Available Online at www.ijcsmc.com

International Journal of Computer Science and Mobile Computing

A Monthly Journal of Computer Science and Information Technology

ISSN 2320–088X

IJCSMC, Vol. 4, Issue. 1, January 2015, pg.510 – 515

SURVEY ARTICLE

Survey Paper on Giving Privacy to Sensitive Labels

[1]Shimpli Dhale, [2]Anuja Ghotkar, [3]Ashish Bundele,

[4]Chiranjivi Kariya, [5]Sandesha Patil, [6]Priyanka Wandile

Computer Engineering Department

Bapurao Deshmukh Collage of Engineering, Sevagram

University of RTMNU at Nagpur

[1] dhaleshimpli@gmail.com, [2] ghotkaranuja@gmail.com, [3] ashishbundile@gmail.com

[4] chiranjivi.kariya@gmail.com, [5] sandesha_2410@rediffmail.com, [6] priyankawandile@gmail.com

Abstract: Sensitive Label and data handling are important issues for social network users. Ideally, access control enforcement should not depend on the social networking provider but should be under the control of the user. We propose a privacy protection scheme that not only prevents the disclosure of identity of users but also the disclosure of selected features in users' profiles also, a graph model where each vertex in the graph is associated with a sensitive label. It makes all requests for private data from third party applications (TPAs) explicit and enables a user to exert fine-grained control over what profile data can be accessed by them. Users can share their access control configurations for TPAs with their friends who can reuse and rate such configurations. The social networks are modeled as graphs in which users are nodes and features are labels. Labels are denoted either as sensitive or as non-sensitive.

Keywords - Sensitive Label, TPAs, Privacy Management, Cluster, fine-grained

I. INTRODUCTION

We must protect the sensitive lables of users on social network site. Here, the challenge is that to devise methods to publish social network data in a form that affords utility without compromising privacy. Various privacy models with the corresponding protection mechanisms that prevent both inadvertent private information leakage and attacks by malicious adversaries have been previously proposed in research. These early privacy models are mostly concerned with identity and link disclosure. The social networks are modeled as graphs in which users are nodes and social connections are edges. The threat Definition and protection

mechanisms leverage structural properties of the graph. This paper is motivated by the recognition of the need for a fine grain and more personalized privacy.

Users entrust social networks such as Facebook and LinkedIn with a wealth of personal information such as their age, address, current location or political orientation. We refer to these details and messages as features in the user's profile. We propose a privacy protection scheme that not only prevents the disclosure of identity of users but also the disclosure of selected features in users' profiles. An individual user can select which features of her profile she wishes to conceal. The social networks are modeled as graphs in which users are nodes and features are labels1. Labels are denoted either as sensitive or as non-sensitive. Figure 1 is a labeled graph representing a small subset of such a social network. Each node in the graph represents a user, and the edge between two nodes represents the fact that the two persons are friends. Labels annotated to the nodes show the locations of users. Each letter represents a city name as a label for each node. Some individuals do not mind their residence being known by the others, but some do, for various reasons. In such case, the privacy of their labels should be protected at data release. Therefore the locations are either sensitive or non-sensitive. (Labels are in red italic in Figure 1). The privacy issue arises from the disclosure of sensitive labels. One might suggest that such labels should be simply deleted.

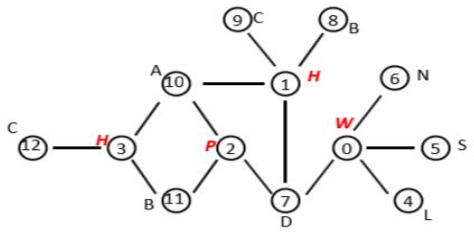

Figure 1. Example of the labeled graph representing a social network

One might suggest that such labels should be simply deleted. Still, such a solution would present an incomplete view of the network and may hide interesting statistical information that does not threaten privacy. A more sophisticated approach consists in releasing information about sensitive labels, while ensuring that the identities of users are protected from privacy threats. We consider such threats as neighborhood attack, in which an adversary find out sensitive information based on prior knowledge of the number of neighbors of a target node and the labels of these neighbors. In the example, if an adversary knows that a user has three friends and that these friends are in A (Alexandria), B (Berlin) and C (Copenhagen), respectively, then she can infer that the user is in H (Helsinki).As shown in the above fig 1.

