www.ijecs.in
International Journal Of Engineering And Computer Science ISSN:2319-7242
Volume – 5 Issue -03 March, 2016 Page No. 15912-15918

Anonymization in Social Networks: A Survey on the issues of Data Privacy in Social Network Sites

A.Praveena, Dr.S.Smys
Assistant Professor, Department of Information Technology
Dr.N.G.P Institute of Technology, Coimbatore
Tamilnadu, India
Professor & Head, Department of Information Technology Karpagam College of Engineering, Coimbatore
Tamilnadu, India

Abstract— **Recent years have seen unprecedented growth in the popularity of social network systems. Social networks are online applications which allow its users to connect by various link types. Online social networks, such as Facebook, LinkedIn are progressively utilized by many people. It makes digital communication technologies sharpening tools for extending the social circle of people. It has already become an important integral part of our daily lives, enabling us to contact our friends and families on time. As social networks have developed rapidly, recent research has begun to explore social networks to understand their structure, advertising and marketing, and data mining online population, representing 1.2 billion users around the world. More and more social network data has been made publicly available and analyzed in one way or another.**

But some of the information revealed is meant to be private hence social network data has led to the risk of leakage of confidential information of individuals. However, privacy concerns can be used to prevent these efforts. Privacy pre- serving publishing of social network data becomes a more and more important concern. This leads for privacy-preserving social network data mining, which is the discovery of information and relationships from social network data without violating privacy. Privacy in online social networks data has been of utmost concern. Hence, the research in this field is still in its early years. Privacy preservation is a significant research issue in social networking. Since more personalized information is shared with the public, violating the privacy of a target user become much easier. We argue that the different privacy problems are entangled and that research on privacy in OSNs would benefit from a more holistic approach. In this paper, after an overview on different research subareas of SNSs, we will get more focused on the subarea of privacy protection in social network and we present a brief yet systematic review of the existing anonymization techniques for privacy preserving publishing of social net- work data.

Keywords—Privacy attacks, social networks, Anonymization

I. INTRODUCTION

Social networks are online web based applications that allow their users to connect with their friends and share common interests, either professionally or personally with whom they share and these networks allow people to list details about themselves that are relevant to the nature of the network. According to Cluley "Social networks are great fun, and can be advantageous but people really need to understand that its complicated world and you need to step wisely". Users post content to the application to update connections and share personal news, accomplishments, interests and more. This content can be in the form of simple text status updates, videos or photos. People may use to find a new job, find new clients or stay in touch with long distance friends and family. Examples are LinkedIn, Facebook, Twitter and YouTube.

Social networks often offer additional applications that extend their functionality through games, quizzes which have built by third party developers and have the potential to introduce the security risks. Facebook is a popular general-use social network where individual users list their favorite activities, books, and movies. Because these sites gather extensive private personal information, application providers have a rare opportunity. That is direct use of this information could be useful to advertisers for direct marketing. However, privacy concerns can be used to prevent these efforts. This conflict between desired use of data and individual privacy presents an opportunity for privacy-preserving social network data mining, that is, discovery of information and relationships from social network data without violating privacy.

Privacy concerns of individuals in a social network can be classified into two categories: privacy after data release, and private information leakage. Examples of privacy after data release involve the identification of specific individuals in a data set subsequent to its release to the general public or to paying customers for a specific usage. Private information leakage deals with the details about an individual that are not specifically stated, but, rather, are inferred through other details released and/ or relationships to individuals who may express that detail. This paper focuses on the problem of private information leakage for individuals as a direct result of their actions as being part of an online social network.

This paper explores how the online social network data could be used to predict some individual private information that a user is not willing to disclose and explore the effect of possible data sanitization approaches on preventing such private information exposure, also letting the receiver of the sanitized data to do inference on the non-exclusive details. We explore how the online social network data could be used to predict some individual private detail that a user is not willing to disclose and explore the effect of possible data sanitization approaches on preventing such private information leakage,

while letting the receiver of the sanitized data to do inference on non private details.

This problem of private information leakage could be an important issue in some cases. Research in these areas has revealed interesting properties of the data and presented efficient ways of maintaining, mining and querying them. However, with the exception of some recent work, the privacy concerns associated with data-analysis over graphs and networks have been largely ignored.

