Online Social Networks: Threats and Solutions

Michael Fire, *Member, IEEE*, Roy Goldschmidt, and Yuval Elovici, *Member, IEEE*

***Abstract*—Many online social network (OSN) users are unaware of the numerous security risks that exist in these networks, including privacy violations, identity theft, and sexual harassment, just to name a few. According to recent studies, OSN users readily expose personal and private details about themselves, such as relationship status, date of birth, school name, email address, phone number, and even home address. This information, if put into the wrong hands, can be used to harm users both in the virtual world and in the real world. These risks become even more severe when the users are children. In this paper, we present a thorough review of the different security and privacy risks, which threaten the well-being of OSN users in general, and children in particular. In addition, we present an overview of existing solutions that can provide better protection, security, and privacy for OSN users. We also offer simple-to-implement recommendations for OSN users, which can improve their security and privacy when using these platforms. Furthermore, we suggest future research directions.**

***Index Terms*—Online social networks, security and privacy, online social network security threats, online social network security solutions.**

I. Introduction

In recent years, global online social network (OSN) usage has increased sharply as these networks have become interwoven into people's everyday lives as virtual meeting places that facilitate communication. OSNs, such as Facebook [1], Google+ [2], LinkedIn [3], Sina Weibo [4], Twitter [5], Tumblr [6], and VKontakte (VK) [7] have hundreds of millions of daily active users (see Fig. 1). Facebook, for example, has more than 1.23 billion monthly active users, 945 million of which are active mobile Facebook users as of December 2013 [8].

Facebook users have a total of over 150 billion friend connections and upload on average more than 350 million photos to Facebook each day [11]. Unfortunately, many OSN users are unaware of the security risks which exist in these types of communications, including privacy risks [12], [13], identity theft [14], malware [15], fake profiles (also in some cases referred to as sybils [16], [17] or socialbots [12], [18], [19]), and sexual harassment [20], [21], among others. A study by Dwyer *et al.* [22] found that Facebook and MySpace [23] users trust these OSNs, and they have trust in other users within these social networks. This trust leads to information sharing and to developing new relationships. Moreover, according to recent studies [12], [24], many OSN users expose personal and intimate details about themselves, their friends, and their relationships, whether by posting photos or by directly providing information such as a home address and a phone number. Furthermore, according to Boshmaf *et al.* [12] and Elyashar *et al.* [19], [25], Facebook users have been shown to accept friendship requests from people whom they do not know but with whom they simply have several friends in common. By accepting these friend requests, users unknowingly disclose their private information to total strangers. This information could be used maliciously, harming users both in the virtual and in the real world. These risks escalate when the users are young children or teenagers who are by nature more exposed and vulnerable than adults.

Fig. 1. **Word Cloud of OSNs with More Than 100 Million Active Users**. This word cloud was constructed using Wordle [9] where the font size of each OSN name is relative to the network's number of active users [10].

As the use of OSNs becomes progressively more embedded in users' daily lives, personal information becomes easily exposed and abused. Information harvesting, by both the OSN operator itself and by third-party commercial companies, has recently been identified as a significant security concern for OSN users. Companies can exploit the harvested personal information for a variety of purposes, all of which can jeopardize a user's privacy. For example, companies can use collected private information to tailor online ads according to a user's profile [26], to gain profitable insights about their customers, or even to share the user's private and personal data with the government [27]. This information may include general data, such as age, gender, and income; however, in some cases more delicate and potentially harmful information can be exposed, such as the user's sexual orientation [28] and if the user has consumed addictive substances [29]. These privacy concerns become more alarming when considering the nature of OSNs: information regarding a network user can be obtained without even directly accessing the individual's online profile; personal details can be inferred solely by collecting data on the user's friends [13].

To cope with the above-mentioned threats, multiple solutions have been offered by OSN operators, security companies, and academic researchers. OSNs, like Facebook, attempt to protect

Manuscript received September 25, 2013; revised March 13, 2014; accepted April 22, 2014. Date of publication May 2, 2014; date of current version November 18, 2014. The associate editor coordinating the review of this paper and approving it for publication was E. Hossain.

M. Fire and Y. Elovici are with the Department of Information Systems Engineering and the Telekom Innovation Laboratories, Ben-Gurion University of the Negev, Be'er Sheva 84105, Israel (e-mail: mickyfi@bgu.ac.il; elovici@bgu.ac.il).

R. Goldschmidt is with The Knesset Research and Information Center, Kiryat Ben Gurion, Jerusalem 91950, Israel (e-mail: rgoldshmidt@knesset.gov.il).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/COMST.2014.2321628

their users by adding authentication processes to ensure that the registered user is a real person [12], [30]–[33]. Moreover, many OSN operators also support a configurable user privacy setting that enables users to protect their personal data from other users within the network [34], [35]. As for privacy settings, OSN operators currently face a conflict of interest: On the one hand, since personal information is a commodity, the more that is shared, the better. On the other hand, a user who is anxious about his or her privacy is a liability and will probably share less information and become consequently less active. Nevertheless, both regulating authorities and public groups try to address privacy concerns and make them a part of public discourse and consideration [36]. Today there are additional protection mechanisms which include defenses against spammers [37]–[43], fake profiles [16], [44]–[52], and other threats. For example, security companies like Check Point [53], Websense [54], and Infoglide [55] offer social tools to protect users in the OSN world. These companies typically offer products which monitor user activity in order to identify and protect users. The modern day threats are so pervasive that even the academic community has addressed this issue by publishing studies which attempt to solve different OSN threats and offer improvements in identity protection [40], [41], [43], [45], [47], [56].

A. Contributions

This paper presents the "big picture" of the current state-of-the-art academic and industry solutions that can protect OSN users from various security and privacy threats. More specifically, this study offers the following contributions: First, we outline the OSN threats that target every user of social networks, with an additional focus on young children and teenagers. Second, we present a thorough overview of the existing solutions to these threats, namely those provided by OSN operators, commercial companies, and academic researchers. Third, we compare and discuss the protection ability of the various solutions. Lastly, we give easy-to-implement recommendations on how OSN users can better protect their security and privacy when using social networks.

B. Organization

The remainder of the paper is organized as follows: In Section II, we introduce insightful statistics on OSNs usage. Next, in Section III, we describe different types of OSN threats. Section IV follows with various solutions to assist in protecting social network users. In Section V, we discuss the various presented threats and their corresponding solutions. In Section VI, we offer recommendations that OSN users can apply in order to improve their online security and privacy. Next, in Section VII, we offer future research directions. Our conclusions are presented in Section VIII.

II. Online Social Network Usage

Today many OSNs have tens of millions of registered users. Facebook, with more than a billion active users, is currently the largest and most popular OSN in the world [57]. Other well-known OSNs are Google+, with over 235 million active users [58]; Twitter, with over 200 million active users [59]; and LinkedIn, with more than 160 million active users [60]. While some experts insist that OSNs are a passing fashion and will eventually be replaced by another Internet fad, current user statistics concur that OSNs are here to stay. A recent survey by the Pew Research Center's Internet and American Life Project [61] revealed that 72% of online American adults use social networking sites, a dramatic increase from the 2005 Pew survey which discovered that just 8% of online American adults used social networking sites. Moreover, the survey revealed that 89% of online American adults between the ages of 18 to 29 use social network sites, while in 2005 only 9% of the survey participants in this age group used this type of site. These survey results are compatible with a previous report published by Nielsen in 2011 [62], disclosing that Americans spent 22.5% of their online time on OSNs and blogs, more than twice the time spent on online games (9.8%). Other common activities that consume Americans' online time include email (7.6%), portals (4.5%), videos and movies (4.4%), searches (4.0%), and instant messaging (3.3%). The amount of collective time spent on OSNs, especially on Facebook, is enormous and ever-growing. U.S. users spent a total of 53.5 billion minutes on Facebook during May 2011, 17.2 billion minutes on Yahoo [63], and 12.5 billion minutes on Google [64].

Mobile devices, or cellular phones, increasingly serve as platforms for Internet usage. According to Facebook's report [8] in December 2013, Facebook had 556 million daily active mobile users, an increase of 49% year over year. Additionally, Facebook and Google+ mobile applications are the second and fourth (respectively) most frequently used smartphone applications [65]. It should be noted that the use of OSNs on mobile devices not only promotes an even "closer relationship" to social networks but also can pose additional privacy concerns, especially regarding the collection of location data and the opportunity for advertisers to identify specific types of users.

Besides being popular among adults, OSNs have become extremely popular with young children and teenagers. A comprehensive study [66] carried out in 25 European countries with 25 000 participants produced the following statistics: 60% of children 9 to 16 years old who access the Internet use it daily (88 minutes of use on average) and 59% of those 9 to 16 years old who use the Internet have a personal OSN site profile (26% of ages 9 to 10; 49% of ages 11 to 12; 73% of ages 13 to 14; 82% of ages 15 to 16). Note that the terms of use governing OSNs do not officially allow users under the age of 13. Furthermore, 26% of the children in this same European study had their social network profile set to "public" (i.e., accessible to strangers), 14% reported having their address or phone number listed on their profile, and 16% admitted that their profile displayed an inaccurate age. In addition, 30% of the children surveyed reported having an online connection with a person they had never met face to face, 9% reported having actually met face to face with someone with whom they had only an online connection, 9% reported experiencing a misuse of personal data, 21% reported encountering one or more types of potentially harmful user-generated content, and 6% reported receiving malicious or hurtful messages on the Internet [66].

These findings reiterate our previous claim: the use of OSNs is embedded in the everyday lives of young children and teenagers, and can result in personal information being exposed, misused, and potentially abused. Interestingly, about a third of the parents in this European study claimed that they filter their children's use of the Internet, while a quarter specifically stated that they use monitoring tools [66].

III. Threats

With the increasing usage of OSNs, many users have unknowingly become exposed to threats both to their privacy and to their security. These threats can be divided into four main categories. The first category contains *classic threats*, namely, privacy and security threats that not only jeopardize OSN users but also Internet users not using social networks (see Section III-A). The second category covers *modern threats*, that is, threats that are mostly unique to the environment of OSNs and which use the OSN infrastructure to endanger user privacy and security (see Section III-B). The third category consists of *combination threats*, where we describe how today's attackers can, and often do, combine various types of attacks in order to create more sophisticated and lethal attacks (see Section III-C). The fourth and last category includes threats specifically targeting children who use social networks (see Section III-D).

Fig. 2 diagrams all the specific threats listed in the following sections. The boundaries between all these categories of threats, however, can become blurred as techniques and targets often overlap.

A. Classic Threats

Classic threats have been a problem ever since the Internet gained widespread usage. Often referred to as malware, spam, cross-site scripting (XSS) attacks, or phishing, they continue to be an ongoing issue. Though these threats have been addressed in the past, they have become increasingly viral due to the structure and nature of OSNs and can spread quickly among network users. Classic threats can take advantage of a user's personal information published in a social network to attack not only the user but also their friends simply by adjusting the threat to accommodate the user's personal information.

For example, an attacker can plant a malicious code inside an attractive spam message that employs a user's details from his or her Facebook profile. Due to the personal nature of this crafted message, the chances that the innocent user will open the message and get infected are likely. In many cases, these threats target essential and everyday user resources such as credit card numbers, account passwords, computing power, and even computer bandwidth (in order to send spam emails). Alarmingly, these types of threats can also exploit the infected user's stolen credentials to post messages on the user's behalf or even change the user's personal information.

The different classic threats are described below, along with real-life scenarios where these types of menaces have jeopardized a real user's privacy and security.

Malware: Malware is malicious software developed to disrupt a computer operation in order to collect a user's credentials and gain access to his or her private information. Malware in social networks uses the OSN structure to propagate itself among users and their friends in the network. In some cases, the malware can use the obtained credentials to impersonatethe user and send contagious messages to the user's online friends. Koobface was the first malware to successfully propagate through OSNs such as Facebook, MySpace, and Twitter. Upon infection, Koobface attempts to collect login information and join the infected computer in order to be part of a botnet [15], a so-called "zombie army" of computers which often is then used for criminal activities, such as sending spam messages and attacking other computers and servers over the Internet.

Phishing Attacks: Phishing attacks are a form of social engineering to acquire user-sensitive and private information by impersonating a trustworthy third party. A recent study [67] showed that users who interact on social networking websites are more likely to fall for phishing scams due to their social and trusting nature. Moreover, in recent years, phishing attempts within OSNs have increased sharply. According to the Microsoft Security Intelligence Report [68], 84.5% of all phishing attacks target social network site users. One such phishing attack occurred on Facebook, luring users onto fake Facebook login pages. Then, the phishing attack spread among Facebook users by inviting friends to click on a link posted on the original user's profile space [69]. Fortunately, Facebook acted to stop this attack.

