Journal of Theoretical and Applied Information Technology
10th April 2014. Vol. 62 No.1

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

AN ANALYSIS OF PRIVACY RISKS AND DESIGN PRINCIPLES FOR DEVELOPING COUNTERMEASURES IN PRIVACY PRESERVING SENSITIVE DATA PUBLISHING

[1]M. PRAKASH, [2]G. SINGARAVEL

[1]Department of Computer Science and Engineering,
K.S.R. College of Engineering, Tiruchengode, Tamilnadu, India
[2]Department of Information Technology,
K.S.R. College of Engineering, Tiruchengode, Tamilnadu, India
E-mail: [1]mmsprakash.research@gmail.com , [2]singaravelg@gmail.com

ABSTRACT

Government Agencies and many other organizations often need to publish sensitive data – tables that contain unaggregated information about individuals. Sensitive data is a valuable source of information for the research and allocation of public funds, trend analysis and medical research. Publishing data about individuals without revealing sensitive information about them is a significant problem. A breach in the security of a sensitive data may expose the private information of an individual, or the interception of a private communication may compromise the security of a sensitive data. Private and Sensitive information is integral to many data repositories. The efficiency of privacy preserving data mining is crucial to many times–sensitive applications like medical data, voter registration data, census data, social network data and customer data. Where information dissemination is quick and easy, both individuals and custodians of data are getting increasingly cautious about privacy, security and ethical issues. In this paper privacy risks in publishing sensitive data and the design principles for developing counter measures are proposed. The main contributions of this study are four folds. First, domain knowledge about the Privacy and related issues is described. Secondly the definition of the utility of released data with reference to social network model is discussed. In the third fold, knowledge based attacks; vulnerabilities and risk analysis are given. Finally, the design considerations for developing countermeasures in privacy preserving sensitive data publishing are presented.

Keywords: *Data Mining, Data Anonymization, Privacy, Privacy Preservation, Data Publishing, Data Fusion, Data Security*

1. INTRODUCTION

Data mining is the method of querying and extracting helpful patterns, trends, knowledge which is previously unknown, from large amount of information or data. This Data mining process is carried out using various techniques such as those form machine learning and pattern recognition. Nowadays the technology is getting used for a good variety of application, like marketing, finance, medicine, biotechnology, multimedia and entertainment. In recent there has been interest in using data mining for privacy and counter terrorism.

A breach in the security of a database could expose the non-public or private information of an individual, or the interception of a private communication could compromise the security of a database. Private and Sensitive information [1], [2] is essential to many data repositories; as an example, medical records of individual patients contain their names, unique identification numbers, age, addresses, phone numbers, history of ailments, medication details and more. In this where information dissemination is fast and easy or simple, both individuals and custodians of data or information are getting increasingly cautious concerning privacy, security and ethical issues. Two or more companies collaborating in their research and analysis efforts could be very sensitive about their trade secrets. Even within the same multinational company, an individual country's legal system may confine sharing customer information between the subsidiary and the parent. There are no easy solutions to these issues. An

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

understanding and the application of these issues would equip a data mining practitioner or researcher to stay in safe waters and perform his or her duties with the best professional standards.

1.1 Privacy Issues

The privacy can be individual or organisational. Individual privacy implies that nobody should know more about an individual after data mining. Organisational privacy implies that the organisational secrets are protected and could not be disclosed at any cost. The privacy of personal information is an undeniable right.

The persistent questions to a data mining practitioner, researcher and indeed to the society in general about privacy are: What information is private? To whom it is sensitive? What are the risks of the privacy breach? What is an acceptable trade-off between privacy of data and the benefits of data mining?

Organisational privacy issues are more complex than individual. Usually data sharing is for the common benefit but an organisation may lose its edge if its commerce secrets are exposed. Furthermore, there are invariably conflicting interests between parties on privacy; for instance, Web user's desire privacy but the advertisers of web content provider look for web user's private information.

Intuitively we tend to might imagine that it would be easy to ensure privacy by removing, as an example, a name, unique identification number, age, address and phone number from a record. But in reality this is not the case because a data mining expert can match information linking with alternate or other sources to acquire the missing information. Removal of many attributes may render the data useless for the data mining purpose. Data Obfuscation and Data Distribution are the two proven ways for maintaining privacy in sensitive data mining.

