International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 10 Issue: 11
DOI: https://doi.org/10.17762/ijritcc.v10i11.5776
***Article Received**: 18 August 2022 **Revised**: 30 September 2022 **Accepted**: 14 October 2022*

Data Security and Anonymization in Neighborhood Attacks in Clustered Network in Internet of Things (NIoT)

Ipseeta Nanda[1], Monika Singh[2], Lizina Khatua[3]
[1]Faculty of Information Technology
Gopal Narayan Singh University
Jamuhar, Rohtas, Bihar, India
ipseeta.nanda@gmail.com
[2]Research Scholar, Faculty of Information Technology
Gopal Narayan Singh University
Jamuhar, Rohtas, Bihar, India
singhmoni@gmail.com
[3]School of Electronics Engineering
KIIT Deemed to be University, Bhubaneswar, India
lizina_khatua@rediffmail.com

Abstract— In this paper author tries to focus on the review on the K Nearest Neighbor (KNN) tied by one or more specific types of inter dependency, such as values, visions, ideas, financial exchange, friendship, conflict, or trade. Social network analysis views social relationships in terms of nodes and ties. It also focuses the network analysis, application as well as problem statement. In this paper presents a outline for the privacy hazard and sharing the anonymized data in the network. This includes a proposed architecture design flow, for which the author considers the several variations and make connections. On several real-world social networks, we show that simple anonymization techniques are inadequate, it results in considerable breaks of privacy for even modestly informed opponents. It also concentrates on a new anonymization technique. It based on the network and validate analytically that leads to saving of the privacy threat. It also analyses the effect that anonymizing the network has on the utility of the data for social network analysis.

Keywords- K Nearest Neighbor, information science, Graph-based structures, Vertex

I. Introduction

A social network is a network of nodes on which people or organizations are typically connected. In a network, nodes are the individual players, while ties are the connections between them. Graph-based structures that are produced are frequently extremely complicated. The relationships between the nodes might take many different forms. Research in a variety of academic subjects has demonstrated that social networks function on many levels, from families up to the level of nations, and play a crucial role in deciding how problems are handled, organizations are run, and the extent to which people succeed in reaching their goals. A social network is essentially a map of all the connections between the nodes under investigation. The social capital of certain players can also be estimated using the network. A social network diagram, in which nodes are the points and ties are the lines, is a common way to illustrate these ideas.

Social Network Analysis

In contemporary sociology, social network analysis has become a critical methodology. Additionally, it has grown significantly in popularity as a subject of discussion and research in the fields of anthropology, biology, communication studies, economics, geography, information science, organisational studies, social psychology, and sociolinguistics.

The value of a social network to its members is influenced by its shape. Smaller, more tightly knit networks may not be as beneficial to their members as networks with many loose links (weak ties) to other people. Compared to closed networks with lots of redundant relationships, more open networks with plenty of weak ties and social connections are more likely to provide new ideas and possibilities to its members. In other words, a group of friends who only engage in activities together already has access to the same opportunities and expertise. A group of people who are connected to different social realms are probably able to get a larger variety of information. More links inside a single network are not better for personal success than several connections to other networks. Similar to this, people can exert influence or serve as middlemen inside their social networks by connecting two unconnected networks (called filling structural holes).

Applications of Social Networks

The strength of social network analysis comes from how it differs from conventional social science research, which make the assumption that individual actors' characteristics—such as whether they are nice or unfriendly, intelligent or foolish, etc.—are what count. A different perspective is produced by social network analysis, one in which connections and links between actors within the network are more significant than individual characteristics. This method has proven to be helpful in describing a wide variety of real-world occurrences, but it leaves less room for human agency—the capacity for people to affect their own success—because so much of it depends on the design of their network.

In order to characterize the numerous informal relationships that bind CEOs together, as well as associations and connections between specific employees at various firms, social networks have also been used to study how organizations interact with one another. For instance, influence within businesses frequently stems less from a person's actual job title and more from how central a role they play in a network of relationships. Additionally, social networks are important for hiring, corporate growth, and job effectiveness. Companies can use networks to acquire information, stifle competition, and coordinate the establishment of prices or regulations. However, the blessing also comes with it privacy and security issues, one of which we address in this project.