In a network, nodes correspond to individuals or other social entities, and edges correspond to relationships between them. The privacy breaches in a network can be grouped to three categories: 1) identity disclosure: the identity of an individual who is associated with a node is revealed; 2) link disclosure: the sensitive relationships between two individuals are disclosed; and 3) content disclosure: the sensitive data associated with each node is compromised, e.g., the email message sent and/or received by the individuals in a email communication network.

A privacy-preservation system over graphs and networks should consider all of these issues. However, compared with existing anonymization and perturbation technique, working with graphs and networks is much more challenging due to the following reasons: 1) It is difficult to model the background knowledge and the capability of an attacker. Hence, it is not clear what are the most appropriate privacy models for graphs and networks, and how to measure the privacy breach in that setting. 2) It is difficult to quantify the information loss. 3) It is even difficult to devise graph-modification algorithms that balance the goals of preserving privacy with the utility of the data. Therefore, the impact of a single change of an edge or a node can spread across the whole network. 4) It is difficult to model the behavior of the participants involved in a network-based collaborative computing environment.

Most of the privacy solutions focus on content privacy rather than contextual privacy (e.g. some users communicate among them). A method that is used to publish data and protect user privacy is called anonymization. Anonymization in context with privacy can be defined as: replacing the information that can damage user privacy with the harmless information. Anonymization is one of the methods mostly used for security purpose such as preventing personal and sensitive information and decrease the success rate of attacks such as context aware spam attack and context aware phishing attack. Anonymization techniques [12] can be classified into four approaches: clustering, clustering with constraints, and modification of graph and hybrid approach.

Various anonymization techniques [19] have been researched to protect the social networking sites from de-anonymization and make it difficult for the adversary to re-identify an anonymized user in online social networks. To combat these challenges, several authors have recently developed different types of privacy models, adversaries, and graph-modification algorithms. Unfortunately, none of the work is likely to solve all the problems in one shot. Protecting against each kind of privacy breaches may require different techniques or a combination of them. In this chapter, we detail a number of recently developed techniques for each type of the disclosure described above. We hope this survey can offer insight into the challenges and therefore opportunities in this emerging area.

II. Possible Attacks in Social Networks

The volume and accessibility of personal information available on social networking sites have attracted malicious people who seek to exploit this information. The same technologies that invite user participation also make the sites easier to infect with malware that can shut down an organization's networks, or keystroke loggers that can steal credentials. Common risks such as spear phishing, social engineering, spoofing, and web application attacks attempt to steal a person's identity. Such attacks are often successful due to assumption of being in a trusting environment.

A. Identity Theft

Identity theft is a form of stealing someone's identity in which someone pretends to be someone else by assuming that individual identity, typically in order to access resources and other benefits in the individual's name. It occurs when someone uses another's personal identifying information, like user name, identifying credit card number, or number without their permission to commit deceit or other crimes.

B. Secondary Data Collection

Secondary data means the data collected by someone other than user. The data published on an OSN helps adversary to guess the additional information like SSN which often act as key to personal information of the victim. Matching the profiles of a person in both: a professional and a more informal online social network for analysis and comparison of the content published in both is another obvious and frequent type of secondary data collection.

C. Communication tracking

It is a threat on communication privacy where malicious social networking sites procurer makes the undue advantage of its set of rights to perform communication to the divulgence of personal communication information to the adversary.

D. Harvesting

In harvesting, an attacker such as social networking site procurer tries to gather information of participant on a large scale for a motive, but user is not vigilant or aware about it.

E. Traffic Analysis

Traffic analysis is the process of intercepting and examining messages in order to deduce information from communication patterns. It can be performed even when the messages are encrypted and cannot be decrypted. Traffic analysis can be used to determine what type of information is being communicated (such as chat, email etc.), even if the data itself is encrypted. This attack would be most effective against encrypted proxies. Ronggong Song and Larry Korba presented different types of traffic analysis attacks:

a) Timing Attack- In this type of attack, the attacker observes incoming and outgoing messages in the network from which it can access useful timing information of route by associating the messages in the two sets. If an attacker has access to one of the communicating parties, he may be able to deduce which route is used with the help of round trip time.

b) Message Delaying Attack- The attacker obtains enough resources and easily monitor network or check whether the receiver is receiving other messages by holding the messages.
c) Replay Attack- In this attack, the attacker examine the incoming and the outgoing messages and then send the same message to the node again and would actuate which message is sent twice from node.
d) Packet Volume and Counting Attack- The attacker observes amount of data being transmitted by sniffing the packets on any router of path between sender and initial node. Due to varying packet size, the attacker can distinguish the message sent by node because size of sent and received message will be same and is correlated by an attacker.
e) Communication Pattern Attack- When user sends or receives messages, a lot of information is received by attacker by observing communication pattern. It's one of menacing attacks and is very effective in real-time environment.

