The CSI Journal on
Computer Science and Engineering
Vol. 14, No. 2, 2017
Pages 11-20
Regular Paper

A New Automatic Method for ICA Detection in Online Social Networks

Maryam Zare Seyed Hossein Khasteh Seyed Ali Khoshroo

Faculty of Computer Engineering, K. N. Toosi University of Technology, Tehran, Iran

Abstract

Online information of social network users, which is shared inside these networks, has started to be interesting for malicious users. One of the threats to users` personal information is called Identity Clone Attack (ICA). In this type of attacks, the malicious user targets an online social network user as a victim and creates a fake but similar profile to his profile. In ICA, most of the time, famous and beloved users of the society like celebrities and politicians are being preyed on. This paper presents a new automated method for detecting fake profiles cloned from celebrities' profiles. In this method, we first cluster the profiles. Subsequently, the profiles which are in the same cluster as the victim's profile with a similarity to the victim`s profile over a predefined threshold are moved to the next phase as suspicious profiles. Then a parameter named "Celebrities' Network" is extracted for the suspected and the victim's profiles. "Celebrities' Network" for a user is a sub graph of the network, whose nodes are popular users following that user. The profile for which this parameter is the highest is considered as the real profile. This is an easily applicable method as it does not require a human agent and extra information. Besides, initial clustering helps the system to save time in searching. The suggested method was applied to face book and Instagram datasets and approximately 100% of the cloned profiles were detected. These findings show that this method is quite promising and the results are comparable to the best studies conducted on ICA.

Keywords: Online Social Network, Fake Profile, ICA, Threat, Popular Users.

1. Introduction

In recent years, online social network websites like Facebook, Twitter, and My space have become very popular among users. Most of the users join these online social networks to connect and share information with the people they know. Sharing the information is actually the reason why they are exposed to various security threats. Different types of security threats that users may encounter are discussed in the following section.

1.1. Stealing Users' Personal Information

In this attack, targets are users' personal information, friends and shared information like posts and status updates. Social Phishing Attack is one example of this type. After the attacker finds the victim's email address from his page on an online social network, he builds a fake website similar to the existing valid ones. He then sends an email to the victim stating that he needs to authenticate his account in the new website; otherwise the account would get blocked or removed. After the victim is lured to the fake website, he gets asked to enter his username and password. As a result, the attacker gets the victim's information and reaches his goal. A lot of studies have been conducted for detecting this kind of attack. The simplest method employed in these studies involves analyzing the emails sent by the attacker. Some words like "login" or "verify" in the body of the email, a link including random and long phrases, or difference between the domains of the sender's address and the link can help detect this kind of attack [1].

Another example is Identity Clone Attack (ICA). In ICA, the attacker abuses the friendship between users and their careless acceptance of friendship requests and forges his

targeted user's profile [2]. This problem is the focus of this paper and methods for detecting this attack is discussed in details in Section 2.

ICA involves a malicious user trying to create a fake profile of the victim on an online social network, connecting with his friends and convincing them that this is actually an authentic profile to get their personal information which is not otherwise publicly available. To this end, the malicious user tries to gather victim's personal information like his name, his living place and workplace, his friends list from his online social network profile or his home page on the internet. He then forges his profile and creates a similar or an identical profile to that and finally sends friendship requests to the victim's friends list [2]. Figure 1 depicts this scenario. When victim's friends list is not available to everyone, the attacker first creates a profile and sends a friendship request to him. Experience has shown that most online social networks users will accept any requests heedless of danger. As soon as the victim accepts the request, the attacker gets access to his personal information and friends list. Finally, the attacker wins and reaches his goal of finding the victim's friends list. He subsequently uses that information to turn his profile into one identical to the victim's and sends friendship requests to all the people in the victim's friends list. This threat can affect popularity, credibility, and value of the resources on an online social network including people, groups, shared applications, etc. It can also have direct effects on choosing friends and establishing a friendship with the victim's friend by building a trust on the victim's faked profile.

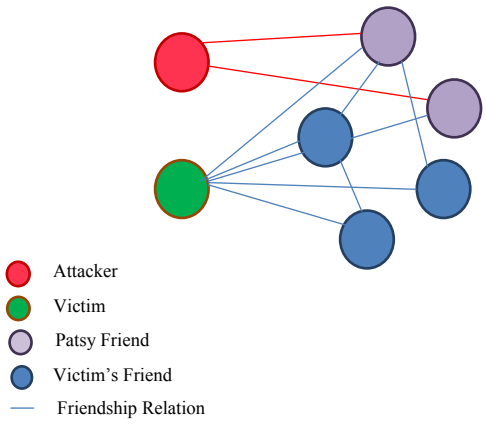

Figure 1. Phenomenon of Identity Cloning

The main purpose of the attacker in this threat is stealing victim's personal information and his friends by cloning his profile and increasing his credibility with the victim's friends to achieve his malicious goals in the future.