II. Motivation

Social networks are being widely used in today's world. Their use has been well established in almost all sectors. Starting from marketing, brand building to investigations social networks have been found to be very useful. Also making the data available publicly added to its uses and advantages. But availability of so much information at hand might prove hazardous. This is where the question of security comes into the picture. The privacy of the users can be easily attacked by an opponent with some network knowledge. Despite the fact that privacy preservation in data publishing has been thoroughly investigated and a number of models and algorithms have been put forth in this field, the majority of the current studies can only deal with relational data, making them inapplicable to social network data. The research therefore focuses on utilizing security techniques in social networks.

III. Aim and Objective of the Proposed Work

Recently as more and more social network data has been made publicly available; preserving the privacy in the social network has become a major concern. The project deals with providing the security in the social networks against a special category of attacks known as the neighborhood attacks. The term" neighborhood attack" means that if the adversary has some knowledge about the neighbors of the target victim and the relationship among the neighbors, then the victim may be re identified from the social network even if the victim's identity is preserved using conventional anonymization techniques.
Hence the project aims at achieving security in social networks against the neighborhood attacks by reducing the probability of an adversary successfully identifying any individual in the network.

IV. Anonymization method

To protect privacy, one way is to guarantee that any individual cannot be identified correctly based on the knowledge about the neighborhoods. This can be achieved by anonymizing the social network. In this method, the anonymization is carried by anonymizing the neighborhood of the vertices taken from the same group.

The process consists of three main phases:

1. Extracting the neighborhood of each vertex.
2. Organize the vertices into groups based on the isomorphism of the neighborhoods
3. Anonymize the vertices in the same group.

V. Requirements

A. *Functional Requirements*

The project requires to make the network anonymized such that, no vertex is unique in the published network according to its structure and hence it cannot be isolated under an attack.

B. *System Requirements*

The algorithm to be used must be efficient as the problem might be involving many nodes.

VI. Problem Formulation

The problem of preventing the neighborhood attacks using anonymization needs the consideration of various factors before the actual procedure. The anonymization algorithm depends on these factors. These factors basically determine the requirements for the algorithm. The steps involved in the formulation of the problem are depicted in the figure 4.

DOI: https://doi.org/10.17762/ijritcc.v10i11.5776
***Article Received**: 18 August 2022 **Revised**: 30 September 2022 **Accepted**: 14 October 2022*

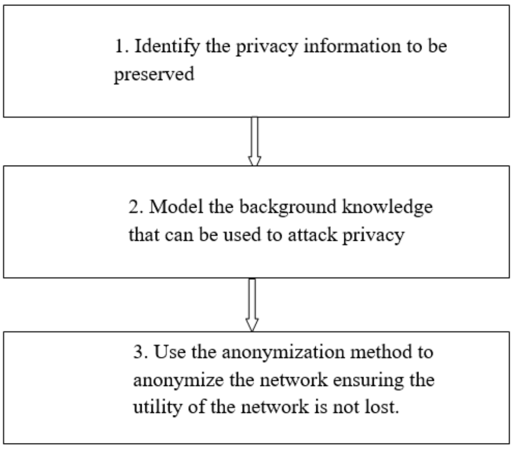
1. Identify the privacy information to be preserved
2. Model the background knowledge that can be used to attack privacy
3. Use the anonymization method to anonymize the network ensuring the utility of the network is not lost.

Figure 4. Problem Formulation

VII. Architecture Design

The overall project architecture design is shown using the figures 5, 6, 7, 8 below using different UML diagrams. The context diagram shows the various systems that are connected to the anonymization algorithm. All these algorithms are required for the execution of the anonymization algorithm.

