Computers and Electrical Engineering 000 (2017) 1–13

Contents lists available at ScienceDirect

Computers and Electrical Engineering

journal homepage: www.elsevier.com/locate/compeleceng

Fake profile detection techniques in large-scale online social networks: A comprehensive review☆

Devakunchari Ramalingam*, Valliyammai Chinnaiah

Department of Computer Technology, MIT campus, Anna University, Chennai, 600044, India

ARTICLE INFO

Article history:
Received 2 December 2016
Revised 4 May 2017
Accepted 6 May 2017
Available online xxx

Keywords:
Fake profile detection
Online social networks
Sybil attacks
Big data

ABSTRACT

In the present era, online social networks are the most popular and rapid information propagation applications on the Internet. People of all ages spend most of their time on social networking sites. Huge volumes of data are being created and shared through social networks around the world. These interests have given rise to illegitimate users who engage in fraudulent activities against social network users. On social networks, fake profile creation is considered to cause more harm than any other form of cyber crime. This crime has to be detected even before the user is notified about the fake profile creation. Many algorithms and methods, most of which use the huge volume of unstructured data generated from social networks, have been proposed for the detection of fake profiles. This study presents a survey of the existing and latest technical work on fake profile detection.

1. Introduction

Social media is growing incredibly fast these days, which is important for marketing campaigns and celebrities who try to promote themselves by growing their base of followers and fans. However, fake profiles, created seemingly on behalf of organizations or people, can damage their reputations and decrease their numbers of likes and followers. They also suffer from fake updates and unnecessary confusion with other people. Fake profiles of all kinds create negative effects that counteract the advantages of social media for businesses in advertising and marketing and pave the way for cyber bullying. The users have different concerns regarding their privacy in an online environment. Fire et al. [1] described the threats of which users are unaware in Online Social Networks (OSNs). These include loss of privacy, identity theft, malware, fake profiles (Sybil's/social bots), and sexual harassment, among others. OSNs have billions of registered users. Facebook is the most famous OSN with more than a billion active users. There are basically four kinds of threats in OSN: classic threats, modern threats, combination threats, and threats targeting children. Several suggested solutions to these threats fall into three categories: operator, commercial, and academic solutions. The mechanisms in each of these categories can help to overcome the security threats in OSNs. Social engineering [2] is the primary cause of many kinds of security and privacy threats in OSNs. The main approaches to social engineering are social-technical, technical, physical, and social, and these are generally carried out using software or humans. The channels for social engineering are e-mail, instant messenger, telephone, Voice over Internet Protocol (VoIP), OSN, cloud, websites and physical channels. The attacks themselves are based on dumpster diving, advanced persistent threats, baiting, and phishing, shoulder surfing, reverse social engineering, and water holing.

☆ Reviews processed and recommended for publication to the Editor-in-Chief by Guest Editor Dr. R. Varatharajan.

* Corresponding author at: No. 3, Kasturibai Street, Muthulakshmi Nagar, Chitlapakkam, Chennai, 600064, India.
E-mail addresses: devakunchari.r@gmail.com (D. Ramalingam), cva@annauniv.edu (V. Chinnaiah).

http://dx.doi.org/10.1016/j.compeleceng.2017.05.020
0045-7906/

There are also state-of-the-art attacks, including social phishing, context aware spam, fake profiles, spear phishing, and fake identities in the cloud. An analysis of security threats to OSNs shows that fake profiles (Sybil's or Social bots) are the most important cause of these threats. Fake profiles must be detected even before such profiles are registered as OSN members. Such detection methods are discussed later in this paper. Many organizations have started to access the unstructured data available in OSNs in order to gain useful insight about the big data available to them. The influence of big data on fake profile detection is also discussed in the following sections.

1.1. Motivation and contribution

There are many social networking sites including Twitter, Facebook, Google+, Myspace, Instagram, Tumblr, Foursquare and LinkedIn. There were 823 million people who used Facebook daily on their mobile devices, which is an increase from the 654 million such users in the previous quarter. Social networking sites such as Facebook cannot yet deliver notifications regarding fake profiles in real-time, and discriminating between real and fake profiles is difficult for non-technically savvy users. Moreover, many big data issues, including data storage, how to handle streaming data, and how to provide immediate responses to users, must be handled while simultaneously operating on large volumes of data to achieve accurate profile identification results.

The main contributions of this paper are an exploration of the various diverse aspects of fake profile detection techniques and models proposed before 2012, as well as a focus on recent OSN Sybil detection studies that have not been previously considered. The year 2012 is chosen because, as stated by Nowotarski [3], there is an increase in social networking patent applications and issued patents every year. The paper identifies the improvements in detection techniques over the years and identifies possible future developments. In addition, several metrics are examined to analyze and compare earlier and more recent models.

1.2. Organization

The rest of this paper is organized as follows. Research related to OSN security threats is discussed in Section 2. Early fake profile and Sybil detection methods are reviewed in Section 3. Recent work and detection models are discussed briefly in Section 4. Based on this review of previous work, future research directions are discussed in Section 5. The study is concluded in Section 6.

2. Research related to OSN security threats

Many fraudulent activities occur on OSNs, and so OSN data must be handled in such a way that it can be made useful for many purposes such as fraud detection, criminal activity, political opinions, and risk management. Viswanath et al. [4] compared the prior Sybil defense algorithms against each other and analyzed, by node ranking, whether community detection algorithms can defend Sybil's with accuracy. Ferrara et al. [5] analyzed the dominant behavioral features that differentiate fake users from human ones and classified the studied literatures into a taxonomy of bot detection approaches, namely, graph-based social bot detection, crowd sourcing-based social bot detection and combinations of multiple approaches. Koll et al. [6] investigated the vulnerability of Sybil detection and Sybil tolerance solutions for the Sybil attacks under classical and modern scenarios. Various existing security issues, privacy leaks and deceptive behavior have been investigated with respect to OSNs. The reputation system describes the techniques involved in attacks and their defense mechanisms.

3. Existing models

This section reviews the literature on social network bot detection and classifies it under three major categories: feature- or content-based defense, network structure- or graph-based defense and hybrid techniques that combine both. Most of the previous Sybil detection works fall under social graph-based defense, which takes into account the link (edges/relations) and node (users) features [7,8]. Private OSN analysis reveals how to ensure privacy by reconstructing the whole graph into private pieces so that malicious users cannot report false information about the graph to the users. Earlier, the link prediction on OSNs can be solved using common neighbors, Jaccard's coefficient and Adamic/Adar. Adamic/Adar is valuable in that it uses preferential attachment and is based on ensemble paths of hit times, page ranks, and other variants.

