Homing Socialbots:
Intrusion on a specific organization’s employee using Socialbots

Aviad Elishar, Michael Fire, Dima Kagan, and Yuval Elovici
Telekom Innovation Laboratories and Information Systems Engineering Department
Ben-Gurion University of the Negev
Beer-Sheva, Israel
{aviade, mickyfi, kagandi, elovici}@bgu.ac.il

Abstract— One dimension on the Internet, which has gained great popularity in recent years are the online social networks (OSNs). Users all over the globe write, share, and publish personal information about themselves, their friends, and their workplace. In this study we present a method for infiltrating specific users in targeted organizations by using organizational social networks topologies, and Socialbots. The targeted organizations, which have been chosen by us, were technology-oriented organizations. Employees from this kind of organization should be more aware of the dangers of exposing private information. An infiltration is defined as accepting a Socialbot’s friend request. Upon accepting a Socialbot’s friend request, users unknowingly expose information about themselves and their workplace. To infiltrate this we had to use our Socialbots in a sophisticated manner. First, we had to gather information and recognize Facebook users who work in targeted organizations. Afterwards, we chose ten Facebook users from every targeted organization randomly. These ten users were chosen to be the specific users from targeted organizations of which we would like to infiltrate. The Socialbots sent friend requests to all specific users’ mutual friends who worked or work in the same targeted organization. The rationale behind this idea was to gain as many mutual friends as possible and through this act increase the probability that our friend requests will be accepted by the targeted users. We tested the proposed method on targeted users from two different organizations. Our method was able to gain a successful percentage of 50% and 70% respectively. The results demonstrate how easily adversaries can infiltrate users they do not know and get full access to personal and valuable information. These results are more surprising when we emphasize the fact that we chose oriented users who should be more aware to the dangers of information leakage for this study on purpose. Moreover, the results indicate once again that users who are interested in protecting themselves should not disclose information in OSNs and should be cautious of accepting friendship requests from unknown persons.

I. INTRODUCTION

As the days pass, Internet becomes more of a central issue in our lives. Today, there are billions of users all over the world who use the Internet [1]. Many of them use it on a daily basis. Nowadays, people can read articles, write talkbacks, buy goods, play games, watch movies, schedule appointments, pay bills, etc. by using the Internet. One sector in the Internet, which has gained great popularity in recent years are online social networks (OSNs) [2] [3]. OSNs like Facebook¹, MySpace², LinkedIn³, etc. allow Internet users to create accounts, to present themselves, and to create or maintain connections with others. There are many reasons given for the growth of OSNs usage. First, many users would like to stay connected to their soundings. With OSNs users can stay updated about their recent friends’ news and maybe be able to maintain close relationship with their close friends [2]. Second, several users’ use these OSNs to create businesses. For years, self-employees, barbers, beauticians, businessmen, etc. use OSNs to publish themselves and to raise the amount of customers they have. Third, some users use OSNs to initiate romantic relationships [4]. Surely, there is great diversity in OSNs. The recent developments in this domain contribute to the creation of many different OSNs in many aspects of users’ lives. One of the biggest social networks in the world with more than one billion users is Facebook [5] [6], but there are many more OSNs, which help users to connect based on shared interests, political views or activities such as LinkedIn, one of the world’s largest professional networks, XING⁴, a social network for business professionals, Academia.edu⁵, a social networking site for academics or researchers, Athlinks⁶, a social networking website aimed at competitive endurance athletes, and many more. The significant problem with these OSNs is the negative effects which may occur when users expose private and sensitive information about themselves, their friends, their workplaces, and other organizations. Malicious users may exploit this crack to gather valuable information about employees, workplaces, and organizations and use it against them. In 2012, Boshmaf et al. [7], described how vulnerable OSNs are to large-scale infiltration campaign run by Socialbots. Their Socialbots were designed to be able to mimic actions of real users. Their mission was to collect data on user behavior in response to a large-scale infiltration campaign. To an adversary, valuable data such as email addresses, phone numbers, and other personal data that have

