Enhancing Privacy in Online Social Networks using Data Analysis

THESIS

Submitted in partial fulfilment

of the requirements for the degree of

DOCTOR OF PHILOSOPHY

by

AGRIMA SRIVASTAVA
ID. No. 2011PHXF413H

Under the supervision of

Dr. G. Geethakumari

BIRLA INSTITUTE OF TECHNOLOGY AND SCIENCE, PILANI

2015

BIRLA INSTITUTE OF TECHNOLOGY AND SCIENCE, PILANI

CERTIFICATE

This is to certify that the thesis entitled **Enhancing Privacy in Online Social Networks using Data Analysis** and submitted by **Agrima Srivastava** ID No **2011PHXF413H** for award of Ph.D. of the Institute embodies the original work done by her under my supervision.

Signature of the Supervisor	
Name in capital letters	**DR. G. GEETHAKUMARI**
Designation	Asst Professor, Dept. of CSIS

Date:

For my Grandfather (Baba)

Acknowledgements

Every mission needs a spirit of hard work and dedication but it also needs to be put on the right path to meet its destination and in my case this credit goes to Dr. G. Geethakumari, my supervisor. I thank her for her constant encouragement, continuous support, valuable suggestions and timely advice. She has been very supportive and has helped me in putting structure in my ideas and work as a whole. I am short of words to thank her for her affectionate behavior and patience throughout the duration of my thesis work.

I would like to thank Dr. Aruna Malapati and Dr. Y. Yoganandam, members of my Doctoral Advisory Committee; for their invaluable suggestions, assistance and support. I would also like to thank the other members of faculty in the Department of Computer Science and Information Systems; Dr. Ray, Dr. Hota, Dr. Gururaj, Dr. Bhanumurthy, Mr. KCS Murthy, Mr. Thakur, Mr. Powar, Mr. Prasanna, Mr. Samant, Mr. Kannan, Mrs. Rakhi and Mrs. Prafulla for their guidance during my course work, various presentations and interactions.

I am grateful to Director Prof. V. S. Rao (Hyderabad campus), for allowing me to carry out my doctoral research work in the institute. I am thankful to Prof. S. K. Verma, Dean, Academic Research Division, BITS-Pilani and Dr. Vidya Rajesh, Associate Dean, Academic Research Division, BITS-Pilani, Hyderabad campus for their co-operation and encouragement at every stage of this research work.

I was fortunate to work with my co-researcher Mr. K.P. Krishna Kumar who was patient enough to listen to all my ideas. We had some interesting discussions over the months that shaped the solution into what it is today. I would like to thank my fellow researchers: Meera, Pavan, Pratik, Neha, Anitha, Kiran, Muthu, Jagan and Balaji. I have learned from each and every one of them, and it has been a pleasure working together. I thank my friends Shalini, Pooja, Jayashree, Rukaiyya, Divyam, Jaya, Saloni, Swati, Harshita, Sabornee and Megha for being so wonderful all throughout.

My sincere gratitude to Mr. Dipesh Kumar Singh, my colleague at TCS for encouraging me to pursue Ph.D. He has helped me in keeping my spirits high on numerous occasions.

I would like to dedicate this piece of work to my family, whose dreams have come to life with me getting the highest degree in education. I would like to thank my Grandfather Mr K.G. Srivastava, my Father Mr. Sanjay Raj and my Mother Mrs. Meenu Srivastava for their efforts and sacrifices.

Above all I thank Lord Almighty for his blessing bestowed upon me.

BITS Pilani, Hyderabad Campus Agrima Srivastava

Abstract

Online Social Networks (OSNs) are widely used social computing platforms where on a daily basis huge quantum of personal information is shared by the users. This personal information is sensitive in nature and could be misused by the adversaries resulting in privacy violations. Hence, we need to study and enhance the present privacy mechanisms in order to reduce the chances of unwanted information disclosures in an OSN. In this research work, we have explored data privacy from the perspectives of OSN users, their online connections and the service providers. We identified the loopholes in the existing mechanisms and proposed efficient data privacy enhancing algorithms. The proposed measures would help the users understand the privacy risks and would enable them to prudently share their data online.

The three main challenges that are being addressed in the thesis are measuring OSN users' data privacy, enabling selective sharing of sensitive OSN data and preventing sensitive OSN data from inference attacks. We measured users' data privacy using the theory of psychometrics. We proposed a privacy settings recommender system to recommend appropriate privacy settings for the OSN users' profile. We analyzed the effect of node's topology on sensitive information spread and proposed the trust enhancing model to refine the existing trusted community. An efficient partial edge set removal algorithm is proposed to reduce the accuracy with which the attacker could infer the sensitive attributes. We also proposed a privacy utility trade-off algorithm that could offer maximum utility and minimum privacy loss in the released anonymized datasets. Our work aims to build efficient data privacy enhancing solutions which could protect the users' data against privacy attacks and make OSNs a privacy preserving platform for data sharing.

Table of Contents

List of Tables

List of Figures

Chapter 1

Introduction

"The question we should all be asking ourselves, our communities, our societies, and our leaders is this: does privacy still matter in the digital age? Yes, privacy still matters in this age of big data and digital devices. But what it means, how we regulate and enforce it, what we are willing to give up for it and how much power we give our governments over it." Terence Craig, Privacy and Big Data

Online Social Networks (OSNs) are one of the biggest advancements that have happened in the past decade. From its inception OSNs have attracted the attention of millions of users. Some of the popular OSNs are *Facebook, Twitter, LinkedIn, Pinterst, MySpace* etc. OSNs have gained tremendous popularity as it supports varying interests and practices of the users [1]. It helps in connecting people and encourages sharing and exchanging of thoughts and ideas. OSN users can create their profiles which are the online representation of themselves using which they can form connections with the other users in the network. These connections or relationships can either be bidirectional as in *Facebook* or can be unidirectional as it is in *Twitter* [2]. The nomenclature of connections for the OSNs varies from one social networking site to the other [3].

J.A Barnes [4] first introduced the concept of Social Networks and described them as connected graphs in which the *nodes* represent entities and the *edges* represent their interdependencies. These entities can be an individual, group or an organization and the edges between them can be their interactions, relationships and connections. Formally an OSN can be represented as a graph G(V,E) where *V* is the set of vertices and *E* is the set of edges [5]. OSNs can be connection based and used for business, enforcing real life relationships, socializing, instant messaging or can be content based and used for content sharing, resource recommendation, advice or news sharing etc. [6].

1.1 Security and Data Privacy Challenges in Online Social Networks

Security and privacy are the major concerns in many areas of Computer Science. Due to the involvement of highly sensitive data it is a topic of interest in OSNs too [7]. An OSN provides a platform to the users to share information with their friends. This gives rise to the problem of personal data security and privacy [8]. Many data security challenges such as *stalking, reputation slandering, private information leakage, personalized spamming, phishing* etc. are commonly observed in an OSN [9]. Some of the popular spamming attacks are the *context aware spamming* and *broadcast spamming attacks*. Structure of the network can also be used to launch the network structural attacks like the *sybil attacks* and the *shilling attacks*. These attacks can even spread worms and botnets which could propagate in the OSN via profile interaction and third party applications [10].

Security is necessary but not the sufficient condition for ensuring privacy i.e. security mechanisms alone do not guarantee the privacy of data. The sensitive data of an individual could be their *name, address, telephone number, E-mail, Social Security Number, credit card information, health and insurance details, financial records, personal photos, videos, notes* etc. Leak in sensitive data could result in lawsuits, loss of customers' confidence, brand damage, erosion of privacy, bad press, loss of revenue etc. [6][7]. In Indian context, the *Indian Information Act - 2000* has the provision for punishment, if people are found guilty of wrongful disclosure of information. [11].

In comparison to the real world where information is ephemeral, the information on the web remains for an infinite time thereby posing a great risk on the privacy of online users. Most of the time users are unaware of the potential risks involved when they are sharing any sensitive information online [12]. Whenever and wherever the *Personal Identifiable Information (PII)* is shared and stored, privacy concerns are bound to arise. Hence, it is difficult to preserve privacy in a domain which is inherently designed for sharing. No unauthorized user should get hold of any sensitive information of the data owner. Therefore, it is a great challenge to protect the confidential and sensitive data from unauthorized users and ensure that the actual data is available to the legitimate users as well [13].

1.1.1 Categories and Effect of Disclosures

The *PII* which is collected during an electronic service is one of the most important assets of information privacy [14]. OSN is enriched with data like the *photos, videos, posts, notes, likes, interests, address, date of birth, gender, education, job details* etc. and any unwanted disclosure of these attributes can harm the privacy of individuals [15]. In this section we will discuss three major types of possible disclosures in an OSN which are *identity disclosure, link disclosure* and *attribute disclosure* [16] [17]. The details of which are as follows:

- **Identity Disclosure:** Identity disclosure results in the disclosure of the identity associated with the entity. It occurs if a mapping from a profile p to a real world entity e is achieved successfully by the adversary.
- **Link Disclosure:** Link disclosure results in the disclosure of sensitive connection or relationship that a particular entity has with the other entities.
- **Attribute Disclosure:** Attribute disclosure is the disclosure of the sensitive attributes of an entity. These sensitive attributes could belong to the node itself, link connecting the node with the other nodes or the affiliations that the node is a part of.

Out of all the three major disclosures our work relates only to the problem of attribute disclosure.

Unauthorized users can significantly breach the data privacy if they gain access to the users' sensitive attributes. Figure 1.1 gives an overview of some of the sensitive attributes and their effects of disclosure on data privacy.

Photos and videos from the profiles could be morphed and used for blackmailing and defaming individuals. Likes and interests reveal a lot about a person and can lead to the formation of controversial opinions. Using address, the location of a person could be known which can result in a criminal attack or burglary [18]. Social Security Number (SSN) of individuals could be determined using a combination of address, date of birth and gender resulting in ID theft or impersonalization [15]. E-mails and phone numbers could be misused for targeted advertising leading to unnecessary interruptions and spam.

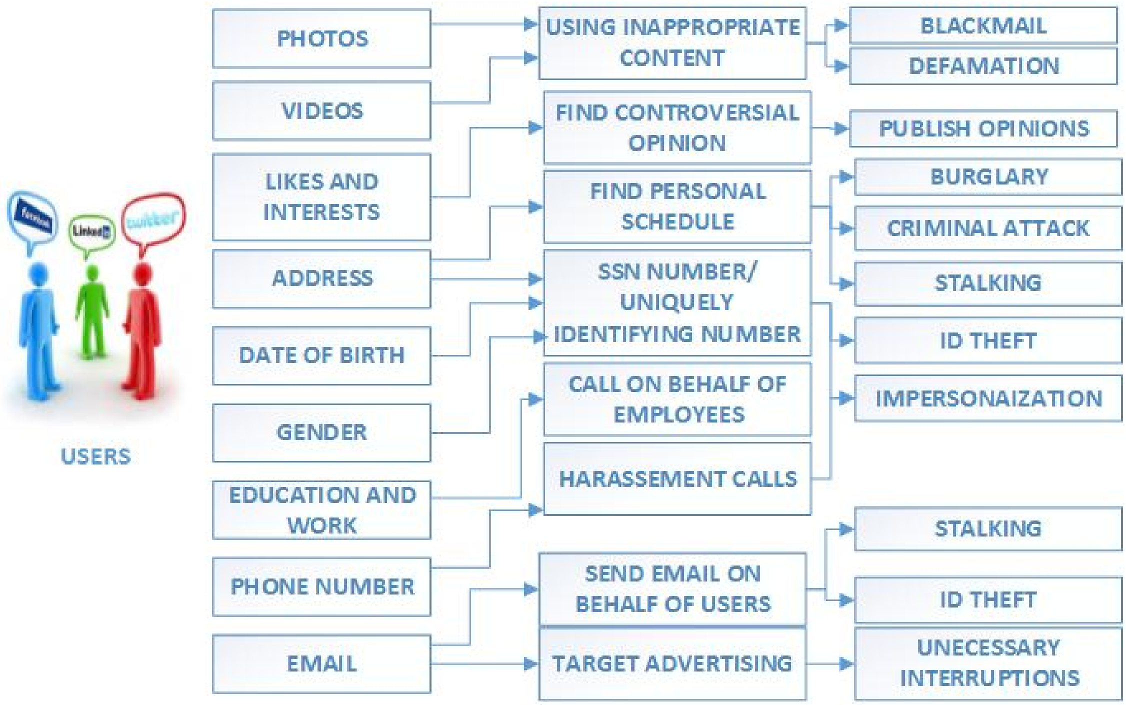


Figure 1.1: Effects of disclosure of sensitive attributes on data privacy

1.2 Objectives of the Research

We have divided the problem into three distinct sub problems and aim to find solutions for them. The three sub problems to be addressed are enlisted below:

- Measuring users' OSN data privacy.
- Enabling selective sharing of sensitive OSN data
- Protecting sensitive OSN data from inference attacks

Figure 1.2 gives an overview of the architecture for the proposed solution. The solution is carried out in four stages namely (i) Data acquisition, (ii) Data pre-treatment (iii) Data analysis for preserving privacy and (iv) Data Output.

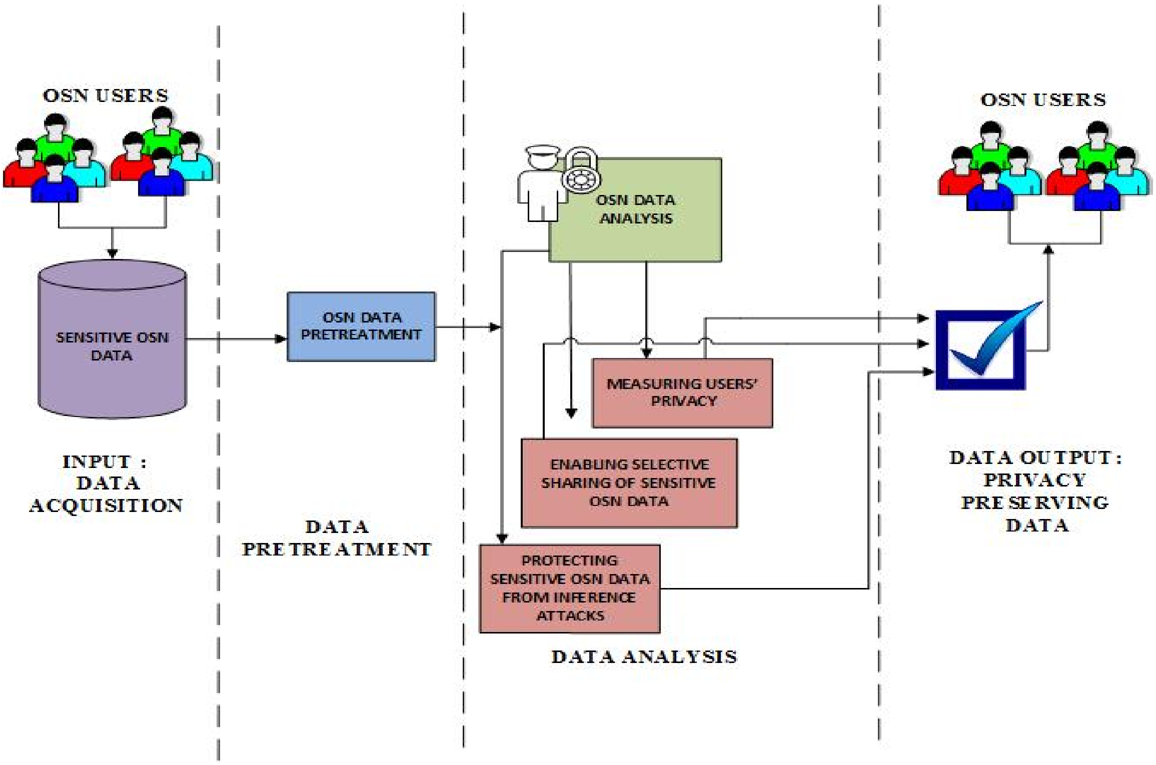


Figure 1.2: Proposed architecture for enhancing OSN data privacy

1.3 Scope and Problem Definition

Privacy and publicity are the two ends of information disclosure, whereas sociality takes the middle path. The path to sociality is taken at the expense of privacy [19]. If in the network there is no flow of information it becomes a static and asocial network. To stay social there should be a sense of digital literacy amongst the users using which they can define the boundaries for their information. *Social capital* [20] of a network is measured by the ability of the users to interact in social networks. With so many privacy concerns in place the users will not exchange ideas with ease and the social capital is bound to decrease.

The main entities responsible for the users' information disclosure are *the users themselves, their online connections and the service providers*. The details of each are stated as follows:

- **Users:** The users themselves could be responsible for sharing a lot of sensitive information about them. This usually happens when they are unaware of the huge privacy risks that follows or if they face difficulty in managing the privacy settings of their profiles.

- **Online Connections:** The users' online connections can further spread their information in the network. This implies that privacy of users alone does not depend upon them but is also linked with its connections. This could also be considered as a case of an *interdependent privacy* [21].

- **OSN Service Providers:** OSN service providers can share users' data to third parties for mining and getting deeper insights in order to improve their business solutions. The untrusted third parties could try to infer other sensitive attributes even if the released data set is anonymized. This can harm the privacy of individuals significantly and the service provider could be held guilty. This would eventually make the customers lose faith in the service providers degrading their market reputation.

In order to provide a holistic solution it is important to address the data privacy issues from the perspective of all the three entities which defines the scope of our research work.

More than a concept, privacy of an individual is a perception and differs from person to person [22] [23]. A lot of OSN users are not aware of the privacy risks in sharing sensitive information. Hence, there is a need to measure data privacy of the user with respect to their connections. Using a privacy measuring scale the individuals could measure their sharing proportions with respect to others in the network. Privacy is an abstract term and cannot be measured directly therefore measuring privacy is a challenging problem.

OSNs are actively being used by a large fraction of people who extensively share a wealth of their information online. If this information is retrieved, stored, processed and spread beyond its scope, without the users' consent, may result in privacy breach. The degree of privacy for an information depends on how it is likely to spread in the network. In a way privacy of a node depends upon its sharing behaviour and its topology in the network. Therefore, it is also important to analyze the network topology of the node in order to measure how privacy preserving it could be. Adopting a coarse grained privacy mechanism such as sharing information to a group of "close friends" or the strong ties of the network is one solution to minimize the risk of unwanted disclosure. However, this does not fully contribute in the process of protecting privacy. There is a high probability for an unwanted information disclosure even if the information is shared with the direct online connections in the profile. Therefore, it is important to quantify users' online trust before sharing sensitive information to them. In most of the privacy literature while building trust the online sharing behavior is never taken into consideration. Hence, there is a need to implement a privacy preserving model where such unwanted and unintentional sensitive information disclosures could be minimized by further refining the trusted community of strong ties with respect to their online privacy behaviour.

To prevent users from privacy threats the data holders such as *Facebook, Twitter, MySpace* etc. provide different privacy controls. Using these controls the users have a choice to hide or disclose the sensitive information in their profile. In order to improve their business solutions data holders often release the social network data and its structure to the third party which undergoes node and attribute anonymization before its release. However, this does not prevent the users from inference attacks which an untrusted third party or an adversary could carry out by analyzing the structure of the graph. Therefore, there is an utmost necessity to not only anonymize the nodes and their attributes but also to anonymize the edge set in the released social network graph. Where anonymization preserves privacy it also reduces the utility of the datasets. Finding an efficient utility based privacy preserving solution to prevent third party inference attacks for an OSN graph is an important challenge in the field [24] which has been addressed in our work.

1.3.1 Organization of our Work

In this work we have mainly modeled the OSN data statistically with a motive to enhance user privacy. Our work spans to three major privacy issues. The first issue relates with privacy concerns caused by the inappropriate sharing behavior of the users. This brings about the importance of measuring users' privacy and comparing it with the privacy of other users in the network. In Chapter 3 we will discuss the algorithms used for measuring and comparing the privacy values.

In Chapter 4 we will discuss about how the network topology contributes to unwanted information disclosure and what are the ways to prevent it. Trust can be used to monitor the private information flow in the network. This chapter also presents algorithmically quantifiable measures of online direct trust which can be calculated by observing patterns of private communication inside the network. Building a refined trusted community on top of the existing community of strong ties is also addressed in the chapter.

The third issue deals with countering the inference attacks. Data miners and researchers mine the social network data to discover hidden knowledge by applying various machine learning and graph mining algorithms. The released data could be exploited by launching the inference attacks in order to predict private attributes in the users' profile. In Chapter 5 we take up the issue of inference and map it as a classification problem. We propose and implement an approach to prevent OSN users from inference attacks and loss of privacy. Graph perturbation would result in the loss of utility of the data sets whereas releasing the actual data would result in complete loss of privacy. Hence, we also discuss an approach to find an optimal trade-off between privacy and utility such that the data holders could choose the best anonymizing scheme ensuring maximum utility.

1.4 Summary

In this chapter we brought about the relevance of data privacy in OSNs. Unwanted disclosure of users' information can result in data privacy breach leaving the users defiled and violated [25]. In OSN the data and identity are very closely linked and every information has a scope which is defined by the people with whom the information is shared. Disclosure of information beyond its scope leads to a privacy breach [6]. More than half a billion users are using OSNs and are sharing their details online [26]. With so many privacy concerns, the OSN users would prevent themselves from sharing information. This would essentially bring down the social capital of the online community making it asocial and stagnant. Therefore, we aim to develop efficient privacy enhancing algorithms and frameworks that could ensure users' information privacy, protect it from unwanted disclosures and maintain the social capital of the community.

Chapter 2

Background and Related Work

"Even if individuals have access to complete information about their privacy risks and modes of protection, they might not be able to process vast amounts of data to formulate a rational privacy-sensitive decision. Human beings rationality is bounded, which limits our ability to acquire and then apply information." - Alessandro Acquisti

2.1 The Concept of Privacy

Privacy is an important area and has its implications in fields like *wireless networks, social networks, healthcare networks, databases, data publishing, data mining* etc [27]. Privacy is derived from the word *privatus* meaning separated from rest. Privacy does not have a specific definition or a universal value which is same across all the contexts. Its value in a particular context depends upon the social importance of the practice of which it is a part of [28]. In a broad sense privacy can be described as the selective revelation about self. According to Warren and Brandeis [29] privacy is *"the right to be let alone"*. Westin's definition says that privacy *"is the claim of individuals, groups or institutions to determine that how, when and up to what extent the information about them are communicated to others."* [30]. Privacy is a sweeping concept and encompasses the freedom of thought, control over the information of oneself, freedom from surveillance, protecting one's reputation, protection from searches, interrogations etc. Problems related to privacy are not well articulated. As a result it is difficult to understand what is at stake when privacy is at risk and what precisely should be done in order to solve these problems [31].

Catering for such issues this chapter first explains the general concept of privacy and introduces its broad taxonomy. We discuss the relevance of data privacy in an OSN and throw light on the various user privacy concerns. We also examine the privacy perceptions of social network users in order to gather basic data privacy awareness amongst them. The chapter also talks about the various mechanisms that are used in order to preserve user privacy and identifies research gaps in the existing literature. In the next section we will discuss the privacy taxonomy and understand the basic concept and different facets of privacy in general.

2.2 Understanding Taxonomy of Privacy

Figure 2.1 shows a privacy taxonomy where we have considered privacy definitions by Altman et al. [32] [33], Palen et al. [34], Solove [35], Gürses et al. [36] and Papacharissi et al. [37]. The details of their classification are described as follows:

Altman privacy theory [32] views privacy management as a dialectic and dynamic boundary disclosure which are explained as follows :

- Privacy as a dialectic: If privacy is considered as a dialectic then it greatly depends upon expectations and experience of users and the ones with whom they interact.

- Privacy as a dynamic boundary disclosure: Privacy as a dynamic boundary regulation process is a continuous process of negotiation to decide a boundary between public and private.

According to Palen et al. [34] at any given time there should be a proper balance between *privacy and publicity, self and others* and *past and future*.

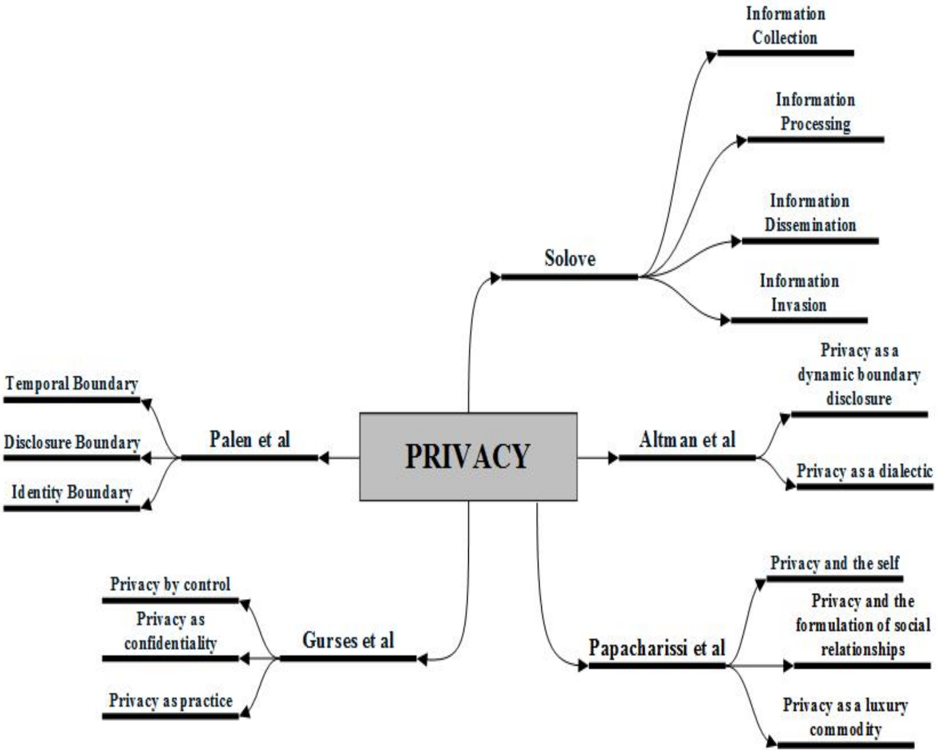


Figure 2.1: The privacy definition taxonomy

On the basis of which they have described three boundaries as the center for characterization of privacy management which are explained as follows:

- Disclosure boundary: Being a part of a society people share information which comes at the cost of privacy. Individuals should decide about an item's visibility before sharing it. The boundary which helps them decide the same is known as the disclosure boundary.

- Identity boundary: The identity boundary for an information is a boundary between self and others.

- Temporal Boundary: According to various observations the critical instances of information disclosures are related to each other i.e. an event in the past affects the present. There should be a temporal boundary for the shared information which is a boundary between the past and future.

According to Gürses classification [36] privacy problems can be divided into three categories which are i) privacy as control ii) privacy as confidentiality and iii) privacy as practice

- Privacy as control: The organizations offering electronic services are responsible to collect *PII* and process it. If this information is disclosed to unauthorized third parties or a broader public then it leads to privacy violation. Privacy is articulated through policies which is defined by users (privacy settings) or organizations (access control). Privacy settings, access control, auditing, purpose based access control are some of the examples of privacy research in this paradigm.

- Privacy as confidentiality: This mainly focuses on the data disclosure problem. Privacy is breached if the information goes beyond its visibility scope. Privacy as confidentiality enables a minimal disclosure such that the information cannot be linked back to the individual. For example, anonymous authentication protocols, anonymous communication networks and private retrieval.

- Privacy as practice: Privacy is not just an individual's matter it is indeed a matter of social concern. Users often decide their privacy and its dimensions based on the community they live in. This is concerned mainly with the feedback and creating general awareness amongst the individuals as well as the data collectors. The main privacy concern is that it becomes difficult for the user to understand how they should control their data and if this information is disclosed what inferences could be made on it. For example, P3P and privacy mirror are the technologies adapting to privacy as a practice.

Autonomy is the ability to build our own path without any external influence or impediment. Based on autonomy Papacharissi et al. [37] view privacy as self, privacy as formulation of social relationship and privacy as luxury commodity.

- Privacy and the self: The identity of an individual is unique but fundamentally social. The sense of self is developed through collaborative and collective experiences of the individual's social interactions. Performance of the self should apply to multiple audiences without compromising the sense of who we truly are.

- Privacy and the formulation of social relationships: Sharing personal information with the public makes it lose its meaning and inherent value. In an OSN platform the individuals form social relationships which encourages the loss of privacy. Hence, it is a challenging task to prevent privacy on such platforms which is explicitly designed for sharing.

- Privacy as a luxury commodity: The web accessible platforms offer services of social nature. They take the personal information and make money out of it. Information is treated as a commodity and therefore, for such web based platforms information privacy is one of the biggest luxury commodities.

The most comprehensive privacy taxonomy so far is the *Soloves's taxonomy* [35] which characterizes the four main stages of information. The privacy issues related to each of them are stated as follows:

- Information Collection: This stage deals with the process of data collection and privacy violation. Surveillance and interrogation are viewed as problematic in this stage.

- Information Processing: This stage deals with the usage, storage and manipulation of collected data. Aggregation, identification, insecurity, secondary use, exclusion are the harms associated with the stage.

- Information Dissemination: This stage deals with the revelation of personal data or the threat of spreading information. Breach of confidentiality, disclosure, exposure, increased accessibility, blackmail, appropriation, distortion are the major privacy concerns with information dissemination.

- Information Invasion: This stage deals with privacy problems when there are attacks on established systems. Intrusion and decisional interference are the privacy issues for the stage.

Our work mainly concentrates on the information dissemination part. Figure 2.2 gives the diagrammatic representation of the Solove's taxonomy.