The different historical Sybil detection techniques, along with their characteristics, assumptions, dataset, detection type and operating threshold, are summarized in Fig. 1. The standard previous works shown in Fig. 1 are graph-based defenses that assume that social networks are fast-mixing. Sybil Guard [9] was developed as the foremost Sybil admission control protocol. For 'n' real users, it limits the Sybil admission by $O(\sqrt{n}\log n)$ per attack edge, but Sybil limit [10] bounds the Sybil admission by $O(\log n)$per attack edge. Sybil Infer [11] allows for a higher degree of attacker nodes, whereas Sybil Limit renders strong guarantees merely on low attack edges. Additionally, Sum-up [12], a Sybil resilient vote aggregation mechanism that exploits user feedback, outperforms Sybil Limit by restricting adversary votes to one per attack edge. Mislove's Algorithm [13] establishes local communities using a greedy approach to partition the social graph into honest and fake regions. However, it costs $O(n^2)$, as shown in Table 1. Gatekeeper [14] allows $O(\log n)$ Sybil's per attack edges, which is similar to

Please cite this article as: D. Ramalingam, V. Chinnaiah, Fake profile detection techniques in large-scale online social networks: A comprehensive review, Computers and Electrical Engineering (2017), http://dx.doi.org/10.1016/j.compeleceng.2017.05.020

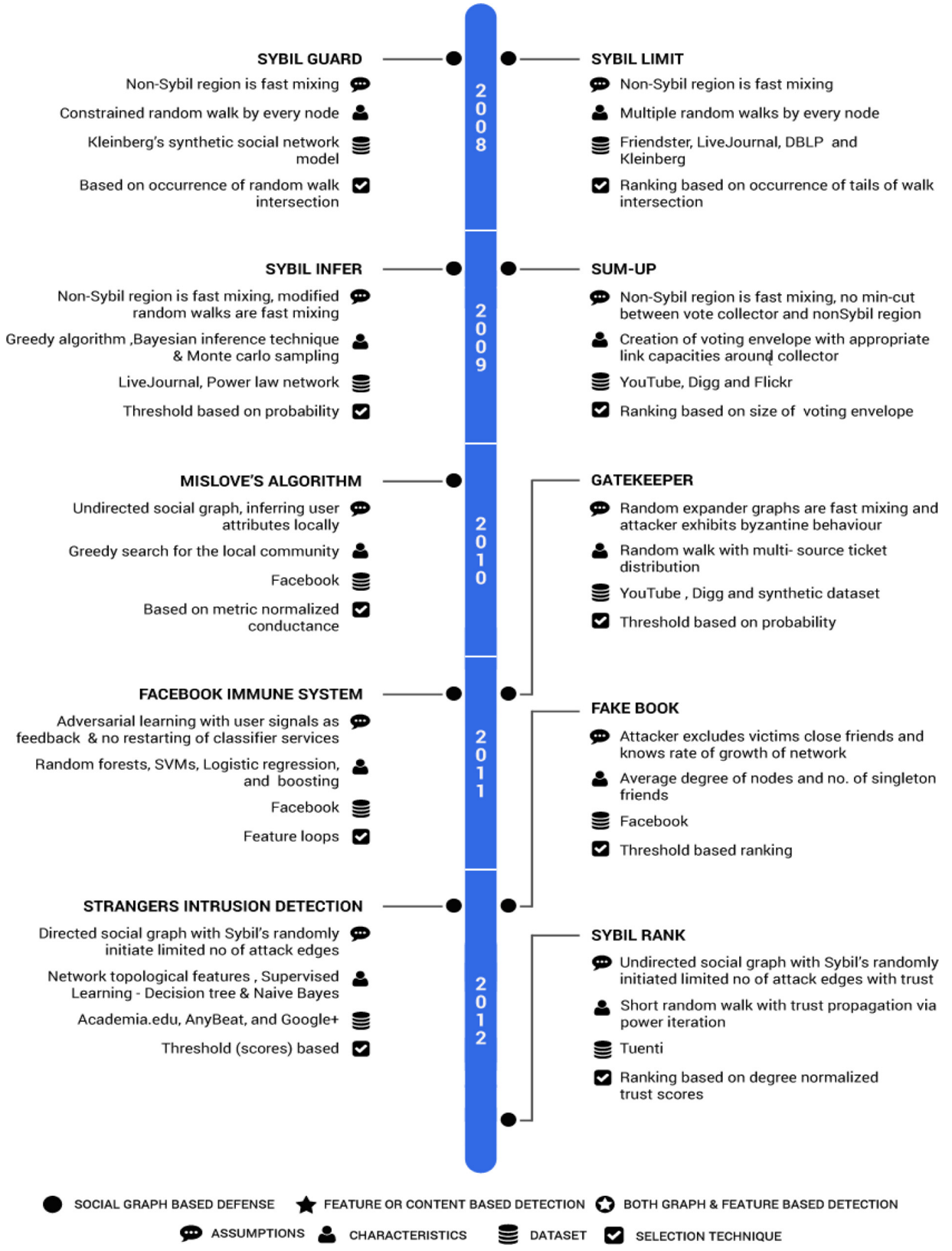

Fig. 1. Previous works on fake profile detection techniques with the corresponding detection models and their features.

Please cite this article as: D. Ramalingam, V. Chinnaiah, Fake profile detection techniques in large-scale online social networks: A comprehensive review, Computers and Electrical Engineering (2017), http://dx.doi.org/10.1016/j.compeleceng.2017.05.020

Table 1
Summary of time complexity of Sybil detection algorithms.

Year	Detection model	Computational Cost of Algorithms
2008	Sybil Guard	$O(\sqrt{mn}\log n)$
2008	Sybil Limit	$O(\sqrt{mn}\log n)$
2009	Sybil Infer	$O(n(logn)^2)$
2009	Sum - Up	$O(\log n)$
2010	Mislove's Algorithm	$O(n^2)$
2011	Gatekeeper	O(nlog n)
2012	Sybil Rank	O(nlog n))

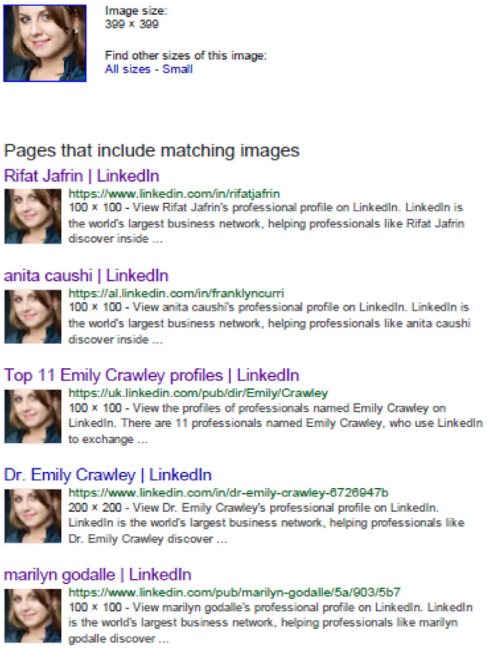

Fig. 2. Fake profile of Emily Crawley on LinkedIn, with similar images in Google image search.