Spammers: Spammers are users who use electronic messaging systems in order to send unwanted messages, like advertisements, to other users. OSN spammers use the social networking platform to send advertisement messages to other users by creating fake profiles [47]. The spammers can also use the OSN platform to add comment messages to pages which are viewed by many users in the network. An example of the prevalence of network spamming can be found on Twitter, which has suffered from a massive amount of spam. In August 2009, 11% of Twitter messages were spam messages. However, by the beginning of 2010, Twitter had successfully cut down the percentage of spam messages to 1% [70]. Nevertheless, a 2013 article [71] states, "Social spam, as it already exists on Twitter, will continue to grow and unless the company addresses the problem quickly, it may be the one thing that sinks it."

Cross-Site Scripting (XSS): An XSS attack is an assault against web applications. The attacker who uses the XSS exploits the trust of the web client in the web application and causes the web client to run malicious code capable of collecting sensitive information. OSNs, which are types of applications, can suffer from XSS attacks. Furthermore, attackers can use an XSS vulnerability combined with the OSN infrastructure to create an XSS worm that can spread virally among social network users [72]. In April 2009, such an XSS worm, called Mikeyy, rapidly transmitted automated tweets across Twitter and infected many users, among them celebrities like Oprah Winfrey and Ashton Kutcher. The Mikeyy worm used an XSS weakness and the Twitter network structure to spread through Twitter user profiles [73].

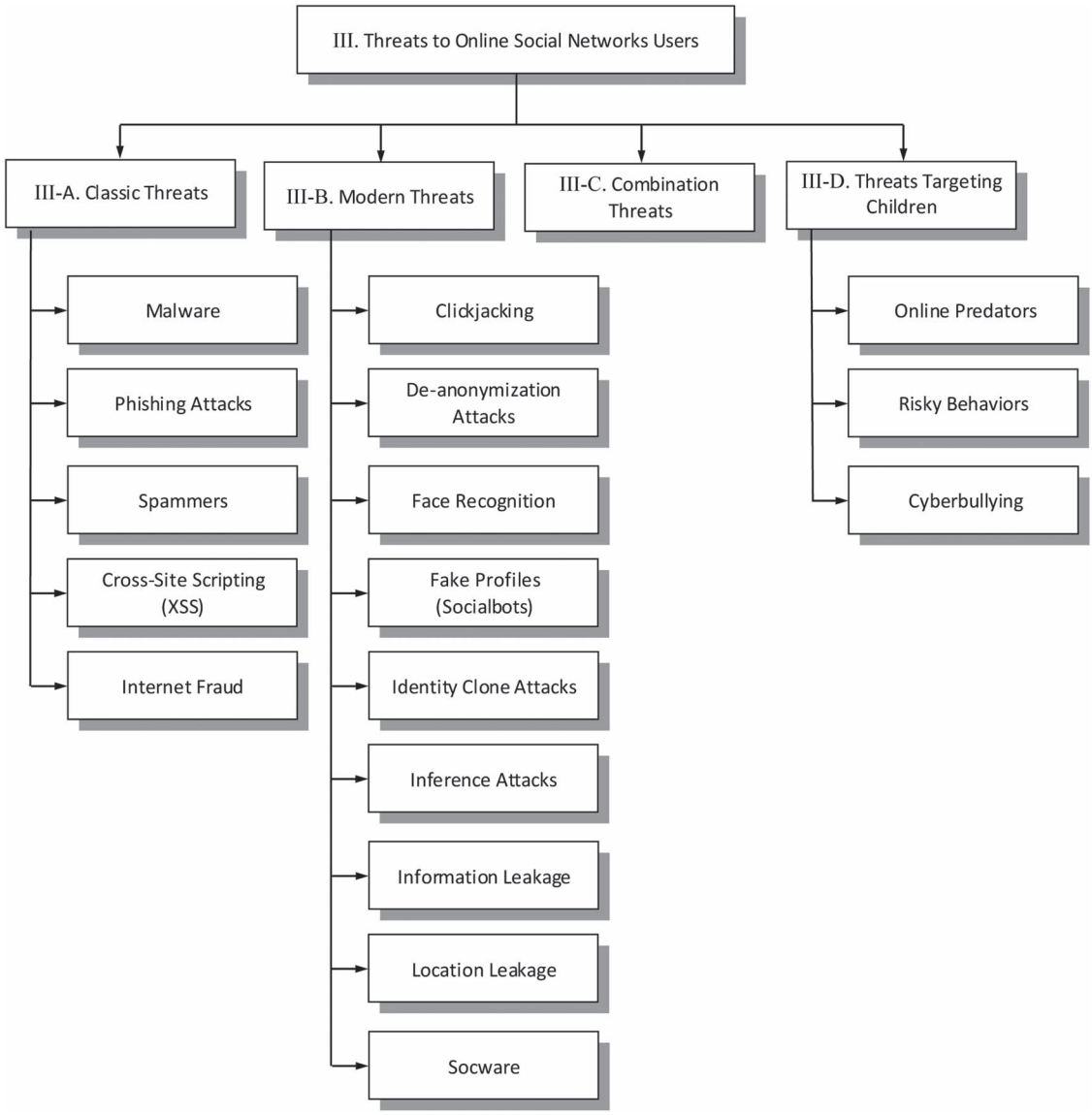

Fig. 2. Threats to online social network users.

Internet Fraud: Internet fraud, also known as cyber fraud, refers to using Internet access to scam or take advantage of people. In the past, con artists used traditional in-person social networks, such as weekly group meetings, to gradually establish strong bonds with their potential victims. Currently, according to the North American Securities Administrators Association (NASAA) [74], with the rising popularity of online networking, con artists have turned to OSNs to establish trust connections with their victims, and then they take advantage of personal data published in the victims' online profiles. In recent years, for example, fraudsters have been hacking into the accounts of Facebook users who travel abroad. Once they manage to log into a user's account, the scammers cunningly ask the user's friends for assistance in transferring money to the scammer's bank account. One victim of this type of fraud was Abigail Pickett. While traveling in Colombia, Abigail discovered that her Facebook account had been hijacked by someone in Nigeria, and it was being used to send requests for money to her network friends on the pretext that she was "stranded" [75].

B. Modern Threats

Modern threats are typically unique to OSN environments. Usually these threats specifically target users' personal information as well as the personal information of their friends. For example, an attacker who is trying to gain access to a Facebook user's high school name—viewable only by the user's Facebook friends—can create a fake profile with pertinent details and initiate a friend request to the targeted user. If the user accepts the friend request, his or her details will be exposed to the attacker. Alternatively, the attacker can collect data from the user's Facebook friends and employ an inference attack to infer the high school name from the data collected from the user's friends.

In what follows, we illustrate the various modern threats and real-life scenarios where these types of threats have jeopardized an OSN user's privacy and security.

Clickjacking: Clickjacking is a malicious technique which tricks users into clicking on something different from what they intended to click. By using clickjacking, the attacker can

manipulate the user into posting spam messages on his or her Facebook timeline, performing "likes" to links unknowingly (also referred as likejacking), and even opening a microphone and web camera to record the user [76]. An example of a clickjacking attack occurred on Twitter in 2009 when Twitter was plagued by a "Don't Click" attack. The attacker tweeted a link with the message "Don't Click" along with a masked URL (the actual URL domain was hidden). When Twitter users clicked on the "Don't Click" message, the message automatically spread virally and was posted onto their Twitter accounts [77].

De-Anonymization Attacks: In many OSNs like Twitter and MySpace, users can protect their privacy and anonymity by using pseudonyms. De-anonymization attacks use techniques such as tracking cookies, network topology, and user group memberships to uncover the user's real identity. An example of de-anonymization was demonstrated by Krishnamurthy and Wills [78], who proved that it is possible for third parties to uncover OSN user identities by linking information leaked via social networking sites. Krishnamurthy and Wills also showed that most users on the studied OSNs were vulnerable to having their OSN identity information leaked via tracking mechanisms, such as tracking cookies. Another example of this type of attack was presented by Wondracek *et al.* [79]; they offered a method to de-anonymize users in OSNs by using only the users' group memberships. Wondracek *et al.* tested their method on the Xing [80] OSN and succeeded in identifying 42% of the users. An additional recent example was presented by Peled *et al.* [81], who introduced a method for matching user profiles across several OSNs. The method was evaluated by matching profiles across Facebook and Xing.

Face Recognition: Many people use OSNs for uploading pictures of themselves and their friends. Millions and millions of photos are uploaded to Facebook each day [11]. Moreover, many Facebook user profile pictures are publicly available to view and download. For instance, the Faces of Facebook website [82] allows Internet users to view the profile images of over 1.2 billion Facebook users. These photos can be used to create a biometric database, which can then be used to identify OSN users without their consent.

In 2011, Acquisti *et al.* [83] demonstrated the threat of face recognition to OSN user privacy by performing three experiments. The first experiment showed that it is possible to match "online to online" image datasets by using publicly accessible Facebook user profile pictures to re-identify profiles on one of the most popular dating sites in the United States. In their second experiment, Acquisti *et al.* demonstrated that "offline to online" image datasets can also be matched. Namely, they used publicly available images from Facebook to identify students strolling through campus. In their third experiment, Acquisti *et al.* illustrated that it is possible to predict personal and sensitive information from a face; an individual's interests, activities, and even his or her social security number could be automatically predicted by matching the face image with the person's Facebook image to obtain the person's full name. Following this action, the attacker could use the obtained name to cross-reference it against other datasets.

Fake Profiles: Fake profiles (also referred to as sybils or socialbots) are automatic or semi-automatic profiles that mimic human behaviors in OSNs. In many cases, fake profiles can be used to harvest users' personal data from social networks. By initiating friend requests to other users in the OSN, who often accept the requests, the socialbots can gather a user's private data which should be exposed only to the user's friends. Moreover, fake profiles can be used to initiate sybil attacks [84], publish spam messages [85], or even manipulate OSN statistics [17], [86]. A recent article asserted that the market of buying fake followers and fake retweets is already a multimillion-dollar business [87]. Additional approaches that generate fake profiles were demonstrated recently by Boshmaf *et al.* [12] when an army of more than a hundred Facebook socialbots was created, which then attempted to infiltrate innocent Facebook profiles by initiating a series of friend requests. The socialbot army succeeded in generating approximately 250 GB of inbound Facebook traffic. Moreover, the socialbot friend acceptance rates climbed to 80% whenever a socialbot and an innocent Facebook user had more than eleven friends in common. In some cases, even one well-manipulated fake profile can cause extensive damage as proven by Thomas Ryan, who assumed the fictional profile of Robin Sage to connect to hundreds of users from various social networking sites [88].

Identity Clone Attacks: Using this technique, attackers duplicate a user's online presence either in the same network, or across different networks, to deceive the cloned user's friends into forming a trusting relationship with the cloned profile. The attacker can use this trust to collect personal information about the user's friends or to perform various types of online fraud. An example of an identity clone attack occurred recently with NATO's most senior commander, Admiral James Stavridis. His profile details were cloned and then used to collect data on defense ministry officials and other government officials by tricking them into becoming friends with the newly cloned Facebook profile [89].

Inference Attacks: Inference attacks in OSNs are used to predict a user's personal, sensitive information that the user has not chosen to disclose, such as religious affiliation or sexual orientation. These types of attacks can be implemented using data mining techniques combined with publicly available OSN data, such as network topology and data from users' friends. An inference attack was demonstrated by Mislove *et al.* [13] who presented techniques for predicting a user's attributes based on other users' attributes in the OSN. They tested their techniques and inferred different Facebook users' attributes, such as educational information, personal preferences, and geographic information. Recently, inference attacks on organizations were explored by Fire *et al.* [90]. They presented an algorithm for inferring the OSN of a targeted organization based solely on publicly available data from social networks. Fire *et al.* tested their algorithm on six organizations of different scales using publicly available data from the Facebook profiles of the organization's employees, resulting in a successful reconstruction of the social networks within these six organizations. Additionally, certain details could be inferred about the targeted organizations, some of which were confidential.

Information Leakage: OSNs allow users to openly share and exchange information with their friends and other users in the

network. In some cases OSN users willingly share sensitive information about themselves and other people, such as health-related information [91], [92] and sobriety status [91]. In a recent study, Torabi and Beznosov [92] observed that 95.8% of 166 participants shared some health-related information through their OSN accounts. Leakage of sensitive and personal information may have negative implications for the social networks users. For example, insurance companies may use OSN data to identify risky clients [93]. These companies can use OSN leaked information to detect clients with medical conditions, consequently increasing their premiums or denying their coverage. Additionally, employers use social networks for screening job applicants [94]. Therefore, leaking personal information, such as drinking habits, on OSNs may jeopardize future chances for finding employment.