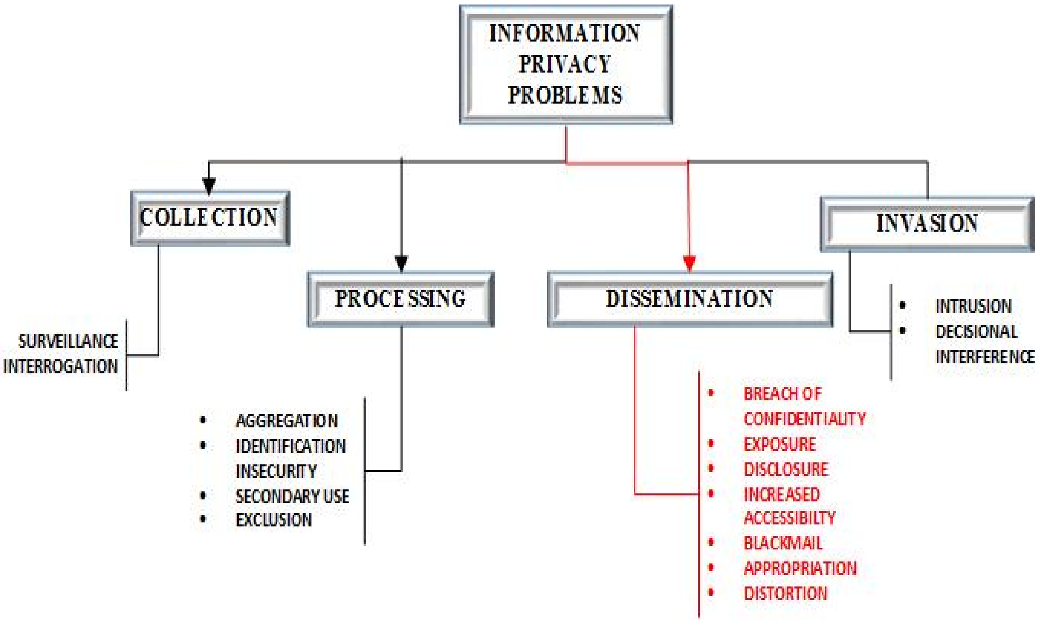


Figure 2.2: The Solove's taxonomy

The definition of data privacy is not limited to the above classification and differs from one context to the other. Data privacy is a serious concern and every individual should understand its importance and protect it. Sharing of information on a platform provided by OSNs have become a common practice [38]. If this information is used in a way which is not desired by the user, their privacy gets affected. Studies were carried out to understand online sharing behavior of people and to know how well do people care about their privacy and the way in which they implement privacy in their offline and online environment. In the next section we will discuss the privacy studies, their methodologies and the results obtained.

2.3 Privacy Perceptions of Users in Online Social Networks

OSN is a web based service that allows individuals to construct their digital profiles within a bounded system. Using such networks users can interact with each other and build and maintain relationships [3]. Data privacy is one of the biggest challenges in OSNs. It is multi-faceted; the individual would want some information to be disclosed but they might not like sharing it to the entire friend list [39]. A wealth of personal information is shared online on a daily basis. Sharing information makes individuals active and popular on these networking sites but a failure to control their PII can lead to privacy violations. In this section we aim to give insights on different studies that were carried out by researchers and understand some of the aspects of privacy like a) relevance of data privacy b) OSN privacy threats and c) reason for sharing personal information etc. Some important studies in this field are discussed as follows:

Liu et al. [40] have measured disparity between actual and the desired privacy settings of objects shared by the users. Their analysis was centered on knowing the ideal privacy settings and the actual privacy settings of the users. The study carried out by them selected 10 photos for query and collected ideal and actual privacy settings for the same. This analysis revealed that the users are uploading significant amount of content online and almost half of the content is shared with the default privacy settings which is desired by just 20% of the users. This study suggested that the default privacy settings are poorly chosen by the users and in most of the cases their expectations did not match the reality. Their work provides a deep statistical insight on the differences between user expectations and reality but does not mention the tools which would help the users bridge this gap.

Stutzman et al. [41] have looked at the association between network compositions, expectancy violation and interpersonal privacy practices of having a *friends only* profile. They drew Petronio's [42] theory of *communications privacy management (CPM)* which discusses iterative process of rule development. It regulates *who* to tell *what* and boundary coordination which develops disclosure ownerships and permeability rules in the network. Boundary turbulence refers to the dynamic process of maintaining and negotiating boundaries to maintain personal disclosures. They identified a range of factors like *gender, network size, weak tie expectancy violations* and *increasing level of interpersonal practices with privacy behavior* in the social network site *Facebook*. Their work concluded that the act of having a *friends only* profile is discretely notable. One of the limitations of their work is that the data is self-reported and has limited accuracy and recall. Their study also has the potential for non-response bias as under representative of males and non-white individuals were considered.

Wang et al. [43] have investigated regrets associated with the posts of users. They targeted sensitive topics, contents with strong sentiments, lies and secrets. They conducted a study and concluded that the participants regretted posting illegal drugs and alcohol use, posting photos depicting a different image of themselves, posting religious and political belief that caused debates. All these practices offended people and damaged relationships. The main reason why people post sensitive content online is either out of depression, frustration or anger. Comments made out of profanity, personal and family issues, expressing work and company in a negative way etc. are too sensitive to be made online. This work was an in depth study but the solutions for the same were not proposed.

Johnson et al. [44] recruited 260 *Facebook* users to install a *Facebook* application that surveyed the users' privacy concerns, their network compositions, the sensitivity of the posted content and their privacy preserving strategies. Their study showed that 86.2% of participants were unconcerned with the threats of strangers viewing their profile content. They could figure out that the threats from inside the network were more of a concern to the users. One drawback with their work was that the sample used in their work was biased toward users who were unconcerned with privacy.

Sleeper et al. [45] have looked at the types of contents that the users were sharing currently. Questions like why they choose not to share different types of content, how much will they share if they know the intended audience and what attributes define the groups with whom the users want to share their content were looked into. The study revealed that the participants are self-centered because they wanted to manage the way they presented themselves to various audiences. The participants indicated that they would have shared about half of the self-censored content if they would have the ability to optimally target audience. One of the biggest demerits of their study was the small sample size because of which the results were difficult to generalize. The method employed to collect data was the diary study which made the data biased.

Asking a user to directly evaluate the privacy concern is related to the emotional response and the results obtained are often biased. Braunstein et al. [46] have proposed an indirect technique for measuring content privacy concerns. A total of three surveys were carried out, the first one being an initial survey that does not mention about privacy or security, the second one to emphasize the security and privacy risks and the third one to explicitly focus on privacy. The privacy ratings were sensitive thus the privacy rankings were used as a ground truth for measuring. They have suggested mechanisms for translating responses to indirect questions into privacy ratings. They proved that this mapping highly preserves the relative rankings of content types from direct privacy surveys. In this study the use of privacy language is made extensively hence, according to the hypothesis the respondents might have adjusted their response to accommodate the goals of the experiment thus making the data biased. In Table 2.1 we show the merits and the demerits of different privacy studies that were carried out to understand the privacy perception of OSN users. The studies reveal that mostly it is the unawareness of privacy and its importance that results in a privacy breach. People find managing privacy settings a time consuming and confusing task. Therefore, there is a dire need for efficient tools and mechanisms to ensure privacy.

Table 2.1: A comparison showing the pros and cons of different privacy studies in the field of OSNs

Author	Pros of the study	Cons of the study
Liu et al.	Measured the disparity between actual and the desired privacy settings	No mention of tool/ method is made
Stutzman et al.	Studied association between network composition, expectancy violation and interpersonal privacy practices of having friends only profile.	Self reported data, Limited accuracy and recall, Non response bias
Wang et al.	Investigated the regrets associated with users' posts	No solutions were provided
Johnson et al.	Study of users' privacy concerns, their network compositions, sensitivity of the posted content privacy preserving strategies.	Sample used was biased toward users who were unconcerned with privacy
Sleeper et al.	Studied the sharing behavior	Small sample size, Difficult to generalize, Diary study resulted in biased data
Braunstein et al.	Indirect technique for measuring content privacy concerns	Use of privacy language was made extensively, respondents might have adjusted their response thus making the data biased.

Many researchers have come up with privacy preserving tools and efficient mechanisms to resolve the problems stated above. In the next section we will discuss different mechanisms in the present state of art using which the users' privacy is enhanced and preserved.

2.4 Mechanisms for Preserving Privacy

In this section we will be explore the tools used for preserving and enhancing privacy in the present state-of-art. We would also discuss the privacy mechanisms for enabling selective sharing and the various algorithms used for preventing sensitive data from inference attacks. Some of the important works in each of these areas are discussed as follows:

2.4.1 Measuring Users' OSN Data Privacy

Privacy does not have an exact measuring scale and its value depends on other hidden traits. Users in an OSN share their digital personal space which gives a way to many privacy risks like the *identity theft, stalking, target advertising, online victimization* etc. Knowing how much should be shared online is an important question in the field of privacy. In order to do so, it is essential to measure privacy and compare it with the other entities in the network. We will now discuss various tools and mechanisms that were proposed in order to provide the solution for measuring privacy in an OSN.

Fang et al. [47] have proposed *"Privacy Wizard"* which automatically configures the privacy settings of users' profile. The fact that the real users conceive their privacy preference according to an implicit set of rules is used to build the model of *privacy wizard*. The users of this tool are asked to assign a privacy label for a profile item with respect to a particular friend. This label takes either the value of *allow* or *deny*. If for a friend f and profile item i the preference is set to *allow* then this signifies that the friend f is allowed to see the profile item i. To intelligently request the user to provide labels to the most informative friends the wizard uses the *uncertainty sampling* as the active learning technique and then the classifier labels rest of the friends automatically. The wizard involves low effort, gives high accuracy, supports graceful degradation and works on limited data. The major limitation of this research is that the wizard was tested on a small data size of 45 *Facebook* users and does not look into the inference and shared data ownership issues.

Ghazinour et al. [48] have presented a tool “YourPrivacyProtector” for recommending privacy settings to the users. The tool collects the personal profile and interests of the users and use it to find the similarities amongst them. The tool uses the concept of *decision tree* to infer the profile type of each user and then makes use of the *k nearest neighbour* classifier for determining the privacy settings of the class the user belongs to. The tool was not tested on a larger dataset therefore, it does not demonstrate the scalability of the model. Also, it does not consider the sensitivity of the data items being shared which would otherwise have given better and improved results.

Mazzia et al. [49] have introduced *“PViz”* which is an interface that corresponds the way an OSN user models groups and privacy policies. The main goal of the tool is to make the users understand the visibility in a natural way. They extracted a hierarchy of communities according to a simple recursive process where the network is partitioned into communities and each community is treated as another network which is partitioned again. This continues until no further partitioning could be done to improve the modularity. *PViz* generates the initial set of labels for the communities and helps the users to visualize and understand their privacy policies. Their study and results show that *PViz* performs better than many tools in the present state of art. The evaluation was carried out on 20 participants which is quite less a number to generalize the efficiency and accuracy with which the tool performs hence, adds to the demerit of the study.

Biczok et al. [21] have defined online privacy interdependence and have modeled its impact through an interprivacy game. They modeled the game for 1 application and 2 players. They assumed that the players are non-cooperative and all players have a 'friend' connection. The strategy is to decide whether the application should be *installed (i)* or *not installed (n)* i.e. S = $\{i, n\}$. Both the positive and the negative externalities could emerge from the decisions of two players. Here the positive externality is having more users who would have installed the applications and the negative externality could be the privacy concerns for the installed application. Privacy interdependence, its effect on the user and the vendor welfare equilibrium was analyzed. This study does not take into consideration the amount and sensitivity of personal data stored in the given OSN user account which would have provided more realistic results.

Kafali et al. [50] have developed *"PROTOSS"* which is a run time tool for detecting privacy leakages in OSNs. The main technique involved here is model checking. *PROTOSS* uses network and agreement information to decide whether agreements are met or not. The two important techniques used in their approach are extracting commitments and checking models for the system. Commitment is an agreement from *debtor* to *creditor* about a property for the specific condition which can be represented as *C (debtor, creditor, condition, proposition)* and is used to verify that a given property holds or not. Here, the system is viewed as a state transition graph and the property as logic formula.

Liu et al [51] have shown a way to measure privacy using two parameters namely i) *sensitivity* and ii)*visibility*. Sensitivity of an attribute is directly associated with the privacy risks i.e. an item that is highly sensitive is not shared frequently and vice versa. Visibility on the other hand is a measure of information spread. They estimated the parameters using the maximum likelihood and expected maximization. Their mathematical model could fit the observed data and was used to calculate the privacy scores. Though there are different psychometric models like the rasch model, the two parameter logistic model, three parameter logistic model etc. but the reason to select the two parameter model is not mentioned in the study. Private information spread does not only depend upon the *sensitivity* and the *visibility* of the items but also depends upon the structure of the network. However, this aspect of privacy is not being considered in the study.

These were some of the tools and mechanisms that were developed and used for measuring, preserving and enhancing privacy in OSNs. The degree of security and privacy also depends upon graph theoretical properties of the social graph. In the next subsection we will discuss some of those mechanisms where the structure of the graph is used to give fundamental insights on the degree of data privacy.

2.4.2 Privacy and Selective Sharing of Sensitive OSN Data

The degree of data privacy in an OSN strongly depends upon the topological properties of the social graph. In this subsection we give an overview of studies where a relation between the structure of a graph and privacy is drawn. We will also discuss the effect of trust on preserving privacy in an OSN.

A social network can be represented using a graph G = (V_G,E_G), with vertices V_G = (v_1,....v_n) and edges E_G = (v_i,v_j) where v_i,v_j $\in$ V_G and i $\neq$ j. Here the nodes are the social actors and edges are the relationships between them [52]. According to Cutillo et al. [53] the achievable security and privacy also depends upon network theoretical properties of the social graph. They analyzed the relationship between the social network topology and the achievable privacy. The three main metrics that they looked into were the *node degree*, *clustering coefficient* and *mixing time*. The effect of each metric is explained as follows:

- Node degree: If the social graph is denoted by *G(V,E)* where *deg(v)* is the degree of the vertex *v*; p_{mal} is the probability that a new friend *n* of *v* is a malicious user then the event of befriending the friends F_{mal}*(v)* of *v* follows a binomial distribution. If user *n* gets access to the sensitive data of *v* then the disclosure could cause a severe damage. This concludes that the out-degree of a node is directly proportional with its usage control.

- Clustering Coefficient: The clustering coefficient *c(v)* is defined as the number of existing links between the nodes $e_{deg(v)}$ divided by the total number of possible links which is equal to $\frac{deg(v)deg(v)-1}{2}$. The tighter the friend set the broader is the disclosure of sensitive data to the users' contact.
- Mixing time: For a random walk the number of hops to reach a steady state distribution is called the mixing time. A small mixing time is required to enhance the security and privacy performance.

The study does not validate the results obtained after comparing the metrics and the reason for selecting the above mentioned metrics is not specified clearly.

Yildiz et al. [54] have proposed a solution to control privacy by automatically detecting *social cliques* amongst friends of the users. This social clique identified is used to create a friend list for the user. Their proposed algorithm starts by an initial clique *C* that consists of the number of participants *P*. The participants are the people who are directly related with the data item shared. In each iteration the clique is expanded by adding the candidates to it. Addition of the candidate maximizes the heuristic function *f*. The algorithm stops when the addition of the candidate does not satisfy the function. They used several clique expansion techniques like the *CLQ*, $BAND_k$, IN_k. They proposed that the $BAND_2$ *and* IN_3 schemes were accurate to form the final exposure set.

Private information spread is an important issue in an OSN. Lind et al. [55] have studied the general model of information spread which is suited for different kinds of social information. They proposed measures like *spreading factor* and *spreading time* which are accessible neighborhood around the node and the minimum time to reach the neighborhood respectively. These properties give an insight about the private information spreading in the network. However factors like decreasing the spreading factor and increasing the spreading time which would be needed to prevent gossip and spread of private information was not discussed.

Sharing a private information by the node *A* to the node *B* about the node *C* (who is not present at the scene) is known as gossiping and can affect the relationship between *A-B, A-C and B-C*. Unlike a rumor which is usually an issue or a matter related to public concern, a gossip deals with the behavior and life of an individual. Any analysis of spreading private information is done at the triad or at a higher level. Shaw et al. [56] have revealed that the information that is passed along one edge can affect the strength of the other edges. They conducted experiments and concluded that gossip decreases the *network clustering* and the *average node degree*. In this work the assumptions made are highly simplistic. The study considers only the negative aspect of gossip whereas gossip could be positive and conducive which is not justified.

Privacy and Trust

Trust has been extensively studied in computer science and other fields [57]. Buskens et al. [58] have discussed the explanations for the emergence of trust where the nodes are regularly informed about the behaviour of the other nodes and label them as untrustworthy. Adali et al. [59] have evaluated trust on the basis of communication and interaction amongst the members in the social network. They calculated the behavioral trust as *conversational* and *propagational trust*. Conversational trust is determined by the frequency of interactions between the nodes and the propagational trust is calculated using how the information obtained from one node is spread to the other nodes in the network. The different hybrid trust models that uses the interactions and social network structure for computing trust are discussed as follows:

Trifunovic et al. [60] have proposed approaches like the explicit and the implicit social trust where explicit social trust is derived from the strong ties based on the frequency of interactions and the implicit social trust is derived on the basis of frequency and the duration of contact between the users. Carminati et al. [61] have introduced a probability based approach that models the likelihood of information propagation from one social network user to the other users who are not authorized to access it. This probabilistic based approach estimates illegal leakage of resources in an OSN where access control is regulated according to the topology based paradigm. Li et al. [62] have proposed a trust aware privacy control approach that utilizes the inter-entity trust relationship to protect privacy. The trust calculated is directly proportional to the weight of the relationship. Therefore, most of the strong ties are considered as the trusted nodes. They use the privacy protocol that consists of audience, action and artifact control to ensure the disclosure of data only to the trusted parties.

2.4.3 Privacy Preserving Data Publishing

Researchers and third party agencies are interested to analyze the OSN data. Releasing this data in its actual form can cause unwanted disclosures. The data owner would want to release the data sets that could be useful for analysis without affecting the privacy of an individual. Basic step towards the release is to remove the unique identifiers in the data sets e.g. *Social Security Number (SSN), name* etc. and replace it with some random identifiers. Such an anonymization method is called the *Naive Anonymization* [63]. Removing identities of the nodes before publishing the graph does not guarantee privacy and the anonymized data set is still prone to attacks. Backstrom et al. [24] have discussed two types of attacks namely the *active and the passive attacks* that can be used on naively anonymized graphs. In active attacks the attacker makes new user accounts and then a unique subgraph *H* out of the newly created nodes, such that the subgraph could be uniquely identified in the anonymized graph. The adversary then identifies the nodes which have to be attacked and make links with them. After the anonymized graph is released the adversary could easily make out the subgraph *H* and identify the links it had set up from the false adversary nodes to the nodes of the genuine users. In the *passive attack* the attacker tries to find their identity in the released graph. Once they identify themselves the data privacy of the nodes with which they are connected is compromised.

One of the most popular algorithms in *PPDP* is the *k anonymity algorithm* proposed by Sweeney et al. [64]. According to this algorithm a data set has *k anonymity* protection if every user in the data set cannot be distinguished from *(k-1)* other users. *K anonymity* had issues like the *homogeneity attack* and the *background knowledge attack* hence, a new algorithm called the *l diversity* [65] came into existence. This algorithm ensured that the sensitive value in each equivalence class remains diversed. An issue with *l diversity* is that it is possible to gain information if the overall distribution of the attribute is known. It also assumes all attributes to be categorical which adds up to another limitation of the algorithm. Li et al. have proposed an algorithm known as the *t closeness* [66]. According to this algorithm the distribution of a sensitive attribute in any equivalence class should be close to the distribution of the attribute throughout the table where the distance between the two distributions should be no more than a threshold *t*. They made the use of the *Earth Mover's Distance* to calculate distance between the distributions.

For *PPDP* a lot of graph modification algorithms have been proposed. Some of the popular ones are *k-degree anonymity [67], k neighborhood anonymity [68], k automorphism*

[69], k symmetry [70], k isomorphism [71], k obfuscation [72] etc. Other methods like the *generalization [73], randomization [74], differential privacy [75] and accurate analysis of private network data methods* were also studied extensively.

Dyadic prediction deals with predicting observations associated with *dyads*. Menon et al. [76] have extended the recent dyadic prediction method for predicting labels for nodes and edges. This method learns latent features within log-linear model in a supervised way which maximizes the predictive accuracy for both dyad observations and item labels. Networks have been explored extensively for the past few decades. Recently, problems like the *influence, attribute prediction, identification of important nodes* in the network etc. have been studied thoroughly. All these issues involve labeling the nodes in the network. He et al. [77] have performed experiments to reveal that the personal attributes of a node could be inferred through their neighbors. They used Bayesian network to model the relationships amongst people in the network. They concluded that even if people hide their private information it could be leaked through other social connections by applying *bayesian inference algorithm.*

Zhelva et al. [78] have shown how an adversary can exploit the public and private user profiles to predict other private attributes of the user. They have used the traditional classification methods for the same. They have also discussed some models using which the sensitive attributes could easily be inferred. Heatherly et al. [79] have proposed three possible sanitization techniques i.e. *manipulating details, choosing details and detail generalization hierarchy (DGH)*. Manipulating details such as adding, modifying and removing details from nodes could be categorized as perturbation and anonymization. It is also important to choose the details that should be removed. They removed the most representative detail which had the highest correlation with the protected class label.

To prevent unwanted privacy breaches the social network graph is anonymized before it is released. Various graph anonymization algorithms could be used for anonymizing the social network graph [80]. These algorithms perturb the actual graph to produce the final graph which could be released for mining. Perturbation reduces the utility of the graph and gives a better privacy protection. The graph released with fewer modifications would have greater utility but would also increase the risk of privacy breaches. Balancing the right combination of privacy-utility is a challenging task. Li et al. [81] have proposed an integrated framework for considering the privacy utility trade-off. They borrowed the concepts from the modern portfolio theory which is mainly used for the financial investment. They have worked out the concept on various node anonymization algorithms such as *k anonymity, l diversity, t closeness and semantic privacy*. In the next section we will discuss some of the important research gaps in the present state of art.

2.5 Identifying the Research Gaps

Data Privacy is a vast area and has many unanswered questions for the researchers all round the globe. The study of privacy in OSN is in the nascent stage and a lot can be explored on the similar lines. We identify the following research gaps in preventing and enhancing privacy in the OSN domain.

- Privacy is abstract hence, measuring privacy is a challenging task. The tools discussed earlier ensure privacy but does not emphasize on measuring the sensitivity of the data being shared which would increase the accuracy and provide more practical results. Measuring privacy using an appropriate model which would incur the least information loss is highly important. There is a need for a holistic system which could recommend an appropriate privacy settings to the users after comparing the users' privacy with their respective friend list. Privacy if viewed from the prism of privacy enhancing algorithms has missing links and is a topic of high relevance.

- Private information spread does not alone depends upon the items' sensitivity and visibility. The extent upto which an information will spread amongst the users also depends upon the nodes' topology in the network. Offline people do not share their information to anyone and everyone. They selectively disclose their private information. The same behavior is difficult to replicate online. To minimize the unwanted information disclosure in an OSN a coarse grained privacy mechanism where the information can be shared to *"close friends"* is followed. This however does not guarantee privacy effectively. Incorporating *trust* before sharing private information in the network would ensure better privacy protection. Privacy and trust go hand in hand and quantifying direct trust would help in preserving the privacy of an individual effectively.

- Data in its original form contains sensitive information about individuals, publishing this unanonymized data leads to privacy violation. Currently there are many practices and policies indicating the types of data that can be published. This practice is not enough to preserve privacy and can result in an unwanted disclosure. *Privacy Preserving Data Publishing (PPDP)* provides ways for publishing useful information and preserving data privacy. One of the biggest challenges in this area is to reduce the possibility of inference attacks using which the value of hidden attributes could be compromised. Perturbation changes the structure of the data and in the bargain the data has low utility. Maintaining a proper trade-off between privacy and utility before releasing the data to the third party is an important research issue. Therefore, a lot can still be explored in this area contributing to the field of privacy preserving social network data publishing.

2.6 Summary

In this chapter we carried out an elaborative study on the privacy landscape where we compared and contrasted the existing privacy literature and discussed different perceptions of privacy amongst users. Intentionally or unintentionally users share a lot of personal information on these networks which results in privacy threats and unwanted privacy breaches. We carried out an extensive literature survey covering the areas of *measuring users' OSN data privacy*, *enabling selective sharing of sensitive OSN data* and *protecting sensitive OSN data from inference attacks*. We analyzed the present state of art for privacy and identified some of the important research gaps in the field. A survey of different data privacy models, aspects and techniques in field of data privacy is published in [Pub1].

Chapter 3

Measuring Users' OSN Data Privacy

"Even though society as a whole is increasing the amount of personal information available to the public, there is still an expectation of privacy. People believe, sometimes falsely, that they can control the personal information they hold out to the public by determining who can access the information and how the information will be used. It is extremely challenging to define a fluid concept like privacy because it touches almost every aspect of a person and society to one degree or another." - Daniel J. Solove

3.1 Introduction

Sharing information in an OSN is a voluntary action on the part of users [82]. It brings along various privacy concerns because most of the data that is shared online is potentially sensitive in nature. To ensure that the user does not become a victim of privacy breach it is important to prevent information from escaping its privacy boundaries. Some of the recent cases that have been reported imply that the current mechanisms are not providing adequate level of data privacy protection due to which privacy in OSN has become a matter of deep concern [39]. It is essential for an OSN user to understand the privacy risks that could follow after carelessly sharing the sensitive data online. In order to make the users aware of the data privacy concerns their information sharing behaviour along with the sharing behavior of their connections should be measured. Measuring this abstract and unobservable trait is a challenging task and is possible only if there is a proper metric of privacy.

Psychometrics is derived from the two words *psycho* and *metrics* which basically means *'mental measurement'* [83]. It is the subfield of psychology and is the science of measuring individuals' unobservable characteristics. It involves the development and analysis of measurement instruments and theoretical approaches. Human beings either make a relative or an absolute judgment while comparing things. Comparing the height of an object with the other object has a direct solution but it is far difficult to compare whether an individual is more proficient than the other [84]. Therefore a measurement tool is needed that could locate and compare individuals' responses or the properties of the object using the metrical system.

Using measurement the researchers could provide a reasonable and consistent way to summarize the responses of people using instruments such as *attitude, achievement tests, questionnaires and surveys*. Instruments are used to relate and map something observed in the real world to something measured as a part of theory [84]. Psychometry gives the'ability' to measure the psychological attributes related to humans. It provides an instrument using which a manifest variable could be mapped to a latent variable [85]. Privacy is unobservable, hidden and an inherent trait. Comparing inherent quality like privacy of a person is a challenging task and requires a measurement tool using which the latent trait of different users could be compared.

OSNs allow their users to control and manage privacy settings on their profile [86]. Often configuring these settings for every item is confusing and a time consuming task. The other issue that is being addressed in the chapter is the development of a context based personalized recommender system which could help the users configure their privacy settings after comparing their sharing behavior with the sharing behavior of other similar users in the network. In this chapter we use the concepts of psychometrics and its various mathematical models to measure and quantify privacy of OSN users. In the later part of the chapter we would also discuss the proposed *privacy settings recommender system* in detail.

The major contributions of the chapter are as follows:

- We studied the issue of measuring privacy of OSN users' profile such that a proper comparison of sharing behaviours between the users and their connections could be drawn and an effort to prevent privacy loss could be made.
- We used psychometric models like the classical test theory and the item response theory to measure privacy of OSN users.
- We proposed and implemented a privacy settings recommender system that could help the users diligently share their data online.

3.2 The Classical Test Theory

For years the classical test theory had formed the formulation for measurement [87]. All the classical test theories follow the concept of total test score which is calculated from the expected value of their raw score of multiple items. Equation 3.2.1 illustrates that if *X* is the raw score obtained by an individual then it is made up of the true component *(T)* and the random error *(E)* component. This random error component has a normal distribution with null expected value and a constant variance [88].

$$X = T + E \tag{3.2.1}$$

Usually a test is made up of a number of items and is administered for a sample of respondents. The responses collected can be dichotomous i.e. can have a value of 1 or 0 or can be polytomous and can have a range of values. To examine important items the descriptive statistics like the mean and variance are used. They provide the information about the usefulness of an item. An item that has a low variance implies less variability and may not be useful for measurement [89]. So generally an item having a high variance and a mean at the center point of the distribution performs better and is chosen. The mean of the dichotomous item is equal to the proportion of individuals who had passed /endorsed an item. Variance of a dichotomous item is calculated by taking the product of p and q. Here p is the proportion of individuals who had passed or endorsed an item and q is the proportion of individuals who failed or did not endorse an item.

In the next section we will discuss the complete procedure to measure data privacy using the classical test theory.

3.3 Calculation of the Privacy Quotient using the Classical Test Theory Approach

3.3.1 Detailed Analysis of the Survey

We used the Classical Test Theory (CTT) model to measure data privacy of OSN users. We carried out an extensive survey which was mainly based on collecting information about the OSN users' in the Indian context. An active participation of people mainly in the age group of 10 - 55 was noted. All the participants had a good understanding of computers and were familiar with the use of social media. A total of 58% of males and 42% of females participated in it. Most of the participants were avid users of OSNs. The questionnaire contained a total of 30 questions. An initial set of questions were set up to understand the general perception of privacy amongst the users. These were mainly related to the questions of privacy in general. The second set of questions were specific to the use of social networks with an intent to capture the sharing behavior of users and measure their online privacy. The third set of questions were set up to collect the general demographics of users. The study was performed on dichotomous items. A total of 600 respondents participated in the survey out of which 118 respondents did not complete the survey. We mainly analyzed the remaining 482 respondents for our study.

OSNs play an active role in building and maintaining relationships. Unintentionally or intentionally people often share personal details like their *phone number, address, Email, date of birth, relationship status, political and religious views* etc. on social networking sites [90]. Around 21.7% of people responded yes to the question where they were asked whether they intentionally share their sensitive information to improve their digital presence in the online social networking world. Posting pictures, videos and tagging friends is a common practice in an OSN. Users become highly uncomfortable when they find that their data is being shared without their consent. Analysis shows that 68.33% of people do not like if some content related to them is made public without their knowledge, out of which 81.18% were females and 18.82% were males.

Social networking sites track and target the users' interests, likes and activities to suggest them the appropriate advertisements [91]. Results show that 81.66% of users are concerned about how the social media is making use of their information whereas 18.33% of the users are not concerned about knowing it. Preserving users' privacy is an important task for the service providers as any loss of users' privacy could earn them a bad name and loss of revenue. According to the results around 83.33% of the people would stop the use of social networking sites if their personal sensitive information were used inappropriately. Changing privacy settings could solve the problem of privacy up to a greater extent but is often a confusing, complicated and time consuming task. A total of 63.33% of the people agreed to it out of which 59.09% were females and 40.91% were males. Surprisingly around 54% of the people skip the step in which they are asked to change their default privacy settings whereas around 24% of users were still using the default privacy settings.