Two types of ICAs are defined: single-site profile cloning attack and cross-site profile cloning attack. In the former, the malicious user clones the victim's profile in the same online social network as the victim's and sends friend request to his friends. On the other hand, in cross-site profile cloning attack, the malicious user creates a fake profile with similar attributes as victim's in another online social network, in which the victim does not have a profile yet, and sends friend requests to his friends who are on both online social networks or new people and pretends to be the victim. The victim's friends that suppose this request comes from their friend will accept that and, in so doing, become vulnerable to this attack, which is very difficult to detect, due to the fact that website administrator or service provider believe that the created profile belongs to a new user [3].

1.2. Building Credibility in Online Social Networks

In this scenario, the malicious user creates many fake profiles or nodes (Sybil Node). One of these profiles acts as a master as is the case in a master-slave structure and the other profiles are controlled by the attacker as followers. The created profiles are connected together and also can connect to other users of the network. Whenever the master node posts something on its own page or sends a comment to a group, it receives support from the followers. This behavior increases reputation of the master profile amongst other users of the online social network and attracts other users to this profile [4].

Past research on detecting this type of attack involves analyzing users' behavior. Normal users behave very differently on the network. Activities like viewing friends' profile, changing profile picture, viewing friends' photos, playing games, posting updates, joining groups, and many other social activities are done by normal users. On the other hand, Sybil nodes are created for some limited activities. The purpose of these nodes is publishing spams. Therefore, action sequences of these nodes are predictable and detection of this attack is feasible [5].

1.3. Spread of Spam Emails

Spams are any poor quality, inappropriate or unwanted content that a user might receive and deal with like text messages or fake profiles. Most spams on online social networks include advertisements or links that try to destruct users' perspective on an issue, social event, popular individuals, or a commercial product. As such, a user might click on and get lured into a fake or unethical websites. Lots of studies have also been conducted on detecting spams like any other fraud. Some of these studies extract the texts inside the messages to detect whether or not it is a spam [6]. For example, in an online store like Amazon, some features like commenters` IP address (which increases the probability of being a spam if many of them come from the same address), time of comments, the comparison of sales and other information about a product, and the number of positive and negative comments about a product could help detecting spam messages posted by malicious users to increase the sale of a product [7].

2. Related Works

This section will present some of the previous studies that have been carried out on detecting cloned profiles based on attribute similarity, friends' network similarity, IP addresses, and users' click pattern. Generally, in initial phases of ICA, attributes of the victim's profile and the suspicious profile (s) will be compared using different similarity measures including Cosine Similarity or Jacard Similarity and whenever the measured value for that specific factor goes

over a predefined threshold, the detection procedure will continue.

2.1. Attribute Similarity

One of the studies in this field is a system to recognize fake profiles based on similarities between profiles attributes on LinkedIn online social network [8]. This system was composed of three parts: Information Distiller, Profile Hunter, and Profile Verifier. Information Distiller extracts key information from the input profile which could be specific for that profile such as email address (excluding information like birth place or nationality that are common among users). Finding cloned profiles would be easier after this. Profile Hunter will use this information to find similar profiles and all the resulting profiles will be passed to the Profile Verifier. This section measures similarities between the input profile and the suspected ones. Finally, all the profiles similar to the input profile will be shown to the user with their measured similarity. This system needs a human confirmation to make the final decision. Also, in this system, the input profile is supposed to be the real valid profile and is not a clone itself. Profile picture comparison is also possible in the suggested system to increase the accuracy of detection. By experimenting this structure on LinkedIn online social network, authors reached the accuracy of 100%.

In a similar study, attribute similarity measure was used to identify fake profiles [9]. In this study, similar profiles were first extracted and the similarity measure was calculated. The relationship strength measure between the victim, suspicious users, and their mutual friends was then used and the profile with a stronger social relationship was considered as the valid user.

Although employing this method to identify fake profiles based on users' social relations is simple, it needs a human agent or a mutual friend who can design question about their friendship history and the user who can answer all questions correctly is recognized as the valid user. In other words, the final verification involves a human agent and is not done by the method discussed. Another study by [3] employed a similar method for ICA detection and only changed the equations for calculating attribute similarity between two profiles, which also needed a human confirmation for the final verification.

2.2. Similarity of Friends' Network

Another method for ICA detection was suggested by [2]. In addition to the comparison of the two profiles` attributes, this method takes advantage of friends' network of the profiles and if these are similar, an ICA will happen with high probability.

In this study, authors assumed that users have three types of lists: Friends List (FL), containing the users who are friends with the user, Recommended Friends List (RFL), the online social network recommendations based on the similarities between the users` attributes or interests, and Excluded Friends List (EFL), which contain users that the user don't want them in his/her friends list. Teenagers use this list to add their parents or employees in to prevent a friendship with them. In case, a malicious user knows enough about the victim, he can send requests to the people in his excluded list. This makes the cloned profile more valid if these requests get accepted. The detection procedure is illustrated in figure 2.