Location Leakage[1]*:* With the increasing use of smart mobile devices that encourage sharing of location information [95], many people use OSNs to willingly share private and sometimes sensitive information about their (or their friends') current or future whereabouts. A study by Humphreys *et al.* [96] found that 20.1% of examined Twitter tweets included information on when people were engaging in certain activities, and 12.1% of the tweets mentioned the person's location. Additionally, a study by Mao *et al.* [91] demonstrated that classifiers can be trained to identify Twitter users' locations in real time. Moreover, Cheng *et al.* [97] presented a framework for estimating a user's city-level location based on the content of the user's tweets. This type of information can be used by criminals and stalkers. For example, Israel Hyman from Arizona tweeted that he was looking forward to his family vacation to St. Louis. He also tweeted again once he had arrived in Missouri. When Hyman returned home, he discovered that his house had been burglarized [96]. An even more disturbing example of location leakage threats is given by the website Pleaserobme.com [98], [99], which shows a way to find the location information of specific Twitter and Foursquare [100] users.

In some cases, OSN users unknowingly share their locations by uploading media items, such as photos and videos, which may be embedded with geotagging information about their current and past locations [101]. For example, Adam Savage, the host of the popular science program *MythBusters*, posted a picture on Twitter of his car parked in front of his house. The uploaded image contained a geotag which exposed the place where the photo was taken [102].

Socware: Socware entails fake and possibly damaging posts and messages from friends in OSNs. Socware may lure victims by offering false rewards to users who install socware-related malicious Facebook applications or visit questionable socware websites. After the users have cruised the socware website or installed the relevant application, the installed socware sends messages on the user's behalf to the user's friends, essentially assisting the socware viral spread [103]. In 2012, Rahman *et al.* [103] investigated over 40 million posts and discovered that 49% of the studied users were exposed to at least one socware post in a four-month period. Moreover, Rahman *et al.* [104] discovered that 13% of 111 000 studied applications were malicious applications that could assist in spreading socware. Additionally, a recent study by Huang *et al.* [105] studied the ecosystem which enables socware to propagate (cascade). By analyzing data from the profile pages of approximately 3 million Facebook users over a period of five months, they discovered that "socware cascades are supported by Facebook applications that are strategically collaborating with each other in large groups."

[1]Location leakage is a private case of information leakage, which was discussed in the previous paragraph. However, due to serious privacy threats that could occur as a result of location leakage, such as location monitoring and stalking, we present this threat in a separate subsection.

C. Combination Threats

Today's attackers can also combine classic and modern threats in order to create a more sophisticated attack. For example, an attacker can use a phishing attack to collect a targeted user's Facebook password and then post a message containing a clickjacking attack on the targeted user's timeline, thus luring the user's Facebook friends to click on the posted message and install a hidden virus onto their own computers. Another example is the use of cloned profiles to collect personal information about friends of the cloned user. Using the friends' personal information, the attacker can send uniquely tailored spam email messages containing a virus. By using personal information, the virus is more likely to be activated.

Note that the recovery processes from classic and modern threats are distinct. In order to recover from a classic attack, like a virus, it is usually possible to simply reinstall the operating system, change the current passwords, and cancel the affected credit cards. However, in order to recover from a modern OSN attack that "steals your reality" [106], more effort must be made because resetting personal information is excessively time consuming and not always possible. For instance, you could change your email address, but it would be much more difficult to change your home address.

D. Threats Targeting Children

Children, whether young children or teenagers, certainly experience the classic and modern threats detailed above, but there are also threats that intentionally and specifically target younger users of OSNs. Due to the critical nature of this topic, this section highlights those threats, as well as describes specific findings from current studies.

Online Predators: The greatest concern regarding the personal information safety of children relates to Internet pedophiles, also referred to as online predators. Livingstone and Haddon [107] of EU Kids Online defined a typology in order to understand the risk and harm related to the following online activities: harm from *content* (a child's exposure to pornography or harmful sexual content), harm from *contact* (a child who is contacted by an adult or another child for the purpose of sexual abuse), and harm from *conduct* (the child as an active initiator of abusive or risky behaviors). Behaviors that are considered to be Internet sexual exploitation of children include adults using children for the production of child pornography and its

distribution, consumption of child porn, and the use of the Internet as a means to initiate online or offline sexual exploitation. In their study from 2008, Wolak *et al.* [20] critically examined the myth and reality of the online predator. The image of an Internet predator in the media is that of an adult man who pretends to be a friend to an innocent young boy or girl through whom he collects personal data; he hides his sexual intentions until the actual meeting, which likely involves rape or kidnapping. According to Wolak *et al.*, however, the truth is far more complex. Wolak *et al.* assert that most Internet-initiated sex crimes indeed start with establishing a relationship between an adult and a child through the use of instant messaging, emails, chats, etc. However, in most cases children are aware of the fact that they are talking to an adult, and if the relationship escalates to attending a real-life meeting, they are aware and to some extent expect to engage in sexual activity. More often than not, the encounter involves non-forcible sexual activity, yet it is with a person under the age of consent and therefore constitutes a crime. Contrary to the common notion, Wolak *et al.* discovered that most victims of Internet-initiated sex crimes were teenagers (aged 13 to 17), and none under age 12 were reported [20]. Therefore, these crimes do not constitute the clinical definition of pedophilia: "the fantasy or act of sexual activity with prepubescent children" [108]. Of course, this does not make the crimes any less distasteful.

Risky Behaviors: Potential risky behaviors of children may include direct online communication with strangers, use of chat rooms for interactions with strangers, sexually explicit talk with strangers, and giving private information and photos to strangers. It should be noted that while each of the above-mentioned behaviors alone poses a risk, the combination of a few of these behaviors can justifiably cause enormous anxiety regarding a child's safety. Wolak *et al.* [20] maintain that risky online behaviors and specific populations who are more exposed to them can be identified. Additionally, there is a well-established link between online and offline behaviors. Researchers contend that victims of Internet abuse are very often vulnerable children, such as youths with a history of physical or sexual abuse or those who suffer from depression or social interaction problems [20]. All children living with these kinds of issues are at a higher risk of sexual abuse on the Internet or through online-initiated encounters [20].

Cyberbullying: Cyberbullying (also referred to as cyber abuse) is bullying that takes place within technological communication platforms, such as emails, chats, phones conversations, and OSNs, by an attacker who uses the platform to harass his victim by sending repeated hurtful messages, sexual remarks, or threats; by publishing embarrassing pictures or videos of the victim; or by engaging in other inappropriate behavior. Today, cyberbullying has become a common phenomenon in OSNs in which the attacker can utilize the network's infrastructure to spread cruel rumors about the victim and share embarrassing pictures with the victim's network of friends [109]. Cyberbullying usually affects children, rather than adults. A recent online survey, which included 18 687 parents from 24 countries, revealed that 12% of parents claim their child has been cyberbullied [110]. Additionally, according to the survey's results, the majority of children experienced this harassing behavior on widely used social networking sites like Facebook. Horrifically, in some cases cyberbullying can cause catastrophic results, as in the cases of Amanda Michelle Todd [109] and Rebecca Ann Sedwick [111], both of whom committed suicide after being cyberbullied on Facebook.

IV. Solutions

In recent years, social network operators, security companies, and academic researchers have tried to deal with the above-mentioned threats by proposing a variety of solutions (see Fig. 3 and Table II). In this section we describe possible solutions which can assist in protecting the security and privacy of OSN users.

A. Social Network Operator Solutions

OSN operators attempt to protect their users by activating safety measures, such as employing user authentication mechanisms and applying user privacy settings. Several of these techniques are described in detail below.

Authentication Mechanisms: In order to make sure the user registering or logging into the social network is a real person and not a socialbot or a compromised user account, OSN operators use authentication mechanisms, such as CAPTCHA [12], photos-of-friends identification [31], multi-factor authentication [33], and in some cases even requesting that the user send a copy of his or her government issued ID [30]. As an example, Twitter recently introduced its two-factor authentication mechanism [32], requiring the user to not only insert a password when logging into Twitter but also provide a verification code that was sent to the user's mobile device.

This mechanism prevents a malicious user from logging in through hijacked accounts and publishing false information through those hijacked accounts. Such a mechanism would thwart incidents such as when hackers hijacked the Associated Press (AP) Twitter account, resulting in the rapid propagation of false information about explosions in the White House, which caused panic on Wall Street [112].

Security and Privacy Settings: Many OSNs support various configurable user privacy settings that enable users to protect their personal data from other users or applications [34], [35]. Facebook users, for example, can customize their privacy settings and choose which other users in the network (such as Friends, Friends of Friends, and Everyone) are able to view their details, pictures, posts, and other personal information [113]. A similar example of customizable privacy settings exists in Google+: users place each one of their friends into groups, also known as circles, such as Best Friends circle, Work circle, and High School Friends circle. Using these circles, Google+ users can better protect their privacy by deliberately choosing which of their posts are exposed to each circle [114]. Moreover, both Facebook and Google+ enable their users to approve or revoke the access of applications to the users' personal data [115], [116].

Some OSNs also support extra security configurations which enable the user to activate secure browsing, receive login

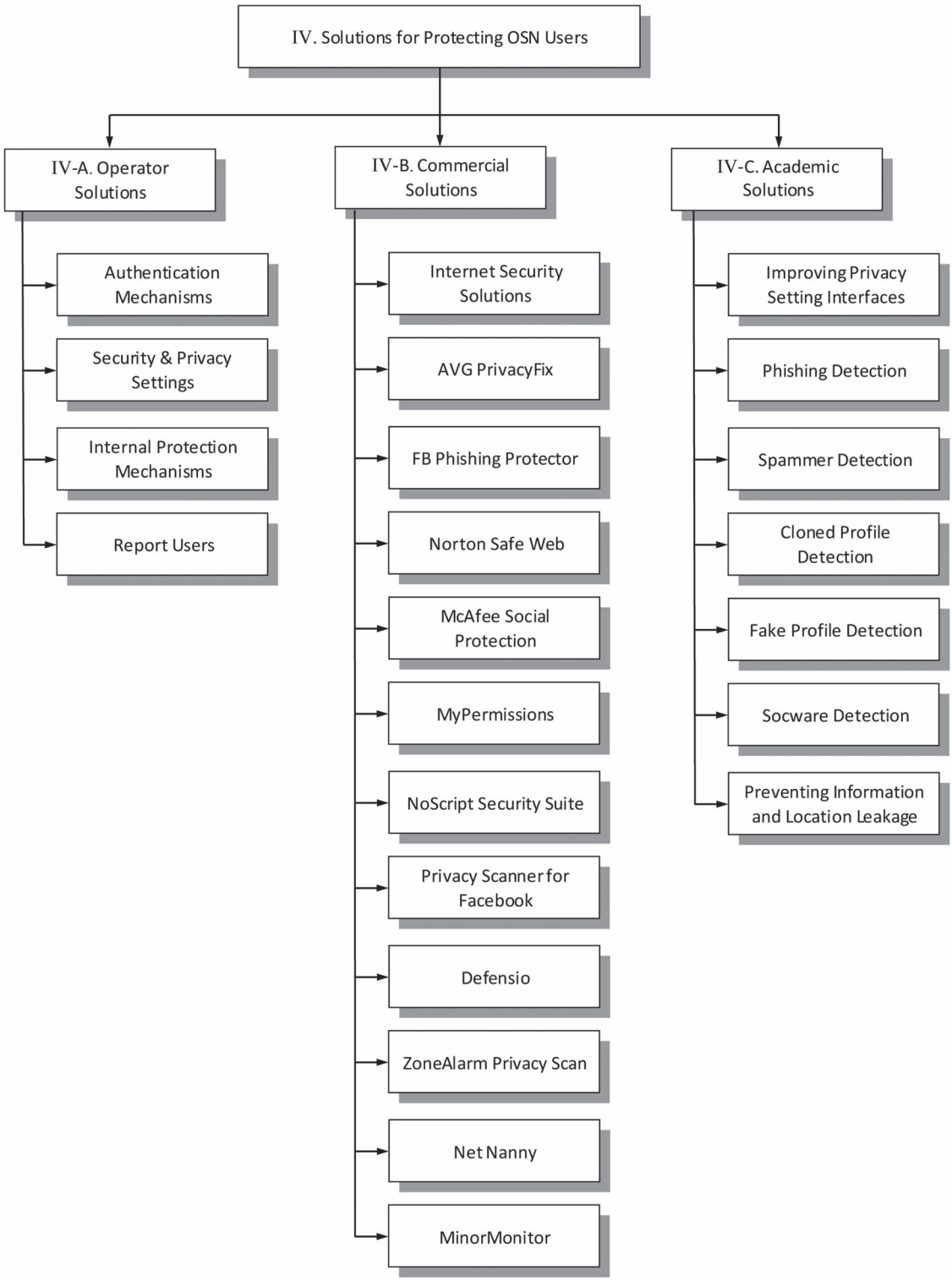

Fig. 3. Security and privacy solutions for online social networks.

notifications, and establish other safety features [117]. However, many OSN users still simply maintain the default privacy settings, letting their data be exposed to strangers [46], [78].

