Spring 2017

Investigation on Security Issues and Features in Social Media Sites (Face Book, Twitter, & Google+)

Puneet Kumar Goud Kandikanti
Governors State University

INVESTIGATION ON SECURITY ISSUES AND FEATURES

IN SOCIAL MEDIA SITES (FACE BOOK, TWITTER & GOOGLE+)

By

Puneet Kumar Goud Kandikanti
Bachelor of Technology,
Jawaharlal Nehru Technological University,2014.

THESIS

Submitted in partial fulfillment of the requirements

For the Degree of Master of Science,
With a Major in Computer Science

Governors State University
University Park, IL 60484

2017

ABSTRACT

Social media sites allow users to communicate and share their information which are a matter of privacy for users, so users should be aware about its limitations and disadvantages to use social media sites. Likewise, there are many social media sites with its different features and it typically works with the latest technology that is provided by the experts to get connected and go along with the flow. The online privacy issues have been a real time problem and these however is the main aim for the experts to reduce the problems while sharing the kind of content that is allowed by the social media sites. There are issues that are general and the public need to oppose for the privacy terms and conditions.

People these days are concerned about the information that they post on the sites such as Face Book, Twitter and Google Sharing the photos and the latest features are now a sign of a problem for many users. In this research paper, Researcher will explore key information about the privacy issues and problems reported by social media users while using social networking sites. Being a personal user of popular social networking sites, researcher faced privacy concerns that initiated me to conduct a research on actual facts and figures behind the privacy issues among the social networking sites.

Nowadays, social media sites are widely used by hackers and un-authorized users where over usage of social media users from different geographic locations lead to increased privacy issues across these sites. In order to resolve the privacy concerns, the social media administrators have implemented many secured anti-privacy attack technique techniques but still they are not totally successfully providing 100% security to the user privacy over social networking sites.

Researcher found this issue to be a serious concern in current cyber crimes for which we have decided to conduct a research on this topic. Researcher had conducted research on privacy

issues and found that there are loads of security concerns in terms of privacy issues such as user privacy while photos and videos upload user privacy during messaging and chatting, user privacy during shares and uploads. All these issues seemed to be critical in terms of security where there is a huge necessity for implementing effective security techniques that are highly capable in reducing privacy issues and ensure 100% privacy to the social media users.

With use of this study researcher will explain the key privacy concerns reported in social media sites and current approaches available to reduce those issues. Researcher will explain the current scope of using those techniques and their limitations in providing privacy to the users. Finally, with use of this paper, researcher will investigate and propose best security technique that can be implemented to reduce the privacy concerns across social media sites.

Contents

CHAPTER1

INTRODUCTION CHAPTER

Aim of the Study

The objective of this study is to investigate and compare the security and privacy features and issues in social medial sites i.e. Face Book, Twitter and Google+ and recommend certain security techniques to overcome those issues.

Objectives of the Study

This study has some major objectives, which are as follows -

- To check the data relating to the security issues in social media sites
- To review the features, pros and cons of social media sites
- To investigate privacy issues seen in Face Book, Twitter & Google+
- To review different security techniques available to detect security issues in social media sites
- To recommend best security techniques to Social Media sites for protecting their online mediums from different security and privacy attacks.

Background Research

Blackmore (2010) stated that internet has become vital tool nowadays in day to day activities of humans and it is acting as key entertainment tool. Social media sites have become most popular with its entertaining support offered in terms of music, movies, audios and videos. The location based networks are that are shown in the smart phones and other such devices are rather a feature, but again an issue for the general public. The social media

websites while analyzing the privacy issues can enable users to be aware with the content they share (Choraria, 2012). There are cyber footprints that are recognized by the hackers and are now becoming a major fraud that realizes an individual on a later stage.

There are again some ethical concerns that are accelerated and these have a major amount of the collection of information by some private companies such as Google, Twitter and Face book (Fogel & Nehmad, 2009). This is a new insight that gives an aspect in routine life with a major element of innovation and creativity by the engineers. As it says that a coin has two sides, similarly using the social media sites is again a boon and a curse for the users.

The social media website is actually a space that is been watched by the companies and it can be protected by the companies (Case, 2012). Again, when it comes to be aware of the hackers and the cybercrime which is prevalent these days, it is basically a fear factor that is reported by different researchers when investigating upon the cases of privacy issues that are faced by the people while using these websites.

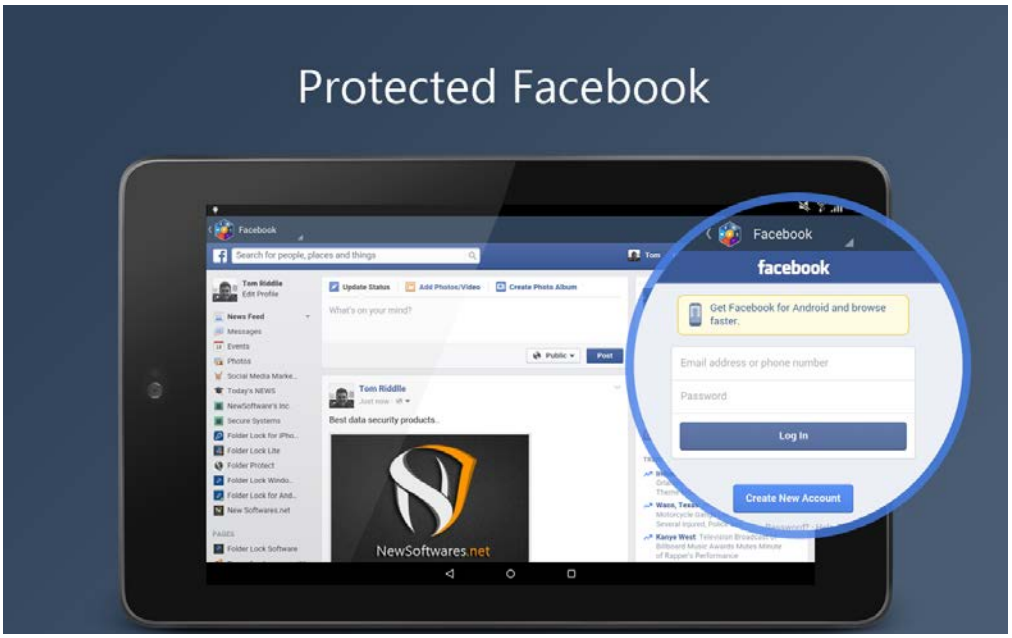

Figure 1: Social Media Security (Case, 2012)

Actually, there is no such privacy that is maintained while using social media websites, so more of the caution is taken by the user when they operate for any social media websites when it comes to sharing pictures and videos that may be a negative behaviour that is caught by the hackers or the cyber attackers (Godbold, 2013). As such the users are now aware and think twice while posting any information that may harm their lifestyle because internet is giving a chance to people who are around to gather personal information which may lead to risk in the near future.