First, a profile as victim and a set of profiles from the network are entered into the system as input. Then, profiles with the same first and last name as victim's are taken to the next phase as candidate profiles. The third phase measures the similarity between the victim's profile and each candidate profile based on attributes similarity and resemblance of friends' network. The profiles whose similarity factor is over a predefined threshold are moved to the verification phase as suspected profiles.

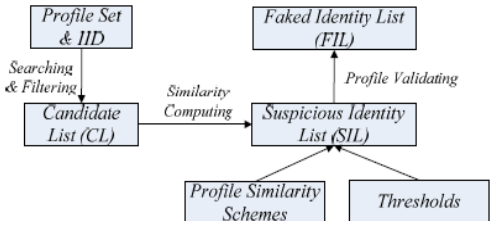

Figure 2. ICA detection process [2]

The input profile is also needed in the verification phase. In this phase, the valid profiles should be identified. To this end, a mutual friend of the victim and the suspected profiles asked questions about their friendship in the past, memories, etc. The user who answers all the questions correctly is recognized as the real user. Finally, all the profiles classified as clone are blocked temporarily and friends of these profiles get warning messages. This method uses friends' network resemblance in addition to attribute similarity method to improve accuracy. However, like previous methods, it needs human judgment in the final stage and the whole recognition system isn`t taken care of automatically.

In another study, [10], used this friends' network similarity to detect ICA. The only difference between their work and the method suggested by [3] is the attribute similarity measure in profiles. In this method like all the previous ones, a human agent is needed for final verification or a mutual friend needs to question about their friendship history and the one with correct answers is recognized as the valid user.

2.3. IP Address Similarity

In addition to the studies mentioned above, another research was conducted by [11] on a Chinese online social network called Ranren. Their suggested tool, "Clone Spotter", is used in servers of online social networks. This recommended tool makes decision based on physical information and real communications. As an example, when user A sends a friendship request to user B, this tool checks whether or not user B has a user C, which has the same name as A in its friends list. If a and C have the same names, similarity is measured for the rest of their profile attributes. If the similarity is higher than a certain threshold, Clone Spotter compares the first 16 bits of the IP addresses of A and C. If both prefixes are alike, it means that both profiles belong to a single person. This situation happens when a user tends to have two accounts in the network. If prefixes are different, that means owners of A and C are not affiliated with each other and in fact an ICA has occurred.

Here the whole detection procedure is automatic and human intervention or a mutual friend is not essential as was the case with previous studies. This is the greatest advantage of this method. It also has better accuracy because of the IP address inspection. However, it needs to be noted that running on online social networks servers for every single friendship requests sent from one user to another, places a heavy burden on servers which can be considered as the downside of the technique.

An optimized version of Clone Spotter was introduced by [12]. In this optimized tool, instead of using IP address, users' click pattern and their activity periods in the online social network are used, which lead to improved accuracy of this tool. This method, like the previous one, doesn't need a human agent for final verification and detection is automated. Despite having a better accuracy and being automatic, this system still puts a lot of load on servers.

Besides online social networks, other networks such as wireless sensor networks also suffer from this threat. Cloning or node replication attack in these networks is almost like ICA in online social networks. This attack happens when an attacker first tries to physically capture victim nodes and after gathering sensitive information such as ID, location information, etc. creates his own low-cost sensor nodes and tries to pass them off as legal nodes.

Wireless sensor networks can be stationary (static) or mobile and a lot of studies have been conducted to detect this type of attack in these networks. The detection process varies depending on WSN type.

Generally, detection methods are divided into two categories: distributed and centralized techniques. In centralized techniques, base station or sink node is responsible for decision making. All nodes in WSN send their ID and location information to it through their neighbors and sink node checks the information received and if it finds two different locations and the same ID, the confliction and replication node attack has occurred[13].Some studies have used this technique (e.g. [14, 15, 16]).

In distributed techniques, there is no return information to a sink node and a special detection method called claimer-reporter-witness is applied. In this technique, the location information is sent to randomly selected node named witness node. Some studies have employed this method to detect attacks (e.g. [17, 18, 19]).

3. Proposed Method

In this part, the process of ICA detection based on the proposed method is described. This process, consists of three main steps. The first step is clustering users. In the second step, candidate profiles are extracted from the victim's cluster and their similarities is measured. Detection of ICA is done in the last step. In this step, "Celebrities, Network" or "CN" for suspected profiles is extracted and accordingly fake profiles are separated from the real ones.CN for a user is a sub graph of the network whose nodes are popular users following that user. Popular individuals of the society in online social networks have lots of followers, but they don't follow just any user. In effect, to maintain their credibility, they don`t follow a user unless they are sure about his credibility. We use this concept and present a new automatic method for detecting these popular users` cloned profiles. In the suggested method, after extracting CN for the victim's and suspected profiles, the real profile will be the one with more celebrities following it. This method has some advantages over the methods employed in previous studies and its results are very promising. In this detection technique, instead of searching the whole network to extract profiles with the same name as the victim's, it is limited to victim's cluster which saves us a lot of time. The main advantage of the suggested method is the automatic detection of fake users. In previous studies, a human agent was needed for final verification (e.g. [2, 3, 8, 9, 10]). It also turns out to be better compared to the study conducted by [11]. In this study, there is no need for human intervention and the suggested method is run for every friend request that users send to each other on anon line social network. Due to easy access to internet and the ever-growing popularity of online social networks, the number of users is increasing. This in turn results in higher rates of friend requests sent and received and more communication among users in general. So it would impose a huge overhead on the server. But in our method, this won't be a problem and there is a low computational overload. The proposed method doesn't need any extra information for detection of ICA and detection process is performed using only the network structure. The three main steps of the method are as follows:

3.1. Clustering the Profiles

In the first Step, profiles are clustered according to the number of their followers. To do so, for every profile in the network, the number of followers is counted and saved as its attribute. Then, by applying one-featured K-means algorithm based on this trait, profiles with approximately the same number of followers are put in a single cluster. We know that famous and renowned people in online social networks have a lot of followers. Therefore, popular people's profiles in the network and their cloned profiles that also have a large number of followers, would appear in the same group. Otherwise, a fake profile could be detected very easily. Given that celebrities have a large number of followers, if a profile is known and doesn't have many followers, its credibility must be questioned.

```
Input: Social Network Graph
N: Set of Nodes in Social Network Graph
Followers Count: Number of Node's Followers
Output: Cluster ID of each Node (ID)

---------------------------------------------------

For each node u in N
    For each node m in N
      If m follows u then
            Followers Count (u) = Followers Count (u) + 1
      End if
    End for
End for
ID = kmeans (Followers Count)
```

Figure 3. The Pseudo Code for the First Step of the Suggested Method (Clustering)

Obviously, such a phenomenon does not occur. When a deceiver forges a celebrity profile, a lot of profiles will follow it automatically. In other words, if a profile is famous

then it has a lot of followers. After the clustering, the rest of clone profile detection is facilitated and it is only limited to a subgroup of profiles. The pseudo code for the first step of the suggested method is available in figure 3.

3.2. Suspicious Profiles Extraction

In this step, we consider a celebrity's profile with a lot of followers as a victim. According to our assumption, any profile whose followers count is greater than a threshold level is a celebrity profile. Note that this profile is selected arbitrary. We don't know whether or not this profile is real and we plan to check the real identity of this profile. All profiles that are not located in the victim's cluster are discarded. The victim's profile and other profiles in his cluster are entered as input into this step of algorithm. Then profiles with the same first name and last name as those of the victim are extracted. These are considered as candidate profiles. Then cosine similarity measure between the victim's profile and each candidate profile (the same equation given in 3.2 was used for this purpose) is calculated. The attributes that are involved in the calculation of similarity measure are common features of users' profiles such as first name, last name, age, education, location, marital status, nationality, and job. Attribute similarity of the two profiles, S_{att}, based on Cosine similarity is defined as equation 3.2:

$$S_{att}(P_c . P_v) = \frac{SA_{cv}}{\sqrt{|A_c| \times |A_v|}} \quad (3.2)$$

If the similarity gets over a certain threshold, these profiles are move to the next phase as suspected profiles for verification. Figure 4 illustrates the pseudo code for this step of the suggested method.

Input: Profiles in Victim's Cluster
P_v: The Victim's Profile
Threshold of Similarity
Candidate List: Candidate profiles with the same name and last name as P_v in the Victim's Cluster
Output: Suspicious profiles the similarity measure of which is larger than a pre-defined threshold for them (Suspicious List)

For each profile u in Candidate List
Similarity = Cosine similarity Value between u and P_v
If Similarity > predefined threshold then
Suspicious List = Suspicious List + u
End if
End for

Figure 4. Extraction of Supicious Profiles

3.3. Detection of Cloned Profiles

Main task of our method is done in this step. To detect cloned profiles, CN for victim and all the other suspicious profiles that were discovered in the previous step is extracted. "Celebrities Network" or "CN" for a profile is a sub graph of the network, whose nodes are popular profiles following it. In this paper, we suppose that any profile whose followers are greater than a threshold is popular. CN includes only celebrity's profiles. So, from all the profiles that follow a profile, celebrity ones are selected.

Popular individuals of the society who are in online social networks have lots of users following them, but these individuals don't follow just any user. It means that to keep their credibility, they won't follow a user unless they are sure about his credibility. For example, a famous person like "Steve Jobs", who is pretty popular among a lot of people of the society and his page has many followers in online social networks, follows only few pages. Because following a user or a page by "Steve Jobs", despite reflecting his interests, could also change his followers' opinion about him. Keeping all these in mind, in the last phase of detection, a profile which has more popular people following it, would be recognized as a valid and real profile from within suspected and the victim's profile, and the rest are considered cloned. Later on, a warning message will be sent to friends and people following these fake profiles.