We present privacy protection algorithms that allow for graph data to be published in a form such that an adversary cannot safely infer the identity and an adversary cannot safely infer the identity and sensitive labels of users. We consider the case in which the adversary possesses structural knowledge and label information. The algorithms that we propose transform the original graph into a graph in which any node with a sensitive label is indistinguishable from at least `−1 other nodes. The probability to infer that any node has a certain sensitive label (we call such nodes Sensitive nodes) is no larger than 1/`. For this purpose we design `-diversity-like model, where we treat node labels as both part of an adversary's background knowledge and as sensitive information that has to be protected. The algorithms are designed to provide privacy protection while losing as little information and while preserving as much utility as possible. In view of the tradeoff between data privacy and utility [16], we evaluate empirically the extent to which the algorithms preserve the original graph's structure and properties such as density, degree distribution and clustering coefficient. We show that our solution is effective, efficient and scalable while offering stronger privacy guarantees than those in previous research, and that our algorithms scale well as data size grows.

II. PROPOSED SYSTEM

In order to concern with social network privacy and data hiding is the major problem. According to existing system, module called profiling and friend request two model are already implemented. Profiling consist user (node) private/public data, friend request contain Third party application which they want to concern.

In our proposed system, we explorer the two new module name as, privacy option and new graph positioning. In real world there were many privacy option as, only me, friends of friends, private/public but, in order to used these option still there is problem to often tagged the information and this is obviously violate privacy and also experienced were more revelation.Here we are providing one more facility i, e make a group within group and finely exchange of data can be possible. Whenever to give privacy in group it will become more secure also help to make sure that uploading text or may called information were use by know third party. It is easy to plot graph in group of the friend relationship.

In the current system profiling and friend requesting these two options are already present and we should develop two new options that are privacy option and new graph positioning option.

Using privacy option user can select any disclosure of his profile which he wants to conceal. And using new graph positioning option we should create a cluster and mountain it on internally plotted graph so that our system work properly and there is no issue of privacy and security. In our proposed system we may add N-number of clients. The following figure shows the proposed system.

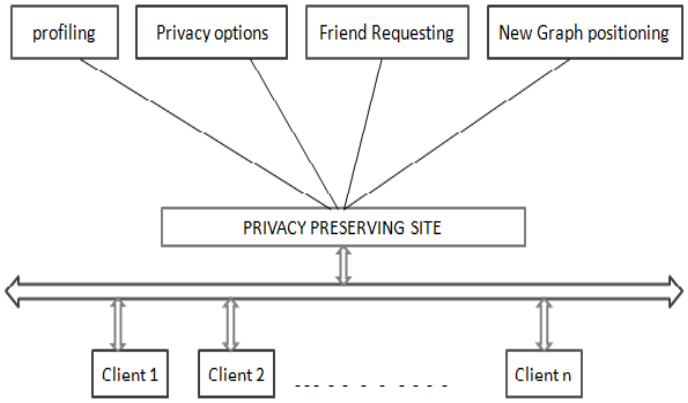

Figure 1 : Privacy preserving network

The social networks are modeled as graphs in which are nodes and features are labels1. Labels are denoted either as sensitive or as non-sensitive.