1.2 Data obfuscation

Data obfuscation is about making data 'fuzzy' by randomisation, anonymisation and data swapping. The values of the attributes are perturbed by adding random number in data randomisation. Data mining experts can do proper data mining with the distorted data if they are also supplied with the probability distribution of the random variable. In this way, the useful information can be obtained from data without compromising privacy, but it is claimed that the mining time increases. Multiplicative randomisation can be more secure than the additive randomisation but is more complex for data mining.

In the anonymisation, intervals are created for attributes which could be either fixed or variable. For example, a person's income can have the intervals 0-20k, 20k-30k, 30k-40k, 40k-50k, 50k-80k and so on. Individual observation can be replaced with class marks, for example, someone earning between 20k and 30k would be given the value of 25k. We use the term k-anonymisation [3][4][10] when each class has at least k number of instance.

Data swapping was first proposed as a method of preserving confidentially in categorical data. Here the technique involves swapping the values of a sensitive attribute between two very similar instances. For example, to preserve group statics like geographical zoning, the US Census Bureau swaps data between blocks instead of instances. The swap rate or the number of times swapping occurs is just enough to mask sensitive information.

1.3 Data distribution

Data distribution uses multiparty computation protocol. The data are divided into multiples and each party is given data on different entities. No communication between the parties is permitted and each party does its own data mining based on the information provided. The individual results from each party are then combined to view the global outcome. It is common in data distribution for each client own a database and not to trust the other. Each does its own Data mining and challenge is to do global mining from these individual results while preserving every client's privacy. This offers opportunities for companies to share information for their common good.

1.4 Data mining for security of life

There is a prevailing perception of missing creep (using data for other purpose) in which data could be used to find illegal immigrants, for example. Multistate Anti Terrorism Information eXchange (MATRIX) was initially developed by Seisint, a Florida-based company, for sharing intelligent information. The analytical core of the MATRIX project is an application known as Factual Analysis Criminal Threat Solution (FACTS). FACTS is a 'technological, investigative tool allowing query-based searches of available or obtainable state and public records in the data reference repositories'. It permits an authorised user to search 'dynamically combined records from disparate datasets - a mixture of more than 3.9 billion public records

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

collected from thousands of sources. Data include FAA pilot licences records and aircraft ownership records, information on vessels registered with coast guard, property ownership records, state sexual offenders lists, corporation filings, federal terrorist watch lists, state-issued professional licences, an information of criminal history, driver's licence information and photo images, motor vehicle registration information, and the information from the commercial sources which are legally permissible under federal law or generally accessible to the public. The data reference repository excludes privacy-invasive matters such as direct mailing lists, telemarketing call lists, airline reservation records, travel records, magazine subscription, telephone records or calling logs, information about purchase made at retailers or over the internet, credit card numbers, debit card numbers, information of mortgage payment or car payment, bank account details, balance information of bank accounts, marriage licences, divorce decrees, birth certificates, or information of utility bill payment.

1.5 Ethical issues

Ethics is understood to be a set of moral principles and values which guides the behaviour of an individual or an organisation. It is the proper way of doing things as judged by a society and often enforced through law. To act ethically implies acting for the greater benefit of the community within one's conscience and set of guidelines of a professional. But laws cannot be universal and therefore it is possible to act unethically yet legally.

It is difficult to define ethics exactly. The perception of ethics varies among people and can include one or more of the virtues of equity, equality, fairness, trustworthiness, honesty, justice and rationality. However, the prime principles of a code of ethics are universally understood to be, to respect the inherent dignity of an individual, to act on the foundation of a well-informed conscience and to act in the concern of the community.This guide provides details to assist authors in preparing a paper for publication in JATIT so that there is a consistency among papers. These instructions give guidance on layout, style, illustrations and references and serve as a model for authors to emulate. Please follow these specifications closely as papers which do not meet the standards laid down, will not be published.

2. RESEARCH CHALLENGES

Data mining technologies are now being applied to many applications. However, are they ready to detect and/or prevent terrorist activities? Is it possible to completely eliminate false negatives as well as false positives? The false positives could be disastrous to the affected individuals. False could increase terrorist activities. The challenge is to find the "needle in the haystack". The need is knowledge-directed data mining to eliminate false negatives as much as possible.

Mining data in real time is another challenge. In present day the tools are available to detect credit-card and calling-card violations in real time. However can able to build models in real time? The research community having the general view that such real-time model building is still quite difficult. Furthermore, to detect counterterrorism activities, good training examples are needed. How can get such examples - especially in an unclassified setting?