Deliverables of the Study

The study has four deliverables that are explained below:

- Literature review on security issues in social media sites
- A hypothesis variables on social media security issues
- Comparison of security features in most popular social media sites like Twitter, Face Book and Google+.
- A project plan in the form of gnat chart for completing the project successfully

Ethical Issues of the Study

This research work was performed and it was successfully accomplished by implementing some principled amendments. Researcher has clearly followed all academic and while using the authors and scholars content in the study, researcher had not performed any kinds of illegal activities like copying or plagiarizing the research data.

Resources of the Study

There are three types of research methods such as quantitative, qualitative and mixed research methods. The narrative re-evaluate section of this study deals with critical analysis and assessment on research methods and their adoption in publishments of the research.

The methodology concerning qualitative research is based on the analysis of research developments. It is because of the statement that this methodology doesn't use any kind of mathematical explanation as well as figures for any kind of information to reach the required outcomes and results (Frels, 2012). As per Bryman (2007) in the quantitative research methodology, the research is totally concentrated on the pragmatic study and investigation.

For conducting the research on marketing strategies of an organization and its effectiveness positivism philosophy has been basically used for acquiring the detailed and necessary information (Kotler and Armstrong, 2012). The method of positivism philosophy method prolifically provides efficient methods for conducting the study as positivism philosophy supports and ensures involvement of huge number of samples which can be derived from mass people.

In this study, the quantitative research method integrated with secondary data collection method was taken into consideration. The information on security and privacy issues in social media sites was gathered from published articles, case studies, journals and web documents and are analyzed critically for getting the final outputs of the project without violating academic ethical rules and regulations.

- Poor reviewed articles
- Journals
- Web documents
- Books and book reviews
- Case studies

CHAPTER2

LITERATURE REVIEW

Industrialization and Need of Social Networking Sites

With the growth in industrialization in the past few decades the build up to technology has been quite astonishingly the most effective aspect for the present generation. With the significant help of technological boost internet has reached every corners of the world making all information flow smoothly. To a major extent the birth of internet has been a boon which has prolifically impacted our daily lives. The revolution of technology especially the birth of internet has carved out the ways to develop different social networking websites. The development of social networking websites such as Face Book, Twitter, Instagram, MySpace, Google+ and LinkedIn is the new trend of the present generation. All these websites are quite popular and are accessed by millions of people all around the world (Abhilasha Singh Rathor, 2013).

Figure 2: Social Networking Sites (Abhilasha Singh Rathor, 2013)

Among all the aforementioned websites Face Book has become the most popular social networking site with approximately one billion people active on Face Book. All these social networking websites have significantly become the most common platform for communication with relatives, family, and friends. These platforms are basically used to share videos, photos, thoughts and various types of information. All these things have also carved out for many notorious crimes which is termed as cybercrime. Presently social networking websites has become an active ground for cybercriminals. Cybercrime has rapidly taken over all the social networking websites and the cybercriminals sensitively exploit all personal information via the implementation of revere social engineering and social engineering (Odoemelam, C 2015).

Need of Security in Social Networking Sites

It is quite obvious that the users of the social networking websites share various information and therefore most of the users lose privacy when they share their important and personal information with strangers. It has also been noticed that sharing of information with the strangers has led many users to fall in honey trap. One of the most important concerns that have emerged with the usage of the social networking websites is privacy. It has been reviewed that most of the users are basically unaware of the fact about the various privacy risks that are prolifically involved with the sharing of sensitive information on the social networking websites. It is quite important that the users must know that the default settings on these websites actually share every bit of personal information (Black, Stone, & Johnson, 2014).

It is critically important for the users to change the default privacy setting options in a customized way by which their personal accounts and personal information will be more secured. Another major issue that has been seen is the security attacks. The security attacks

continue to be one of the major concerns for all the social networking websites users. Significant challenges of keeping the social networking websites more private and secure have become the most important concern for every single user. It is next to impossible for the users to make the social networking websites more secure by adjusting the privacy control without the effective and practical implementation of identifying the privacy evaluation (O'Keeffe, G. 2011).

Out of all the users in the current generation, many users are not having proper knowledge on protecting their personal information published on social networking websites. So the problems on privacy issues are often arising. To avoid this type of discrepancies each and every user must have the basic knowledge on the privacy related precautions as it will also decide their standard in using social networking websites (Sirohi, M. 2015).

But in practical, this is not happening due to the lack of basic knowledge in using the social networking websites among the users and many users are affected by the hacking groups who always look for such kind of opportunities to steal the useful information of other people. Hence, it is very important to put a check to these hacking organisations and this can be done by conducting a specific survey that helps in finding the users of the social networking websites who are facing difficulties and helps them to improve the awareness of protecting the information published in social networking websites mainly in face book and twitter as they are having maximum number of users (Dasgupta, n.d.2014).

Privacy Issues and Concerns in Social Networking Sites

Social networking privacy issues along with associated concerns in privacy has become one of the most debatable topic in the present age as participation is social networking sites have rapidly increased in the past couple of years (Stockman, J. 2012). Everyday

various articles and journals are coming up displaying the issue of online crimes associated with social networking websites as the usage of these websites are increasing at a rapid rate. Dashburst a prolifically renowned website in its blog mentions about the rise in cyber crimes quoting “privacy related issues are increasing day by day with the increase in the development of social media.” It basically deals with the major issues and dangers that an amateur can face while using the social networking websites like FaceBook and Twitter etc.

Therefore it is quite essential for an amateur user to read all the privacy policies, regulations and measures that are offered by these social networking websites for having a safe online networking. Moreover the blog also tells about the significant changes in privacy policies that are implemented by the most renowned social networking websites like Twitter and FaceBook. It also mentions clearly about how these changes in privacy policies has affected the users. One of the key reasons that have been cited by various online researchers about the tremendous increase in the users of social networking websites is that the users can be able to create strong bondage with the help of these social networking websites. This led the websites gain huge popularity within the people especially the young ones (Dejong, 2013).

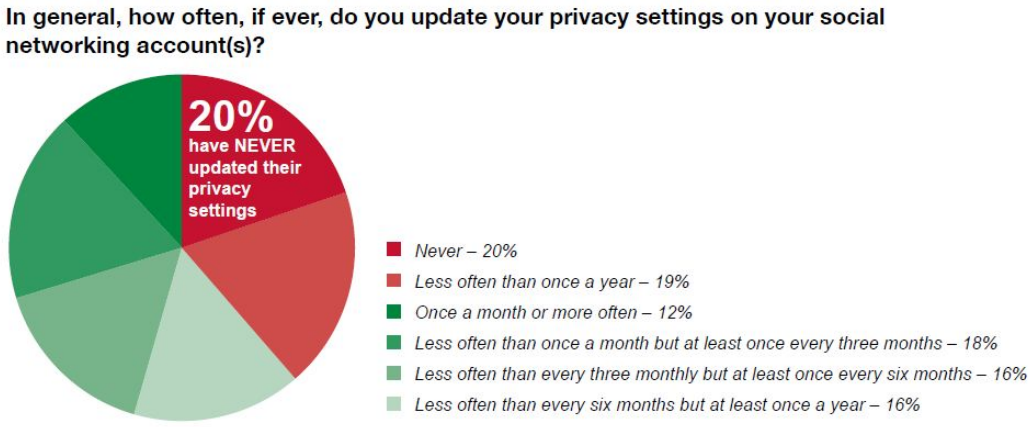

Figure: Facts and Figures on Social media privacy issues

Source: (Young, S. & Jordan, 2013).