Sybil limit, but it does not depend on the size of the social network. It costs O(nlog n) per honestnode. Another graph partitioning approach, Sybil Rank [15] outperforms other previous approaches by employing a seed selection method and early terminated random walks to propagate trust through O(log n)Power iterations. The cost of a single power iteration is O(n). The total cost of the algorithm is O(nlog n), independent of the number of honest seeds. The overall computation cost of Sybil detection algorithms is summarized in the Table 1.

Previous work on fake identities addresses experiments that analyze countermeasures against the behaviors of fake users. A Facebook social engineering experiment using the Facebook Graph API analyses qualitative data, such as the number of female and male friends, user data records, mutual friend cluster analysis, work and education information, location information, and common interests of Facebook users. The counter measures for protecting users from attackers include privacy awareness, privacy regulations, privacy enhancing strategies, and awareness training.

Facebook uses the Facebook Immune System [16] to identify spam, suspicious links, and patterns of user behavior in their social network. However, spammers often attempt to elude these security measures and spread malware or inappropriate content to users. Though OSNs employ measures to protect the user's information, it is the responsibility of the users themselves to accept friend requests that are real profiles because their privacy is also their friend's privacy. Social bots are the fastest means through which fake owners can initiate connections with real users in an OSN. This activity is detected using the assumption that the attacker knows the defense strategy employed and eventually tries to reduce the number of random friend requests that are accepted.

4. Recent work on privacy-preserving fake profile detection

Fake profiles have been detected in the LinkedIn dataset from a set of profiles available on the network [17], as in Fig. 2.

The detection process involves a few steps. First, profiles are processed to extract features using principal component analysis. These features are used to develop a training model, using the Resilient Back Propagation algorithm in neural

Please cite this article as: D. Ramalingam, V. Chinnaiah, Fake profile detection techniques in large-scale online social networks: A comprehensive review, Computers and Electrical Engineering (2017), http://dx.doi.org/10.1016/j.compeleceng.2017.05.020

networks and Support Vector Machines (SVMs), for classification. Test data is given to the neural network and SVM to classify fake profiles. Finally, the detection accuracy of this model is compared with other methods.

Sybil Frame [18] uses a multi-stage level classification mechanism to detect Sybil's on Facebook and Twitter. There are two types of approaches to address Sybil regions, content-based and structure-based approaches. Stage 1 explores a dataset and extracts information used to calculate prior information about its nodes and edges. Stage 2 correlates the nodes using a Markov random field and loopy belief propagation, which employs posterior information. Sybil detection is also done by analyzing user click streams. Friend recommendation schemes are used to detect independent sybils and sybils that collude with identified sybils.

Vote Trust [19] reveals the detection mechanism for the classification of benign and fake user accounts on OSNs. This detection depends on the prediction that victims use user-level activities. Unique features of the user accounts are extracted and applied to a classifier. An OSN graph is analyzed with the assumption that fake accounts have very few edges. A friend invitation graph is drawn based on incoming and outgoing links between nodes. Trust is then calculated using the votes posted for acceptance and rejection of the requests from the users based on personal influence. The trust is also globally calculated based on the votes from the nodes in the complete network. The trust is propagated to the entire network and is used as a basic criterion for detecting sybils. The community detection identifies the victims around the identified sybils. However, on Twitter, it is possible to sell legitimate accounts that have been compromised. Furthermore, attackers can increase the number of links in order to increase the acceptance of sybils. Hence, vote trust is only the first level of defense. Fig. 3 provides a summary of recent studies on fake profile identification and privacy preservation, along with their characteristics, assumptions, operating thresholds and OSN datasets.

4.1. Analysis of abnormal user behavior in social networks

User influence mining plays a vital role in immediately reporting noticed abnormalities to the user. This social influence of the user is assessed through two essential factors: the user's impact on others and the user's importance. The evaluation is done using a fine-grained feature-based social influence (FBI) model [20]. A page rank algorithm is used to analyze the extent of the user's influence on his or her friends. The user's impact is then assessed using affinity, either directly or indirectly. The FBI model is a scalable and generalized model for calculating influence among users in OSNs. Anomalous behavior was discovered using statistical methods to detect the intrusion of criminals in temporal networks [21]. Two datasets, from Twitter and VAST, were monitored over 10 days, during which all tweets and telephone calls, respectively, were analyzed to classify the observed behavior as normal or abnormal. The statistical method uses the binomial distribution of a discrete time model to calculate p-values and thus detect abnormality. The experimental results showed that all abnormal nodes were detected in the VAST dataset, but with a large number of false positives. Because false positives reduce the accuracy of detection, this issue must be addressed.

Structured Learning analyzes user behavior and its corresponding linkage information across different social platforms. The linking of data for the same user over different platforms results in fewer benefits with respect to data completeness, consistency, and continuity, in addition to unreliable attributes, data misalignment, platform difference, behavior asynchrony, and data imbalance. Heterogeneous behavior can be assessed against user attributes, generated content, behavior trajectory, and core OSN features. The evaluation is based on precision and recall values and total execution time. Hydra, a multi-objective learning framework across heterogeneous social platforms, is more efficient than other existing methods.

4.2. Privacy preserving profile matching schemes

Private matching protocols [22] use an asymmetric social proximity measure to identify matches between initiator and responder profiles without revealing the private information of one user to the other before confirming a friend request. There are four protocols used, namely, L1P, L2P, EL2P, and L3P. These private matching protocols analyze the overlap of similar interests between the two users based on their similar communities. All protocols are capable of providing at least some security to OSN users. For all protocols, both the computational cost and the communication cost regarding the exchange of keys are high. Trustworthiness analysis uses OSNs for cloud computing and grid computing aspects of distributed computing. A social cloud implies that the resource provider and the resource consumer are involved in a relationship with some kind of OSN, which helps to generate trust between them. This further helps with scheduling efficiency in terms of both task and trust, which improves computing efficiency in general. Such a social cloud is capable of improving computational task processing, as in grid and cloud computing.