¹ http://www.facebook.com
² http://www.myspace.com
³ http://www.linkedin.com
⁴ http://www.xing.com
⁵ http://academia.edu
⁶ http://athlinks.com
monetary value can be used for online profiling and large-scale email spamming and phishing campaigns. Moreover, In 2010 Kwak et al. [8], described how they crawled the entire Twitter site and obtained billions of social relations, tens of millions of user profiles, hundreds of millions of tweets, and thousands of trending topics. In 2012, Fire et al.[9], presented algorithms for constructing organization crawlers. The crawlers collected data from Facebook in order to gather public information of employees who work or worked in a specific organization. By only using publicly available data they were able to restructure parts of the targeted organization and discover hidden departments, leadership roles, etc. Furthermore, in 2012 Elyashar et al. [10], presented a method for the mining of information of organizations through the use of social networks. Their designated Socialbots sent friend requests to Facebook users who work or worked in a targeted organization. Upon accepting a Socialbot's friend request, users unknowingly expose valuable information about themselves and about their workplace. They tested their method on two organizations and as a result, were able to infiltrate them. With their method they were able to discover more employees and more informal organizational links. In this study, we were influenced by the study of Boshmaf et al. [7] regarding the infiltration of online social networks by Socialbots, from the study of Fire et al. [9], which discovered publicly valuable information about the organizations' structure, and from Elyashar et al. [10], which were able to discover more hidden links and users in targeted organizations. We combined all three approaches and used public Facebook profiles in order to infiltrate specific users in targeted organizations, but as opposed to Boshmaf et al. [7], who infiltrate different users, we tried to infiltrate to users who are employees in technology-oriented organizations. These kinds of users should be more aware of the dangers of information leakage and may avoid our infiltration. We tested the new infiltration method on targeted users from two high tech organizations. Our method was able to gain a successful percentage of 50% and 70% respectively (See Section V). These results demonstrate how easily adversaries can infiltrate users they do not know and get full access to valuable information. These results are more surprising when we emphasize the fact that for this study we purposely chose technology oriented users who should be more aware of the dangers of information leakage. Moreover, the results once again indicate that users who are interested in protecting themselves should not disclose information in OSNs and should be more cautious of accepting friendship requests from unknown persons. The remainder of this study is organized as follows. In Section II we give a brief overview of online social network studies whose issues were focused on this study. Section III describes the experimental framework and a method we used in order to infiltrate specific users in targeted organizations. Section IV presents an algorithm we developed to infiltrate specific users, and Section V presents our numeric results. Finally, in Section VI we present our conclusion and future directions.

7 https://twitter.com

II. RELATED WORK

During the last few years, there were many studies that handled the problems of users' privacy in online social networks.

In 2007 Dwyer et al. [11], focused on the issues of trust and privacy in social network sites. This issue was emphasized by a comparison between the two well-known social networks, Facebook, and MySpace. The comparison they made lead them to a conclusion that in online interaction, trust is not as necessary in the building of new relationships as it is in face to face encounters. Furthermore, they were able to show that trust and the willingness to share information do not automatically translate into new social interactions in an online site. Finally, they concluded that online relationships can develop successfully in sites where perceived trust and privacy safeguards are weak.

Also in 2007 Chau et al. [12], emphasized how easy it is to retrieve information from social networks. Chau et al. described their implementation of the crawlers for online social networks. By using the implemented crawlers, they were able to visit a total of approximately eleven million auction users, and about 66,000 of which were completely crawled.

In 2007 Mislove et al. [13], examined information gathered from four popular online social network sites. The OSNs were Orkut', YouTube9, Flickr10, and LiveJournal11. They tried to study the characteristics of online social network graphs on a large scale. Finally, they reached a data set that contained over 11.3 million users and 328 million links.

In 2009 Lindamood et al. [14], was concerned about the information revealed inside a social network. They claimed that some of the information revealed inside social networks is private. Moreover, they noted that corporations can use learning algorithms on the released data to predict undisclosed private information. In order to predict undisclosed private information about users they launched inference attacks using released data. Finally, they discussed the effectiveness of possible sanitization activities, which may help defend against inference attacks.

In 2010, Kwak et al. [8], crawled Twitter social networks and gathered 41.7 million user profiles, 1.47 billion social relations, 4,262 trending topics, and 106 million tweets. Kwak et al. made an analysis of the tweets of top trending topics. Eventually, they were able to classify the trending topics based on the active period and the tweets and showed that the majority of topics are headlining news or persistent news in nature. Moreover, Kwak et al. revealed that any re-tweeted tweet is to reach an average of 1,000 users no matter what the number of followers is of the original tweet.

