www.ijiarec.com

ISSN:2348-2079

Volume-6 Issue-2

International Journal of Intellectual Advancements and Research in Engineering Computations

Detect Malware and Search Rank Fraud in Google Play

[1]Mr.S.Sambasivam, M.C.A., MPhil., Associate professor,
[2]Ms. M.Viveka Final MCA,
Department of MCA,Nandha Engineering College(Autonomous),Erode-52.
E-Mail ID: sammy2173@gmail.com, viveka1324@gmail.com

Abstract—Theprotection 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. 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.The privacy risks of identity disclosures in sequential releases of a dynamic network.It also presents a heuristic algorithm for generating releases satisfying S-K^{w} -structural diversity anonymity so that the adversary cannot utilize his knowledge to re-identify the victim and take advantages. The evaluations on both real and synthetic data sets show that the proposed algorithm can retain much of the characteristics of the networks while confirming the privacy protection.

Index Terms-Privacy Protection, Anonymity,clustering.

I. INTRODUCTION

The social identity approach explains trust in strangers as a function of group-based stereotypes or in-group favouringbehaviours based on salient group memberships. With regard to ingroup favoritism, people generally think well of strangers but expect better treatment from in-group members in comparison to out-group members. This greater expectation then translates into a higher propensity to trust an in-group rather than out-group member.It has been pointed out that it is only advantageous to form such expectations of an in-group stranger if they too know the group membership of the recipient. There is considerable empirical activity related to the social identity approach.

Allocate studies have frequently been employed to understand group-based trust in strangers. They may be operation as unilateral or bilateral relationships of exchange. Any amount given would be tripled and the receiver would then decide on whether they would return the favour by giving money back to the sender.

II.RELATED WORKS

*Graham Cormode et al [1]*describe private data often come in the form of associations between entities, such as customers and products bought from a pharmacy, which are naturally represented in the form of a large, sparse bipartite graph. As with tabular data, it is desirable to be able to publish anonymized versions of such data, to allow others to perform ad hoc analysis of aggregate graph properties. However, existing tabular anonymization techniques do not give useful or meaningful results when applied to graphs: small changes or masking of the edge structure can radically change aggregate graph properties. We introduce a new family of groupings. These groupings preserve the underlying graph structure perfectly, and instead anonymize the mapping from entities to nodes of the graph. We identify a class of "safe" *(k,i)*-groupings that have provable guarantees to resist a variety of attacks, and show how to find such safe groupings. We perform experiments on real bipartite graph data to study the utility of the anonymized version, and the impact of publishing alternate groupings of the same graph data.

Min-Soo Kim et al [2] describe the dynamic networks and attracting increasing interest due to their high potential in capturing natural and social phenomena

Copyrights © International Journal of Intellectual Advancements and Research in Engineering Computations, www.ijiarec.com

over time. Discovery of evolutionary communities in dynamic networks has become a critical task. The previous evolutionary clustering methods usually adopt the temporal smoothness framework, which has a desirable feature of controlling the balance between temporal noise and true concept drift of communities. They, however, have some major drawbacks: (1) assuming only a fixed number of communities over time; and (2) not allowing arbitrary start/stop of community over time.
The forming of new communities and dissolving of existing communities are very common phenomena in real dynamic networks. In this paper, we propose a new particle-and-density based evolutionary clustering method that efficiently discovers a variable number of communities of arbitrary forming and dissolving. We first model a dynamic network as a collection of lots of particles called nano-communities, and a community as a densely connected subset of particles, called a quasi l-clique-by-clique (shortly, l-KK). Each particle contains a small amount of information about the evolution of data or patterns, and the quasi l-KKs inherent in a given dynamic network provide us with guidance on how to find a variable number of communities of arbitrary forming and dissolving.