3.3.2 Organizing the Data in Response Matrix

Second set of questions in the survey were designed to measure users' data privacy in an OSN. Questions like *Have you shared your date of birth, contact number, Email in an OSN* etc. were asked. For *N* number of users and *n* number of profile items, a response matrix of the order *N X n* is formed. Here, the range of items is signified by *i* where *i* varies from $1 \leq i \leq n$ and the range of users is signified by *j* where *j* varies from $1 \leq j \leq N$. Figure 3.1 shows the response matrix of *N* number of users and *n* number of profile items. Each row represents the users and each column the profile item to be considered. We dealt with the dichotomous response matrix which can either take the value of 0 or 1. If a user *j* has shared the information about the profile item *i*, then the value of *R(i,j) = 1*. If a user *j* has not shared the information about the profile item *i*, then the value of *R(i,j) = 0*. Table 3.1

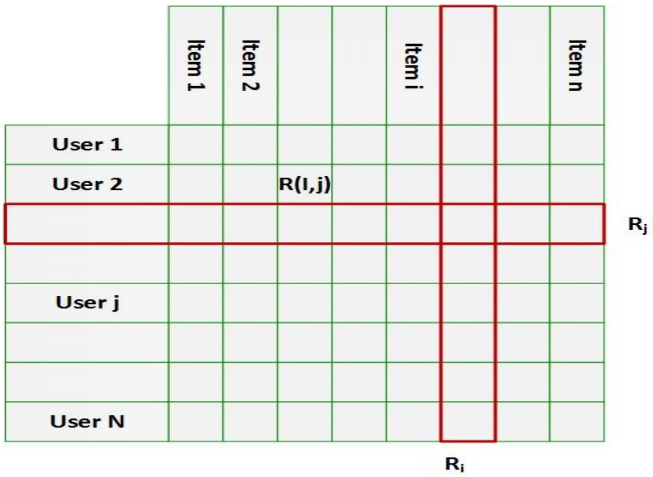


Figure 3.1: A N X n response matrix

lists 11 profile items that we considered for measuring users' sharing behavior and their privacy.

Table 3.1: List of profile items considered for measuring privacy in an OSN

SNo	Profile Items
1	Contact Number
2	E mail
3	Address
4	Birthdate
5	Hometown
6	Current Town
7	Job Details
8	Relationship status
9	Interests
10	Religious Views
11	Political Views

3.3.3 Calculation of Sensitivity

Sensitivity (β) is the risk in sharing a profile item. As sensitivity increases, privacy risks involved in sharing the item also increases. Hiding a more sensitive item makes the user more private than hiding a less sensitive item. The difficulty level *p* in the CTT is denoted by the number of individuals who endorse a particular item. An item that has a high *p* value is considered as an easy item and the items that have a low *p* value is the difficult item. The difficulty level *p* is calculated by $\frac{R(i)}{N}$ and sensitivity is calculated using equation 3.3.1.

$$\beta = 1 - p \tag{3.3.1}$$

$$\beta = \frac{N - R_i}{N} \tag{3.3.2}$$

where R_i = $\sum_j$ R(i,j) and R_j = $\sum_i$ R(i,j)

3.3.4 Calculation of Visibility

Visibility is the property of an information that captures the popularity of an item in the network. The wider is the spread of information the more visible it is. $V_{(i,j)}$ i.e. the visibility of a profile item *i* by the user *j* is calculated using the following equations :

$V_{(i,j)}$ = $Pr_{(R(i,j)=1)} \times 1$ + $Pr_{(R(i,j)=0)} \times 0$;
$V_{(i,j)}$ = $Pr_{(R(i,j)=1)}$ + 0;
$V_{(i,j)}$ = $Pr_{(R(i,j)=1)}$;
where $Pr_{(R(i,j)=1)}$ is the probability that the user *j* shares the profile item *i*.
and $Pr_{(R(i,j)=0)}$ is the probability that the user *j* does not share the profile item *i*.
Visibility $V_{(i,j)}$ can be calculated using equation 3.3.3.

$$V_{(i,j)} = \frac{R_i}{N} \times \frac{R_j}{n} \tag{3.3.3}$$

Here, R_i are the total number of users who have shared an item *i* and R_j are the total number of items shared by the user *j*.

3.3.5 Calculation of Privacy Quotient

Privacy quotient (PQ) is the score given to the user by analyzing their sharing behaviors. It is the potential risk that is caused by the users' participation in the network. If β_i is the sensitivity of the profile item *i* and $V_{(i,j)}$ is the visibility of the profile item *i* for a user *j* then the privacy quotient $PQ_{(i,j)}$, for a profile item *i* for user *j* can be calculated as specified using equation 3.3.4

$$PQ_{(i,j)} = \beta_i * V_{(i,j)} \tag{3.3.4}$$

Equation 3.3.5 calculates the overall privacy quotient of the user *j* for all profile items *i*

$$PQ_{(j)} = \sum_{i}^{n} PQ_{(i,j)} = \sum_{i}^{n} \beta_i * V_{(i,j)} \tag{3.3.5}$$

where the range of items i.e. *i* varies from 1 $\leq$ i $\leq$ n.

3.3.6 Result and Analysis for Privacy Quotient Calculation using the Classical Test Theory

On the basis of the data collected from 482 respondents we have calculated the sensitivity of 11 profile items using equation 3.3.1. In Figure 3.2 we can clearly see that *address* is the most sensitive attribute followed by *political views, contact number, religious views* and *relationship status*. Whereas, *birthdate, current town* and *hometown* details are shared by most of the users and are comparatively less sensitive. Figure 3.3 shows a bar graph representing the number of users in the specific range of privacy quotient. We normalize the privacy quotient such that the minimum value is 0 and the maximum value is 1. We see that most of the respondents have a privacy quotient in the range of 0.7 - 0.8. This indicates that these users have greater privacy risks because of their excessive sharing behaviours. For measuring privacy we used CTT model which is one of the widely used psychometric models. These models are also referred as the *weak models* because the assumption of these models can easily meet the test data. These models are strictly linear and the item difficulty and the discrimination are group and sample dependent. Hence, there is a need to apply a better psychometric model to measure the users' privacy.

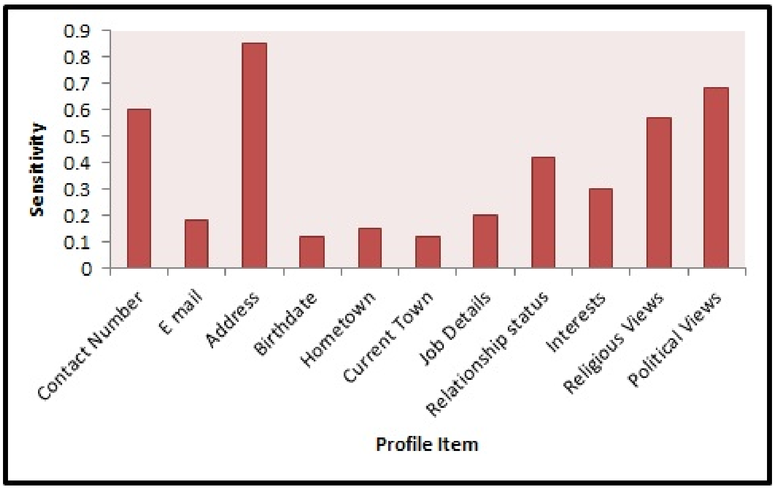


Figure 3.2: Sensitivity of the various profile items

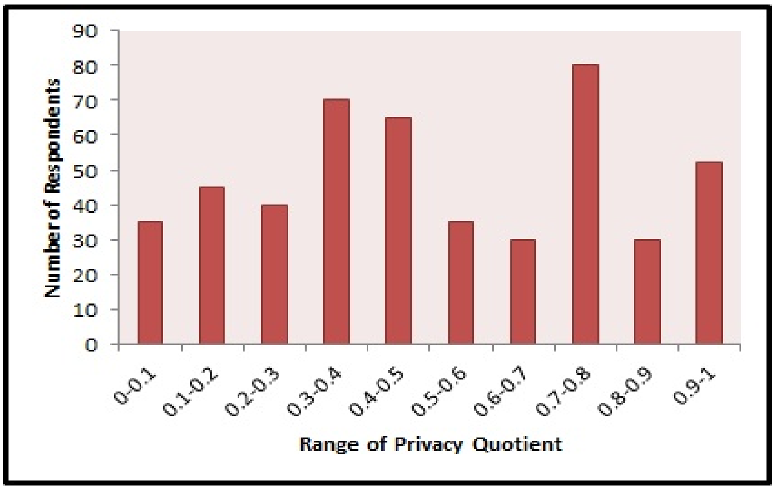


Figure 3.3: Number of users and the range of privacy quotient

In contrast to CTT the *Item Response Theory (IRT)* models are considered as *strong models* because their assumptions are stringent. The IRT model is a nonlinear model, is greatly flexible and permits the calculation of one's probability of answering a question or sharing an item correctly. It fits the observed data well and computes intuitive values of *ability, sensitivity and visibility* [92]. In the next section we will be discussing about the IRT model in detail.

3.4 Item Response Theory

In order to measure the latent trait it is important to have a measurement scale but defining the scale of measurement and intervals on the scale is a difficult task to undertake. In order to measure the ability in psychometrics generally a test is developed which contains a list of items. Each item contributes in measuring some aspect of ability of interest. Mainly, this concept is used in competitive exams where the respondents are free to write any response that would seem appropriate to them [93]. The items are dichotomously scored where for every correct response the examinee gets a score of 1 and for every incorrect answer they get a score of 0.

In CTT the scores are decided by summing the individual scores for each item. In IRT the emphasis is on the individual items rather than the aggregate of the item responses. Each respondent possesses a certain amount of the underlying ability. This ability score is denoted by the letter θ. At each ability level there is a probability that an examinee would answer the question correctly. This probability is denoted by $P(\theta)$. Respondents having a low value of θ have a low value of $P(\theta)$ and vice versa. If an attempt is made to plot the value of θ against P(θ), the result is a smooth S shaped curve. This curve is known as the *Item characteristic Curve (ICC)*. Figure 3.4 shows an item characteristic curve with θ (ability) and β (difficulty) on the x axis and probability i.e. PROB on the y axis. At the low level of ability the probability of giving the correct response is nearly 0. At the highest level of ability this probability value $P(\theta)$ reaches 1. Every ICC has two technical parameters i.e. the difficulty and the discrimination which are described as follows:

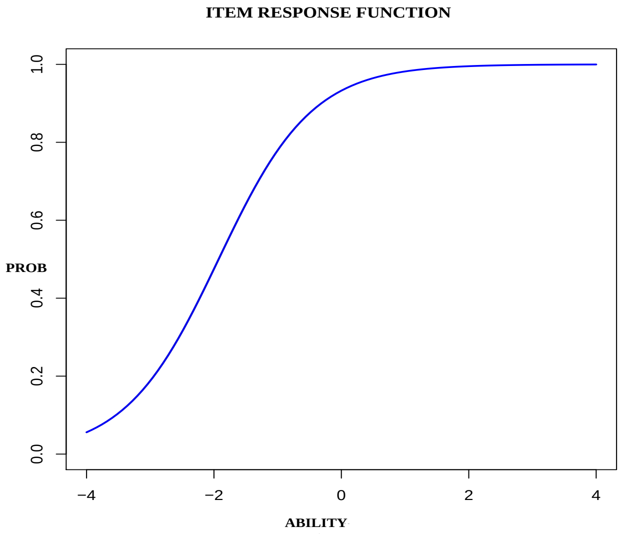


Figure 3.4: The Item Characteristic Curve

- Difficulty: The difficulty essentially describes where exactly the item functions along the ability scale. The difficult item functions along the high ability examinees and the easy item functions along the low ability values.

- Discrimination: The discrimination describes how well an item can differentiate between individuals or examinees having ability above and below the item location. The steepness of the slope is directly proportional to the discrimination value of the curve. The steeper the curve the better it can discriminate between the items. A flat curve would have 0 discrimination factor.

The three mathematical models defined for the ICC are the one parameter logistic model (Constrained and Unconstrained), two parameter logistic model and the three parameter logistic model. The standard mathematical model for the ICC is the cumulative form of the logistic function.

$$P(\theta) = \frac{1}{1 + e^{-L}} \tag{3.4.1}$$

Here L = $\alpha(\theta$ - $\beta)$ which is the logistic deviate (logit).
e is a constant with the value of 2.718.
θ is the ability of the respondents.
α is the discrimination parameter of the item.
β is the difficulty value of the item.

Equation 3.4.2 shows that the probability of an item being shared is a function of the model's item parameters and ability of the individual.

$$Prob(Sharing) = f(parameters, ability) \tag{3.4.2}$$

The *three parameter logistic model* is not in the scope of our work hence, we will not be going into its details. IRT can be used to find the privacy strength of the users' profile using which they can manage their privacy settings in an effective way. We will discuss the proposed framework for measuring privacy strength in an OSN users' profile in the next section.

3.5 A Framework to Measure the Privacy Strength of OSN Users' Profile

Measuring users' data privacy and ranking them according to their privacy quotient greatly helps in comparing the sharing behaviors of online users. In order to enhance the privacy quotient there is a need to compare the users' privacy quotient with the other users in the network. Customizing privacy settings of the users' profile by looking at such a diversified data is not feasible as different people have different requirements and understanding of privacy. This essentially happens as data privacy greatly depends upon the context. What is private to a particular individual may not be private to others. The privacy requirements are a function of demography of the users as well as their social ties. For measurement it is important that the objects to be measured should be similar in most respect, so while measuring the data privacy of a user a comparison should be drawn with their connections with whom they bear great similarity. In the next subsection we will discuss the steps for the proposed framework that would help the users' in knowing their OSN profile's privacy strength using which they can improve their overall sharing behavior in the network.

3.5.1 Outline of the Steps used in the Framework

Figure 3.5 shows the complete framework to calculate privacy strength of the users' profile. The steps involved in the calculation of privacy strength are as follows:

- Input to the framework in the form of response matrix.
- Calculate the model parameters for the profile items using IRT models.
- Select the best model that will fit the data.
- Calculate the privacy quotient.
- Decide the range of privacy quotients.

- Set up privacy strength labels for each range.
- Output the privacy strength of the user's profile.

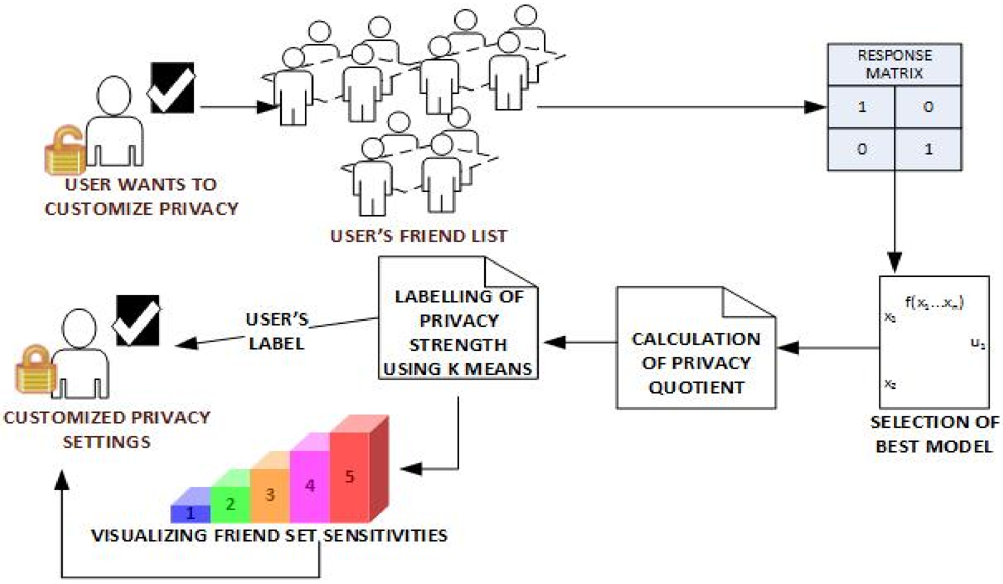

USER WANTS TO CUSTOMIZE PRIVACY
USER'S FRIEND LIST
RESPONSE MATRIX
1
0
0
1
$f(x_1...x_n)$
x_1
u_1
x_2
SELECTION OF BEST MODEL
CALCULATION OF PRIVACY QUOTIENT
LABELLING OF PRIVACY STRENGTH USING K MEANS
USER'S LABEL
CUSTOMIZED PRIVACY SETTINGS
1
2
3
4
5
VISUALIZING FRIEND SET SENSITIVITIES

Figure 3.5: The proposed framework for measuring privacy strength of the users' profile

3.5.2 Input to the Framework

The framework was tested on the data sets obtained from *Facebook* which is one of the popular OSNs. In order to collect the data from Facebook we mainly used the *Facebook Query Language (FQL)* and *Graph API*. Figure 3.6 shows the data collection diagram where an application makes a request to the server by making use of FQL. A unique access token is generated which is verified at the server and the result is sent back to the application. The framework extracts the data of the users' connection. If we take n dichotomous profile items for N users then we can generate a $N\times n$ dichotomous response matrix. The range of i and j are $1 \leq i \leq n$ and $1 \leq j \leq N$ respectively where i represents the number of items and j represents the number of users. If an item i is being shared by an individual j then the value of *jth* row and *ith* column is marked as 1 otherwise is marked as 0. The response

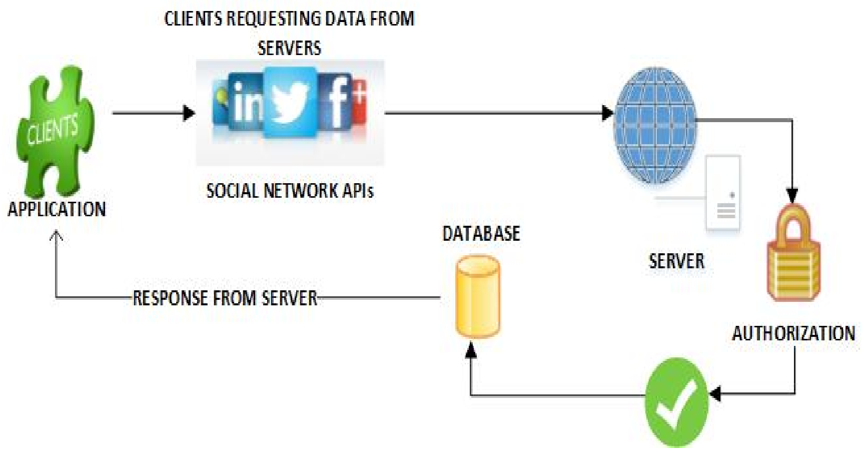


Figure 3.6: The data collection module

matrix is similar to the matrix representation in Figure 3.1. The cronbach's alpha for the profile items was found to be 0.7228, which is ≥ 0.7 and signifies a greater reliability of the profile items. We extracted the data from Facebook profiles of 150 users. Each user on an average had 245 connections. Altogether we analyzed the data for 36,845 users. For each user a response matrix was built and the model parameters were calculated.

3.5.3 Calculation of the Model Parameters using the One Parameter Logistic Model

The *one parameter logistic model* is the simplest of all the IRT models. They can be categorized as constrained and unconstrained models. For the *unconstrained one parameter logistic model* the value of α_i is a constant which is same for every item. Having a constant α means that all the items are equally discriminating. The ICC depends on β_i that denotes the sensitivity of an item *i*. The *constrained one parameter logistic model* is same as the *unconstrained one parameter logistic model* but the only difference is that the value of α is set to 1 for each profile item.

We plot an ICC between the ability of the respondents and their probability of sharing an item. Figure 3.7 represent the *item characteristic curve* for the *unconstrained and constrained one parameter logistic model*.

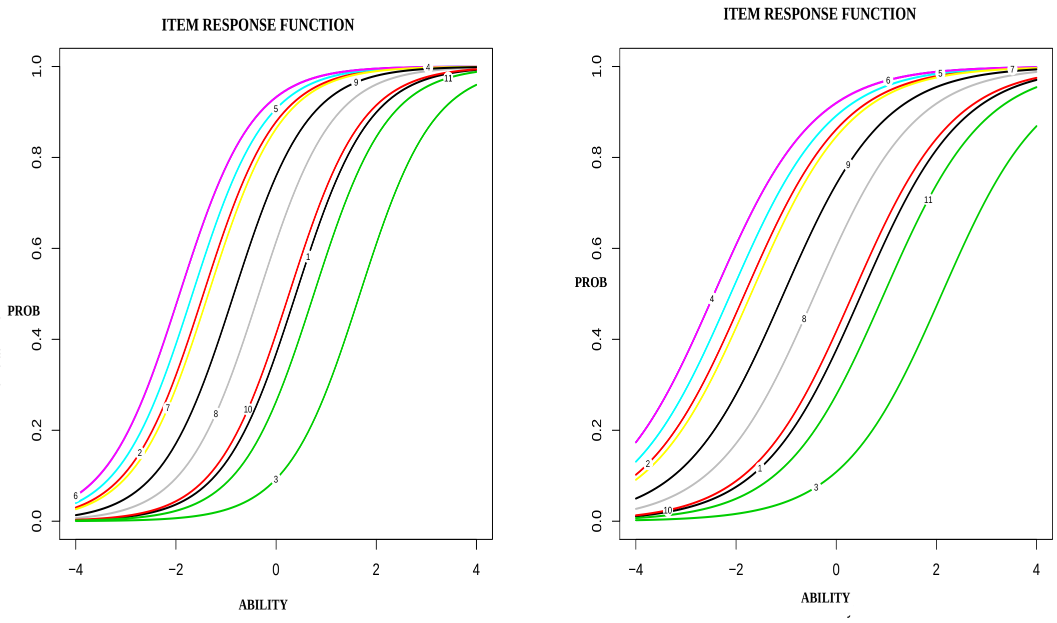


Figure 3.7: Item characteristic curve for the unconstrained and the constrained one parameter logistic model

At any given instance for a person with a specific ability, the probability of sharing *date of birth* (curve number 4) is much higher than the probability of sharing the *address* (curve no 3) which essentially means that the *address* is more sensitive than the *date of birth*.

Table 3.2 and 3.3 gives us the calculated average values of model parameters for all the 11 profile items using the *unconstrained and constrained one parameter logistic model* respectively. We see that the *address* is considered as the highly sensitive profile item followed by *political views* and *contact number*.

Table 3.2: Discrimination and Difficulty using one parameter unconstrained model

SNo	Profile Item	Discrimination	Sensitivity
1	Contact Number	1.363163	2.323667e+00
2	Email	1.363163	4760522e-01
3	Address	1.363163	3.601252e+00
4	Birthdate	1.363163	3.453648e-06
5	Home Town	1.363163	2.573421e-01
6	Current Town	1.363163	0.000000e+00
7	Job Details	1.363163	5.753196e-01
8	Relationship Status	1.363163	1.585066e+00
9	Interests	1.363163	1.085306e+00
10	Religious Views	1.363163	2.187068e+00
11	Political Views	1.363163	2.683194e+00

Table 3.3: Discrimination and Difficulty using one parameter constrained model

SNo	Profile Item	Discrimination	Sensitivity
1	Contact Number	1	2.9446946645
2	Email	1	0.6117859481
3	Address	1	4.5480914380
4	Birthdate	1	0.0000000000
5	Home Town	1	0.3316536983
6	Current Town	1	0.0003948577
7	Job Details	1	0.7386141987
8	Relationship Status	1	2.0182711543
9	Interests	1	1.3872733265
10	Religious Views	1	2.7736557816
11	Political Views	1	3.3946345724

3.5.4 Calculation of Parameters using the Two Parameter Logistic Model

The two parameter model calculates the probability of the user for sharing an item based on the values of sensitivity (β) as well as discrimination (α). The probability of a j^{th} individual who is having an ability of θ_j for sharing an i^{th} profile item is given by equation 3.5.1.

$$PR_{(\theta_{ij}=1)} = \frac{1}{1 + e^{\alpha_i(\theta_j - \beta_i)}} \tag{3.5.1}$$

where θ_j is the ability of the j^{th} user, β_i is the sensitivity of the i^{th} profile item, α_i is the discrimination constant of the i^{th} profile item. In Figure 3.8 we see that the *political view* (curve no 11) has got the highest discrimination value and *address* (curve no 3) has got the lowest of all. Table 3.4 gives the calculated average values of model parameters of all the

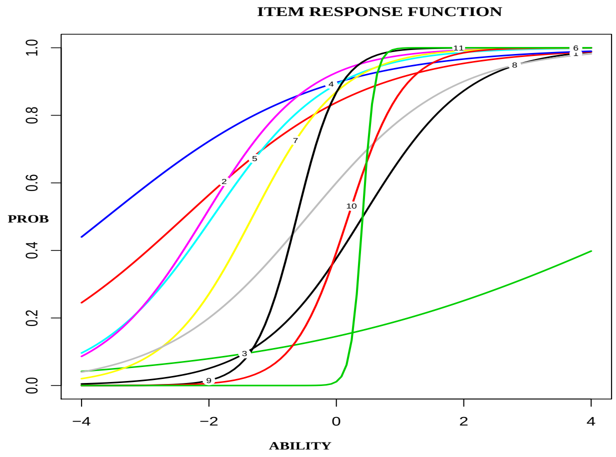


Figure 3.8: Item characteristic curve for two parameter logistic model.

11 profile items for 150 Facebook user profiles using the two parameter logistic model.

Table 3.4: Discrimination and Difficulty using the two parameter logistic model

SNo	Profile Item	Discrimination	Difficulty
1	Contact Number	0.8744772	4.012698
2	Email	0.3523724	1.223004
3	Address	0.0000000	8.815682
4	Birthdate	0.2619243	0.000000
5	Home Town	0.7471443	1.658689
6	Current Town	0.8839744	1.527580
7	Job Details	1.1031453	2.283478
8	Relationship Status	0.5572550	3.144633
9	Interests	2.6932438	2.977822
10	Religious Views	1.9696065	3.783610
11	Political Views	10.5183067	4.009395

3.5.5 Selecting the Best Model

While modeling the data set to calculate the results we may not get the exact expected value. There is always a difference between the expected and the observed result. This happens because the selected model does not fit the data completely and gives erroneous results. The model which would incur the least information loss is the best model out of all. We would be selecting the model which would minimize the loss of information. We will be using the concept of *Akaike Information Criterion (AIC)* and *Bayesian Information Criterion (BIC)* for model selection. AIC and BIC can be calculated using Equation 3.5.2 and 3.5.3 respectively.

$$AIC = -2(log - likelihood) + 2K \tag{3.5.2}$$

$$BIC = -2(log - likelihood) + Kln(N) \tag{3.5.3}$$

where *K* is the number of parameters used in the model and *N* is the total number of instances. The model having the least value of *AIC* and *BIC* is preferred over all the models. In Table 3.5 we have compared the *constrained one parameter logistic model, unconstrained one parameter logistic model* and the *two parameter logistic model.* We observed that the *two parameter logistic model* gives the lowest *AIC* and *BIC* values [94] [95]. Hence, we select the two parameter logistic model for calculating the privacy quotient of the OSN users.

Table 3.5: AIC and BIC values for different models

Model	AIC	BIC	log-likelihood
Constrained 1PL	53498.80	53570.49	-26738.40
Unconstrained 1PL	53290.42	53368.63	-26633.21
2PL	51356.44	51499.81	-25656.22

3.5.6 Mapping of Privacy Quotient with the Privacy Strength

After selecting the best model the privacy quotient (PQ) of OSN users could be calculated. Using the PQ, users could measure their privacy and rank themselves in the privacy measuring scale. The PQ is calculated using Equation 3.5.4

$$PQ(j) = \sum_{i} \beta_i * P_{ij}(\theta_j) \tag{3.5.4}$$

where β_i is the sensitivity and $P_{ij}(\theta_j)$ is the probability of sharing an item *i* by an individual *j* with an ability of θ_j. Table 3.6 shows the mapping of PQ with the respective privacy strength. The PQ of all the users were normalized and then the *k means clustering* algorithm was run for 1000 iterations to determine five clusters. A separate label is given to represent each range of privacy quotient which we refer as the *privacy strength* of the profile. We categorize the labels as *High, Good, Average, Below Average and Poor.*

Table 3.6: Mapping of privacy quotient with privacy strength

Range of PQ	Privacy strength
0 ≥ pq ≥ 0.373010	High
.373010 ≥ pq ≥ .544307	Good
.544307 ≥ pq ≥ .713378	Average
.713378 ≥ pq ≥ .913382	Below Average
.913382 ≥ pq ≥ 1	Poor

Figure 3.9 represents the general statistics of the users' profile privacy strength.

We see that a majority of users' profile privacy strength was either average or below average and very few profiles had a high privacy strength. Using the framework the users will know the privacy strength of their online profile and the privacy strength of their online connections. The framework outputs the list of users whose profiles have better privacy strengths than the user and also displays their sharing patterns. It allows the user to view the sensitivities of various profile items so that they can manage and enhance their privacy quotient and improve the privacy strengths of their OSN profile. Understanding data privacy

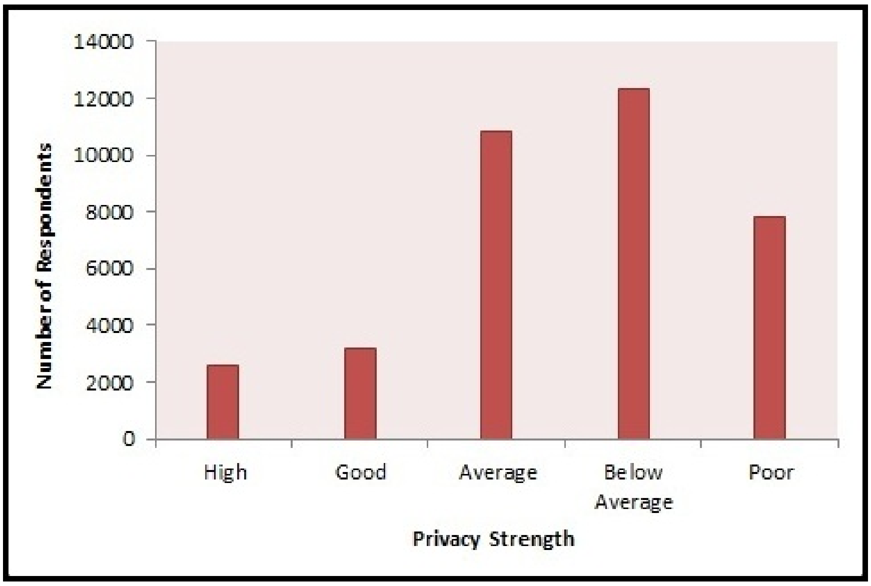


Figure 3.9: Number of respondents and their privacy strengths

evolves over a period of time and so does the PQ of OSN users [96]. Privacy is an inherent attribute which is related to many factors like the users' age, relationship status, demographics etc. According to Palen et al. [34], the problem of disclosure boundary drives the interpersonal privacy management in a user's life. The disclosure boundary emphasizes on the fact that participation in any social network requires the user to go for selective disclosure of personal information.