In 2011, Stein et al. [15] described the Facebook Immune System (FIS). Like the human immune system, the FIS is described as a system, which protects Facebook from adversaries and malicious attacks. FIS is an adversarial

---

8 http://www.orkut.com
9 http://www.youtube.com
10 http://www.flickr.com
11 http://www.livejournal.com
learning system, which performs real-time checks and classifications on every read and writes actions on Facebook’s database. These checks exist in order to protect Facebook users and the entire social graph from malicious activities. Furthermore, Stein et al. described the design of the FIS, the challenges FIS faced, etc.

In 2011, Boshmaf et al. [16], described methods of infiltration on a targeted online social network on a large scale by Socialbots. Boshmaf et al. presented methods to build Socialbots army’s in order to infiltrate users. During their study, Boshmaf et al. operated the Socialbot on the Facebook social network for eight weeks. They were able to collect data according to users’ behavior. The results included three main conclusions. First, OSNs, such as Facebook, can be infiltrated with a success rate of up to 80%. They were able to demonstrate that the more friends a user has, the more likely the user is to accept new friendships. Second, depending on users’ privacy settings, a successful infiltration can result in privacy breaches, where even more users’ data are exposed when compared to a purely public access. Third, in practice, online social networks security defenses are not effective enough in detecting or stopping a large-scale infiltration as it occurs. In 2012, Boshmaf et al. [7], continued with their infiltration process and added two more new major conclusions. First, running a large-scale infiltration campaign might be profitable in the economics of today’s markets, but not for stable and independent businesses. Second, the protection against Socialbots raises many challenges that relate to issues such as web automation, online-offline identity binding, and usable security. Moreover, in 2012, Boshmaf et al. [17], were concerned with how easily cyber criminals can create online personas and relationships through social networks. The recent developments in the field of artificial intelligence helped adversaries design and build bots which can mimic human behavior. Bots of this kind can be used to infiltrate online communities, build up trust over time and then send personalized messages to elicit information, sway opinions, etc. Furthermore, they expect that defending against such malicious bots raises new challenges such as web automation, online-offline identity binding and usable security.

Likewise in 2012, Fire et al. [9], analyzed different types of organizations through data mining. The analysis was based on organizations’ employees who had been exposed on Facebook, LinkedIn, Google12 search results, the company’s web page and other publicly available sources. To accomplish their goal they designed and built a web crawler. The web crawler extracted a network of informal social relationships of employees of a given target organization. In contrast to standard crawlers, which were found insufficient for performing data collection because they collected many irrelevant profiles, and skipped Facebook users who worked in a target organization, the designated web crawler optimized data collection from users associated with a specific group or organization. They collected publicly available data from six well-known hi-tech companies on three different scales using the designated crawler.

In 2012, Elyashar et al. [10], presented a method for the mining of data of a targeted organization by using OSNs and Socialbots. Their method was based on accepting a Socialbot’s friend request. By accepting friend requests, users unknowingly exposed information about themselves and about their workplace. They tested the described method on two organizations. They succeeded to discover up to 13.55% more employees and up to 18.29% more informal organizational links.

Moreover, in 2012 Magdon-Ismail et al. [18], studied infiltration of trust based on a network. They used an agent, which sent friend requests. Its mission was to amass as many connections as possible. Magdon-Ismail et al. described a model for infiltration based on two properties of actors in the network. First, actors would like as many links to others as possible. Second, actors are more likely to connect to trusted nodes. From their research they established a number of conclusions. First, the trust effect is crucial. If an agent does not trust enough, then it will be difficult to infiltrate a network because of its robustness. Second, the network structure is very important. If the trust effect is minor, then it is easier to infiltrate. If the trust effect is larger, then large expansion networks are easier to infiltrate. Third, the algorithm used by the agent is crucial for the success of infiltration. Probably, random requests will be much less successful to infiltrate than complicated ones.