2.1 Information Related Terrorism

Information related terrorism; it means security violations and cyber terrorism through access control and other means. Trojan horses as well as viruses are also information related security violations, which first group it as information-related terrorism activities.

2.2 Cyber Terrorism

The major terrorist threats the society face today is Cyber Terrorism. Because so much of the information is now available electronically, and much of it is on the web, attacks on the databases, computers, networks, and the internet can be devastating to businesses. One of the estimated reports reveals that cyber terrorism could cause billions of dollars worth of damage to businesses. For example consider the banking system. If terrorists attack such a bank's information system and deplete accounts of the funds, the bank could lose millions if not billions of dollars. By crippling the bank's computer system millions of hours of productivity can be lost. Several hours of productively could be loss and result in a major financial loss by a simple power outage at work through some accident. It is therefore critical that the information systems be secure.

2.3 Malicious Intrusions

Malicious intrusions could involve networks, servers, web clients, operating systems, databases and more. In a network intrusion, intruders try to tap into a network and intercept transmitted information. The intruders may be human, or they may be Trojan horses created by humans. Intrusions can also happen on files. For example, a person

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

masquerading as someone else might log into some other's computer and access their files. Intrusions can also occur on databases. Like legitimate users, intruders can pose SQL queries or other queries, and access data that they are not authorised to see.

Essentially, cyber terrorism includes malicious inclusions as well as sabotage through malicious intrusions or otherwise. Cyber security consists of security mechanisms that attempt to provide solutions to cyber terrorism or cyber attacks.

2.4 Credit Card Fraud and Identity Theft

In credit card fraud, a thief gets hold of a person's credit card and makes one or more unauthorized purchases. By the time the owner of the card finds out about the unauthorized activity, it may be too late. The thief may have left the country by them. A similar problem occurs with telephone calling cards too.

2.5 Information security violation

Access control violations are the causes for Information security violations. The users are granted access depending on their roles known as "role-based access control" or their clearance level called "multilevel access control" or on a need-to-know basis. Access controls are violated usually because of either poor design or designer errors. For instance, suppose Paul does not have rights to access salary data. By some error this rule is not enforced and Paul gains rights to access to salary values.

Access control violation can also occur because of malicious attacks. In a malicious attack, a person might enter the system by pretending to be the system administrator, and then delete the access control rule that Paul does not have access to salaries data. Another mode is for a Trojan horse to operate on behalf of malicious user. In this scenario, each time Paul makes a request, the malicious code can make certain that the access control is bypassed.

2.6 Security Problems for the Web

Since the web is the major means of information transportation, web security threats merit special consideration. The web threats here are applicable through any information system including computer networks, operating systems and databases. These threats include integrity violations, access control violations, sabotage, fraud, infrastructure attacks and denial of service.

Traditional access control violations can be extended to the web. Users may access unauthorized data across the web. There is so much data in so many places on the web that controlling access poses quit a challenge. Data on the web may also be subject unauthorized modification. This makes it is easier to corrupt the data. Also, data can originate from anywhere. Consequently, producers of data may not be trustworthy. Incorrect data can cause serious damages such as incorrect bank accounts, which might result in incorrect transactions.

The hackers can break into systems and posting inappropriate messages. Without proper controls, internet fraud can cause business to lose millions of dollars since so much of business and commerce is carried out on the web. Intruders can obtain the identity of legitimate users and might empty bank accounts. The hackers can brought down the infrastructures like the telecommunication system, the power system, and the heating system. This system are controlled by computers and often accessed through the internet. Such attacks can cause denial of service.

Other threats include authenticity, violations to confidentiality, and no repudiation. Confidentiality violations enable intruders to listen in no messages. Authentication violations include using passwords without permissions, and non repudiation violations enable a person to deny the given message. The web threats discussed here occur because of insecure clients, server, and network. To have complete security, one needs end-to-end security; that means secure servers, secure clients, secure middleware, secure operating systems, secure databases, and secure networks.

2.7 Challenging Issues

A major challenge for counter terrorism data mining is privacy. The challenge is to extract useful information while, at the same time, maintaining privacy. Several efforts are under way that attempt to preserve privacy throughout data mining. The different techniques, like randomization, cover stories, or multiparty policy enforcement, can be used to preserve privacy while data mining. While there is some progress in this area, the effectiveness of such techniques needs further evaluation.

3. RESEARCH CHALLENGES

Some issues impacting privacy constraints on data mining are discussed so far. Here the key is an ability to measure privacy. Since privacy has many meanings, it required a set of metrics. In this section several suggestions are proposed.