III. Related Work

Recent years have seen unprecedented growth in the popularity of Social Network Systems (SNSs), with stories concerning the privacy and security of such household names as Facebook and MySpace appearing repeatedly in mainstream media. Mobile social networks as emerging social communication platforms have attracted great attention recently, and their mobile applications have been developed and implemented pervasively. Privacy is one of the major concerns when sharing data on network for business analysis. Privacy preservation is the most important area in today's computing field. The area of privacy inside a social network encompasses a large breadth, based on how privacy is defined.

In paper "Privacy Preservation by *k*-Anonymization of Weighted Social Networks", proposed an anonymization technique for weighted graphs, i.e. for social networks. They proposed a method that provides *k*-anonymity of nodes against attacks where the adversary has information about the structure of the network, including its edge weights. They mainly consider prevention of identity disclosure, but they also touch on edge and edge weight disclosure in weighted graphs. The advantage of this method, it efficiently work in a weighted graph whereas drawback of this method, to preserve utility of the graph.

In paper "Anonymizing Classification Data for Privacy Preservation" discussed method TDR, which having the large scale of data sets and complex anonymity requirements. There research objective in this is to evaluate the method, that is, TDR, for preserving the usefulness of classification and the scalability on large data sets. In past researched work some model the classification metric on the masked table, the optimality of such metrics does not translate into the optimality of classifiers. According to author knowledge, classification of anonymity on the basis of single dimensional generalization and on basis of this impact is evaluated. Because of these reasons, there evaluation used baseline of unmodified data and reported results. All experiments on TDR were conducted on an Intel PIV 2.6-GHz PC with 1-GB RAM.

The proposed TDR method having many advantages, it should produced comparable accuracy, TDR much more efficient than previously researched and TDR have also found better anonymization solution for classification.

In paper "Protecting Sensitive Labels in Social Network Data Anonymization", proposed technique trust-based privacy preservation method for data sharing in P2P. In a P2P system, in which privacy is different from the traditional node anonymity problem and the identities of the participants are known. During data acquirement a peer acted as the proxy server. To get the data through proxy server, requester sends the request first which made it difficult for the eavesdroppers. A privacy measuring method is given to evaluate the proposed research. A recommendation caused the largest change in the trust value is predetermined by the peers. Least-square-error method are determined the parameters. When a threshold exceeds then the difference between the predicted trust value and the observed value and the algorithm will change the values of the parameters. The advantages of this method are efficiency/accuracy trade-off in trustworthiness assessment. The drawback of this method is lack of privacy.

In paper *"*Trust-Based Privacy Preservation for Peer-to-Peer Data Sharing*"* suggested technique privacy preserving repository to integrate data from various data sharing services. The performance of decryption depends on these parameter repository only collects the minimum amount of information from data sharing services based on users' integration requests, and data sharing services can restrict our repository to use their shared information only for users' integration requests, but not other purposes. In this research they have only focused on matching operations and additive homomorphism encryption schemes there repository could be easily extended to support SUM and AVG aggregate operations. The drawback is it cannot worked for large scale data sets and enable there repository to support more types of data integration operations.

In paper "Inferring Privacy Information from Social Networks", He et al. consider ways to infer private information via friendship links by creating a Bayesian network from the links inside a social network. This work focuses on social network data classification and inferring the individual's private information. More private information is inferred by applying collective classification algorithm. This system uses a collective classification algorithm for classifying the social network data. It showed that, user's private information can be inferred via social relations and release of personal information in social network. To protect the individual's private information the system either hide our friendship relations or ask our friends to hide their attributes.

Research efforts [13], [14], [17] have been put on identity presentation and privacy concerns in social networking sites. Gross and Acquisti argued that users are putting themselves at risk both offline (e.g., stalking) and online (e.g., identity theft) based on a behavior analysis of more than 4,000 students who have joined a popular social networking site. Stutzman presented a quantitative analysis of identity information disclosure in social network communities and subjective opinions from students regarding identity protection and information disclosure.