Internal Protection Mechanisms: Several OSNs protect their users by implementing additional internal protection mechanisms for defense against spammers, fake profiles, scams, and other threats [70], [118]. Facebook, for example, protects its users from malicious attacks and information collecting by activating the Facebook Immune System (FIS). The FIS is described as an adversarial learning system that performs real-time checks and classifications on read-and-write actions on Facebook's database [118].

Report Users: OSN operators can attempt to protect young children and teenage users from harassment by adding an option to report abuse or policy violations by other users in the network [119]. In some countries, social networks like

TABLE I
COMMERCIAL SOLUTIONS OVERVIEW

Commercial Solutions	Company	Platform	Pricing	Description
Internet Security Solutions	Many security companies	Mainly PC	Usually requires licensing fees with free trial period	Includes anti-virus, firewall, and other Internet protection layers which assist OSN users in shielding their computers from various threats.
AVG PrivacyFix	AVG	Mobile application or web browser add-on	Free	Assists users in managing their privacy settings on Facebook, LinkedIn, and Google.
FB Phishing Protector	Diego Casorran	Browser add-on	Free	Provides Facebook users protection against various phishing attacks.
Norton Safe Web	Symantec	Facebook application	Free	Warns users about unsafe links and sites in their Facebook News Feed.
McAfee Social Protection	Intel Security	Mobile application	Free	Enables Facebook users to safeguard their uploaded photos.
MyPermissions	Online Permissions Technologies	Web service	Free	Provides its users convenient links to the permissions pages for many OSNs, such as Facebook and Twitter.
NoScript Security Suite	Giorgio Maone	Browser add-on	Free	Allows executable web content, such as JavaScript and Flash, to run only from trusted domains of the user's choice.
Privacy Scanner for Facebook	Trend Micro	Mobile application	Free	Scans the user's privacy settings and identifies risky settings which may lead to privacy concerns.
Defensio	Websense	Web service	Free	Helps protect from threats like links to malware that could be posted on the user's Facebook page. Also assists in preventing information leakage.
ZoneAlarm Privacy Scan	Check Point	Facebook application	Free	Scans recent activity in the user's Facebook account to identify privacy concerns and to control what others can see.
Net Nanny	ContentWatch	PC and mobile application	Paid Software	Allows parents to monitor their children's social media activity.
MinorMonitor	Infoglide	Web service	Free	Gives parents a quick dashboard view of their child's Facebook activities and online friends.

Facebook and Bebo [120] have also added a "Panic Button" to better protect children [121].

B. Commercial Solutions

Various commercial companies have expanded their traditional Internet security options and now offer software solutions specifically for OSN users to better protect themselves against threats. In this section, we present mainstream software and application-protection solutions which were developed by well-known security companies, such as Symantec and Check Point, as well as solutions which were created by several startup companies, such as Online Permissions Technologies, and open-source solutions, such as NoScript Security Suite (see also Table I).

Internet Security Solutions: Many security companies, such as AVG, Avira, Kaspersky, Panda, McAfee, and Symantec [122], offer OSN users Internet security solutions. These software suites typically include anti-virus, firewall, and other Internet protection layers which assist OSN users in shielding their computers against threats such as malware, clickjacking, and phishing attacks. For example, McAfee Internet Security software [123] provides its users with protection against various threats such as malware, botnets, and inappropriate sites.

AVG PrivacyFix: AVG PrivacyFix [124] is software available as a mobile application or a web browser add-on which offers its users a simple way to manage their privacy settings on Facebook, LinkedIn, and Google. Additionally, PrivacyFix helps its users block over 1200 trackers by following their movements online. The software also tells its users how much revenue they are generating for Facebook and Google.

FB Phishing Protector: FB Phishing Protector [125] is a Firefox add-on which warns Facebook users when a suspicious activity is detected, such as a script-injection attempt. This add-on provides protection against various phishing attacks.

Norton Safe Web: Symantec's Norton Safe Web [126] is a Facebook application with more than 500 000 users. It scans the Facebook user's News Feed and warns the user about unsafe links and sites.

McAfee Social Protection: McAfee Social Protection [127] is a mobile application which enables Facebook users to safeguard their uploaded photos by letting users control precisely who can view and download their images.

MyPermissions: Online Permissions Technologies' MyPermissions [128] is a web service that provides its users with convenient links to the permissions pages for many OSNs, such as Facebook, Twitter, and LinkedIn. These links can help users view and revoke the permissions they had given in the past to various applications, thus better protecting their privacy. Additionally, MyPermissions offers periodic email reminders that prompt users to check their OSN permissions settings.

NoScript Security Suite: NoScript Security Suite [129] is an open-source extension to Mozilla-based web browsers like Firefox, which allows executable web content such as JavaScript, Java, and Flash to run only from trusted domains of the user's choice. Blocking executable web content running from untrusted sites can protect OSN users from clickjacking and XSS attacks.

Privacy Scanner for Facebook: Trend Micro's Privacy Scanner for Facebook [130] is an Android application which scans the user's privacy settings and identifies risky settings which may lead to privacy concerns. It then assists the user in fixing the settings.

Defensio: Websense's Defensio web service [131] helps protect social network users from threats like links to malware that could be posted on the user's Facebook page. The Defensio service also assists in preventing information leakage by controlling the user's published content by removing certain words from posts or filtering specific comments.

ZoneAlarm Privacy Scan: Check Point's ZoneAlarm Privacy Scan [132] is a Facebook application which scans recent activity in the user's Facebook account to identify privacy concerns and to control what others can see. For instance, ZoneAlarm Privacy Scan can identify posts that expose the user's private information.

Net Nanny: ContentWatch's Net Nanny [133] is software which assists parents in protecting their children from harmful content. Net Nanny lets parents monitor their children's social media activity on different OSN websites, such as Facebook, Twitter, and Flickr [134].

MinorMonitor: Infoglide's MinorMonitor [55] is a parental control web service which gives parents a quick dashboard view of their child's Facebook activities and online friends. By using MinorMonitor, parents can be informed about questionable content that may have been revealed to their child, and they can identify over-age friends in their child's Facebook friends list.

C. Academic Solutions

Several recently published studies have proposed solutions to various OSN threats. These solutions have primarily focused on identifying malicious users and applications. In this section, we present studies which provide solutions for improving OSN users' privacy settings; for detecting phishing, spammers, cloned and fake profiles, and socware; and for preventing information and location leakage.[2] These academic solutions provide cutting-edge insight into dealing with social network threats. They can be used by OSN operators to improve their users' security and privacy, by security companies to offer the customers better OSN protection, or by early-adopter OSN users who want to better protect themselves.

Improving Privacy Setting Interfaces: In recent years several studies have offered OSN users methods and applications to help them better understand and improve their social network privacy settings. In 2008, Lipford *et al.* [135] introduced the Audience View interface for Facebook which enables users to view their profiles from the point of view of other Facebook users, whether from the point of view of a friend or that of a complete stranger. This type of interface can help OSN users know exactly which personal details are visible to other users and then change their privacy settings accordingly. In 2010, Fang and LeFevre [136] presented a template for the design of a social networking privacy wizard for OSNs to automatically configure the user's privacy settings with minimal effort from the user. Fang and LeFevre also presented a sample privacy wizard based on their generic template. The sample wizard used active learning algorithms and was found to be "quite effective in reducing the amount of user effort, while still producing high-accuracy settings" [136]. In 2012, Fire *et al.* [45] presented The Social Privacy Protector add-on which can assist Facebook users in adjusting their privacy settings with just one simple click, according to predefined various privacy setting usage templates. Also in 2012, Paul *et al.* [137] offered the C4PS privacy interface which utilizes simple principles of color coding to highlight each attribute in the user's profile with a particular color, depending on the group of people who have access to this attribute. Moreover, the interface enables users to change privacy settings for a specific attribute by simply clicking on buttons located near the specific attribute.

Phishing Detection: Many researchers have suggested anti-phishing methods to identify and prevent phishing attacks; most of these methods have been based on techniques that attempt to identify phishing websites and phishing URLs [138]–[140]. With the increasing number of phishing attacks on OSNs [68], several researchers have suggested dedicated solutions for identifying social network phishing attacks. In 2012, Lee *et al.* [141]

[2]Many of these solutions overlap and can assist in preventing more than one threat. For example, algorithms for identifying fake profiles can also help identify spammers and phishing attacks.

introduced WarningBird, a suspicious URL detection system for Twitter which can handle phishing attacks that conceal themselves by using conditional redirection URLs. Later in the same year, Aggarwal *et al.* [142] presented the PhishAri technique, which can detect whether or not a tweet posted with a URL is phishing by utilizing specific Twitter features such as the account age and the number of followers of the user who posted the suspicious tweet.

Spammer Detection: Many researchers have recently proposed solutions for spammer detection in OSNs. In 2009, Benevenuto *et al.* [38] offered algorithms for detecting video spammers which succeeded in identifying spammers among YouTube [143] users. In 2010, DeBarr and Wechsler [40] used the graph centrality measure to predict if a user is likely to send spam messages. Wang [43] proposed a method to classify spammers on Twitter by using content and social network graph properties. Stringhini *et al.* [42] created more than 300 fake profiles (also referred to as "honey-profiles") on Twitter, Facebook, and MySpace and successfully identified spammers who sent spam messages to the fake profiles. Lee *et al.* [41] also presented a method for detecting social spammers of different types by using honeypots combined with machine learning algorithms. In 2013, Aggarwal *et al.* [37] presented machine learning algorithms for detecting various type of spammers in Foursquare. Recently, Bhat and Abulaish [39] introduced a community-based framework to identify OSN spammers. Also, Verma *et al.* [144] presented a survey which reviews existing techniques for detecting spam users on Twitter.

Cloned Profile Detection: In 2011, Kontaxis *et al.* [56] proposed a methodology for detecting social network profile cloning. They designed and implemented a prototype which can be employed to investigate whether or not users have fallen victim to clone attacks. In 2013, Shan *et al.* [145] presented the CloneSpotter which can be deployed into the OSN infrastructure and can detect cloning attacks by using users' data records, such as a user's login IP records that are available to the OSN operator.

Fake Profile Detection: In recent years, researchers have developed algorithms, techniques, and tools to identify fake profiles and prevent various sybil attacks via OSNs.[3] In 2006, Yu *et al.* [52] presented the SybilGuard decentralized protocol that assists in preventing sybil attacks. Later, in 2008, Yu *et al.* [51] also presented the SybilLimit protocol, a near-optimal defense against sybil attacks using social networks. In 2009, Danezis and Mittal [44] offered the SybilInfer defense algorithm which can distinguish between "honest" and "dishonest" users. In the same year, Tran *et al.* [48] presented the SumUp sybil defense system to limit the number of fake votes cast by sybils.

In 2012, Cao *et al.* [16] introduced the SybilRank tool which utilizes OSN graph properties to rank users according to their perceived likelihood of being fake. Later, they deployed SybilRank in the operation center of Tuenti [146], the largest OSN in Spain, and estimated that about 90% of the 200 000 users who received the lowest rank were actually fake profiles. In the same year, Wang *et al.* [50] proposed a crowdsourced fake profiles detection system and evaluated it using data from Facebook and from Renren [147], a Chinese OSN. Also, in 2012, Fire *et al.* [47] presented an algorithm for identifying malicious profiles using the social network's own topological features. They evaluated their methods on three directed OSNs—Academia.edu [148], Anybeat,[4] and Google+—and succeeded in identifying fake profiles and spammers. Fire *et al.* [45] also presented The Social Privacy Protector application which assists Facebook users in identifying fake profiles among their friends. They used the dataset created by The Social Privacy Protector application and developed machine learning classifiers which can identify fake profiles on Facebook [46]. Recently, Wang *et al.* [49] presented a system which can detect fake profiles based on analyzing clickstream models. Additional surveys regarding solutions to sybil attacks have also been presented by Levine *et al.* [149] and by Hoffman *et al.* [150].

Socware Detection: In the last few years, several studies have tried to better understand and identify socware. In 2012, Rahman *et al.* [103] presented the MyPageKeeper Facebook application that aims to protect Facebook users from damaging posts on their timelines. Rahman *et al.* also presented Facebook's Rigorous Application Evaluator (FRAppE) for detecting malicious applications on Facebook [104]. In 2013, Huang *et al.* [105] studied the socware ecosystem and discovered several insights about socware propagation characteristics that can assist in future research on the detection and prevention of socware propagation.