OSNs have static objects like the *Email, address, date of birth* etc. which do not change too often. They also have dynamic contents like *photos, videos, notes* etc. which are uploaded frequently and have a visibility range. Knowing the privacy strength the users can modify the privacy settings for the static objects in their profiles as discussed earlier but changing the privacy settings for the dynamic object is a cumbersome task. OSNs permit users to use privacy controls for the information that they want to share but managing and configuring privacy settings for every asset is tedious. Hence, there is a need for a system which could not only compare the users' privacy with their connections but could also recommend appropriate privacy settings for dynamic objects to them. In the next section we will discuss about the proposed privacy settings recommender system in OSN and give an overview of the steps used to build it.

3.6 Privacy Settings Recommender System for Online Social Networks

On a day to day basis we rely on recommendations from others either by *word of mouth* or by general surveys and reviews in newspapers etc [97]. A recommender system automates this natural social process. It uses data mining techniques and generates meaningful recommendations to the users for items or products of their interests [98]. A more mathematical explanation for the recommender system could be as follows: Let C be the set of all the users and S be the set of all possible items to be recommended. Let there be a function u which measures the usefulness of an item s for user c. For every user $c \in C$, we would be choosing items $s \in S$ that would maximize the users' utility. The recommender systems are usually classified as *content based and collaborative recommender systems* [99]. In a content based recommender system the recommendation engine recommends items similar to the ones that the users have preferred in the past. In the collaborative recommender system the recommendation engine recommends items, that people with the same preference have liked in the past. Another approach is the hybrid approach in which the recommendation engine combines the content based and the collaborative based methods to provide recommendations to the users [100].

3.6.1 Overview of the Privacy Recommender System

Figure 3.10 shows the privacy settings recommender system which recommends an optimal and meaningful privacy settings to the target user. The solution goes through various stages

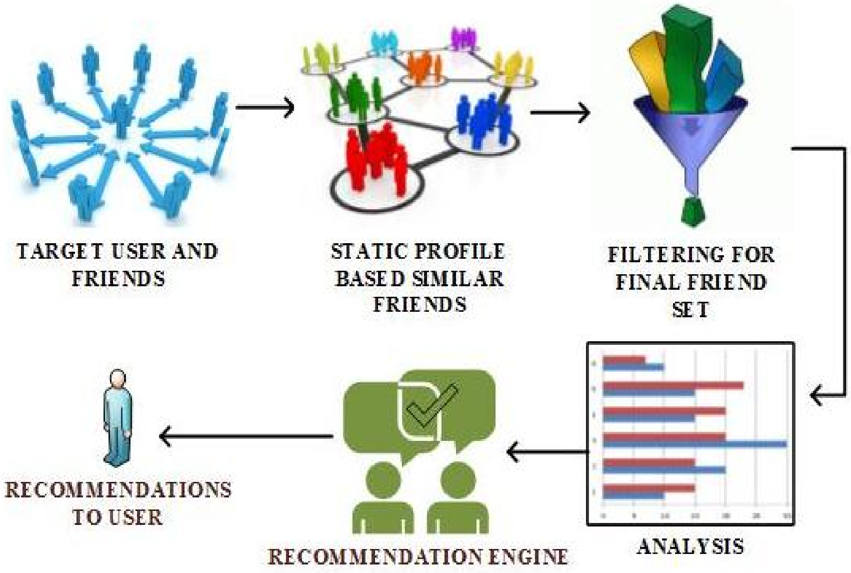


Figure 3.10: OSN privacy settings recommender system

and the output of one stage becomes the input for the next stage. Various stages of the privacy settings recommender system are as follows:

- Data collection of target users and their friends.
- Formation of a user object matrix for dynamic contents.
- Static profile based similarity and filtering.
- Forming a refined Friend Set.
- A context based personalized privacy settings recommendation for the target user.

Data Collection

We selected *Facebook* which is one of the popular OSNs and is widely used. A plethora of information about the user is available in it and the information on *Facebook* is personally identified. We used the *Facebook Query Language (FQL)* and *Graph API* for data collection. We collected the data for static as well as dynamic contents in the OSN. The input to the system are the static profile items of the target user and its connections. There are various static and dynamic contents in an OSN profile but our scope is mainly limited to the profile items listed in Table 3.7 which shows the static profile items that were extracted from the users' profile. Here age is a derived attribute from the birthdate. *Photos, videos, link* and *notes* are the list of the dynamic contents in an OSN profile. These dynamic contents can have different visibility levels in *Facebook*. Visibility level signifies the number of hops through which an information could travel from the target user to the network. Table 3.8 shows various visibility levels that a dynamic object can have in *Facebook*.

Table 3.7: List of static profile items

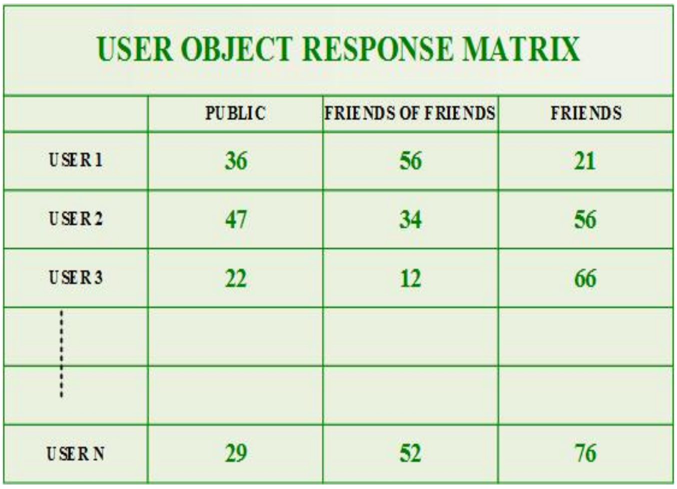
SNo	Profile Items
1	Gender
2	Age
3	Relationship Status
4	Home town
5	Home state
6	Home country
7	Current town
8	Current state
9	Current country
10	Education
11	Work

Table 3.8: Visibility level for dynamic profile items in Facebook

Profile Items	Explanation
Public	Share open to anyone
Friends	Share to friends only
Friends of Friends	Share to friends and their friends

Formation of a User Object Response Matrix for Dynamic Profile Items

In this stage we will extract the privacy settings for all the dynamic objects such as the *photos, videos, links, notes* etc. With this data we construct a $N \times 3$ user object response matrix. A sample user object response matrix is shown in Figure 3.11. Each object has its own user object response matrix. Here each cell can be defined as *R(j,i) = k* where the range of *j* is $1 \leq j \leq N$ and *N* are the total number of users. Here *i* can take any value out of all the visibility levels *public, friends of friends, friends* and *k* are the number of objects set to visibility *i* by the user *j*. In order to measure the users' privacy quotient with people who are aware of disclosure boundary we filter out all the users who have not shared any dynamic object on their OSN profiles. We also filter out the users who have shared all the dynamic objects publicly. If the user's connections have not shared anything or have shared everything publicly then comparing user's privacy with these connections would not benefit in any way. Users who have not shared anything could be the ones who are less active on OSNs or lack knowledge of the sharing feature. Hence, we filter out these sets of users and send the remaining users in the user object matrix to the next stage.



USER OBJECT RESPONSE MATRIX

	PUBLIC	FRIENDS OF FRIENDS	FRIENDS
USER 1	36	56	21
USER 2	47	34	56
USER 3	22	12	66
⋮			
⋮			
USER N	29	52	76

Figure 3.11: A user object response matrix

Static Profile Item Based Similarity and First Level Filtering

An interaction between similar people occurs at a higher rate than the dissimilar ones. This principle is referred as *homophily* [101]. We make use of the concept of homophily to find out the users who are similar to the target user. Each individual has different characteristics like *gender, age, relationship status, home town, current town, education, work* etc. Any social entity localized to a social space will adhere to its norms and dynamics [102] [98]. The lesser the distance between the characteristics of two individuals the more similar they would be. We first calculate the measure of similarity between the target user and all its connections. Table 3.7 shows the parameters that we have considered to calculate the similarity between the users. Figure 3.12 shows that we calculate the similarity between the target user and all its connections using the cosine similarity metric and filter out the less similar users from the user object response matrix. The two data sets collected can be viewed as two vectors of non negative values. If v_1 and v_2 are the two vectors, we find out the cosine angle between the two vectors using equation 3.6.1.

$$SIM_C(v_1, v_2) = \frac{\overrightarrow{v_1}.\overrightarrow{v_2}}{|\overrightarrow{v_1}|.|\overrightarrow{v_2}|} \tag{3.6.1}$$

The less similar users are decided on the basis of the value which could be altered using the framework. We divided the entire set of similarity values into three clusters using the *k means clustering algorithm* and labeled the users as *Highly Similar, Average Similar and Less Similar*. Each cluster represents the similarity degree. Using the framework the similarity degree could be chosen as *Highly Similar or Average Similar*. The framework does not allow to select the users labeled as *Less Similar* as they are the ones who bear the least similarity with the target user. If the similarity degree chosen is *Highly Similar* then the application would select the users labeled as *Highly Similar* and would filter out the ones labeled as *Average Similar* and *Less Similar*. Similarly, if the similarity degree chosen is *Average Similar* then the application would select the users labeled as *Average Similar* and would filter out the ones labeled as *Less Similar*.

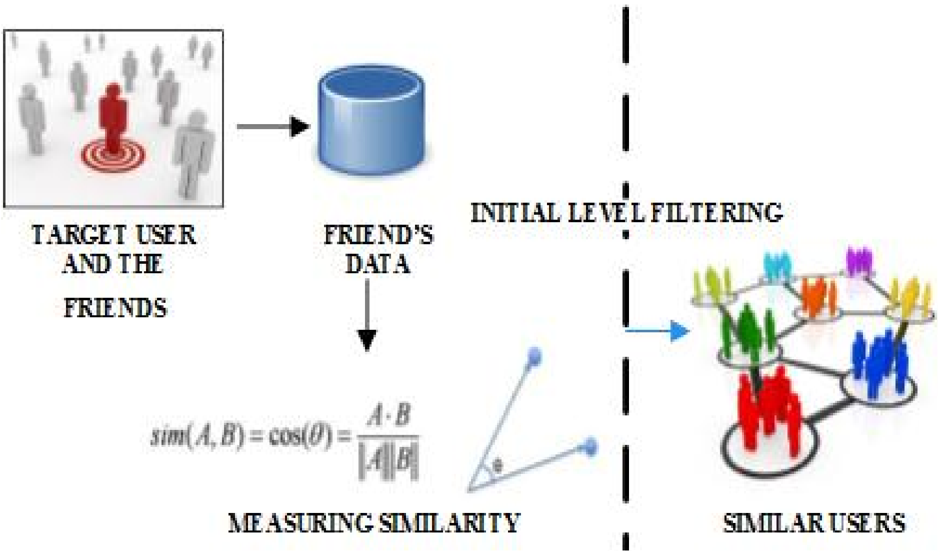

TARGET USER AND THE FRIENDS
FRIEND'S DATA
INITIAL LEVEL FILTERING
$sim(A,B) = \cos(\theta) = \frac{A \cdot B}{\|A\|\|B\|}$
θ
MEASURING SIMILARITY
SIMILAR USERS

Figure 3.12: Measuring the similarity and initial level filtering

Forming a Refined Friend Set

In order to calculate the refined *Friend Set* we calculate the private public ratio (PrPu) for all the remaining users in the *user object response matrix* using Equation 3.6.2.

$$PrPu = \frac{C_{Pr}}{C_P} \tag{3.6.2}$$

here $C_{Pr} = C_{Fr} + C_{FoF}$
where C_{Fr} is the count of objects that are visible to the user's friends. In such a privacy setting an information is visible to maximum 1 hop from the target user. C_{FoF} is the count of objects that are visible to the user's friends where an information is visible to maximum 2 hops away from the target user. C_P is the count of objects that are visible to everyone and this visibility level can make the information reach anywhere in the network.

Mostly users who have the PrPu ratio of 1 or greater than 1 are observed to have some basic understanding of privacy whereas the users having PrPu ratio less than 1 are generally observed sharing a high percentage to public hence, we do not include them in the final

Friend Set which will be used for recommendation. We calculate the percentage of objects that the users in the Friend Set have shared to *Public*, to *Friends of Friends* and to *Friends* and the same is calculated for the target user as well. Each user is given a score on the basis of the objects that they have shared and its visibility level in the network. This score is referred as PQ_D which is the privacy quotient for the dynamic objects.

The visibility levels are given a rank which is calculated using the maximum number of hops upto which an information is visible from the target node. If a user had shared an object to their immediate connections i.e. with the visibility label *Friends* then the chances of it being breached is very less. Therefore from a scale of 1 to 3, such a setting gets the highest rank. Similarly, Table 3.9 shows the other visibility levels and their respective ranks. The users in the *Friend Set* are then scored using Equation 3.6.3.

Table 3.9: Ranks of the visibility level

Visibility	Rank
Public	1
Friends of Friends	2
Friends	3

$$PQ_D = 1 * P_P + 2 * P_{FoF} + 3 * P_F \tag{3.6.3}$$

where P_P, P_{FoF}, P_F are the percentage of objects whose visibility is set to *Public, Friends of Friends and Friends* respectively. We filter out the users who have a lesser PQ_D than the target user as it implies that the target user is better in comparison with these users in terms of privacy. At this stage we will be left with all the users in the *Friend Set* who would be having a better PQ_D than that of the target user.

We calculate the average percentage of objects set by the users as *Public, Friends of Friends and Friends* by the users in the final *Friend Set*. Based on the results obtained we compare the differences in the percentage of objects being shared by the *Friend Set* and the percentage of objects shared by the target user. Figure 3.13 explains this fact clearly. Here A is the average percentage of objects set to different visibilities by the users in the *Friend Set* and B is the average percentage of objects set to different visibilities by the target user. A comparison between A and B is made and a suitable recommendation is given to the target user to improve the privacy settings of their OSN profile.

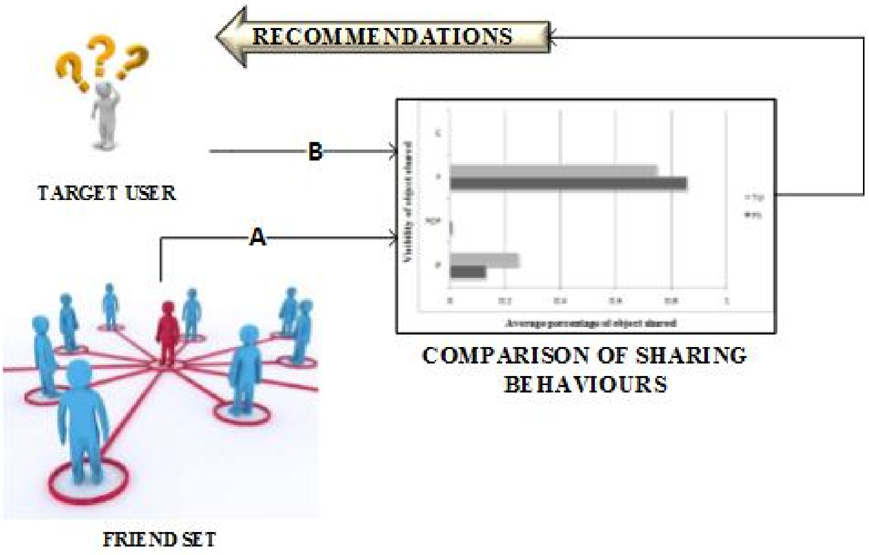

RECOMMENDATIONS
TARGET USER
B
A
Visibility of object shared
C
F
FOF
P
0
0.2
0.4
0.6
0.8
1
Average percentage of object shared
COMPARISON OF SHARING BEHAVIOURS
FRIEND SET

Figure 3.13: A context based personalized privacy settings recommendation system

3.6.2 Results of the Recommender System

We selected target users from *Facebook* and applied our solution on their list of friends. We tested the solution on 150 *Facebook* users where each user had different number of connections. Figure 3.14 shows the average percentage of objects shared by the *Friend Set* and compares it with the percentage of objects shared by the target user for different visibility levels. The *Friend Set* is a group of users who have a better privacy quotient (PQ_D) than the target user hence, an appropriate recommendation goes to the target user for modifying their privacy settings such that they can improve their scores and prevent themselves from privacy breach. For example, the first sub-figure in Figure 3.14 shows

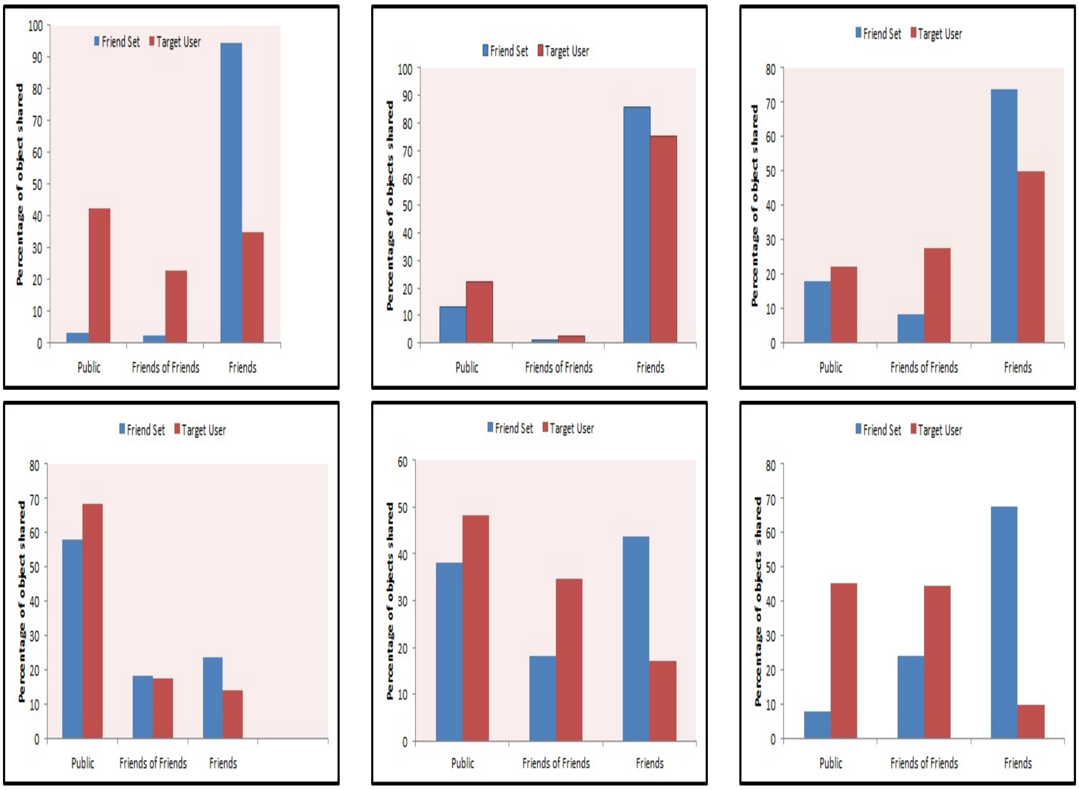


Figure 3.14: Comparison between the sharing behavior of target users and users in the Friend Set

that the average percentage of objects that the users in the *Friend Set* share to public are 3.09% and to *Friends* are 94.62% whereas the target user shares 42.345% of the objects to *public* and 35% to *Friends*. Hence, the framework picks certain objects which are shared to *public* and recommends the users to change its visibility level to *Friends*. It does so until the target user is at par with the other users in the *Friend Set*. The entire mechanism measures privacy dynamically and the results vary with the addition and deletion of any object in the users' and their friends' profile. The user has the option to either customize the profile accordingly or simply ignore the recommendation.

3.7 Summary

In this chapter we analyzed the problem of measuring privacy in OSNs and proposed an effective way to automatically configure privacy settings for the dynamic contents in the users' profiles. At first we carried out a detailed analysis of the survey to understand the privacy perception of users in general. The present state of art lacks a framework using which the users could measure their privacy and compare the privacy strength of their profiles with their connections. Hence, we measured privacy of the users' OSN profile by applying the theory of psychometrics. We made use of models like the CTT and the IRT to provide a solution to the problem of measuring privacy. Using the framework OSN users would be able to know and enhance the privacy strength of their profile. We also proposed and implemented an effective privacy settings recommender system which uses the concept of homophily to compare the users and recommends appropriate privacy settings for the dynamic objects that they share in an OSN. The use of psychometrics to measure and enhance privacy using the CTT is published in [Pub4], [Pub5] and [Pub11]. The use of privacy settings recommender system is published in [Pub6].

Chapter 4

Enabling Selective Sharing of Sensitive OSN Data

"We can think of privacy in concentric circles, with ourselves in the center. In the middle, held closest to us, are the secrets, thoughts and rituals that we keep entirely to ourselves and share with no one. Further out are the conversations we have and the actions we take that involve others but that we expect to remain private. We also expect a measure of privacy toward the outer circles, as some issues are kept inside our company without further publication" - Theresa and Ted

4.1 Introduction

The fast adoption of OSN is accompanied with problems like disclosure and unwanted access of sensitive data which results in an individual's privacy breach [39]. Privacy is a matter of deep concern and managing it in an open platform like an OSN is not an easy task. Using OSNs the users can not only share information about themselves but can also re-share the information shared by others in the network. This could be better understood by the *share* and *retweet* feature in *Facebook* and *Twitter* respectively. In this chapter we will take up the privacy issue related to sharing of sensitive data in the network and help the users selectively share their information online. We will analyze the extent upto which a node will spread the sensitive information it receives from its connections. The spread of sensitive information not only depends on the sharing behavior of the node but also upon its topology. Hence, we will measure the sharing behavior of the node with respect to its topological features in the network [53]. The collective value of a social network defines its *social capital* [103]. Preventing oneself from sharing information online will result in low social capital and a community of limited value. Therefore, while sharing information there has to be a trade-off between the social capital of the community and an individual's privacy. This according to Palen et al. [34] is defined as the problem of *disclosure boundary* which emphasizes on the fact that participation in any social network requires the user to go for *selective sharing* of personal information.

Trust is a terminology which is widely used in the field of *Sociology, Psychology, Economics, Computer Science* etc [104]. Privacy and trust are highly related to each other. Trust is defined as a measure of confidence that an entity or entities will behave in an expected way despite the lack of ability to monitor or control the environment in which it operates. It plays a role in the formation of social communities as well as helps to determine how information flows in the network [105]. In order to prevent unwanted disclosures it is important to form a trusted community where all the members can share their thoughts and opinions without the fear of privacy breach. In an offline environment the users share their sensitive information to the trustworthy connections. These trustworthy connections do not spread the users' sensitive information further in the network. However, this behavior is very difficult to map online. Many OSNs have a feature where the users can customize their friend list as close friends, family etc. All close friends may not be privacy conscious and might have a position in the network such that the chances of spreading sensitive information through them is very high. Therefore, there is a need to refine the existing online trusted community which could preserve the users' privacy upto a great extent.

The main contributions of the chapter are as follows:

- We analyze the spread of users' sensitive information to the rest of the network and prove that the nodes' sensitive information spread depends on its sharing behavior as well as its topology.
- We study the issues of forming an online trusted community and compute the *direct trust* for the target users with respect to every node in their network. We also refine the trusted communities of *strong ties* to reduce *insider attacks* and prevent the sensitive information leak in an OSN.
- We propose efficient models to enable *selective sharing* of sensitive information in an OSN to prevent users from unwanted and unintentional privacy breaches.

4.2 Preserving Privacy in OSNs using Graph Structural Analysis

Graph theory has proved to be extremely useful in fields like *networking, sociology, communication* etc [106]. Analyzing some of the basic graph properties helps in better evaluation of solutions and aids in performance improvement. The privacy of a node depends on how an information will spread through it. This can be understood by drawing the relation of the node's privacy with its topology in the network. We require a metric using which the information flow through the node can easily be determined. Hence, we take up *centrality* which is one of the most important structural properties of the network. In the next section we will discuss the proposed privacy preserving model using which the nodes having a high probability for spreading sensitive information could be identified. Once these nodes are identified, preventing sensitive information from them would highly reduce the chances of information disclosure.

4.2.1 Proposed Privacy Preserving Model

In order to preserve privacy of the nodes we propose a model that makes use of *PQ* and *centrality* values of the node. The steps in the model execution are as follows:

A: Make a $N \times n$ response matrix where *N* is the number of nodes and *n* is the count of the total profile items to be considered. If a node *j* shares a profile item *i* then it is marked as 1 or else is marked as 0.

B: Calculate the *sensitivity*, *visibility* of profile items and *privacy quotient (*PQ*)* of the nodes followed by normalization.

C: Measure the *betweenness centrality* and *closeness centrality* of each node followed by normalization.

D: Identify the nodes having a high probability for spreading sensitive information.

E: Implement algorithm for enabling selective sharing in an OSN.

4.2.2 Detailed Explanation of the Steps used in the Model

The procedure for steps A and B of the algorithm are the same as explained in Chapter 3. The sensitivity and visibility of the items in the profile as well as *PQ* of the user are calculated using the *IRT* models. In the next subsection we will discuss the procedure to calculate *centrality* of the nodes in the network.

4.2.3 Measure Centrality of Nodes in the Network

Centrality ranks the nodes according to their importance in the network. The traffic flow through the node is one of the factors responsible to determine the importance of nodes. Centrality determines the way in which information or the traffic flows in a network [107].

Where privacy is a concern, the sensitive information of the individual should not be subjected to unwanted disclosures after it is being shared to any connection in the network. Hence, the lesser the traffic flow through a node, the more privacy preserving it is. Therefore we need to measure centrality along with *PQ* of the node to find out the best combination which would incur the least information disclosure in the network. In our work we consider two of the most important centralities which are *betweenness centrality* and *closeness centrality*. For our problem *degree centrality* is less relevant in comparison to the betweenness and closeness centrality because it is the measure of the number of links a node has. A node with a higher degree having a few high influential neighbors may have much higher influence than a node having a larger number of less influential neighbors [108]. Degree centrality would give us highly biased results hence, betweenness and closeness centrality were chosen for measuring the importance of a node.

Measure Betweenness and Closeness Centrality of a Node in an OSN

According to Freeman [109] ***betweenness centrality*** of a node *v* is the number of times a node *s* passes through node *v* in order to reach the node *t* if it follows the shortest path. Given a graph *G (V,E)* where *V* and *E* are the number of vertices and edges respectively, then the betweenness centrality C$_B(v)$ is defined as

$$C_B(v) = \sum_{s \neq v \neq t \in V} \frac{\sigma_{st}(v)}{\sigma_{st}} \tag{4.2.1}$$

where σ_{st} denotes the total number of shortest paths for each pair of *s* $\in$ *V* and *t* $\in$ *V*, and $\sigma_{st}(v)$ denotes the number of those paths that pass through *v*.

Closeness centrality of a node *v* is the sum of the graph theoretic distances from all the other nodes to the node *v* and the distance is defined in terms of the shortest distance from one node to the other. Closeness centrality C$_C(v)$ is defined as

$$C_C(v) = \frac{1}{\sum_{y \in V} d_G(v, y)} \tag{4.2.2}$$

where d$_G(v, y)$ is the shortest distance between node *v* and node *y*.

4.2.4 Identify the Nodes having a High Probability for Spreading Sensitive Information

We simulated *random networks* following the principle of *preferential attachment* [110] to measure the relation between the centrality of nodes and privacy. Table 4.1 lists some of the properties like the *number of nodes, clustering coefficient, average node degree* etc. for the simulated networks. Each node was given a *PQ* following a normal distribution i.e. very few users had a good and a poor *PQ* whereas most of the users had an average value of *PQ*. Closeness centrality and betweenness centrality of the nodes followed a normal and a power law distribution respectively.

Table 4.1: Number of nodes, clustering coefficient and average node degree of the simulated networks

No. of Nodes	Clustering Coefficient	Average Node Degree
50	.344	3.85
100	.340	4.16
150	.245	3.68
200	.264	3.54
250	.194	3.185
300	.168	3.03
350	.136	2.87
400	.113	2.76
450	.091	2.64
500	.076	2.56

Each node has a *PQ, B and C* value. Here *PQ, B and C* represent the privacy quotient, betweenness centrality and closeness centrality respectively. We take three variables *(PQ, B, C)* and cluster the values of each into a range of high and low by making use of the *k-means clustering* algorithm. A total of 8 combinations i.e. $\{\{L,L,L\},\{L,L,H\},\{L,H,L\}, \{L,H,H\},\{H,L,L\},\{H,L,H\}, \{H,H,L\},\{H,H,H\}\}$ are possible where *H* and *L* stands for High and Low respectively. We choose one of the nodes as the *source* or the *originator* of the information and followed the *ICM* [111] to spread this information from the source to the other nodes in the network. In the first wave the victim node shares its information with some probability to the ones it is directly connected. This probability of sharing is proportional to *(1 - PQ)* of the source node. In the second wave some of the immediate neighbors of the source node will know the sensitive information and will share it with their neighbors with a probability proportional to their *(1 - PQ)*. This continues until all the nodes who know about the sensitive information of the source have got at least one chance to spread this sensitive information with their neighbours.

Neighbors of the source node can have one property from the set of $\{\{L,L,L\}, \{L,L,H\}, \{L,H,L\}, \{L,H,H\}, \{H,L,L\}, \{H,L,H\}, \{H,H,L\}, \{H,H,H\}\}$. In order to identify the nodes having a high probability for spreading sensitive information we ran simulations on networks having properties as specified in Table 4.1. Figure 4.1 shows the graph between the total number of nodes and the number of nodes knowing about the sensitive information of the source node. The first graph shows the sensitive information spread when the information was not shared with the nodes having the property of $\{L,L,L\}$ but was allowed to pass to other nodes using *ICM*. The second graph shows the sensitive information spread when the information was not shared to the nodes having the property of $\{L,L,H\}$ whereas allowed to pass through other nodes using *ICM*. Similarly sensitive information spread was modeled for all the other cases as well.

Observation: We observed that in all cases the maximum sensitive information flow is through nodes having the property of $\{L,H,H\}$ and $\{H,H,H\}$. This is because preventing information flow through these nodes reduced the number of sensitive information aware nodes upto a great extent. An average spread of sensitive information was observed by the nodes having $\{L,H,L\}$ and $\{L,L,H\}$ property. Minimum sensitive information spread was observed through nodes having the property of $\{H,L,L\}$, $\{L,L,L\}$, $\{H,L,H\}$ and $\{H,H,L\}$. Preventing the information flow through these nodes did not have much effect on the count of sensitive information-aware nodes.