IV. Categories of Privacy Beach

The privacy breaches in social networks can be categorized into three types [9-10]: i. Identity disclosure - Identity disclosure occurs when an individual behind a record is exposed. This type of breach leads to the revelation of

information of a user and relationship he/she shares with other individuals in the network. ii. Sensitive link disclosure - Sensitive link disclosure occurs when the associations between two individuals are revealed. Social activities generate this type of information when social media services are utilized by users. iii. Sensitive attribute disclosure – Sensitive attribute disclosure takes place when an attacker obtains the information of a sensitive and confidential user attribute. Sensitive attributes may be linked with an entity and link relationship.

All these mentioned privacy breaches pose severe threats like stalking, blackmailing and robbery because users expect privacy of their data from the service provider end. Besides that it damages the image and reputation of an individual. There are many examples of accidental disclosure of private information of users' data that causes organizations to be conservative in releasing the network data, such as the AOL search data example [11] and attacks on Netflix data [12]. Therefore, data needs to be released to third parties in such a way that ensures the privacy of the users. Thus data should be anonymized before releasing or publishing to third parties. But preserving privacy in social networks is difficult.

V. Modelling privacy preservation in Social networks

To battle privacy attacks and develop protection techniques in social networks, we need to model three aspects. First, we need to identify the *privacy information* which may be under attack. Second, we need to model the *background knowledge* that an adversary may use to attack the privacy of target individuals. Last, we need to specify the *usage* of the published social network data so that an Anonymization method can try to retain the utility of the data as much as possible while the privacy information is fully preserved. Generally, we model a social network as a simple graph $G =(V;E; L;LV ;LE)$, where V is a set of vertices, $E \mu V £ V$ is a set of edges, L is a set of labels, and a labeling function $LV : V ! L$ assigns each vertex a label and a labeling function $LE : E ! L$ assigns each edge a label. For a graph G, $V(G)$, $E(G)$, $L(G)$, $LV(G)$, and $LE(G)$ are the set of vertices, the set of edges, the set of labels, the vertex labeling function in G, and the edge labeling function in G, respectively.

VI. Categories of Anonymization Methods

In privacy preserving data publishing, in order to prevent privacy attacks, data should be anonymized properly before it is released. Anonymization methods should take into account the privacy models of the data and the utility of the data. Generalization and perturbation are the two popular anonymization approaches for relational data. Although privacy preservation in social network data is a relatively new problem, several privacy preserving methods have been developed. Similar to privacy preservation methods in relational data, specific anonymization methods are developed for specific privacy models of social networks and specific utility goals of anonymized data.

Anonymization methods on social network data are categorized as follows. Clustering-based approaches: It clusters vertices and edges into groups and anonymizes a subgraph into a super-vertex. In such a way, the details about individuals can be hidden properly. The methods in this category can be further divided into vertex clustering methods, edge clustering methods, vertex and edge clustering methods, and vertex-attribute mapping clustering methods.

Graph modification approaches: A graph modification method anonymizes a graph by modifying (that is, inserting and/or deleting) edges and vertices in a graph. The modification can be conducted in three ways and correspondingly there are three sub-categories of the methods. Last, the greedy graph modification approaches greedily introduce modification to meet the privacy preservation requirement and optimize the data utility objectives.

VII. Challenges in Preserving Privacy in Social Network Data Publishing of Privacy Beach

Ensuring privacy for social network data is difficult than the tabular micro-data because [13]:

1) Modeling of background knowledge of adversaries is difficult in social network data than tabular micro-data. In tabular micro-data, users are identified by linking quasi-identifiers from whereas in social network information from various sources such as labels of vertices and edges, subgraphs, and neighborhood graphs can be used to identify individuals.

2) Information loss is the metric which measures the amount of distortion. In tabular micro-data information loss can be measured using the sum of information loss in individual records. Since, a social network is a graphical structure with a set of vertices and edges hence it is difficult to compare two social networks by comparing the vertices and edges individually.

3) Development of privacy preserving techniques in social network data is difficult than for relational data. The methods proposed for tabular micro-data cannot be directly applied to social network data due to the connectivity between vertices in the graph network as compared to independent nodes in tabular data. An adversary can use the information regarding network structure to violate privacy of users. So there is a need is to develop a technique that can ensure the privacy of entities in network data publishing.