All these social networking websites helps users to get in touch with the dear and near ones. Services offered by Face Book, Twitter, and LinkedIn etc. help the users to get connected to the virtual world. The latest trends in the social networking websites is updating and sharing private information about the day to day happenings. All these social networking websites basically targets the young individuals especially who are in the age group above 15 and below 25 years. Also it was noticed, most of the times the young users spend most of the day in accessing these social networking websites. According to reports most of the young users access these websites for the sake of connectivity and sharing major news or information rather than sharing it personally (Young, S. & Jordan, 2013).

Most of the young users use social networking websites to share all information by updating them. The young brigade has been so much influenced by the likes of accessing social networking websites that if they meet their friends in college in the morning they basically use the social networking websites to share any prolific news or information rather than personally sharing it via verbal communication. This significantly provides a clear understanding about the increasing dependency on such websites. It is quite obviously strange enough to know that the communication process in a broader way is getting hampered due to the advancement of the internet and the usage of social networking sites (Denise E. Agosto., 2011).

In 2009 Face Book, the most popular social networking website has implemented many strategies that are affecting the privacy of personal information of the users. This is recognised by checking the personal information published by the users such as their hobbies as well as their interests are the opportunities taken by the attackers to steal their information. It can be easily explained with the simple illustration that the personal information published by the users such as their drinking and smoking habits along with their most personalised information involving their sexual aspects can be targeted by the attackers without any complexity (Fly, R. 2011).

Figure 3: Security Issues (Fly, R. 2011)

Publishing the personal information in social networking websites will result in attracting the attention of the attackers that in turn results in hacking the user's information. Creating and organising an account is a very simple process that can be performed by any individual having basic computer knowledge. But it is very essential to have proper knowledge on protecting the information published by the user in the process of creating the account by the users in social networking websites. Otherwise the information will be hacked by the attackers who tries to steal others information all the time.

Issues with the sites of Social Media

Two issues regarding privacy of the Social Medias are discussed in this segment. The identity of the user is all known as the anonymity of the user. It has two approaches that are the identification of user identity on the social network and the issue of information leakage from the profiles of users (O'Keeffe, G, 2011).

Hacking Anatomy

Hackers have exceptional skills and are trained for the attacks which need to be understood clearly. Thus it is too easy for hackers to attack the user. While creating accounts on the social networking sites, users enter their personal information on various sites which are not mandatory for the account creation. Such non-essential information provides scope to the hackers to attack their personal information (Engdahl, 2010).

Attackers search for their victims through search process on social media. For stealing such personal information from the social media users hackers use different strategies. Here are the two methods that are typically used to attack social media accounts of the users:

De-Anonymization Attack:

The techniques of stealing from history and using information from group membership were mainly used by the Gilbert Wondracek and team members with him. The anonymity of the account holders of social media can be easily reviewed by the hackers. To use this methods attacker need to study users account on the social media site precisely, according to Gilbert Wondracek. Social networking groups concentrated because groups have more users than the single users over the social network (Partridge, K 2011). Thus it is a simple task for an attacker to access single users by concentrating on the group. These methods are often deployed to steal the information from the users who visit the site often.

The two unique links available on social media sites are mainly used by the attackers to implement this method (Espejo, n.d 2011.). With this method the section of the home of a user account is used and the link which is dynamically associated with the link which varies with each group or each single user.

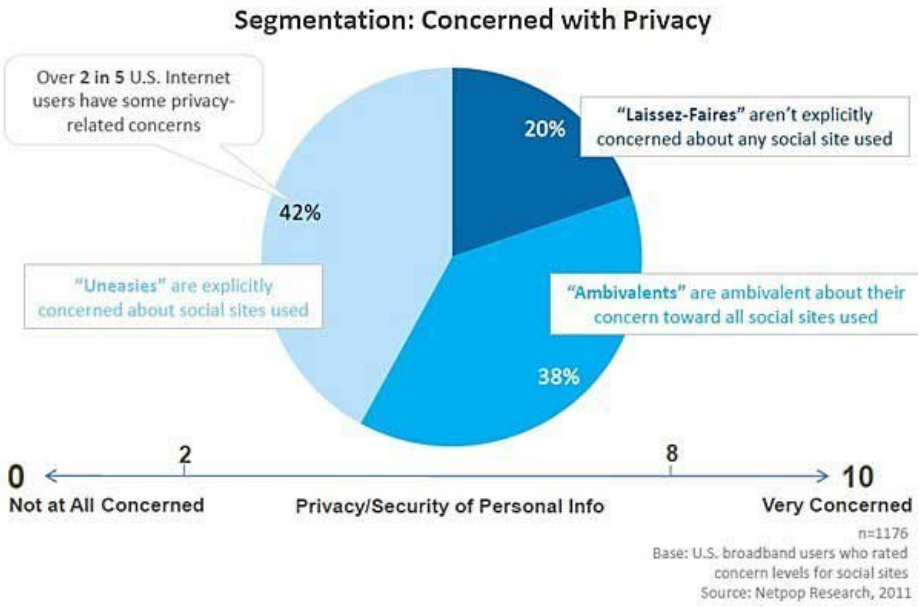

Figure: Facts on Privacy issues in social media sites

Example link: http://www.FaceBook.com/groups/groupID/

Attackers attract a number of users visit their developed website who implement their method for stealing personal information of the account holders. With such visit, attackers extract users browsing history by giving different URLs to them. Group directories can be used to obtain such URLs available on the social media sites. Attackers can go through browsing history of the victim and find out if the visited URL is on the list. To attract users on the developed sites conditional logic in Cascading Style Sheet (CSS) can be used by the attackers (Engdahl, S.2011).

With the use of the method of history stealing easy access can be performed by the attacker on the history list of the user. This list is used to find out often visited sites from the user account to find out associated activities of the victim on the internal. Mailing list of all the members in the group is specifically provided by a number of groups (Espejo, n.d.2011). With such obtained images to identify the person with his personal information.

Neighborhood Attack:

Outlines are used for addressing for social networking sites where social networking site users are addressed has a center point and user coordination of social media sites are addressed by the edge. Attackers efficiency is the prime factor to get information of such center point of victim's neighbor and coordination between the victim and neighbor to discover victim's center point easily (Abhilasha Singh Rathore, A 2013).

Information of the user's profile

Social networking site's sign up process asks professional and personal information of the user. In such data, personal information includes e-mail id, name, contact number, address and other details. In this personal information, most of the sites are the process of contact verification which makes contact number to be correct mandatorily whereas other information may be asked before publishing. Professional information includes work experience, the status of current work or past job that is essential. For accessing the account provision of required information is recommended.

