CHAPTER 2

LITERATURE REVIEW

2.1 Brief History of Social Networking Sites

Social Networking Site is a service present online that focuses specifically in building and developing social networks or create a new relationship amongst people. The people may share some common platform by sharing certain similarity in their interests, backgrounds or activities they perform. The services or sites allow the users to exchange a large amount of information like their views on a particular subject, their likes or dislikes, interests, etc and also provide some additional services.  

The development of first social networking site took place in the year 1994. Instant messaging service which became popular was AOL in the year 1997. Since then a number of social networking sites have become popular and developed drastically. Examples of some networking sites which are presently popular are: Twitter, Facebook, Google+, Friendster, My Space, Hi5, Pinterest etc.

A brief review of the history of the social networking sites evolution is given as follows:

The web-based social networking site which was developed first was Geocities in 1994. The users could create their own websites here. SIX DEGREES.COM and THEGLOBE.COM both emerged in 1997. They allowed users to publish content, create profiles and share similar interests.

In 2002 Friendster was developed as an initiator of utilizing the idea of web systems administration with certifiable companions. The number of users rose to 3 million. In 2003 MYSPACE was launched and truly like Friendster however has presented new elements. The other person to person communication locales like LINKEDIN, RIBE.NET, CLASSMATES.COM and so on were propelled in the next years.

In 2004 FACEBOOK was driven at Harvard University as a method for uniting all the U.S. understudies. Furthermore, with a crevice of two years, that is in 2006 Twitter was developed. It was a massive success with 300 million clients.

In 2008 MySpace was overwhelmed by Facebook and turn into the pioneer among the long range interpersonal communication locales. Shortly Facebook has a month to month dynamic client base of more than 1.44 billion [1].

![Evolution of Different SNS from 1997 - till date](image-url)
Figure 2.1: Launch Year and Registration Status of Different SNS.

Figure 2.2: Total Number of Registered Users in Different SNS.

2.2 Concepts and Applications of Social Networking Sites

Social Networking Sites revolutionized computers and communications around the world and paved the way for many technological evolutions to follow. They provide a platform to connect with people and it became even easier to search, communicate and stay connected. Also with internet services now available on mobile phones, connecting through social networking sites has become just a finger tap away.

In social networking sites, the users make a customized account known as profile that incorporates the data like date of conception, interests, inclinations, instruction status, relationship status and individual interest and so forth and add one another to their relating informal organization. It is a site that gives a spot to individuals to post on a specific subject or arbitrarily at worldwide stage (Tuunainen et al., 2007).

Hence the popularity of ‘check-in’ feature wherein a user can get his current position identified using GPS and then posts the location on SNS. This feature is used mainly when the users are travelling and want to keep their friends posted about new expeditions. Social networking sites are now used by millions of people around the world. Ellison (2007) suggested that, social networking sites are Web based services that allow individuals to construct a public or semi public profile within a bounded system.

Prior examination proposes that the primary inspiration to utilize online social networking is to impart and to look after connections. It was found that distinctive cooperation ceremonies are performed on a SNS for remaking the built up of informal communities. Mainstream exercises incorporate overhauling individual data, sharing photographs, getting overhauls on the exercises by companions, sending messages secretly or posting open testimonials (Ellison, 2007; Lehtinen et al., 2009; Zaideh, 2012).
These long range social networking sites have turned into one of the greatest commercial enterprises. With a specific end goal to stay at the top, new applications and components are added from time to time, to make it easier to understand else its client base can fall radically and piece of the overall industry can lessen as it happened if there should arise an occurrence of Orkut (Kaur,2015).

Researcher Li et al, 2013 states how the data can be prevented from malicious users for users privacy preservation using HBC model.

The applications of a designed framework by Shaon, Hasan & Khan, 2015 is used to recommend friend, group that matches with the users behavior or interest. The behavior of the user is completely related to the activities of the users. Users can do different activities to determine a new recommendation system having more accuracy.