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

3.1 Bounded Knowledge

The data obscuration techniques guide to a bounded knowledge metric. Bounded knowledge implies that some information about a protected attribute may be revealed, but the actual value can only be estimated. Many researchers [5] given a good measure for quantifying privacy based on such bounded estimation. Based on the differential entropy [11][12] of a random variable the measures are proposed. The differential entropy h(A) is a measure of a uncertainty inherent in A. Their metric for privacy is 2h(A). Specifically, if add noise from a random variable A, the privacy is

$$\Pi(A) = 2^{(-\int_{\Omega A} f_A(a)\log_2 f_A(a)da)}$$

Where Ω_A is the domain of A

This metric has several good and unique features. It is intuitively satisfying for simple causes. For noise generated from A, a uniformly distributed random variable between 0 and a, Π(A) = a. Thus this privacy metric is exactly the width of the unknown region. Furthermore, if a sequence of random variables Ai converges to Π(B). For most random variables, for example a Gaussian, the notion of width of the unknown region does not make any sense. However, it can be calculated Π for such random variables, and above properties allows making the case that the privacy which is equivalent to having no knowledge of the value except that it is within a region of width Π. This gives an intuitively satisfying way of comparing the privacy of different methods of adding random noise.

The authors extend this definition to condition privacy, capturing the possibility that the inherent privacy from obscuring data can be reduced by what learning from a collection. The conditional privacy Π(A|B) is given below which is derived from the conditional entropy definition.

$$\Pi(A|B) = 2^{(-\int_{\Omega A,B} f_A(a,b)\log_2 f_{A|B=b}(a)dadb)}$$

It shows how this can be applied to measure the actual privacy after reconstructing distribution of the original data to improve the accuracy of decision trees build on the obscured data. The result is that data obscuration techniques do not provide as much privacy as it might naively expect, as the ability to use them to produce valid data mining results also decreases the effectiveness to addition noise. Similar analysis on other data obscuration techniques would provide an effective way to compare those techniques.

Another use of this metric is to evaluate the inherent loss of privacy caused by data mining results or outcomes. The use of conditional privacy enables to estimate how much privacy is lost by knowing the data mining results even with a "perfect" privacy-preserving technique such as secure multiparty computation. The literature has not yet addressed this issue; the assumption has generally been that the data mining results do not of themselves violate privacy.

In this study the social network model is taken for further discussions with respect to the privacy requirements, risks and countermeasures. In the context of social network data publishing, many researchers have proposed anonymization techniques for sanitizing social network data before releasing it for third party mining services [6][7][8]. It is widely recognized that the primary privacy risks in publishing social network data is centred on the inference attacks, namely high confidence inference of associations of published data with some background identity information of known individuals, thus leading to intrusion of privacy of such individuals.

For example, one may infer salary information for an individual in census data, or infer individual's viewing history in video or individual's search history in search. All of these risks are inferred by linking sensitive information to some identity of an individual that is available as background knowledge or common sense or domain-specific knowledge. Example for background knowledge, which includes information about the released data set such as X dined in restaurant M, or P has disease Q. Example for common sense, which includes common knowledge about entities involved in the data set such as teen is children between 13 and 18.

Thus, privacy preserving publishing of social network data should aim at preserving privacy required by users (social network entities) while maintaining the maximum utility of released data. However, if the released social network data is used for community detection [5], then preserving the original graph structure is perhaps the most important utility.

4. SOCIAL NETWORK DATA PUBLISHING MODELS

4.1 Social Network Model

Social relationship can be represented in two ways, Friendships and activity groups. In friend relationships, people get connected through the

Journal of Theoretical and Applied Information Technology
10th April 2014. Vol. 62 No.1

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

friend request and agreement protocol to establish the mutual confirmation of friendship. In activity group relationships, people are getting connected through engaging in the same type of activities or events. In fact, friendship can be viewed as one specific type of activity group. One of the big advantages of promoting activity based classification of social ties is to facilitate finer granularity of exploration on social relationships among people to enhance personal and business experiences of social networks.

4.2 Sensitive Attributes in Social Network

Social network data publishing usually addresses the privacy of users by removing user identity attributes in user profiles, such as user ID or user name which is used to identify a user uniquely in a social network. However, recent privacy research has pointed out numerous privacy attacks over perturbed social network datasets where user IDs and names are completely removed. Most of such privacy risks are due to the inference attacks over attributes of quasi-identifier and sensitive attributes of the user profiles. Thus it is described that the social network data publishing model in terms of profile attributes and structured attributes, categorize all attributes into the following three types: (1) Identifier attributes (2) Quasi-identifier attributes, and (3) Sensitive attributes.

