ISSN 2319–7757 EISSN 2319–7765

Indian Journal of Engineering

Collaborative User Security Using Nearest Neighbor Approach in OSN

Chander Kumar M

Assistant Professor, Department of Computer Science & Engineering, Fatima Michael College of Engineering & Technology, Madurai, Tamilnadu, India; E-mail: chanderkumarmca@gmail.com

Publication History

Received: 12 November 2015

Accepted: 16 December 2015

Published: January-March 2016

Citation

Chander Kumar M. Collaborative User Security Using Nearest Neighbor Approach in OSN. *Indian Journal of Engineering*, 2016, 13(31), 75-81

General Note

Article is recommended to print as digital color version in recycled paper.

ABSTRACT

Online Social Network (OSN) connects users to share information from one to another or group. Information sharing process requires high security. But in OSN, information security and trusted users identification is in low level. Because information shared with friends of friends affects security policy mechanisms. This work uses a classification algorithm for identifying users who are sharing information securely in social networks based on some constraint and implement the policies to the group. The First Module of this work is to create new user and validate that user whether their profile is correct or not. Whenever user enters the login form an additional security mechanism is implemented. This system initiates Profile creation process that represents the user account creation process and content uploading process. After profile creation process friends list are categorized. Category may be based on user's neighbors, colleagues, etc. information are shared to one or more users specified in the category, based on the shared users list, this work identifies the specific category of users shared the information to more number of users. Nearest neighbor approach performs to identify the shared user's category. Voting schemes are also used to identify shared users category. Threshold value is calculated from the voting schemes and compares the results for identifying the trusted users information shared in the network. In this, user may share their data to others in a secure manner and reduces the security issues in the social network.

Keywords: OSN, kNN, Social Network

1. INTRODUCTION

Social networking service is a platform to build social networks or social relations among people who, for example, share interests, activities, backgrounds or real-life connections. A social network service consists of a representation of each user (often a profile), his social links, and a variety of additional services. Social networking is a web-based service that allows individuals to create a public profile, to create a list of users with whom to share connection, and view and cross the connections within the system. Online social network is the sub part of social networking service.

The term is used to describe a social structure determined by such interactions in OSN. An axiom of the social network approach to understand social interaction is that social phenomena should be primarily conceived and investigated through the properties of relations between and within units, precisely because many different types of relations, singular or in combination form these network configurations, network analytics are useful to a broad range of research enterprises.

For solving critical issues, preliminary protection mechanisms have been offered by existing OSNs. For example, Facebook allows tagged users to remove the tags linked to their profiles or report violations asking Facebook manager to remove the contents that they do not want to share with the public. However, these simple protection mechanisms suffer from several limitations.

On one hand, removing a tag from a photo can only prevent other members from seeing a user's profile by means of the association link, but the user's image is still contained in the photo. Since original access control policies cannot be changed, the user's image continues to be revealed to all authorized users. On the other hand, reporting to OSNs only allows us to either keep or delete the content. Such a binary decision from OSN managers is either too loose or too restrictive, relying on the OSN's administration and requiring several people to report their request on the same content. OSN specifically analyze three scenarios—profile sharing, relationship sharing, and content sharing—to understand the risks posted by the lack of collaborative control in OSNs. Profile sharing implements an appealing feature of some OSNs is to support social applications written by third-party developers to create additional functionalities built on the top of users' profile for OSNs.

To provide meaningful and attractive services, these social applications consume user profile attributes, such as name, birthday, activities, interests, and so on. To make matters more complicated, social applications on current OSN platforms can also consume the profile attributes of a user's friends. In this case, users can select particular pieces of profile attributes they are willing to share with the applications when their friends use the applications.

User access the profile attributes of user's friend such as friends personal details and pictorial information with user's authentication. If user's another friend user also access this type information then information may be shared to unauthorized users. To prevent this type access from un authorized users user may fix the policies that specify the access control mechanism and also specify sharing restrictions to others or groups.

User relationship is implemented to create the policies and relationships are either unidirectional or bidirectional. Information is shared on both directions. Therefore policies represent the access control mechanisms on both directions. Some policies regulate the display of user's information to others.

2. PROBLEM STATEMENT

The basic concept is the evaluation of the various social network information analysis techniques and is also referred to solve the security concerns in the OSN. There are various policies to identify the user information flow in the network. Selecting the appropriate method plays a major role getting the desired output. The matching methods tend to be problem specific. In this dissertation, a study is made on the various fingerprint matching techniques.