Preventing Information and Location Leakage: In their study on privacy leaks on Twitter, Mao *et al.* [91] offered a "guardian angel service" that can monitor users' tweets and alert users to potential privacy violations. Their offered solution can be based on classifiers they constructed throughout their study which can identify tweets containing private information, such as vacation plans. Moreover, Gómez-Hidalgo *et al.* [151] used Named Entity Recognition (NER) algorithms to prevent data leakage. In their study, they implemented a prototype to demonstrate how their methods can prevent data leakage. Their methods may also be used to prevent OSN users from exposing their locations. Recently, Ghiglieri *et al.* [152] presented the Personal DLP tool to help OSN users better understand and evaluate the sensitivity of their posted statuses. The study included 221 participants, and the developed Personal DLP prototype was found to have a positive impact on users' privacy awareness.

V. Discussion

In Section III, we presented the many threats that can jeopardize OSN users' security and privacy. These threats attempt to achieve one or more of the following goals: (a) gain access to the user's resources, such as passwords and credit card numbers

[3] Although the common goal of both fake profile algorithms and sybil defense algorithms is to identify fake profiles, a difference exists: Fake profile detection algorithms seek to identify fake profiles in general, including cases of cyber predators which hold only a few fake profiles in the OSN; sybil defense algorithms are a private case of fake profile detection algorithms and are usually intended to identify attackers who create a large number of fake profiles in the OSN.

[4] As of May 2012, the Anybeat OSN has been shutdown.

(see Section III-A); (b) gain access to the user's private and sensitive information, such as age, political views, and current or future whereabouts (see Section III-B); (c) utilize the gained control over the user's OSN profile as a spreading platform to attack his or her trusting online friends; and (d) locate future potential victims (see Sections III-B and III-D). Some of these threats are passive; they use only the user's lack of awareness or knowledge to achieve their goals. For example, the face recognition threat introduced in Section III-B can simply utilize the user's public profile photos to create a biometric database. Other threats are active, and their goal is to try and set up the users. For example, the clickjacking threat tries to trick OSN users into clicking on something different from what they had intended to click (see Section III-B). Alarmingly, many of the presented threats are not limited to cyberspace but have the potential to threaten the user's well-being in the real world as well. For example, it has been suggested that most burglars use OSNs such as Facebook and Twitter to target their victims [153].

To better protect OSN users from the above mentioned threats, OSN operators, commercial security companies, and academic researchers offer OSN users a variety of security and privacy solutions which are presented in Section IV. Similar to real-world security solutions, these solutions can provide OSN users with several layers of protection against these threats. The first protection layer, which parallels the functionality of a *door lock*, strives to prevent unwelcome intruders from entering and viewing OSN users' personal posts and details. This layer consists of different security and privacy settings offered by various OSN operators. However, in many cases the average OSN user does not know or is unaware of the best way to "lock" his or her profile, instead leaving the privacy settings on default, which often provides insufficient protection [46], [78]. To assist such users, security companies and academic researchers have developed solutions, such as Privacy Scanner for Facebook [130], ZoneAlarm Privacy Scan [132], and The Social Privacy Protector [45], all of which can assist OSN users in improving their privacy settings. However, much like in real life, sometimes OSN users can forget to "lock their door," and consequently they may leak sensitive information about themselves, such as their future vacation plans or their medical condition [91]. To prevent this type of exposure, researchers [91], [97] and security companies [131] have offered solutions that automatically scan the users' posted information and prevent them from uploading posts that contain their sensitive information.

The second protection layer parallels the functionality of a *security alarm*, and it aims to prevent malicious users from collecting OSN users' personal posts and details, that is, to prevent these malicious users from hacking into the innocent users' devices and social network accounts. This layer consists of the different commercial Internet security solutions (see Section IV-B), as well as the various phishing, fake profile, and socware detection solutions offered by academic researchers that the OSN users can install by themselves (see Section IV-C). These types of solutions can be very effective in identifying active threats, which in many cases attempt to infect as many OSN users as possible. In most cases, however, these solutions are insufficient for identifying more targeted threats, such as de-anonymization attacks, identity clone attacks, inference attacks, and online predators, all of which choose to target individuals using an OSN.

The third protection layer, which functions as a *security camera*, is a special layer specific to children and their OSN use. This layer aims to protect both young children and teenagers by enabling parents to monitor online activity primarily via various monitoring software such as Net Nanny [133] and MinorMonitor [55]. This solution can help parents protect their children from targeted threats such as online predators and cyberbullying.

The fourth protection layer, which can be likened to the functionality of a *neighborhood watch*, uses wisdom of the crowd to pinpoint malicious users in the OSN. This layer consists of various solutions such as the option to report other social network users to an OSN operator. OSN users can work together to identify threats such as fake profiles, clickjacking, internet fraud, socware, and cyberbullying, and report them to the OSN operator.

The fifth protection layer, which parallels the functionality of a *police force*, includes authentication mechanisms which are responsible for making sure that only real people can log into the OSN. The authentication mechanisms can assist in identifying malicious users, such as socialbots, and prevent them from logging into the OSN and attacking other social network users. Additionally, due to its almost unlimited access to OSN users' data, metadata, and activities, the OSN operator can identify many potential threats based on the full social network topology, along with users' IP addresses, login times, and behavioral patterns, which in most cases are accessible only to the OSN operator. Moreover, as demonstrated in Sections IV-A and IV-C, utilizing these unique datasets can help protect OSN users from threats such as phishing attacks [142], spammers [42], cloning attacks [145], and fake profiles [47]. Fire *et al.* [47] showed how the OSN operator can utilize the full social network graph topology in order to identify fake profiles and spammers. Furthermore, as demonstrated by Stringhini *et al.* [42], the OSN operator can use its control over the network to scatter many "honey-profiles" that can assist in identifying malicious users, such as spammers.

These five protection layers can give OSN users sufficient protection against almost all of the threats described in Section III (also see Table II). Moreover, if the OSN users choose to enable only the first three protection layers, they are still safeguarded from most of the described threats. Nevertheless, OSN operators—due to their control of the network, their unique access to all users' data and metadata, and their ability to monitor users' activities OSN operators—are in the best position to improve their users' security and privacy.

VI. Recommendations

As we have demonstrated throughout this study, OSN users are facing prevalent and varied security and privacy threats. Fortunately, there are many software solutions and techniques that exist today which can assist OSN users in better defending themselves against these threats. In this section, we provide several easy-to-apply methods which can help OSN users improve their security and privacy in social networks such as

TABLE II
ONLINE SOCIAL NETWORK THREATS AND THEIR CORRESPONDING SOLUTIONS

Solutions \ Threats	Malware	Phishing Attacks	Spammers	Cross–Site Scripting (XSS)	Internet Fraud	Clickjacking	De-anonymization	Face Recognition	Fake Profiles (Socialbots)	Identity Clone Attacks	Inference Attacks	Information Leakage	Location Leakage	Socware	Online Predators	Risky Behaviors	Cyberbullying
Authentication Mechanisms		X	X						X	X					X		
Security & Privacy Settings			X		X		X	X	X		X	X	X		X		X
Internal Protection Mechanisms	X	X	X	X	X	X			X	X				X	X		X
Report Users	X	X	X	X	X	X			X	X				X	X		X
Internet Security Solutions	X	X	X	X	X	X								X			
AVG PrivacyFix			X				X	X	X		X	X	X		X		X
FB Phishing Protector		X		X		X											
Norton Safe Web	X	X	X	X	X	X								X			
McAfee Social Protection								X				X	X				
MyPermissions			X									X	X	X			
NoScript Security Suite		X	X	X		X								X			
Privacy Scanner for Facebook			X				X	X	X		X	X	X		X		X
Defensio	X	X	X	X	X	X							X	X			
ZoneAlarm Privacy Scan							X	X	X		X	X	X		X		X
Net Nanny					X				X						X	X	X
MinorMonitor					X				X						X	X	X
Improving Privacy Setting Interfaces			X				X	X	X		X	X	X		X		X
Phishing Detection		X	X		X	X			X					X			
Spammer Detection		X	X		X	X			X					X			
Cloned Profile Detection								X	X	X		X	X				
Fake Profile Detection			X		X	X			X	X				X			X
Socware Detection			X			X		X	X			X	X	X			
Preventing Information and Location Leakage							X	X			X	X	X				

Facebook and Twitter. We advise OSN users who want to better protect themselves in these platforms to implement the following eight recommendations in each of their OSN accounts:

1) **Remove Unnecessary Personal Information**. We advise OSN users to review the details they have inserted into their OSN accounts and remove extraneous information about themselves, their family, and their friends. It is also recommended that users hide their friends list if possible, to prevent inference attacks. Additionally, we advise users not to use their full name when using OSNs, and in order to prevent face recognition, we highly recommend users not to use an identifiable image as their profile picture.
2) **Adjust Privacy and Security Settings**. In many social networks, like Facebook, the default privacy settings are insufficient. Yet a recent study has shown that many Facebook users tend to stay with their default privacy settings [46]. In order for users to better protect themselves on Facebook and in other OSNs, we recommend modifying the privacy settings so that users' personal data will be exposed only to themselves, or at most to their friends only (for example, see Fig. 4). Additionally, if possible, we advise users to activate the secure browsing option and any other available authentication mechanisms (see Section IV), such as Twitter's two-factor authentication [32].
3) **Do Not Accept Friend Requests From Strangers**. As we demonstrated in Section III, fake profiles are quite common and often dangerous. Therefore, if a user receives a friend request from an unknown person, we recommend ignoring such a request. If the user is uncertain and is considering approving the friend request, we recommend performing a short background check on the new "friend" and, at a minimum, insert the friend's profile image into Google Images search [154] and submit the friend's full name and other details to other search engines in order to validate the authenticity of

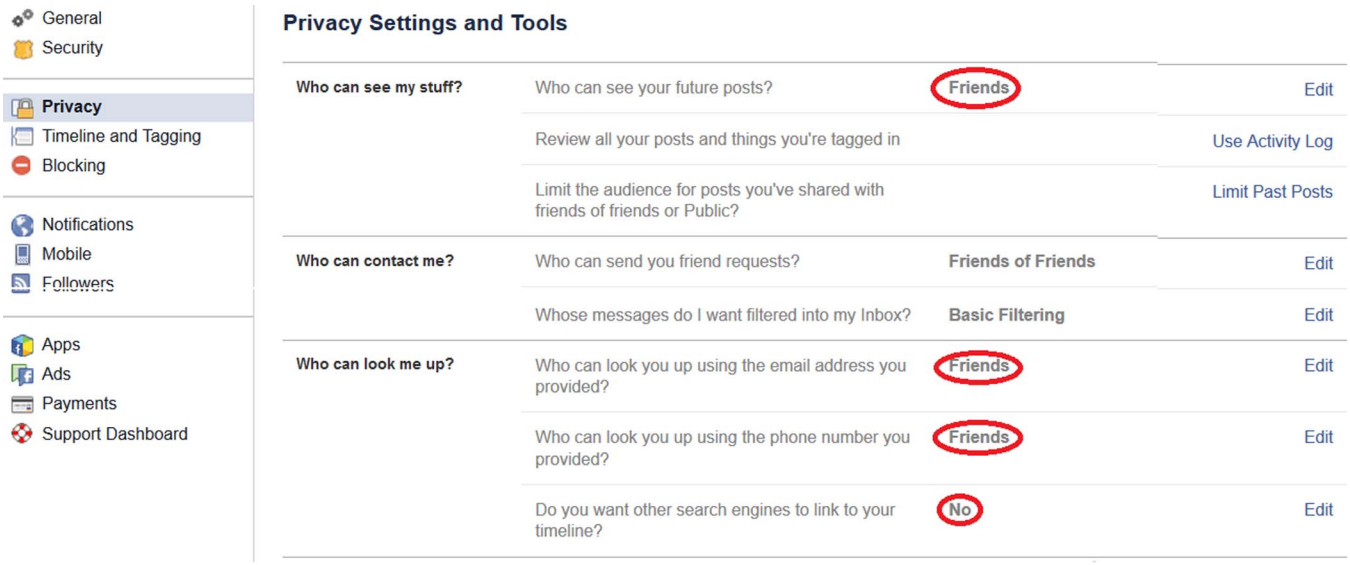

Fig. 4. **An Example of Recommended Privacy Settings on Facebook**. Only friends can have access to the user's private information.

the individual. In order to identify and remove strangers who are already listed as friends with the user, we recommend OSN users examine their friends list or use applications such as The Social Privacy Protector [45] and periodically remove friends with whom they are not familiar or friends who should not have access to personal information.