4.3. Utilizing machine learning techniques

The previous approaches assume that the machine learning techniques are too challenging because the attackers create patterns that cannot be trained by machines. But recent works have applied many standard machine learning algorithms, such as ensemble of classifiers, Random Forests, SVM, Adaboost and Naïve Bayes, to adversarial learning. Several machine learning algorithms are utilized for the clustering and classification of profiles based on their attributes. The survey on efficient machine learning introduces several machine learning algorithms and discusses their big data processing ability with respect to prediction accuracy. Efficiency is calculated based on a model's computational requirements, least memory

Please cite this article as: D. Ramalingam, V. Chinnaiah, Fake profile detection techniques in large-scale online social networks: A comprehensive review, Computers and Electrical Engineering (2017), http://dx.doi.org/10.1016/j.compeleceng.2017.05.020

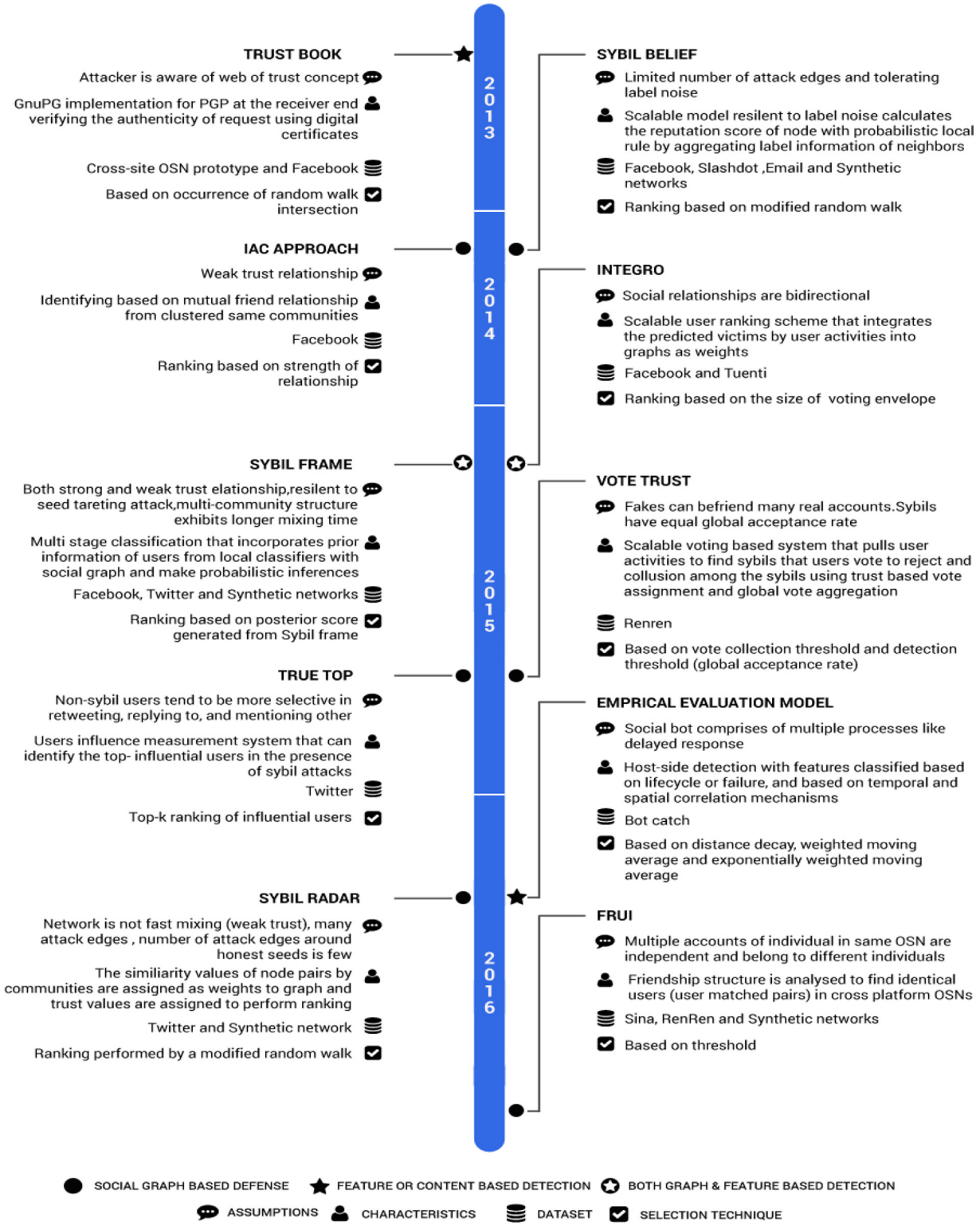

Fig. 3. Recent works on fake profile detection techniques with the corresponding detection models and their features.

Please cite this article as: D. Ramalingam, V. Chinnaiah, Fake profile detection techniques in large-scale online social networks: A comprehensive review, Computers and Electrical Engineering (2017), http://dx.doi.org/10.1016/j.compeleceng.2017.05.020

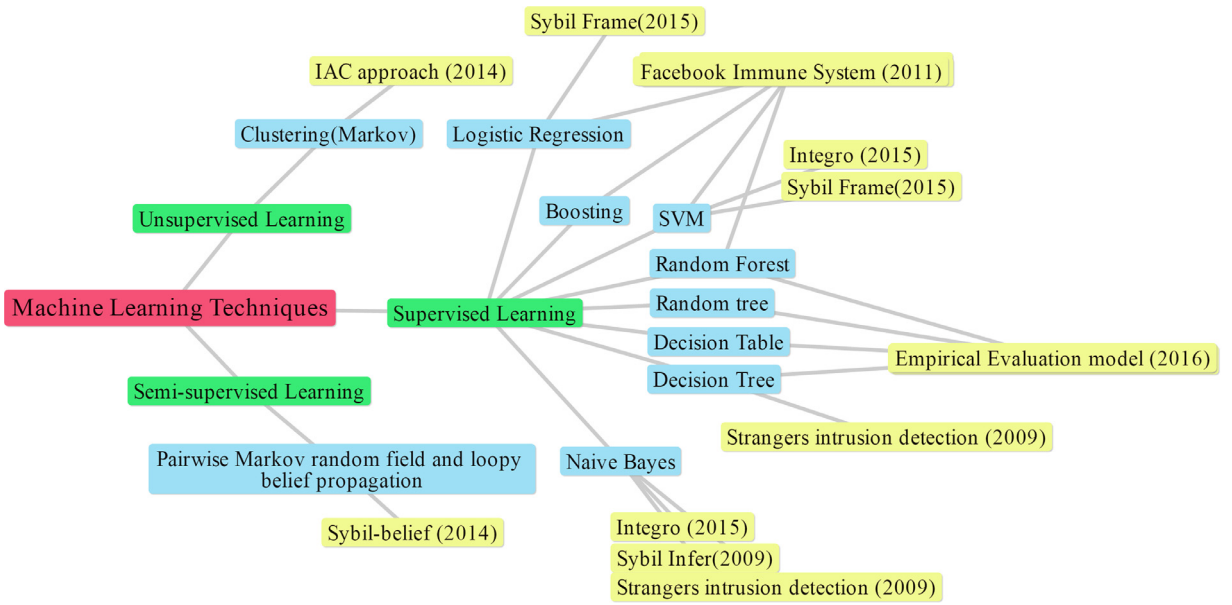

Fig. 4. Machine learning techniques used in recent works.

requirements, and ratio of computational cost to prediction accuracy. Sybil communities are detected from social network graphs using clustering algorithms. The different machine learning techniques employed by various works for their evaluation are shown in Fig. 4.