VIII. Privacy preserving Techniques in Social Networks

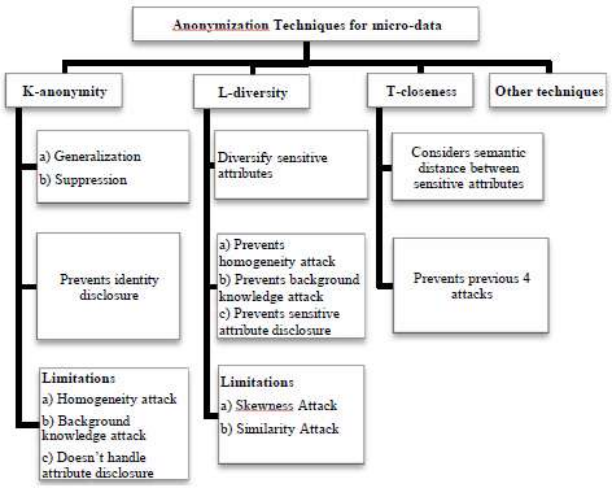

Fig. 1. Existing Privacy Preserving Techniques in Social Networks

Significant work has been done for preserving privacy in tabular micro-data. Models like K-anonymity [14], L-diversity [15], T-closeness [16] have been proposed which have shown good results in anonymization. Fig. 1 briefs three models, their

properties and drawbacks. Wei et al. [21] considered privacy disclosure in online social network data publishing. It has been assumed that adversaries have knowledge of degree of a target individual and target's immediate neighbours. A solution to defend against background knowledge attacks was proposed.

Anonymized social networks obtained by proposed method can be used to answer aggregate network queries with high accuracy. Zou et al. [22] proposed k-automorphism based on the assumption that the adversary has knowledge about degree, subgraph and neighbor of the target node. Cheng et al. [24] used K-isomorphism to preserve privacy when adversary has subgraph knowledge. Wu et al. [25] proposed k-symmetry technique to protect privacy against re-identification using subgraph information. Lan et al.[26] developed an algorithm called KNAP against 1-neighborhood attack for publishing social networks data. Skarkala et al. [27] applied K-anonymity to weighted social networks. Liu et al. [28] proposed the concept of k-degree to prevent vertex re-identification through the information of vertex degree.

Preserving privacy in social networks using k-anonymity protects against linking disclosure but still it may leak privacy under the cases of homogeneity and background knowledge attacks. Moreover, it doesn't protect against attribute disclosure. So, L-diversity was developed by Machanavajjhala [15] in year 2007. Panda et al. [29] used a new practical and efficient definition of privacy called l-diversity on preserving privacy in collaborative social network data and the effect on the utility of the data for social network analysis has been seen. It has been identified that l-diversity social network still may leak privacy as an adversary may have some prior knowledge about the sensitive attribute value of an individual before seeing the released table. Li et al. [30] proposed to preserve relationship privacy between two users one of whom can be identified in released network data.

Kavianpour et al. [31] proposed an integrated algorithm that takes the advantages of K-anonymity and l-diversity algorithm then evaluated the effectiveness of the combined strengths. Proposed algorithm has been able to increase the level of privacy for social network users by anonymizing and diversifying disclosed information.