Reasons of leakage of users profile

- Most of the users of social networking sites do not have knowledge of all the privacy settings available for them. Such users use their personal account on the social site through a personal computer which increases the chance for hackers to get access to

their personal data. Most of the social networking sites lack privacy settings like the FaceBook which make them less secure (Black, S., Stone, 2014.). Most secured social sites also have some flaws like friends of a friend of a friend can access personal information displayed on the user account.

- The social site like Twitter and FaceBook have APIs (Application Programming Interface) to create apps for third party developers to scuttle their policies. Such third party applications are popular among the social media users. When the user installs third party application on some occasion, they give permission of accessing user data automatically. Such applications have authority to publish their data on the timeline of user's friend (Fly, 2011). User's important information like their personal detail can be used without user's knowledge.
- With the user of third party domain services social networking sites activities of the user. Such access allows advertisement partners in aggregating and accessing information of the user to be used for their personal task.

Issue of identity theft

To steal personal information of someone can be turned as identity theft. Such personal information may include vital information or identity proof of the user. This identity can be used by someone else to perform malicious tasks that are undesirable. The social networking site has the potential of targeting such user information that attracts attackers as social networking sites have a tremendous number of user's information (GOYAL, 2012). Identity theft technique has a number of methods among which profile cloning is mostly used. In this method attacker deal with the reliance of friends on their friends account which is gain by making except them a friend request from the attacker. By the attackers for which they may not cross check user account.

The method of profile cloning:

This is the technique in which identity of the user is stolen which is defined as profile cloning. The main aim of profile cloning is attacking the account holders who have set their privacy setting to public so that all can see their information on the social site. Such public settings make it easier for hackers to attack information on the profile. They can duplicate or copy that user information to create their false identity. Here are the two major categories of profile cloning.

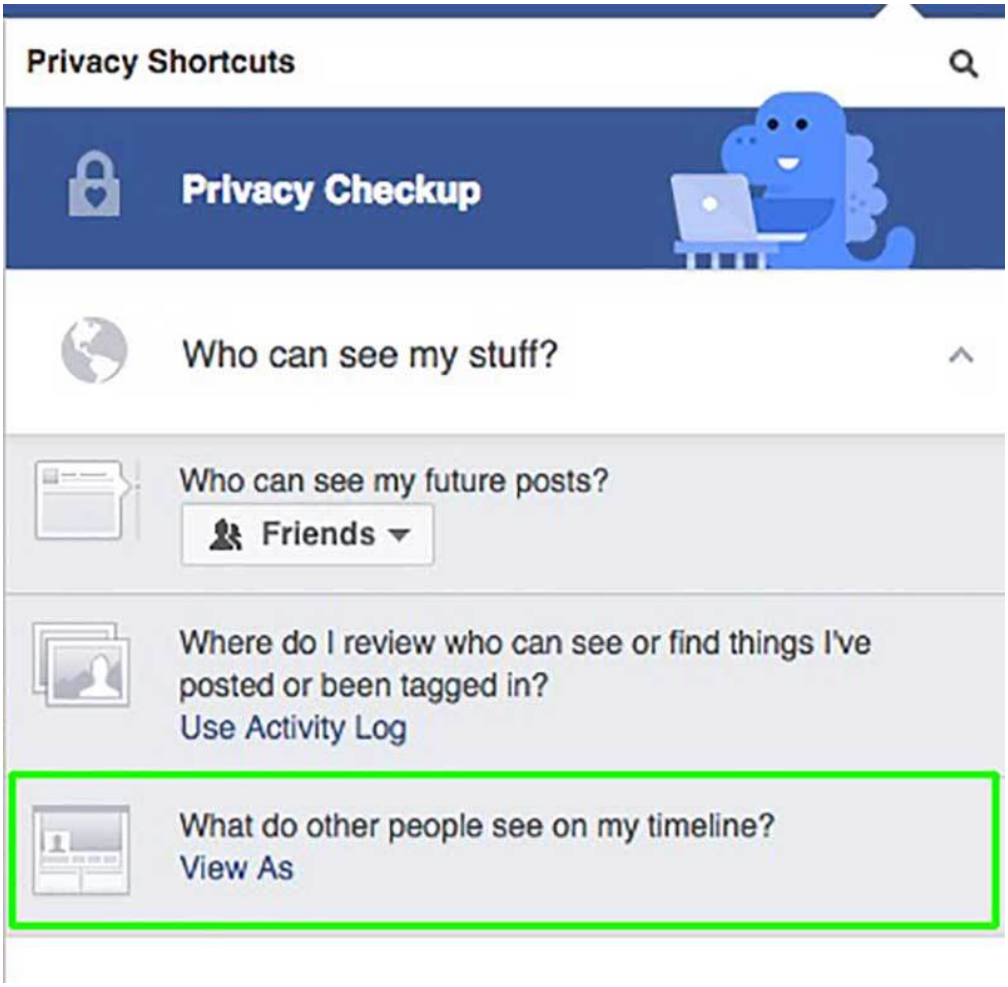

Figure: Profile Cloning issue

Source: (O'Keeffe, 2011)

- Existing profile cloning: The technique of existing profile cloning consist of the creation of a profile of already register users by the attacker. With the same name and other personal data, attacker steals the profile picture also which makes victim's friend relay on victims false account. With such false account, the attacker sends a friend request to victims friends of which attacker created the account (O'Keeffe, 2011). These friend requests can be accepted by victims friends easily by which attacker get the permission of accessing all the important information of that user.
- Cross Site profile Cloning: In this technique of cross-site profile cloning user profile is stolen by the attackers from the social media platform. Attackers register their account on some other social networking site with same information of the user account where he doesn't have his account. To create such account on other social networking site attacker use contact information. Attackers access the contact list of the same user to same friend request to victim friends on other social networking sites where the false account is created (GOYAL, 2012). When such friend requests are accepted by victim's friend attacker can hack their accounts also to steal their professional and personal information and even their sensitive information like identity.

Issue of Social Phishing:

Attackers design and develop an unorganized website in the method of social phishing. This unorganized website created by attacker has same interface and appearance like a real website. Such duplicate website attracts the victims with the provision of their personal information. This personal information may consist of their personal pictures, financial activity, their profile picture and even the sensitive information like password and security question.

Figure: Sample Social Phishing

Source: (Dasgupta, S. 2015).

The method of social engineering is used by the attackers especially to gather all the required information of the victim from their profile on social networking. This information is then used by the attacker to perform data extraction process automatically (Dasgupta, S. 2015).

Issue of Spamming:

What comes in different forms, “Spamming” is an indecent process that involves the sharing of unwanted messages. In most of the cases, it’s mode of propagation are emails. There might be a single person or a group of many such people that send irrelevant messages. Such people are called Spammers or spam attackers. While this process of spamming is enabled only through electronic media, social networking sites are the ultimate destination for spammers.