4.3 Identifier data

The most identifier attributes which is used to represent the user are user names and unique user identification number or social security number (SSN). The identifier attributes is removed in all data publishing techniques from the raw data set prior to release. In many cases, removing of identifier attributes may not be sufficient.

4.4 Quasi-identifier data

Quasi-identifier attributes is used in combination to uniquely identify a profile, though used in separation causes no disclosure danger. Examples of quasi-identifiers are birthdate, residence zip code, city code, gender of users. It is well known that by combining birthday, sex and zip code, one can identify large population of people living in United States without knowing their identifier data [3][9]. For social network data, structural features of a node, like degree and neighbourhood sub-graph, can be used as quasi-identifiers to uniquely identify a node. This can lead to the privacy risk of linking a user ID with the sensitive information contained in the published dataset when certain background knowledge is made publically available. Quasi-identifier data often contain good utility of released data for both safe sharing and retention.

4.5 Sensitive data

In social network context sensitive data refers to those user attributes in their profiles which are considered private and have controlled access, like specific activities of a user. Also certain relationships between nodes can be sensitive data. For example, an activity that Harshith and Jeni had a date on 10/10/2013 in Bangalore can be sensitive information for both of them. Thus, publishing this activity simply by replacing user IDs of Harshith and Jeni with randomly generated numbers can be risky due to possible inference attack on the subgraph of two user nodes connected to the same group node which represents a dating activity on 10/10/2013 in Bangalore.

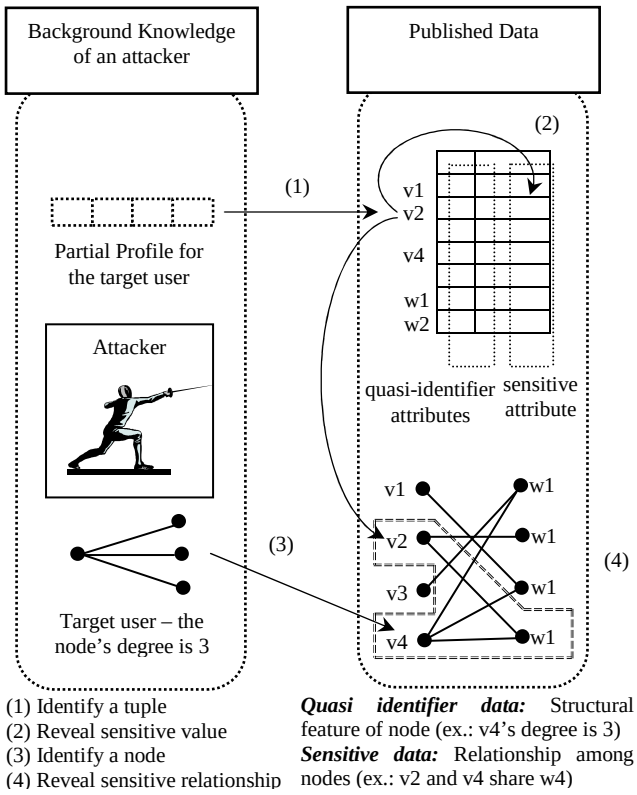

Figure 1: Quasi-identifying data and Sensitive data in Social Network

Furthermore, the participation of certain event or specific interest group of a user can be highly sensitive to some users but not considered sensitive to others. Thus, data sensitivity is a very personal issue and may differ from event to event, user to user, location to location, and time to time.

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

Sensitive attributes of entities, on one hand, are the part of data with highest utility and on the other hand, present the sensitive associations that need to hide or prevent identity linking attacks.

Figure 1 shows the intimate relations among the three types of attributes for social network datasets and how an attacker reveals sensitive data of some target users. Basically, with certain background knowledge acquired by an attacker, such as partial profile of a target user or partial structure/relationship pattern, the attacker may identify the profile of the target user with high probability as shown in Figure. 1(1), then the attacker can infer sensitive attribute values of the target user in the corresponding profile, depicted in Figure 1(2). In addition, the attacker can also reveal the relationship of the target user in the graph structure, shown in Figure 1(3), and reveal sensitive relationship of the target user, in Figure 1(4). Figure 2 shows the Permutation technique applied for user-group affiliation network.

