Detection Techniques of Clone Attack on Online Social Networks: Survey and Analysis

Deepti Dave, Nishchol Mishra and Sanjeev Sharma*

School of Information Technology, Rajiv Gandhi Proudyogi Vishwavidyalaya, Bhopal, M.P., India

Abstract

Online social network presents a wide variety of online uses that exhibit a series of challenges to the security. The security and privacy threats increase, as the amount of personal information posted by users in their profile is made public. Social networking provides interaction and socialization among users which generates large amount of data. This necessitates the adoption of new tools and techniques to provide security of online data. Online social network users are unaware of the numerous security risks that exist in these networks. This paper presents a meticulous survey on the classification of different attacks and defence mechanisms which threaten the well-being of online social network. In addition, the paper presents a classification and analysis of detection mechanisms of clone attacks on online social network, based on attribute similarity, friend network similarity, profile analysis for a time interval and record of Internet Protocol sequences.

Keywords: Online Social networks; attacks; defence mechanisms; profile clone attack; cross-site profile clone attack.

1. Introduction

With the evolution of World Wide Web in 1990 and the development of information exchange, social networking has grown rapidly. Millions of people around the world are connected to each other by Online Social Networks (OSN) which uses web 2.0 technologies. A social networking site is a website where every user has its own profile and has contacts with its friends, family members, employees, share their updates and join new communities and groups which belong to user’s interests. Most of the people interact with those persons on OSN, whom they have contact in offline world. Hence, their profile indicates actual representation of user’s individuality of real life [1].

OSN’s stores massive amount of sensitive as well as private information of users and their conversations. As the content of information is increasing, the social network providers and security companies are bound to provide advanced security features in their social network. Some of them provide protection against hackers, spammers, identity cloning, social bots, phishing, and many more threats. But a larger part of online users are not aware with privacy schemes and they often reveal a huge amount of personal data on their profiles that can be seen by anyone in the network. Many third party applications of various OSNs sites have proposed and employed defence mechanisms against adversaries. Products like Websense’s Defensio, Checkpoint’s SocialGuard, Reclaim Privacy, UnitedParents and PrivAware application which are open source and commercial, provide online social network users tools for safeguarding themselves. Till 20th century, very few cases of spam were reported, but now spam is becoming an increasing problem.

Moreover, administrators often face problems in differentiating between Sybil attacks and Identity Clone attacks and they fail to deploy suitable and effective defence approaches against them. Lei Jin, Xuelian Long, Hassan Takabi and James B.D. Joshi [2] presented a paper and distinguished these two attacks on the basis of pre-requirements, network topology and attack impacts. In Sybil attack, adversary’s main aim is to create multiple online user identities (Sybil identities) and tries to achieve
vital information with the help of these identities. Where as in an identity Clone attack (Profile Cloning attack), an attacker first creates similar or even identical profiles to imitate victims on OSN system where genuine user is already registered.

Social network managers have added various authentication procedures to protect their genuine users from fake ones. SybilGuard, SybilLimit, SybilInfer, SybilDefender and SumUp are the major approaches to defend against Sybil attack. SybilGuard presents a decentralised protocol by confining both the number and size of Sybil group. In comparison to SybilGuard, SybilLimit ensures more optimal and acceptable limits for the number of Sybil nodes in the network. Both SybilGuard and SybilLimit are good solutions for detecting Sybil nodes. However, these techniques are not sufficient to detect clone accounts.

Cloning attack is one of the major attack patterns towards online social networks, where the adversary conceals fake accounts details by thieving and copying genuine user profiles and sends friend requests to the friends of the cloned victim. It is difficult for ordinary users to detect these fake identities because of the identical names and similar profile information. This paper discusses several approaches of detecting clone profiles on OSN.

The rest of the paper is organized as follows. Section 2 defines social network and evolution of social network. It also examines the attacks and defence mechanisms applied on OSN. Section 3 describes various techniques for detecting clone attack and includes analysis table which illustrates the summarized form of all the techniques. Section 4 concludes the paper and discusses future work.

2. Online Social Network

In this segment, we mainly focus on definition and evolution of OSN. It also describes different types of attacks and defence mechanism applied on OSN.