Input: Suspicious List
P_v
Output: Cloned Profiles

CN = Celebrities Network for victim
CN _ P_v = Number of Nodes in CN
For each profile P_s in Suspicious List
CN = Celebrities Network for P_s
CN_P_s (P_s) = Number of Nodes in CN
End for
CNs = Merge CN_P_s with CN _ P_v
Max CN = Max(CNs)
Real Profile = Profile With the max CN value
Cloned Profiles = Other Profiles else real Profile in Suspicious List

Send a warning message to followers and friends of other Profile in Suspicious List: "Your Friend/Follower is a Cloned Profile".

Figure 5. Detection of Cloned Profiles in Suspicious List

Figure 5 illustrates the pseudo code for the third step of the suggested method.

4. Experimental Results

In this section, we present our experimental results to show how the suggested method can be used in ICA detection. In this paper, the offline datasets from Facebook [20] and Instagram [21] online social network were used. The number of users and their links are shown in table 1. In these datasets, each user has 13 attributes e.g. first name, last name, age, academic and work history, place of living, nationality, and so on. Links on both networks are directed and half duplex.

Table 1. Users and their relationships count in tested datasets

Dataset	Users	Directed Links
Facebook	45,222	1,469,518
Instagram	34,048	281,080

Due to the inaccessibility of real datasets in which fake profiles have been marked [2, 9, 10], the test datasets are constructed and used in this way.

First, famous profiles of each networks were selected and for each of them, some cloned profiles were created and added to the network. These profiles were different from the

real ones in only a few attributes. For each fake profile, we produced a random number C between 1 to 4. Then we omit C attributes from each profile`s attributes. To connect these fake profiles to the graph of the online social network, a random follower Count number was produced for each of the cloned profiles. The fake profiles then follow as many as follower Count other profiles in the network. Also, another random number follower Count was produced and which is indicative of the number of followers for each fake profile. To produce follower Count value, the threshold that leads to popularity was considered since these profiles were cloned from popular profiles and should have many followers as the real ones.

Afterwards, the first step of detection process was performed. In this step, after calculating the number of followers for each profile, they were clustered by the K-means algorithm. The output of this step is the cluster ID of each profile in the network. Then, the second phase of the process was run to extract the suspected profiles. In this step, Cosine similarity measure was calculated between the victim's profile and those users' profiles that had the same first name and last name and sat in the victim's cluster. In every running of algorithm, we provided the victim's profile as the input of this step. Output is the suspicious profiles that are very similar to the victim's profile. In the last phase, CN was extracted for the victim and suspicious profiles. In this method, the number of profiles' followers is considered as the measure of popularity. We assume that each profile with lots of followers is highly likely to be a famous and beloved person. Therefore, CN for a given profile refers to the popular users with a lot of followers who follow that profile. Hence, each victim or suspicious user with more popular followers were recognized as real profiles. When a profile was recognized as cloned, a warning message was sent to its followers and friends. The output of the last step constitutes the cloned profiles.

Table 2. Experimental results for executing the suggested method on Facebook dataset, K = 2

Threshold of Similarity	S = 1	S = 5	S = 10	S = 20	S = 100
	Accuracy	Accuracy	Accuracy	Accuracy	Accuracy
0.5	97%	92%	83%	72%	52%
0.6	97%	92%	83%	72%	50%
0.7	95%	87%	80%	67%	49%
0.8	93%	86%	76%	61%	49%
0.9	90%	82%	73%	59%	47%

Table 3. Experimental results for executing the suggested method on Facebook dataset, K = 3

Threshold of Similarity	S = 1	S = 5	S = 10	S = 20	S = 100
	Accuracy	Accuracy	Accuracy	Accuracy	Accuracy
0.5	98%	95%	83%	70%	55%
0.6	98%	95%	84%	70%	55%
0.7	95%	91%	82%	66%	44%
0.8	94%	90%	78%	58%	46%
0.9	88%	85%	73%	54%	42%

Table 4. Experimental results for executing the suggested method on Facebook dataset, K = 5

Threshold of Similarity	S = 1	S = 5	S = 10	S = 20	S = 100
	Accuracy	Accuracy	Accuracy	Accuracy	Accuracy
0.5	95%	94.3%	87%	69%	57%
0.6	95%	95 %	87%	69%	57%
0.7	94%	94%	84%	64%	52%
0.8	92%	88%	79%	61%	49%
0.9	89.5%	88%	78%	60%	45%

Table 5. Experimental results for executing the suggested method on Instagram dataset, K = 2

Threshold of Similarity	S = 1	S = 5	S = 10	S = 20	S = 100
	Accuracy	Accuracy	Accuracy	Accuracy	Accuracy
0.5	97%	95%	88%	85%	53%
0.6	97%	95%	88%	85%	54%
0.7	96%	94%	86%	79%	50%
0.8	95%	92%	83%	74%	50%
0.9	93%	89%	80%	76%	47 %

Table 6. Experimental results for executing the suggested method on Instagram dataset, K = 3