III. ALGORITHM

The main objective of the algorithms that we propose is to make suitable group- ing of nodes, and appropriate modification of neighbors' labels of nodes of each group to satisfy the l-sensitive-label-diversity requirement. We want to group nodes with as similar neighborhood information as possible so that we can change as few labels as possible and add as few noisy nodes as possible. We propose an algorithm, Global-similarity-based Indirect Noise Node (GINN) that does not attempt to heuristically prune the similarity computation as the other two algorithms, Direct Noisy Node Algorithm (DNN) and Indirect Noisy Node Algorithm (INN) do. Algorithm DNN and INN, which we devise first, sort nodes by degree and compare neighbourhood information of nodes with similar degree. Details about algorithm DNN and INN please refer to [15]

Larger value indicates larger similarity of the two neighbourhoods. Then nodes having the maximum similarity with any node in the group are clustered into the group till the group has ℓ nodes with different sensitive labels. Thereafter, the algorithm proceeds to create the next group. If fewer than ℓ nodes are left after the last group's formation, these remainder nodes are clustered into Existing groups according to the similarities between nodes and groups. After having formed these groups, we need to ensure that each group's members are indistinguishable in terms of neighbourhood information. Thus, neighbourhood labels are modified after every grouping operation, so that labels of
Nodes can be accordingly updated immediately for the next grouping operation. This modification process ensures that all nodes in a group have the same neighbourhood information.
The objective is achieved by a series of modification Operations. To modify graph with as low information loss as possible, we devise three modification operations: label union, edge insertion and noise node addition.
Label union and edge insertion among nearby nodes are preferred to node addition, as they incur less alteration to the overall graph structure. If there are nodes in a group still having different neighbourhood information, noise nodes with non-sensitive labels are added into the graph so as to render the nodes in group indistinguishable in terms of their neighbours' labels?

Global-Similarity-based Indirect Noise Node Algorithm

```
Input: graph G(V,E,L,Ls), parameter l;
Result: Modified Graph G'
1   while Vleft > 0 do
2       if |Vleft| ≥ l then
3           compute pairwise node similarities;
4           group G ← v1,v2 with Maxsimilarity;
5           Modify neighbors of G;
6           while |G| < l do
7               dissimilarity(Vleft,G);
8               group G ← v with Maxsimilarity;
9                 Modify neighbors of G without actually adding noisy nodes ;
10      else if |Vleft| < l then
11          for each v ∈ Vleft do
12              similarity(v,Gs);
13              GMax similarity ← v;
14            Modify neighbors of GMax similarity without actually adding noisy  nodes;
15  Add expected noisy nodes;
16  Return G'(V' ,E' ,L');
```

In this algorithm, noise node addition operation that is expected to make the nodes inside each group satisfies sensitive-label-diversity are recorded, but not performed right away. Only after all the preliminary grouping operations are performed, the algorithm proceeds to process the expected node addition operation at the final step. Then, if two nodes are expected to have the same labels of neighbors and are within two hops (having common neighbors), only one node is added. In other words, we merge some noisy nodes with the same label, thus resulting in fewer noisy nodes.

IV. LITERATURE REVIEW

2012 IEEE International Conference on Pervasive Computing and Communications Workshops[1] on Access Control in Decentralized Online Social Networks: Applying a Policy-Hiding Cryptographic Scheme and Evaluating Its Performance presented by Oleksandr Bodriagov, Gunnar Kreitz, and Sonja Buchegger suggested that Privacy concerns in online social networking services had prompted a number of proposals for decentralized online social networks (DOSN) that remove the central provider and aim at giving the users control over their data and who can access it. This was usually done by cryptographic means. Existing DOSNs used cryptographic primitives that hide the data but revealed the access policies. At the same time, there were privacy-preserving variants of those cryptographic primitives that did not reveal access policies. They were, however, not suitable for usage in the DOSN context because of performance or storage constraints.

33rd International Conference on Distributed Computing System Year 2012[2] on Message in a Sealed Bottle: Privacy Preserving Friending in Social Networks presented by Lan Zhangff, Xiang-Yang Li proposed that in this paper, they designed novel mechanisms, when given a preference-profile submitted by a user, that searched a person with matching-profile in decentralized multi-hop mobile social networks. There mechanisms were privacy-preserving: no participants' profile and the submitted preference-profile were exposed. There mechanisms established a secured communication channel between the initiator and matching users at the time when the matching user was found. There rigorous analysis shows that there mechanism was secured, privacy-preserving, verifiable, and efficient both in communication and computation. Extensive evaluations using real social network data, and actual system implementation on smart phones show that there mechanisms were significantly more efficient then existing solutions.

