Sybilframe: A defense-in-depth framework for structure-based Sybil detection." [Online]. Available: https://arxiv.org/abs/1503.02985

[66] Q. Cao, M. Sirivianos, X. Yang, and T. Pregueiro, "Aiding the detection of fake accounts in large scale social online services," in *Proc. 9th USENIX Symp. Netw. Syst. Design Implement. (NSDI)*, 2012, pp. 197–210.

[67] Q. Cao and X. Yang. (2013). "Sybilfence: Improving social-graph-based Sybil defenses with user negative feedback." [Online]. Available: https://arxiv.org/abs/1304.3819

[68] Q. Cao, M. Sirivianos, X. Yang, and K. Munagala, "Combating friend spam using social rejections," in *Proc. IEEE 35th Int. Conf. Distrib. Comput. Syst. (ICDCS)*, Jun./Jul. 2015, pp. 235–244.

[69] Y. Boshmaf *et al.*, "Integro: Leveraging victim prediction for robust fake account detection in OSNs," in *Proc. NDSS*, 2015, pp. 8–11.

[70] A. Mislove, A. Post, P. Druschel, and P. K. Gummadi, "Ostra: Leveraging trust to thwart unwanted communication," in *Proc. NSDI*, 2008, pp. 15–30.

[71] B. Viswanath, M. Mondal, K. P. Gummadi, A. Mislove, and A. Post, "Canal: Scaling social network-based Sybil tolerance schemes," in *Proc. 7th ACM Eur. Conf. Comput. Syst.*, 2012, pp. 309–322.

[72] J. Zhang, R. Zhang, J. Sun, Y. Zhang, and C. Zhang, "TrueTop: A Sybil-resilient system for user influence measurement on Twitter," *IEEE/ACM Trans. Netw.*, vol. 24, no. 5, pp. 2834–2846, Oct. 2015.

[73] M. Anitha, A. Kanchana, R. Padmapriya, and N. Malathi, "Simulation of semi Markov process to detect mimicking attacks based on user behavior," *Int. J. Adv. Res. Comput. Commun. Eng.*, vol. 5, no. 1, pp. 471–474, 2016.

[74] C. Schafer, "Detection of compromised email accounts used for spamming in correlation with mail user agent access activities extracted from metadata," in *Proc. IEEE Symp. Comput. Intell. Secur. Defense Appl. (CISDA)*, May 2015, pp. 1–6.

[75] D. S. Sisodia and S. Verma, "Analysis of spamming threats and some possible solutions for online social networking sites (OSNS)," *Analysis*, vol. 1, p. 27207, Mar. 2015.

[76] K. P. Murphy, *Machine Learning—A Probabilistic Perspective*. Cambridge, MA, USA: MIT Press, 2012.

[77] B. Batrinca and P. C. Treleaven, "Social media analytics: A survey of techniques, tools and platforms," *AI Socur.*, vol. 30, no. 1, pp. 89–116, 2015.

[78] A. C. Bahnsen, D. Aouada, A. Stojanovic, and B. Ottersten, "Feature engineering strategies for credit card fraud detection," *Expert Syst. Appl.*, vol. 51, Jun. 2016, pp. 134–142.

[79] RenRen Home page. (2016). [Online]. Available: http://www.renren-inc.com/en/

[80] G. Wang *et al.* (2012). "Social turing tests: Crowdsourcing Sybil detection." [online]. Available: https://arxiv.org/abs/1205.3856

[81] K. R. Canini, B. Suh, and P. L. Pirolli, "Finding credible information sources in social networks based on content and social structure," in *Proc. IEEE 3rd Inernat. Conf. Social Comput. Privacy, Secur., Risk Trust (PASSAT) (SocialCom)*, Oct. 2011, pp. 1–8.

[82] Z. Cai and C. Jermaine, "The latent community model for detecting Sybil attacks in social networks," in *Proc. NDSS*, 2012.

