Detecting Social Network Profile Cloning

Georgios Kontaxis, Iasonas Polakis, Sotiris Ioannidis and Evangelos P. Markatos
Institute of Computer Science
Foundation for Research and Technology Hellas
{kondax, polakis, sotiris, markatos}@ics.forth.gr

***Abstract*—Social networking is one of the most popular Internet activities, with millions of users from around the world. The time spent on sites like Facebook or LinkedIn is constantly increasing at an impressive rate. At the same time, users populate their online profile with a plethora of information that aims at providing a complete and accurate representation of themselves. Attackers may duplicate a user's online presence in the same or across different social networks and, therefore, fool other users into forming trusting social relations with the fake profile. By abusing that implicit trust transferred from the concept of relations in the physical world, they can launch phishing attacks, harvest sensitive user information, or cause unfavorable repercussions to the legitimate profile's owner.**

In this paper we propose a methodology for detecting social network profile cloning. We present the architectural design and implementation details of a prototype system that can be employed by users to investigate whether they have fallen victims to such an attack. Our experimental results from the use of this prototype system prove its efficiency and also demonstrate its simplicity in terms of deployment by everyday users. Finally, we present the findings from a short study in terms of profile information exposed by social network users.

I. Introduction

Social networking has become a prevalent activity in the Internet today, attracting hundreds of millions of users, spending billions of minutes on such services. Facebook has more than 500 million users [1] and recently surpassed Google in visits [2]. At the same time, LinkedIn hosts profiles for more than 70 million people and 1 million companies [3]. As the majority of users are not familiar with privacy issues, they often expose a large amount of personal information on their profiles that can be viewed by anyone in the network. In [4] the authors demonstrate an attack of profile cloning, where someone other than the legitimate owner of a profile creates a new profile in the same or different social network in which he copies the original information. By doing so, he creates a fake profile impersonating the legitimate owner using the cloned information. Since users may maintain profiles in more than one social networks, their contacts, especially the more distant ones, have no way of knowing if a profile encountered in a social networking site has been created by the same person who created the profile in the other site.

The usual assumption is that a new profile, claiming to be related to a pre-existing contact, is a valid profile; either a new or secondary one. Unsuspecting users tend to trust the new profile and actions initiated from it. This can be exploited by attackers to lure victims into clicking links contained in messages that can lead to phishing or drive-by-download sites. Furthermore, a cloned profile could be used to send falsified messages in order to harm the original user. The victimized user has no way of knowing the existence of the fake profiles (especially if across social networks). For that matter, we believe profile cloning is a silent but serious threat in today's world of social networks, where people might face consequences in the real world for actions of their (counterfeit) electronic profiles [5], [6].

In this paper, we propose a tool that automatically seeks and identifies cloned profiles in social networks. The key concept behind its logic is that it employs user-specific (or user-identifying) information, collected from the user's original social network profile to locate similar profiles across social networks. Any returned results, depending on how rare the common profile information is considered to be, are deemed suspicious and further inspection is performed. Finally, the user is presented with a list of possible profile clones and a score indicating their degree of similarity with his own profile.

The contributions of this paper are the following.

- We design and implement a tool that can detect cloned profiles in social domains, and conduct a case study in LinkedIn.
- We present a study of the information publicly available on a large number of social network profiles we collected.

II. Related work

Social networks have gained the attention of the research community that tries to understand, among other, their structure and user interconnection [7], [8] as well as interactions among users [9] and how user privacy is compromised [10]. In [4], the authors demonstrate the feasibility of automated identity theft attacks in social networks, both inside the same network and different ones. They are able to create forged user profiles and invite the victims' contacts to form social links or open direct messages. By establishing a social link with the forged profile, the attacker has full access to the other party's protected profile information. Furthermore, direct messages, originating from the stolen and implicitly

trusted identity, may contain malicious HTTP links to phishing or malware web sites. This attack is possible mostly due to users revealing a large amount of information on their profiles that can be accessed by practically anyone. A study conducted by Gross et al [11] revealed that only 0.06% of the users hide the visibility of information such as interests and relationships, while in [7] the authors report that 99% of the Twitter users that they checked retained the default privacy settings. Therefore, a first defense measure against such attacks could be employed by social networking sites if they promoted more strict default privacy policies. Baden et al [12] argue that by using exclusive shared knowledge for identification, two friends can verify the true identity of each other in social networks. This can enable the detection of impersonation attacks in such networks, as attackers that impersonate users will not be able to answer questions. Once a user's identity has been verified, public encryption keys can be exchanged. Furthermore, by using a web of trust one can discover many keys of friends-of-friends and verify the legitimacy of user profiles that they don't know in the real world and don't share any secret knowledge.

