SAJET

South Asian Journal of Engineering and Technology Vol.2, No.17 (2016) 133–137

ISSN No: 2454-9614

Session Based Protection Identification Model for Dynamic Network using Threshold Monitoring Algorithm

T. Karthika, T. Sathishkumari*

Department of Computer Science and Engineering Vellalar College of Engineering and Technology Erode,Tamilnadu, India.
b

*Corresponding Author: T. Karthika
E-mail: karthigathangaraj2@gmail.com

Received: 10/11/2015, Revised: 15/12/2015 and Accepted: 07/03/2016

Abstract

This p0roject considers the protection of vertex and community identities of individuals in a dynamic network. A simple approach for this problem is to anonymize each release to satisfy some privacy model before a network is published. However, due to the lack of consideration in sequential releases, adversaries can have chances to launch attacks and get advantages by gathering victim's information continuously and comparing the multiple releases. In this project to show that an adversary can successfully infer a victim's vertex identity and community identity by the knowledge of degrees within a time period. In Social networks model the social activities between individuals. In light of useful information from such dynamic networks, here is a continuous demand for privacy preserving data sharing with analyzers, collaborators or customers. This project addresses the privacy risks of identity disclosures in sequential releases of a dynamic network. To prevent the privacy breaches, this project proposes novel K^w-structural diversity anonymity, where k is an appreciated privacy level and w is a time period that an adversary can monitor a victim to collect the attack knowledge. This project also presents a heuristic algorithm for generating releases satisfying K^w-structural diversity anonymity so that the adversary cannot utilize his knowledge to re-identify the victim and take advantages.

Reviewed by **ICETSET'16** *organizing committee*

.

1. Introduction

Today, mobile users interact with each other and share files via an infrastructure formed by geographically distributed base stations. However, users may find themselves in an area without wireless service (e.g., mountain areas and rural areas). Moreover, users may hope to reduce the cost on the expensive infrastructure network data.

P2P file sharing model makes large-scale networks a blessing instead of a curse, in which nodes share files directly with each other without a centralized server. Wired P2P file sharing systems like Bit Torrent have already become a popular and successful paradigm for file sharing among millions of users.

The successful deployment of P2P file sharing systems and the aforementioned impediments to file sharing

in MANETs make the P2P file sharing over MANETs (P2P MANETs in short) a promising complement to current infrastructure model to realize pervasive file sharing for mobile users.

The mobile digital devices are carried by people that usually belong to certain social relationships. So this thesis is focused on the P2P file sharing in a disconnected MANET community consisting of mobile users with social network properties. In such a file sharing system, nodes meet and exchange requests and files in the format of text, short videos, and voice clips in different interest categories.

One typical scenario is a course material sharing system in a college/school campus. Such scenarios ensure for the most that nodes sharing the same interests carry corresponding files and meet regularly.

A collective of nodes is referred that share common interests and meet frequently as a community. According to P3, a node has high probability to find interested files in its community. If this fails, based on P1, the node can rely on nodes that frequently travel to other communities for file searching. Thus, the community construction algorithm is proposed to build communities to enable efficient file retrieval.

According to P1, a node role assignment algorithm is proposed that takes advantage of node mobility for efficient file searching. The algorithm designates a stable node that has the tightest connections with others in its community as the community coordinator guide intra community searching. For each known foreign community, a node that frequently travels to it is designated as the community ambassador for intercommunity searching.

SPOON is novel in that it leverages social network properties of both node interest and movement pattern. First, it classifies common-interest and frequently encountered nodes into social communities. Second, it considers the frequency at which a node meets different interests rather than different nodes in file searching. It chooses stable nodes in a community as coordinators and highly mobile nodes that travel frequently to foreign communities as ambassadors. Such a structure ensures that a query can be forwarded to the community of the queried file quickly. SPOON also incorporates additional strategies for file perfecting, querying completion and loop-prevention, and node churn consideration to further enhance file searching efficiency.

2. Related Work

Morvarid Sehatkar and Stan Matwin Sequence data mining has many interesting applications in a large number of domains including finance, medicine, and business. Sequence data often contains sensitive information about individuals and improper release and usage of this data may lead to privacy violation. In this paper, we study the privacy issues in publishing multidimensional sequence data. We propose an anonymization algorithm, using hierarchical clustering alignment techniques, which is capable of preventing both identity disclosure and sensitive information inference. The empirical results show that our approach can effectively preserve data utility as much as possible, while preserving privacy.