Threshold of Similarity	S = 1	S = 5	S = 10	S = 20	S = 100
	Accuracy	Accuracy	Accuracy	Accuracy	Accuracy
0.5	96%	94%	88%	81%	60%
0.6	96%	94%	88%	81%	60%
0.7	95%	93%	85%	77%	56%
0.8	95%	91%	78%	76%	41%
0.9	92%	90%	78%	75%	44%

Table 7. - Experimental results for executing the suggested method on Instagram dataset, K = 5

Threshold of Similarity	S = 1	S = 5	S = 10	S = 20	S = 100
	Accuracy	Accuracy	Accuracy	Accuracy	Accuracy
0.5	94%	93%	89%	68%	49%
0.6	95%	93%	89%	68%	49%
0.7	93%	91%	85%	66%	44%
0.8	93%	90%	84%	60%	42%
0.9	91%	88%	81%	61%	43%

Table 8. Results of Clone Spotter and Enhanced Clone Spotter experiments

Attack Pattern	Clone Spotter					Enhanced Clone Spotter				
	t_1	t_2	t_3	t_4	t_5	t_1	t_2	t_3	t_4	t_5
Accepted Request	5	6	8	12	14	5	6	8	12	14
Clone Request	2	0	1	3	4	2	0	1	3	4
Clone Detector	0	0	1	2	3	2	0	1	2	4

We ran our proposed method 10 times. The mean results are presented in table 2 to table 7. In these tables, the parameter S is the number of cloned profiles for each of the popular user in the network. K is the number of clusters in K-means and Accuracy parameter shows the detection accuracy of the algorithms.

In this method, we ran the algorithm with various threshold levels from 0.5 to 0.9. It should be noted that the minimum level for similarity factor of the two profiles in this study was 0.6 and the maximum level was 1. As was mentioned earlier, to save time, we used clustering algorithm based on the number of users' followers. The suggested method was also run with different K's in K-means algorithm. Past research in clustering field has shown that determining a precise and optimal value for K isn't an easy task and hence, we chose some values empirically. According to the results, when the K parameter is larger than 5, the accuracy decreases. When the number of clusters increases in a network, it is quite likely that recently created cloned profiles with fewer followers sit in a cluster other than the victim`s cluster. Therefore, the proposed algorithm doesn`t consider them in the remaining detection stages. This is a shortcoming of the method which highlights the need for a more precise clustering method.

Moreover, in the experiments in which the S parameter is held constant, fewer profiles gets over the threshold level and are sent to the next phase when the threshold level increases. But, all of these cloned profiles that pass all the phase in algorithm, are detected in the last phase. So our method is sensitive to both K and similarity threshold. When these parameters increased, the accuracy decreased. To examine the efficiency of the algorithm in detection process, data was gathered from different cases of each popular user with one or more cloned profiles. According to the results, when there are few cloned profiles, algorithm differentiates the real profile from clones quite precisely. However, as the number of cloned profiles increases for each famous user, the accuracy of the algorithm decreases. So, if a group of malicious users use this method as an opening to other harmful threats like Sybil Attack, the detection percentage of the suggested method in this situation will fall.

Because this method is an automatic detection method, we compared it to 2 other automatic methods Clone Spotter and Enhanced Clone Spotter proposed by [11] and [12], respectively. To detect ICA, their proposed algorithm was run per friend request from a user to another at a time. [12] designed an experiment to compare two Clone Spotter tools. They considered 5 different time slices and ran both methods for each friend request in those intervals. The number of clone requests varied from 0 to 4 among time slices. But they didn't report the threshold of attribute similarity. Thus, we consider the experiments in which we have 5 cloned profiles per user. The result of their experiments is presented in table 8.

To compare the results of the method proposed in this study with those of the two above-mentioned methods

(i.e. Clone Spotter and Enhanced Clone Spotter), a two-sample T-test was run. The null hypothesis (H_0) posited that there was not any significant difference between the results obtained from our method and those obtained using Enhanced Clone Spotter and Clone Spotter. The alternative hypothesis (H_1), however, stated that the results of our method is superior to those from these methods:

$$H_0: \mu1 - \mu2 = 0$$
$$H_1: \mu1 - \mu2 > 0$$

This test was used to reject the null hypothesis and to prove the alternative hypothesis (H_1) .

T-test was run to calculate the t statistics based on the following equations:

$$t = \frac{\mu1 - \mu2 - \Delta}{\sqrt{\frac{s1}{n1} + \frac{s2}{n2}}}$$

Where μ1 is the mean results obtained from our proposed method, μ2 is the mean samples obtained from Enhanced Clone Spotter and Clone Spotter, and Δ is 0 given that the null hypothesis posits that the two means are not significantly different. n1 and n2 are the number of samples of the first and second groups and s1 and s2 are the variance of the samples of the two groups, respectively. As was mentioned earlier, we repeated the test 10 times. As such, n1 = 10 and, according to table 8, n2 = 5, and, therefore, the degree of freedom is n2-1 = 4. Also, α value was set at 0.05. Furthermore, to run the T-test for this method, the K was set at 3, similarity threshold = 0.6, and S = 5. The T-test was run in four modes. The results are presented in table 9. As shown in the T-table, t-statistic is 2.13.