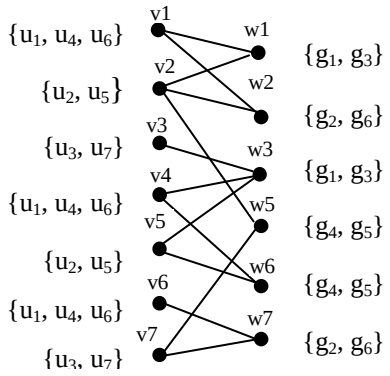

Figure 2: Permutation technique applied for user-group affiliation network

5. VULNERABILITIES AND RISK ANALYSIS

One way to look at anonymization is to contemplate it as an obfuscation methodology of adding uncertainty to certain data such that an attacker cannot be sure about the presence or the associations of released data to some known identity. The major point for quality measure of privacy is the confidence level in which all possible interpretations of data being released have equal probability of unauthorized disclosure under association attacks and or linking attacks. A key quality measure of utility is the range of services that can be performed over the released data with high confidence and accuracy. The detailed description about the vulnerabilities are given in this section which are found in existing social network data anonymizaiton mechanisms and provide better understanding of privacy risks involved in publishing social network data as well as the importance of realistic assumptions in designing effective countermeasures.

5.1 User-Group Constraint Violation Attack

An adversary makes use of one's background knowledge to define a set of constraints between user nodes and group nodes and between user nodes that participate in the same group activity. Such graph structure based constraints can be a powerful tool for launching inference attacks. For example, in Japan a woman who is less than 15 years old cannot get marriage is a well-known custom constraint on marriage relationship. The one who is a vegetarian does not eat meat is another general or common-sense constraint on user group. An adversary can utilize this type of user-group constraint violation attacks to identify and eliminate those possible worlds that clearly do not make sense or impossible to be true.

The adversary will select the right set of background knowledge in order to isolate those possible worlds that have low probability from those high probability ones. Thus the adversary can bring in a time-difference constraint between user and group. Referring the Figure 2, and using this constraint, easily can detect that (*u1*, *g3*) and (*u2*, *g3*) are violating this constraint since *u1*, *u2* has Japan as its current residence country in their profiles. Based on this analysis, the adversary can easily identify those user-group relationships which violate the constraint. In this example, (*u1*, *g3*) and (*u2*, *g3*) violate the constraint meetingtime_const, thus the attacker can eliminate those possible worlds which include these pairs of nodes. There are 4 possible worlds remaining after removing the inappropriate possible worlds, and they are shown as follows

$$PW(G, G') = \{ pw(u_3, u_4, u_5, g_3), pw(u_3, u_6, u_5, g_3), pw(u_7, u_6, u_5, g_3), pw(u_7, u_4, u_5, g_3)\}$$

The attacker can detect that (u_5, g_3) has higher probability to be the true user-group link by eliminating those obvious false possible worlds.

5.2 Probability Skew Attack

The Probability Skew Attack deals the situation like an adversary may not be able to determine with high confidence which possible worlds to eliminate. Often in such situations, if an adversary uncovers the skewed probability distribution on the possible set of worlds for anonymized SN graph, the adversary may influence the probabilities for

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

skewed distribution to instigate a successful inference attack.

For example, an adversary may define a scoring function *f(u,g)* between a possible user node *u* and a possible activity group node *g* based on his background knowledge. This function calculates the probability of this association to be true in the original SN graph.

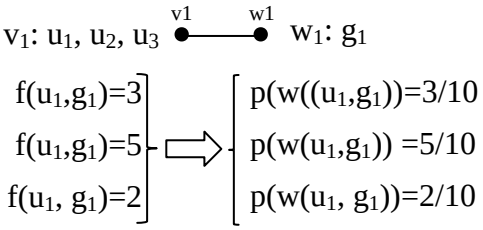

Figure 3: Skewed probability distribution attack and an example scoring function

The Figure 3 shows an example of such scoring function. Here is an anonymized user-group association *(v_1, w1)*, where $v_1 = (u_1; u_2; u_3)$ and $w_1 = \{g_1\}$. Thus the three possible worlds are: *(u_1; g_1), (u_2; g_1), (u_3, g_1)*. Assume that the adversary make use of his background knowledge to acquire the scoring function and the scores for *(u_1, g_1), (u_2, g_1)* and *(u_3, g_1)* are 3, 5 and 2 respectively. The probability for each possible world can be easily computed. By using value of function, the attacker can infer the probability of each possible world.