Recent advances in information technology have enabled public organizations and corporations to collect and store huge amounts of individuals' data in data repositories. Such data are powerful sources of information about an individual's life such as interests, activities, and finances. Corporations can employ data mining techniques

to extract useful knowledge from individuals' data and exploit this knowledge to improve their strategic decision making, enhance business performance, and improve services.

As a result, the demand for collecting and sharing data has been rapidly increased. Among the types of individuals' data, event sequence data mining has many interesting applications in a large number of domains. Sequence data mining enables us discover behaviour patterns of individuals through temporal activities.

Such knowledge is precious for planning, detecting behavioural changes, and commercial purposes. Longitudinal medical records of patients can be used to analyze patients' reactions to a new drug or to support a diagnosis. However, despite all benefits of analyzing event sequence data, this data often contain sensitive information and may violate privacy of individuals if published. In event sequence data, every event may have a number of attributes that act as quasi-identifiers (QIs). Due to temporal correlation among the events of each sequence, in addition to the values of QIs within an event, any combination of QIs values across events along with the temporal information about these values might lead to privacy breach. Including admission year (AdmYr), ZIP code, number of days since the first visit in each year (DSFC), and the length of stay in the hospital (LOS), which all act as QIs, as well as one sensitive attribute diagnosis.An adversary with some background knowledge about visits of a target individual is able to launch two types of privacy attacks: identity disclosure and attribute disclosure.

Praveena Devi K. and Sri Priya P Social networks have become the universal consumer phenomena and have emerged with increasing popularity nowadays. The amount of network data grows enormously due to the increase of networking websites. The development of social networks has led to the increasing demand for the protection of privacy in publishing the social network data as the social network sites are accumulated with large amount of sensitive information of individuals. So preserving the privacy in publishing social network data has become an important concern these days. The recent rise in popularity of social networks has created large quantities of data about interactions. Such data may have many secret details about individuals so anonymization has to be done earlier to attempts to make the data more extensively available for scientific research. Data anonymization protects the sensitive information from unsolicited access and ensures privacy protection.

Preserving privacy in publishing social network data has become more essential these days because the sensitive information of the individuals may be disclosed. With some local knowledge about the individuals in social network, an adversary can attack the privacy of the victims easily. Once the identity of an individual is leaked, then automatically the individual is re-identified and the corresponding relationship with others and their sensitive data are also exposed. Hence clever adversaries usually try to launch identity disclosure attacks on targeted victims. Privacy preserving of the data mining has continuously gained attention in the data mining community in studying the difficulty of the problem and proposing various approaches for data anonymization. It is one of the significant areas of data mining that aims to provide security for sensitive information from unwanted disclosure.

Many techniques for privacy preserving data mining like statistical, randomization methods, k-anonymity model, l- diversity and etc., have come up over the last decade. But most of the existing techniques holds good only

for relational databases. The main drawback is that the earlier studies on privacy protection concentrated only on static network. They can deal with simple graphs only, and cannot be applied to large graphs. Whereas the cryptographic techniques holds good to deal with the privacy of a file system or a network.

3. Social Network

The users are added in the table. Then friend users are added for the users. So a social network is constructed with users as nodes and their relationship to other users as edges. So a graph is being constructed. This will become a time-stamped graph of time t=0.The friend users are added and removed for the users. The social network is modified with users as nodes and their relationship to other users as edges. This will become a time-stamped graph at time $t_{(t=1)}$. Whenever the users added, removed or relationship changed, the time-stamped graph is assumed that it is being generated at time t_{t+1}.

4. Anonymization Algorithms

The CS-Table is constructed. The CS-Table is a table consisting of three columns vertex v, $v \in V$, the degree sequence and the sequence of multi community identities. The CS-Table is built according to the degree sequences of vertices. The table is not built at once since the anonymization of a dynamic graph is a continuous process. The construction of the CS-Table is achieved together with the anonymizations of the first w releases. This involves in sorting all the vertices. When anonymizing G^1, the CS-Table is simultaneously modified. Then, each vertex is in a k-shielding group after the anonymization of G^1. Later given G^2, the vertex information of G^2 is attached behind the corresponding records.