2.1. Definition of Online Social Network

The offline social network depicts a relationships and communication between folks, groups or other knowledge processing entities. Boyd and Ellison [1] [3] defined Online Social network site as “web based services that allow individuals to (1) construct a public or semi-public profile within a bounded system, (2) articulate a list of other users with whom they share a connection and (3) view and traverse their list of connections and those made by others within the system”. A pair of users in an online social network can have more than one kind of relationship. These relationships or links can be of two types: user-user link and user-group link. An online user is connected to his/her friends by user-user link and is associated with any group or committee with user-group link. The number of links may vary from profile to profile.

2.2. Evolution of Social Network

Before the invention of Internet in sixties, S. Milgram [4] speculated the likelihood that any pair of actors on the planet are separated by at most six degrees of separation. SixDegrees, launched in 1997 was the first example of a social network. SixDegrees gave users the facility to create their own profiles and connect to the people with similar interests and hobbies. Currently, SixDegrees is open to only previously registered members. New members are only permitted if they are invited. Friendster embarked in 2002 is the next social network that is worth to be mentioned. Friendster security features were designed such that, it restricted the appearance of the profiles to other users, who were more than four degrees faraway (friends-of-friends-of-friends-of-friends).

From 2002 to 2006, a great boom of social networking sites came. MySpace launched in August 2003, grew quickly and has become a meeting point for sharing music, videos and photos. The next social network that catches people's eye is Facebook. Facebook was created by Harvard undergraduate Mark Zuckerberg in 2004. At that time, this social network was thought to be a means for only Harvard students to keep up with the dating games among friends, but nowadays it has become a worldwide social network.

There are many other social networks which came into existence and had made a mark in social media. They are LinkedIn, Hi5, Orkut, Bebo, Twitter, Pinterest and Google+. Recently Microsoft made its own public intention for a social network site called Wallop.

2.3. Privacy in Social Network

Privacy is an important issue in OSNs, since the illegal disclosure and improper use of user’s personal information can cause undesirable consequences in personal life. Social network service providers regularly update their security policies to provide safety measures to user’s profiles. This section focuses on the area of privacy in SNSs (social network systems), describing the broader overview of different attacks performed on various social networking sites and also the defence mechanisms applied by OSN providers.

2.3.1. Attacks

Even though social networking application managers supply sufficient amount of security, but the interested hackers still try to invade unauthorized information. Attacks can be grouped in two forms: anonymized data attack and attacks on access limited information.

In anonymized data attack, adversary gathers some identifiable information about the target user and is able to recognise target among the anonymized information. In [5], these types of attacks are further presented as active attacks, passive attacks and semi-passive attacks. In active attack, the adversary attempts to make new profiles on sites and creates new links in the network so that these links will be present in the anonymized form of information. While in
passive attack, no new creation or modifications are done, the adversary only tries to detect the peculiar properties of users after anonymized information is published. On the other hand in semi passive attack some links are made with the target user before the information is anonymized.

Different social networking sites provide the security feature of adjusting profile details of user. For example, a user can set his personal details like picture, date of birth, phone number, and address to limited access only which is then seen by his/her friends only. Likewise different accessibility options can be applied on profiles. But the intension of attackers is always to break these security features. Attacks such as exploiting privacy breaches, automated identity theft, node identity forging, and social network link forging comes under attacks on access limited information. Automated identity theft has its subcategory as profile clone attack and cross site profile clone attack. The main aim of an adversary in profile clone attack is to obtain personal information about the targets friends after successfully forging the target and to make a trustworthy network within the targets social network for future deception. Fig 1: Depicts clone attack in a social network, where attacker creates clone profile of victim after sniffing its profile and then sends friend request to victim’s friends. Some friends (green ones) accept clones friend request. As a result certain confidential information is leaked to the attacker.