Andrew Besmer and Heather Richter Lipford [1] develop privacy concerns and mechanisms surrounding these tagged images. Using a focus group, this paper explored the needs and concerns of users, resulting in a set of design considerations for tagged photo privacy. This system then designed a privacy enhancing mechanism based on our findings, and validated it using a mixed methods approach.

Bennet Hammer and James Parrish [2] represents a study to examine the effects that personal images posted to an individual's SNS and the comments associated with the image have on the interpretation of those images. It builds on prior studies done in this area by specifically examining SNS images and not an entire SNS profile.

Bert Huang and Angelika Kimmig [3] develops a flexible framework for probabilistic modeling of social networks that allows one to represent these different models and more. The framework, probabilistic soft logic (PSL), is particularly well-suited for this domain, as it combines a declarative, first-order logic-based syntax for describing relational models with a soft-logic representation, which maps naturally to the non-discrete strength of social trust.

David Zejda [4] focuses to bring overview on state of the art in main ideas behind a trust processing in online social networking systems. There had been impressive visions of trustworthy Internet, such as Augmented Social Network, where internet-wide persistent online identity across systems would facilitate reliable interactions of so called 'citizens of the Net'. A lot of work has been done to make Internet more trusty space already. We have security and trust authorities, security certificates, great algorithms, trust-related ontologies, whole area of trust management, some great systems.

Trust was discussed as something dynamic, continuously influenced by various factors. Further we outlined basic ideas of trust processing and inference of indirect trust and explained that subjective trust, either explicit or inferred, may be used as a source of objective trust metrics, such as community-wide reputation or system-wide trustworthiness. Trust systems may be used in many ways, e.g. to foster reliable interactions among users, to augment utility of shared content providing a property of reliability, as a major source of information for access control systems and for recommender systems.

Ed Novak and Qun Li [5] highlight the major issues concerning privacy and security in online social networks. This concept aims to protect user data from the various attack vantage points including other users, advertisers, third party application developers, and the online social network provider itself. Then the concept covers social network inference of user attributes, locating hubs, and link prediction.

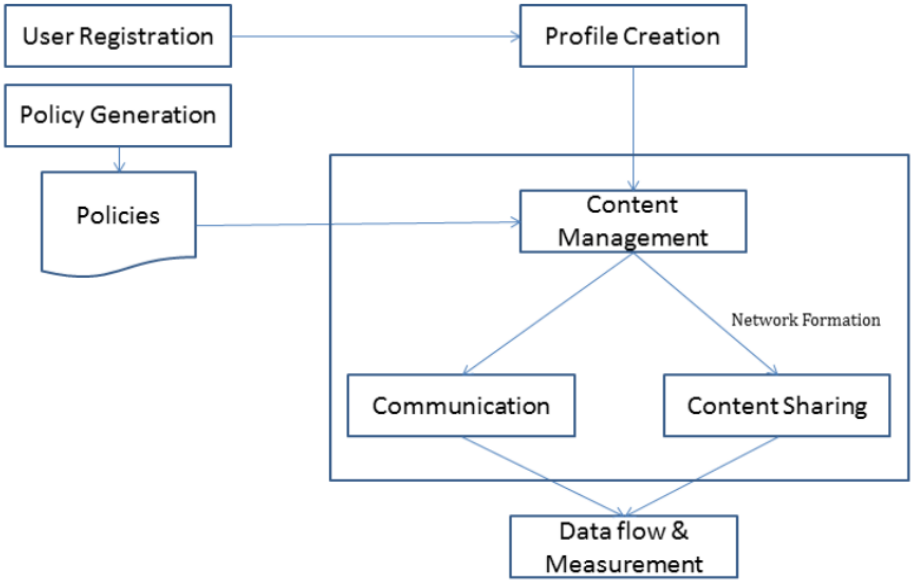

Figure 1 Architecture Diagram

3. PROFILE MANAGEMENT MODEL

OSN specifically analyze three scenarios—profile sharing, relationship sharing, and content sharing—to understand the risks posted by the lack of collaborative control in OSNs. Profile sharing implements an appealing feature of some OSNs is to support social applications written by third-party developers to create additional functionalities built on the top of users' profile for OSNs [1]. To provide meaningful and attractive services, these social applications consume user profile attributes, such as name, birthday, activities, interests, and so on. To make matters more complicated, social applications on current OSN platforms [7] can also consume the profile attributes of a user's friends. In this case, users can select particular pieces of profile attributes they are willing to share with the applications when their friends use the applications.