From now on, rather than sort all vertices, we only need to sort the vertices in the same groups since the vertices are already in decreasing order of their previous degrees. Thus, the sorting time can be reduced. After the anonymization of G^2, a similar process is executed until G^w is anonymized.

The CS-Table incremental update is made. When a new snapshot G^t comes, the CS-Table has to be updated to maintain the correct information corresponding to the concerned period w. For this purpose, before attaching the vertex information of G^t to the CS-Table.

It is required to remove the information of b G'^{t-w} and re-sort the vertices according to the degree sequences instead. Therefore, rather than re-sort all vertices, we can re-sort the k-shielding consistent groups.

5. Anonymization Process

The anonymization process is carried out. For anonymization, three operations are used to adjust the degree of a vertex.

1) Operation Adding Edge connects two vertices in the same community. The connection between different communities is forbidden because it may destroy the distinction between different communities. The reason for adding edges alone but not removing edges is that removing edges can severely destroy the community structural

information of a graph than adding edges.

2) Operation Redirecting Edge increases the degree of a vertex v by changing the not-yet-anonymized end-point of a previously added edge to vertex v, i.e., $E = E / (x, y) \cup (x, v)$, where vertex x is anonymized, y has not yet been anonymized. This operation is satisfactory since it does not change the degree of an anonymized vertex and allows increasing the degree of a vertex without adding additional edges.

3) Operation Adding Vertex connects a vertex and an additional fake vertex, and Note that adding vertices is usually less preferred because it changes the vertex set of a graph and causes more information distortion. So Adding Edge and Redirecting Edge is always prior to Adding Vertex.

6. Conclusion

Through this project, the problem of identity protection is solved. In addition, the application required less working experience in systems to run the software. The application is tested well so that the end users use this software for their whole operations. It is believed that almost all the system objectives that have been planned at the commencement of the software development have been net th and the implementation process of the project is completed.

A trial run of the system has been made and is giving good results the procedures for processing is simple and regular order. The process of preparing plans been missed out which might be considered for further modification of the application. The project effectively stores and retrieves the records from the cloud space database server. The records are encrypted and decrypted whenever necessary so that they are secure.

References

[1] Bhadri Raju, and Chinta Someswara Rao (2014),'Privacy Measure for Publishing the Data- A Case Study', ISSN: 2278-3075, Vol.3, Issue 10, pp.65-67
[2] Chih-Hua Tai, Peng Jui Tseng, Philip Yu S., and Ming-Syan Chen(2014),'Identity Protection In Sequential Releases Of Dynamic Network' Transactions On Knowledge And Data Engineering, Vol.26, No.3, pp.635-651.
[3] Dr. Jeffrey Zients, Scott Crosier and John Snow (2013), 'Big Data: Seizing Opportunities, Preserving Values',: The London Cholera Epidemic of 1854, Center for Spatially Integrated Social Science, University of California, pp.1-85.
[4] Lakshmi M. and Mala R., Rajakumar R... (2014), 'Reduce cost and efficient access for , cloud storage Using Intermediate Cloud Datasets', Vol.4, Issue.9, ISSN 2250-3153,pp.1-5.
[5] Mayil S. and Vanitha M.(2014) 'A Survey on Privacy Preserving Data Mining Techniques', Vol.5, pp.6054-5056
[6] Morvarid Sehatkar and Stan Matwin (2014), 'Clustering-based Multidimensional Sequence Data anonymization', the Workshop Proceedings of the EDBT/ICDT Joint Conference on ISSN 1613-0073,pp.385-389.
[7] Praveena Devi K., Sri Priya P.(2015), 'Data Anonymization For Privacy Protection', Vol.3, Issue 1, pp.33-34.
[8] Sabrina De Capitani di Vimercati, Sara Foresti ,Giovanni Livraga, And Pierangela amarati(2013), 'Protecting Privacy In Data Release',Journal Of Privacy Technology.
[9] Suma Reddy and Shilpa G.V. (2015), 'Privacy Preserving Publishing of Social Network Data Privacy and Big Data Mining', International Conference on Advances in Computer and Communication Engineering, ISSN 2250-2459 Vol.5, Special Issue 2,pp.126-135.
[10] Sweta Parmar and Tamanna Kachwala (2014), 'An Approach for Preserving Privacy in Data Mining',Vol.4, Issue 9 ,pp.370-373.