2011 IEEE International Conference[3] on Collaborative Privacy Management for Third-Party Applications in Online Social Networks presented by Pauline Anthonysamy, Awais Rashid proposed that Privacy control mechanisms for online social networks (OSNs) offered little by way of managing access to a user's personal information by third-party applications (TPAs). Most OSNs provide an "accept all or nothing" mechanism for managing permissions from TPAs to access a user's private data.

In this paper, they proposed an approach that makes all requests for private data from TPAs explicit and enables a user to exert fine-grained access control over what profile data can be accessed by individual applications. Equally importantly, there approach also allows users to share their access control configurations for TPAs with their friends who can reuse and rate such configurations. This was particularly beneficial to novice users or those new to a particular TPA or an OSN. They presented an implementation of there approach for managing privacy for third-party Facebook applications.

2011 IEEE International Conference[4] on Measuring Privacy Risk in Online Social Networks presented by Justin Becker, Hao Chen suggested in that paper the PrivAware, a tool to detected and report unintended information loss in online social networks. There goal is to provide a rudimentary framework to identify privacy risk and provide solutions to reduce information loss. The first instance of the software is focused on information loss attributed to social circles. In subsequent released they intended to incorporate additional capabilities to captured ancillary threat models. From there initial results, they quantify the privacy risk attributed to friend relationships in Facebook. They show that for each user in our study a majority of their personal attributes can be derived from social contacts. Moreover, they present results denoting the number of friends contributing to a correctly inferred attribute.

2012 IEEE International Conference[5] on Enforcing Access Control in Social Network Sites presented by Filipe Beato, Markulf Kohlweiss, and Karel Wouters proposed that l, SNS platform-independent solution, for social network users to control their data. We develop con- cepts that are general enough to describe access control restrictions for different SNS platforms. Our architecture uses encryption to enforce ac- cess control for users' private information based on their privacy prefer- ences. We have implemented our model as a Firefox extension.

2009 IEEE International Conference [6] on Preserving Privacy in Social Networks against Neighborhood Attacks presented by Bin Zhou Jian Pei, they take an initiative towards preserving privacy in social network data. They identify an essential type of privacy attacks: neighborhood attacks. If an adversary had some knowledge about the neighbors of a target victim and the relationship among the neighbors, the victim may be re-identified from a social network even if the victim's identity is preserved using the conventional anonymization techniques. They show that the problem is challenging, and present a practical solution to battle neighborhood attacks. The empirical study indicates that anonymized social networks generated by our method can still be used to answer aggregate network queries with high accuracy.

2009 IEEE International Conference [7] on Privacy-Preserving P2P Data Sharing with OneSwarm presented by Tomas Isdal, Thomas Anderson suggested a new design point in trade off between privacy and performance. They describe the design and implementation of a new P2P data sharing protocol, called OneSwarm, that provides users much better privacy than BitTorrent and much better performance than Tor or Freenet. A key aspect of the ones warm design is that users have explicit configurable control over the amount of trust they place in peers and in the sharing model for their data: the same data can be shared publicly, anonymously, or with access control, with both trusted and untrusted peers. One-

Swarm's novel lookup and transfered techniques yield a median factor of 3.4 improvements in download times relative to Tor and a factor of 6.9 improvements relative to Freenet. OneSwarm is publicly available and has been downloaded by hundreds of thousands of users since its released.