4) **Install Internet Security Software**. We advise OSN users to install at least one of the many commercial Internet security software products; Facebook offers several free security downloads [122]. We also encourage users to install other security and privacy products as described in Section IV-B.
5) **Remove Installed Third-Party Applications**. Unbeknown to many users, third-party applications frequently collect online personal data. A recent study showed that 30% of an examined group of Facebook users had at least forty applications installed on their accounts [46]. It is recommended that users do not install new, unnecessary applications on their accounts. Moreover, users are advised to periodically go over their list of installed applications and remove any unnecessary applications.
6) **Do Not Publish Your Location**. As we described in Section III-B, many users publish their current or future location in multiple OSNs, and this information can be used by criminals or stalkers. It is recommended that users avoid publishing any geographic location whatsoever in their accounts. Moreover, users are advised to disable geotagging on their mobile devices and cameras to prevent uploading of photos and videos that may contain location information.
7) **Do Not Trust Your OSN Friends**. As we described in Section III, OSN users tend to trust their friends in the social network. Since this trust can be misplaced, we recommend OSN users take extra precautions when communicating with their online friends. We also recommend that users think twice before offering any personal and sensitive information about themselves, even when posting photos. OSN users should definitely avoid revealing their home address, phone number, or credit cards numbers.
8) **Monitor Your Children's OSN Activity**. We strongly advise parents to apply all the above mentioned recommendations to their children's OSN profiles. Additionally, we recommend parents monitor their children's online activity in OSNs. This monitoring can be done manually or by using one of the monitoring software products which we reviewed at the end of Section IV-B. Moreover, we highly recommend that parents and their children periodically scan the friends list together in order to remove unwelcome "friends."

VII. Future Research Directions

The field of OSN security and privacy is a new and emerging one, offering many directions to pursue. Security researchers can continually provide better solutions to online threats; they can also discover new security threats to address. We believe that in order to improve the present solutions, the next step is to create synergy among the different security solutions which were presented in Section IV-C. This will create more robust and effective security solutions for detecting fake profiles, spammers, phishing attacks, socware, and other threats.

Besides the creation of synergy, another worthwhile direction is to apply various algorithms to enhance OSN security. A variety of Natural Language Processing (NLP) techniques and temporal analysis algorithms can be utilized; combining these with existing solutions would provide better and more accurate protection against social network threats. For example, researchers can predict many users' private traits, such as age and gender, based on their Facebook likes [29]. Combining this algorithm with other topological-based fake profile detection methods (see Section IV-C) can assist in spotting phony details, such as a false age, thus identifying fake profiles. Other algorithms also can be utilized: Various Data Leak Prevention (DLP) algorithms can analyze and monitor OSN users' posted information, recommending to the users which of their

posted information might be sensitive and therefore advised to be removed from social network. Additionally, state-of-the-art anomaly detection algorithms could be used to develop solutions for identifying fake OSN user accounts or OSN user accounts that have been compromised.

A further research direction for improving OSN users' privacy is to analyze and evaluate the different existing privacy solutions offered by OSN operators, pinpointing their shortcomings and suggesting methods for improving privacy solutions. Research that develops techniques to better educate users about these solutions would also be of value, as would techniques to make users more aware of existing OSN threats.

Additional possible future research directions include developing privacy-preserving OSNs, such as Safebook [155], and developing solutions for privacy-preserving *ad hoc* social networks (i.e., self-configuring social networks that connect users using mobile devices [156]), such as the semantics-based mobile social network (SMSN) framework [157]. As SMSN grows in popularity, addressing security concerns will be increasingly important.

One additional possible future research direction includes studying the emerging security threats due to the increasing popularity of geo-location tagging of social network users [158] in order to offer solutions for threats with geosocial specificity.

VIII. Conclusion

OSNs have become part of our everyday life and, on average, most Internet users spend more time on social networks than in any other online activity (see Section II). We enjoy using OSNs to interact with other people through the sharing of experiences, pictures, and videos. Nevertheless, social networks have a dark side ripe with hackers, fraudsters, and online predators, all of whom are capable of using OSNs as a platform for procuring their future victims. In this paper, we have presented scenarios which threaten OSN users and can jeopardize their identities, privacy, and well-being in both the virtual world as well as the real world (see Section IV-C). Furthermore, we have provided examples of many of the presented threats in order to demonstrate that these threats are real and can endanger every user. We have also emphasized certain threats which challenge the safety of young children and teenagers across the OSN cyberspace.

There are remedies to these threats, and we have offered a range of solutions which help protect an OSN user's privacy and security (see Section IV). However, as demonstrated in Table II, the presented solutions are not magical antidotes that will provide full protection to a user's privacy and security. In order to be well protected against the various online threats, users must stay attentive to the information they post online, and they must employ more than one solution. In many cases, the users should seek the OSN operator's assistance in providing tools (see Section IV-A) both to better protect their privacy and to identify potential threats.

We have outlined eight recommendations that are simple to implement for OSN users to better protect themselves (see Section VI). We advise OSN users to not only adopt our recommendations but also to educate themselves and their loved ones regarding online threats. All social network users must consider very carefully what personal information is being revealed about themselves, about their friends, and about their workplaces. Users should also know that the information they post in OSNs can be cross-referenced with other data sources [159] and could be used to infer their personal and intimate details. If a user's personal information falls into the wrong hands, it could potentially cause a vast amount of damage, and in many cases there is no way to recapture what has been lost.

In addition, parents must monitor their children's activity in these social platforms. As parents, we cannot be naïve; we need to recognize the enticements of social networks and be aware of hidden dangers. We are obligated to educate our children to be aware of potential threats, and we must teach them not to engage with strangers either in the real world or in the cyber world.

As far as future research (see Section VII), OSNs offer fertile ground for new and interesting research with many opportunities to pursue, such as improving the current state-of-the-art security products, discovering new types of security and privacy threats, and developing and evaluating new privacy solutions and schemes. Overall, researchers can play a significant role by recognizing the value of solution synergies and by applying useful techniques and algorithms. Social networks can enhance our lives, but we must take the correct precautions to preserve our security and privacy.

Acknowledgment

We would like to thank Jennifer Brill and Liza Futerman for proofreading this article. Especially, we want to thank Carol Teegarden for her editing expertise and endless helpful advice which guided this article to completion. We also want to thank the anonymous reviewers for their helpful comments.

References

[1] Facebook, accessed Jan. 14, 2014. [Online]. Available: http://www.facebook.com/

[2] Google+, accessed Jan. 14, 2014. [Online]. Available: https://plus.google.com/

[3] LinkedIn, accessed Jan. 14, 2014. [Online]. Available: http://www.linkedin.com/

[4] Sina Weibo, accessed Jan. 14, 2014. [Online]. Available: http://www.weibo.com/

[5] Twitter, accessed Jan. 14, 2014. [Online]. Available: http://www.twitter.com/

[6] Tumblr, accessed Jan. 14, 2014. [Online]. Available: http://www.tumblr.com/

[7] VKontakte, accessed Jan. 14, 2014. [Online]. Available: http://www.vk.com/

[8] Facebook, *Facebook Reports Fourth Quarter and Full year 2013 Results*, accessed Jan. 14, 2014. [Online]. Available: http://investor.fb.com/releasedetail.cfm?ReleaseID=821954

[9] J. Feinberg, accessed Jan. 14, 2014. [Online]. Available: http://www.wordle.net/

[10] Wikipedia, *List of Virtual Communities With More Than 100 Million Active Users*, accessed Sep. 8, 2013. [Online]. Available: http://en.wikipedia.org/wiki/List_of_virtual_communities_with_more_than_100_million_active_users

[11] Facebook, *Form 10-k (Annual Report)—Filed 02/01/13 for the Period Ending 12/31/12*, 2013, accessed Jan. 9, 2014. [Online]. Available: http://files.shareholder.com/downloads/AMDA-NJ5DZ/2301311196x0xS1326801-13-3/1326801/1326801-13-3.pdf

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

[13] A. Mislove, B. Viswanath, K. Gummadi, and P. Druschel, "You are who you know: Inferring user profiles in online social networks," in *Proc. 3rd ACM Int. Conf. Web Search Data Mining*, 2010, pp. 251–260.

[14] L. Bilge, T. Strufe, D. Balzarotti, and E. Kirda, "All your contacts are belong to us: Automated identity theft attacks on social networks," in *Proc. 18th Int. Conf. World Wide Web*, 2009, pp. 551–560.

[15] J. Baltazar, J. Costoya, and R. Flores, "The real face of koobface: The largest web 2.0 botnet explained," *Trend Micro Res.*, vol. 5, no. 9, p. 10, 2009.

[16] 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 Conf. NSDI*, 2012, p. 15. [Online]. Available: http://dl.acm.org/citation.cfm?id=2228298.2228319

[17] G. Stringhini *et al.*, "Follow the green: Growth and dynamics in twitter follower markets," in *Proc. Conf. Internet Meas. Conf.*, 2013, pp. 163–176.

[18] Y. Boshmaf, I. Muslukhov, K. Beznosov, and M. Ripeanu (2013, Feb.). Design and analysis of a social botnet. *Comput. Netw., Int. J. Comput. Telecommun. Netw.* [Online]. *57(2)*, pp. 556–578. Available: http://dx.doi.org/10.1016/j.comnet.2012.06.006

[19] A. Elyashar, M. Fire, D. Kagan, and Y. Elovici, "Organizational intrusion: Organization mining using socialbots," in *Proc. IEEE/ASE Int. Cyber Security Conf.*, 2012, pp. 7–12.

[20] J. Wolak, D. Finkelhor, K. Mitchell, and M. Ybarra, "Online "predators" and their victims," *Psychol. Violence*, vol. 1, pp. 13–35, 2010.

[21] M. Ybarra and K. Mitchell, "How risky are social networking sites? A comparison of places online where youth sexual solicitation and harassment occurs," *Pediatrics*, vol. 121, no. 2, pp. e350–e357, Feb. 2008.

[22] C. Dwyer, S. R. Hiltz, and K. Passerini, "Trust and privacy concern within social networking sites: A comparison of facebook and myspace," presented at the 13th Americas Conf. Information Systems (AMCIS), Keystone, CO, USA, 2007, Paper 339. [Online]. Available: http://aisel.aisnet.org/amcis2007/339/

[23] MySpace, accessed Jan. 14, 2014. [Online]. Available: http://www.myspace.com

[24] A. Acquisti and R. Gross, "Imagined communities: Awareness, information sharing, privacy on the facebook," in *Privacy Enhancing Technologies*. New York, NY, USA: Springer-Verlag, 2006, pp. 36–58.

[25] A. Elyashar, M. Fire, D. Kagan, and Y. Elovici, "Homing socialbots: Intrusion on a specific organization's employee using socialbots," in *Proc. IEEE/ACM Int. Conf. Adv. Social Netw. Anal. Mining*, 2013, pp. 1358–1365.

[26] C. Tucker, "Social networks, personalized advertising, perceptions of privacy control," presented at the 10th Workshop Economics Information Security (WEIS), Fairfax, VA, USA, 2011.

[27] C. C. Miller, The New York Times, *Tech Companies Concede to Surveillance Program*, Jun. 2013, accessed Jan. 14, 2014. [Online]. Available: http://www.nytimes.com/2013/06/08/technology/tech-companies-bristling-concede-to-government-surveillance-efforts.html

[28] C. Jernigan and B. F. Mistree, "Gaydar: Facebook friendships expose sexual orientation," *First Monday*, vol. 14, no. 10, Oct. 2009.

[29] M. Kosinski, D. Stillwell, and T. Graepel, "Private traits and attributes are predictable from digital records of human behavior," *Proc. Nat. Acad. Sci. USA*, vol. 110, no. 15, pp. 5802–5805, Apr. 2013.

[30] J. Constine, *Facebook Launches Verified Accounts and Pseudonyms*, Feb. 2012, accessed Jan. 14, 2014. [Online]. Available: http://techcrunch.com/2012/02/15/facebook-verified-accounts-alternate-names/

[31] A. Jeffries, *Facebook's Security Check Asks Users to Identify Photos of Friends' Dogs, Gummi Bears*, 2010, accessed Feb. 1, 2014.

[32] J. O'Leary, *Getting Started With Login Verification*, May 2013, accessed Jan. 14, 2014. [Online]. Available: https://blog.twitter.com/2013/getting-started-login-verification

[33] A. Song, *Introducing Login Approvals*, May 2011, accessed Jan. 14, 2014. [Online]. Available: https://www.facebook.com/note.php?note_id=10150172618258920

[34] Y. Liu, K. Gummadi, B. Krishnamurthy, and A. Mislove, "Analyzing facebook privacy settings: User expectations vs. reality," in *Proc. ACM SIGCOMM Conf. Internet Meas. Conf.*, 2011, pp. 61–70.

[35] S. Mahmood and Y. Desmedt, "Poster: Preliminary analysis of google+'s privacy," in *Proc. 18th ACM Conf. Comput. Commun. Security*, 2011, pp. 809–812.