![Fig.1. Profile clone attack.](image)

In the cross site profile clone attack, adversary gathers information about his targets profile and makes new profile on some other OSN where the target is not registered. Now the motive of attacker is to establish a full reliable network by adding friends of the target on a new OSN. Fig 2: shows cross site profile clone attack where clone identity is built by an attacker on another social network. The attacker copies victim’s profile details which are public and makes a new profile. The private attributes of victim’s profile (shown by *) are not displayed in faked identity (victim’s clone profile). Friends of victim can only see public attributes of fake identity that are not connected to it. So they might face difficulty in differentiating between genuine and fake identity. Thus, there is a high possibility of accepting clone’s friend request.

![Fig.2. Cross site profile clone attack.](image)

2.3.2. Defence Mechanism

A number of defence methods have been proposed. Privacy related texts like privacy policies should be unambiguous, as short as possible, and simply understandable to users. Help desk for users is one of the main methods. Users are provided with friendly privacy related text. Mohammad Soryani and Behrooz Minaei [6] examined that, Online Social managers should help users towards regulating their privacy settings and presenting informative matrices; informing users that which part of their information is accessible and to whom; providing some facilities to control the way that people can access their information.

In [7], a few risky behaviours are mentioned: incautious in accepting friendship requests, clicking on links received from others without enough vigilance, reacting to suspicious friendship requests after accepting them (not before) and therefore letting suspicious users to access information, interaction with fake profiles and overall, high implicit trust that exists in SNSs. A study was cited there in which, 41% of 200 users whom a friendship request were randomly sent to, accepted it, and most of the users had not put limited access to their personal profile information. Some of the attackers try to take advantage of publically available data, either in a legal or illegal way. An important tool that is used to acquire the online information of SNSs is a bot, with which intruders are able to accomplish their actions automatically and with a large scale. Among the
actions that could be done with these automatic systems are: crawling, creating fake profiles and sending forged friendship requests. To avoid this, many of the social networking provider’s presents safeguard mechanism of online information of SNSs. They identify breaches, based on anomaly detection techniques. Automatic blocking systems are based on IP monitoring and Completely Automated Public Turing test to tell Computers and Humans Apart (CAPTCHA). A CAPTCHA is a program that protects websites against bots by creating tests that human can pass but programs cannot.

Many social networking suppliers offer protection while advertisement, through smart publishing and anonymization. An interesting approach is presented [8] in which, only some results of Social Network Analysis (SNA) is presented in the form of two centrality metrics to investigators and gives them the opportunity to send queries without violating privacy measures. Anonymization is a technique that aims to protect the user’s privacy and publish information at the same time. The study [9] of anonymization methods have been categorized according to clustering based approaches and graph modification approaches. Clustering based approaches can be further subdivided into four subcategories of vertex clustering, edge clustering, vertex-edge clustering and vertex-attribute mapping clustering. Graph modification approaches comprises of three subcategories: optimization, randomized graph modification and greedy graph modification. In order to grant protection against social networking providers, encryption and P2P architecture were also proposed.

3. Related Work

Most of the business networking or friendship social networking sites have the feature of adding friends in their network. When user X receives friend request from another user Y, or when user X obtains user Y’s link during friend search or if it is recommended by some mutual friend, there is no instant approach for user X to verify the authenticity of user Y. If user Y has applied limited accessibility feature on his profile, then user X can see only publically available information of user Y. At this time adversary takes the advantage of this situation and creates fake profile of the target by including the public attributes and adds the individuals of the targets friend circle into its own circle. In some cases, adversary first tries to become friend with the target so as to gain more information that are visible to friends, which can build a more trustful profile. As most of the features between fake profile and real profile are same, it becomes difficult to make out which profile is genuine or which one is faked for the users.

The detection mechanisms of identifying clones on online social network can be divided into two categories: content related based and content free. The following three subsections (3.1, 3.2 and 3.3) describes detection of fake profiles based on attribute similarity, friend network similarity and profile analysis for a time interval respectively, which comes under content related information and section 3.4 describes detection of clone attack based on content free information which uses recent record of IP (Internet Protocol) addresses.

3.1. Attribute Similarity

Attackers often target social networking sites, as they contain sensitive and personal information of registered users. For registering any of the social networking sites, the registration form includes: username, password, gender, date of birth, email address, professional background and some- times even their past relationship. Hence, it attracts attackers as they can get access to these types of details. Personal information could also be used for social engineering attacks. Some sites, like Facebook has the default setting that only friend can view personal details.