Figure: Social Spamming in Social media sites

Source: (O'Keeffe, 2011).

In this method, attackers access the emails of the user. Such standard spam attack may be non-beneficial for attackers as the email addresses are made indiscriminately or it can also be used to move toward open websites for scanning the email ids. To make this spam attack most of the social network display their strategy on their site which has proved very successful (O'Keeffe, 2011). HTTP grabbing method is used to make the site more powerful for protection from spamming.

This article will rightly deal with the phenomena of spamming, issues related to it, its effects on social media, spamming through emails and how HTTP emerges as a boon to spammers.

Spamming and social media:

Spam attacks have brought a kind of realization that social networking media namely Twitter and FaceBook are not really reliable when it comes to issues concerning spamming. Spammers or spam attackers can easily make their way into the account of other people. In a very convenient and secretive way, they hack accounts and share unsolicited messages and links to the user's contacts that mostly includes friends and relatives. When a follow request or a friend request is accepted by the user, things turn easy for a hacker. The hacker or the spammer can tour through the user's account and gain knowledge of it.

These days, people spend more time on social media. As a result, their activities increase. The hacker can then access through and begin his job. Mostly, spam messages include links and they also consist of messages that are designed in an attractive and an appealing way. The underlying goal is to attract as many users as possible to easily facilitate their advertising or malware practices. Nothing stays hidden on social media and while hackers or spammers always have a keen eye in order to hunt their victims, social media is a convenient platform for spam attackers.

Spamming through E-mail:

This form of spamming arises when unwanted and irrelevant messages, mostly commercial or advertising are shared to a large group of unknown recipients. In the present times, email spam is facilitated and spread by means of home computers that are infected by viruses. In such infected computers, a backdoor is already installed that provides a spammer with an easy access to the user's computer. Since in this case, the spamming problem is somewhat in – built, it becomes difficult for the user to control spamming. Social media provides the e-mail ID of users and as a result, getting to know about the e-mail address and other personal information becomes easy and handy.

Indiscriminate messages are then sent to the list of email addresses that are obtained. Some sets of messages are not clearly understandable and hence the user deletes them the very next minute once he realises that they are spam messages. Such type of spamming is termed as Broadcast Spam. In another type of spamming, the spam attacker gains information about the user through social media that states his date of birth, relationship status and other personal information. The spam attacker then makes matches thus obtaining email addresses. The job of spamming then continues in the same way. This type is the Context Aware Spam.

Hijacking of HTTP sessions:

This can be done conveniently and works effectively. On social websites, this hijacking is brought about which then helps the spammer in gaining knowledge about a particular victim. Things go the same way and then with the same information, the victim's partner is targeted. The process goes same as in case of Context Aware Spam.

Malware related issues:

Different types of techniques that are used by attackers or the ones that attackers opt for are discussed here in detail. (Paravastu Pattarabhiran 2012) A section of this part will also involve Twitter Worm and Koobface which are some notable malware issues. Till date, malwares and issues related to them spread efficiently across the networking sites.

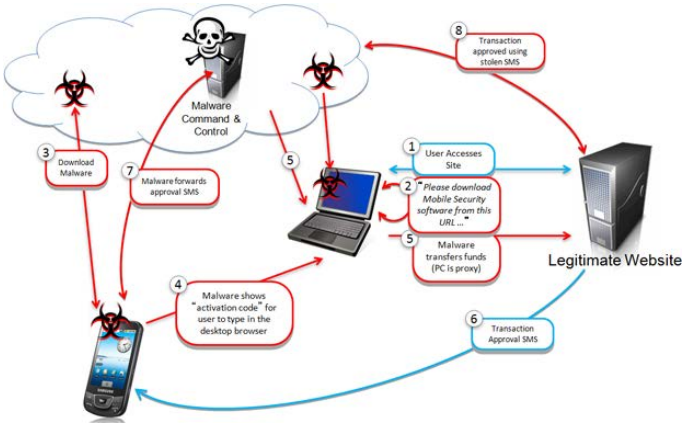

Figure: Malware Attack

Source: (Lan, Jin, Lu 2011).

Modes of Malware-spread:

The basic aspect over which the social networking sites or media are connected is interconnection between many internet systems. Co-ordination between these systems create conditions favourable for the spread of malware and issues related to them. There are a number of websites that still lack the parameters for the detection of harmful URLs or embedded links. This forms the best opportunity for the spam attackers. These defects in the social networks

sometimes direct the users to malicious websites that are capable enough of stealing information about the user.

Formation of fake profiles:

Attackers or spammers create fake accounts or profiles and design them in an attractive way that looks appealing. This draws the attention of many users thus making them victims and prone to spamming. Fake profiles are an effective way in which attackers can spread malwares across and amongst users.

Spread of malwares through applications:

There are many available applications everyone can have access to. When a user installs this application, the perfect looking application might hide with itself unwanted links. These unwanted or malicious links become the source of many malware issues. And unknowingly, the user becomes an effective mode of propagation of malwares here.

The “Malvertising” attack:

The term rightly denotes the spread of malware by means of advertisement. Hackers or spammers utilise this very well. They post attractive advertisements which are actually harmful. When the user clicks on the advertisement to have a look, he gets redirected towards harmful and malicious sites. And unintentionally then, the user who is now a victim installs harmful codes namely Java or ActiveX.

Condensed links

Mostly everyone looks out for shortcuts today. In this case, if you take a shortcut, you end up becoming a victim to many malwares. In condensed links, where the website names are shortened, hackers or spam attackers post attractive spamming or advertising content. Hackers make the use of fake URLs and trap users easily.

The attack of the XSS worm:

This attack is termed as the Cross-site scripting attack. The procedure of the hacker here is in a fixed pattern. They implement HTML codes and when the user uses these codes, the attacker gets the help as the harmful website sends in the information about the victim to the hacker. With such codes, a web program implemented which directs the virus into the user's computer. Once the computer gets infected with virus, number of ways gets open up for the hacker.

The "Click" trap:

Here, the hacker or the attacker makes it a point to attract the user or victim towards clicking a button or a website or a link. Once the button or the website or the link is clicked by the user, the underlying code that was hidden until now receives a trigger. This trigger gives the code a stimulus to facilitate some forms of malicious or harmful actions which the user is completely unaware of. Hence, malware spreads.

Notable examples of Malware spread:

Twitter Worm

Like every disease needs a mode of transmission, this worm also known as the Twitter Worm spreads through Twitter. In order to have access to a particular link, when the user tweets it further, the user then unknowingly is inviting malware towards it. In the same way, there is another type of a worm termed as Goo.gl worm. This worm binds the user towards the utilisation of the condensed forms of URL of the Google application.