4.4. Big data in fake profile analysis

Big data concerns play a key role in the empirical analysis of data from OSNs. The analysis of big data in OSNs helps communication media in various respects. Some challenges regarding big data are explained using a few examples. Algorithms capable of processing big data are as follows: ensemble methods for improving performance by reducing computational cost, model complexity methods to optimize time complexity, local learning strategies to reduce computational complexity, semi-parametric approximations for global models, deep learning to increase chip processing capabilities and lower computing hardware cost, big data computing with Map Reduce, Graph Lab for batch processing, and Storm and SAMOA for stream processing.

Integro [23] concentrates on complete community detection solutions for large networks using Map Reduce and takes O(nlog n) time to complete its computation. Photo-based social authentication is used to automatically flag fake or suspicious accounts. A ranking system is employed such that fake accounts are given low ranks and benign accounts are given high ranks. Integro is limited to undirected graphs, and it has a delay when considering new accounts. Scalable community detection uses an attribute-based recommendation of friends for social relationships between strangers using Map Reduce technology. Shi et al. introduced DElayed Processing Of Large Degree nodes (DEPOLD) [24] for the efficient processing of OSNs with millions of nodes in the network. It makes use of social attributes to find the matched friends and a multi-hop trust chain for making friends between strangers in an OSN.DEPOLD filters out high degree nodes and applies Map Reduce-based detection algorithms on the remaining nodes. A community is detected based on similarity metrics for a number of common communities between the nodes of the OSN. Compared to the efficiency and consistency of Integro, DEPOLD saves computation time and reduces complexity.

Recent studies on memory reveal the difficulties of in-memory management for big data. In-memory data storage systems include H-Store/VoltDB, Hekaton, Hyper/ScyPer, and SAP HANA. In-memory No SQL databases include MemepiC, MongoDB, RAMCloud, and Redis. In-memory data processing includes big data analytics systems such as Main Memory Map Reduce (M3R), Spark/RDD, and real-time processing systems such as Spark Streaming and the Yahoo Simple Scalable Streaming System (S4). The challenges in data management are to determine optimization possibilities in indexing, data layouts, parallelism, and concurrency control, query processing, fault tolerance, and data overflow.

Scalable methodology, which consists of batch and stream processing, is a real-time model of scalable processing of social data for classification and prediction. Batch processing is used to train machine learning models, which is not adequate for real-time applications. Stream processing is used to test the data model and to predict the output based on the batch processing classification model. The batch processing model performs query attributes and data clustering, and it builds machine learning models. The stream processing model carries out prediction, performs classification or segmentation, and

Please cite this article as: D. Ramalingam, V. Chinnaiah, Fake profile detection techniques in large-scale online social networks: A comprehensive review, Computers and Electrical Engineering (2017), http://dx.doi.org/10.1016/j.compeleceng.2017.05.020

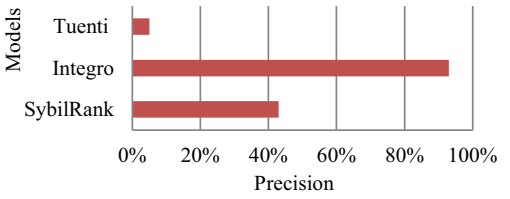

Fig. 5. Comparison of precision results among sybil detection schemes.

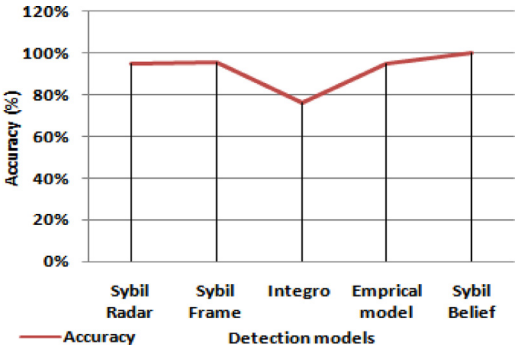

Fig. 6. Comparison of accuracy of detection models.

provides recommendations. Storage uses the Hadoop distributed file system. This prototype is used for the Web recommender system and user behavior prediction system.

5. Comparison of current techniques

Sybil defense works can be centralized or decentralized protocols. Centralized protocols are based on admission control by a central authority, which bounds the sybils and keeps the bogus accounts within a defined operating threshold. Decentralized protocols distribute the jobs to the nodes in the system. Table 2 lists the studies on protocol design of the Sybil detection model, their network type (synthetic /real) that is used (along with the size of dataset) for evaluation of the detection model, and the attacks that each type addresses in addition to the Sybil attacks. Performance metrics are given for scalability, robustness, reliability and accuracy: if the authors of the model intended to specifically address the metric, it is given a '√'.

Most of the existing works merely use graph features of networks to identify real and fake nodes and fail to adequately address scalability and user metadata analysis or employ machine learning techniques. The early bots are simple to spot by monitoring, for example, bulk content posted routinely from a node. However, modern large-scale bot infiltration poses challenges that require active machine learning algorithms. The existing approaches such as SybilGuard provide assurance on the number of sybils admitted per attack edge. If the number of attack edges p in a network with n real nodes is at most $O(\sqrt{n}/\log n)$, SybilGuard accepts at most $O(\sqrt{n}\log n)$ sybil nodes per attack edge. If p increases, SybilGuard is not able to restrict the number of Sybil nodes. SybilLimit improves on SybilGuard by allowing at most $O(\log n)$ Sybils per attack edge regardless of the number of attack edges. For example, in a million node network, it allows for ~10 to ~40 sybil nodes per attack edge. It overcomes SybilGuard's limitation that if the number of attack edges exceeds a fixed threshold, false negatives are produced. Sybil rank outperforms all existing works, but Vote trust improves over Sybil rank by 10%in false positive and false negative rates. Integro improves over Sybil rank by 30% in node ranking. This improvement refers to the probability of ranking an arbitrary honest user higher than an arbitrary corrupt node. Fig. 5 shows that Integro detects sybils with greater precision (95%) than both Tuenti's user reporting system (5%) and Sybil Rank (42%), for 20k low ranking accounts.

Sybil Frame and Integro are hybrid techniques that combine user profile features and graph properties for detection. Sybil Radar also improves over Sybil Rank and attains Sybil detection accuracy identical to that of Integro without features. Similar to previously discussed methods, Sybil Belief uses a restricted amount of attack edges for finding fake identities. For 1000 real and fake randomly selected nodes on Twitter, Sybil Frame detects 68.2% sybils using node prior classifier, which helps in reducing the false positive rate from 8.5% to 4.2% (with 51% detection rate), whereas Sybil Belief identifies all nodes as sybils. Sybil Frame also ranks 500 fake accounts, which is a 12-fold and 35-fold improvement over Sybil Belief and Sybil Rank, respectively. Fig. 6 compares the accuracy of Sybil Radar, Sybil Frame, Integro, Empirical evaluation model, Integro and Sybil Belief. Sybil Belief achieves accuracy closest to 100% when compared with other methods (each with their own datasets).