[36] J. Angwin and J. Singer-Vine, "Selling you on facebook," *The Wall Street Journal*, Apr. 2012, accessed Jan.14, 2014. [Online]. Available: http://online.wsj.com/news/articles/SB10001424052702303302504577327744009046230

[37] A. Aggarwal, J. Almeida, and P. Kumaraguru, "Detection of spam tipping behaviour on foursquare," in *Proc. 22nd Int. Conf. World Wide Web Companion*, 2013, pp. 641–648.

[38] F. Benevenuto, T. Rodrigues, V. Almeida, J. Almeida, and M. Gonçalves, "Detecting spammers and content promoters in online video social networks," in *Proc. 32nd Int. ACM SIGIR Conf. Research Dev. Inf. Retrieval*, 2009, pp. 620–627.

[39] S. Y. Bhat and M. Abulaish, "Community-based features for identifying spammers in online social networks," in *Proc. IEEE/ACM Int. Conf. Adv. Social Netw. Anal. Mining*, 2013, pp. 100–107.

[40] D. DeBarr and H. Wechsler, "Using social network analysis for spam detection," in *Proc. Adv. Social Comput.*, 2010, pp. 62–69.

[41] K. Lee, J. Caverlee, and S. Webb, "Uncovering social spammers: Social honeypots+ machine learning," in *Proc. 33rd Int. ACM SIGIR Conf. Res. Dev. Inf. Retrieval*, 2010, pp. 435–442.

[42] G. Stringhini, C. Kruegel, and G. Vigna, "Detecting spammers on social networks," in *Proc. 26th Annu. Comput. Security Appl. Conf.*, 2010, pp. 1–9.

[43] A. Wang, "Don't follow me: Spam detection in twitter," in *Proc. Int. Conf. SECRYPT*, 2010, pp. 1–10.

[44] G. Danezis and P. Mittal, "Sybilinfer: Detecting sybil nodes using social networks," in *Proc. NDSS*, 2009, The Internet Society.

[45] M. Fire, D. Kagan, A. Elyashar, and Y. Elovici, "Social privacy protector-protecting users' privacy in social networks," in *Proc. 2nd Int. Conf. SOTICS*, 2012, pp. 46–50.

[46] M. Fire, D. Kagan, A. Elyashar, and Y. Elovici, "Friend or foe? Fake profile identification in online social networks," *arXiv preprint arXiv:1303.3751*, 2013.

[47] M. Fire, G. Katz, and Y. Elovici, "Strangers intrusion detection-detecting spammers and fake profiles in social networks based on topology anomalies," *Human J.*, vol. 1, no. 1, pp. 26–39, 2012.

[48] D. N. Tran, B. Min, J. Li, and L. Subramanian, "Sybil-resilient online content voting," in *Proc. NSDI*, 2009, vol. 9, pp. 15–28.

[49] G. Wang *et al.*, "You are how you click: Clickstream analysis for sybil detection," in *Proc. 22Nd USENIX Conf. SEC*, 2013, pp. 241–256. [Online]. Available: http://dl.acm.org/citation.cfm?id=2534766.2534788

[50] G. Wang *et al.*, "Social turing tests: Crowdsourcing sybil detection," *arXiv preprint arXiv:1205.3856*, 2012.

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

[52] H. Yu, M. Kaminsky, P. Gibbons, and A. Flaxman, "Sybilguard: Defending against sybil attacks via social networks," *Proc. ACM SIGCOMM Comput. Commun. Rev.*, vol. 36, no. 4, pp. 267–278, Oct. 2006.

[53] Check Point, accessed Jan. 14, 2014. [Online]. Available: http://www.checkpoint.com/

[54] Websense, accessed Jan. 14, 2014. [Online]. Available: http://www.websense.com/

[55] Infoglide, *Minormonitor—Facebook Monitoring and Parental Control Software*, accessed Jan. 14, 2014. [Online]. Available: http://www.minormonitor.com/

[56] G. Kontaxis, I. Polakis, S. Ioannidis, and E. Markatos, "Detecting social network profile cloning," in *Proc. IEEE Int. Conf. PERCOM Workshops*, 2011, pp. 295–300.

[57] Facebook, *Facebook Newsroom*, 2013, accessed Jan. 14, 2014. [Online]. Available: http://newsroom.fb.com/Key-Facts

[58] V. Gundotra, *Google+: Communities and Photos*, Dec. 2012, accessed Jan. 14, 2014. [Online]. Available: http://googleblog.blogspot.co.il/2012/12/google-communities-and-photos.html

[59] S. Fiegerman, *Twitter now has More Than 200 Million Monthly Active Users*, Dec. 2012, accessed Jan. 14, 2014. [Online]. Available: http://mashable.com/2012/12/18/twitter-200-million-active-users/

[60] Linkedin, *About Linkedin*, 2013, accessed Jan. 14, 2014. [Online]. Available: http://press.linkedin.com/about

[61] J. Brenner and S. Aaron, *72% of Online Adults are Social Networking Site Users*, Aug. 2013, accessed Jan. 9, 2014. [Online]. Available: http://www.pewinternet.org/2013/08/05/72-of-online-adults-are-social-networking-site-users

[62] Nielsen, *State of the Media: The Social Media Report (q3 2011)*, 2011, accessed Jan. 9, 2014. [Online]. Available: http://cn.nielsen.com/documents/Nielsen-Social-Media-Report_FINAL_090911.pdf

[63] Yahoo, accessed Jan. 14, 2014. [Online]. Available: http://www.yahoo.com/

[64] Google, accessed Jan. 14, 2014. [Online]. Available: http://www.google.com/

[65] Z. Fox, *The 10 Most Frequently Used Smartphone Apps*, Aug. 2013, accessed Mar. 3, 2014. [Online]. Available: http://mashable.com/2013/08/05/most-used-smartphone-apps/

[66] S. Livingstone, L. Haddon, and K. Ólafsson, *Eu Kids Online: Final Report*, 2011.

[67] T. Amin, O. Okhiria, J. Lu, and J. An, *Facebook: A Comprehensive Analysis of Phishing on a Social System*, 2010, accessed Feb. 1, 2014. [Online]. Available: https://courses.ece.ubc.ca/412/term_project/reports/2010/facebook.pdf

[68] D. Cavit *et al.*, *Microsoft Security Intelligence Report Volume 10*, 2010, accessed Mar. 11, 2014. [Online]. Available: http://www.microsoft.com/en-us/download/details.aspx?id=17030

[69] E. Mills, *Facebook Hit by Phishing Attacks for a Second Day*, Apr. 2009, accessed Jan. 14, 2014. [Online]. Available: http://news.cnet.com/8301-1009_3-10230980-83.html

[70] A. Chowdhury, *State of Twitter Spam*, Mar. 2010, accessed Jan. 14, 2014. [Online]. Available: https://blog.twitter.com/2010/state-twitter-spam

[71] L. Tristan, *Twitter's Growing Spam Problem*, Forbes, Jul. 2013, accessed Mar. 3, 2014. [Online]. Available: http://www.forbes.com/sites/tristanlouis/2013/04/07/twitters-growing-spam-problem/

[72] B. Livshits and W. Cui, "Spectator: Detection and containment of javascript worms," in *Proc. USENIX Annu. Tech. Conf.*, 2008, pp. 335–348.

[73] I. Paul, "Twitter worm: A closer look at what happened," PCWorld, San Francisco, CA, USA, Apr. 2009. [Online]. Available: http://www.pcworld.com/article/163054/twitter_mikeyy_worm_stalkdaily.html

[74] *Informed Investor Advisory: Social Networking*, North American Securities Administrators Association (NASAA), Washington, DC, USA, Sep. 2011. [Online]. Available: http://www.nasaa.org/5568/informed-investor-advisory-social-networking/

[75] J. Halliday, "Facebook fraud a 'Major Issue'," The Guardian, London, U.K., Sep. 2010. [Online]. Available: http://www.theguardian.com/technology/2010/sep/20/facebook-fraud-security

[76] R. Lundeen, J. Ou, and T. Rhodes, "New ways I'm going to hack your web app," in *Proc. Blackhat AD*, 2011, pp. 1–11.

[77] R. McMillan, "Researchers make wormy twitter attack," PCWorld, San Francisco, CA, USA, Mar. 2009. [Online]. Available: http://www.pcworld.idg.com.au/article/296382/researchers_make_wormy_twitter_attack/

[78] B. Krishnamurthy and C. E. Wills, "On the leakage of personally identifiable information via online social networks," in *Proc. 2nd ACM Workshop Online Social Netw.*, 2009, pp. 7–12.

[79] G. Wondracek, T. Holz, E. Kirda, and C. Kruegel, "A practical attack to de-anonymize social network users," in *Proc. IEEE Symp. SP*, 2010, pp. 223–238.

[80] Xing, accessed Jan. 14, 2014. [Online]. Available: http://www.xing.com/

[81] O. Peled, M. Fire, L. Rokach, and Y. Elovici, "Entity matching in online social networks," in *Proc. Int. Conf. SocialCom*, 2013, pp. 339–344.

[82] *The Faces of Facebook*. [Online]. Available: http://app.thefacesoffacebook.com/

[83] A. Acquisti, R. Gross, and F. Stutzman, "Faces of facebook: Privacy in the age of augmented reality," in *Proc. BlackHat USA*, 2011, pp. 1–56.

[84] J. R. Douceur, "The sybil attack," in *Proc. 1st Int. Workshop IPTPS*, 2002, pp. 251–260. [Online]. Available: http://dl.acm.org/citation.cfm?id=646334.687813

[85] H. Gao *et al.*, "Detecting and characterizing social spam campaigns," in *Proc. 10th ACM SIGCOMM Conf. Internet Meas.*, 2010, pp. 35–47.

[86] C. Taylor, *Startup Claims 80% of its Facebook ad Clicks Are Coming From Bots*, Jul. 2012. [Online]. Available: http://techcrunch.com/2012/07/30/startup-claims-80-of-its-facebook-ad-clicks-are-coming-from-bots/

[87] N. Perlroth, "Fake twitter followers become multimillion-dollar business," The New York Times, New York, NY, USA, Apr. 2013. [Online]. Available: http://bits.blogs.nytimes.com/2013/04/05/fake-twitter-followers-becomes-multimillion-dollar-business/

[88] T. Ryan, "Getting in bed with Robin Sage," in *Proc. Black Hat Conf.*, 2010, pp. 1–8.

[89] J. Lewis, "How spies used facebook to steal NATO chiefs' details," The Telegraph, London, U.K., Mar. 2012. [Online]. Available: http://www.telegraph.co.uk/technology/9136029/How-spies-used-Facebook-to-steal-Nato-chiefs-details.html

[90] M. Fire, R. Puzis, and Y. Elovici, "Organization mining using online social networks," *arXiv preprint arXiv:1303.3741*, 2013.

[91] H. Mao, X. Shuai, and A. Kapadia, "Loose tweets: An analysis of privacy leaks on twitter," in *Proc. 10th Annu. ACM Workshop Privacy Electron. Society*, 2011, pp. 1–12.

[92] S. Torabi and K. Beznosov, "Privacy aspects of health related information sharing in online social networks," presented at the USENIX Workshop Health Information Technologies, Washington, DC, USA, 2013. [Online]. Available: https://www.usenix.org/conference/healthtech13/workshop-program/presentation/torabi

[93] L. Scism and M. Maremont, "Insurers test data profiles to identify risky clients," *Wall Street J.*, Nov. 2010. [Online]. Available: http://online.wsj.com/news/articles/SB10001424052748704648604575620750998072986

[94] J. Vicknair, D. Elkersh, K. Yancey, and M. C. Budden, "The use of social networking websites as a recruiting tool for employers," *Amer. J. Bus. Educ.*, vol. 3, no. 11, pp. 7–12, Nov. 2010.

[95] L. Humphreys, "Mobile social networks and social practice: A case study of dodgeball," *J. Comput.-Mediated Commun.*, vol. 13, no. 1, pp. 341–360, Oct. 2007.

[96] L. Humphreys, P. Gill, and B. Krishnamurthy, "How much is too much? Privacy issues on twitter," in *Proc. Conf. Int. Commun. Assoc.*, 2010, pp. 1–29.

[97] Z. Cheng, J. Caverlee, and K. Lee, "You are where you tweet: A content-based approach to geo-locating twitter users," in *Proc. 19th ACM Int. CIKM*, 2010, pp. 759–768. [Online]. Available: http://doi.acm.org/10.1145/1871437.1871535

[98] Pleaserobme.com. [Online]. Available: http://pleaserobme.com/

[99] J. Van Grove, "Are we all asking to be robbed?" Mashable, New York, NY, USA, Feb. 2010. [Online]. Available: http://mashable.com/2010/02/17/pleaserobme/

[100] FourSquare. [Online]. Available: http://www.foursquare.com/

[101] G. Friedland and R. Sommer, "Cybercasing the joint: On the privacy implications of geo-tagging," in *Proc. 5th USENIX Conf. HotSec*, 2010, pp. 1–8. [Online]. Available: http://dl.acm.org/citation.cfm?id=1924931.1924933

[102] K. Murphy, "Web Photos that reveal secrets, like where you live," The New York Times, New York, NY, USA, Aug. 2010. [Online]. Available: http://www.nytimes.com/2010/08/12/technology/personaltech/12basics.html

[103] M. Rahman, T. Huang, H. Madhyastha, and M. Faloutsos, "Efficient and scalable socware detection in online social networks," in *Proc. 21st USENIX Conf. Security Symp.*, 2012, pp. 32–32.