Authors [10], at the beginning described about profile cloning attack and cross site profile cloning attack and further on, illustrated their detection technique. A prototype system called iCloner (identity clone) was designed which consisted of four main parts: crawler, identity matcher, profile creator and message sender.

The crawler component crawls five social networking sites (XING, LinkedIn, StudiVZ, Facebook and MeinVZ) and stores all the publicly available information about the users in the database. The information stored in database is then analysed by identity matcher and tries to identify profiles, in different social network that relates to same person. Then profile creator uses this information to create existing profile within the same social network (profile clone attack) or create profiles in another social network where the target has not registered yet (cross site profile clone attack). At last, the message sender sign in into fake accounts and send friend requests to all known contacts of the target along with a message. Due to this personal message, authenticity of fake profile also increases. To make these attacks feasible, CAPTCHA analysis tool was also added to iCloner with its breaking capability. CAPTCHA algorithms are mainly used to distinguish between human and computer application. During creation of profile on MeinVZ, StudiVZ and Facebook require users to solve CAPTCHA.

In MeinVZ and StudiVZ, iCloner was able to solve CAPTCHA with a probability of 99.8% in one of the three consecutive attempts. Facebook uses a more advance version of CAPTCHA called reCAPTCHA. To make CAPTCHA more difficult to break, images are displayed in distorted form, as some noise is added to it and is covered by a curve line. In first 100 attempts, iCloner was able to solve only 4 to 7 CAPTCHAs. In addition, iCloner was able to crawl more than 5 million user profiles along with the contact information and more than 1.2 million profiles with full user information on MeinVZ and StudiVZ. On XING, iCloner was able to crawl 2000 profiles easily as it didn’t contained CAPTCHA algorithm to prevent automated crawling. Also, iCloner supports profile cloning attack on Facebook and automatically compares and falsify accounts from XING to LinkedIn based on attribute similarity.
A scoring system and threshold was decided to verify that two accounts with the same name on different social networking sites relate to the same user or not. Two points were allotted if education field match, two more points were assigned if company for which user is working is same and one point was given if the country and the city where users are living are also same. All the assigned points are added up and if the total score comes to at least three, then it’s concluded that two profiles belongs to same user. Sometimes the details to users don’t match as way of writing them varies. For example, Tata Consultancy Services can also be written as TCS. Here the entries of work place are same, but string comparison would fail in this case. The solution to this problem was to search for both these entries on Google. And if both these entries appear in top 3 hits, it is concluded as equivalent.

The shortcoming of this paper was that it concentrates only on the cloned targets friend list. Also it does not provide a method to integrate multiple Sybil users into a specific online social network.

In 2011, Georgios Kontaxis, Iasonas Polakis, Sotiris Ioannidis and Evangelos P. Markatos [11] offered a methodology of detecting clones which was also based on attribute value similarity in LinkedIn social network.

Attribute similarity of two profiles $S_{aa}$ is calculated as [12]:

$$S_{aa}(P_c, P_v) = \frac{SA_{cv}}{\sqrt{|A_c| \times |A_v|}}$$

Where

- $P_c =$ public profile of a candidate identity $c$,
- $P_v =$ public profile of victim $v$,
- $SA_{cv} =$ the number of the attribute for which $P_c$ and $P_v$ have similar values,
- $|A_c|$ and $|A_v|$ represents the number of attributes in $P_c$ and $P_v$ respectively.

The architecture was built for identifying forged profiles which can be used by LinkedIn users who wish to check for clones of their profile. This system constituted of Information Distiller, Profile Hunter and Profile Verifier. The Information Distiller extracts the user identifying information from genuine social network profile. Then profile hunter uses this information to search for similar profiles on social network. All the selected profiles are forwarded to profile verifier. Profile verifier computes a similarity score by comparing the genuine profile and the selected profile record, based on the common values of certain fields. Image comparison feature was also setup in their prototype implementation. At last, an indication is given to user of which profiles are most likely to be cloned along with the similarity score. 10 existing LinkedIn profiles of members of their lab were cloned in the same network and examined by the architecture defined above.