Unlike the historical random walk-based methods, Trust book [27] uses digital certificates and GNU Privacy Guard to identify trusted friend requests. True top [28] uses weighted eigenvector centrality for measuring the influence of a user in

Please cite this article as: D. Ramalingam, V. Chinnaiah, Fake profile detection techniques in large-scale online social networks: A comprehensive review, Computers and Electrical Engineering (2017), http://dx.doi.org/10.1016/j.compeleceng.2017.05.020

Table 2
Overview of the dataset and evaluation network of Sybil detection models.

Model & year	Kind of synthetic social network	Dataset statistics		Protocol design	Type of attacks addressed	Performance metrics			
		Nodes	Edges			Robustness	Scalability	Reliability	Accuracy
Sybil Guard	Kleinberg's model	1×10^6 & 1×10^4 100	Node degree of 24 Node degree of 12	De-centralized	Sybil attack, IP harvesting	√	√	–	√
Sybil Limit	Kleinberg model	**Friendster** 9.3×10^5	7.8×10^6	De-centralized	Sybil attacks	–	√	–	√
		Live Journal 9×10^5	8737,636						
		DBLP 1×10^5	625,932						
		Kleinberg 1×10^6	10,935,294						
Sybil Infer	Scale-free (or power law) topology with preferential attachment	**Scale-free** 1000	–	Centralized	Sybil attacks	√	–	√	√
		Live Journal 33,000	–						
Sum-up	Not mentioned	**YouTube** 4,46,000	3458,000	Centralized & De-centralized	Sybil attacks	–	–	√	√
		Flickr 1530,000	21,399,000						
		Synthetic 300,0000	24,248,000						
Mislove's algorithm	Barabasi–Albert (BA) random synthetic network	**Facebook regional network** 63,731	816,886	De-centralized	Community detection (Sybil)	√	–	–	√
		BA 512	Degree, m = 8						
Gate Keeper	Random graphs with average node degree of 6	**YouTube** 446,181	1728,948	De-centralized	Sybil attacks	–	√	–	√
		Digg 539,242	4035,247						
Sybil Rank	Barabasi's scale-free model	**BA** 10,000	39,399	De-centralized	Sybil attacks	√	√	–	√
		Facebook 10,000	40,013						
Sybil Belief	Not mentioned	**Facebook** 43,953	182,384	De-centralized	Sybil attacks	√	√	–	√
		Slashdot 82,168	504,230						
Integro	Small-world graph model	**Facebook –2 samples** 2991 6136	13,952 38,144	De-centralized	Sybil attacks	√	√	–	√
Sybil Frame	Preferential Attachment (PA) model	**Twitter** 20×10^6	265×10^6	De-centralized	Sybil attacks	√	√	√	√
		Facebook Ego-4039	88,234						
Vote Trust	No synthetic network.	**Renren** 200×10^3	5.01×10^6	De-centralized	Sybil attacks, spammers, collusion attack	√	√	–	–
Sybil Radar [25]	Power law model	**Twitter** 469,506	2153,427	De-centralized	Sybil attacks	√	–	–	√
Friend Relation-ship-based User Identification (FRUI) [26]	Erdos–Renyi (ER) random networks,Watts-Strogatz (WS) small-world networks and BA networks	**Sina** 1.17×10^6	1.9×10^6	De-centralized	Profile cloning, Sybil attack (fake profile), Identity theft	–	√	√	–
		Renren 5.5×10^6	14.6×10^6						
		ER and WS- 5 networks with5,000 nodes and another 5 networks with 10,000 nodes **BA** - 5 networks with 10,000 nodes and another 5 networks with 20,000 nodes							

Please cite this article as: D. Ramalingam, V. Chinnaiah, Fake profile detection techniques in large-scale online social networks: A comprehensive review, Computers and Electrical Engineering (2017), http://dx.doi.org/10.1016/j.compeleceng.2017.05.020

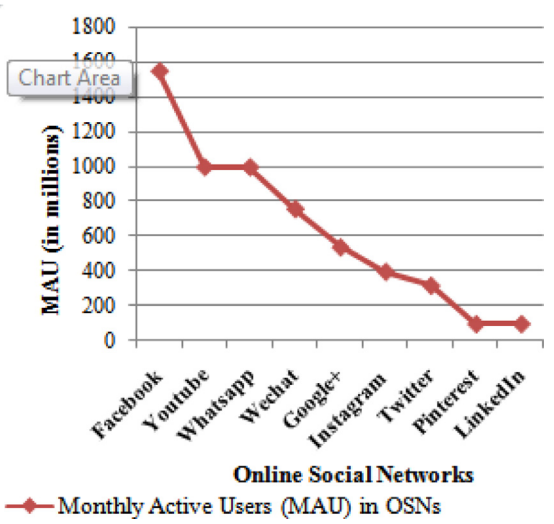

Fig. 7. Top 10 social networks based on monthly active users (in millions).

the presence of Sybil attacks on Twitter. It is accurate, ranks top-k influential user probability and bounds fake accounts. The IAC approach (Markov clustering) is used by Morteza et.al [29] to detect profile cloning where the real profiles are clustered together with similar profiles and the strength of relationship between the profiles is calculated. The empirical evaluation model [30] exploits the features in existing techniques and adds 9 new features with spatial and temporal correlations. The detection model uses Random Forest Classifier and achieves a 0.3% false positive rate and a 4.6% false negative rate with a 99.2% detection rate, which is higher than previous approaches. FRUI identifies fake identities across heterogeneous multiple social networks using mutual friend relationships. Using the sub-graphs of Sina and Renren, it calculates user matching pairs and obtains a recall rate of 0.5 with low computational cost.

6. Open issues

Despite various works in the Sybil detection domain, a number of limitations remain that could be addressed and worked on in the future.

First, it is necessary to prevent the creation of fake accounts, rather than just detecting them. Therefore, alerting the user about incoming malicious friend requests is crucial. Most of the works done in the past and in recent years only address detection and fail to address prevention. Second, little work has been done on integrating the feature-based and graph-based approaches for fake account detection, which also needs to be addressed.

Third, real-time recommendations concerning fake friend requests that utilize online machine learning algorithms need to be more focused: many works involve only offline processing of user information. Fourth, the reviewed works produced only a small difference between real and fake user. Hence, novel and efficient techniques need to be formulated. Fifth, for each of the detection techniques, namely, feature-based and graph-based, innovative features and graph metrics should perhaps be used to distinguish the real or fake user.