[83] H. Gao, J. Hu, C. Wilson, Z. Li, Y. Chen, and B. Y. Zhao, "Detecting and characterizing social spam campaigns," in *Proc. 10th ACM SIGCOMM Conf. Internet Meas.*, 2010, pp. 35–47.

[84] Z. Yanbin, "Detecting and characterizing spam campaigns in online social networks," in *Fall LERSAIS IA Seminar*, Univ. Pittsburgh, Pittsburgh, PA, USA, 2010.

[85] S. Qian *et al.*, "Social event classification via boosted multi-modal supervised latent dirichlet allocation," *ACM Trans. Multimedia Comput. Commun. Appl. (ACM TOMM.)*, vol. 11, no. 2, Jan. 2015, pp. 27.1–27.22.

[86] A. Mukherjee *et al.*, "Spotting opinion spammers using behavioral footprints," in *Proc. 19th ACM SIGKDD Int. Conf. Knowl. Discovery Data Mining*, vol. 2013, pp. 632–640

[87] G. Wang, T. Konolige, C. Wilson, X. Wang, H. Zheng, and B. Y. Zhao, "You are how you click: Clickstream analysis for Sybil detection," in *Proc. USENIX Secur.*, 2013, pp. 1-15.

[88] S. Misra, A. S. M. Tayeen, and W. Xu, "SybilExposer: An effective scheme to detect Sybil communities in online social networks," in *Proc. IEEE Int. Conf. Commun. (ICC)*, Kuala Lumpur, Malaysia, May 2016, pp. 1–6, doi: 10.1109/ICC.2016.7511603.

[89] K. Krombholz, D. Merkl, and E. Weippl, "Fake identities in social media: A case study on the sustainability of the facebook business model," *J. Service Sci. Res.*, vol. 4, no. 2, pp. 175–212, 2012.

[90] Y. Zhang, M. Chen, S. Mao, L. Hu, and V. Leung, "CAP: Crowd activity prediction based on big data analysis," *IEEE Netw.*, vol. 28, no. 4, pp. 52–57, Jul. 2014.

[91] F. Li, P. Mittal, M. Caesar, and N. Borisov, "SybilControl: Practical Sybil defense with computational puzzles," in *Proc. 7th ACM Workshop Scalable Trusted Comput.*, Raleigh, NC, USA, 2012, pp. 67–78.

[92] L. V. Ahn, M. Blum, and J. Langford, "Telling humans and computers apart automatically," *Commun. ACM*, vol. 47, no. 2, pp. 56–60, Feb. 2004.

[93] N. Borisov, "Computational puzzles as Sybil defenses," in *Proc. 7th IEEE Int. Conf. Peer-to-Peer Comput. (P2P)*, Sep. 2006, pp. 171–176.

[94] S. Pise and R. Kumar, "Recent trends in Sybil attacks and defense techniques in social networks," *Int. J. Eng. Res. Technol.*, vol. 3, no. 1, pp. 1147–1153, 2014.

[95] J. Jing, Z.-F. Shan, X. Wang, L. Zhang, and Y.-F. Dai, "Understanding Sybil groups in the wild," *J. Comput. Sci. Technol.*, vol. 30, no. 6, pp. 1344–1357, 2015.

[96] S. Gurajala, J. S. White, B. Hudson, and J. N. Matthews, "Fake Twitter accounts: Profile characteristics obtained using an activity-based pattern detection approach," in *Proc. Int. Conf. Social Media Soc.*, 2015, Art. no. 9.

[97] C. Xiao, D. M. Freeman, and T. Hwa, "Detecting clusters of fake accounts in online social networks," in *Proc. 8th ACM Workshop Artif. Intell. Secur.*, 2015, pp. 91–101.

[98] H. Zhang, J. Zhang, C. Fung, and C. Xu, "Improving Sybil detection via graph pruning and regularization techniques," in *Proc. 7th Asian Conf. Mach. Learn.*, 2015, pp. 189–204.