Koobface

The mode of transmission of this type of worm is through messages. When users share or send messages with friends and relatives, this worm transmits itself and generates issues

related to malware. This worm significantly attacks many social media like FaceBook. These forms of messages consist of links that once clicked direct the users towards a video. In an attempt to watch video when the user clicks on the link, he sees another link that says that the video will play only when a “Flash Player” is installed. When the user does the same as said, he gives way to the entrance of viruses in his PC. Once the computer is infected, the attacker then gains information and details about the user and gradually advances towards the PCs of other users. As a result, this worm is potential enough to spread malwares across a large number of users making them victims.

You are physically threatened too:

The idea of internet or online is that the person is not physically harmed. In this case, the user may face online as well as physical threats. Physical threats here refer to stealing, stalking, harassment or robbery. All these forms can be done easily because when a hacker hacks an account, he is accessible to the user’s personal information. And sometimes he gains the user’s identity, his address, picture and other information too. So stealing, robbery and their allied forms become easy. Physical damages can be easily facilitated. The user never knows or gains knowledge of the hacker. As a result, anonymity is another advantage of social networking websites for hackers or attackers and not for the users or victims who really do not know that they are, in a way, spreading malware.

DATA ANALYSIS & FINDINGS CHAPTER

Current Research and Scope of Security in FaceBook, Twitter and Google+

With the rapid proliferation of internet and mobile internet devices, billions of people all over the world subscribe to different online social networks. The concept on online social network is so interesting and easy to use that the subscriber base of all the popular social networking sites have been increasing in leaps and bounds. Naturally, a huge data has been generating in these social networking sites on a regular basis. Technically this data is called “Social Network Data”.

Though the social networking sites are legally bound to preserve the data from theft and hacking, in some extraordinary conations they can share the data with a third party. Firstly, they can share the data with a research team who is researching on different aspects of social networking behavior of general public. Secondly, they can share the data with their advertising partners. This is a part of their general policy which is normally accepted by the registered members at large. The related advertising partners need such data to design and air the most target oriented social media advertisement.

In a multiple stream in social science, like sociology, geography, economics, etc. social network data analysis plays a vital role in many research methodologies. Researchers use the data they obtain from a social networking site in various types of research purpose, like a government welfare department use the data to understand the need for certain welfare service in a specific area, or for a specific demography.

So, in the above-mentioned situations, data needs to be published and it is not illegal, but in some situations, it may be a serious threat for the registered member of the concerned social

networking site. It is obvious that the demand for such a huge data stock is too high these days. Many categories of third parties, like researchers as stated above, advertisers, and app

developers are keenly interested in gather data from the social networking sites, and such demand has been increasing in leaps and bounds.

Privacy protection in FaceBook, Twitter and Google+

The raw data preserved automatically in the system of a social networking site contains many private information that a registered member may not want to get published. So, the members need to learn privacy preservation technique in his or her preferred social networking sites. Presently there are a few well-established micro-data preservation techniques available for the social sites. Privacy setting techniques like K-Anonymity, I-Diversity, T-Closeness, etc. are much used these days. Social network data are available in graphical form where the nodes represent an individual and the edges represent the link between the two nodes.

Privacy protection techniques have been developed keeping in the following aspects: the knowledge of the adversary and what utility will the data bear if published. Researchers have developed a special technique of protecting a member's data by using the concept of K-Anonymity. After much investigation, the researchers have come to the conclusion that an adversary possesses the same levels of knowledge like a targeted individual or his or her neighbours. Simultaneously, experts have proposed a practical solution against unwanted knowledge disclosure. The "Anonymized Social Network" obtained through this process is now utilized to answer average network related issues with the highest levels of accuracy. Here, a social network is modelled in the form of "k-subgraph" which anundirected labelled graph.

Techniques to secure social media

Privacy is a big thing today however protecting one's original work on the internet is not easy. With so many people sitting to copy content and publish it as their own, it becomes important to protect data in the best way possible.

In order to protect all the data, various researchers have put in lot of hard work to protect the privacy of publishers' contents. Some of the methods used are K-anonymity, L-diversity and sometimes integration of both.

In order to protect the data, two things are kept in mind and those are the knowledge of the adversary and how the data will be used after it will be released.

K-anonymity

Privacy disclosure in online social networking platforms matters a lot. When it comes to hackers and other adversaries, they have immense knowledge. They not only know the target's location and other information, but they also know about their neighbour's information as well. In order to reduce the public disclosure of important information, there has to be a practical approach and this has been devised by the name of k-subgraph.

However, this solution doesn't work in all the situations, even after reforming it again and again and this is where L-diversity comes in the picture which was developed in the year 2007 (Machanavajjhala et al.2007).

The privacy may be attacked by an adversary with the use of neighborhoods.

For a social network G, suppose an adversary knows Neighbor G(u) for a vertex u ∈ V(G). If Neighbor G (u) has k instances in G where G is an anonymization of G, then u can be

re-identified in G with confidence 1 k. Similar to the philosophy in the k-anonymity model [35], to protect the privacy of vertices sufficiently, we want to keep the re-identification confidence lower than a threshold. Let k be a positive integer. For a vertex u ∈ V(G), u is k-anonymousin anonymization G if there are at least (k −1) other vertices v1,...,vk−1 ∈ V(G)such that NeighborG (A(u)), NeighborG (A(v1)), . . . , NeighborG (A(vk−1)) are isomorphic. G is k-anonymous if every vertex in G is k-anonymous. Analogous to the correctness of k-anonymity model on relational data, following claim can be found. Property 1 (K - anonymity) Let G be a social network and G an anonymization of G. If G is k-anonymous, then with the neighborhood background knowledge, any vertex in G cannot be re-identified in G with confidence larger than 1 k.

Example of k-Anonymous Table:

	Race	Birth	Gender	ZIP	Problem
t1	Black	1965	m	02141	short breath
t2	Black	1965	m	02141	chest pain
t3	Black	1964	f	02138	obesity
t4	Black	1964	f	02138	chest pain
t5	White	1964	m	02138	chest pain
t6	White	1964	m	02138	obesity
t7	White	1964	m	02138	short breath

where k=2 and QI= {Birth, Race, Gender, ZIP} (Machanavajjhala et al.2007)

The above mentioned table shows an example of k-anonymity. Here, k=2 and QI (quasi identifier) = {Birth, Race, Gender, ZIP}. For each of the tuples given in the table T, the values of the quasi identifier appear minimum 2 times in T. Specifically, t1[QIT]= t2 [QIT], t3[QIT]=t4[QIT, and t5[QIT]=t6[QIT]= t7[QIT].

As per a theorem, let RT(A1.,,,,An) be a table, QIRT=)Ai,…,Aj) be the quasi-identifier related with RT, Ai,….,Aj⊆A1,….,Each sequence of values in RT [Ax] appears with at least k occurrence in RT [QIRT] for x=I,..j.

In the k-anonymity table T, occurrence of k in each value adheres to k-anonymity. Hence, every value is related with an attribute of QI in T and appears at least k times. |T[Race ="black"]| = 4. |T[Race ="white"]| = 3. |T[Birth ="1964"]| = 5. |T[Birth ="1965"]| = 2. |T[Gender ="m"]| = 5. |T[Gender ="f"]| = 2. |T[ZIP ="02138"]| = 5. And, |T[ZIP ="02141"]| = 2. It is hence proved that if k-anonymity is satisfied with released data RT with respect to the quasi-identifier QIPT, then the combination of the released data RT and the external sources on which QIPT was based, cannot link on QIPT or a subset of its attributes to match fewer than k individuals.