TABLE I. OTHER PRIVACY PRESERVING TECHNIQUES IN OSN

Year	Author	Brief
2008	Zhou et al. [21]	Reviewed existing anonymization techniques for privacy preserving publishing of social network data.
2008	Guha et al. [22]	Encryption has been used to provide privacy and only authorized users can decode and decrypt the result.
2008	Blosser et al. [23]	Proposed protocols to create and interact with privacy preserving collaborative social networks that combines small networks together while retaining the purity of data for the owners.
2008	Campan et al. [24]	Greedy approach to optimize utility using the attribute and structural information simultaneously has been used. Structural Information loss has been introduced. SANGREEA (Social Network Greedy Anonymization).
2008	Zheleva et al[25]	How to preserve sensitive relationships.
2009	Ford et al. [26]	A new algorithm for enforcing p-sensitive k-anonymity on social network data based on greedy clustering approach was proposed.
2009	Narayanan et al. [27]	Developed Re-identification algo for anonymized graphs. Validated for Flickr and Twitter.
2009	Lijie et al. [28]	Studied link identification attack in which the adversary attacks using linking probability, *t*-confidence has been proposed. Dataset: EPINON, COA
2009	Ying et al. [29]	Considered edge re-identification attacks when the adversary has no background knowledge Dataset: Enron, Email,Polblogs.
2009	Tootoonchian et al.[30]	Presented Lockr, a system that improves the privacy of centralized online system like Flickr and decentralized online content sharing systems like BitTorrent.
2009	Fong et al. [31]	Proposed an access control model that generalizes the privacy preservation mechanism of Facebook.
2010	Tang et al. [32]	Introduced KNN and EBB algorithm for constructing generalized subgraphs before sharing social network with other parties and mechanism to integrate generalized information
2010	Lan et al. [33]	Proposed an approach for preserving privacy of social networks which can be represented as bipartite graphs. Synthetic dataset
2010	Ding et al. [34]	Presented a systematic review of the existing de-anonymization attacks in online social networks.
2010	Sun et al. [35]	Proposed a privacy-preserving method for sharing data in social networks, with efficient revocation for preventing a contact's access right to the private data once the contact is removed from the social group and can be used as a plug-in for Facebook.
2010	Beach et al. [36]	q-Anon model has been presented to measure the probability of an attacker to identify unknown information from a social network API with the assumption that the data being protected may already be public.
2010	Zhu et al. [37]	Proposed a collaborative framework for access control in social networks through an innovative key management.
2010	Wu et al. [38]	Categorized the existing anonymization methods on simple graphs in 3 main categories: K-anonymity based privacy preservation via edge modification, probabilistic privacy preservation via edge randomization, privacy reservation via generalization.
2011	Zheleva et al. [40]	Surveyed literature on privacy in social networks, Possible privacy breaches has been defined and possible privacy attacks..
2012	Fire et al. [41]	Developed Social Privacy Protector, software which aims to improve the security & privacy of Facebook users.
2013	Tassa et al. [43]	The first study of privacy preservation in distributed social networks which shown to outperform SaNGreeA algorithm which is leading algorithm for achieving anonymity in networks by means of clustering
2013	Heathely et al. [44]	Examined that friendship links and details altogether provide better predictability than details alone, effect of removing details and links in preventing sensitive information leakage has been explored.
2013	Cheng et al. [45]	Proposed a framework to provide users controls over how third party applications can access their data and activities in social networks while still retaining the functionality of third party applications.

IX. RESEARCH DIRECTIONS

Following are the few inferences drawn from literature survey:

1) To preserve usefulness(utility) of anonymized data is an important aspect while applying techniques for privacy

preservation. So, there is a need to develop methodologies that can quantitatively measure utility of data. There is need to evaluate various techniques in terms of tradeoff between privacy and utility.

*2)*Many algorithms like k-anonymity, L-diversity, integrated approach of k-anonymity & L-diversity have been developed for preserving privacy of social network user data but existing techniques leads to substantial information loss.

3) Anonymization techniques have been developed for one time released network data. But many applications require publishing data periodically so there is a need to develop techniques that can preserve privacy of dynamic releases.

4) Techniques are available for preserving privacy in case of distributed tabular data e.g. [59]. However, in case of social network distributed privacy preserving techniques are not well reported in literature except [56].

5) Existing privacy preserving approaches for social networks have been evaluated using either small datasets or synthetic datasets. There is need to conduct empirical experiments on large datasets.

6) There is no existing technique which can prevent homogeneity attacks, background knowledge attacks, attacks arising due to distance between sensitive values.

X. CONCLUSION

Emergences of social networks and increasing participation of people in activities in these sites have attracted other parties to have access to information or to the results of analyzing it. Potential access to the private data of users, such as profiles and contact lists, and possible misuse of such information is viewed as the highest privacy exposure. People understand the importance of privacy in social network that's why researchers are working hard on providing different approaches to protect the user's personal data from adversaries. To hide the data, various Anonymization techniques are discovered and we surveyed a few recent studies on anonymization techniques but still the adversaries are able to read the communication between users through traffic analysis. Although privacy preserving data publishing and analysis techniques in relational data have been well explored, the research and development of anonymization techniques on social network data is still in its infancy. As social network data is much more complicated than relational data, privacy preserving in social networks is much more challenging and needs many serious efforts in the near future. Techniques like K-anonymity, L-diversity, integrated K-anonymity L-diversity have been used till now but these techniques lead to substantial information loss. So, there is a scope of improvement of the techniques that provide privacy preservation with minimum information loss and better utility of released data. Particularly, modeling adversarial attacks and developing corresponding privacy preservation strategies are critical. In this survey we focused on the issue of protecting users' privacy and presented a categorization of different aspects of this area.