Finally, the collaboration and cross-platform analysis of heterogeneous social networks and big data analysis in social networks for the presence of Sybil is the focus of the present work, and very few works have focused on this. Data science is an emerging field that is expected to become the largest area of research in the next decade. Social data is the primary source of big data on the Internet. Processing and analyzing big data helps various fields. Fake profile detection can be made simple through big data analytics. The unstructured behavior data and sentiment analysis of user social behavior can be leveraged rather than simply using data with predefined labels or observing the behavior of a randomly chosen node in social networks.

Facebook can provide assured privacy levels to protect each user's personal and private information. As shown in Fig. 7, Facebook holds the highest position in number of monthly active users. Facebook has stated that approximately 46 million of its accounts are false accounts. Additionally, the number of users that are misclassified is 23 million for personal profiles for an organization, business, or non-human entity, and 14 million for undesirable accounts. This means that the number of fake accounts totals approximately 83 million. The motives of some imposters may be whimsical, but the ultimate aim behind the majority of fake profiles is malicious. A phony user may try to phish individual users into bogus relations that will lead to malware downloads, redirection of Java script code, or an increase of their audience, which leads to the potential for teens to become victims. This issue needs to be considered seriously, as OSN users, their connections, and their data sharing increase day by day.

Please cite this article as: D. Ramalingam, V. Chinnaiah, Fake profile detection techniques in large-scale online social networks: A comprehensive review, Computers and Electrical Engineering (2017), http://dx.doi.org/10.1016/j.compeleceng.2017.05.020

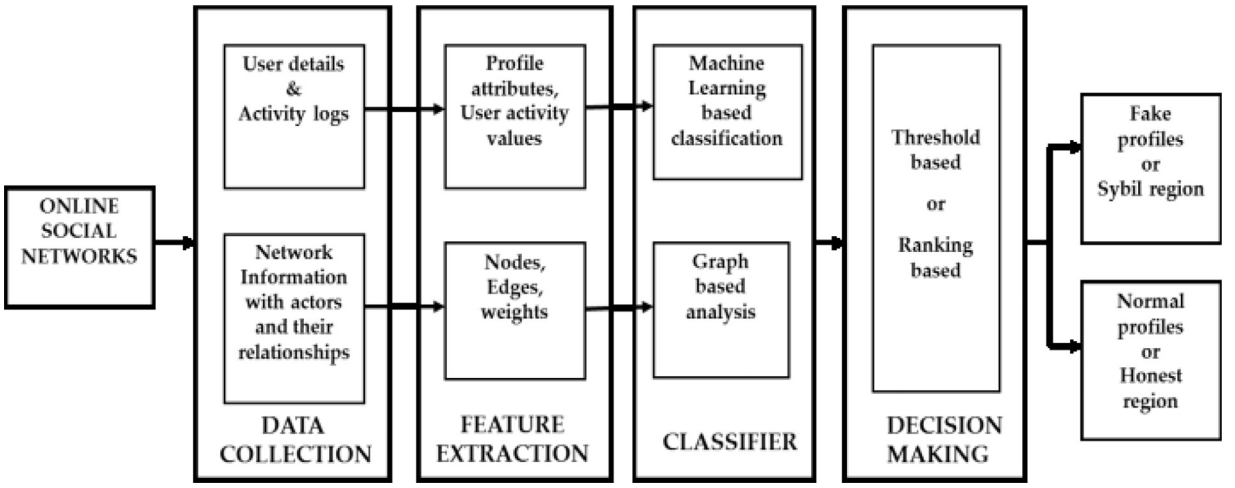

Fig. 8. Generic process flow in identification of the real or fake profile.

This research also needs to address the challenge of increasing volumes of data with high velocity and variety. Fig. 8 illustrates the generic process flow in identifying real or fake profiles. The big data generated by OSNs can be used effectively to differentiate fake and genuine profiles and subsequently recommend requests only from real profiles. This is important for non-technical users, teens and children who are unaware of privacy settings. The research should focus on helping the user to distinguish between real and fake profiles in real-time, as well as guiding them as to whether to report the profiles and warn their connected friends.

7. Conclusion

This survey provides a comprehensive review of important techniques for fake profile detection in OSNs. The paper explores the most prominent historical approaches and focuses on state of the art works for detecting Sybil or fake accounts in social networks. The various approaches, along with their synthetic network type and dataset statistics, are compared and tabulated. We also focused on recently proposed schemes and their strengths and drawbacks. These schemes are compared based on their qualitative performance. The open issues in the domain of fake profile detection in OSNs are stated. We conclude that, despite numerous existing schemes, there is still no systematic solution for fake profile detection in OSNs that can provide efficient, fast and reliable recognition of user information. Furthermore, in this paper, it is asserted that fast big data technologies such as Hadoop and Spark will definitely be part of the solution for rapidly accessing large amounts of social network data for the security analysis of user profiles. Scalable algorithms need to be designed with concurrency so that they can run on fast data systems and stream input data at line speeds rather than carrying out batch analyses.

References

[1] Fire M, Goldschmidt R, Elovici Y. Online social networks: threats and solutions. IEEE Commun Surv Tut 2014;16(4):2019–36. https://doi.org/10.1109/COMST.2014.2321628.
[2] Krombholz K, Hobel H, Huber M, Weippl E. Advanced social engineering attacks. J Inf Secur Appl 2015;22:113–22. https://doi.org/10.1016/j.jisa.2014.09.005.
[3] Nowotarski M. Don't steal my avatar! challenges of social networking patents. [accessed 12. 06. 2016] http://www.ipwatchdog.com/2011/01/23/dont-steal-my-avatar-challenges-of-social-networking-patents/id=14531/.
[4] Viswanath B, Post A, Gummadi KP, Mislove A. An analysis of social network-based sybil defenses. ACM SIGCOMM Comput Commun Rev 2010;40(4):363–74. https://doi.org/10.1145/1851275.1851226.
[5] Ferrara E, Varol O, Davis C, Menczer F, Flammini A. The rise of social bots. Commun ACM 2016;59(7):96–104. https://doi.org/10.1145/2818717.
[6] Koll D, Li J, Stein J, Fu X. On the state of OSN-based Sybil defenses. In: Proceedings of the networking conference, 2014 June 2-4; IFIP, Trondheim, Norway. IEEE; 2014. p. 1–9. https://doi.org/10.1109/IFIPNetworking.2014.6857128.
[7] Conti M, Poovendran R, Secchiero M. Fakebook: detecting fake profiles in on-line social networks. In: Proceedings of the international conference on advances in social networks analysis and mining, 2012 August 26. IEEE; 2012. p. 1071–8. https://doi.org/10.1109/ASONAM.2012.185.
[8] Fire M, Katz G, Elovici Y. Strangers intrusion detection-detecting spammers and fake profiles in social networks based on topology anomalies. Hum J 2012;1(1):26–39.
[9] Yu H, Kaminsky M, Gibbons PB, Flaxman AD. Sybilguard: defending against sybil attacks via social networks. IEEE/ACM Trans Networking 2008;16(3):576–89. https://doi.org/10.1109/TNET.2008.923723.
[10] Yu H, Gibbons PB, Kaminsky M, Xiao F. Sybillimit: a near-optimal social network defense against sybil attacks. In: IEEE symposium in security and privacy, 2008 May 18. IEEE; 2008. p. 3–17. https://doi.org/10.1109/SP.2008.13.
[11] Danezis G, Mittal P. Sybilinfer: detecting sybil nodes using social networks. In: Proceedings of NDSS, February. The Internet Society; 2009.
[12] Tran DN, Min B, Li J, Subramanian L. Sybil-resilient online content voting. In: Proceedings of the 6th USENIX symposium on networked systems design and implementation. 2009 April 22, 9. Berkeley, CA, USA: USENIX Association; 2009. p. 15–28.
[13] Viswanath B, Bashir MA, Crovella M, Guha S, Gummadi KP, Krishnamurthy B, et al. Towards detecting anomalous user behavior in online social networks. In: Proceedings of the 23rd USENIX conference on security symposium. 2014 August 20, 14. Berkeley, CA, USA: USENIX Association; 2014. p. 223–38.