L-diversity

Panda et al. (2010) proposed an entirely new approach which was based on the amount of knowledge an adversary had on a potential target. Some privacy leaks are inevitable as many adversaries do have some background knowledge and it might not play out in the favour of the people who are using social media on a daily basis. To tackle this problem, the concept of t-closeness has been brought up and is being worked on.

L-diversity is a way to preserve the relationship privacy of the users. All the types of methods, k-anonymity, t-closeness, l-diversity and the integration of all these has been in use. However, they do lead to severe information loss and this is why better and more refined frameworks are being tested and will soon be launched to protect the privacy of the customers.

L-Diversity can be expalined with the following theorm.

Let q be a value of the attribute Q which is nonsensitive, in the base table T; let s be a possible value of the sensitive attribute; let q* be the generalised value of q in the published table T*; let $n_{(q^*,s')}$ be the number of tuples t* € T*

Where t*[Q] = q* and t*[S] = s'; and let f(s'|q*) be the conditional probability of the sensitive attribute conditioned on the fact that the nonsensitive attribute Q can be generalized to q* (Panda et al, 2010). Then the relationship is as follows:

$$\beta_{(q,s,T^\star)} = \frac{n_{(q^\star,s)} \frac{f(s|q)}{f(s|q^\star)}}{\sum_{s' \in S} n_{(q^\star,s')} \frac{f(s'|q)}{f(s'|q^\star)}} \qquad (1)$$

(Panda et al, 2010)

The Theorem facilitates to calculate the observed belief of the adversary. Let's define a q ⋆ -block to be the set of tuples in T ⋆ whose nonsensitive attribute values generalize to q ⋆ . Consider the case of positive disclosures; i.e., Alice wants to determine that Bob has t[S] = s with very high probability. From Theorem, this can happen only when:

$$\exists s,\ \forall s' \neq s, \quad n_{(q^\star,s')} \frac{f(s'|q)}{f(s'|q^\star)} \ll n_{(q^\star,s)} \frac{f(s|q)}{f(s|q^\star)} \qquad (2)$$

(Panda et al, 2010)

Contribution of researchers

To diminish the risk of privacy disclosure due to authorized data publication, the famous researcher Zou and his team has proposed k-automorphism (Zou, 2009). He and his team opined in their work that the antagonist has the knowledge about the degree, subgraph, and

neighbour of the concerned member. Tripathy and his team (Tripathy, Panda, 2010) invented an algorithm based on “Adjacent Matrix” that will help in the graph isomorphism. Lan contributed significantly when he developed KNAP, an algorithm to protect 1-neighborhood attack (Lan, Jin, Lu 2011). The famous researcher, Skarkala and his team used K-Anonymity to weigh social network.

K-Anonymity is widely popular, yet it has some disadvantages too and in some cases data may be leaked. Again, it is also proved by the researchers that it can’t protect data against attribute disclosure. In order o develop more authentic data privacy system, L-Diversity was developed by Machanavajjhala in 2007 and I-Diversity by Tripathy and Panda (Tripathy, Panda, 2010). Though it was though initially that I-Diversity is a good provision against the leakage of data, later it was proved that it has a few drawbacks. If the adversary possesses some knowledge regarding the sensitive attribute value of a member, some data may be leaked. The difference between the posterior knowledge, i.e. the knowledge gained by an adversary after seeing the release table and prior knowledge is the principal factor in leaking privacy.

Keeoing this in focus, Li and his team (Li, Zang, Das, 2011) proposed T-Closeness to keep the relatuionship privacy between the two users of the same social network. At the same time, I-Diversity anonymization model has been proposed to completely protect a user’s personal relationship privacy. On the other hand, to protect the individual privacy more efficiently, Kavianpour and his team (Kavianpour, Hossein, Ismail, 2011) proposed an algorithm to combine the best of K-Anonymity and I-Diversity. Tipathy and his team of researchers (Tripathy, Panda, 2010) proposed another algorithm that combines K-Anonymity and I-Diversity and offers multi-sensitive traits during the anonymization process. In their proposition, Tripathy and his team modified the algorithm for anonymization against neighborhood attacks. The proposed algorithm is a bit complex in nature and requires further improvement.

Yuan distinct a combined k-degree I-diversity anonymity design for the safety of structural information and people keeping highly sensitive data in a social network. Like K-Anonymity, many other privacy models are proposed that can prevent node re-identification in the data graph, but data of a person is not still 100% protected. As an improvement over structural anonymization method, an anonymization method has been invented by introducing the noise nodes by attributing the least possible distortion to the original graphical properties.

The above discussion reveals that much work has been done in the field of privacy protection methodologies for social networking sites. Many other research works are still underway and some of which have been showing a sign of good improvement in this realm. Here are some of those research works:

- One to one data encryption process was introduced by Guha et al in 2008. It enables only authorized users to decrypt or decode the result.
- A protocol was proposed by Blosser et al in 2008 where small networks will be able to transfer data while keeping the data of the owner completely safe.
- Social Network Greedy Anonymization was introduced by Campan et al in 2008
- Zheleva et al. proposed a way to preserve highly sensitive relationships in 2008.
- Ford et al proposed an mathematical equation for implementing p-sensitive k-anonymity on any data which is based on a greedy grouping method.
- Narayanan et al. in 2009 proposed new re-identification algorithm was developed particularly for anonymized graph.
- 2009 Lijie et al. showed a newer way of link detection attack where the adversary targets using linking opportunity, and t-confidence.

- 2009 Tootoonchia n et al. showed the Lockr, which is a system that develops the privacy level of the centralized online systems like Flickr and limited online content sharing systems like Torrents.
- 2009 Fong et al. projected an access control method which simplifies the privacy preservation system of FaceBook.
- The KNN and EBB is the algorithm which was introduced by 2010 Tang et al.
- 2010 Lan et al. proposed a method that can protect privacy of the social networks that can be distinguish as bipartite graphs
- Artificial dataset 2010 Ding et al. gave an organized overview of the de-anonymization attacks in social networking websites.
- A privacy conservation process for sharing data online with well-organized revocation for stopping a contact's admission right to the personal information was proposed by 2010 Sun et al.
- Q-Anon model was presented by 2010 Beach et al. to measure the possibility of getting hacked and data beaching in social networking sites by hackers.
- The existing anonymization process was further categorized into three simple graphs as Kanonymity, probabilistic privacy preservation and privacy reservation 2010 Wu et al.
- A cristical analysis was done on privacy and data breaches on social networking sites by surveying the literature 2011 Zhelva.
- The Social Privacy Protector software was developed to modify and upgrade the data security of FaceBook users 2012 Fire et al.
- A edgeperturbing anonymization enhancement technique was proposed which was primarily based on the structural role and frame among social networking theory.