In our study, we attempted to infiltrate specific users in targeted organizations. We combined all three approaches. First, we started with Boshmaf et al. [7], which discussed about infiltration of OSNs via Socialbots. Like them, we used Socialbots to send friend requests to users in Facebook. However, as opposed to Boshmaf et al. who infiltrated different users without mentioning any connection between user to organization, place, etc., we tried to infiltrate users who are employees in technology-oriented organizations. This kind of user should be more aware of the dangers of information leakage and may avoid our infiltration. Secondly, Fire et al. [9], which discovered publicly valuable information about an organizations’ structure. These structures were the basis platform in which we act. The rule of thumb in our algorithm was to gather users who are employees in a targeted organization and to analyze the organization’s structure. Thirdly, Elyashar et al. [10], which were able to discover more hidden links and users in targeted organizations. We used similar methods to Elyashar et al., but this time to infiltrate specific users and not to infiltrate targeted organizations.

III. METHODS AND EXPERIMENTS

The main goal of this study was to emphasize how easily it is to infiltrate the private information of a specific organization employee. In other words, given an employee X who works in a secret organization, the probability to get access to the employees' private information increases.

For this purpose, we based our study on the Facebook social network. Facebook is one of the biggest and most

12 https://www.google.com
popular social networks in the world. It includes over one billion monthly active users as of October 2012 and 584 million daily active users on average in September 2012 [6]. Users in Facebook must create a user account, and then they are ready to create connections with existing friends as well as connecting with strangers. In many cases, a Facebook user's account may include photographs, birthday, hometown, religion, ethnicity, and personal interests [16]. In such an undirected network as Facebook, users connect to one another by initiating a friend request. The received party must accept the friend request to establish a friendship link with the requested sender. When the process of link establishment has finished, the two parties receive the privilege to access each other's profile details whenever they please. Therefore, we define accepting friend requests as an infiltration. After accepting the Socialbots' friend requests, we receive full access to the users profile and can gather valuable information about the user and the user's surroundings.

To infiltrate specific users, we had to take several actions. First, for the infiltration process, we had to crawl on targeted organizations and gather public information regarding its employees who have a Facebook user account and declared that they work or worked in the targeted organizations. For the crawling process, we created a public user account in Facebook. We used it and a crawler similar to the crawler which had been introduced by Fire et al. [9], in order to crawl the Facebook network and to find an organization's employees.

Second, at the end of the crawling process we gained sufficient information about Facebook users who work or worked in targeted organizations and its connections, i.e. intelligence on the targeted organizations' employees. We used this intelligence in order to choose ten users to be a target for infiltration randomly, and their mutual friends who will help us to infiltrate to them.

Third, in the analysis process we created a Facebook Socialbot account for every organization we would like to infiltrate. Before the infiltration of a targeted organization, we designed these user profiles to look like reliable profiles of real persons. At the beginning, we suggested friend requests to random users who had more than 1,000 friends regardless of any organizational affiliation. The idea behind this action was "...the more friends they had, the higher the chance was that they accepted a friendship request from a Socialbot (i.e., a stranger)" [16].

After a Socialbot succeeded to gain the acceptance of fifty random users to its friend requests, it automatically sent friend requests to targeted users' mutual friends who are employees in the same organization. The reason we waited for our Socialbot to gain fifty random users' friendships was because we wanted the Socialbots to look as much like real users as possible. For this reason we could not send friend requests while our Socialbot has a very small number of friends. This act could in most cases bring users to reject our Socialbot's friend requests because of the low amount of members. Users tend to be suspicious when a user with a small amount of friends requests to be their friend. The full process has been described by Elyashar et al. [10].

Fourth, our goal was to use friend requests of our Socialbots in order to become friends with specific users in the targeted organizations. In this study we defined the platform in which each Socialbot should act, i.e., a different targeted organization for every Socialbot. A Socialbot called S will try to infiltrate specific users in an organization called O, from which S currently has no friends. The main goal was not to infiltrate a new organization, but rather to infiltrate specific users inside a targeted organization.

Fifth, for each chosen targeted user, we marked their friends inside the organization and sent friend requests to each of targeted user's friends inside the organization. The idea was to gain as many mutual friends as possible. This kind of accomplishment will increase the probability that our friend request will finally be accepted by the targeted user. It is important to mention that we did not send a friend request to the targeted user until we finished sending friend requests to all its friends inside the organization.

Sixth, after we completed the process of sending friend requests to targeted users' mutual friends, we sent friend requests to the ten targeted users.