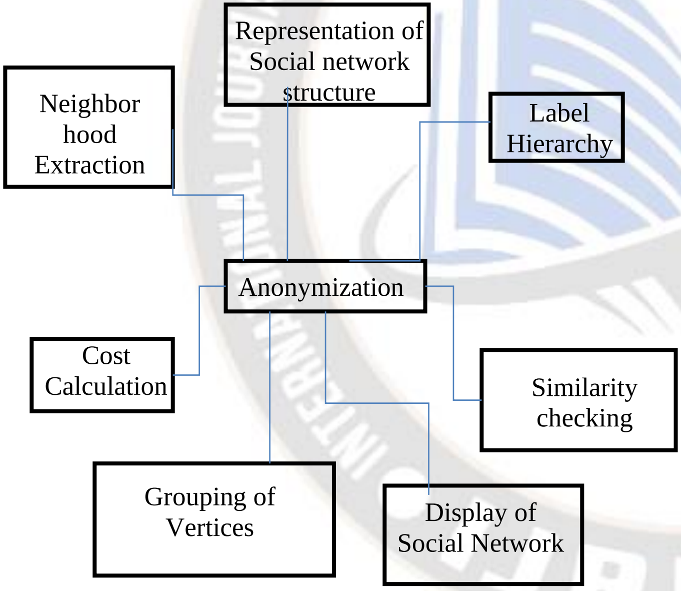

Figure 5. shows the classes and how they are connected to each other

The Label class is associated with the Vertex and Anonymize class for its various functions. Each vertex has its specific label. The class check similarity is required during anonymization procedure for checking the similarity between each pair of vertices and then anonymizing the pair with least cost. The class digraph is for providing a user interface where the graph entered by the user can be displayed. Also, the finally anonymized graph can be displayed using the adjacency matrix.

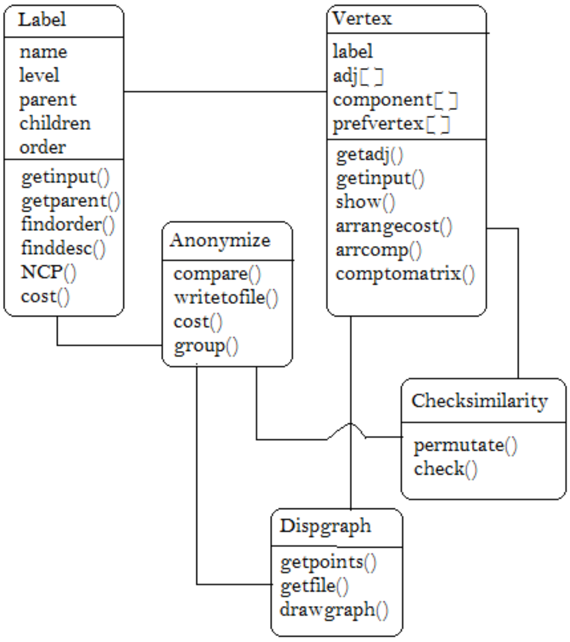

Figure 6. Class diagram

Figure 7 gives the overall design and flow of the project along with the basic inputs and constraints. The first step is to input the social network and extract the neighborhood of each vertex. Then the cost of anonymization is calculated for every vertex pair. For the calculation of cost various parameters like the addition of vertices, edges and label hierarchy are considered. Based on the cost factor the vertices are grouped together. Now the similarity between the different vertices is checked. After finding out the similar vertices anonymization algorithm is applied on them.