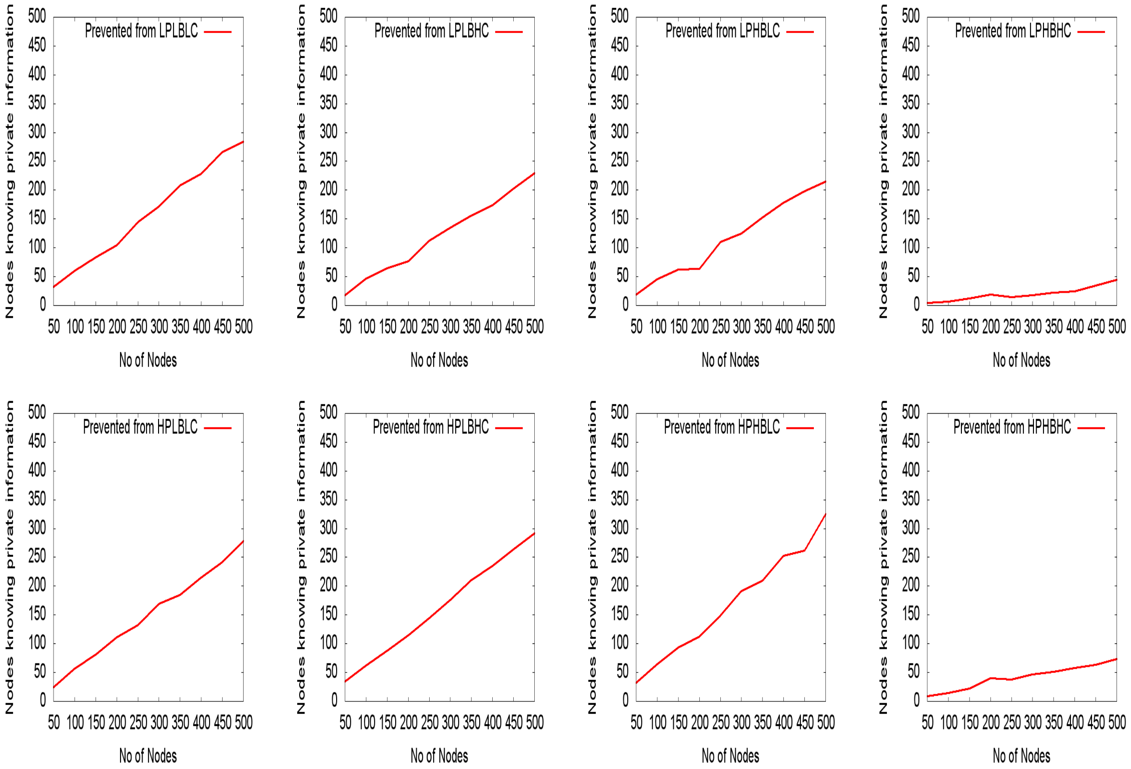


Figure 4.1: The spread of sensitive information after simulation

4.3 Algorithm for Enabling Selective Information Sharing in OSNs

Contents like *photos, videos, likes, dislikes* etc. are shared in an OSN. These contents have different levels of sensitivity. Personal photos and videos etc. can be highly sensitive in nature whereas some contents like the photos of an actor, nature etc. are less sensitive. Other than being high or low in sensitivity, a content can have an average sensitivity too. Generally we would like to share the sensitive content to our close friends and people whom we trust offline [34]. This restricted sharing does not guarantee data privacy as the close friends may or may not have a good *PQ* or their position in the network may be such that the chances of the sensitive information being leaked through them is high. Hence, sharing information to such nodes could result in an unwanted privacy breach.

The content shared in an OSN can be categorized into three categories on the basis of their sensitivities.

- Content having high sensitivity
- Content having average sensitivity
- Content having low sensitivity

Table 4.2 summarizes the relation amongst the sensitivity of the shared profile item, property of the node, its group and the sharing suggestions given by the proposed privacy preserving algorithm.

Table 4.2: Sensitivity of the item, property of nodes, their groups and the sharing suggestions by the algorithm

Sensitivity of profile item	Property of node	Group	Sharing Suggestions
High	{L,H,H},{H,H,H}, {L,L,H},{L,H,L}	Group 1	Do not share sensitive information
Average	{L,H,H}	Group 2	Do not share sensitive information
Less	Any property	-	Share sensitive information following ICM.

Given below are the three algorithms proposed for each of the above categories. Initially the value of *chance* for all the nodes is equal to *false* which implies that none of the nodes have got a chance to spread the sensitive information in the network. The variable *knowPI* denotes whether the neighbor knows the sensitive information or not.

Algorithm 4.3.1: HIGHLYSENSITIVE(V)

```
if $chance[V] = False$
 and $Neighbor[V] =$ not $knowPI$ and $haspropertyof =$ "$Group1$"
   then $share =$ false
   else $share =$ true and $Neighbour[V] = knowPI$ and $chance[V] = true$
```

Algorithm 4.3.1 is applied if the item that is to be shared is highly sensitive. If the neighbors of the source node *V* do not know the personal information, and belongs to *Group 1* then the information is not shared with them.
Otherwise the source follows ICM and passes its information to rest of the neighbors which do not belong to *Group 1*

Algorithm 4.3.2: AVERAGESENSITIVE(V)

```
if $chance[V] = False$
if $Neighbor[V] =$ not $knowPI$ and $haspropertyof =$ "$Group2$"
   then $share =$ false
   else $share =$ true and $Neighbour[V] = knowPI$ and $chance[V] = true$
```

Algorithm 4.3.2 is applied if the item that is to be shared is of average sensitivity. If the neighbors of the source node *V* do not know the personal information, and belongs to *Group 2* then the information is not shared with them.
Otherwise the source follows ICM and passes its information to rest of the neighbors not belonging to *Group 2*. Here we do not impose too many restrictions as in the previous algorithm.

Algorithm 4.3.3: LESSSENSITIVE(V)

```
if $chance[V] = False$
if $Neighbor[V] =$ not $knowPI$
   then $share =$ true and $Neighbour[V] = knowPI$ and $chance[V] = true$
```

Algorithm 4.3.3 is applied if the item that is to be shared is less sensitive in nature. If the neighbors of the source node *V* do not know the personal information then the information is passed to their neighbors using the ICM without applying any restrictions as in the previous algorithms.

Figure 4.2 represent the snapshots from NetLogo [112] for the sensitive information spread of less sensitive, average sensitive and the highly sensitive information. Here the blue nodes are the information aware nodes and the magenta nodes are the node that do not know the information.

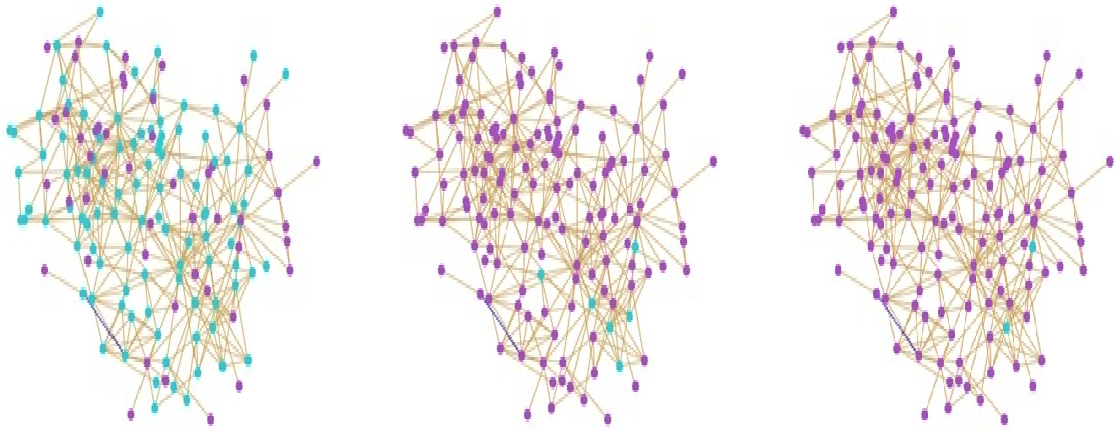


Figure 4.2: Snapshot from NetLogo of the spread of less sensitive information, average sensitive information and highly sensitive information

Figure 4.3, *LSI* indicates the Less Sensitive Information spread, *ASI(WM), ASI(M), HSI(WM) and HSI(M)* indicates the *Average Sensitive Information spread* and *High Sensitive Information spread.* Here *WM* and *M* indicates Without Model and with the implementation of Model respectively. The lines in red indicates the sensitive information spread

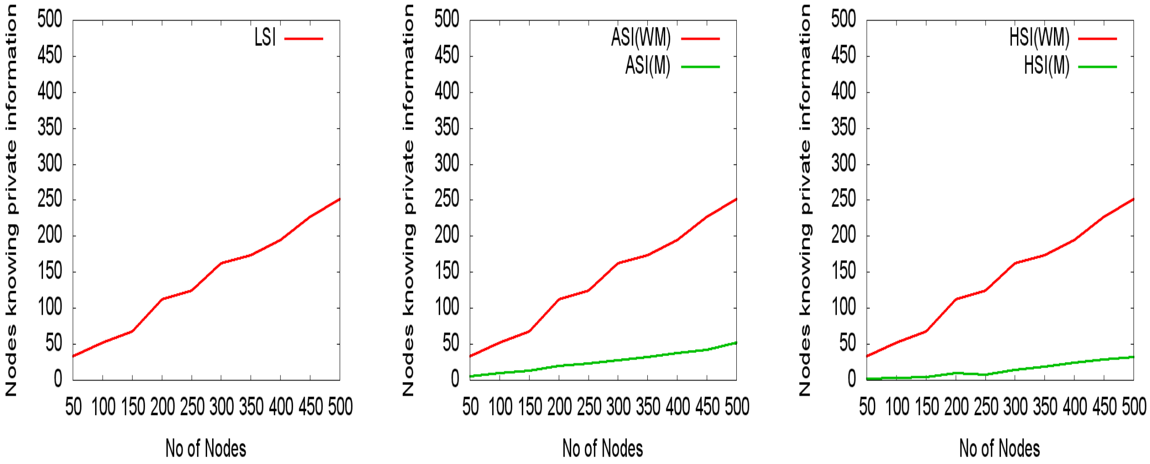


Figure 4.3: Simulation results for the spread of less sensitive information, average sensitive information and highly sensitive information

without the model implementation and the lines in green represent the sensitive information spread after the model was implemented. We see that the sensitive information spread reduces greatly with the implementation of the privacy preserving model.

4.4 Comparison with Cutillo's Approach

Cutillo et al. [53] have defined Q_v which is a measure of the average ratio of a node *v's* friends that can obtain sensitive information disclosed by a malicious friend. They found out that the network having a low Q_v will ensure the best privacy protection. Here Q_v and p_v are defined as

$$Q_v = p_v * c(v) \ \ and \ \ p_v = 1 - p_{mal}^{deg(v)} \tag{4.4.1}$$

They assumed p_{mal} to be .01, *deg(v)* is the degree and *c(v)* is the clustering coefficient of the node v. In Table 4.3, *X* is the percentage of sensitive information aware nodes without implementing the model and *Y* is the percentage of sensitive information aware nodes with the implementation of model.

Observation : Table 4.3 shows the results where we simulated four random networks

Table 4.3: Comparison with the analysis of Cutillo et al.

Network No	$\overline{deg_v}$	$\overline{pv}$	C(G)	Q$_v$	X	Y
Network 1	5.521	.0539	.5034	.0271	74.6	49.5
Network 2	5.20	.0509	.4861	.0247	78.1	41.2
Network 3	3.51	.0346	.2968	.01029	58	21.8
Network 4	3.571	.0352	.2762	.0097	54	17.6

following preferential attachment. We see that out of all the four networks, *Network 4* is having the least value of Q_v and is the most privacy preserving. Privacy in this network can be further preserved if we consider the sensitivity of the shared item and then apply the proposed model. The sensitive information aware nodes were 54% before the implementation of the model and was reduced to 17.6% after the model was implemented. The same behavior was observed with the other three networks also. After modeling and analyzing our algorithm we conclude that privacy of a node depends on the sensitive information sharing behavior as well as the position of the node in the network. We measured privacy of a node using the privacy quotient and its betweenness and closeness centrality in the network. Therefore, in order to prevent the sensitive information spread in an online social network the user should decide the scope of the shared item by analyzing the sensitivity of the item. Information disclosure can be reduced by preventing sharing of the highly sensitive information from the nodes belonging to Group 1.

Experiments and research have proved that when a trust bond exists between the individuals in an environment, the perceived risks involved in revealing personal information is diminished [113]. In an OSN if the sensitive content is shared to the selected online trustworthy connections then the chances of unwanted disclosure is reduced greatly. The major challenge here is to build and establish trust in an OSN environment. Hence, in the next section we will calculate *direct trust* for the online users in order to form an online trusted community which would prevent the users' sensitive information from unwanted disclosures.

4.5 Quantifying Direct Trust for Sensitive Information Sharing in OSN

In order to form an online trusted community we need to quantify and measure the trust for each individual. The amount of trust user X has on user Y can easily be collected by monitoring *X's* experience with Y, measuring *X's* attitude information for Y or by tracking *X's* behavior for Y over time [114]. Interactions are a measure of trust [59]. Users often interact with their *strong ties* and share most of their information with them [115]. Though a community of people with strong ties is a highly trusted one, but when privacy is a concern sharing information with these *strong ties* without considering their online sharing behavior might result in an unwanted and unintentional information disclosure. Hence, there is a need to quantify *online direct trust* for the strong ties to minimize the insider privacy attacks and prevent information disclosure.

4.5.1 Proposed Model for Trust Building

We give an overview of the proposed trust building model for preventing unwanted information disclosures. Steps involved in the model are stated below and are explained in detail in the subsequent subsections.

- Measuring the *PQ* of each node in the network.
- Forming a trusted community T_c of *strong ties* based on the *interactions*.
- Computing the *direct trust* of all the members in T_c after considering their online sharing behavior.
- Classifying the strong ties as *unconcerned, pragmatics and fundamentalist*.
- Filtering out the *unconcerned strong ties* from T_c to form a *refined trusted community*.

4.5.2 Measuring Privacy Quotient for the Static Objects

In an OSN the static profile items are the ones that are not changed too frequently and have a limited set of values. Some examples of static profile items are *hometown, current-town, date-of-birth, relationship status, contact number, job details, e-mail* etc. In order to calculate the *Static Privacy Quotient* i.e. *PQ(Static)*, we calculate the sensitivity and visibility of objects using the *IRT* [116]. *PQ(Static)* is the product of *sensitivity* and *visibility* of the profile items and can be calculated as Equation 4.5.1 [51].

$$PQ(Static)_j = \sum_i \beta_i * \frac{1}{1 + e^{\alpha_i(\theta_j - \beta_i)}} \tag{4.5.1}$$

where β_i is the sensitivity and $\frac{1}{1+e^{\alpha_i(\theta_j-\beta_i)}}$ is the visibility of the i^{th} profile item and α_i is the discrimination constant of the i^{th} profile item.

Dynamic contents are the ones that are frequently uploaded and can have different levels of sensitivity. Users' *photos, videos, links, posts* etc. are some of the dynamic contents. In order to calculate the *PQ* for dynamic objects we need to look at the ratio of objects kept as private to objects shared to public.

$$PQ(dynamic) = \frac{C_{Pr}}{C_P} \tag{4.5.2}$$

Here C_{Pr} is the count of objects kept as private and C_P is the count of objects shared to Public. We can either use *PQ(Static)* or *PQ(dynamic)* based on the requirement of the target user. The procedure to calculate the *privacy quotient* for the static and dynamic objects is explained in detail in Chapter 3.

4.5.3 Forming a Trusted Community of Strong Ties Based on Interactions

The nodes in an OSN exchange information with each other in order to build and maintain the *social capital*. The combination of amount of time, the emotional intensity, the intimacy and the reciprocal service characterizes a tie. There are *strong ties* and *weak ties* in the network. The connection between the individuals who belong to distant areas of the social graph are known as the *weak ties*. In contrast to this the *strong ties* are the trusted and known individuals. Any OSN comprises of sets of nodes and edges where the nodes are the actors or individuals and the edges signify the relationship between them. The nodes interact and communicate with each other. These interactions are a good indicator of social relationships and conversational trust [59]. The *strong ties* of any node *X* are the nodes in the network with whom the node *X* interacts frequently. On the basis of interactions these nodes usually form a community of close and trustworthy friends of the user. This community would also have people who might not be privacy conscious and their profiles would be having a weak privacy strength. Therefore in order to form a community of privacy aware strong ties we need to quantify the *direct trust* by taking the sharing behavior of the nodes into consideration. In the next subsections we explain the procedure to quantify the *direct trust* for each node and the methodology to refine the trusted community.

4.5.4 Quantifying the Direct Trust for all the Strong Ties in the Network

A lot of OSNs use the concept of *"friends"* or *"close-friends"* to provide privacy control. Though this measure reduces the risk of disclosure but is extremely coarse. So we require a fine-grained privacy control mechanism using which the users would have the privilege to share objects to their privacy aware strong ties. This can help the users regulate their online data without limiting the value of the *social capital*. In order to achieve this we need to compute the *direct trust* between the nodes having an edge [62] with the target user. Direct trust is the trust value obtained with the direct relationship. If *s* and *t* are the two nodes then the direct trust between *s* and *t* can be calculated as

$$DT(s,t) = \frac{W(R_i)}{N_s} * pfactor(s,t) \tag{4.5.3}$$

Here $W(R_i)$ is the weight of the relationship between *s* and *t*. A strong tie interacts more with the target user in comparison with a weak one hence the strong tie will have a higher value of $W(R_i)$ than the weak tie.
N_s are the total number of edges a node *s* has.
pfactor of a node *s* for node *t* can be defined as

$$pfactor(s,t) = PQ(t) * (1 - PQ(s)) \tag{4.5.4}$$

Derivation of *pfactor*

The probability that the node *t* will hide the information that a node *s* shares to it is the pfactor(s,t) which is given as:

$$pfactor(s,t) = Prob(t\,hides) \bigcap Prob(s\,shares) \tag{4.5.5}$$

$$Prob(t\,hides) \bigcap Prob(s\,shares) = (Prob(t\,hides)|P(s\,shares)) * Prob(s\,shares) \tag{4.5.6}$$

Event 1 : Node *s* sharing the information to *t*
Event 2 : Node *t* hiding the information shared by *s*
Event 1 and Event 2 are independent events. So the final equation can be rewritten as

$$Prob(t\,hides) \bigcap Prob(s\,shares) = Prob(t\,hides) * P(s\,shares) \tag{4.5.7}$$

Privacy quotient (PQ) is directly proportional to the probability that determines the intensity of information hiding. We can replace the probabilities with the respective privacy quotients of the nodes.
The probability of the node *t* hiding the information is the *PQ* of *t* and the probability of *s* sharing an information is *(1 - PQ(s))*

$$Prob(t\,hides) \bigcap Prob(s\,shares) = PQ(t) * (1 - PQ(s)) \tag{4.5.8}$$

Adding the pfactor while computing the direct trust will also consider the online sharing behavior of the node *t* before s builds a trust value for it.

4.5.5 Classifying Nodes on the basis of Privacy

Dr. Alan Westin [30] conducted over 30 privacy surveys and created one or more privacy indexes to summarize his results and show the trends in privacy concerns. Westin has classified the public into three categories and each of which is explained as follows:

- ***High and the Fundamentalist:*** These are the people who are cautious and worried about their privacy. They are distrustful of the organizations and choose privacy over any form of consumer service benefits.
- ***Medium and the Pragmatics:*** These are the ones who believe that business organizations should earn the trust rather than assuming that they have it. They look at the benefits provided to them in comparison with the degree of intrusiveness of their personal information.
- ***Low and the Unconcerned:*** These people trust the organizations with the collection of their private information and are ready to use the customer service benefits in exchange of their personal information.

4.5.6 Filtering out the Unconcerned Users from the Trusted Community

Using the *direct trust (DT)* calculated from equation 4.5.3 we categorize all the strong ties as *unconcerned, pragmatics and fundamentalist*. We use *k means clustering* to determine the two thresholds. Table 4.4 shows the classification of strong ties as unconcered, pragmatics and fundamentalist. After identifying the unconcerned nodes within the trusted

Table 4.4: Classification of strong ties according to Westin

DT $\geq$ 0 and DT $<$ threshold 1	Unconcerned Nodes
DT $\geq$ threshold 1 and DT $<$ threshold 2	Pragmatics Nodes
DT $\geq$ threshold 2	Fundamentalist Nodes

community of friends, we refine the community by filtering them out.

The entire concept could be better explained by the figures below. In Figure 4.4 we show a trust aware model that calculates trusts on the basis of the weight of the relationship. The ties having a higher relationship value are categorized as the strong ties and information is shared with them without considering their online sharing behavior. The red and the blue nodes indicate the strong and the weak ties respectively. This could also result in an unwanted information disclosure. Figure 4.5 shows that before sharing this information with the strong ties in the network their online sharing behavior is also taken into consideration. All the strong ties are classified as *unconcerned, pragmatics and fundamentalist*. We represent them as *U, P* and *F* respectively. After classification the unconcerned strong ties are filtered and the information is shared with the refined trusted community of strong ties. Using this approach the unwanted disclosures are reduced to a great extent. In the following sections we prove the same.

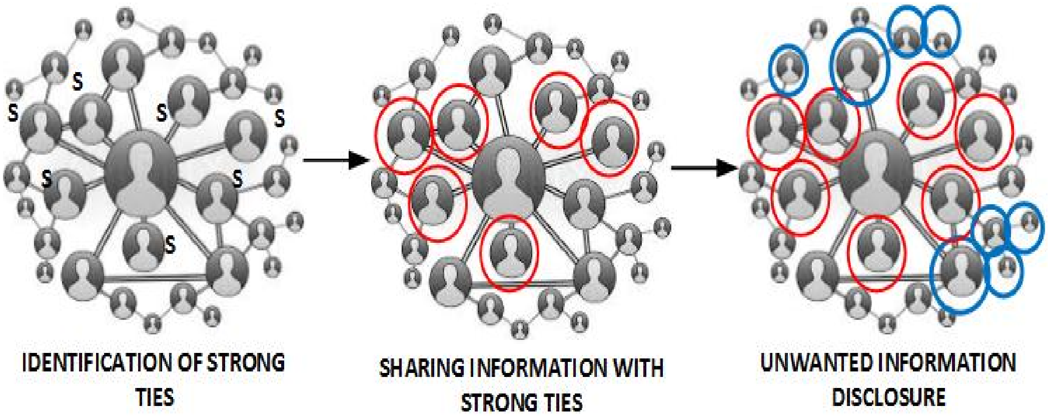

S
S
S
S
S
S
S
IDENTIFICATION OF STRONG TIES
SHARING INFORMATION WITH STRONG TIES
UNWANTED INFORMATION DISCLOSURE

Figure 4.4: Model that does not consider the online sharing behavior of strong ties

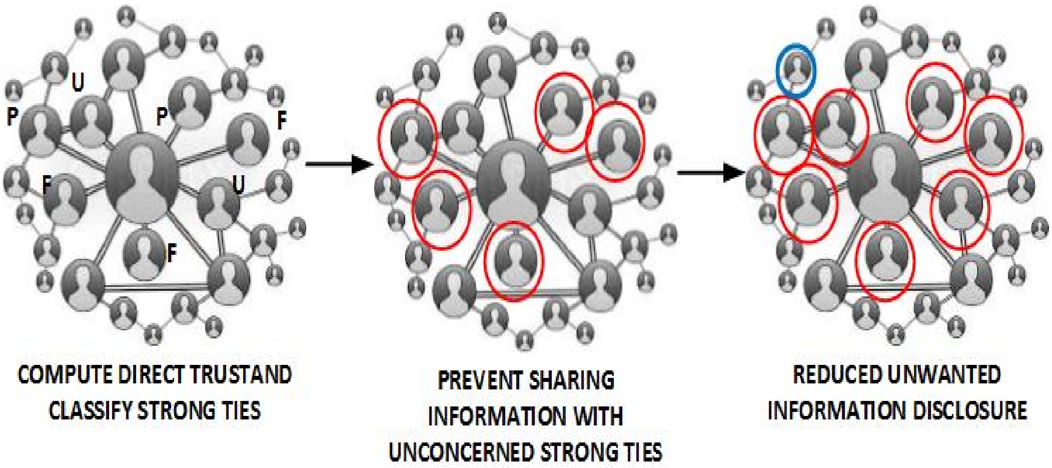

U
P
P
F
F
U
F
COMPUTE DIRECT TRUSTAND CLASSIFY STRONG TIES
PREVENT SHARING INFORMATION WITH UNCONCERNED STRONG TIES
REDUCED UNWANTED INFORMATION DISCLOSURE

Figure 4.5: The proposed privacy preserving model that considers the online sharing behavior of strong ties

4.6 Results and Discussion

4.6.1 Simulation Set Up

We obtained the data sets from *Stanford Network Analysis Project* [117] for *Facebook* which is a popular social networking site. Table 4.5 lists the complete statistics of different networks that were studied. Table 4.6 lists some of the properties like the *number of nodes, clustering coefficient, average node degree* etc. for each network. Here *WCC* and

Table 4.5: Statistics of the collected Facebook data

Nodes	4039
Edges	88234
Nodes in largest WCC	4039
Edges in largest WCC	88234
Nodes in largest SCC	4039
Edges in largest SCC	88234
Average clustering coefficient	0.6055
Number of triangles	1612010
Fraction of closed triangles	0.2647
Diameter (longest shortest path)	8

SCC represents the *weakly connected components* and the *strongly connected components* respectively. Depending upon the requirement the appropriate category of *(PQ(static)* or *PQ(dynamic)* was selected. We have experimented with the static as well as with the dynamic values of *PQ* which we measured using equations 4.5.1 and 4.5.2. The results shown below are for the dynamic *PQ* and the algorithm performs equally well for the static *PQ* as well.

Table 4.6: Properties of the different network data sets

Number of Nodes	Average Degree	Network Diameter	Average Clustering Coefficient
Network 1	1034	51.739	.534
Network 2	224	28.5	.544
Network 3	150	22.573	.67
Network 4	168	19.764	.534
Network 5	786	35.684	.476
Network 6	747	80.388	.635
Network 7	534	18.026	.544
Network 8	52	5.615	.462
Network 9	333	15.129	.444

We performed two sets of simulations with the aim of measuring the private information spread.

- ***Simulation 1 :*** Measure the sensitive information spread through the *unconcerned, pragmatics* and *fundamentalist strong ties* of the target user.

- ***Simulation 2 :*** Measure the sensitive information spread through the *pragmatics* and *fundamentalist strong ties* of the target user.

In Figure 4.6 the *blue (rhombus)* line indicates the spread of sensitive information including the unconcerned nodes and the *red line (square)* indicates the spread of sensitive information without including the *unconcerned nodes*. We measured the sensitive information spread with the list of networks mentioned in Table 4.6. The graphs show the percentage of nodes knowing the sensitive information with every iteration. *Independent Cascade Model*

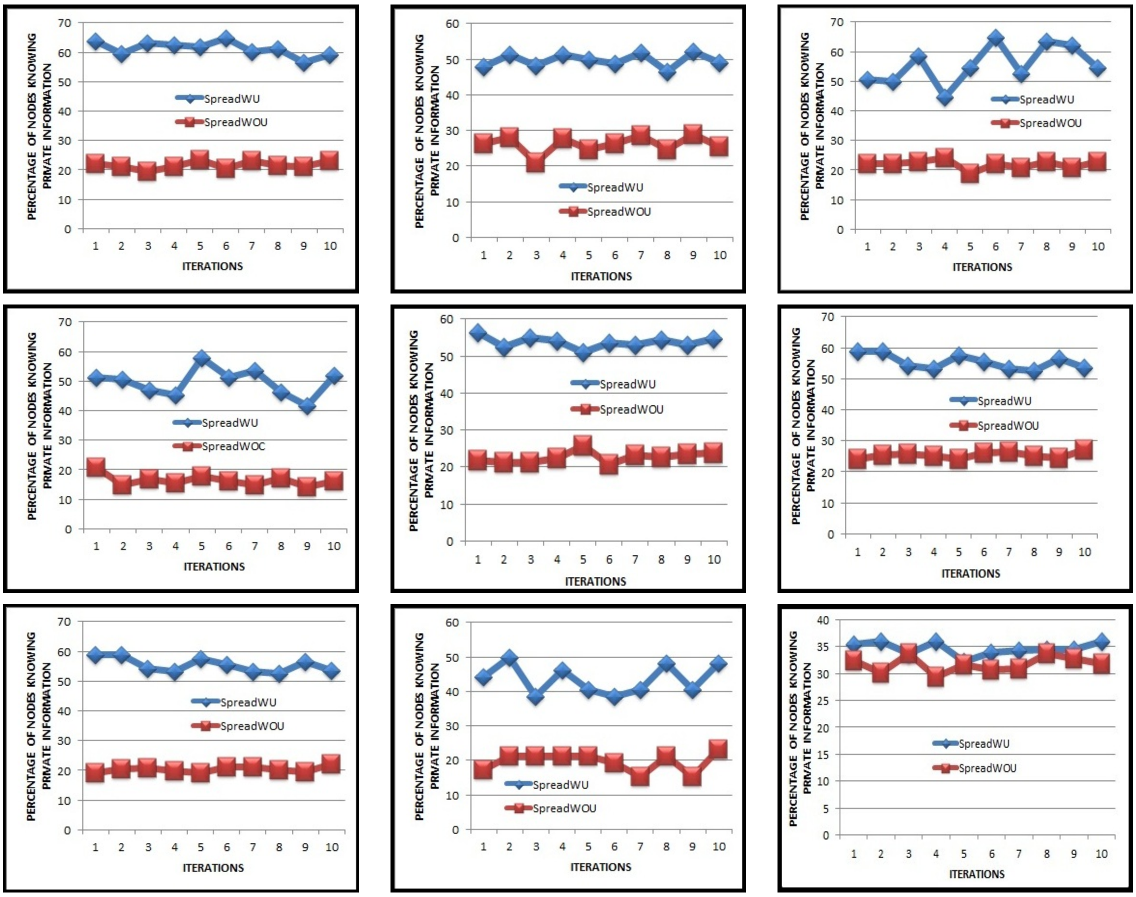


Figure 4.6: Comparison of sensitive information spread with and without including the unconcerned strong ties

(ICM)[111] was used to model this information spread from the source to the other nodes in the network. In the first wave all the strong ties share their information with the nodes they are directly connected to with the probability proportional to (1 - PQ). In the second wave the sensitive information aware neighbors might spread the information further with a probability proportional to (1 - PQ). This continues until all the information aware nodes have got at least one chance to spread this information with their neighbors.

Observation: It was observed that while sharing the sensitive information with the trusted strong ties if the unconcerned users are excluded from the trusted community of strong ties the sensitive information-aware nodes are reduced to a great extent.

