Sensitive Label Privacy Protection on Social Network Data

Parveenbano Khan,Mr.Rajat Gabhne ,Shilpa Bisen,Mr.Hitesh Nirwan,Sneha Nagarkar

vertex of graph is linked with sensitive

Abstact **: This paper is motivated by the recognition of the need for a finer grain and more personalized privacy in data publication of social networks. 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 profiles she wishes to conceal. 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. We treat node labels both as background knowledge an adversary may possess, and as sensitive information that has to be protected. We present privacy protection algorithms that allow for graph data to be published in a form such that an adversary who possesses information about a node's neighborhood cannot safely infer its identity and its sensitive labels.**

INTRODUCTION

Protecting the privacy of personal information is one of the biggest challenges facing website developers, especially social network providers. Several researchers have discussed the issue of privacy. In today's internet determined the people we have witnessed the rapid growth of online social networking sites (OSN) as well as their integration into our everyday life. OSN such as Facebook (FB), Twitter, LinkedIn, Myspace etc. now represent a fundamental shift in the way that we communicate in our personal and working live. With the sharing nature of OSN's and the sites' control of posted information and personal

The relationships, concerns have developed regarding trust and privacy issues within social networking. Mainly, the data may contain sensitive information about individuals that cannot be disclosed without compromising their confidentiality. This paper we use AES algorithm to encrypt labels or private information. We develop a new algorithm (heuristic search) by adding noise nodes into the original graph without change original graph drastically, and provide security of each user & its sensitive data.

The publication of social network data entails a privacy threat for their users. Sensitive information about users of the social networks should be protected. The challenge is to devise methods to publish social network data in a form that affords utility without compromising privacy. Previous research has pro-posed various privacy models with the corresponding protection mechanisms that prevent both inadvertent private information leakage and attacks by malicious adversaries. 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 definitions and protection Sensitive Label Privacy Protection on Social Network Data mechanisms leverage structural properties of the graph. This project is motivated by the recognition of the need for a finer 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 labels. Labels are denoted either as sensitive or as non-sensitive. 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.

Scope-

- Privacy is one of the major concerns when publishing or sharing social network data forsocial science research and business analysis.
- Privacy models similar to k-anonymity to prevent node reidentification through structureinformation. However, even when these privacy models are enforced, an attacker maystill be able to infer other private information if a group of nodes largely share the samesensitive labels.
- Proposed approach definethe k-degree-l-diversity anonymity model that considers the protectionof structural information as well as sensitive labels of individuals.
- Proposed method will produce anonymization methodology based on adding noise nodes.It develop a new algorithm by adding noise nodes into the original graph with the considerationof introducing the least distortion to graph properties

METHODOLOGY

System Design:

Data Flow Diagram / Sequence Diagram / Component Diagram:

The DFD is also called as bubble chart. It is a simple graphical formalism that can be used to represent a system in terms of the input data to the system, various processing carried out on these data, and the output data is generated by the system.

We propose a k-degree-l-diversity model for privacy preserving social network data publishing. We implement distinct K-degree, l-diversity and Anonymization. We design an algorithm to preserving privacy of user on social network.

. We are using heuristic search strategy that will search the input phase with minimum overhead. With give approximate answer within polynomial time. We give a rigorous analysis of the theoretical bounds on the minimum number of noise nodes added. The software industry includes many different processes, for example, analysis, development, maintenance and publication of software. This industry also includes software services, such as training, documentation, and consulting.

Extensive experimental results demonstrate that the add minimum noise node AES algorithms and heuristic strategy can achieve a better result than the previous work using edge editing only and noise node adding attractive direction to study clever algorithms which can reduce the reduction of noise nodes with anonymization and diversity. Privacy is key matter when sharing social network data for organization and personal. It is necessary of today's large use of social network to provide privacy and security of private information. We present new technique that will reduce noise nodes in our model

1.Add minimum no of nodes & improve anonymization technique.

2.We implementing privacy-preserving approach.

