Sanitization Techniques for Protecting Social Networks from Inference Attacks

JAYASREE DASARI\textsuperscript{1}, K.R.KOTESWARA RAO\textsuperscript{2}

Assistant Professor, Department of CSE, Gokul Institute of Technology and Sciences, Piridi village
Bobbili mandalam, Vizianagaram dt. Jntu kakinada university A.P, India\textsuperscript{1}
Student, Department of CSE, Gokul Institute of Technology and Sciences, Piridi village
Bobbili mandalam, Vizianagaram dt. Jntu kakinada university A.P, India\textsuperscript{2}

jayashree058@gmail.com\textsuperscript{1}, koteswarao.gokul0@gmail.com\textsuperscript{2}

Abstract— Social networking applications became popular as they provide virtual platform that can help in making people of all walks of life to have various communication channels over Internet. These applications allow users to publish their data and share it with other users as well. In the process, it is possible that adversaries can launch private information inference attacks to know sensitive information. In this paper we focus on finding how such applications are vulnerable to private information inference attacks and devise preventing mechanisms. We employ different sanitization techniques that prevent private information inference attacks. The sensitive information is thus prevented from being disclosed to unauthorized people. We built a prototype system that can demonstrate the proof of concept. We used the Census dataset in order to test the application. The empirical results are encouraging.

Index Terms – Social networks, data mining, inference attacks, sanitization techniques

I. INTRODUCTION

Social networking has been around for many years. People of all walks of life depend on Internet for obtaining various kinds of information. However, the usage of social networks like Facebook became increasingly more by people of all demographics. In this context, lot of information is being made available over social networks. When sensitive information is disclosed that might be misused by unknown people. Moreover the security settings provided by social networks are inadequate. In this context, users’ private information is not fully protected. When some data is known, the missing data can be obtained from other data sources. This is possible as the data is replicated in multiple sites or places. The kind of attack that infers unknown information from the known information is called as private information inference attack.
As can be seen in Figure 1, the information available over Internet and especially with social networks is pertaining to alliances, contacts, internet, social media, career, connections, social skills, relationships, virtual community, friendship, online, and business opportunities. Out of this information, sensitive information disclosure in the problem to be addressed.

In this paper we used various sanitization mechanisms in order to prevent private information attacks. The remainder of the paper is structured as follows. Section II provides review of literature. Section II reviews literature on prior works. Section III presents the proposed methodology. Section IV provides implementation and results. Section V concludes the paper besides providing directions for future work.

II. RELATED WORKS

This section provides review of literature on prior works. In [5] attacks are explored on anonymized networks. The goal of such attacks was to know the identity of people. In [7] many anonymizing techniques were explored. They tried to resolve the problem of sensitive information disclosure. He et al. [8] focused on creating Bayesian network by using the links available in social network. They did their research using hypothetical attributes. Many methods are proposed in [9] also anonymization techniques are explored using a technique known as k-anonymity. In [10] specific usage scenarios are explored in social networks. They also focused on various attacks such as stalking, re-identification and so on. When social networking web applications publish public data even after anonymization there are many possible attacks that can disclose sensitive information. One such attack is private information inference attack. Inference attack refers to an attack in which an attacks uses different sources of data to obtain unknown data from known data by matching known attributes. In [11] usage trends of Facebook data was explored.

Link based classification has been studied in order to devise solutions for the problem. Various methods related to link based classification are studied in [12]. The methods include mean-field relaxation, loopy belief propagation, and link-based classification. However, these techniques were not considered in [13] where alternative approach using Markov networks is proposed. In order to predict class labels for unknown data dynamic methods are employed in [14]. Predicting private attributes is carried out in [15]. Classification of data available over social network can provide useful insights into the attributes and it is possible to devise plans for protecting data as well. When data is published it is to be ensured that such exposed data do not allow adversaries to perform inference attacks.

III. PROPOSED SOLUTION

In this paper we proposed a solution for preventing private information inference attacks in social networks. An inference attack is the attack used to obtain private and sensitive information from the known data. Though sensitive information is not directly disclosed, it is possible to match the known information with other data sources available and successfully complete inference attacks. This can be prevented by proposing new sanitization techniques. In this
paper, we proposed a framework that has provision for making inference attacks and also prevention mechanisms in the form of sanitization techniques.

![Diagram](image)

**Figure 2 – Outline of the proposed architecture**

As can be seen in Figure 2, it is evident that the proposed architecture enables possible inference attacks in order to demonstrate the usefulness of the proposed solution to prevent inference attacks in social networks. Sanitization techniques are very useful are used to combat such attacks in social networks. Naïve Bayes Classification, Network Classification are used to achieve inference based solution. More details on the sanitization techniques can be found in [50]. Generalized information loss and structured information loss are the metrics used to make use of sanitization techniques.

**IV. PROTOTYPE AND RESULTS**

We built a prototype application using Java programming language. Intuitive application demonstrates the inference attacks, information loss procedures and the results. The environment used for the experiments is a PC with dual core processor, 4 GB RAM running Windows 7 operating system. The application is tested using census data that is collection from Internet sources.

![Image](image)

**Figure 3 – Sanitization process**
As can be seen in Figure 3, it is evident that the data is split into multiple result tables and they are subjected to sanitization. The reason behind sanitization is that the data is exposed to public without disclosing sensitive information.

Figure 4 – Clustering process

As can be seen in Figure 4, it is evident that there is provision for loading data, anonymizing it, clustering and viewing table. Two clusters are formed and age column is used as criterion for forming clusters.

Figure 5 – Zipcode process for anonymization

As can be seen in Figure 5, it is evident that the zipcode of the census data is used for anonymization. The UI provides provision for loading clustered data, generating zip code, viewing zip code and reziping the zipcode.
Figure 6 – General information loss process

General information loss is computed as shown in Figure 6. The UI facilitates loading clustered data and performing general information loss process besides allowing users to view total general information loss and natural general information loss.

Figure 7 – Sensitive value secure process

As can be seen in Figure 7, it is evident that the sensitive value secure process has provision for processing, sensitive data calculation, securing sensitive data and viewing the results.
As can be seen in Figure 8, it is evident that the total data loss is presented. It provides various details and visualize them using demographic features.

The generation information loss is computed and presented in Figure 9. From the result it shows the general information loss pertaining to cluster 1, cluster 2 and the total loss incurred.
Figure 10 – Shows structural information loss

The generation information loss is computed and presented in Figure 10. From the result it shows the structural information loss pertaining to cluster 1, cluster 2 and the total loss incurred.

V. CONCLUSIONS AND FUTURE WORK

In this paper we studied the need for protecting private information in social networks. Especially we focused on the private inference attacks on social networks. Sensitive information disclosure attacks are launched by hackers in which they infer unknown information from the known information. Information inference attacks can disclose the weakness of security mechanism in social networking systems. In this paper we employed such information inference attacks and devised sanitization techniques in order to prevent them. We built a prototype application that demonstrates the proof of concept. We tested the application using Census dataset obtained from Internet sources. The empirical results are encouraging. In future we continue our research to investigate more into privacy information inference attacks for building fool proof security with respect to preserving privacy.

REFERENCES


AUTHORS

JAYASREE DASARI is currently working towards her M.Tech degree in Gokul Institute of Technology and Sciences, Piridi village, Bobbili mandalam, Vizianagaram dt, A.P, India. Her research interests include Networking and cloud computing.
K.R. KOTESWARA RAO is working as an Assistant professor in Gokul Institute of Technology and Sciences, Piridi village, Bobbili mandalam, Vizianagaram dt, A.P, India. His main research interests are data mining and big data mining.