4.7 Summary

In this chapter we analyzed the problem of unwanted information disclosure through the users' connection in an OSN. At first we emphasized and proved the fact that the sensitive information spread not only depends on the sharing behavior of the nodes but also on its topological features. We studied the relationship between centrality and the information flow through the node in a network. Using the proposed privacy preserving algorithm we could prevent the sensitive information to pass through the nodes that have a high probability of spreading information. With such selective sharing we could reduce the unwanted and unintentional disclosures to a great extent. Secondly, we studied the relation between privacy and trust in an OSN. We proposed and implemented a trust building model using which we calculated the direct trust between two nodes after considering their online sharing behavior. We identified and filtered out the unconcerned strong ties within the network which reduced the unwanted spread of sensitive information. The proposed algorithms could be used to build a third party application or a browser plugin which could additionally help the users in managing the sharing of information online. The relation between privacy and graph structural analysis is published in [Pub7]. The use of direct trust to form the refined trusted community of privacy aware nodes is published in [Pub8].

Chapter 5

Protecting Sensitive OSN Data from Inference Attacks

"The question isn't, What do we want to know about people?, It's, What do people want to tell about themselves?" - Mark Zuckerberg

5.1 Introduction

OSNs have provided a powerful means for sharing, organizing and finding contacts. Information in an OSN has a scope defined by the set of authorized entities. These entities have the right to know and view the information. Information that goes out of its scope can lead to privacy breach. In order to avoid these privacy breaches it is important that the sensitive information should be prevented from unwanted disclosures caused by unauthorized entities. Social Network data mining is about extracting important, meaningful and previously unknown information from social network data. In order to obtain better insights the OSN data owners share the social network data and its structure with the third party [118]. This detailed data in its actual form contains some sensitive information about the OSN users. Users have the right to make the sensitive items in their profile as private and selectively disclose the data to a section of connections. Policies and data protection strategies guarantee that only the data that is made publicly available by its owner gets published to the external third party entities.

The main stake-holders in the entire process are the *data owners, data holders* and the *data recipient*. The description of each of them is as follows:

- **Data Owner**: These are the users of an OSN who share their *personal identifiable information(PII)*, express their feelings and build and maintain relationship with other users in the network.
- **Data Holder**: These are the social networking sites such as *Facebook, Twitter, LinkedIn* etc. They collect, store and process the data obtained from the data owner.
- **Data Recipient**: These are the third parties to which the data holder releases the data collected from the data owner so that the data holder can improve and enhance their business solutions.

Data holders can release the data in the form of attribute list and edge list. This data before its release can be anonymized by using sophisticated methods from the field of *privacy preserving data publishing (PPDP)* [64] [65]. This however does not protect users from *inference attacks*. Using an *inference attack* the private attributes of an individual could be inferred by analyzing the public attributes of their associations even if the released data set is anonymized. Another relevant issue is that the data which is not perturbed carefully would have less utility and will not remain meaningful to the data recipients. Hence, it becomes an important problem to design efficient privacy preserving algorithms which could maintain utility and reduce the accuracy with which an attacker could infer sensitive attributes. In this chapter we propose and implement a *partial edge set removal* algorithm which modifies the structure of the graph such that the released graph becomes privacy preserving and has enough utility. The proposed algorithm also notably reduces the accuracy with which the adversary would infer the sensitive attributes. We propose a *privacy-utility trade-off* algorithm using which the *data holders* would be able to decide the best edge anonymization strategy to follow before releasing the data to the *data recipient*.

The major contributions of the chapter are as follows:

- We give an overview of some of the possible *inference attacks* due to *network classification* in an OSN.
- We describe the proposed efficient *partial edge set removal* algorithm.
- We measure the *utility loss* by taking different *structural metrics* for the released graph structure.
- We propose a *utility privacy trade-off* algorithm which would help the data holders decide the *optimal edge anonymization scheme* before releasing the data.

5.2 The Preliminaries

In order to study social networks extensively the *data holders* share data and structural information to the third party data *recipients*. The data recipient then analyzes the interactions between nodes, looks at the structural patterns and concludes certain findings out of the same [119]. They analyze the data for text analysis, search, image analysis, detecting patterns, target advertising etc.

The two main categories of data that can be released in an OSN are *linkage based* and *content based* [120]. The explanation for each of them is as follows:

- **Linkage Based:** In this category the data holder releases the structure of the social network depicting who is connected to whom.

- **Content Based:** A lot of content like the *PII, photos, videos, user generated content* etc. is generated online and releasing this data enables content based analysis.

For better results the researchers combine linkage based and content based analysis. Individuals who are aware of privacy issues do not disclose their sensitive attributes and keep it private. An important issue here is that after knowing the content and the structure of the social network graph the hidden attributes could be easily inferred by using some sophisticated probabilistic models. Consider a marketing application where the users with a particular label say U are interested to buy a product X. Other nodes that are connected to the set of users U, who have not explicitly shared the label could be targeted by inferring the hidden labels. This is done by making use of the correlation property which says that if two nodes are correlated then their node labels are correlated too [121]. Figure 5.1 explains the problem statement clearly.

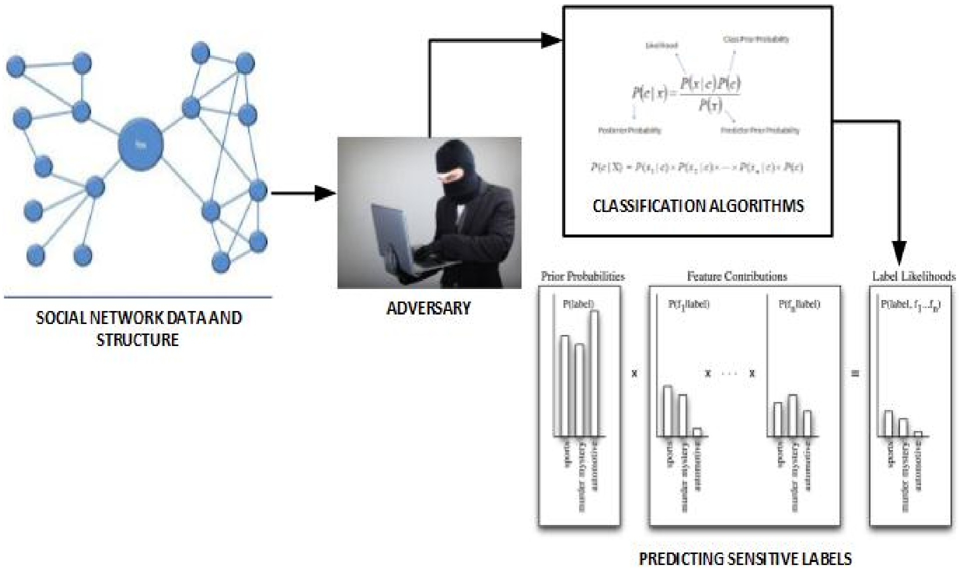

SOCIAL NETWORK DATA AND STRUCTURE
ADVERSARY
Likelihood
Class Prior Probability
$P(c|x) = \frac{P(x|c)P(c)}{P(x)}$
Posterior Probability
Predictor Prior Probability
$P(c|X) = P(x_1|c) \times P(x_2|c) \times \cdots \times P(x_n|c) \times P(c)$
CLASSIFICATION ALGORITHMS
Prior Probabilities
P(label)
Feature Contributions
P(f$_1$|label)
P(f$_n$|label)
x
x · · · x
=
Label Likelihoods
P(label, f$_1$...f$_n$)
sports
murder mystery
adventure
PREDICTING SENSITIVE LABELS

Figure 5.1: Inference attacks on the data set and structure using classification algorithms

The data holder releases the attribute list and edge list to the third parties. The untrusted third party could use network classification algorithms that involve the use of *local* *classifiers, relational classifiers and the inference algorithm* to infer private sensitive attributes from the available set of public and private attributes. In the next section we will discuss the adversary model in detail.

5.3 The Adversary Model

In this section we will discuss different inference models that could be used by the attackers to infer sensitive attributes from the users' profile. Individuals in an OSN share a lot of data about themselves. This data could be seen as *labels*. Some of the examples of these labels are *location, age, gender, political views, religious views, interests, hobbies, education details, work details* etc. The third party would want to know the value of the labels on all the nodes to carry out various data analysis tasks. Most of the users do not share their labels and keep it private due to different privacy concerns. This degrades the performance of the inference model drastically. In an OSN the links between the nodes are not random. They represent some relationship between the individuals sharing an edge. Therefore, there is a similarity between the nodes that are linked to each other. Inferring private labels by analyzing the public labels is a *classification problem*. This classification problem actually differs from the traditional classification problem where the nodes are independent of each other whereas in *collective classification* [122] the features of the object depends greatly upon the nodes with which they are linked to.

5.3.1 Categorizing Inference Attacks Based on Link Information

Depending on whether the link information is provided in the released data set or not the inference attacks could be divided into two broad categories.

- Attack without the link information
- Attack with the link information

Attack Without the Link Information

This attack comes into play when the *data holder* does not release the link information of the social network graph but instead releases the OSN data in an anonymized form. Even if the data is anonymized, attacks such as the *basic attack* could be used for inference. The *basic attack* considers the sensitive attribute value as the *random variable*. This disclosure problem can now be modeled using a *probabilistic model M* for predicting the sensitive value of the node. If the overall distribution of the sensitive attribute is known using the public profile then the private attributes could be inferred. The problem can be explained using equation 5.3.1.

$$P_{BASIC}(v_p = a_i) = P(v_p.a = a_i|V_o) = \frac{|V_o.a_i|}{|V_o|} \tag{5.3.1}$$

According to equation 5.3.1 the probability that a node *v* will have a private label a_i is the ratio of total number of nodes having the observed label as a_i i.e. $V_o.a_i$ to the total number of observed labels V_o. Here, the observed labels are synonymous to the labels that are shared publicly in the network.

Resolving the inference attacks without links is not in the scope of our work. Moreover many well known algorithms already exist in place that modifies the distribution of the attributes in the data set in order to prevent inference. Any information that is explicitly shared to public by the user could easily be obtained from their respective profiles so modifying or anonymizing information which is publicly available would not be of much help. The major concern is to prevent those users from *inference attacks* who are cautious about privacy and did not share their sensitive attributes publicly.

Attack with the Link Information

In this type of attack the adversary uses the data information along with the structure of the social network graph to infer the hidden sensitive attributes of the users. In the next section we will discuss one such algorithm which is the *collective classification algorithm* using which the adversary could carry out the inference attacks in the released OSN data.

5.4 Collective Classification Algorithm

The collective classification [122] makes use of both the labels and the link structure in the graph. It has a *local classifier, a relational classifier* and an *inference algorithm*. The local classifiers are traditional classifiers. They build models using the labels of the training set and then use it to predict the unknown labels for the test set. *Naive bayes, SVM, decision trees* etc. are some of the popular ones in this category. The relational classifiers analyze the links of the graph and train the model according to the labels in the training set and then is used to predict the unknown ones in the test set. *Class distribution relational neighbor (cdRN), weighted-vote relational neighbor (wvRN), network only link based classification (nLB)* [123] etc. are some of the examples for the same. We now discuss the third component i.e. the inference algorithms in detail.

Inference Algorithm

Given a network *G* having a node *v*, three distinct types of correlations can be used for classification.

- Correlation between the labels of the node *v* and the observed attributes.
- Correlation between the labels of the node *v* and the observed attributes that also include observed labels of the neighboring nodes.
- Correlation between the label of node *v* and the private labels of the neighborhood.

Collective classification involves all the three correlations for classification. It is a combinatorial problem where $V = v_1, v_2, ...v_n$ are the set of nodes. Each node is a random variable *v* that can take any value from the list of available values in the domain. The set of nodes *V* are divided into two categories i.e. type *X* and type *Y*. Type *X*, where $X \subseteq V$ are the nodes for which we know the correct values of labels and type *Y* where $Y \subseteq V$ are the nodes for which the actual labels are hidden and unknown. Therefore, the task is to label all the *Y* nodes with the appropriate label from the entire list of the available labels $L = l_1, l_2, ...l_q$.

We will discuss some efficient collective classification algorithms such as the *Iterative Classification Algorithm (ICA) [124], Gibbs Sampling [125] and Relaxation Labelling* [123] subsequently.

Algorithm 5.4.1: ITERATIVE CLASSIFICATION ALGORITHM(G)

```
for all nodes $Y_i \epsilon$ Y // Initial bootstapping
Predict labels using $X_i \epsilon N_i$.
repeat
  Generate Ordering O for all nodes in $Y_i$
  for each node $Y_i \epsilon$ O
  Predict $Y_i$ using the new estimates of $N_i$
until all labels stabilize or a fixed number of iterations have been achieved
```

In order to determine the hidden label for a node Y_i the Algorithm 5.4.1 first assigns a prior probability to all the unknown nodes *Y*. It does so by training the initial classifier using the features of all the nodes $X_i \ \epsilon \ N_i$. Here N_i are the set of nodes in the neighborhood of Y_i. This classifier which is now trained is used to determine the label for Y_i. The process is repeated for all the *Y* nodes. Here *Y* are the set of all the nodes for which the private information is not known. Either the best suited label out of the group of fixed labels *L* is returned as an output by the classifier or the label in each iteration is assigned by applying the maximum likelihood estimation. The *ICA* takes the set of attributes for each node as an input. These attributes are the feature vectors for each node. As the nodes in a graph are linked so the link features also go as input to the classifier. After the first iteration the value of the link feature changes and some of the unobserved nodes get a label. Therefore, there is a need to iterate the algorithm until all the labels stabilizes or a fixed number of iterations are achieved.

Algorithm 5.4.2: GIBBS SAMPLING ALGORITHM(G)

```
**for** all nodes $Y_i \epsilon$ Y
Predict labels using $X_i \epsilon N_i$. // Initial Bootstrapping
**for** $i \leftarrow 1$ **to** $Burn - in$
    Generate Ordering O
    **for** each node $Y_i \epsilon$ O
    Predict val i.e. the label of $Y_i$ using the new estimates of $N_i$
**for** each node $Y_i \epsilon$ Y
    **for** each label $l_i \epsilon$ L
    $count[i, l] \leftarrow 0$
**for** $p \leftarrow 1$ **to** $Sample$
    Generate Ordering O
    **for** each node $Y_i \epsilon$ O
    Predict val i.e. the label of $Y_i$ using the new estimates of $N_i$
    $count[i, val] \leftarrow c[i, val] + 1$
**for** each node $Y_i \epsilon$ Y
$val \leftarrow argmax_{l \epsilon L} c[i, l]$
```

In Algorithm 5.4.2 initially the labels for each node Y_i ϵ Y are assigned based on the labels of nodes X_i ϵ X where X_i ϵ N_i. The labels are then assigned to each node Y_i using a local classifier for a fixed number of iterations called the “burn-in”. After which a sampling is done to calculate the best estimate of the label. A count is maintained to keep track of the number of times the label *l* was sampled for Y_i. The label that was assigned the maximum number of times to Y_i is selected as the final label. This is one of the most accurate approximate inference algorithms. *Gibbs Sampling Algorithm* is slow and it is difficult to determine the value at which it converges which adds to the drawback of this algorithm.

Algorithm 5.4.3: RELAXATION LABELLING ALGORITHM(G)

```
for all nodes $Y_j \epsilon$ Y // Initializing messages
  {for each $val_j \epsilon$ L
  {$b_j(val_j) \leftarrow 1$
repeat
  for all nodes $Y_j \epsilon$ Y
    {for each $val_j \epsilon$ L
    {calculate $b_j(val_j)$ the marginal probability distribution of assigning $Y_j$ with label $val_j$
until all $b_j(val_j)$ stabilizes
```

In Algorithm 5.4.3 a different approach for collective classification algorithm is followed where a global objective function is defined and is then optimized. Relaxation labelling through mean field is one such algorithm to do the same. Here we calculate $b_j(val_j)$ which is the marginal probability of assigning Y_j with label val_j. This process is repeated for every node Y_j until all the $b_j(val_j)$ values are stabilized.

5.5 Motivation and Basic Idea for Preventing Inference Attacks

Figure 5.2 shows the proposed algorithm which acts as a middleware. It reduces the accuracy with which the adversary would predict or infer the sensitive attributes. The exact details of the algorithm is explained later in the chapter.

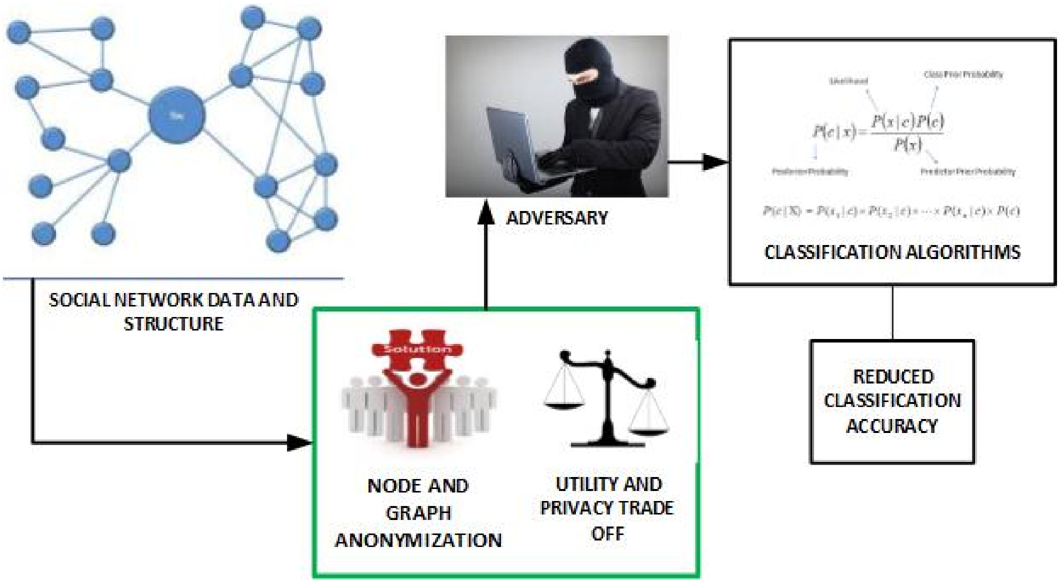

SOCIAL NETWORK DATA AND STRUCTURE
ADVERSARY
Likelihood
Class Prior Probability
$P(c|x)=\frac{P(x|c)P(c)}{P(x)}$
Posterior Probability
Predictor Prior Probability
$P(c|X)=P(x_1|c)\times P(x_2|c)\times\cdots\times P(x_n|c)\times P(c)$
CLASSIFICATION ALGORITHMS
REDUCED CLASSIFICATION ACCURACY
Solution
NODE AND GRAPH ANONYMIZATION
UTILITY AND PRIVACY TRADE OFF

Figure 5.2: Preventing inference attacks on the data set and structure using the proposed algorithm

The edges in an OSN are not random. Any two people who are connected with an edge share some common characteristics. According to Social Sciences, if two people are related then primarily the principles of homophily and co-citation regularity [124] hold good there. According to the concept of *co-citation regularity* when two similar nodes are connected to the same objects it implies that they share common interests or behavior. For example, If two people have same tastes for music, movies, hobbies then according to the *co-citation regularity* they have similar interests. According to *homophily* if two nodes are connected then they would have some common characteristics between them. This is also known as “birds of a feather” phenomenon. For example, people who are connected to each other and share an informal or formal relation have some common characteristics like *age, education, gender* etc. We use the knowledge of *co-citation regularity* and *homophily* and propose the following algorithm:

For an attribute A_i the nodes can have a value displayed to public or it could be hidden. This value could be any label out of the given set of labels *L*. A node can possibly have four associations between them i.e. *a private node linking to a private node, a private node linking to a public node, a public node linking to a public node and a public node linking to another private node*. In Algorithm 5.5.1 the complete graph *(G)* with a set of nodes *(V)* and the set of edges *(E)* will go as an input. This graph will have a set of *X* nodes i.e. the nodes whose sensitive attribute is set to public. It will also have a set of *Y* nodes i.e. the nodes whose sensitive attribute is hidden and is private. At first the algorithm takes the list of nodes and labels them as Private if V_i ϵ *Y* and Public if V_i ϵ *X*.

Algorithm 5.5.1: RELEASING PARTIAL EDGE SET(V, E)

```
for all nodes $V_i \epsilon$ V
  { Label $V_i$ as Private if $V_i \epsilon Y_i$.
  { Label $V_i$ as Public if $V_i \epsilon X_i$.

for all edges $E_i \epsilon$ E formed by the nodes $V_i$ and $V_j$ make a labeled edge set $E_{lab}$
  { Label $E_i$ as Private-Public if $V_i$ is Private and $V_j$ is Public
  { Label $E_i$ as Private-Private if $V_i$ is Private and $V_j$ is Private
  { Label $E_i$ as Public-Private if $V_i$ is Public and $V_j$ is Private
  { Label $E_i$ as Public-Public if $V_i$ is Public and $V_j$ is Public

for all edges in $E_{lab}(i) \epsilon$ $E_{lab}$
  { Do not include the edge in $E_{final}$ if $E_i$ is labeled as Private-Public
  { Include the edge in $E_{final}$ if $E_i$ is labeled as Private-Private
  { Do not include the edge in $E_{final}$ if $E_i$ is labeled as Public-Private
  { Include the edge in $E_{final}$ if $E_i$ is labeled as Public-Public

Release $E_{final}$
```

The algorithm takes the edge set *E* and analyzes the nodes V_i and V_j that forms an edge E_{ij}. Depending upon the visibility of the nodes the edges are labeled as *Private-Private, Private-Public, Public-Private and Public-Public*. This produces a new labeled edge set that we define as E_{lab}. To get to the final edge list we will remove all the edges labeled as *Private-Public* and *Public-Private* and keep rest of the edges intact. The algorithm will output E_{final} which is the final edge list that should be released.

No extra information could be inferred from the structure of the graph which was earlier not evident when only the data sets were being released. This could be properly explained using the equation 5.5.1.

$$I_{DS} - I_D \approx 0 \tag{5.5.1}$$

In equation 5.5.1, I_{DS} is the information known if the data and the structure both are released and I_D is the information known if only the data is released. This essentially means that the difference of the two should be approximately equal to zero i.e. the data holders can release more data having less privacy loss and more utility gain.

5.6 Results obtained for the Partial Edge Set Algorithm

We experimented the proposed solution on four categories of network data sets namely *CoRA, IMDb, Industry and WebKB* [126] and also verified the proposed solution on the data sets extracted from 15000 *Facebook profiles*. For implementation we used *NetKit-SRL* [126] which is an open source statistical relation learning toolkit. We will give a detailed explanation on the results obtained from the *CoRA* data sets.

The CoRA dataset comprises research papers from the field of Computer Science and includes the full citation graph as well as labels for the topic of each paper. There are seven labels namely *Case Based, Genetic Algorithm, Probabilistic Methods, Neural Networks,* *Reinforcement Learning, Rule Learning and Theory*. Table 5.1 refers to some of the important statistics of the *CoRA* data set. After applying the network classification algorithm on this data set the accuracy for the same was recorded. We used the *Receiver Operating Characteristics (ROC) curve* which is an important visual tool to compare the accuracy of the classification model. For a given model an *ROC* curve shows a trade-off between the *true positive rate (TPR)* and the *false positive rate (FPR)*. For a two class problem it helps to visualize the trade-off between the rate at which the model correctly classifies the positive cases versus the rate at which it misclassifies the negative cases as positive. For the *CoRA* data sets we had seven labels hence, had seven *ROC* curves, each for one of the seven labels respectively.

Table 5.1: Statistics for the CoRA data set

Description	**Count**
Total No of Nodes	4240
Total No of Edges	22514
Total No of Private Nodes	848
Total No of Public Nodes	3392
Total No of Private Private Links	840
Total No of Public Public Links	14596
Total No of Private Public Links	3539
Total No of Public Private Links	3539

The network classification algorithm uses *a local classifier, a relational classifier and an inference algorithm*. A lot of combinations are possible with different choices of local classifiers, relational classifiers and inference algorithm. Out of all those, we experimented with three such combinations, one with each of the inference algorithms discussed above. Table 5.2 lists all the three combinations that were used for classifying the hidden labels.

All the three combinations for each of the four data sets were experimented. For demon-

Table 5.2: The three combinations of local, relational and inference classifiers

Configuration No.	**Local Classifier**	**Relational Classifier**	**Inference Algorithm**
Configuration 1	Class Prior	wvRN classifier	Iterative Classification
Configuration 2	Class Prior	wvRN classifier	Relaxation Labelling
Configuration 3	Class Prior	wvRN classifier	Gibbs Sampling

stration we will discuss the results obtained for the second configuration which used the *class prior algorithm* as the *local classifier*, *wvRN algorithm* as the *relational classifier* and *relaxation labelling* as the *Inference algorithm*. Figure 5.3 shows an *ROC* curve for the label *"Case Based"*. The curve clearly indicates that the accuracy with which the adversary would classify the hidden labels decreases when the edges modified by the *partial edge set algorithm* is being released. Here the red line represents the accuracy of the network classification algorithms when all the edges are shared and the blue line indicates the accuracy of the algorithm using the *partial edge set algorithm*.

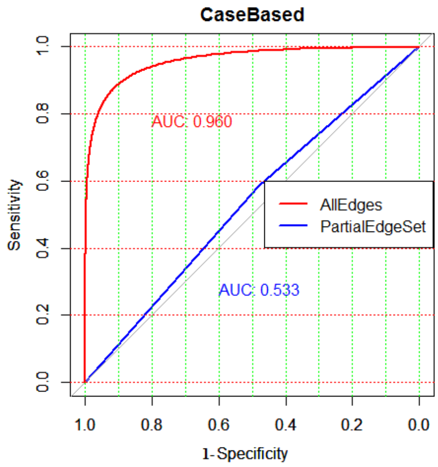


Figure 5.3: An ROC curve for the Case Based label of the CoRA data set

The area under curve for the modified structure is reduced to 53.3% in comparison to 96% which was observed for the original graph. Similarly, Figure 5.4 shows the ROC curves

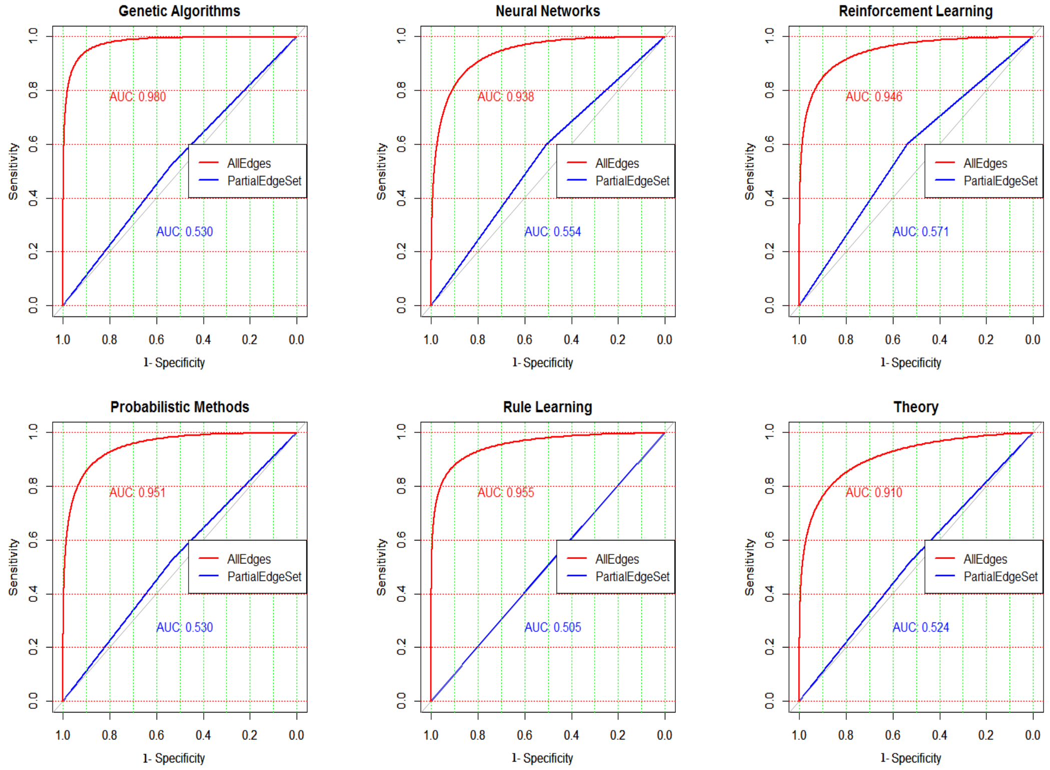


Figure 5.4: ROC curves for different labels of the CoRA data set

that we have obtained for the other six labels as well. Table 5.3 compares the performance of the same network classification algorithm on the original as well as the modified graph. We considered some of the performance metrics like *accuracy, precision, recall, false positive rate, true negative rate, false negative rate, F measure and kappa*. After applying the algorithm; *accuracy, precision, recall, true negative rate and kappa* showed a decrease in the value whereas the *false positive rate, false negative rate, F measure and misclassification rate* showed an increase. Thus, the *partial edge set algorithm* was able to reduce the accuracy of the classifier greatly.

Table 5.3: Values of different performance metrics for the CoRA data set

Measure	Publishing All Edges	Publishing Partial Edge Set
ACCURACY	90.21%	68.98%
PRECISION	53.90%	12.10%
RECALL	80.85%	28.72%
FALSE POSITIVE RATE	8.62%	26%
TRUE NEGATIVE RATE	91.37%	74%
FALSE NEGATIVE RATE	19.14%	71.27%
F MEASURE	.6468	.1702
KAPPA	.593	.017
MISCLASSIFICATION RATE	9.78%	31.07%

5.6.1 Results on Facebook Data Sets

We extracted the *Facebook* data sets for 15000 user profiles and segregated the data into sensitive and non-sensitive data. We modeled the data set to predict the sensitive data of the users. We experimented the algorithm on predicting the *relationship status*, *users' age group* and *users' political views*. Table 5.4 lists down the comparison between the metrics where the *relationship status* was considered as sensitive. The possible labels of the relationship status were *Single, Married, Engaged, In a Relationship and Divorced*.

Table 5.4: Values of different performance metrics for the Facebook data

Measure	Publishing All Edges	Publishing Partial Edge Set
ACCURACY	72.667%	54%
PRECISION	61.176%	43.077%
RECALL	86.667%	46.667%
FALSE POSITIVE RATE	36.66%	41.11%
TRUE NEGATIVE RATE	63.33%	58.88%
FALSE NEGATIVE RATE	13.33%	53.33%
F MEASURE	.71	.448
KAPPA	.468	.055
MISCLASSIFICATION RATE	27.33%	31.07%

Overall Deduction: The use of *partial edge set algorithm* could reduce the classification accuracy and increase the misclassification rate significantly.