And detection result was 100% that is; all cloned profiles were detected without any false positive and negative.

The drawback of this architecture was that, it was compatible with only LinkedIn network and it missed cloned profiles where the adversary purposefully inserted errors so as to avoid detection. This tool can further be improved by implementing fuzzy string matching for comparing user details.

### 3.2. Friend Network Similarity

Further on, Lei Jin, Hassan Takabi, James B.D. Joshi in 2011 [12] put forward two approaches for calculating profile similarity between two identities. Along with attribute similarity and another parameter friend network similarity was added for detecting fake profiles on OSN. They assumed that each user has three different kinds of lists; a friend list (FL) that includes users friends in his circle, a recommended friend list (RFL) that OSN manager creates to recommend possible new friends based on common interests or likes and an excluded friend list (EFL) that includes those users which user wants to avoid adding them into his circle. Sometimes, attacker creates victims fake profile and send friend request to those users whom victim has included in its EFL. If these friend requests are acknowledged, then clone profile becomes more real and turns out to be simpler for attacker to include target’s friends whom EFL is not known.

The detection process consisted of an input identity (IID) and a profile set. Those profiles which have same name to that of IID were stored as Candidate list (CL). When the resemblance of a candidate identity exceeds a predefined threshold, it was added into Suspicous Identity List (SIL). All IIDs were also included into SIL as it is not known whether IID fake or not. That identity which comes out to be faked was putted into Faked Identity List (FIL).

Similarity between FLs in two identities is calculated as $S_{ff}$ [12].

$$S_{ff}(P_c, P_v) = \frac{|MFF_{cv}|}{\sqrt{|F_c| \times |F_v|}}$$

Similarity between FL of $P_c$ and RFL of $P_v$ is defined as $S_{frf}$ [12].

$$S_{frf}(P_c, P_v) = \frac{|MFRRF_{cv}|}{\sqrt{|F_c| \times |RF_v|}}$$

Similarity between FL of $P_c$ and EFL of $P_v$ is defined as $S_{gef}$ [12].

$$S_{gef}(P_c, P_v) = \frac{|MFEEF_{cv}|}{\sqrt{|F_c| \times |EF_v|}}$$

Where
\[ |F_c| \text{ and } |F_v| \text{ represents the number of friends in } P_c \text{ and } P_v \text{ respectively.} \]
\[ |RF_v| = \text{ the number of recommended friends in } P_v. \]
\[ |EF_v| = \text{ the number of excluded friends in } P_v. \]
\[ MFF_{cv} = \text{ set of mutual friends common in the FLs of } P_c \text{ and } P_v. \]
\[ MFRF_{cv} = \text{ set of mutual friends common in FL of } P_c \text{ and RFL of } P_v. \]
\[ MFEF_{cv} = \text{ set of mutual friends common between FL of } P_c \text{ and EFL of } P_v. \]
\[ |X| = \text{ the number of elements in set } X. \]

If the identity verified to be authentic then its trust value was increased to avoid future multiple validations. In the end, faked identities were temporarily closed or blocked and friends of faked identities received notification that their friends are identified as fake. Authors [12] had used offline dataset of Facebook such that every user has at least 2 to 10 non hidden attributes and added 10 to 42 recommended friends and 5 to 40 excluded friends of users in them, for the detection process.

The above mentioned approach had some limitations: like it did not validate the detection architecture in real system as it used only offline dataset of Facebook. Also this mechanism was not able to detect fake profiles of those users which never used social networking sites. In this case, adversary gathers enough personal information and creates fake profiles. Sometimes, attackers creates profiles of those person also who do not exists in this world. This detection approach can further be expanded to real base OSN by developing third party fake identity detection application.