Anonymization technique introduces uncertainty into certain data. A desired property of anonymization is to ensure that all possible worlds have equal or very similar probability to be the true world. However, by exposing information in level 2 and 3, such ideal condition is no longer valid because different possible worlds may have different probabilities for being the true world. In this attack, an adversary tries to select a world which is the closest to the true world based on his background knowledge. Concretely, one approach to conducting such inference is to define a score function *f(pw(G, G'))* that can produce a ranking of possible worlds in terms of how closely each matches with the true world entity. Such scoring function should take into account of as much background knowledge as possible to improve the attack-resilience of the published SN graph.

For example, g_3 in Figure 4 refers to a meeting, and then an attacker may use his background knowledge to assume that people who attend the meeting have the same or similar professional profile with each other. The attacker defines a score function based on this assumption, so that the possible world that closely matches with one another will have higher probability to be mapped to the true world. To define the background knowledge, the example score function can be introduced. For example, the adversary can select a set *U* of attributes that are representative of user's professional attributes. The maximum number of people can be counted for each selected attribute of people who have the same value and we regard the max value as the similarity of the attribute sim_{attr}.

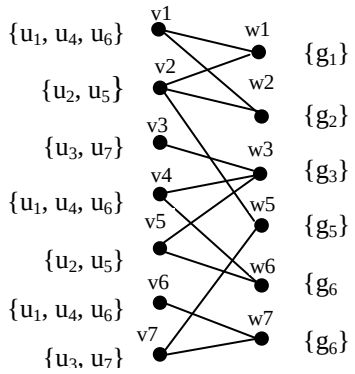

Figure 4: Graph structure of an example dataset

Consider an example, u_3 and u_5 have A as their residence city in a possible world *pw(u_3, u_6, u_5, g_3)*, but the city of u_6 is B then sim_{city} is 2. The score function *f(pw(G, G'))* for each possible world can be defined by calculating the sum of the values for all attributes is as follows

$$f(pw(G, G'), U) = \sum_{a \in ATTR(U)} sim_a$$

The score for *pw(u_3, u_6, u_5, g_3)* is calculated as follows

$$f(pw(u_3, u_6, u_5, g_3)) = sim_{age} + sim_{job} + sim_{city} + sim_{country}$$
$$= 1 + 2 + 1 + 2 + 2 = 8$$

All other possible world's scores are given under;

$f(pw(u_3, u_4, u_5, g_3)) = 1 + 2 + 2 + 3 + 3 = 11$

$f(pw(u_7, u_4, u_5, g_3)) = 2 + 3 + 1 + 2 + 2 = 10$

$f(pw(u_7, u_6, u_5, g_3)) = 2 + 2 + 1 + 2 + 2 = 9$

Based on the scoring function and the results, the attacker identifies the possible world with the highest similarity score as the most probable matching to the true world. From the above example, given that *pw(u3; u4; u5)* has the highest similarity score of 11=(11 + 10 + 9 + 8) = 11=38, thus it is identified by the attacker as most likely the true world. When an attacker calculates the

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

possibility of the true entity based on the scoring function *f(pw(G, G'), U)*, the *highest* possibility calculated by the following formula is given below

$$\frac{max_{pw \in PW(G, G')} f(pw)}{\sum_{pw \in PW(G, G')} f(pw)}$$

which is greater than $\frac{1}{|PW(G, G')|}$

6. DISCUSSIONS AND COUNTERMEASURES

The privacy risks in publishing anonymized social network data are described. The two types of background knowledge based attacks are constraint violation attack and probability skew attack. One of the fundamental vulnerabilities in the design of graph permutation techniques is the lack of consideration of background knowledge and the risks of combining background knowledge with published profile data and graph structure data.

Concretely, take *(k, l)* grouping permutation approach as an example, the algorithm for composing *l* activity groups and *k* user groups from the input social network *G = (V, W, E)* focuses on meeting the safe condition that nodes in same group of *V* have no common neighbours in *W*, which is a higher utility condition, but it does not assure background knowledge attack resilience.

For example, the *(k, l)*-grouping algorithm first sorts the nodes by groups in the sense that user nodes that connect to the same group node are queued together. To simplify the discussion, let us set *l* = 1. Then consider each group as a class c and for each node in the sorted list, it checks whether the node and each class *c* satisfy the safety condition and if yes, this node is added into class *c*. obviously, the sorting condition can be revised. Instead of sorting nodes according to groups, the nodes are sorted in terms of both groups and attribute similarity.