III. Design

In this section we outline the design of our approach for detecting forged profiles across the Web. Our system is comprised of three main components and we describe the functionality of each one.

1) **Information Distiller**. This component is responsible for extracting information from the legitimate social network profile. Initially, it analyzes the user's profile and identifies which pieces of information on that profile could be regarded as rare or user-specific and may therefore be labeled as user-identifying terms. The information extracted from the profile is used to construct test queries in search engines and social network search services. The number (count) of results returned for each query is used as a heuristic and those pieces of information that stand out, having yielded significantly fewer results than the rest of the information on the user's profile, are taken into account by the distiller. Such pieces of information are labeled as user-identifying terms and used to create a user-record for our system along with the user's full name (as it appears in his profile). The record is passed on to the next system component that uses the information to detect other potential social network profiles of the user.
2) **Profile Hunter**. This component processes user-records and uses the user-identifying terms to locate social network profiles that may potentially belong to the user. Profiles are harvested from social-network-specific queries using each network's search mechanism that contain these terms and the user's real name. All the returned results are combined and a profile-record is created. Profile-records contain a link to the user's legitimate profile along with links to all the profiles returned in the results.
3) **Profile Verifier**. This component processes profile-records and extracts the information available in the harvested social profiles. Each profile is then examined in regards to its similarity to the user's original profile. A similarity score is calculated based on the common values of information fields. Furthermore, profile pictures are compared, as cloned profiles will use the victim's photo to look more legitimate. After all the harvested profiles have been compared to the legitimate one, the user is presented with a list of all the profiles along with a similarity score.

We can see a diagram of our system in Figure 1. In step (1) the Information Distiller extracts the user-identifying information from the legitimate social network profile. This is used to create a user-record which is passed on to the Profile Hunter in Step (2). Next, Profile Hunter searches online social networks for profiles using the information from the user-record in step (3). All returned profiles are inserted in a profile-record and passed on to the Profile Verifier in step (4). The Profile Verifier compares all the profiles from the profile-record to the original legitimate profile and calculates a similarity score based on the common values of certain fields. In step (5) the profiles are presented to the user, along with the similarity scores, and an indication of which profiles are most likely to be cloned.

IV. Implementation

In this section we provide details of the proof-of-concept implementation of our approach. We use the social network LinkedIn [13] as the basis for developing our proposed design. LinkedIn is a business-oriented social networking site, hosting profiles for more than 70 million registered users and 1 million companies. As profiles are created mostly for professional reasons, users tend to make their profiles viewable by almost all other LinkedIn users, or at least all other users in the same network. Thus, an adversary can easily find a large amount of information for a specific user. For that matter, we consider it a good candidate for investigating the feasibility of an attack and developing our proposed detection tool.

A. Automated Profile Cloning Attacks

We investigate the feasibility of an automated profile cloning attack in LinkedIn. Bilge et al. [4] have demonstrated that scripted profile cloning is possible in Facebook, XING and the German sites StudiVZ and MeinVZ. In all these services but XING, CAPTCHAs were employed and CAPTCHA-breaking techniques were required. In the case of LinkedIn CAPTCHA mechanisms are not in place.