[99] M. Conti, R. Poovendran, and M. Secchiero, "FakeBook: Detecting fake profiles in on-line social networks," in *Proc. IEEE Comput. Society Int. Conf. Adv. Social Netw. Anal. Mining (ASONAM)*, Aug. 2012, pp. 1071–1078.

[100] G. Schoenebeck, A. Snook, and F.-Yi Yu, "Sybil Detection using latent network structure," in *Proc. ACM Conf. Economics Comput.*, 2016, pp. 739–756.

[101] M. Alrubaian, M. Al-Qurishi, M. Al-Rakhami, M. M. Hassan, and A. Alamri. (2016). "Reputation-Based Credibility Analysis of Twitter Social Network Users," *Concurrency Computation: Practice and Experience*. [Online]. Available: http://onlinelibrary.wiley.com/doi/10.1002/cpe.3873/full

[102] M. Chen, Y. Zhang, Y. Li, S. Mao, and V. C. M. Leung, "EMC: Emotion-aware mobile cloud computing in 5G," *IEEE Netw.*, vol. 29, no. 2, pp. 32–38, Mar./Apr. 2015.

[103] M. Guerini and J. Staiano, "Deep feelings: A massive cross-lingual study on the relation between emotions and virality," in *Proc. 24th Int. Conf. World Wide Web*, 2015, pp. 299–305.

[104] M. S. Hossain, G. Muhammad, M. F. Alhamid, B. Song, and K. Al-Mutib "Audio-visual emotion recognition using big data towards 5G," *Mobile Netw. Appl.*, vol. 21, no. 5, pp. 753–763, Oct. 2016.

[105] M. S. Hossain and G. Muhammad, "Audio-visual emotion recognition using multi-directional regression and ridgelet transform," *J. Multimodal User Inter.*, vol. 10, no. 4, pp. 325–333, Dec. 2016.

[106] H. Lin, W. Tov, and L. Qiu, "Emotional disclosure on social networking sites: The role of network structure and psychological needs," *Comput. Human Behavior*, vol. 41, pp. 342–350, Dec. 2014.

[107] J. Bollen, B. Gonçalves, G. Ruan, and H. Mao, "Happ iness is assortative in online social netwoks," *Artif. Life*, vol. 17, no. 3, pp. 237–251, 2011.

[108] M. Chen *et al.*, "CP-Robot: Cloud-assisted pillow robot for emotion sensing and interaction," in *Proc. Int. Conf. Ind. IoT Technol. Appl.*, 2016, pp. 81–93.

[109] T. Simonite, (2015). *Bots Infiltrate Social Networks, Push Ideas, Products*. [Online]. Available: https://www.technologyreview.com/s/535901/fake-persuaders/

[110] M. S. Hossain, G. Muhammad, B. Song, M. M. Hassan, A. Alelaiwi, and A. Almari, "Audio-visual emotion-aware cloud gaming framework," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 25, no. 12, pp. 2105–2118, Dec. 2015.

[111] Q. Fang *et al.*, "Relational user attribute inference in social media," *IEEE Trans. Multimedia*, vol. 17, no. 7, pp. 1031–1044, Jul. 2015.

[112] M. Chen, Y. Zhang, Y. Li, M. M. Hassan, and A. Alamri, "AIWAC: Affective interaction through wearable computing and cloud technology," *IEEE Wireless Commun.*, vol. 22, no. 1, pp. 20–27, Feb. 2015.

MUHAMMAD AL-QURISHI received the master's degree in information systems from King Saud University, Riyadh, Saudi Arabia, where he is currently pursuing the Ph.D. degree with the Information Systems Department, College of Computer and Information Sciences. He has published several papers in refereed IEEE/ACM/Springer journals and conferences. His research interests include online social networks, social media analysis and mining, human-computer interaction, and health technology.

MABROOK AL-RAKHAMI received the master's degree in information systems from King Saud University, Riyadh, Saudi Arabia, where he is currently pursuing the Ph.D. degree with the Computer Sciences Department, College of Computer and Information Sciences. He has authored several papers in refereed IEEE/ACM/Springer conferences and journals. His research interests include social networks, cloud computing, and health technology.