References

[1] R. Gross, A. Acquisti, and H. J. H. III, "Information revelation and privacy in online social networks," in *WPES*, 2005, pp. 71–80.

[2] K. P. N. Puttaswamy, A. Sala, and B. Y. Zhao, "Starclique: guaranteeing user privacy in social networks against intersection attacks," in *CoNEXT*, 2009, pp. 157–168.

[3] J. He, W. Chu, and V. Liu, *"Inferring Privacy Information from Social Networks*," Proc. Intelligence and Security Informatics, 2006.

[4] L. Sweeney, "k-anonymity: A model for protecting privacy," *International* Journal of Uncertainty, Fuzziness and Knowledge-Based Systems, vol. 10, no. 5, pp. 557–570, 2002.

[8] A.Annapoorani, Ms.P.Indira Priya, Inferring Private Information from Social Network Using Collective Classification, International Journal of Innovative Research in Computer and Communication Engineering, Volume 4, Special Issue 1, March 2014

[9] Jayasree Dasari, K.R.Koteswara Rao, "Sanitization Techniques for Protecting Social Networks from Inference Attacks," International Journal of Computer Science and Mobile Computing, IJCSMC, Vol. 3, Issue. 12, December 2014, pg.236 –244.

[10] Chethana Nair, Neethu Krishna, Siby Abraham, "Generalization Algorithm For Prevent Inference Attacks in Social Network Data", International Journal of Research in Computer and Communication Technology", IJRCCT, December 2014.

[11] Divya.R, B.Mahesh and R.Ushasree, "Data Implication Attacks on Social Networks with Data Sanitization", International Journal of Current Engineering and Technology, Vol.4, No.3 (June 2014).

[12] Brinal Colaco, Shamsuddin Khan, "Privacy Preserving Data Mining for Social Networks", International Journal of Engineering Research & Technology, Vol. 3 - Issue 8, August – 2014.

[13] Pooja Shelke, Ashish Badiye, "Social Networking: Its Uses and Abuses", Research Journal of Forensic Sciences, Maharashtra, (2013): 2-7.

[14] Ahmadinejad, Seyed Hossein, and Philip WL Fong. "On the feasibility of inference attacks by third-party extensions to social network systems." Proceedings of the 8th ACM SIGSAC symposium on Information, computer and communications security.ACM, 2013.

[17] R.Pranay, P. Pavan Kumar, "A Survey on Obstruction of Confidential Information Attacks in Social Networks", International Journal of Research in Information Technology, Volume 2, Issue 6, June 2014.

[20] B. Zhou, Jian Pei, Wo-Shun Luk, "A brief survey on anonymization techniques for privacy preserving publishing of social network data," ACM SIGKDD Explorations Newsletter, Vol. 10, pp. 12-22, 2008.

[21] Saikat Guha, Kevin Tang, Paul Francis, "NOYB: Privacy in Online Social Networks", in Proc. of first workshop on Online social networks WOSN'08, ACM New York, NY, USA, pp 49-54, 2008.

[22] Gary Blosser, Justin Zhan, "Privacy Preserving Collaborative Social Network", In Proc. of International Conference on Information Security and Assurance ISA 2008, Busan, pp, 543 - 548, 2008.

[23] Alina Campan, Traian Marius Truta, Nicholas Cooper, "P-Sensitive K-Anonymity with Generalization Constraints", In: Transactions on Data Privacy archive, Vol. 3 Issue 2, pp 65-89, 2010.

[24] Elena Zheleva, Lise Getoor, "Privacy in Social Networks: A Survey", In: Social Network Data Analytics, Springer US, pp 277-306, 2011.

[25] Roy Ford, Traian Marius Truta, and Alina Campan, "P-Sensitive K-Anonymity for Social Networks".

[26] A. Narayanan, V. Shmatikov, "De-anonymizing social networks", In Proc of 30th IEEE Symposium on Security and Privacy, Berkeley, CA, pp 173-187, 2009.

[27] Z. Lijie and Z. Weining, "Edge Anonymity in Social Network Graphs," in Proc. of International Conference on Computational Science and Engineering CSE '09, pp 1-8, 2009.