3.3. Profile Analysis for a time interval

Authors [13] discovered a novel approach based on growth rate of social network graph and on the interactions among social networking users for detecting fake profiles. They examined online social network graphs from dynamic perspective, keeping privacy issues in mind. Mauro Conti, Radha Poovendran and Marco Secchiero [13] had applied time evolution method on OSN by formulating Facebook application for collecting information from genuine users. A normal user on any online social network has initial fast growth region trailed by slow growth region and which is expected to some level. The attackers who create fake profile try to avoid sending friend request to close friends of the victim as they have the highest probability of alerting the target in minimum amount of time. This approach uses three features- evolution over time of the number of OSN friends, real social interaction and evolution over time of the structure of OSN graph. The methodology they adopted for detecting fake profiles was applied where victims account does not existed on the social network.

The drawback with this approach was that, it was not able to completely recreate day-by-day graph of the user’s network starting from the profile creation. This work can further be extended by including features of online interactions among users based on tags, friendship requests, joining new communities and posting data of some related content on profiles.

Mainly, profile analysis is performed to detect compromised accounts. Chao Michael Zhang and Vern Paxson [14] did a study on this and proposed an approach for finding out compromised profiles on twitter based on tweet timestamps. Twitter provides a facility of verified account to users, with a verified badge that appears in the top-right portion of a users profile just above the name, location and bio information. This verified badge signifies that the profile belongs to legitimate user. Till 2011, 1738 verified profiles existed on twitter. Twitter limits the user with 1000 tweets per day. So, if there is automation of tweets at times which are either too uniform or not uniform within a threshold limit then compromised profile can be detected.

They [14] used Pearson’s \( \chi^2 \) test to measure the uniformity of second-of-the-minute and minute-of-the-hour distributions by analysing consistency of update time pattern of tweets. By applying \( \chi^2 \) test, p-value is returned, which signifies the probability of the observed distribution of time arising if the profile is really publishing updates constantly across seconds-of-the-minute or minutes-of-the-hour.

3.4. IP address similarity

A recent study was conducted by Zifei Shan, Haowen Cao, Jason Lv, Cong Yan and Annie Liu [15] in 2013 on Renren and presented content free approach for detecting clones in OSN. The detector was named as CloneSpotter and was deployed into social network servers. They also discovered novel threatening attack pattern called snowball sampling technique allowing clones to add more friends and iteration attack offering an approach to insert fake profiles into the social network, thus increasing the influence of the attacker.

Snowball sampling is an iterative sampling technique where the adversary sends friend request to users, who are friends of those users who have accepted the friend request of the cloned account. As the number of common friends increases, the reliability increases and it enhances the speed of acceptance of friend request. In Iteration attack, the adversary makes new cloning profiles using the detailed profile by copying the user’s information, picture and posts to become more authentic to other users. Several Sybil profiles can also be introduced into the network by using this iteration technique. By doing so, the cloned profiles achieves more power and they can be further used for spreading spams, phishing attack, malware and other types of theft.
The experiment conducted on Renren contained clones of three levels:
- The first level of clones have common name as of the target.
- The second level of clones has common name, birthday and school.
- The third level of clones uses old profile picture of target along with the matching personal details of the user.

A comparison was made between traditional attack pattern and new attack pattern and it was observed that enhanced cloning pattern is more hazardous. On an average, out of 24 requests traditional attack pattern received only 2.7 acceptances, while cloned profiles received at least 5 positive answers out of 24 requests. By using original clone pattern attack in level 1, only 6.3 acceptance were received, on the other hand an average of 11 requests were acknowledged using Snowball sampling optimization. This survey indicates that snowball sampling technique improves the chances of accepting the cloned friend request as common friends play a major role when normal users deal with it.

Authors [15] have proposed a real time based content free detector called CloneSpotter which makes decision according to physical information and real connections, neglecting the user generated data. CloneSpotter is a lightweight and speedy detector deployed on the server side of OSNs. When user A sends friend request to user B, it checks whether user B has a friend C with same name as A or not? Similarity between attributes like gender, school, city etc. is calculated. After exceeding a threshold attribute similarity, it compares first 16 bits of each login IP address stored in recent list of each user to distinguish between them. If two users comes out to have same IP prefix then it is concluded that both accounts are owned by same user or it has the highest probability that attacker has cloned the victims account. Finally, the server sends warning to victim’s friend to be cautious and block the cloned profile if the result of CloneSpotter detects the adversary. The false positive rate is less in content free approach than content related approach.