Yuan et al , 2014 implemented an image-centric privacy social framework to expand user’s friends with common interests effectively and securely into cloud as storage back end. Our system is deployed under modern architecture, which leverages cloud as image storage back end.

### 2.3 Benefits of Using Social Networking Sites

Due to huge popularity of SNS’s, they are being used for marketing and advertising purposes. Such is their impact; even information posted on these sites is being used in computer forensics for legal and criminal investigations. With growing usage, come many benefits but also many concerns. Furthermore, social networking sites are also used by a number of organizations for marketing campaigns. It is considered as an opportunity to present oneself and interact with many people worldwide. Though most common social networking site like Facebook, provides data access and privacy control but, user's information is not sufficiently protected, which results in information leakage and thus affects privacy in general which is hard to measure and examined (Krombholz et al, 2012; Stutzman et al, 2011; Acquisti and Gross, 2006)

There was an ongoing debate that whether social networking sites actually bring people
closer or take them away from each other by isolating them. Researchers have investigated how much a student is aware of the available privacy issues and the protection offered by few SNS. It is observed that most of the respondents are aware of the consequences of providing personal information but they are comfortable in posting and do not take any initiative to protect the information. Also, participant’s personal information can be accessed by three group of stakeholders in a SNSs, namely the network, the hosting site, and the third parties, thus from here the information is knowingly or unknowingly revealed by the participant (Govani and Pashley, 2005).

The analysts found that clients who upgrade their status are spurred primarily by a yearning for expressive data sharing, though people who post remarks do as such for unwinding diversion, brotherhood, and social connection. Two intentions, proficient headway and social connection, were found as fundamental reasons why user’s sent private messages. And thus, the utilization of gatherings on the site was absolutely impacted by expressive data sharing (Smock, 2011).

Ellison et al., (2007) concluded that there are advantages in social investment as a consequence of sharing data in a SNS that may constrain the attraction of broad security controls. The methodology displayed client’s availability to shared information by all the user’s security intrigues.

2.4 Previous Similar Research

Some of past work that is most significant to aggregate protection administration for Social Networking Sites is examined herewith. A few studies have led to protection and conceivable dangers which users face when ineffectively securing their own information (Acquisti and Gross, 2006) in SNS. Carminati and Ferrari (2008); Felt and Evas(2008) & Carminati et al.,(2006) gave an intriguing examination of user’s security disposition over SNS and highlighted that on-line companionships can bring about a larger amount of revelations because of absence of genuine contact.
To adapt the security and protection issues, SNS are right now augmenting their entrance towards control-based components, to enhance in adaptability and point of confinement for undesired data exposure. There is a general agreement that in SNS, another standard of access control should be created [Felt and Evas(2008); Carminati et al.,(2006); Gates (2007)]. A first endeavour along this bearing has been taken (Golllu et al., 2007), where an informal communication based access control plan suitable for internet sharing was exhibited. They proposed a methodology that considered ways of life as key sets and social relationship on the premise of social validations. Access control records are utilized to characterize the entrance arrangements of clients. Analysts have proposed a telnet based access control instrument for SNS for account authorization of complex approaches. Furthermore, issues authentications was developed for conceding connections, legitimacy and the customer side requirement of access control (Carminati et al.,2006). Our investigation of aggregate security administration does not identify with the protection of user’s connections. We concentrate on aggregate methodologies for security insurance of user’s shared substance. Another model proposed a substance based access control model, which makes utilization of relationship data accessible in SNS for protection issues (Hart et al., 2007). Another intriguing business related is HomeViews (Mannan & Oorschot, 2008), incorporated a framework for substance sharing supporting a light-weight access control system. HomeViews encourages specially appointed, distributed sharing of information between unmanaged home PCs. In the underlying sections we have streamlined the gaps in existing system and talk about recent research work done in these verticals of graphical authentication, biometric techniques, encryption and watermarking.