**Algorithm 1 : Socialbot Organizational Infiltration**

| Input: Uids - A set of seed URLs to Facebook profile pages of organization’s employees, S - Socialbot, O - targeted organization. OrgPublicGraph ← Organizational-Crawler(S, Uids, O) i ← 0 while (i < 10) do TU ← ChooseRandomTargetedUsers(i) i ← i + 1 end while (NumOfFriends(S) ≤ 50) SendFriendRequestToRandomUsers(S) TargetedUserFriends ← FindOrgFriends(O, TU, OrgPublicGraph) for (f ∈ TargetedUserFriends) SendFriendRequest(f) end for (TU ∈ TargetedUsers) SendFriendRequest(TU) end

It is important to emphasize that Algorithm 1 described how we infiltrate specific users in targeted organizations. In contrast to Elyashar et al. who tried to infiltrate as many users as possible by sending friend requests to targeted organization's employees who have the highest number of mutual friends in the same organization, we sent friend requests to mutual friends of the targeted users to, at last, get the friendship of the targeted users themselves.

**IV. ETHICS CONSIDERATION**

Given the nature of OSNs, we should ask ourselves a legitimate question: Is it ethical to perform such research? We believe that the answer to this question is positive for several reasons. First, similar previous researches, which were conducted by Boshmaf et al. [7] [16], received the approval of their university's behavioral research ethics board. In this research we aimed to conduct similar research with a small amount of Facebook public user accounts. Second, in this...
research we mostly focused on the structure of the social network and not on the public information, which Facebook users provided us. Moreover, we showed the structures of the social networks we found without mentioning the names of the targeted organizations. Third, as opposed to previous researches, we avoided using profile pictures, which include users’ faces for our Socialbots as much as possible. We chose profile pictures, which did not include any user's faces, but profile pictures in which it is hard to identify to whom the picture belongs. Finally, we are aware that this kind of research is problematic and has a certain degree of contravention; however, it is necessary to study the existing dangers for users in OSNs. Facebook estimated that 8.7% of its accounts are defined as fake. This means that Facebook includes 83.09 million fake accounts [19]. This enormous number should emphasize that we are faced with an acute problem which should be studied in order to be solved.

V. RESULTS

We used three Socialbots for specific user intrusion on three different organizations. To achieve these goals, our Socialbots sent friend requests to all the targeted users' mutual friends in order to gain as many mutual friends as possible so that they can eventually be accepted by the targeted users. We only presented the results we gained from the S1 and S2 Socialbots. Unfortunately, Facebook FIS disabled S3 because it sent too many friend requests and did not receive an adequate number of users that accepted these requests. The main reason for this unaccepted failure is based on the targeted organization's location and the identity of our Socialbot. In the O3 organization, most of the employees are from a foreign country, therefore many users, who received friend requests from S3, refused them. Eventually, S1 sent 124 friend requests to 124 different users (including the ten targeted users). Among them 46 users accepted, and 78 users rejected S1's requests (See Figure 1, 2).

S2 sent 114 friend requests to 114 different users (including the ten targeted users). Among them 38 users accepted, and 76 users rejected S2's requests (See Figure 3, 4).
i. O1 Intrusion Results

First, we randomly chose ten users who declared that they worked or still work in the O1 organization on their Facebook profiles. We then collected all the friends of the ten targeted users who also worked in the O1 and sent them friend requests. Next, Socialbot S1 sent friend requests to the ten targeted users (TU 1 - TU 10). In total, Socialbot S1 sent 124 friend requests and succeeded to connect to 46 different users (See Table 2). In regards to targeted users, S1 was able to be a friend of five targeted users (TU 1, TU 5, TU 6, TU 8, TU 10), with a successful percentage of 50% (See Figure 5). Moreover, S1 was able to be a friend of 37.09% of all users who received friend requests (See Table 1).

Figure 5: Users who received S1’s friend requests. Green presents user who accepted, red presents users who rejected, yellow triangle presents targeted user who accepted, and black diamond presents targeted user who rejected.