Nigel Medforth et al [3] describe the identity of the participants (nodes) must be anonymized to protect the privacy of the individuals and their relationships (edges) to the other members in the social network. We identify a new form of privacy attack, which we name the *degree-trail attack*. This attack re-identifies the nodes belonging to a target participant from a sequence of published graphs by comparing the degree of the nodes in the published graphs with the degree evolution of a target. The power of this attack is that the adversary can actively influence the degree of the target individual by interacting with the social network. We show that the adversary can succeed with a high probability even if published graphs are anonymized by strongest known privacy preserving techniques in the literature. Moreover, this success does not depend on the distinctiveness of the target nodes nor require the adversary to behave differently from a normal participant. One of our contributions is a formal method to assess the privacy risk of this type of attacks and empirically study the severity on real social network data.

Lei Tang et al [4] describe a multi-mode network typically consists of multiple heterogeneous social actors among which various types of interactions could occur. Identifying communities in a multi-mode network can help understand the structural properties of the network, address the data shortage and unbalanced problems, and assist tasks like targeted marketing and finding influential actors within or between groups. In general, a network and the membership of groups often evolve gradually. In a dynamic multi-mode network, both actor membership and interactions can evolve, which poses a challenging problem of identifying community evolution. In this work, we try to address this issue by employing the temporal information to analyze a multi-mode network. A spectral framework and its scalability issue are carefully studied. .

Chih-Hua Tai et al [5] describe the social networks model the social activities between individuals, which change as time goes by. In light of useful information from such dynamic networks, there is a continuous demand for privacy-preserving data sharing with analyzers, collaborators or customers. In this paper, we address the privacy risks of identity disclosures in sequential releases of a dynamic network. To prevent privacy breaches, we proposed novel kw-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. We also present a heuristic algorithm for generating releases satisfying kw-structural diversity anonymity so that the adversary cannot utilize his knowledge to reidentify the victim and take advantages. The evaluations on both real and synthetic data sets show that the proposed algorithm can retain much of the characteristics of the networks while confirming the privacy protection.

III.SYSTEM METHODOLOGY

The k-structural diversity anonymity defined the adversary with degree knowledge cannot recover the community identity as well as the vertex identity of an individual. These prior works mainly focus on static network. However, note that due to the rapid growth of the network data, the up-to-date data are usually regarded as much more valuable for the reasons of commerce and research. It is then not realistic to only release one snapshot at a certain time and be left behind to the changes after the release time. It presents a privacy model for protecting multi-community identity. This paper introduces a new privacy model, dynamic k^w-SDA, to ensure the protection for both vertex and multi-community identities in sequential publications.

To achieve k^w-SDA, we develop an efficient solution for anonymizing large-scale dynamic networks with limited information distortion. For better execution efficiency, we propose to construct a table, named the Cluster Sequence Table (CS-Table), to summarize

Copyrights © International Journal of Intellectual Advancements and Research in Engineering Computations, www.ijiarec.com

the vertex information of sequential releases and avoid the need of scanning all the releases for anonymization.

B	**3->3**	**{1}->{1}**
D	**3->2**	**{2}->{2}**
C	**2->2**	**{1,2}->{1}**
F	**2->1**	**{3}->{3}**
E	**1->3**	**{2}->{2,3}**
A	**1->1**	**{1}->{1}**

A.CS-TABLE CONSTRUCTION

The CS-Table is constructed. The CS-Table is a table consisting of three columns: vertex v, v ∈ V, the degree sequence and the sequence of multicommunity identities

$$A_{V = DV1+------+DV2}$$

Construction of the CS-Table

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.

Specifically, as the procedure in the following Figure shows, given $\mathbf{G^1}$, the CS-Table contains the vertices of $\mathbf{G^1}$ in decreasing order of their degrees. This involves in sorting all the vertices. When anonymizing $\mathbf{G^1}$, the CS-Table is simultaneously modified.

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 $\mathbf{G^2}$, a similar process is executed until $\mathbf{G^w}$ is anonymized.