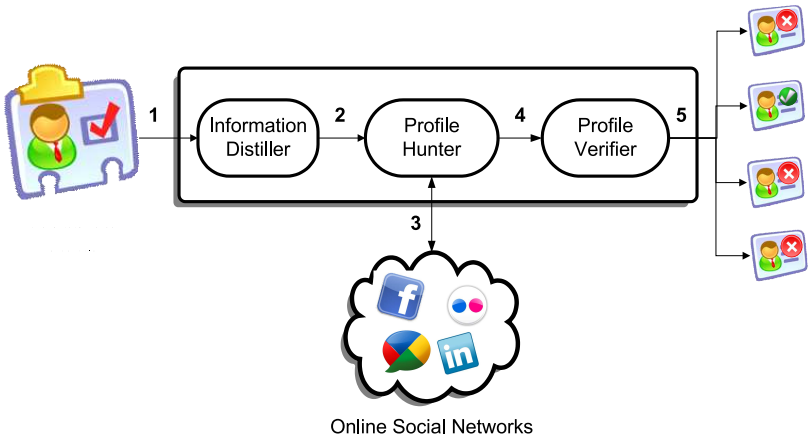

Figure 1: Diagram of our system architecture.

The user is initially prompted for his real name, valid e-mail address and a password. This suffices for creating a provisionary account in the service, which needs to be verified by accessing a private URL, sent to the user via e-mail, and entering the account's password. Receiving such messages and completing the verification process is trivial to be scripted and therefore can be carried out without human intervention. To address the need for different valid e-mail addresses, we have employed services such as 10MinuteMail [14] that provide disposable e-mail inbox accounts for a short period of time. Once the account has been verified, the user is asked to provide optional information that will populate his profile.

We have implemented the automated profile creation tool and all subsequent experiments detailed in this paper rely on this tool and not manual input from a human. This was done to test its operation under real-world conditions. Let it be noted that all accounts created for the purposes of testing automated profile creation and carrying out subsequent experiments have been now removed from LinkedIn, and during their time of activity we did not interact with any other users of the service. Furthermore, due to ethical reasons, in the case where existing profiles were duplicated, they belonged to members of our lab, whose consent we had first acquired.

B. Detecting Forged Profiles

In this section we present the details of implementing our proposed detection design in Linkedin. We employ the cURL [15] command-line tool to handle HTTP communication with the service and implement the logic of the various components of our tool using Unix bash shell scripts.

1) **Information Distiller**. This component requires the credentials of the LinkedIn user, who wishes to check for clones of his profile information, as input. The component's output is a *user-record* which contains a group of keywords, corresponding to pieces of information from the user's profile, that individually or as a combination identify that profile. After logging in with the service, this component parses the HTML tags present in the user's profile to identify the different types of information present. Consequently, it employs the Advanced Search feature of LinkedIn to perform queries that aim to identify those keywords that yield fewer results that the rest [1]. Our goal is to use the minimum number of fields. If no results are returned, we include more fields in an incremental basis, according to the number of results they yield. In our prototype implementation, we identify the number of results returned for information corresponding to a person's present title, current and past company and education. We insert the person's name along with the other information in a record and provide that data to the next component.
2) **Profile Hunter**. This component employs the user-record, which contains a person's name and information identified as rare, to search LinkedIn for similar user profiles. We employ the service's Advanced Search feature to initially find out the number of returned matches and subsequently use the protected and, if available, public links to those profiles to create a *profile-record* which is passed on to the next component. The upper limit of 100 results per query is not a problem since at this point queries are designed to be quite specific and yield at least an order of magnitude less results, an assumption which has been validated during our tests.
3) **Profile Verifier**. This component receives a *profile-record* which is a list of HTTP links pointing to protected or public profiles that are returned when we search for user information similar to the original

[1]Those that yield a number of results in the lowest order of magnitude or, in the worst case, the one with the least results.

user. Subsequently, it accesses those profiles, uses the HTML tags of those pages to identify the different types of information and performs one to one string matching with the profile of the original user. This approach is generic and not limited to a specific social network, as the verifier can look for specific fields according to each network. In our prototype implementation, we also employ naive image comparison. We assume that the attacker will have copied the image from the original profile. We use the convert tool, part of the ImageMagick suite, to perform our comparisons. In detail, to discover how much image 'A' looks like image 'B', we calculate the absolute error count (i.e. number of different pixels) between them and then compare image 'A' with an image of random noise (i.e. random RGB pixels). The two error counts give the distance between 'A' and something completely random and the distance between 'A' and 'B'. This way we can infer how much 'A' and 'B' look alike. To correctly estimate the threshold of error that can be tolerated, we plan on conducting a study where images will be manipulated so as to differ from the original photo but remain recognizable. The component outputs a similarity score between the original profile and each of the other profiles.