ATIF ALAMRI is currently an Associate Professor with the Information Systems Department, College of Computer and Information Sciences, King Saud University. Riyadh, Saudi Arabia. His research interests include multimedia-assisted health systems, ambient intelligence, and service-oriented architecture.

Prof. Alamri was a Guest Associate Editor of the IEEE Transactions on Instrumentation and Measurement, a Co-Chair of the first IEEE International Workshop on Multimedia Services and Technologies for E-health, a Technical Program Co-Chair at the 10th IEEE International Symposium on Haptic Audio Visual Environments and Games He also serves as a Program Committee Member for many conferences in multimedia, virtual environments, and medical applications.

MAJED ALRUBAIAN (S'15) received the master's degree in information systems from King Saud University, Riyadh, Saudi Arabia, where he is currently pursuing the Ph.D. degree with the Information Systems Department, College of Computer and Information Sciences. He has authored several papers in the refereed IEEE/ACM/ Springer journals and conferences. He is a Student Member of the ACM. His research interests include social media analysis, data analytics and mining, social computing, information credibility, and cyber security.

SK MD MIZANUR RAHMAN received the Ph.D. degree in risk engineering (Major: cyber security engineering), Laboratory of Cryptography and Information Security, Department of Risk Engineering, University of Tsukuba, Japan, in 2007.

He was involved in cryptography and security engineering in the high-tech industry in Ottawa, Canada. He was also a Post-Doctoral Researcher with the University of Ottawa, the University of Ontario Institute of Technology, and the University of Guelph, Canada. He is currently an Assistant Professor with the Information Systems Department, College of Computer and Information Sciences, King Saud University, Saudi Arabia. His research interests include cryptography, software and network security, privacy, cloud-security, sensor network security, white box-cryptography, and Internet of Things security.

Dr. Rahman was awarded the IPSJ Digital Courier Funai Young Researcher Encouragement Award from the Information Processing Society of Japan for his excellent contribution in IT security research. He received the "Gold Medal" Award for the marks of distinction in his undergraduate and graduate program.

M. SHAMIM HOSSAIN (SM'09) received the Ph.D. degree in electrical and computer engineering from the University of Ottawa, Canada. He is currently an Associate Professor with King Saud University, Riyadh, Saudi Arabia. His research interests include serious games, social media, IoT, cloud and multimedia for health care, smart health, and resource provisioning for big data processing on media clouds. He has authored and co-authored around 120 publications, including refereed IEEE/ACM/Springer/Elsevier journals, conference papers, books, and book chapters. He has served as a member of the organizing and technical committees of several international conferences and workshops.

Prof. Hossain has served as a Co-Chair, a General Chair, a Workshop Chair, a Publication Chair, and a TPC for more than 12 IEEE and ACM conferences and workshops. He currently serves as a Co-Chair of the 7th IEEE ICME Workshop on Multimedia Services and Tools for E-health MUST-EH 2017. He received a number of awards, including the Best Conference Paper Award, the 2016 *ACM Transactions on Multimedia Computing, Communications and Applications* Nicolas D. Georganas Best Paper Award, and the Research in Excellence Award from King Saud University. He is on the Editorial Board of the IEEE Access, *Computers and Electrical Engineering* (Elsevier), *Games for Health Journal*, and the *International Journal of Multimedia Tools and Applications* (Springer). Previously, he served as a Guest Editor of the IEEE Transactions on Information Technology in Biomedicine (currently JBHI), the *International Journal of Multimedia Tools and Applications* (Springer), *Cluster Computing* (Springer), *Future Generation Computer Systems* (Elsevier), *Computers and Electrical Engineering* (Elsevier), and the *International Journal of Distributed Sensor Networks*. He currently serves as a Lead Guest Editor of the *IEEE Communication Magazine*, the IEEE Transactions on Cloud Computing, and the IEEE Access and Sensors (MDPI). He is a member of ACM and ACM SIGMM.