Table 1: Targeted users – summary results

<table>
<thead>
<tr>
<th>Organization</th>
<th>Targeted Users</th>
<th>Accepted/Total</th>
<th>Acceptance Percentage</th>
<th>Accepted?</th>
</tr>
</thead>
<tbody>
<tr>
<td>O1</td>
<td>TU 1</td>
<td>5/13</td>
<td>38.46%</td>
<td>Yes</td>
</tr>
<tr>
<td></td>
<td>TU 2</td>
<td>4/13</td>
<td>30.76%</td>
<td>No</td>
</tr>
<tr>
<td></td>
<td>TU 3</td>
<td>5/16</td>
<td>31.25%</td>
<td>No</td>
</tr>
<tr>
<td></td>
<td>TU 4</td>
<td>6/16</td>
<td>37.5%</td>
<td>No</td>
</tr>
<tr>
<td></td>
<td>TU 5</td>
<td>6/17</td>
<td>35.29%</td>
<td>Yes</td>
</tr>
<tr>
<td></td>
<td>TU 6</td>
<td>21/42</td>
<td>50%</td>
<td>Yes</td>
</tr>
<tr>
<td></td>
<td>TU 7</td>
<td>7/21</td>
<td>33.33%</td>
<td>No</td>
</tr>
<tr>
<td></td>
<td>TU 8</td>
<td>4/14</td>
<td>28.57%</td>
<td>Yes</td>
</tr>
<tr>
<td></td>
<td>TU 9</td>
<td>7/13</td>
<td>53.84%</td>
<td>No</td>
</tr>
<tr>
<td></td>
<td>TU 10</td>
<td>13/32</td>
<td>40.62%</td>
<td>Yes</td>
</tr>
<tr>
<td>Total</td>
<td>46/124</td>
<td></td>
<td>37.09%</td>
<td>50%</td>
</tr>
</tbody>
</table>

ii. O2 Intrusion Results

Just as in the previous process, here we also randomly chose ten users who declared, on their Facebook profiles, that they worked or still work in the O2 organization. We then collected all the friends of the ten targeted users who also worked in the O2 and sent them friend requests. Next, Socialbot S2 sent friend requests to the ten targeted users (TU 1 - TU 10). In total, Socialbot S2 sent 114 friend requests and succeeded to be connected to 38 different users (See Table 2). Regarding the targeted users, Socialbot S2 was able to be a friend of seven targeted users (TU 1, TU 2, TU 3, TU 5, TU 7, TU 9, TU 10), with a successful percentage of 70% (See Figure 6). Moreover, S2 was able to be a friend of 33.33% of all the users who got friend requests (See Table 1).

Figure 6: Users who received S2’s friend requests. Green presents user who accepted, red presents users who rejected, yellow triangle presents targeted user who accepted, and black diamond presents targeted user who rejected.

Table 2: Socialbots’ total requests

<table>
<thead>
<tr>
<th>Intrusion Users</th>
<th>Mutual Friends</th>
<th>Targeted Users Accepted/Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>S1</td>
<td>46/124</td>
<td>5/10</td>
</tr>
<tr>
<td>S2</td>
<td>38/114</td>
<td>7/10</td>
</tr>
</tbody>
</table>
would be more difficult to infiltrate privacy-oriented users; however, our two Socialbots succeeded in accumulating 50% and 70% of the targeted users respectively. This surprising fact emphasizes that even people who are aware to information leakage dangers are exposed themselves and so are their organizations. Fifth, we can conclude that there is a huge link between mutual friends and the percentage of acceptance to friend requests. There is no doubt that if our Socialbots sent friend requests to targeted users in targeted organizations without gaining friendships of their mutual friends who work in the same organization, the percentage of acceptance was decreased. Sixth, according to a correlation test between the amount of mutual friends who accepted friend requests and targeted user decisions, we can conclude that S1 was able to infiltrate a targeted user when it gains more than six mutual friends (See Figure 7). S2 was able to do the same when it gains seven or more mutual friends (See Figure 8). We can conclude that the more that mutual friends accept the Socialbot's requests, the more likely the targeted user is to accept Socialbot's friend request. Furthermore, most of our conclusions are consistent with the conclusions that Boshmaf et al. have reached.

We believe that this study has several future research directions. A possible direction is to use the same described algorithm, but this time on more than two organizations in order to examine our conclusions. One more possible direction can be a study, which is focused on the question "What is the minimal amount of mutual friend's that a Socialbot should gain in order to infiltrate specific user?"

REFERENCES