- 2013 Heathely et al. showed the privacy preservation on circulated social networks which revealed it outperformed the primary SaNGreeA algorithm.
- 2013 Cheng et al. projected a structure to allow users to be in command over how third party applications can admittance their information in social networks even running the third party applications.

CONCLUSION AND RECOMMENDATIONS

By analyzing and examining current research process the key conclusion and recommendations has drawn the following assumption-

- *Promoting use of reputation techniques*: In order to detrmine the authenticity of the users and claims made by them, the use of reputation techniques can be a very effective approach which can help to prevent the possible threats to security. Possible uses may include reporting of identity theft, filteration of comments on basis of quality so as to enhance quality, filtration of spam comments, reporting of copyrighted or inappropriate comments and increasing reliability of widgets of third party.
- *Need of getting consent for including tags in images*: Tagging of images with profile tags or email address tags without taking consent of the subject image leads to violation of user's right of having a control over who publishes their data. The users of SNS should implement a system for giving control to users over who tags them in images.
- *Building automated filters*: Smart filters should be used. Filters should not be replaced from human intervention as they fail to understand cultural sensitivities or slang trends. Automated filters can determine if a specific piece of content that is driving high volume of traffic, is a legitimate content. When used in combination with reputed systems, they can prove to be highly effective.
- For privacy preservation techniques it is important to understand the worth of anonymized data. As a result for quantitatively measure utility of data an effective methodology is needed to be developed. There is requirement to analyze a variety of method for transaction among privacy and utility.

- For preserving the security of social media, techniques like K-anonymity and l-diversity can be used. However, both techniques possess certain limitations also, hence the decision of using a technique can be as per the requirement.
- Although there are many mathematical equations like L-diversity, integrated approach of k-anonymity & L-diversity that was improved to preserve confidentiality of sensitive user information in social networks however offered systems leads to considerable information loss.
- There is an urgent need for developing strategies for preserving confidentiality of dynamic releases for applications that need data publishing from time to time. The current Anonymization techniques are best for one time publishing of network data.
- There is a lack of social network circulated privacy preserving methods so it requires to be upgraded as safeguarding privacy for distributed tabular data techniques are available.
- By analyzing small datasets or synthetic datasets the current privacy preserving methods for social networks has been evaluated but there is a lack of empirical experiments on large data base systems which needs to be conducted.
- There is a lack of current and available methods and systems that can stop homogeneity cyber attacks.

REFERENCES OF THE STUDY

Abhilasha Singh Rathor, A. (2013). Social Networking Websites and Image Privacy. *IOSR Journal Of Computer Engineering*, *10*(6), 59-65. http://dx.doi.org/10.9790/0661-1065965

Black, S., Stone, D., & Johnson, A. (2014). Use of Social Networking Websites on Applicants' Privacy. *Employee Responsibilities And Rights Journal*, *27*(2), 115-159. http://dx.doi.org/10.1007/s10672-014-9245-2

Balmaceda, J.M., Schiaffino, S., & Godoy, D. (2014). How do personality traits affect communication among users in online social networks? Online Information Review, 38(1), 136-153

Case, D.O. (2012). Looking for information: A survey of research on information seeking, needs, and behavior (3rd ed.). New York: Academic Press.

Choraria, S. (2012). Factors determining the flow of information among the online community users. Journal of Systems and Information Technology, 14(2), 105-122.

Dasgupta, S. *International Journal of Virtual Communities and Social Networking (IJVCSN)* (1st ed.).

Dejong, S. (2013). *Blogs and tweets, texting and friending* (1st ed.). Amsterdam: Academic Press Inc.

Denise E. Agosto.,. (2011). *Teens, Libraries, and Social Networking* (1st ed.). Libraries Unlimited.

Dolvara Gunatilaka (2015). A Survey of Privacy and Security Issues in Social Networks. Retrived on 01-05-2017. Accessed from: http://www.cs.wustl.edu/~jain/cse571-11/ftp/social/index.html#sec7http://www.cs.wustl.edu/~jain/cse571-11/ftp/social/index.html%23sec7

Engdahl, S. (2010). *Online social networking* (1st ed.). Farmington Hills, MI: Greenhaven Press.

Espejo, R. *Location-based social networking and services* (1st ed.).

Fly, R. (2011). Detecting Fraud on Websites. *IEEE Security & Privacy Magazine*, *9*(6), 80-85. http://dx.doi.org/10.1109/msp.2011.161

Fogel, J., & Nehmad, E. (2009). Internet social network communities: Risk taking, trust, and privacy concerns. Computers in Human Behavior, 25, 153-160.

Godbold, N. (2013). An information need for emotional cues: Unpacking the role of emotions in sense making. Information Research, 18(1), paper 561.

GOYAL, S. (2012). FaceBook, Twitter, Google+: Social Networking. *International Journal Of Social Networking And Virtual Communities*, *1*(1). http://dx.doi.org/10.11591/socnetvircom.v1i1.732

Odoemelam, C. (2015). Adapting to Surveillance and Privacy Issues in the Era of Technological and Social Networking. *International Journal Of Social Science And Humanity*, *5*(6), 572-577. http://dx.doi.org/10.7763/ijssh.2015.v5.520

Machanavajjhala, A.; Kifer, D. & Gehrke, J. (2007). l-diversity: Privacy beyond k-anonymity," In: ACM Transactions on Knowledg e Discovery from Data (TKDD),

O'Keeffe, G. (2011). *CyberSafe* (1st ed.). Elk Grove Village, Ill.: American Academy of Pediatrics.

Panda, G.; Mitra, A.; Prasad, A.; Singh, A. & Gour, D. (2010). Applying l-Diversity in anonymizing collaborative social network" In: International Journal of Computer Science and Information Security, 8(2), pp. 324 - 329.

Paravastu Pattarabhiran, L. (2012). *Ads in FaceBook* (1st ed.).

Partridge, K. (2011). *Social networking* (1st ed.). New York: H.W. Wilson.

Shipley, T. & Bowker, A. (2014). *Investigating internet crimes* (1st ed.). Waltham, MA: Syngress.

Sirohi, M. *Transformational dimensions of cyber crime* (1st ed.).

Social networking advertisement. (2010). *IEEE Security & Privacy Magazine*, *8*(6), 63-63. http://dx.doi.org/10.1109/msp.2010.189

Stockman, J. (2012). Use of Social Networking Sites in Online Sex Crimes Against Minors: An Examination of National Incidence and Means of Utilization. *Yearbook Of Pediatrics*, *2012*, 11-12. http://dx.doi.org/10.1016/j.yped.2011.03.006

Young, S. & Jordan, A. (2013). The Influence of Social Networking Photos on Social Norms and Sexual Health Behaviors. *Cyberpsychology, Behavior, And Social Networking*, *16*(4), 243-247. http://dx.doi.org/10.1089/cyber.2012.0080