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