[104] M. S. Rahman, T.-K. Huang, H. V. Madhyastha, and M. Faloutsos, "FRAppE: Detecting malicious facebook applications," in *Proc. 8th Int. Conf. Emerging Netw. Exp. Technol.*, 2012, pp. 313–324.

[105] T.-K. Huang, M. S. Rahman, H. V. Madhyastha, and M. Faloutsos, "An analysis of socware cascades in online social networks," in *Proc. 22nd Int. Conf. World Wide Web*, 2013, pp. 619–630.

[106] Y. Altshuler, N. Aharony, Y. Elovici, A. Pentland, and M. Cebrian, "Stealing Reality: When Criminals Become Data Scientists (or Vice Versa)," in *Security and Privacy in Social Networks*, Y. Altshuler, Y. Elovici, A. B. Cremers, N. Aharony, and A. Pentland, Eds. New York, NY, USA: Springer-Verlag, 2013, pp. 133–151. [Online]. Available: http://dx.doi.org/10.1007/978-1-4614-4139-7_7

[107] S. Livingstone and L. Haddon, *Child Safety Online: Global Challenges and Strategies*, 2011.

[108] Psychology Today Diagnosis Dictionary, *Pedophilia*, 2010. [Online]. Available: http://www.psychologytoday.com/conditions/pedophilia

[109] M. Deans, *The Story of Amanda Todd*, The New Yorker, Oct. 2012. [Online]. Available: http://www.newyorker.com/online/blogs/culture/2012/10/amanda-todd-michael-brutsch-and-free-speech-online.html

[110] Lpsos, *One in Ten (12%) Parents Online, Around the World Say Their Child has Been Cyberbullied, 24% Say They Know of a Child Who has Experienced Same in Their Community*, Jan. 2012. [Online]. Available: http://www.ipsos-na.com/news-polls/pressrelease.aspx?id=5462

[111] M. Pearce, "Florida girl, 12, found dead after bullies said 'Kill Yourself'," Los Angeles Times, Los Angeles, CA, USA, Sep. 2013. [Online]. Available: http://articles.latimes.com/2013/sep/12/nation/la-na-nn-florida-cyberbullying-20130912

[112] H. Moore and D. Roberts, "AP twitter hack causes panic on wall street and sends dow plunging," The Guardian, London, U.K., Apr. 2013. [Online]. Available: http://www.theguardian.com/business/2013/apr/23/ap-tweet-hack-wall-street-freefall

[113] Facebook, *Facebook Help Center: Privacy*. [Online]. Available: http://www.facebook.com/help/privacy/

[114] D. Pogue, "Google+ improves on facebook," The New York Times, New York, NY, USA, Jul. 2011. [Online]. Available: http://www.nytimes.com/2011/07/14/technology/personaltech/google-gets-a-leg-up-on-facebook.html

[115] Facebook, *How do I Restrict an app From Accessing my Information?*. [Online]. Available: https://www.facebook.com/help/151008078302798

[116] Google, *Checking Security Settings and Revoking Access*. [Online]. Available: https://support.google.com/a/answer/2537800

[117] L. Popov, *Staying in Control of Your Facebook Logins*, May 2010. [Online]. Available: https://www.facebook.com/blog/blog.php?post=389991097130

[118] T. Stein, E. Chen, and K. Mangla, "Facebook immune system," in *Proc. 4th Workshop SNS*, 2011, pp. 1–8. [Online]. Available: http://doi.acm.org/10.1145/1989656.1989664
[119] Facebook, *Report Abuse or Policy Violations*. [Online]. Available: https://www.facebook.com/report/
[120] Bebo. [Online]. Available: http://www.bebo.com
[121] S. Axon, *Facebook Will Add a "Panic Button" for U.K. Teens*, Jul. 2010. [Online]. Available: http://mashable.com/2010/07/11/facebook-panic-button-ceop/
[122] Facebook, *Facebook Security Products: Protect Your Computer With Free Security Software Downloads From Your Friends at Facebook*. [Online]. Available: https://www.facebook.com/security/app_360406100715618
[123] Macfee, *Macfee Internet Security*. [Online]. Available: http://home.mcafee.com/store/internet-security
[124] AVG, *Avg Privacyfix*. [Online]. Available: https://www.privacyfix.com/
[125] Diego Casorran, *Facebook Phishing Protector*. [Online]. Available: https://addons.mozilla.org/en-US/firefox/addon/facebook-phishing-protector/
[126] Symantec, *Norton Safe Web*. [Online]. Available: https://www.facebook.com/appcenter/nortonsafeweb
[127] McAfee, *Mcafee Social Protection Beta*. [Online]. Available: https://www.protectmediaonline.com
[128] Online Permissions Technologies, *Mypermissions*. [Online]. Available: http://mypermissions.org/
[129] NoScript Security Suite. [Online]. Available: http://noscript.net/
[130] T. Micro, *Privacy Scanner for Facebook*. [Online]. Available: https://play.google.com/store/apps/details?id=com.trendmicro.socialprivacyscanner
[131] Websense, *The Defensio Web Service*. [Online]. Available: http://www.defensio.com/
[132] CheckPoint, *Zonealarm Privacy Scan*. [Online]. Available: https://www.facebook.com/appcenter/sgprivacy
[133] ContentWatch, *Net Nanny*. [Online]. Available: http://www.netnanny.com/
[134] Flickr. [Online]. Available: http://www.flickr.com/
[135] H. R. Lipford, A. Besmer, and J. Watson, "Understanding privacy settings in facebook with an audience view," in *Proc. 1st Conf. UPSEC*, 2008, pp. 2:1–2:8. [Online]. Available: http://dl.acm.org/citation.cfm?id=1387649.1387651
[136] L. Fang and K. LeFevre, "Privacy wizards for social networking sites," in *Proc. 19th Int. Conf. World Wide Web*, 2010, pp. 351–360.
[137] T. Paul, M. Stopczynski, D. Puscher, M. Volkamer, and T. Strufe, "C4ps: Colors for privacy settings," in *Proc. 21st Int. Conf. WWW Companion*, 2012, pp. 585–586. [Online]. Available: http://doi.acm.org/10.1145/2187980.2188139
[138] S. Garera, N. Provos, M. Chew, and A. D. Rubin, "A framework for detection and measurement of phishing attacks," in *Proc. ACM Workshop Recurring Malcode*, 2007, pp. 1–8.
[139] J. Ma, L. K. Saul, S. Savage, and G. M. Voelker, "Beyond blacklists: Learning to detect malicious web sites from suspicious urls," in *Proc. 15th ACM SIGKDD Int. Conf. Mining*, 2009, pp. 1245–1254.
[140] G. Xiang, J. Hong, C. P. Rose, and L. Cranor (2011, Sep.). CANTINA+: A feature-rich machine learning framework for detecting phishing web sites. *ACM Trans. Inf. Syst. Security* [Online]. *14(2)*, pp. 1–28. Available: http://doi.acm.org/10.1145/2019599.2019606
[141] S. Lee and J. Kim, "Warningbird: Detecting suspicious urls in twitter stream," in *Proc. Symp. NDSS*, 2012, pp. 1–13.
[142] A. Aggarwal, A. Rajadesingan, and P. Kumaraguru, "Phishari: Automatic realtime phishing detection on twitter," in *Proc. eCrime Res. Summit*, 2012, pp. 1–12.
[143] YouTube. [Online]. Available: http://www.youtube.com/
[144] M. Verma, Divya, and S. Sofat, "Article: Techniques to detect spammers in twitter—A survey," *Int. J. Comput. Appl.*, vol. 85, no. 10, pp. 27–32, Jan. 2014.
[145] Z. Shan, H. Cao, J. Lv, C. Yan, and A. Liu, "Enhancing and identifying cloning attacks in online social networks," in *Proc. 7th Int. Conf. Ubiquitous Inf. Manage. Commun.*, 2013, p. 59.
[146] Tuenti. [Online]. Available: http://www.tuenti.com/
[147] Renren. [Online]. Available: http://www.renren.com/
[148] Academia.edu. [Online]. Available: http://www.academia.edu/
[149] B. Levine, C. Shields, and N. Margolin, "A survey of solutions to the sybil attack," Univ. Massachusetts Amherst, Amherst, MA, USA, 2006.
[150] K. Hoffman, D. Zage, and C. Nita-Rotaru, "A survey of attack and defense techniques for reputation systems," *ACM Comput. Surveys*, vol. 42, no. 1, p. 1, Dec. 2009.
[151] J. M. Gómez-Hidalgo *et al.*, "Data leak prevention through named entity recognition," in *Proc. IEEE 2nd Int. Conf. SocialCom*, 2010, pp. 1129–1134.
[152] M. Ghiglieri, M. Stopczynski, and M. Waidner, "Personal DLP for facebook," in *Proc. 6th IEEE Workshop SESOC*, Budapest, Hungary, 2014.
[153] A. Bloxham, "Most burglars using facebook and twitter to target victims, survey suggests," The Telegraph, London, U.K., Sep. 2011. [Online]. Available: http://www.telegraph.co.uk/technology/news/8789538/Most-burglars-using-Facebook-and-Twitter-to-target-victims-survey-suggests.html
[154] Google, *Google Search by Image*. [Online]. Available: http://www.google.com/insidesearch/searchbyimage.html
[155] L. A. Cutillo, R. Molva, and T. Strufe, "Safebook: A privacy-preserving online social network leveraging on real-life trust," *IEEE Commun. Mag.*, vol. 47, no. 12, pp. 94–101, Dec. 2009.
[156] J. Li and S. U. Khan, "MobiSN: Semantics-based mobile ad hoc social network framework," in *Proc. IEEE GLOBECOM*, 2009, pp. 1–6.
[157] J. Li, H. Wang, and S. U. Khan, "A semantics-based approach to large-scale mobile social networking," *Mobile Netw. Appl.*, vol. 17, no. 2, pp. 192–205, Apr. 2012.
[158] C. Ruiz Vicente, D. Freni, C. Bettini, and C. S. Jensen, "Location-related privacy in geo-social networks," *IEEE Internet Comput.*, vol. 15, no. 3, pp. 20–27, May/Jun. 2011.
[159] B. Krishnamurthy, "Privacy and online social networks: Can colorless green ideas sleep furiously?" *IEEE Security Privacy*, vol. 11, no. 3, pp. 14–20, May/Jun. 2013.

Michael Fire (M'12) received the M.Sc. degree (*magna cum laude*) in mathematics from the Bar-Ilan University, Ramat Gan, Israel, and the Ph.D. degree (*summa cum laude*) in information system engineering from Ben-Gurion University of the Negev, Be'er Sheva, Israel.

He is currently a Postdoctoral Fellow with the Department of Information System Engineering, Ben-Gurion University of the Negev. He won the Kreitman Prize for excellence in the Ph.D. studies. He has published dozens of papers for prestigious conferences and journals in the fields of social network analysis and data mining. He also has extensive experience as a Data Scientist working for several companies and organizations.

Roy Goldschmidt received the M.A degree in sociology and anthropology from the Hebrew University, Jerusalem, Israel.

He is currently a Researcher with the Knesset Research and Information Center, Kiryat Ben Gurion, Jerusalem, Israel. He writes policy papers for MPs and for Knesset committees, mainly on issues of science, technology, and ICT.

Yuval Elovici (M'08) received the B.Sc. and M.Sc. degrees in computer and electrical engineering from Ben-Gurion University of the Negev (BGU), Be'er Sheva, Israel, and the Ph.D. degree in information systems from Tel Aviv University, Tel Aviv, Israel.

He is the Director of the Telekom Innovation Laboratories, Head of the Cyber Security Laboratory, and a Professor with the Department of Information Systems Engineering, BGU. He has served as the head of the software engineering program at BGU for two and a half years. For the past ten years, he has led the cooperation between BGU and Deutsche Telekom. He has published more than 60 articles in leading peer-reviewed journals and published over 100 papers in various peer-reviewed conferences. In addition, he has co-authored books on social network security and on information leakage detection and prevention. His primary research interests include computer and network security, cyber security, web intelligence, information warfare, social network analysis, and machine learning. He also consults professionally in the area of cyber security.