2.4.1 Graphical Authentication

Pictures are generally easier to be remembered or recognized than text. In addition, if the array of possible pictures is suitably large, then the space of a graphical password scheme may exceed that of text based schemes. The graphical authentication
presumably offer better resistance to dictionary attacks, shoulder surfing and brute force.

Table 2.1: Similar Previous Research on Password Authentication

<table>
<thead>
<tr>
<th>Technique/Author</th>
<th>Conclusion</th>
<th>Limitations</th>
</tr>
</thead>
<tbody>
<tr>
<td>Text-based Password</td>
<td>1. Long and random passwords needs to be created.</td>
<td>94^K (total 94 characters is the length of the password).</td>
</tr>
<tr>
<td>Zhao and Li (2007)</td>
<td>2. The typing of password should be fast.</td>
<td>Attacks- Dictionary attack, brute force search, guess, spyware, shoulder surfing continued on page no18</td>
</tr>
<tr>
<td>Perrig and Song (1999)</td>
<td>1. Select n number of image from the picture pool.</td>
<td>Takes longer to time to create than text password.</td>
</tr>
<tr>
<td></td>
<td>2. Study showed that many people remember pictures than text-based passwords.</td>
<td>N!/K!(N-K)! (N = total number of pictures; K = number of pictures in the graphical password)</td>
</tr>
<tr>
<td>Sobrado and Birget (2002)</td>
<td>1. Click within an area covered by pre-registered picture objects.</td>
<td>Hard to remember when large numbers of objects are involved.</td>
</tr>
<tr>
<td></td>
<td>2. It is very fast.</td>
<td>N!/K!(N-K)! (N =total number of picture objects; K = number of pre-registered objects) Attacks- Brute force search, guess</td>
</tr>
<tr>
<td>Man et al. (2003), Hong et al. (2004)</td>
<td>1) User has to type in the code of pre registered picture objects.</td>
<td>Users have to memorize both picture objects and their codes.</td>
</tr>
<tr>
<td></td>
<td>2) It can be very fast.</td>
<td>N!/K!(N-K)! (N = total number of picture objects; K= number of pre-registered objects) Attacks- Brute force search, spyware</td>
</tr>
<tr>
<td><strong>Passface [38]</strong></td>
<td><strong>Table 2.1 (continued)</strong></td>
<td></td>
</tr>
<tr>
<td>---</td>
<td>---</td>
<td>---</td>
</tr>
</tbody>
</table>
| 1) Recognize and click the preregistered pictures. | Includes more than text-based password as the choices for selection are somewhat predictable. \(N^K \) (\(K\) is the number of rounds of authentication, \(N\) is the total number of pictures at each round) | **Attacks**: Brute force search, guess, shoulder attacks  
Continued on page no 18 |
| **Jansen (2003:2004)** | **Table 2.1 (continued)** |  |
| 1) User registers set of pictures or a sequence of images based on different themes. | Slower than text-based password. \(N^K \) (\(N\) = total number of pictures, \(K\) = number of pictures in the graphical password. \(N\) is small due the size limit of mobile devices) | **Attacks**: Brute force search, guess, shoulder surfing |
| **Takada and Koike (2003)** |  |
| 1) Recognize and click on the pre-registered images.  
2) Users can use their favorite images and is easy to remember than system assigned pictures. | Slower than text-based password.  
**Attacks**: Brute force search, guess, shoulder surfing |
| **Jermyn et al. (1999), Thorpe and Oorschot (2004)** |  |
| 1) On a 2D grid users draw something. | Studies showed that drawing sequence is hard to remember.  
Password space is larger than text based password. But the size of DAS space decreases with fewer strokes for fixed password length.  
**Attacks**: Dictionary attack, shoulder surfing |
| **Syukri et al. (1998)** |  |
| 1) Draw signature using mouse. Needs a reliable signature recognition program. | It is hard to recognize.  
Infinite password space.  
**Attacks**: Guess, dictionary attack, shoulder surfing |
<table>
<thead>
<tr>
<th>Goldberg et al. (2002)</th>
<th>User draws something with a stylus onto a touch sensitive screen.</th>
<th>Very easy to remember, Infinite password space Attacks- Guess, dictionary attack, shoulder surfing</th>
</tr>
</thead>
<tbody>
<tr>
<td>Blonder[1], Passlogix [2], Paulson(2002); Wiedenbeck et al. (2005)</td>
<td>User needs to click on several preregistered locations of a picture in the specific sequence.</td>
<td>Is hard to remember. $N^K$ (N = number of pixels or smallest units of a picture, K = number of locations to be clicked on) Attacks- Guess, brute force search, shoulder surfing</td>
</tr>
</tbody>
</table>