V. Evaluation

In this section we evaluate the efficiency of our proposed approach for detecting forged social network profiles. First, we provide data from a study on LinkedIn regarding the amount of information exposed in public or protected [2] user profiles.

A. LinkedIn Study

In order to understand how much information is exposed in public profiles of LinkedIn users, we compiled three distinct datasets of profiles and studied their nature. The idea is that an adversary seeking to perform automated profile cloning, can create such datasets and copy their information. Here we study the type and amount of information available for such an attack.

Table I presents those three distinct datasets. To do so, we created a fake LinkedIn account, that contains no information at all, and used the service's search feature to locate profiles that matched our search terms. In the free version of the service, the number of search results is bound to 100 but one can slightly modify his queries to count as different searches and at the same time return complementary sets of results. In our case, we used three lists as search terms to retrieve user profiles; one with the most common English surnames, one with the top companies according to Fortune Magazine [16] and one with the top U.S. universities.

[2]To view the profile information, a service account is required.

Trace Name	Description	Profiles
surnames	Popular 100 English names	11281
companies	Fortune 100 companies	9527
universities	Top 100 U.S. universities	8811

Table I: Summary of data collected.

Each of the ∼30K search results returned a summary of the user's profile, which we consider adequate information to convincingly clone a profile. As we can see in table II, almost one out of every three returned search results is public and contains the user's name, along with current location and current title or affiliation. These profiles are accessible by anyone on the web, without the need for a LinkedIn account. In detail, in the *surnames* dataset 89% of the profiles has a public presence on the web. On the other hand, for profiles collected from the *companies* and *universities* datasets, public presence is merely 2.3% and 1.6% respectively. The big discrepancy is probably due to the fact that users from the industry and academia use LinkedIn for professional purposes and therefore set their profiles as viewable by other LinkedIn users only.

Table III presents the core profile information in all the profiles that are publicly available. Interestingly, besides the person's name, almost all public profiles carry information about the present location and relative industry. Additionally, about half of the profiles include a person's photo, current title or affiliation and education information.

In Table IV we can see the information available in all the profiles that require a LinkedIn account for viewing. While the percentage of profiles from which we can access the user's photo is smaller compared to the public profiles, all the important information fields present a much higher availability. The fact that we cannot access the photos in many profiles is due to default privacy setting of LinkedIn where a user's photo is viewable only to other users from the same network. Nonetheless, an adversary could set his account to the specific network of the targeted victims in order to harvest the photo. Furthermore, all users reveal their location, and connections, and almost all their industry field. Most profiles from the surname dataset contain information regarding the user's current work status and education (86% and 70% respectively). The other datasets have an even larger percentage verifying the professional usage orientation of the users. Specifically, 99% of the profiles from the companies dataset contained information on current status and 92% revealed the user's education, and profiles from the universities dataset stated that information in 94% and 99% of the cases. Therefore, any user with a LinkedIn account can gain access to user-identifying information from profiles in the vast majority of cases.

A short study by Polakis et. al [17] concerning the type and amount of information publicly available in Facebook profiles, demonstrated a similar availability of personal in-

	surnames	*companies*	*universities*
Public Name	90.5%	2.5%	2.0%
Public Profile	89%	2.3%	1.6%

Table II: Exposure of user names and profile information.

	surnames	*companies*	*universities*
Photos	47%	59%	44%
Location	98%	99%	99%
Industry	85%	97%	98%
Current Status	70%	86%	72%
Education	53%	66%	82%
Past Status	42%	54%	63%
Website	36%	50%	39%
Activities / Societies	21%	22%	55%

Table III: Information available in public LinkedIn profiles for each dataset.

formation. While their results show a lower percentage of Facebook users sharing their information publicly, close to 25% of the users revealed their high school, college and employment affiliation, and over 40% revealed their current location.