Table 9. Calculated t-statistic value

	Proposed Method on Facebook Dataset	**Proposed Method on Instagram Dataset**
Clone Spotter	**2.37**	**2.3**
Enhanced Clone Spotter	**1.73**	**1.12**

According to table 9, the t-statistic calculated for the samples of the method proposed in this study and the samples of Clone Spotter is larger than the value form the T-table (2.13). Thus, it can be concluded that H1 is accepted and H0 is rejected with a confidence level of 95%. In other words, our proposed method was significantly more efficient compared to Clone Spotter. Moreover, the t-statistic calculated for the samples of our method and the samples of Enhanced Clone Spotter is smaller than the value form the T-table (2.13). Therefore, the H0is accepted and H_1 is rejected with a confidence level of 95%. This can be taken to mean that the proposed methods and Enhanced Clone Spotter are equally efficient.

5. Conclusion and Future Work

An online social network, with all its benefits and applications, have its own special flaws. In this paper, one of these flaws, namely Identity Cloning Attack, its various types, and the people who are more likely to face this threat were discussed. A three step process for detecting this attack and extracting cloned profiles was then described. In this method, the users were clustered by K-means based on the number of their followers in the first phase so as to reduce the search time. Then the profiles in the same cluster as the victim's cluster and with a similarity higher than a predefined threshold to the victim's profile were sent to the final phase. In the last phase, to complete the final recognition of the cloned profiles, a new method was employed which was fast, automatic, without the need of human agent, simple and applicable. At the end, the experimental results of this process on two datasets of two popular online social networks and its features and accuracy on tracking ICA were presented. The main advantage of our method is that it does not need the judgement of a human agent for final verification like some previous studies and detects cloned profiles automatically. Besides, this method saves time in searching as it clusters users.

Moreover, this method doesn`t need any extra information like user's IP address or click pattern and the detection process is performed only using network structure and users' profiles. According to the experimental results, more than 90% of cloned profiles were detected on both online social networks when a small number of the cloned profiles of a single popular user were created. So our method is comparable with other studies. This algorithm, however, has its own weaknesses. When a large number of cloned profiles of a single popular user is created, e.g. 100 (this value is related to the type and internal links of online social networks) cloned profile in our test, its accuracy drops. In this paper, the parameter showing users popularity was the number of their followers.

Deciding on a threshold for this could be difficult. K-means algorithm is used for clustering in this paper. It is also not easy to figure out the number of clusters. Future work can be carried out to improve this method through using a different factor for users` popularity on online social networks. For example, we can use Google's page rank algorithm. According to this algorithm, the profiles which are more popular and credible have higher ranks. Also, we can use different clustering algorithms and compare the results with this method. Given the growing number of online social networks and their users, it is essential to implement online social networks analysis based on big data techniques and this can be a future work to improve the performance of the method suggested in this study.

References

[1] R. Banest, S. Mukkamala, and A. H. Sung, "Detection of Phishing Attacks: A Machine Learning Approach," *Soft Computing Applications in Industry, STUDFUZZ 226*, vol. 226, pp. 373-383, 2008.

[2] L. Jin, B. D. J. Joshi, and H. Takabi, "Towards Active Detection of Identity Clone Attacks on Online Social Networks," *Proceedings of the 1st ACM conference on Data and application security and privacy*, USA, pp. 27-38, 2011.

[3] S. RiziF, Y. M. Kharaji, and R. KhayyambashiM, "A New Approach for Finding Cloned Profiles in Online Social Networks," *International Journal of Network Security*, vol. 6, pp. 25-37, 2014.

[4] S. M. Shariff, and X. Zhang, "A Survey on Deceptions in Online Social Networks," *International Conference on Computer and Information Sciences (ICCOINS)*, 2014.

[5] K. Zhang, "Sybil attacks and their defenses in the Internet of things," *IEEE Internet of Things Journal*, vol. 1, no.5, pp. 372-383, 2014.

[6] L. Song, R. C. W. Kwok, S. S. Y. Liao, and W. Zhang, "A Critical Analysis of the State-Of-The-Art on Automated Detection of Deceptive Behavior in Social Media," *Proceedings of the Pacific Asia Conference on Information Systems*, Paper 168, pp. 1-15, 2012.

[7] A. Heydari, Z. Heydari, N. Salim, and M. A. Tavakoi, "Detection of review spam: A survey," *Expert Systems with Applications 42*, pp. 3634-3642, 2015.

[8] G. Kontaxis, S. Ioannidis, E. P. Markatos, and I. Polakis, "Detecting Social Network Profile Cloning," *Proceedings of IEEE International Conference on Pervasive Computing and Communications*, pp. 295-300, 2011.

[9] M. Y. Kharaji, and F. S. Rizi, "An IAC Approach for Detecting Profile Cloning in Online Social Networks," *International Journal of Network Security & Its Applications (IJNSA)*, vol. 6, no. 1, pp. 75-90, 2014.