2.4.2 Biometric Authentication

While the overall human structure is the same and it puts biometrics in great demand in the constantly updating field of security. Though the approach is still in its infancy, many people believe that biometrics will play a critical role in future computers, and especially in Social Networking Sites (Sukhai et al, 2004).

The lack of open source tools to research the sharing of biometric data forces the researcher to dedicate resources is difficult (Curmi, 2014). Biometrics includes a level of invasiveness that is obvious to the user. Through the use of vignettes, the study encompasses a systematically varied set of usage contexts for biometric devices to provide a generalized view on the model which constructs the need for privacy, security and perceived physical invasiveness of biometric devices as factors that influence intention to use (James et al, 2006).

In one of the researches, a hierarchical model is presented to compute the biometric sample quality at three levels: database, class and image quality levels. The method was based on the quantities of genuine and impostor matching score distributions (He et al, 2008). Zhang and Wang (2009), presented an asymmetry-based quality assessment method of face images. The method uses SIFT descriptor for quality assessment. The presented method has shown its robustness against illumination and
pose variations. Another asymmetry-based method is presented in [Gao et al. (2007); Sang et al. (2009) & El-Abed et al. (2011)] suggested a quality assessment method based on the use of two types of information: 1) image quality and 2) pattern-based quality using the SIFT descriptor. The presented metric has the advantages of being multimodal (face, fingerprint and hand veins), and independent from the user authentication system.

Table 2.2: Similar Previous Research on Biometric Technique

<table>
<thead>
<tr>
<th>Technique/Author</th>
<th>Conclusion</th>
<th>Limitation</th>
</tr>
</thead>
</table>
| RFID and EPC Hansen et al.(2008) | 1. RFID transponders with EPCs allow identification of individuals.  
                                   | 2. Individuals are tracked and targeted for advertising purposes but their names are not required.  
                                   | 3. Provide precautions to protect the privacy of individuals to prevent misuse by private or public entities. | Personal data may only be processed an shared on the legal grounds. |
| Fuzzy Scheme Sohn et al.(2009) | 1. A home network is constructed for sharing the content with other home networks. | No centralized authority is given. |

20
2. Encrypt the content by secret key and open it to the public.

| Trusted Social Network | 1. User satisfaction depends on privacy and data protection.  
<table>
<thead>
<tr>
<th>Chang et al. (2010)</th>
<th>2. An authenticate person is linked to a digital identity by a strong multimodal biometric authentication.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Security risk in online social network</td>
<td>Table 2.2 (continued)</td>
</tr>
</tbody>
</table>
| Creese et al. (2012) | 1. Key points and attributed data points are shared in social networking sites.  
| | 2. The potential for negative privacy and security model elucidates potential linkages between data exposed on social networking sites |
| Mobile social network system | 1. New mobile SNS system combines LBS, RS, and user facial authentication technology.  
| Kwon et al. (2013) | 2. The user authentication is a secure way to protect personal information on a mobile SNS |
| The technology described here could have a limited effect on user. | continued on page 21 |
| The system access memory more frequently due to real time image processing which involves complex calculations. |
Biometric in cyber security.
Gavriloa (2014)

1. The real and virtual worlds, based on multi-modal information fusion, template protection and social networks.

1. Intruder detection and prevention
2. Changes in databases over time
3. “Big data” challenges that demand real-time performance with high recognition rates on very large data repositories

2.4.3 Encryption Techniques

Encryption of typed messages and generalized suppression techniques are applied to provide security to the user’s data. The post or the message typed by the user is shared on social networking sites based on user choice for protection of personal information to safeguard privacy leakage.