As demonstrated from both of these studies, it is trivial for an adversary to gather information from social network accounts that will allow him to successfully clone user profiles. With the creation of a single fake account, an adversary can gain access to a plethora of details that we consider sufficient for deploying a very convincing impersonation attack. Even so, this information is also sufficient for the detection and matching of a duplicate profile from our tool.

B. Detection Efficiency

Initially, we evaluated our hypothesis that different pieces of information from a user profile yield a variable number of results when used as search terms, for instance in a social network's search engine. To do so, for each profile in our datasets, we extracted the values from different types of information and used them as search terms in the Advanced Search feature of the service. Next, we recorded the minimum and maximum number of results returned by any given term. Finally we calculated the range (maximum - minimum) of search results for information on that profile. Figure 2 presents the CDF of the range of search results returned for each profile in our dataset. One may observe a median range value of $\sim$1000 and also that only 10% of profiles had a range of search results lower that 20. Overall, we can see that the majority of profiles exhibited diversity in the number of search results returned by different pieces of information, and by leveraging this can be uniquely identified by the carefully crafted queries of our system.

Next, we conducted a controlled experiment to test the efficiency of our tool. Due to obvious ethical reasons we were not able to deploy a massive profile cloning attack in the wild. Thus, we selected a set of 10 existing LinkedIn profiles, that belong to members of our lab, and cloned them inside the same social network using the automated method described in IV. We then employed our tool to try and find the duplicates. Overall, we were able to detect all the profile clones without any false positives or negatives.

	surnames	*companies*	*universities*
Photos	22%	52%	26%
Location	100%	100%	100%
Industry	94%	100%	100%
Connections	100%	100%	100%
Current Status	86%	99%	94%
Education	70%	92%	99%
Past Status	58%	96%	95%
Twitter Username	13%	0%	1%
Websites	41%	2%	1%

Table IV: Information available in protected LinkedIn profiles.

Finally, we used public user profiles as seeds into our system to try and detect existing duplicates inside LinkedIn. The Information Distiller produced user-records using information from current or past employment and education fields. Overall, we used 1,120 public profiles with 756 being from the surnames dataset, the 224 public profiles from the companies dataset and the 140 public profiles from the universities dataset. The Profile Hunter component returned at least one clone for 7.5% of the user profiles (in 3 cases our tool discovered 2 cloned instances of the profile). Our prototype system relied on the exact matching of fields and did not employ our image comparison technique to detect cloned profiles. Furthermore, similarity scores were based on the number of fields that contained information on both profiles (in several cases, one profile had less fields that contained information). After manual inspection, we verified that all detected profiles pointed to the actual person and that the score produced by the Profile Verifier was accurate. We cannot be certain if those clones are the result of a malicious act or can be attributed to misconfiguration. Furthermore, our prototype may have missed cloned profiles where the attacker deliberately injected mistakes so as to avoid detection. We discuss how our system can be improved in Section VI.

VI. Future Work

In this section we discuss the future approaches we will take to improve our approach. An important drawback of our system is that it currently uses only the LinkedIn social network. Our next step is to extend its functionality to utilize other popular social networks and create a profile parser for each network.

The next axis upon which our tool can be improved lies in the accuracy of comparing two profiles and assigning a similarity score. Our current implementation of the Profile Verifier looks for exact string matches in information fields when comparing two profiles. Instead of looking for exact

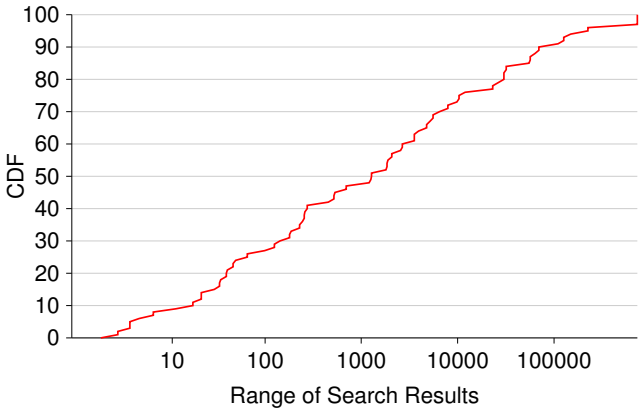

Figure 2: CDF of the range of search results returned for different pieces of information on a user profile.

matches we can use fuzzy string matching to overcome wrongly typed information, or deliberately injected mistakes. An important aspect of this is to correctly tune the fuzzy matching algorithm to match our needs. Since the presentation of the same information across OSNs may vary, we must implement information extracting functions specific for each social network, that extract the information and convert it to a custom representation format. Finally, we want to conduct a study to calculate the error threshold for the image comparison.