2011 IEEE International Conference[8] on Multiparty Access Control for Online Social Networks: Model and Mechanisms presented by Hongxin Hu, Gail-Joon Ahn proposed in this paper, that we proposed an approach to enable the protection of shared data associated with multiple users in OSNs. They formulate an access control model to capture the essence of multiparty authorization requirements, along with a multiparty policy specification scheme and a policy enforcement mechanism. Besides, they present a logical representation of our access control model which allows them to leverage the features of existing logic solvers to perform various analysis tasks on our model. They also discuss a proof-of-concept prototype of our approach as part of an application in Facebook and provide usability study and system evaluation of our method.

IEEE Transaction on parallael and distrubuted system vol:24 No:12year 2011[9], on Sybil Defender: Defend Against Sybil Attacks in Large Social Networks presented by WeiWei∗, Fengyuan Xu∗, Chiu C. Tan†, Qun Li∗ suggest that they present SybilDefender, a scheme that leverages the network topologies to defend against sybil attacks in large social networks. There evaluation shows that SybilDefender can correctly identify the sybil nodes even when the number of sybil nodes introduced by each attack edge approaches the theoretically detectable lower bound, and it can effectively detect the sybil community surrounding a sybil node with different sizes and structures.

2008 IEEE International Conference[10] CoPE: Enabling Collaborative Privacy Management in Online Social Networks presented by Anna Squicciarini, Xiaolong (Luke) Zhang, suggested that This paper presents a prototype system and a preliminary study on the perceptions of the usefulness and usage of the system. In addition to extending the design space of image-related privacy management, this research also suggests a general approach for privacy protection in online social networks (OSNs).

V. CONCLUSION

In this paper we have investigated the protection of private label information in social network data publication. We consider graphs with rich label information, which are categorized to be either sensitive or non-sensitive. We assume that adversaries possess prior knowledge about a node's degree and the labels of its neighbors, and can use that to infer the sensitive labels of targets. We suggested a model for attaining privacy while publishing the data, in which node labels are both part of adversaries' background knowledge and sensitive information that has to be protected. We accompany our model with algorithms that transform a network graph before publication, so as to limit adversaries' confidence about sensitive label data. Our experiments on both real and synthetic data sets con- firm the effectiveness, effciency and scalability of our approach in maintaining critical graph properties while providing a comprehensible privacy guarantee.

REFERENCES

[1] L. A. Adamic and N. Glance. The political blogosphere and the 2004 U.S. election: divided they blog. In LinkKDD, 2005.

[2] L. Backstrom, C. Dwork, and J. M. Kleinberg. Wherefore art thou R3579X?: anonymized social networks, hidden patterns, and structural steganography. Com- mun. ACM, 54(12), 2011.

[3] S. Bhagat, G. Cormode, B. Krishnamurthy, and D. S. and. Class-based graph anonymization for social network data. PVLDB, 2(1), 2009.

[4] A. Campan and T. M. Truta. A clustering approach for data and structural anonymity in social networks. In PinKDD, 2008.

[5] J. Cheng, A. W.-C. Fu, and J. Liu. K-isomorphism: privacy-preserving network publication against structural attacks. In SIGMOD, 2010.

[6] G. Cormode, D. Srivastava, T. Yu, and Q. Zhang. Anonymizing bipartite graph data using safe groupings. PVLDB, 19(1), 2010.

[7] S. Das, ¨O. Egecioglu, and A. E. Abbadi. Anonymizing weighted social network graphs. In ICDE, 2010.

[8] A. G. Francesco Bonchi and T. Tassa. Identity obfuscation in graphs through the information theoretic lens. In ICDE, 2011.

[9] M. Hay, G. Miklau, D. Jensen, D. Towsley, and P. Weis. Resisting structural re-identification in anonymized social networks. PVLDB, 1(1), 2008.

[10] Y. Li and H. Shen. Anonymizing graphs against weight-based attacks. In ICDM Workshops, 2010.

[11] K. Liu and E. Terzi. Towards identity anonymization on graphs. In SIGMOD, 2008.

[12] L. Liu, J. Wang, J. Liu, and J. Zhang. Privacy preserving in social networks against sensitive edge disclosure. In SIAM International Conference on Data Mining, 2009.