***Article Received**: 18 August 2022 **Revised**: 30 September 2022 **Accepted**: 14 October 2022*

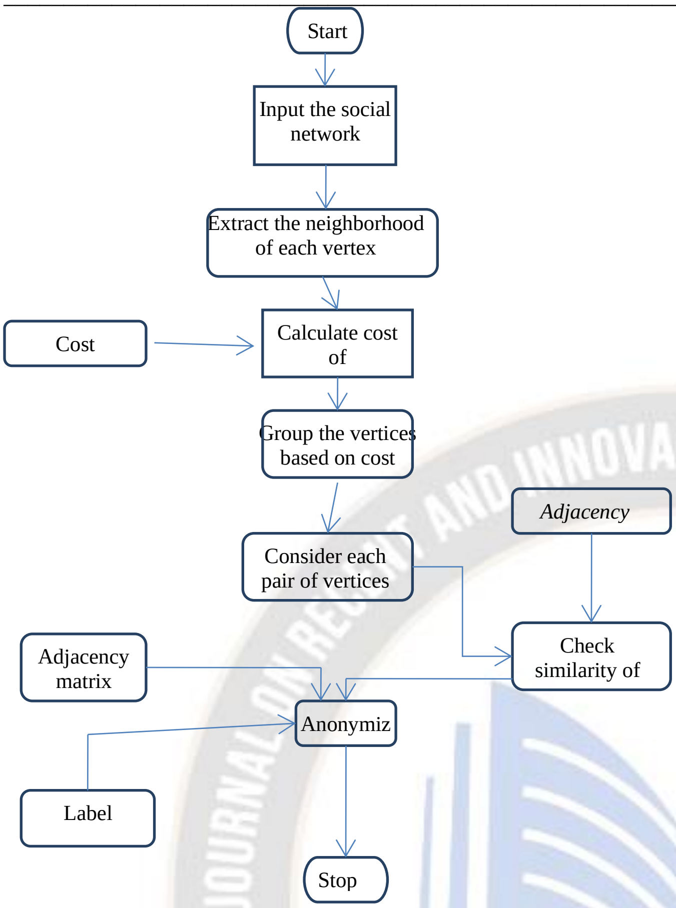

Figure 7. Overall Architecture

The following figure depicts the way in which the data flows in the anonymization algorithm.

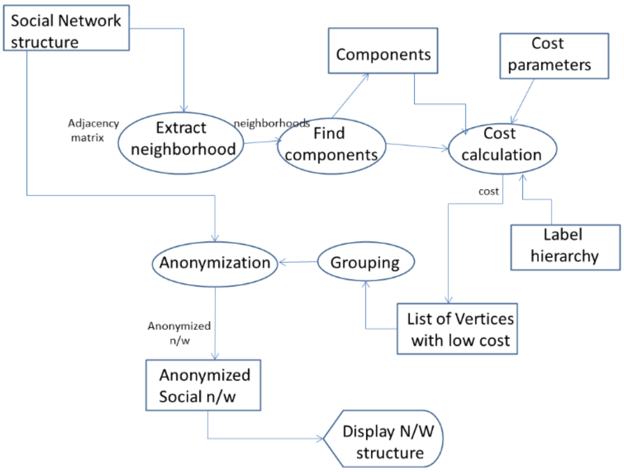

Figure 8. Data Flow Diagram

VIII. Modular Division

The project has been divided into 4 modules for the sake of easy implementation. The modules are as follows.

1. Modeling - Representation of the network using Graph Terminology.
2. Neighborhood Extraction - Collects the 1-neighborhood of each vertex.
3. Cost Calculation - Calculates the cost of anonymization.
4. Grouping and Anonymization - Groups the vertices into groups based on the similarity of the neighborhoods and Anonymizes the vertices in the same group or with minimum cost difference.

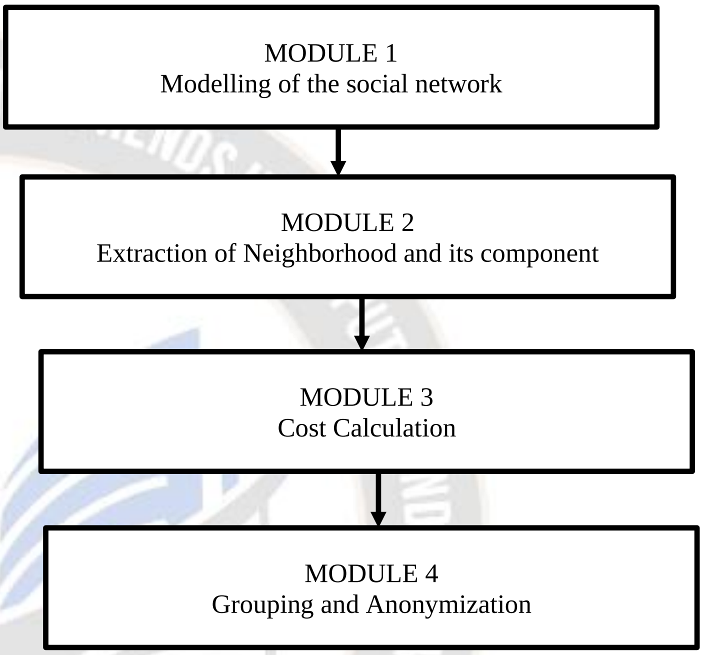
MODULE 1
Modelling of the social network
MODULE 2
Extraction of Neighborhood and its component
MODULE 3
Cost Calculation
MODULE 4
Grouping and Anonymization