VII. Conclusions

In this paper we propose a methodology for detecting social network profile cloning. We first present the design and prototype implementation of a tool that can be employed by users to investigate whether they have fallen victims to such an attack. The core idea behind our tool is to identify any information contained in a user's profile that can uniquely identify him. We evaluate our assumption regarding the effectiveness of such a tool and find that user profiles usually reveal information that is rare and, when combined with a name, can uniquely identify a profile and thereby any existing clones. In that light, we present the findings from a study regarding the type and amount of information exposed by social network users and conclude that the same user-identifying information which allows an attacker to clone a profile also assists us in identifying the clone. This is demonstrated by a test deployment of our tool, in which we search LinkedIn for duplicate profiles, and find that for 7% of the user profiles checked, we discover a duplicate profile in the same social network.

Acknowledgments

The research leading to these results has received funding from the European Union Seventh Framework Programme (FP7/2007-2013) under grant agreement 257007. This work was supported in part by the Marie Curie Actions – Reintegration Grants project PASS. We thank the anonymous reviewers for their valuable comments. Georgios Kontaxis, Iasonas Polakis, Sotiris Ioannidis and Evangelos P. Markatos are also with the University of Crete.

References

[1] "Facebook statistics," http://www.facebook.com/press/info.php?statistics.

[2] "Facebook traffic tops google for the week," http://money.cnn.com/2010/03/16/technology/facebook_most_visited/.

[3] http://techcrunch.com/2010/06/20/linkedin-tops-70-million-users-includes-over-one-million-company-profiles/.

[4] L. Bilge, T. Strufe, D. Balzarotti, and E. Kirda, "All your contacts are belong to us: automated identity theft attacks on social networks," in *WWW '09: Proceedings of the 18th international conference on World wide web*.

[5] "BBC News - Facebook prison officer is sacked," http://news.bbc.co.uk/2/hi/uk_news/england/leicestershire/7959063.stm.

[6] "BBC News - "Ill" worker fired over Facebook," http://news.bbc.co.uk/2/hi/8018329.stm.

[7] B. Krishnamurthy, P. Gill, and M. Arlitt, "A few chirps about twitter," in *WOSN '08: Proceedings of the first ACM workshop on Online social networks*.

[8] J. Tang, M. Musolesi, C. Mascolo, and V. Latora, "Temporal distance metrics for social network analysis," in *WOSN '09: Proceedings of the 2nd ACM workshop on Online social networks*.

[9] B. Viswanath, A. Mislove, M. Cha, and K. P. Gummadi, "On the evolution of user interaction in facebook," in *WOSN '09: Proceedings of the 2nd ACM workshop on Online social networks*.

[10] B. Krishnamurthy and C. E. Wills, "On the leakage of personally identifiable information via online social networks," in *WOSN '09: Proceedings of the 2nd ACM workshop on Online social networks*.

[11] R. Gross and A. Acquisti, "Information revelation and privacy in online social networks," in *WPES '05: Proceedings of the ACM workshop on Privacy in the electronic society*.

[12] R. Baden, N. Spring, and B. Bhattacharjee, "Identifying close friends on the internet," in *Proc. of workshop on Hot Topics in Networks (HotNets-VIII)*, 2009.

[13] "LinkedIn," http://www.linkedin.com/.

[14] "10 Minute Mail," http://10minutemail.com/.

[15] "cURL," http://curl.haxx.se/.

[16] "Fortune magazine," http://money.cnn.com/magazines/fortune/.

[17] I. Polakis, G. Kontaxis, S. Antonatos, E. Gessiou, T. Petsas, and E. P. Markatos, "Using social networks to harvest email addresses," in *WPES '10: Proceedings of the 9th annual ACM workshop on Privacy in the electronic society*.