In the next section we will measure the utility of the published graph to ensure that the proposed algorithm does not only preserve the privacy but also ensures utility for the released data sets.

5.7 Measuring Utility of the Published Graph

Graph perturbation for anonymization would result in the loss of utility but would also ensure a better protection for the individuals' privacy. A proper balance between privacy and utility is needed for a good practice of privacy preserving data publishing. Releasing a heavily perturbed graph would bring down the utility and ensure greater privacy to users. Abstaining data from sharing to the third party would be an ideal and unrealistic privacy protection strategy as it would ensure a 100% privacy protection but would have 0% utility gain. Contrary to this approach in order to achieve maximum utility for the data sets the data owners could release the entire graph structure which would result in a 100% utility gain and will have a 0% privacy protection for the data set. Therefore, we need to have an algorithm using which a trade-off between privacy and utility could be maintained and the entire purpose of publishing the data to the third party could be served.

For measuring utility we considered some important metrics such as *average path length, clustering coefficient, transitivity, modularity, graph density* and *reciprocity*. By removing the edges connecting the *Private-Public* associations for a sensitive attribute, we modify the structure of the actual graph. The more similar the modified graph is with the actual graph the more utility it is bound to have.

Table 5.5 compares the value of different metrics for the graph before and after perturbation along with the percentage of the graph perturbed. Figure 5.5 shows the comparison of different metrics for the actual and the modified graph.

Table 5.5: Values of different performance metrics for measuring utility for the CoRA data set

Measure	Modified Graph	Original Graph	% of perturbation
AVERAGE PATH LENGTH	2.3722	2.3214	2.189%
CLUSTERING COEFFICIENT	.30393	.30398	1.57%
TRANSITIVITY	.30393	.30398	1.57%
MODULARITY	.31433	.29786	5.529%
GRAPH DENSITY	.01873	.02717	31.059%
RECIPROCITY	.090671	.089664	1.122%

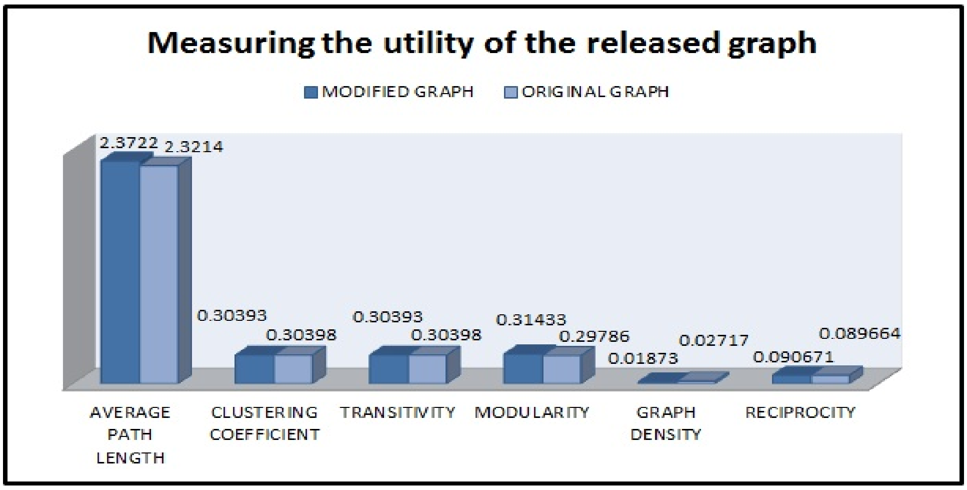


Figure 5.5: Comparison of different performance metrics for measuring utility for the CoRA data set

Observation: We observed that there was a small difference between the structural metrics of the data sets before and after perturbation. The minimum perturbation obtained in the CoRA data set was observed to be 1.57% and the maximum was 31.059%.

Different edge anonymization algorithms follow different perturbation mechanisms hence, it becomes a tough task for the *data holders* to decide the optimal edge anonymization algorithm that could ensure them the highest *privacy protection* and *maximum utility gain*. In the next section we will explore the issue of finding an optimal *trade-off* between *privacy* and *utility* for the released data sets. We draw motivation from the *Modern portfolio theory* [127] which is used in finance to maximize the expected return for a given amount of risk. It emphasizes on the mathematical formulation where while selecting the collection of investment assets the one with the lower risk is preferred. We draw similarity between privacy and *portfolio theory* to obtain optimal trade-off between privacy and utility. We map *risk* with *privacy* and *return* with *utility* to calculate the *trade-off* for various edge anonymization algorithms.

5.8 The Modern Portfolio Theory

In order to determine the trade-off between privacy and utility we draw concepts from *Modern Portfolio Theory*. Portfolio theory [128] is mainly about maximizing the return while minimizing the risk. A portfolio is essentially a combination of different assets. Figure 5.6 shows the mean-variance plot which demonstrates this concept. The vertical axis represents the expected rate of return and the horizontal axis signifies the investor's risk tolerance. According to the mean-variance model the agents base their investments mainly on the expected return and variance of the portfolio. Every investor would decide to choose a portfolio that offers higher returns and lesser risks. This could be depicted using the mean return of every asset against their standard deviation. A feasible set of portfolios out of all the given portfolios exist on the efficient frontier where it is possible to obtain a very low standard deviation for a given value of mean or a very high mean for a given value of standard deviation. Figure 5.6 represents the plot depicting the concept of efficient frontier. Here, out of all the points the green points are the points that represent an efficient frontier.

A combination of assets which form a portfolio is referred as "efficient" if it can give the best return for a certain level of risk. All the optimum portfolios should lie on the curve. Any portfolio that lies under the region represents a less than ideal investment. The reason

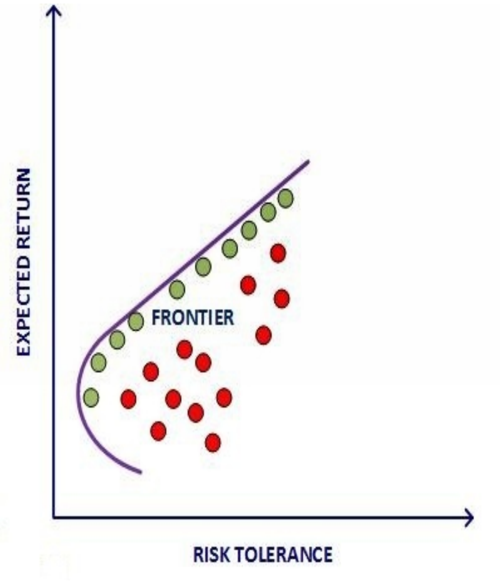


Figure 5.6: The mean variance plot between the risk and return

is specified in Figure 5.7. Consider the assets *A, B* and *C*. For a given value of return we can decide to choose between the assets *A* and *C*. Asset *A* will give us the same return but with a higher risk than asset *C*. Hence, asset *C* is an efficient selection. For a given value of risk we can decide to choose between the assets *A* and *B*. Here asset B will give a higher return than asset *A* for a fixed value of risk. Hence, asset *B* will be an efficient selection.

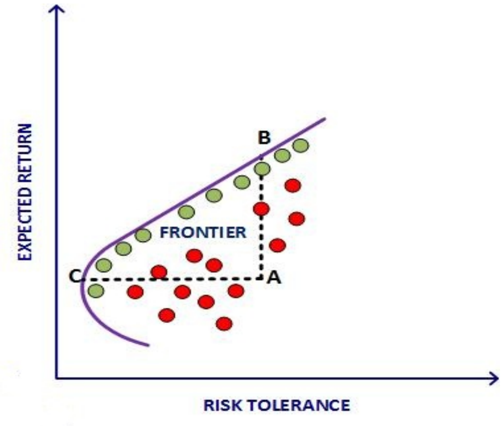


Figure 5.7: The mean variance plot depicting the selection of assets

For our work we map utility with the rate of return and privacy with the *risk tolerance*.

Each mechanism perturbs the data set *D* and changes it to the new data set D_A. This data set D_A will have a utility *u* and privacy *p*. Given different mechanisms and their respective *p* and *u* values the *data holders* should be able to select the mechanism which would ensure maximum utility and minimum privacy loss helping them to decide on the trade-off between privacy and utility. In the next subsection we will be discussing some edge anonymization algorithms [129] as our further analysis would be based on them.

5.8.1 Edge Anonymization Algorithms

Our main focus is to select the best edge anonymization algorithm for the OSN graph which could provide maximum utility and minimum privacy loss. For our work we will consider the following four edge anonymization algorithms as described by Zhelva et al. [129].

- Intact Edge Algorithm
- Partial Edge Removal Algorithm
- Cluster Edge Anonymization Algorithm
- Removed Edges Algorithm

Intact edges: It is the basic edge removal algorithm. The input to the algorithm is the list of edges (*E*), list of sensitive edges E_S and the percentage of the sensitive edges to be removed (Per). An edge is randomly picked from the list of sensitive edges E_S and is removed from the list of E_{final} which is the final list of edges to be released. This is repeated *EdgeRemoved* number of times where *EdgeRemoved* is the percentage of the edges to be removed.

Algorithm 5.8.1: INTACT EDGE ALGORITHM(E, E_S, Per)

```
EdgeRemoved = Per * count($E_S$)
for edgecount = 1 to EdgeRemoved
    Pick an edge $E_{ij} \epsilon E_S$ randomly
    $E_{final} = E - E_{ij}$
$E_{final}$ is released
```

Partial Edge Removal: In this type of anonymization algorithm the input to the algorithm is the list of edges (*E*), list of parameters List(Par) and the percentage of sensitive edges to be removed (Per). There could be different criteria to remove the partial edges from the edge set *E*. This criteria varies from one algorithm to the other. It could be removing the *Public-Private* associations as described earlier in the chapter, removing the edges connecting the high degree nodes, high betweenness, high centrality etc. Hence, only that percentage of the list of edges *E* are removed which fulfill the specified criteria. This is basically repeated *EdgeRemoved* number of times where *EdgeRemoved* is the percentage of the edges to be removed.

Algorithm 5.8.2: PARTIAL EDGE REMOVAL ALGORITHM($E, List(Par), Per$)

```
Label edges as sensitive i.e. $E_S$
that satisfies List(Par)
EdgeRemoved = Per * count($E_S$)
for edgecount = 1 to EdgeRemoved
    for each edge $E_{ij} \epsilon$ E
    if $E_{ij}$ fulfills the specified criteria
    $E_{final} = E - E_{ij}$
$E_{final}$ is released
```

Cluster Edge Anonymization: In this approach at first the anonymized nodes form different clusters using any standard clustering algorithm. The number of clusters is V_C which is an input to the algorithm. Each edge E_{ij} from the edge set *E* is taken and the vertices V_1 and V_2 forming the edge E_{ij} are identified. If the vertices belong to different clusters C_1 and C_2 then an edge is formed between the clusters.

Algorithm 5.8.3: CLUSTER EDGE ANONYMIZATION(V_A, V_C, E)

```
Apply clustering algorithm to form $V_C$ number of clusters
for each edge $E_{ij}\epsilon$ E and between the vertices $V_i$ and $V_j$
  { Locate the respective clusters $C_1$ and $C_2$ for vertices $V_i$ and $V_j$
  { If $C_1$ and $C_2\epsilon$ different clusters then add the edge $E_{ij}$ between $C_1$ and $C_2$
Release the final graph $Efinal$
```

Removed edges: This type of edge removal algorithm intends to randomly remove the edges from the graph. The percentage of the edges that have to be removed goes as an input and the edges are removed accordingly. Unlike the *intact edges* algorithm the *removed edges* algorithm removes both the sensitive and the non-sensitive edges randomly from the graph.

Algorithm 5.8.4: REMOVED EDGES ALGORITHM (E, Per)

```
EdgeRemoved = Per * count(E)
for edgecount = 1 to EdgeRemoved
  { for each edge $E_{ij}\epsilon$ E picked randomly
  { $E_{final} = E - E_{ij}$
Release $E_{final}$
```

5.8.2 Measuring Privacy and Utility

The lesser the accuracy with which the attacker would be able to infer the private attributes the higher would be the privacy of the data set against inference. We quantify how privacy preserving a data set is, using Equation 5.8.1.

$$\mathit{Privacy} = (1 - \mathit{AC}) \tag{5.8.1}$$

Here *AC* is the accuracy of inference with which the attacker could correctly infer the hidden or sensitive attribute. If the accuracy of inference is 0% then the data set is completely privacy preserving against the inference attacks and if the data set has an accuracy of 100% then the privacy preservation factor is absolutely nil.

We measure *utility* by comparing the structure of the anonymized graph with the actual graph. The more the anonymized graph is similar to the actual graph the lesser perturbations and modifications it has gone through. In order to measure utility we calculate the structural similarity between the actual and the perturbed graph. The structural similarity can be measured by comparing the neighborhoods of all nodes for the two graphs to be compared. Let G_1 = (V, E_1) and G_2 = (V, E_2) be the two graphs having a common set of vertices and different edge sets E_1 and E_2. Then the similarity of G_1 an G_2 can be calculated by summing the similarity of each neighborhood as follows:

$$Utility = Similarity(G_1, G_2) = \frac{1}{|V|} \sum_{v \epsilon V} \frac{|N_{G_1} \bigcap N_{G_2}|}{|N_{G_1} \bigcup N_{G_2}|} \tag{5.8.2}$$

In Equation 5.8.2 utility is measured by analyzing the similarity between the neighborhoods of each node of the two graphs G_1 and G_2 where G_1 is the actual graph G_2 is the perturbed graph. N_{G_1} and N_{G_2} are the neighborhoods of the graphs G_1 and G_2 respectively. The value of *utility* will vary between 0 and 1, where 1 denotes structural identity i.e. the highest utility and 0 denotes a completely different structure i.e. the lowest utility. In the next section we will be discussing the proposed *trade-off* algorithm using which the data holders would be able to decide on an appropriate mechanism that could guarantee maximum utility with minimum privacy loss.

5.9 The Proposed Utility Privacy Trade-off Algorithm

We will now discuss the complete algorithm used in the proposed solution.

Algorithm 5.9.1: TRADE-OFF ALGORITHM($V, E, NM, EM, C, TList, UReq$)

```
for all nodes $V_i \epsilon$ V
Anonymize each node using the node anonymization mechanism NM
for each $EM_i \epsilon$ EM
  for each threshold $TList_i \epsilon$ TList
    Anonymize edges using the edge anonymization mechanism $EM_i$ and threshold $TList_i$
    Using the collective classification algorithm C, calculate the accuracy of inference
    Privacy = 1 - Accuracy

    $N_{G_1}$ = Neighborhood of all the vertices in graph $G_1$
    $N_{G_2}$ = neighborhood of all the vertices in graph $G_2$
    Find out the similarity between the two graphs $G_1$ and $G_2$
    Utility = Similarity ($G_1, G_2$)
    Plot the utility, privacy pair on the graph
Determine the points on the efficient frontier
Output the best privacy preserving EM for a given value of UReq
```

In the proposed Algorithm 5.9.1 *V* and *E* are the set of nodes and edges in the actual graph. V_A and E_A are the set of nodes and the set of edges in the anonymized graph. *NM* and *EM* are the mechanisms using which the nodes and edges in the actual graph are anonymized respectively. *C* is any collective classification algorithm and *TList* is the list of thresholds. At first all the nodes are anonymized using any standard node anonymization algorithm say *NM*. For every edge anonymization algorithm and a specific threshold value from the list of *TList* the edge set *E* is modified using an edge anonymization algorithm *EM*. Using the *collective classification* algorithm *C* the accuracy with which an attacker could infer the sensitive attributes is calculated. From equation 5.8.1 privacy against an inference algorithm could be calculated. Utility for the same modified graph can be calculated by measuring the similarity between the two graphs. This is obtained by comparing the neighborhoods of every vertex of the graph G_1 with the corresponding vertex in the graph G_2. For every edge anonymization algorithm *EM* and each value of threshold *TList* we obtain a pair of *utility-privacy* values. These values are plotted and for a given value of utility *UReq* the algorithm uses the concept of *efficient frontier* and returns the best edge anonymization algorithm that gives the maximum privacy protection.

5.10 Results for the Utility Privacy Trade-off Algorithm

Table 5.6 represents the utility privacy value pairs. Here $\{U_1, P_1\}$, $\{U_2, P_2\}$, $\{U_3, P_3\}$, $\{U_4, P_4\}$, $\{U_5, P_5\}$ are the utility and privacy values when 10%, 30%, 50%, 70% and 90% of all the edges belonging to the edge set were being modified respectively. We have experimented the *trade-off algorithm* on the *CoRA* data set which we have used earlier for implementation of the *partial edge set removal* algorithm. We divided the range of threshold i.e. (0 - 100)% into 5 intervals of 10%, 30%, 50%, 70% and 90% respectively. The data sets were perturbed using all the four edge anonymization algorithms discussed above. After plotting these *utility-privacy* value pairs on the graph we obtain four curves for

Table 5.6: Utility and privacy value pair for different edge anonymization algorithms at different thresholds

Algorithm	U_1	P_1	U_2	P_2	U_3	P_3	U_4	P_4	U_5	P_5
Intact Edge	.73	.1949	.59	.2023	.43	.2771	.35	.2824	.25	.2995
Partial Edge Set	.75	.2121	.68	.2689	.52	.2923	.33	.3385	.25	.4102
Cluster Edge Anonymization	.55	.1413	.42	.1778	.37	.2929	.32	.3723	.21	.4344
Removed Edges	.85	.1015	.64	.1321	.35	.2774	.28	.3491	14.	.4005

each edge anonymization algorithm respectively. In Figure 5.8 we show the plot between the values of utility and privacy.

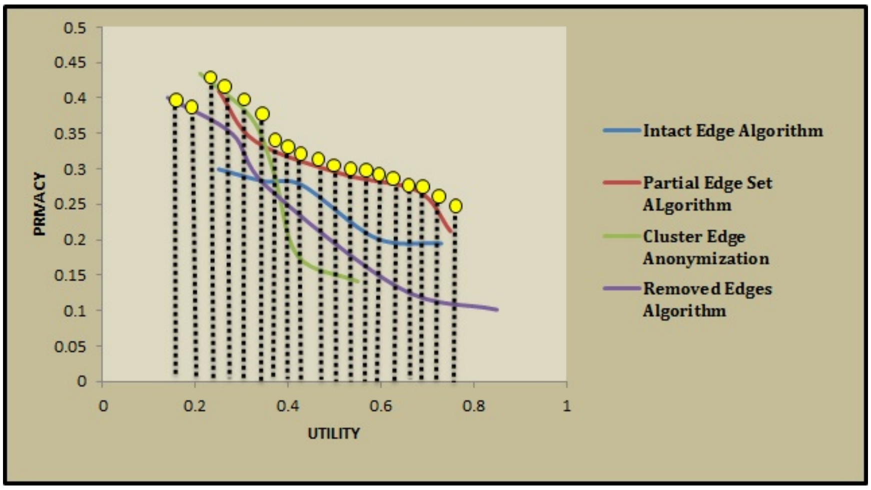


Figure 5.8: The utility privacy trade-off curve for fixed utility

We plot the utility on the *x axis* and privacy on the *y axis* to make a *privacy-utility trade-off graph*. Each graph will have curves equal to the number of anonymization mechanisms that have been followed in the experiment. The circles indicate the point on the efficient frontier. These points have the maximum privacy that one could get for a particular fixed value of utility. For example for a utility value of 0.4 the maximum privacy could be achieved using the *partial edge set algorithm* whereas other algorithms like the *intact edge, cluster edge anonymization* and *removed edge algorithm* would under-perform.

5.10.1 Results on the Facebook Dataset

We extracted Facebook data sets for 15000 user profiles. The results are shown in Figure 5.9. We anonymized the edges of the graph using the four algorithms mentioned above. Our aim was to model the data set to predict the relationship status and the users' year of birth. Using network classification we modeled our data set and applied the edge anonymizations as mentioned above and predicted the *relationship status* and the *year of birth* from the

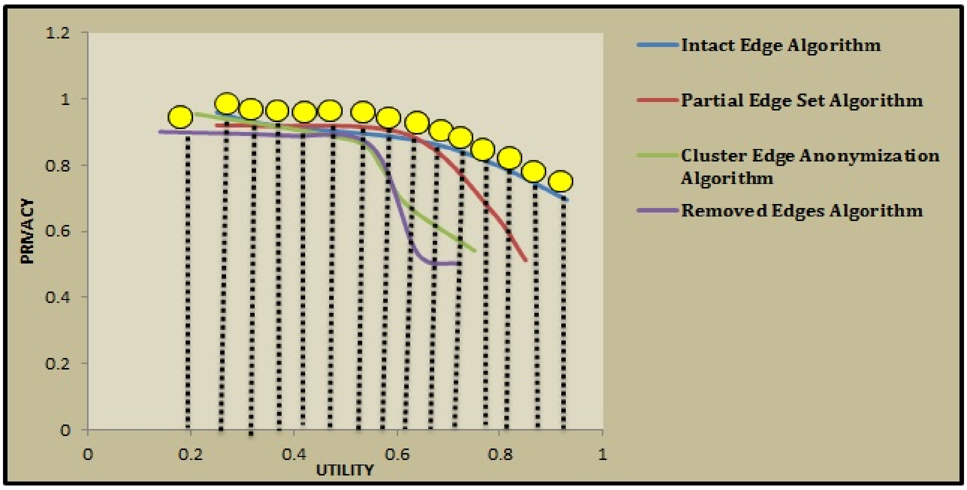


Figure 5.9: The utility privacy trade-off curve for the Facebook dataset

perturbed graphs. In Figure 5.9 we show the results obtained for the *relationship status*. Therefore, before the data holder releases the data set using various anonymization algorithms they should decide on the scheme which would give them the maximum utility for a given value of privacy or maximum privacy for a given value of utility. Overall, on an average the performance of the partial edge set anonymization algorithm is better than the other three algorithms in place.

5.11 Summary

In this chapter, we proposed a solution to prevent the released OSN sensitive attributes against the inference attacks. We labeled the original graph and constructed a new labeled graph where every edge had a label. These labels showed the association between the nodes which could either be private or public. Our algorithm changes the labeled graph to the final graph that should be released to the third party by partially removing certain set of edges from the original graph. The edges removed were the ones that showed the relationship between a private and a public node. We proved that this algorithm ensures the decrease in the accuracy with which the adversary would label the unknown nodes. We have also measured the utility loss which comes by perturbing the original graph by considering different structural metrics. The major challenge for the data holders is to follow such an anonymization algorithm that could guarantee them minimum privacy loss and maximum utility. We had also worked out and solved the problem of deciding the trade-off between privacy and utility in an OSN graph using the concept of portfolio theory. The approaches discussed above would help the data providers decide the best edge anonymization scheme which could provide the maximum utility and minimum privacy loss and would help the users keep their trust intact with the data holders. The implementation of *partial edge set algorithm* and *utility-privacy trade-off algorithm* is published in [Pub2] and [Pub10] respectively.

Chapter 6

Conclusions and Future Scope

In this research work we have analyzed data privacy in OSN from the perspectives of the user, their online connections and the service providers. We have identified the loopholes in the existing privacy solutions specific to the areas of measuring the OSN users' data privacy, enabling selective sharing of sensitive OSN data and protecting sensitive OSN data from inference attacks. To overcome the identified research gaps, we proposed a OSN data privacy solution which comprises effective privacy preserving frameworks and efficient privacy enhancing algorithms. Using this solution, we can reduce the chances of OSN data breach to a great extent. We summarize the important deductions drawn from our work in the next section.

6.1 Summary of Deductions

Privacy is an abstract term hence, measuring users' data privacy in Online Social Networks was identified as a challenge. Privacy Quotient (PQ) which is a score given to the users' on the basis of their sharing behavior was used to measure data privacy in OSNs. At first we carried out an extensive survey with a total of 600 respondents and used the CTT model to evaluate PQ for each subject. We also used the IRT model to implement a framework to measure the privacy strength of the users' profile for static profile items. We proposed and built a privacy settings recommender system which would compare the users privacy quotient with their connections and also would recommend appropriate privacy settings for the dynamic profile items.

Users' data privacy also depends on their connections and the connections' privacy quotient hence, mechanisms for enabling selective sharing of the sensitive data were explored. We proved that the privacy of a node not only depends on its sharing behavior but also on its topology in the network. A relationship between centrality and private information spread was drawn and it was observed that the nodes having low privacy quotient, high betweenness and high closeness centrality cause the maximum private information spread in the network. An information sharing algorithm was proposed which could prevent sensitive information from passing through nodes which have a high probability of spreading information. The association between privacy and trust was closely studied and a trust enhancing model was proposed to refine the existing trusted community of strong ties after considering their online sharing behaviour.

Data privacy from the perspective of inference attacks was studied thoroughly and a partial edge set anonymization algorithm was proposed that could reduce the accuracy with which the attackers could infer the sensitive attributes. Data perturbation results in the loss of utility. Hence, we measured the utility loss in the released dataset considering different structural metrics. We proposed a utility privacy trade-off algorithm for the data sets anonymized using the edge anonymization algorithms. We used portfolio theory and the concept of efficient frontier to determine an appropriate trade-off such that the released datasets would have maximum utility and minimum privacy loss.

6.2 Future Scope of the Work

With this work we hope to motivate advanced research in the field of data privacy specifically in the area of measuring user privacy, enabling selective sharing of sensitive data and protecting sensitive data from inference attacks in OSNs. In future we would like to resolve the issue of detecting privacy leaks in the unstructured data as it is usually sparse, computationally hard and more vulnerable to attacks. As an extension to our work the proposed recommender system could be made more robust where it can filter out the sensitive contents like users' personal photos and videos from the general content before computing the privacy quotient. Privacy and network topology studies are in the nascent stage and a lot could be explored on similar lines. The graph structural property like assortativity and conductance could be explored in detail. In our work, we have explored the problem of measuring direct trust for privacy whereas the effect of indirect and bootstrap trust on privacy could also be explored. In the area of preventing sensitive data from inference attacks we would like to address the issue where more than one attributes could be considered sensitive. Preserving privacy in OSN is an important issue. With our work, we envision a number of privacy preserving applications which could be built using the proposed algorithms to make Online Social Networks a better platform for sharing and exchanging information.

List of Publications

Peer Reviewed Journals

[Pub1] Agrima Srivastava and G Geethakumari (2014), “Privacy Landscape in Online Social Networks”, Published in the *International Journal of Trust Management in Computing and Communications, Inderscience Publishers*, ISSN online: 2048-8386, ISSN print: 2048-8378, 2014.

[Pub2] Agrima Srivastava and G Geethakumari (2015), “Privacy preserving solution to prevent Inference Attacks in Online Social Networks”, *International Journal of Computational Science and Engineering, Inderscience publishers*, 2015.

[Pub3] KP Krishna Kumar, Agrima Srivastava and G Geethakumari (2014), “A Psychometric Analysis of Information Propagation in Online Social Networks using Latent Trait Theory”, Published in the Journal of ‘Computing’, Springer publishers, 2014.

International Conferences

[Pub4] Agrima Srivastava and G Geethakumari (2013), “Measuring Privacy Leaks in Online Social Networks”, *Proceedings of the IEEE ICACCI-2013: Proceedings of the International Symposium on Women in Computing and Informatics (WCI-2013)*, August 22 - 25, Mysore, India, 2013, pp 95-100.

[Pub5] Agrima Srivastava and G Geethakumari (2013), “A Framework to Customize Privacy Settings of Online Social Network Users”, *Proceedings of the IEEE International*

Conference on Recent Advances in Intelligent Computational Systems (RAICS) 2013, December 19 - 21, 2013, India, pp 187-192.

[Pub6] Agrima Srivastava and G Geethakumari (2014), "A Privacy Settings Recommender System for Online Social Networks", *Proceedings of the IEEE International Conference on Recent Advances and Innovations in Engineering - (ICRAIE-2014)*, May 09-11, 2014, India, pp 1-6.

[Pub7] Agrima Srivastava, K P Krishna Kumar and G Geethakumari (2014), "Preserving privacy in online social networks using the graph structural analysis", *Proceedings of the International Conference on Advances in Computing, Communications and Information Science, ACCIS - 14*, June 26 - 28, 2014, India. Proceedings in Elsevier India, pp 219-228.

[Pub8] Agrima Srivastava and G Geethakumari (2014), "Quantifying direct trust for private information sharing in an Online Social Network", *Proceedings of the International Symposium on Intelligent Informatics (ISI14)*, September 24-27, 2014, India; Proceedings in the Journal Advances in Intelligent and Soft Computing (Springer) Series, pp 21-30.

[Pub9] KP Krishna Kumar, Agrima Srivastava and G. Geethakumari (2014). "Preventing Disinformation Cascades using Behavioural Trust in Online Social Networks", *Proceedings of the International Conference on Advances in Computing, Communications and Information Science, ACCIS - 14* ISBN: 9789351072478, pp 113-123.

[Pub10] Agrima Srivastava and G Geethakumari (2015), "An efficient privacy utility approach for Online Social Network data publishing", accepted at the *IEEE 12th International Conference INDICON*, December 17-20,2015, India.

[Pub11] Agrima Srivastava and G Geethakumari (2015), "A framework for improving privacy strength of Online Social Network user profile", accepted at the *Grace Hopper Celebration of Women in Computing, India* 2015, December 2-4, 2015.

Bibliography

[1] Michalis Faloutsos, Thomas Karagiannis, and Seung-Hyun Moon. Online social networks. *Network, IEEE*, 24(5):4–5, 2010.

[2] Laura Garton, Caroline Haythornthwaite, and Barry Wellman. Studying online social networks. *Journal of Computer-Mediated Communication*, 3(1):0–0, 1997.

[3] Nicole B Ellison. Social network sites: Definition, history, and scholarship. *Journal of Computer-Mediated Communication, Wiley Online Library, volume*, 13(1):210–230, 2007.

[4] John A Barnes. Graph theory and social networks: A technical comment on connectedness and connectivity. *Sociology, Sage Publications, volume*, 3(2):215–232, 1969.

[5] Ravi Kumar, Jasmine Novak, and Andrew Tomkins. Structure and evolution of online social networks. In *Link mining: models, algorithms, and applications*, pages 337–357. Springer, 2010.

[6] Michael Beye, Arjan Jeckmans, Zekeriya Erkin, Pieter Hartel, Reginald Lagendijk, and Qiang Tang. Literature overview-privacy in online social networks. 2010.

[7] Abdullah Al Hasib. Threats of online social networks. *IJCSNS International Journal of Computer Science and Network Security*, 9(11):288–93, 2009.

[8] Lei Jin, Hassan Takabi, and James BD Joshi. Towards active detection of identity clone attacks on online social networks. In *Proceedings of the first ACM conference on Data and application security and privacy*, pages 27–38. ACM, 2011.