Please cite this article as: D. Ramalingam, V. Chinnaiah, Fake profile detection techniques in large-scale online social networks: A comprehensive review, Computers and Electrical Engineering (2017), http://dx.doi.org/10.1016/j.compeleceng.2017.05.020

[14] Tran N, Li J, Subramanian L, Chow SS. Optimal sybil-resilient node admission control. In: Proceedings of IEEE INFOCOM. 2011 April 10. Shanghai: IEEE; 2011. p. 3218–26.
[15] Cao Q, Sirivianos M, Yang X, Pregueiro T. Aiding the detection of fake accounts in large scale social online services. In: Proceedings of the 9th USENIX conference on networked systems design and implementation, April 25. USENIX Association; 2012 15–15.
[16] Stein T, Chen E, Mangla K. Facebook immune system. In: Proceedings of the 4th workshop on social network systems, April 10-13. Salzburg, Austria: ACM; 2011. p. 1–8. https://doi.org/10.1145/1989656.1989664.
[17] Adikari S, Dutta K. Identifying fake profiles in LinkedIn. In: Proceedings of PACIS 2014; 2014 Paper 278. http://aisel.aisnet.org/pacis2014/278.
[18] Gao P, Gong NZ, Kulkarni S, Thomas K, Mittal P. Sybilframe: a defense-in-depth framework for structure-based sybil detection. 2015 March 10. arXiv preprint arXiv:1503.02985. [cs.SI]. https://arxiv.org/abs/1503.02985.
[19] Yang Z, Xue J, Yang X, Wang X, Dai Y. VoteTrust: leveraging friend invitation graph to defend against social network sybils. IEEE Trans Dependable Secure Comput 2016;13(4):488–501. https://doi.org/10.1109/TDSC.2015.2410792.
[20] Wang G, Jiang W, Wu J, Xiong Z. Fine-grained feature-based social influence evaluation in online social networks. IEEE Trans Parallel Distrib Syst 2014;25(9):2286–96. https://doi.org/10.1109/TPDS.2013.135.
[21] Vigliotti MG, Hankin C. Discovery of anomalous behaviour in temporal networks. Sci Direct 2015;41:18–25. https://doi.org/10.1016/j.socnet.2014.12.001.
[22] Thapa A, Li M, Salinas S, Li P. Asymmetric social proximity based private matching protocols for online social networks. IEEE Trans Parallel Distrib Syst 2015;26(6):1547–59. https://doi.org/10.1109/TPDS.2014.2329016.
[23] Boshmaf Y, Logothetis D, Siganos G, Lería J, Lorenzo J, Ripeanu M, et al. Íntegro: leveraging victim prediction for robust fake account detection in large scale OSNs. Comput Secur 2016;61:142–68. Elsevier https://doi.org/10.1016/j.cose.2016.05.005.
[24] Shi J, Xue W, Wang W, Zhang Y, Yang B, Li J. Scalable community detection in massive social networks using MapReduce. IBM J Res Dev 2013;57(3/4). 12-1https://doi.org/10.1147/JRD.2013.2251982.
[25] Mulamba D, Ray I, Ray I. SybilRadar: a graph-structure based framework for sybil detection in on-line social networks. In: Proceedings of IFIP international information security and privacy conference. 2016 May 30, 5. Springer International Publishing; 2016. p. 179–93. https://doi.org/10.1007/978-3-319-33630-5_13.
[26] Zhou X, Liang X, Zhang H, Ma Y. Cross-platform identification of anonymous identical users in multiple social media networks. IEEE Trans Knowl Data Eng 2016;28(2):411–24. https://doi.org/10.1109/TKDE.2015.2485222.
[27] Noor U, Anwar Z, Mehmood Y, Aslam W. TrustBook: web of trust based relationship establishment in online social networks. In: Proceedings of 11th international conference on frontiers of information technology. 2013 December 16. IEEE; 2016. p. 223–8. https://doi.org/10.1109/FIT.2013.48.
[28] Zhang J, Zhang R, Sun J, Zhang Y, Zhang C. Truetop: a sybil-resilient system for user influence measurement on twitter. IEEE/ACM Trans Networking 2016;24(5):2834–46. https://doi.org/10.1109/TNET.2015.2494059.
[29] Kharaji MY, Rizi FS. An IAC approach for detecting profile cloning in online social networks. 2014 March 8. arXiv preprint arXiv:1403.2006.
[30] Ji Y, He Y, Jiang X, Cao J, Li Q. Combating the evasion mechanisms of social bots. Comput Secur 2016;58:230–49 Elsevier.

Please cite this article as: D. Ramalingam, V. Chinnaiah, Fake profile detection techniques in large-scale online social networks: A comprehensive review, Computers and Electrical Engineering (2017), http://dx.doi.org/10.1016/j.compeleceng.2017.05.020

Devakunchari Ramalingam, M. Tech, is a Research scholar at the Department of Computer Technology, Anna University, India. Her research interests include Social Network Analysis and Big Data Analytics. She has published around 8 papers in National and International conferences and journals. She is currently pursuing her Ph.D. in Social Big Data Analytics, Anna University, India.

Valliyammai Chinnaiah M. Tech, Ph.D., is the Senior Grade Assistant Professor at the Department of Computer Technology, Anna University, India. She received her Ph.D. in computer science and engineering at Anna University. She has 15 years of teaching experience. Her areas of interest include Cloud computing, Big Data, Network management, Grid computing and Mobile agents. She has published around 40 papers in National and International conferences and journals.

Please cite this article as: D. Ramalingam, V. Chinnaiah, Fake profile detection techniques in large-scale online social networks: A comprehensive review, Computers and Electrical Engineering (2017), http://dx.doi.org/10.1016/j.compeleceng.2017.05.020