Table 2.3: Similar Previous Research on Encryption Technique

<table>
<thead>
<tr>
<th>Technique/Author</th>
<th>Conclusion</th>
<th>Limitation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Face Cloak Luo et al. (2009)</td>
<td>1. The sensitive information about the user is stored in encrypted form on separate server. 2. Firefox profile is shared by only one user.</td>
<td>1. If keys are leaked, no revocation method is used.</td>
</tr>
<tr>
<td>Flybynight Lucas et al. (2008)</td>
<td>1. Public key encryption algorithm to encrypt the user message before sending it to the server. 2. Users private key is encrypted with a password</td>
<td>1. Facebook own application, so it can be removed at any point of time</td>
</tr>
</tbody>
</table>
| NOYB | Divides each users private information into atoms like name and age grouped together which is then replaced with the corresponding atom from a randomly selected other users. | 1. Deals only with profile information  
2. The number of users impacts its effectiveness.  
Enough information is required to recover friend information |
|---|---|---|
| Persona | 1. Effective means of creating applications, where user defines policy over access to private data not the SSN service providers.  
2. ABE and PKC functions are used. Also XML based formats are used to transfer data.  
3. User generate asymmetric key pair and distributes the public key to the user with whom they want to share | 1. Allow user to store private data persistently with intermediaries, and are not trusted by users.  
2. ABE operations are 1000 times slower than RSA.  
3. Worst when users fetch and decrypt 100 of data items continued on Page no 23 |
| BlogCrypt | Table 2.3 (continued)  
1. The logic and the user interface were programmed in Java Script and XUL. Also, AES encryption was implemented for easy retrievals and substitution which effects disclosure of private data.  
2. Security of data It does away with the use of trusted third parties, delegating the responsibility of the protection of the information entirely to a client-side application. | 1. No focus is given on the issues of key management and distribution.  
2. The important modules encryption, decryption and key management, are loosely coupled. |

### 2.5 Summary

A summary of main research ideas includes making of personal data flows more comprehensive for individuals that provide means and motivation for individuals to protect their privacy in this new evolving personal data ecosystem. We finish up from this review of existing research that the security prerequisites definition, assurance and investigation issues have yet to be tended to. Existing writing on the different
protection thoughts should be incorporated to give a system that can focus on pertinent security ideas and evoke security attentiveness toward a framework. Further, we have to assess whether existing routines for prerequisites separated are fitting for evoking protection concerns in given social connections. At last, we find that none of the proposed framework provides security necessities that can give the arrangement of ideas through which the different protection concerns, prerequisites, and the security arrangements can be connected. These are all deficiencies in existing research that we go to with our commitments in this theory.

Other works required the data to be re-encrypted and the keys to be re-distributed in a secure method and this mainly occurred with works that used ABE techniques. Data sharing and collaboration is still currently a strong focus of research today and in particular many works are focusing on solving the user information leakage problem as well as ways to manage the sharing and collaboration of user’s private data.

User's online practices and comprehension of social networking sites changed broadly. Stages utilized and the reasons these were utilized for was firmly identified with the courses in which users drew in with social networking sites, from making substance to just watching substance on the web. It is nothing unexpected then, with this variety in sorts of engagement that the recurrence of utilization shifted. Various methodologies for overseeing such online dangers were accounted for, including confining the sort and substance of data shared, and altering security settings.

Our thesis investigates the issues of privacy related to SNSs and presents the results of the study among users. This study was focused on, the use of Social Networking Sites and related privacy awareness to protect users information.

The study employs users of SNSs, and acquired data based on users response. We concluded that most of the respondents with their knowledge disclose certain amount of private information about them and are not aware of the visibility and leakage of information to unknown people and third party service providers. Also the privacy policy and terms of use are either not known or understood by the respondents.