- **Data Flow Diagram**

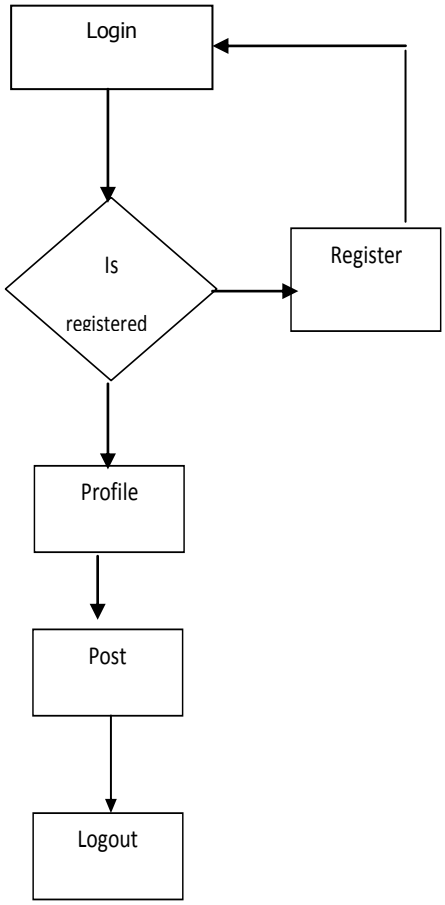


Fig .user flow

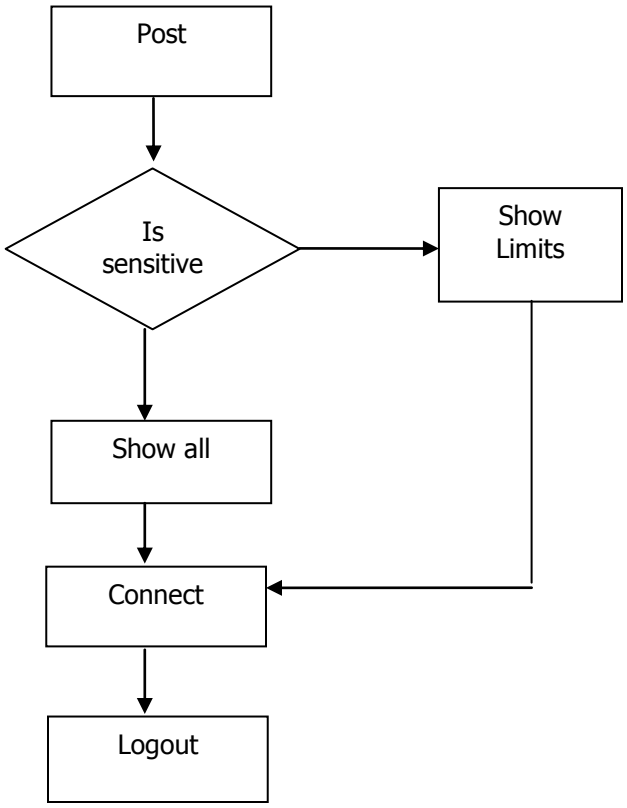


Fig .Post flow

Algorithm GINN:

The algorithm starts out with group formation, during which all nodes that Have not yet been grouped are taken into consideration, in clustering-like fashion. In the _rst run, two nodes with the maximum similarity of their neighborhood labels are grouped together. Their neighbor labels are modied to be the same immediately so that nodes in one group always have the same neighbor labels.

For two nodes, v1 with neighborhood label set (LSv1), and v2 with neighborhood label set (LSv2), we calculate neighborhood label similarity (NLS) as follows:

$$NLS(v1; v2) = \frac{jLSv1 \setminus LSv2\ j}{jLSv1\ [\ LSv2\ j}$$

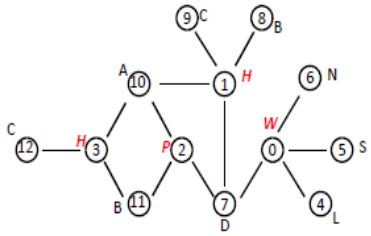
Larger value indicates larger similarity of the two neighborhoods.

Then nodes having the maximum similarity with any node in the group are clustered into the group till the group has ` nodes with different sensitive labels.

Thereafter, the algorithm proceeds to create the next group. If fewer than ` nodes are left after thelast 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 mem-bers are indistinguishable in terms of neighborhood information. Thus, neigh-borhood labels are modi_ed after every grouping operation, so that labels of nodes can be accordingly updated immediately for the next grouping opera-tion. This modi_cation process ensures that all nodes in a group have the same neighborhood information. The objective is achieved by a series of modi_cation operations. To modify graph with as low information loss as possible, we devise three modi_cation operations: label union, edge insertion and noise node addi-tion. Label union and edge insertion among nearby nodes are preferred to node addition, as they incur less alteration to the overall graph structure.Edge insertion is to complement for both a missing label and insu_cient degree value. A node is linked to an existing nearby (two-hop away) node with that label.

Way of Social Network



Example of the labeled graph representing a social network

Conclusion:

- We suggested a model for attaining privacy while publishing the data.
- In which node, labels are both part of adversaries background knowledge.
- Sensitive information that has to be protected.

References

[1] Mingxuan Yuan, Lei Chen, "Protecting Sensitive Labels in Social Network Data",IEEE transaction on knowledge and data engineering, Vol. 25, No. 3,March 2013.

[2] Chongjing Sun, Philip S. Yuz, Xiangnan Kong and Yan Fu, "Privacy Preserving SocialNetwork Publication Against Mutual Friend Attacks," arXiv:1401.3201v1 [cs.DB] 11 Oct2013.

[3] Mr. A.Stalin Irudhaya Raj, Ms. N.Radhika, "Securing Sensitive Information in Social NetworkData ," A.Stalin Irudhaya Raj et al, International Journal of ComputerScience and Mobile Applications, Vol.2 Issue. 1, January- 2012.

[4] Lijie Zhang andWeining Zhang , "Privacy Protection of Social Network Graphs," 2010.

[5]. C.-H. Tai, P. S. Yu, D.-N. Yang, and M.-S. Chen. Privacy-preserving social network publication against friendship attacks. In SIGKDD, 2011.