Also the safety condition is revised such that nodes in similar or same group of *V* cannot have any common neighbours in *W* and additionally their aggregated attribute similarity ought to be higher than a system-defined threshold to ensure user nodes that are too dissimilar should not be placed within the same cluster. The intuition behind the design of a new countermeasure is two folds: First, for those user nodes that are very similar with respect to the group nodes they are related to, then putting them into one cluster will create the anonymized graph safer and more resilient to background knowledge attacks. Second, putting those user nodes that are more likely to join a particular group but have not nonetheless because the most eligible candidates to be added into the group cluster. This will instantly increase the resilience to the both attacks, namely violation of user-group constraint attack and probability skew attack.

Potentially interesting another countermeasure is to combine *k*-anonymity based algorithm with *(k, l)*-grouping permutation. For instance, *k*-anonymization can be applied over the set of nodes to construct *k*-anonymized groups. Such group can then be tuned and revised to obey the safety conditions.

7. CONCLUSION

Privacy risks in publishing sensitive data and the design principles for developing counter measures are discussed. The social network model is taken as platform for the analysis of privacy breach. The main contributions of this study are four folds. The domain knowledge of the privacy and its related issues are described first. Second, two types of background knowledge based inference attacks are identified with reference to the social network model. Then, the knowledge based attacks are given followed by vulnerabilities and risk analyses are presented. Finally, some of the design principles are discussed for developing countermeasures in privacy preserving sensitive data publishing.

REFERENCES

[1] C. Agarwal and D. Agarwal, "On the Design and Quantification of Privacy Preserving Data Mining," *Proc. of ACM SIGACT – SIGMOD - SIGART Symposium on Principles of Database Systems*, pp. 247-255, 2001.

[2] M. Kantarcioglu and J. Vaidya, "An Architecture for Privacy - Preserving Mining of Client Information," *Proc. of IEEE Int'l Conf. on Privacy, Security, and Data Mining*, pp. 37-42, 2002.

[3] L. Sweeney, "*k*-Anonymity: A Model for Protecting Privacy," *Int'l Journal of Uncertainty, Fuzziness and Knowledge-Based Systems*, vol. 10, no. 5, pp. 557-570, 2002.

[4] K. LeFevre, D. J. DeWitt and R. Ramakrithnan, "Mondrian: Multidimensional *k*-Anonymity," *Proc. of ICDE*, 2006.

[5] S. Bhagat, G. Cormode, B. Krishnamurthy and D. Srivastava, "Class-based graph anonymization for social network data," *Proc.*

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

of the VLDB Endowment, vol. 2, issue 1, pp.766-777, 2009.

[6] M. Hay, G. Miklau, D. Jensen, D. Towsley and C. Li, "Resisting structural re-identification in anonymized social networks," *The VLDB Journal*, vol. 2010, No.19, pp.797 – 823, 2010.

[7] K. Liu and E. Terzi, "Towards identity anonymization on graphs," *Proc. of the 2008 ACM SIGMOD international conference on Management of data (SIGMOD-2008)*, pp.93–106, 2008.

[8] L. Zou, L. Chen and M. T. Ozsu, "K-automorphism: A general framework for privacy preserving network publication," *Proc. of the VLDB Endowment*, vol.2, issue 1, pp. 946–957, 2009.

[9] A. Machanavajjhala, J. Gehrke, D. Kifer and M. Venkitasubramaniam, "*l*-Diversity: Privacy Beyond *k*-Anonymity", *Proc. of the Int'l Conf. on Data Eng. (ICDE)*, p. 24, 2006.

[10] Ninghui Li, Tiancheng Li, and Suresh Venkatasubramanian, "t-Closeness: Privacy Beyond k-Anonymity and l-Diversity," *Proc. of the Int'l Conf. on Data Engg.(ICDE)*, pp. 106-115, 2007.

[11] M. Prakash and G. Singaravel, "A new model for privacy preserving sensitive Data Mining," *Proc. of IEEE Int'l Conf. on Computing Communication & Networking Technologies (ICCCNT'12)*, pp. 1-8, 2012.

[12] M. Prakash and G. Singaravel, "A Study of Privacy Risks and Countermeasures in Mining and Publishing of Sensitive Data," *Proc. Int'l Conf. Innovation in Communication, Information and Computing (ICICIC'13)*, ISBN - 978-81-925286-87, 2013.