[28] X. Ying and X. Wu, "On link privacy in randomizing social networks," In: Advances in Knowledge Discovery and Data Mining, pp.28-39, 2009

[29] Amin Tootoonchian, Stefan Saroiu, Yashar Ganjali, Alec Wolman, "Lockr: Better Privacy for Social Networks", in Proc. of the 5th ACM International Conference on Emerging Networking Experiments and Technologies (CoNEXT), 2009.

[30] Philip W. L. Fong, Mohd Anwar, and Zhen Zhao, "A Privacy Preservation Model for Facebook-Style Social Network Systems", In: Computer Security - ESORICS 2009, Lecture Notes in Computer Science, Vol. 5789, 2009, pp 303-320, 2009.

[31] X. Tang and C.C. Yang, "Generalizing Terrorist Social Networks with K-Nearest Neighbor and Edge Betweenness for Social Network Integration and Privacy Preservation," In Proc. of IEEE International Conference on Intelligence and Security Informatics, 2010.

[32] Lihui Lan, Shiguang Ju Hua Jin, "Anonymizing Social Network using Bipartite Graph", In Proc. of International Conference on Computational and Information Sciences (ICCIS), Chengdu, pp 993 - 996, 2010.

[33] Xuan Ding, Lan Zhang, Zhiguo Wan, and Ming Gu, "A Brief Survey on De-anonymization Attacks in Online Social Networks", In Proc. of International Conference on Computational Aspects of Social Networks, Taiyuan, pp 611 - 615, 2010.

[34] Jinyuan Sun, Xiaoyan Zhu, and Yuguang Fang, "A Privacy-Preserving Scheme for Online Social Networks with Efficient Revocation", In Proc. of INFOCOM, IEEE, San Diego, CA, pp 1-9, 2010.

[35] Aaron Beach, Mike Gartrell, Richard Han, "q-Anon: Rethinking Anonymity for Social Networks", In Proc. of IEEE Second International Conference on Social Computing (SocialCom), Minneapolis, MN, pp 185 – 192, 2010.

[36] Yan Zhu, Zexing Hu, Huaixi Wang, Hongxin Hu, Gail-Joon Ahn, "A Collaborative Framework: for Privacy Protection in Online Social Networks", In Proc. of 6th International Conference on Collaborative Computing: Networking, Applications and Worksharing (CollaborateCom), Chicago, IL, pp 1 – 10, 2010.

[37] Xintao Wu, Xiaowei Ying, Kun Liu, ,Lei Chen, "A Survey of Privacy-Preservation of Graphs and Social Networks", In : Managing and Mining Graph Data, Advances in Database Systems, Vol. 40, pp 421-453, 2010.

[38] Christopher C. Yang , "Preserving privacy in social network integration with τ-tolerance", In Proc. of IEEE International Conference on Intelligence and Security Informatics (ISI), Beijing, pp 198 – 200, 2011 .

[39] Elena Zheleva, Lise Getoor, "Privacy in Social Networks: A Survey", In: Social Network Data Analytics, Springer US, pp 277-306, 2011.

[40] Michael Fire, Dima Kagan, Aviad Elishar, and Yuval Elovici, "Social Privacy Protector - Protecting Users' Privacy in Social Networks,",In Proc. of the Second International Conference on Social Eco-Informatics (SOTICS), Venice, Italy, 2012.

[41] Amirreza Masoumzadeh, James Joshi, "Preserving Structural Properties in Edge-Perturbing Anonymization Techniques for Social Networks", In: IEEE Transactions On Dependable And Secure Computing, Vol. 9, No. 6, pp 877-889, 2012.

[42] Tamir Tassa, Dror J. Cohen, "Anonymization of Centralized and Distributed Social Networks by Sequential Clustering", In: IEEE Transactions On Knowledge And Data Engineering, Vol. 25, No. 2, pp 311- 324, 2013.

[43] Raymond Heatherly, Murat Kantarcioglu, and Bhavani Thuraisingham, "Preventing Private Information Inference Attacks on Social Networks", In: IEEE Transactions On Knowledge And Data Engineering, Vol. 25, No. 8, pp 1849-1862, 2013.

[44] Yuan Cheng, Ravi Sandhu, "Preserving User Privacy from Third-party Applications in Online Social Networks, In Proc. of 22nd international conference on World Wide Web companion, Geneva, Switzerland, pp 723-728, 2013.