Figure 9. Module Diagram

Conclusion

Privacy protection when sharing information via social networks has grown to be a major concern. An opponent may simply invade the privacy of some victims if they have some local knowledge about the users of a social network. As a result, numerous strategies have been devised to address this issue in the past. However, the majority of these research on privacy preservation primarily address relational data and do not apply to data from social networks. This issue is addressed by a recent approach suggested and investigated by Zhou and Pei using the k-anonymization method. This approach, however, has a lot of drawbacks.

References

[1] Machanavajjhala et al., “L-diversity: Privacy beyond k-anonymity,” in ICDE’06.

[2] Bin Zhou and Jian Pei,” Preserving privacy in Social networks against neighbourhood Attacks” Simon Fraser University, IEEE 2008.

[3] Mohsen Jamali and Hassan Abolhassani, “Different aspects of Social network analysis” in IEEE/ACM/WIC 2006.

DOI: https://doi.org/10.17762/ijritcc.v10i11.5776
***Article Received**: 18 August 2022 **Revised**: 30 September 2022 **Accepted**: 14 October 2022*

[4] L. Getoor and C. P. Diehl, "Link mining: a survey," ACM SIGKDD Explorations Newsletter, vol. 7, no. 2, pp. 3–12, 2005.

[5] L. Adamic and E. Adar, "How to search a social network," Social Networks, vol. 27, no. 3, pp. 187–203, July 2005.

[6] D.-W. Wang et al., "Privacy protection in social network data disclosure based on granular computing," in Proceedings of the 2006 IEEE International Conference on Fuzzy Systems, 2006, pp. 997–1003.

[7] M. Hay et al., "Anonymizing social networks," University of Massachusetts Amherst, Tech. Rep. 07-19, 2007.

[8] P. Samarati, "Protecting respondents' identities in microdata release," IEEE Transactions on Knowledge and Data Engineering, vol. 13, no. 6, pp. 1010–1027, 2001.

[9] P. Samarati and L. Sweeney, "Generalizing data to provide anonymity when disclosing information," in PODS'98.

[10] X. Xiao and Y. Tao, "Personalized privacy preservation," in SIGMOD'06.

[11] S. Wasserman and K. Faust, Social Network Analysis. Cambridge University Press, 1994.

[12] L. Backstrom et al., "Wherefore art thou r3579x?: anonymized social networks, hidden patterns, and structural steganography," in WWW'07.

[13] G. Kossinets and D. J. Watts, "Empirical analysis of an evolving social network," Science, vol. 311, no. 5757, pp. 88–90, January 2006.

[14] L. Backstrom et al., "Group formation in large social networks: membership, growth, and evolution," in KDD'06.

[15] R. Kumar et al., "Structure and evolution of online social networks," in KDD'06.

[16] L. Sweeney, "K-anonymity: a model for protecting privacy," International Journal on uncertainty, Fuzziness and Knowledge-based System, vol. 10, no. 5, pp. 557–570, 2002.

[17] M. Faloutsos et al., "On power-law relationships of the internet topology," in SIGCOMM'99.

[18] J. Xu et al., "Utility-based anonymization using local recoding," in KDD'06.

[19] Jonathan Green, Jayellen, " Graph theory and its applications", chp 2, pp 48-8, 2006.