[9] Chi Zhang, Jinyuan Sun, Xiaoyan Zhu, and Yuguang Fang. Privacy and security for online social networks: challenges and opportunities. *Network, IEEE, volume*, 24(4):13–18, 2010.

[10] Hongyu Gao, Jun Hu, Tuo Huang, Jingnan Wang, and Yan Chen. Security issues in online social networks. *Internet Computing, IEEE, volume*, 15(4):56–63, 2011.

[11] Judith DeCew. Privacy. In Edward N. Zalta, editor, *The Stanford Encyclopedia of Philosophy*, 2012.

[12] Gail-Joon Ahn, Mohamed Shehab, and Anna Squicciarini. Security and privacy in social networks. *Internet Computing, IEEE*, 15(3):10–12, 2011.

[13] Prateek Joshi and CC Jay Kuo. Security and privacy in online social networks: A survey. In *Multimedia and Expo (ICME), 2011 IEEE International Conference on*, pages 1–6. IEEE, 2011.

[14] Pritam Gundecha and Huan Liu. Mining social media: A brief introduction.

[15] Alessandro Acquisti and Ralph Gross. Predicting social security numbers from public data. *Proceedings of the National academy of sciences, National Acad Sciences, volume*, 106(27):10975–10980, 2009.

[16] George Duncan and Diane Lambert. The risk of disclosure for microdata. *Journal of Business & Economic Statistics*, 7(2):207–217, 1989.

[17] Kun Liu, Kamalika Das, Tyrone Grandison, and Hillol Kargupta. Privacy-preserving data analysis on graphs and social networks. Next Generation Data Mining. CRC Press, 2008.

[18] Joseph Bonneau and Sören Preibusch. The privacy jungle: On the market for data protection in social networks. In *Economics of information security and privacy*, pages 121–167. Springer, 2010.

[19] Joshua Fogel and Elham Nehmad. Internet social network communities: Risk taking, trust, and privacy concerns. *Computers in human behavior*, 25(1):153–160, 2009.

[20] Alejandro Portes. Social capital: Its origins and applications in modern sociology. *LESSER, Eric L. Knowledge and Social Capital. Boston: Butterworth-Heinemann*, pages 43–67, 2000.

[21] Gergely Biczók and Pern Hui Chia. Interdependent privacy: Let me share your data. In *Financial Cryptography and Data Security*, pages 338–353. Springer, 2013.

[22] David J Houghton and Adam N Joinson. Privacy, social network sites, and social relations. *Journal of Technology in Human Services*, 28(1-2):74–94, 2010.

[23] Hanna Krasnova, Oliver Günther, Sarah Spiekermann, and Ksenia Koroleva. Privacy concerns and identity in online social networks. *Identity in the Information Society*, 2(1):39–63, 2009.

[24] Lars Backstrom, Cynthia Dwork, and Jon Kleinberg. Wherefore art thou r3579x?: anonymized social networks, hidden patterns, and structural steganography. In *Proceedings of the 16th international conference on World Wide Web*, pages 181–190. ACM, 2007.

[25] Sabine Trepte and Leonard Reinecke. The social web as a shelter for privacy and authentic living. In *Privacy Online*, pages 61–73. Springer, 2011.

[26] Balachander Krishnamurthy and Craig E Wills. On the leakage of personally identifiable information via online social networks. In *Proceedings of the 2nd ACM workshop on Online social networks*, pages 7–12. ACM, 2009.

[27] Mary J Culnan and Pamela K Armstrong. Information privacy concerns, procedural fairness, and impersonal trust: An empirical investigation. *Organization science*, 10(1):104–115, 1999.

[28] Daniel J Solove. Conceptualizing privacy. *California Law Review, JSTOR, volume*, pages 1087–1155, 2002.

[29] Samuel D Warren and Louis D Brandeis. The right to privacy. *Harvard law review, JSTOR, volume*, 4(5):193–220, 1890.

[30] Alan F Westin and Louis Blom-Cooper. *Privacy and freedom*, volume 67. Atheneum New York, 1970.

[31] Ruth Gavison. Privacy and the limits of law. *Yale law journal*, pages 421–471, 1980.

[32] Irwin Altman. Privacy: A conceptual analysis. *Environment and behavior, ERIC, volume*, 8(1):7–29, 1976.

[33] Irwin Altman, Anne Vinsel, and Barbara B Brown. Dialectic conceptions in social psychology: An application to social penetration and privacy regulation. *Advances in experimental social psychology, volume*, 14:107–160, 1981.

[34] Leysia Palen and Paul Dourish. Unpacking privacy for a networked world. In *Proceedings of the SIGCHI conference on Human factors in computing systems*, pages 129–136. ACM, 2003.

[35] Daniel J Solove. A taxonomy of privacy. *University of Pennsylvania Law Review, JSTOR*, pages 477–564, 2006.

[36] Claudia Diaz and Seda Gürses. Understanding the landscape of privacy technologies1, 2012.

[37] Zizi Papacharissi and Paige L Gibson. Fifteen minutes of privacy: Privacy, sociality, and publicity on social network sites. In *Privacy Online*, pages 75–89. Springer, 2011.

[38] Mani R Subramani and Balaji Rajagopalan. Knowledge-sharing and influence in online social networks via viral marketing. *Communications of the ACM*, 46(12):300–307, 2003.

[39] Ralph Gross and Alessandro Acquisti. Information revelation and privacy in online social networks. In *Proceedings of the 2005 ACM workshop on Privacy in the electronic society*, pages 71–80. ACM, 2005.

[40] Yabing Liu, Krishna P Gummadi, Balachander Krishnamurthy, and Alan Mislove. Analyzing facebook privacy settings: User expectations vs. reality. In *Proceedings of the 2011 ACM SIGCOMM conference on Internet measurement conference*, pages 61–70. ACM, 2011.

[41] Fred Stutzman and Jacob Kramer-Duffield. Friends only: examining a privacy-enhancing behavior in facebook. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pages 1553–1562. ACM, 2010.

[42] Sandra Sporbert Petronio. *Boundaries of privacy*, volume 2002. State University of New York Press Albany, NY, 2002.

[43] Yang Wang, Gregory Norcie, Saranga Komanduri, Alessandro Acquisti, Pedro Giovanni Leon, and Lorrie Faith Cranor. I regretted the minute i pressed share: A qualitative study of regrets on facebook. In *Proceedings of the Seventh Symposium on Usable Privacy and Security*, page 10. ACM, 2011.

[44] Maritza Johnson, Serge Egelman, and Steven M Bellovin. Facebook and privacy: it's complicated. In *Proceedings of the Eighth Symposium on Usable Privacy and Security*, page 9. ACM, 2012.

[45] Manya Sleeper, Rebecca Balebako, Sauvik Das, Amber Lynn McConahy, Jason Wiese, and Lorrie Faith Cranor. The post that wasn't: exploring self-censorship on facebook. In *Proceedings of the 2013 conference on Computer supported cooperative work*, pages 793–802. ACM, 2013.

[46] Alex Braunstein, Laura Granka, and Jessica Staddon. Indirect content privacy surveys: measuring privacy without asking about it. In *Proceedings of the Seventh Symposium on Usable Privacy and Security*, page 15. ACM, 2011.

[47] Lujun Fang and Kristen LeFevre. Privacy wizards for social networking sites. In *Proceedings of the 19th international conference on World wide web*, pages 351–360. ACM, 2010.

[48] Kambiz Ghazinour, Stan Matwin, and Marina Sokolova. Yourprivacyprotector: Arecommender system for privacy settings in social networks. *International Journal of Security*, 2013.

[49] Alessandra Mazzia, Kristen LeFevre, and Eytan Adar. The pviz comprehension tool for social network privacy settings. In *Proceedings of the Eighth Symposium on Usable Privacy and Security*, page 13. ACM, 2012.

[50] Ozgur Kafali, Akin Gunay, and Pinar Yolum. Protoss: A run time tool for detecting privacy violations in online social networks. In *Advances in Social Networks Analysis and Mining (ASONAM), International Conference on*, pages 429–433. IEEE, 2012.

[51] Kun Liu and Evimaria Terzi. A framework for computing the privacy scores of users in online social networks. In *Data Mining, 2009. ICDM'09. Ninth IEEE International Conference on*, pages 288–297. IEEE, 2009.

[52] Michelle Girvan and Mark EJ Newman. Community structure in social and biological networks. *Proceedings of the national academy of sciences*, 99(12):7821–7826, 2002.

[53] Leucio Antonio Cutillo, Refik Molva, and Melek Onen. Analysis of privacy in online social networks from the graph theory perspective. In *Global Telecommunications Conference (GLOBECOM)*, pages 1–5. IEEE, 2011.

[54] Hakan Yildiz and Christopher Kruegel. Detecting social cliques for automated privacy control in online social networks. In *Pervasive Computing and Communications Workshops (PERCOM Workshops), 2012 IEEE International Conference on*, pages 353–359. IEEE, 2012.

[55] Pedro G Lind, Luciano R da Silva, José S Andrade Jr, and Hans J Herrmann. Spreading gossip in social networks. *Physical Review E, APS, volume*, 76(3):036117, 2007.

[56] Allison K Shaw, Milena Tsvetkova, and Roozbeh Daneshvar. The effect of gossip on social networks. *Complexity*, 16(4):39–47, 2011.

[57] Donovan Artz and Yolanda Gil. A survey of trust in computer science and the semantic web. *Web Semantics: Science, Services and Agents on the World Wide Web*, 5(2):58–71, 2007.

[58] Vincent Buskens. *Social networks and trust*, volume 30. Springer, 2002.

[59] Sibel Adali, Robert Escriva, Mark K Goldberg, Mykola Hayvanovych, Malik Magdon-Ismail, Boleslaw K Szymanski, William A Wallace, and Gregory Williams. Measuring behavioral trust in social networks. In *Intelligence and Security Informatics (ISI)*, pages 150–152. IEEE, 2010.

[60] Sacha Trifunovic, Franck Legendre, and Carlos Anastasiades. Social trust in opportunistic networks. In *INFOCOM IEEE Conference on Computer Communications Workshops, 2010*, pages 1–6. IEEE, 2010.

[61] Barbara Carminati, Elena Ferrari, Sandro Morasca, and Davide Taibi. A probability-based approach to modeling the risk of unauthorized propagation of information in on-line social networks. In *Proceedings of the first ACM conference on Data and application security and privacy*, pages 51–62. ACM, 2011.

[62] Na Li, Maryam Najafian Razavi, and Denis Gillet. Trust-aware privacy control for social media. In *CHI Extended Abstracts on Human Factors in Computing Systems*, pages 1597–1602. ACM, 2011.

[63] Michael Hay, Gerome Miklau, David Jensen, Philipp Weis, and Siddharth Srivastava. Anonymizing social networks. *Computer Science Department Faculty Publication Series*, page 180, 2007.

[64] Latanya Sweeney. k-anonymity: A model for protecting privacy. *International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems, World Scientific, volume*, 10(05):557–570, 2002.

[65] Ashwin Machanavajjhala, Daniel Kifer, Johannes Gehrke, and Muthuramakrishnan Venkitasubramaniam. l-diversity: Privacy beyond k-anonymity. *ACM Transactions on Knowledge Discovery from Data (TKDD), ACM, volume*, 1(1):3, 2007.

[66] Ninghui Li, Tiancheng Li, and Suresh Venkatasubramanian. t-closeness: Privacy beyond k-anonymity and l-diversity. In *23rd International Conference on Data ENgineering*, pages 106–115. IEEE, 2007.

[67] Kun Liu and Evimaria Terzi. Towards identity anonymization on graphs. In *Proceedings of the 2008 ACM SIGMOD international conference on Management of data*, pages 93–106. ACM, 2008.

[68] Bin Zhou and Jian Pei. Preserving privacy in social networks against neighborhood attacks. In *Data Engineering, 2008. ICDE 2008. IEEE 24th International Conference on*, pages 506–515. IEEE, 2008.

[69] Lei Zou, Lei Chen, and M Tamer Özsu. K-automorphism: A general framework for privacy preserving network publication. *Proceedings of the VLDB Endowment, VLDB Endowment, volume*, 2(1):946–957, 2009.

[70] Wentao Wu, Yanghua Xiao, Wei Wang, Zhenying He, and Zhihui Wang. k-symmetry model for identity anonymization in social networks. In *Proceedings of the 13th international conference on extending database technology*, pages 111–122. ACM, 2010.

[71] James Cheng, Ada Wai-chee Fu, and Jia Liu. K-isomorphism: privacy preserving network publication against structural attacks. In *Proceedings of the 2010 international conference on Management of data*, pages 459–470. ACM, 2010.

[72] Francesco Bonchi, Aristides Gionis, and Tamir Tassa. Identity obfuscation in graphs through the information theoretic lens. In *Data Engineering (ICDE), 2011 IEEE 27th International Conference on*, pages 924–935. IEEE, 2011.

[73] Michael Hay, Gerome Miklau, David Jensen, Don Towsley, and Chao Li. Resisting structural re-identification in anonymized social networks. *The VLDB JournalThe International Journal on Very Large Data Bases, Springer-Verlag New York, Inc., volume*, 19(6):797–823, 2010.

[74] Rakesh Agrawal and Ramakrishnan Srikant. Privacy-preserving data mining. In *ACM Sigmod Record, ACM, volume*, volume 29, pages 439–450. ACM, 2000.

[75] Johannes Gehrke, Edward Lui, and Rafael Pass. Towards privacy for social networks: a zero-knowledge based definition of privacy. In *Theory of Cryptography*, pages 432–449. Springer, 2011.

[76] Aditya Krishna Menon and Charles Elkan. Predicting labels for dyadic data. *Data Mining and Knowledge Discovery*, 21(2):327–343, 2010.

[77] Jianming He, Wesley W Chu, and Zhenyu Victor Liu. Inferring privacy information from social networks. In *Intelligence and Security Informatics*, pages 154–165. Springer, 2006.

[78] Elena Zheleva and Lise Getoor. To join or not to join: the illusion of privacy in social networks with mixed public and private user profiles. In *Proceedings of the 18th international conference on World wide web*, pages 531–540. ACM, 2009.

[79] Raymond Heatherly, Murat Kantarcioglu, and Bhavani Thuraisingham. Preventing private information inference attacks on social networks. *IEEE Transactions on Knowledge and Data Engineering, volume*, 25(8):1849–1862, 2013.

[80] Alexandros Labrinidis. Privacy-preserving data publishing. *Foundations and Trends in Databases*, 2(3):169–266, 2009.

[81] Tiancheng Li and Ninghui Li. On the tradeoff between privacy and utility in data publishing. In *Proceedings of the 15th ACM SIGKDD international conference on Knowledge discovery and data mining*, pages 517–526. ACM, 2009.

[82] Frederic Stutzman. An evaluation of identity-sharing behavior in social network communities. *Journal of the International Digital Media and Arts Association*, 3(1):10–18, 2006.

[83] Ronald Jay Cohen and Mark E Swerdlik. *Psychological testing and assessment: An introduction to tests and measurement* . McGraw-Hill, 2002.

[84] John Rust and Susan Golombok. *Modern psychometrics: The science of psychological assessment*. Routledge, 2014.

[85] Ronald Jay Cohen, Mark E Swerdlik, and Suzanne M Phillips. *Psychological testing and assessment: An introduction to tests and measurement* . Mayfield Publishing Co, 1996.

[86] Kevin Lewis, Jason Kaufman, and Nicholas Christakis. The taste for privacy: An analysis of college student privacy settings in an online social network. *Journal of Computer-Mediated Communication*, 14(1):79–100, 2008.

[87] Sivakumar Alagumalai and David D Curtis. *Classical test theory*. Springer, 2005.

[88] Ross E Traub. Classical test theory in historical perspective. *Educational Measurement: Issues and Practice, Wiley Online Library, volume*, 16(4):8–14, 1997.

[89] Melvin R Novick. The axioms and principal results of classical test theory. *Journal of Mathematical Psychology*, 3(1):1–18, 1966.

[90] Wing S Chow and Lai Sheung Chan. Social network, social trust and shared goals in organizational knowledge sharing. *Information & Management*, 45(7):458–465, 2008.

[91] Andrew Martin Turpin and Jeffrey B Katz. System and method for implementing advertising in an online social network, January 30 2008. US Patent App. 12/011,880.

[92] Xitao Fan. Item response theory and classical test theory: An empirical comparison of their item/person statistics. *Educational and psychological measurement, Sage Publications, volume*, 58(3):357–381, 1998.

[93] Robert J Harvey and Allen L Hammer. Item response theory. *The Counseling Psychologist*, 27(3):353–383, 1999.

[94] Shuhua Hu. Akaike information criterion. *Center for Research in Scientific Computation*, 2007.

[95] David L Weakliem. A critique of the bayesian information criterion for model selection. *Sociological Methods & Research*, 27(3):359–397, 1999.

[96] Mark Ackerman, Trevor Darrell, and Daniel J Weitzner. Privacy in context. *Human–Computer Interaction, Taylor & Francis, volume*, 16(2-4):167–176, 2001.

[97] Bradley N Miller, Istvan Albert, Shyong K Lam, Joseph A Konstan, and John Riedl. Movielens unplugged: experiences with an occasionally connected recommender system. In *Proceedings of the 8th international conference on Intelligent user interfaces*, pages 263–266. ACM, 2003.

[98] Paul Resnick and Hal R Varian. Recommender systems. *Communications of the ACM, ACM, volume*, 40(3):56–58, 1997.

[99] J Ben Schafer, Joseph Konstan, and John Riedl. Recommender systems in e-commerce. In *Proceedings of the 1st ACM conference on Electronic commerce*, pages 158–166. ACM, 1999.

[100] Dan Cosley, Shyong K Lam, Istvan Albert, Joseph A Konstan, and John Riedl. Is seeing believing?: how recommender system interfaces affect users' opinions. In *Proceedings of the SIGCHI conference on Human factors in computing systems*, pages 585–592. ACM, 2003.

[101] Miller McPherson, Lynn Smith-Lovin, and James M Cook. Birds of a feather: Homophily in social networks. *Annual review of sociology*, pages 415–444, 2001.

[102] Mariea Grubbs Hoy and George Milne. Gender differences in privacy-related measures for young adult facebook users. *Journal of Interactive Advertising, The Canadian Press, volume*, 10(2):28–45, 2010.

[103] Janine Nahapiet and Sumantra Ghoshal. Social capital, intellectual capital, and the organizational advantage. *Academy of management review, Academy of Management, volume*, 23(2):242–266, 1998.

[104] Xiao-Yong Li and Xiao-Lin Gui. Research on dynamic trust model for large scale distributed environment. *Ruan Jian Xue Bao(Journal of Software)*, 18(6):1510–1521, 2007.

[105] Sarbjeet Singh and Seema Bawa. A privacy, trust and policy based authorization framework for services in distributed environments. *International Journal of Computer Science, volume*, 2(2), 2007.

[106] John Adrian Bondy and Uppaluri Siva Ramachandra Murty. *Graph theory with applications*, volume 290. Macmillan London, 1976.

[107] Stephen P Borgatti. Centrality and network flow. *Social networks, Elsevier, volume*, 27(1):55–71, 2005.

[108] Duanbing Chen, Linyuan Lü, Ming-Sheng Shang, Yi-Cheng Zhang, and Tao Zhou. Identifying influential nodes in complex networks. *Physica A: Statistical Mechanics and its Applications, Elsevier, volume*, 391(4):1777–1787, 2012.

[109] Linton C Freeman. Centrality in social networks conceptual clarification. *Social networks, Elsevier, volume*, 1(3):215–239, 1979.

[110] Mark EJ Newman. Clustering and preferential attachment in growing networks. *Physical Review E, APS, volume*, 64(2):025102, 2001.

[111] David Kempe, Jon Kleinberg, and Éva Tardos. Maximizing the spread of influence through a social network. In *Proceedings of the ninth ACM SIGKDD international conference on Knowledge discovery and data mining*, pages 137–146. ACM, 2003.

[112] Elizabeth Sklar. Netlogo, a multi-agent simulation environment. *Artificial life, MIT Press, volume*, 13(3):303–311, 2007.

[113] Miriam J Metzger. Privacy, trust, and disclosure: Exploring barriers to electronic commerce. *Journal of Computer-Mediated Communication, Wiley Online Library, volume*, 9(4):00–00, 2004.

[114] Wanita Sherchan, Surya Nepal, and Cecile Paris. A survey of trust in social networks. *ACM Computing Surveys (CSUR), ACM, volume*, 45(4):47, 2013.

[115] Yanjie Bian. Bringing strong ties back in: Indirect ties, network bridges, and job searches in china. *American sociological review*, pages 366–385, 1997.

[116] Frank B Baker and Seock-Ho Kim. *Item response theory: Parameter estimation techniques*. CRC Press, 2004.

[117] Julian J McAuley and Jure Leskovec. Learning to discover social circles in ego networks. In *NIPS*, volume 272, pages 548–556, 2012.

[118] David J Martin, Daniel Kifer, Ashwin Machanavajjhala, Johannes Gehrke, and Joseph Y Halpern. Worst-case background knowledge for privacy-preserving data publishing. In *Data Engineering, 2007. ICDE 2007. IEEE 23rd International Conference on*, pages 126–135. IEEE, 2007.

[119] Ruilin Liu and Hui Wang. Privacy-preserving data publishing. In *Data Engineering Workshops (ICDEW), 2010 IEEE 26th International Conference on*, pages 305–308. IEEE, 2010.

[120] Smriti Bhagat, Graham Cormode, Balachander Krishnamurthy, and Divesh Srivastava. Class-based graph anonymization for social network data. *Proceedings of the VLDB Endowment, VLDB Endowment, volume*, 2(1):766–777, 2009.

[121] Aris Anagnostopoulos, Ravi Kumar, and Mohammad Mahdian. Influence and correlation in social networks. In *Proceedings of the 14th ACM SIGKDD international conference on Knowledge discovery and data mining*, pages 7–15. ACM, 2008.

[122] Prithviraj Sen, Galileo Namata, Mustafa Bilgic, Lise Getoor, Brian Galligher, and Tina Eliassi-Rad. Collective classification in network data. *AI magazine*, 29(3):93, 2008.

[123] Soumen Chakrabarti, Byron Dom, and Piotr Indyk. Enhanced hypertext categorization using hyperlinks. In *ACM SIGMOD Record*, volume 27, pages 307–318. ACM, 1998.

[124] Qing Lu and Lise Getoor. Link-based classification. In *ICML*, volume 3, pages 496–503, 2003.

[125] Stuart Geman and Donald Geman. Stochastic relaxation, gibbs distributions, and the bayesian restoration of images. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, (6):721–741, 1984.

[126] Sofus A Macskassy, Foster Provost, and Foster Provost. Netkit-srl: A toolkit for network learning and inference. In *Proceeding of the NAACSOS Conference*, 2005.

[127] Edwin J Elton, Martin J Gruber, Stephen J Brown, and William N Goetzmann. *Modern portfolio theory and investment analysis*. John Wiley & Sons, 2009.

[128] Harry M Markowitz. Foundations of portfolio theory. *The journal of finance*, 46(2):469–477, 1991.

[129] Elena Zheleva and Lise Getoor. Preserving the privacy of sensitive relationships in graph data. In *Privacy, security, and trust in KDD*, pages 153–171. Springer, 2008.

Glossary of terms used in the thesis

Classical Test Theory (CTT). Classical test theory is a part of related psychometric theory that predicts outcomes of psychological testing such as the difficulty of items or the ability of test-takers.

Cronbach's alpha. Cronbach's alpha is used as a lower bound estimate of the reliability of a psychometric test.

Diary Study. In Human-Computer Interaction, a diary study is a qualitative technique for collecting data on what users have done or experienced.

Dyadic Predictions. In dyadic prediction, the training set consists of pairs of objects called dyads, with associated labels. The task is to predict labels for unobserved dyads that do not appear in the training set.

F Measure. The traditional F-measure or balanced F-score is the harmonic mean of precision and recall.

Facebook. It is a popular free social networking website that allows registered users to create profiles, upload photos and video, send messages and keep in touch with friends, family and colleagues.

False Positive (FP). It measures the proportion of negatives that are wrongly identified as positives.

Inference Attacks. An Inference Attack is a data mining technique performed by analyzing data in order to illegitimately gain knowledge about a subject or database.

Interdependent Privacy. Privacy concerns arise along with data sharing. In such an intertwined setting, the privacy of individual users is bound to be affected by the decisions of others, and could be out of their own control. This gives rise to the phenomenon which we term as interdependent privacy.

Item Response Theory (IRT). In psychometrics, item response theory (IRT) also known as latent trait theory, strong true score theory, or modern mental test theory, is a paradigm for the design, analysis, and scoring of tests, questionnaires, and similar instruments measuring abilities, attitudes, or other variables.

Kappa. Cohen's kappa is a measure of the agreement between two raters who determine which category a finite number of subjects belong to whereby agreement due to chance is factored out.

MySpace. It is an online social network. MySpace caters to artists and bands, who enjoy the flexibility of creating an individual look for their page. MySpace allows users to friend each other and create groups.

Online Social Networks. An online community of people who are socializing with each other via a particular web site. It helps to connect socially or professionally with other people.

Personal Identifying Information (PII). Personally identifiable information (PII) is any data that could potentially identify a specific individual.

Preferential Attachment It is a process in which units of objects are distributed amongst individuals according to the units they already have.

Privacy Preserving Data Publishing (PPDP). Privacy-preserving data publishing (PPDP) provides methods and tools for publishing useful information while preserving data privacy.

Privacy Quotient (PQ). It is the score given to the OSN user by analyzing their sharing behaviors. It is the potential risk that is caused by the users participation in the network.

Privacy Settings. Privacy settings are controls available on many social networking and other web-sites that allow users to limit who can access their profile and what information visitors can see.

Profiles. Profiles are the information that you provide about yourself when signing up for a social networking site.

Psychometrics. Psychometrics is a field of study concerned with the theory and technique of psychological measurement.

Recommender System. An information filtering technology, commonly used on e-commerce web sites that uses a collaborative filtering to present information on items and products that are likely to be of interest to the reader.

Social Capital. Social capital is a concept used in business, non profits and other arenas that refers to the good will and positive reputation that flows to a person through his or her relationships with others in social networks.

Social Media. Social media refers to works of user-created video, audio, text or multimedia that are published and shared in a social environment, such as a blog, podcast, forum, wiki or video hosting site. More broadly, social media refers to any online technology that lets people publish, converse and share content online.

Social Security Number (SSN). In the United States, a Social Security number is a nine-digit number issued to U.S. citizens, permanent residents and temporary (working) residents.

Social Spam. Social spam is an unwanted spam content appearing on social networks and any website with user-generated content (comments, chat, etc.). It can be manifested in many ways, including bulk messages, profanity, insults, hate speech, malicious links, fraudulent reviews, fake friends, and personally identifiable information.

Tags. Keywords that describe the content of a web site, bookmark, photo or blog post. Multiple tags can be assigned to the same online resource.

Taxonomy. Taxonomy is an organised way of classifying content, as in a library.

True Positive (TP). It measures the proportion of positives that are correctly identified.

Trust. Trust is defined as a measure of confidence that an entity or entities will behave in an expected way despite the lack of ability to monitor or control the environment in which it operates.

Acronyms

AIC:	Akaike Information Criterion
BIC:	Bayesian Information Criterion
cdRN:	Class Distribution Relational Neighbour
CTT:	Classical Test Theory
FPR:	False Positive Rate
ICM:	Independent Cascade Model
IRT:	Item Response Theory
ICA:	Iterative Classification Algorithm
nLB:	Network only Link Based
OSN:	Online Social Network
PII:	Personal Identifying Information
PPDP:	Privacy Preserving Data Publishing
PQ:	Privacy Quotient
ROC:	Receiver Operating Characteristics
SSN:	Social Security Number
SCC:	Strongly Connected Components
TPR:	True Positive Rate
WCC:	Weakly Connected Components
wvRN:	Weighted Vote relational Neighbour

Biography: Agrima Srivastava

Agrima Srivastava completed her B.Tech in Computer Science and Engineering (CSE) from Banasthali University, Jaipur (Rajasthan, India) in 2011. After completing B.Tech she has worked as an Assistant Systems Engineer at Tata Consultancy Services (TCS), Gandinagar. She enrolled for Ph.D. at the Department of Computer Science and Information Systems of BITS Pilani, Hyderabad Campus in the year 2012. Her research interests are algorithms for preserving user privacy in online social networks, misinformation detection in online social networks, computational social science, data analysis and machine learning. She is passionate about studying social networks and applying data analysis and statistical techniques for understanding and evaluating human behavior.

Biography: Dr. G. Geethakumari

Dr. G Geethakumari is Asst.Professor, Dept. of Computer Science and Information Systems at BITS Pilani, Hyderabad Campus. Before joining BITS, she worked as a faculty in the CSE Dept. at the National Institute of Technology, Warangal. Dr Geetha received her Ph.D. from University of Hyderabad. Her Ph.D. thesis was titled 'Grid Computing Security through Access Control Modelling'. She has many international publications to her credit. Her areas of research interests include: Information security, cloud computing and security, cloud forensics, enterprise security challenges and data analysis, cloud authentication techniques, cyber security, semantic attacks and privacy in online social networks. She has been the Faculty Advisor for Computer Science Association during 2008-2011. Presently she is the IEEE Student Branch Counselor, BITS-Pilani, Hyderabad Campus. She is also the Coordinator for the Linux User Group, BITS Pilani, Hyderabad Campus. Dr. Geetha is a Member, IEEE as well as Member, IEEE Computer Society. She is also a Professional Member, ACM. Dr. Geetha was the Publicity Co-Chair for the IEEE Prime Asia Conference hosted by BITS Pilani, Hyderabad Campus during December 5-7, 2012.

Dr. Geetha was the Publicity Co-Chair for the IEEE Prime Asia Conference hosted by BITS Pilani, Hyderabad Campus during December 5-7, 2012. She was the Organizing Committee Member for the Workshop on Advances in Image Processing and Applications held in BITS Pilani, Hyderabad Campus during October 26 - 27, 2013. She was part of the Organizing Committee for the National Seminar on Indian Space Technology - Present and Future (NSIST-2014) held at BITS Pilani Hyderabad Campus on 1st May, 2014. She has given many guest lectures on topics in emerging areas such as cyber security, cloud computing and cloud security. She has been a member of the Technical Program Committees of various IEEE International Conferences. An extract from the paper 'A taxonomy for modelling and analysis of diffusion of (mis)information in social networks', co-authored by Dr. Geetha and published in the International Journal of Communication Networks and Distributed Systems, Vol. 13, No. 2, 2014, pp.119-143, by Inderscience Publishers, was selected for a press release on 'Semantic attacks in online social media'.