B	**3**	**{1}**
D	**3**	**{2}**
F	**2**	**{3}**
C	**2**	**{1,2}**
E	**1**	**{2}**
A	**1**	**{1}**

T=1

T=2

B	**3->3**	**{1}->{1}**
D	**3->3**	**{2}->{2}**
C	**2->2**	**{1,2}->{1}**
F	**2->2**	**{3}->{3}**
E	**1->3**	**{2}->{2,3}**
A	**1->3**	**{1}->{1}**

T=3

B	**3->3**	**{1}->{1}**
D	**3->3**	**{2}->{2}**
E	**2->2**	**{2,3}->{2}**
A	**3->2**	**{1}->{1}**
C	**2->2**	**{1}->{1}**
F	**2->1**	**{3}->{3}**

T=4

Copyrights © International Journal of Intellectual Advancements and Research in Engineering Computations, www.ijiarec.com

Procedure CS-TableConstruction

Input: G^t $(1 \leq t \leq w)$, CS-Table

Output: CS-Table

1. **For** $v \in G^t$
2. AttachInfo(v, CS-Table)
3. **For** $\Theta_{\Delta}^{[1,t-1]} \in$ CS-Table
4. RankVertices($\Theta_{\Delta}^{[1,t-1]}$, $\Delta^{[t,t]}$)
5. **Return** CS-Table

B. CS-TABLE INCREMENTAL UPDATE

The CS-Table incremental update is made. Incremental Update. When a new snapshot $\mathbf{G^t}$ 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 $\mathbf{G^t}$ to the CS-Table, it is required to remove the information of b $\mathbf{G^{'t-w}}$ and re-sort the vertices according to the degree sequences

The anonymization process is carried out. For anonymization, three operations are used to adjust the degree of a vertex. Operation AddingEdge connects two vertices in the same community, i.e., 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.

Operation RedirectingEdge 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) ∪ (x, v), where vertex x is anonymized, y has not yet been anonymized, and $C_1 \cap C_2 \neq 0$.
This operation is satisfactory since it does not change the degree of an anonymized vertex and allows to increase the degree of a vertex without adding additional edges. 3) Operation AddingVertex connects a vertex and an additional fake vertex, i.e., V = V ∪ {u} and $Ct_1 \cap Ct_2 \neq 0$.

Algorithm Graph Anonymization

Input: G^t, CS-Table

Output: $\widetilde{G}^t$, CS-Table

1. **If** $t \leq w$
2. CS-TableConstruction(G^t, CS-Table)
3. **Else**
4. CS-TableIncrementalUpdate(G^t, CS-Table)
5. **While** ∃ not-yet-anonymized vertex in CS-Table
6. $v \leftarrow$ highest-ranked not-yet-anonymized vertex
7. $\Delta_M \leftarrow \Delta^w : \min MergeCost_V(v, \Delta^w)$
8. $cost_{MV} = MergeCost_V(v, \Delta_M)$
9. $cost_{ME} = MergeCost_E(v, \Delta_M)$
10. $cost_{CV} = CreateCost_V(v)$
11. $cost_{CE} = CreateCost_E(v)$
12. **If** (v is a new vertex) or
    $((cost_{MV} = \infty)$ and $(cost_{CV} = \infty))$
13. SpecialCase()
14. **Else If** $(cost_{MV} < cost_{CV})$ or
    $((cost_{MV} = cost_{CV})$ and $(cost_{ME} < cost_{CE}))$
15. MergeApproach(v)
16. **Else**
17. CreateApproach(v)
18. SynchronousChangeCS-Table( )
19. **Return** $\widetilde{G}^t$, CS-Table

The problem systemsolves the problem of protection of vertex and community identities of individuals in a dynamic network like existing system. In addition, implementation is carried out with real social network data. Any number of vertices can be set with any number of edges. 'w' value (time period during which an adversary can monitor a victim v.) can be set dynamically**.**

- To analyze the attack model and consider the protection of vertex and community identities of individuals in a dynamic network.
- To anonymize each release to satisfy some privacy model before a network is published.
- To prevent the changes to launch attacks made by adversaries due to the lack of consideration in sequential releases
- To avoid adversaries get advantages by gathering victim's information continuously and comparing the multiple releases.
- 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.
- To presents a privacy model for protecting multi-community identity.