[10] M. R. Khayyambashi, and F. S. Rizi, "An Approach for Detecting Profile Cloning in Online Social Networks," In *Proceedings of the 7th Internationl Conference on e-Commerce in Developing Countries with Focus on e-Security*, pp. 1-12, 2013.

[11] Z. Shan Z, H. Caao, and J. Lv, "Enhancing and Identifying Cloning Attacks in Online Social Networks," *Proceedings of the 7th international conference on Ubiquitous Information Management and Communication, ACM, Kota Kinabalu, Malaysia*, no. 59, pp 17-19, 2013.

[12] S. Kiruthiga, A. Kannan, and S. P. Kola, "Detecting Cloning Attack in Social Networks Using Classification and Clustering Techniques," International *Conference on Recent Trends in Information Technology*, 2014.

[13] W. Z. Khan, M. Y. Aalsalem, M. N. B. M. Saad, and Y. Xiang, "Detection and Mitigation of Node Replication Attacks in Wireless Sensor Networks: A Survey," *International Journal of Distributed Sensor Networks*, no. 149023, pp. 1-22, 2013.

[14] H. Choi, T. F. L. Porta, and S. Zhu, "SET: Detecting Node Clones in Sensor Networks," *Proceedings of the 3rd International Conference on Security and Privacy in Communication Networks (Secure Comm '07)*, pp. 341-350, 2007.

[15] R. Brooks, P. Y. Govindaraju, M. T. Kandemir, M. Pirretti, and N. Vijaykrishnan, "On The Detection of Clones in Sensor Networks Using Random key Predistribution," *IEEE Transactions on Systems, Man and Cybernetics C*, vol. 37, no. 6, pp. 1246–1258, 2007.

[16] K. Xing, X. Cheng, D. H. C. Du, and F. Liu, "Real-time Detection of Clone Attacks in Wireless Sensor Networks," *Proceedings of the 28th International Conference on Distributed Computing Systems (ICDCS '08)*, pp. 3-10, 2008.

[17] B. Zhu, V. G. K. Addada, S. Jajodia, S. Roy, and S. Setia, "Efficient Distributed Detection of Node Replication Attacks in Sensor Networks," *Proceedings of the 23rd Annual Computer Security Applications Conference (ACSAC '07)*, pp. 257-266, 2007.

[18] B. Zhu, S. Jajodia, S. Roy, S. Setia, and L. Wang, "Localizedmulticast: Efficient and Distributed Replica Detection in Large-Scale Sensor Networks," *IEEE Transactions on Mobile Computing*, vol. 9, no. 7, pp. 913-926, 2007.

[19] L. C. Ko, H. Y. Chen, and G. R. Lin, "Aneighbor-based Detection Scheme for Wireless Sensor Networks Against Node Replication Attacks," *Proceedings of the International Conference on Ultra Modern Telecommunications and Workshops (ICUMT '09),* pp. 1-6, 2009.

[20] B. Viswanath, M. Cha, K. P. Gummaldi, and A. Mislove, "On the Evolution of User Interaction in Facebook," *Proceedings of the 2nd ACM workshop on Online Social Networks*, pp. 37-42, 2009.

[21] E. Ferrara, R. Interdonato, and A. Tagarelli, "Online Popularity and Topical Interests Through The Lens of Instagram," *Proceeding of the 25th ACM Conference on Hypertext and Social Media, September 1-4*, 2014.

Maryam Zare received the B.Sc. degree in Computer Software Engineering from Shariaty College of Technology, Tehran, Iran. She is currently studying M.Sc. in Artificial Intelligence at K. N. Toosi University of Technology, Tehran, Iran. Her current research include social network analysis and big data analysis.
Email: zare.maryam1992@gmail.com

Seyed Hossein Khasteh received the B.Sc. degree in electrical engineering, the M.Sc. degree in Artificial Intelligence and the Ph.D. degree in Artificial Intelligence all from the Sharif University of Technology, Tehran, Iran. He is currently an assistant Professor with the Computer Engineering Department, K. N. Toosi, University of Technology, Tehran, Iran. His current research interests include social network analysis, machine learning and big data analysis.
Email: khasteh@kntu.ac.ir

Sayed Ali Khoshroo received the B.Sc. degree in Computer Hardware Engineering from AmirKabir University of Technology, Tehran, Iran. He is currently studying M.Sc. in Artificial Intelligence at K. N. Toosi University of Technology, Tehran, Iran. His current research interests are concerned primarily with Machine learning, Big data analysis and Blockchain technology.
Email: alikhoshroo@email.kntu.ac.ir

Paper Handling Data:

Submitted: 11.05.2017
Received in revised form: 19.02.2018
Accepted: 02.03.2018
Corresponding author: Dr. Seyed Hossein Khasteh, Faculty of Computer Engineering, K. N. Toosi University of Technology, Tehran, Iran.