Clone detection techniques which were content related based, made decision by analysing and comparing the profile details, pictures, messages and many more attributes of genuine user with the attackers profile. Though these approaches required less authority permission and were easily experimented on client side, but still they were highly dependent on user generated information. The content related approaches which included analysis of profile were efficient enough to detect clones, but by the time detection is complete, the attacker would had already gained vital information about the victim. The latest mechanism proposed in 2013 was content free approach that is more powerful for protection against clone attack, as CloneSpotter detector is real time based and requires less cost. The table-1 given below briefly describes the summarized form of content based and content free techniques which various researchers in recent years had proposed, including the parameters used for detection of clones, Website on which they were implemented, special features involved and their experimental outcomes.

<table>
<thead>
<tr>
<th>Techniques</th>
<th>Approaches used</th>
<th>Parameters for detection</th>
<th>Implemented on</th>
<th>Advanced Features</th>
<th>Outcomes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Content Based.</td>
<td>Attribute Similarity.</td>
<td>First Name, Surname, Education, Company, City, Country.</td>
<td>Facebook, LinkedIn, MeinVZ, StudiVZ, XING.</td>
<td>Offer Automated crawling and CAPTCHA.</td>
<td>Friend request acceptance of profile cloning: over 60% and cross site profile cloning: 56.4%.</td>
</tr>
<tr>
<td>Content Based.</td>
<td>Attribute similarity and Friend Network Similarity.</td>
<td>Surname, current and past employment, education, Image comparison.</td>
<td>LinkedIn.</td>
<td>Image comparison.</td>
<td>100% detection of clone profiles.</td>
</tr>
<tr>
<td>Profile Analysis for a Time Interval.</td>
<td>Number of friends, real social interaction, structure of online social network graph.</td>
<td>Facebook.</td>
<td>Provide detection on Basic Profile Similarity (BPS) and Multiple-Faked Identities Profile Similarity (MFIPS).</td>
<td>114 more faked identities were detected in MFIPS approach than BPS approach.</td>
<td></td>
</tr>
<tr>
<td>Content Free.</td>
<td>Attribute similarity and Record of IP address.</td>
<td>User icon, background image, screen name, number of followers and friends, status updates (number of updates, sources, frequency and contents).</td>
<td>Twitter.</td>
<td>Pearson’s $\chi^2$ test on number of tweets: minutes-in-the-hour and seconds-in-the-minute.</td>
<td>Detected 16% accounts which publish tweets that exhibit discernible automation.</td>
</tr>
</tbody>
</table>

| Content Free. | Attribute similarity and Record of IP address. | Name, gender, school, living city, first four recent 16bits of IP address. | Renren. | Snowball Sampling and Iteration attack. | Suffers from IP Spoofing. |
4. Conclusion and Future Work

This paper presents a brief knowledge about the attacks and defence mechanisms which are prominent on Online Social networks. It also explains the work which had been performed in the field of detecting clone profiles and cross site clones on OSN. Content related approaches exploits user generated information and is quite simple for detection. Whereas content free method, makes use of information which is stored on OSN server. Hence, it is seen that content free detection approach is more efficient than content based, resulting in less false positive profiles.

In future, a new hybrid dynamic approach can be developed by integrating the content free approach with content based. Additional features like evaluation of login action time pattern and click pattern of users or observing users activities (content free) can be included, along with the threshold put on similarity of users value attributes (content based). Another framework can be implemented by analysing the genuine user activity of sending friend request and putting a threshold on suspicious cloned profile (having same name and other attributes to that of user) for sending invitation to join clone account. Moreover, another criterion can be added to the future work which can first do the analysis on the type of messages sent by user and then compare them with the type of messages sent by the suspicious user. This message similarity feature (content based) can be combined with content free approach of record of recent IP sequence. Since OSNs server store many variants of cookies, therefore this information can be utilized for detecting clones. Operating system information, browser used, location records, pages visited, etc. are some of them. These parameters can prove to be beneficial for efficient detection of clones in Online Social Networks.

References