Copyrights © International Journal of Intellectual Advancements and Research in Engineering Computations, www.ijiarec.com

- To introduce a new privacy model, dynamic kw-SDA, to ensure the protection for both vertex and multi-community identities in sequential publications.
- To achieve kw-SDA, thereby develop an efficient solution for anonymizing large-scale dynamic networks with limited information distortion.
- For better execution efficiency, to construct the table, named the Cluster Sequence Table (CS-Table).
- To summarize the vertex information of sequential releases and to avoid the need of scanning all the releases for anonymization.

IV.CONCLUSION

The problem of identity protection is solved. It proposed an efficient solution for anonymizing large-scale dynamic networks with limited information distortion. In addition, a summary table, the CS-Table is developed to summarize the vertex information of sequential releases and improve the efficiency. The project retains much of the characteristics of a dynamic network while confirming the privacy protection. In addition, the application required less working experience to construct the graph as well as construct CS Table and then anonymized the graph. The application is tested well so that the end users use this software for their whole operations. A trial run of the project has been made and is giving good results the procedures for processing the time stamped records in 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 database server. The following enhancements are should be in future.

The application if developed as web services, then many applications can make use of the records.

Multi Threading approach can be used so that the protection speed is increased.

The web site and database can be hosted in real servers during the implementation.

REFERENCES

[1] G. Cormode, D. Srivastava, T. Yu, and Q. Zhang, "Anonymizing Bipartite Graph Data Using Safe Groupings," Proc. VLDB Endowment, vol. 1, pp. 833-844, 2008.

[2] M.-S. Kim and J. Han, "A Particle-and-Density Based Evolu- tionary Clustering Method for Dynamic Networks," Proc. VLDB Endowment, vol. 2, pp. 622-633, 2009.

[3] S. Bhagat, B. Krishnamurthy, G. Cormode, and D. Srivastava, "Prediction Promotes Privacy in Dynamic Networks," Proc. Third Conf. Online Social Networks (WOSN), 2010.

[4] J.W. Byun, Y. Sohn, E. Bertino, and N. Li, "Secure Anonymization for Incremental Data Sets," Proc. Third VLDB Int'l Conf. Secure Data Management (SDM), 2006.

[5] J. Cheng, A.W.-C. Fu, and J. Liu, "K-Isomorphism: Privacy Preserving Network Publication against Structural Attacks," Proc. ACM SIGMOD Int'l Conf. Management of Data (SIGMOD), 2010.

[6] S. Fortunato, "Community Detection in Graphs," Physics Report, vol. 486, nos. 3-5, pp. 75-174, 2010.

[7] M. Girvan and M.E.J. Newman, "Community Structure in Social and Biological Networks," Proc. Nat'l Academy Science USA, vol. 99, pp. 7821-7826, 1999.

[8] S. Gregory, "An Algorithm to Find Overlapping Community Structure in Networks," Proc. 11th European Conf. Principles and Practice of Knowledge Discovery in Databases (PKDD), 2007.

[9] M. Hay, G. Miklau, D. Jensen, D. Towsley, and P. Weis, "Resisting Structural Re-Identification in Anonymized Networks," Proc. VLDB Endowment, vol. 1, pp. 102-114, 2008.

[10] L. Backstrom, D.P. Huttenlocher, J.M. Kleinberg, and X. Lan, "Group Formation in Large Networks: Membership, Growth, and Evolution," Proc. 12th ACM SIGKDD Int'l Conf. Knowledge Discovery and Data Mining (KDD), 2006.

Copyrights © International Journal of Intellectual Advancements and Research in Engineering Computations, www.ijiarec.com