User accesses the profile attributes [6] of user's friend such as friends personal details and pictorial information with user's authentication. If user's another friend user also access this type information then information may be shared to unauthorized users. To prevent this type access [9] from un authorized users user may fix the policies that specify the access control mechanism and also specify sharing restrictions to others or groups. User relationship is implemented to create the policies and relationships are either unidirectional or bidirectional. Information is shared on both directions. Therefore policies represent the access control mechanisms on both directions. Some policies regulate the display of user's information to others.

The solution for sharing problem in online social network is to implements user registration and creates the profile f or that user. A new policy mechanism is implemented by each user and shared data must be secured manner.

A. Profile Creation

Create a New user with necessary details (User may be in the organization). This system creates a social network in any organization. Profile creation module contains the following sub modules

1. Content Management
2. Content Sharing
    a. Public sharing
    b. Private or limited sharing
    c. Timeline

Information may be transferred in one-to-one or one-to-many communication. This process is the essential process in any organization. User may share information to other people in both private and public. Information is maintained by separate user profile. Security policies are created based data on clustering algorithm. In the clustering approach k-means partitioning algorithm[1] is implemented. In every user this method is applied. In this approach partition the friends list based on factors (k) such as categories. Each factor groups the users list and implements separate policies for every group. When policies are created, information can be shared to many users and groups. Using that information, dissertation transfers information to numerical values for convenient input of the Nearest Neighbor Approach and find out the most relevant category users or groups accessed in that information. This system analyzes the security violation based on information flow of each user. Whenever a user updates or shares the information to his friends or colleagues then records are updated to every user who is viewing that information. This record immediately passes to that user and the user may identify the unknown user using that record. Numerical results will be calculated and the accuracy will be estimated.

4. NEAREST NEIGHBOR APPROACH

The Nearest Neighbors approach (or ***k*-NN** for short) is a non-parametric method used for classification and regression. In both cases, the input consists of the *k* closest training examples in the feature space. The output depends on whether *k*-NN is used for classification or regression:

- In *k-NN classification*, the output is a class membership. An object is classified by a majority vote of its neighbors, with the object being assigned to the class most common among its *k* nearest neighbors (*k* is a positive integer, typically small). If *k* = 1, then the object is simply assigned to the class of that single nearest neighbor.
- In *k-NN regression*, the output is the property value for the object. This value is the average of the values of its *k* nearest neighbors.

The neighbors are taken from a set of objects for which the class (for *k*-NN classification) or the object property value (for *k*-NN regression) is known. This can be thought of as the training set for the algorithm, though no explicit training step is required. A shortcoming of the *k*-NN algorithm is that it is sensitive to the local structure of the data.

In the classification phase, *k* is a user-defined constant, and an unlabeled vector (a query or test point) is classified by assigning the label which is most frequent among the *k* training samples nearest to that query point. A commonly used distance metric for continuous variables is Euclidean distance[8]. For discrete variables, such as for text classification, another metric can be used, such as the overlap metric[9] (or Hamming distance). Often, the classification accuracy of *k*-NN can be improved significantly if the distance metric is learned with specialized algorithms such as Large Margin Nearest Neighbor or Neighborhood components analysis. This dissertation can compute the distance between two scenarios using some distance function, where are scenarios composed of features, such that. Two distance functions are discussed in this summary:

Absolute distance measuring:

$$D_A (x,y) = \sum |x_i - y_i| \qquad (1)$$

Where i value is starting from 1 to n

Euclidean distance measuring

$$D_A (x,y) = \sum \sqrt{} \ (x_i^2 - y_i^2) \qquad (2)$$

Where i value is starting from 1 to n. This work consists an Euclidean distance measurement only. Because there are two users participated in each sharing process. If data is shared in multiple user's at a time or shared in a group that have more number of users at a time then this dissertation considers every shared user in a group, so every shared user has an owner of controller and achieved the two participants in the shared data.

This system finds the Euclidean distance from two different points. The Algorithm steps given as follows
Algorithm: To Find Euclidean Distance from neighbors
Input: Neighbor's distance n, r,t
Output: Euclidean distance D_a

1. To find receivers who are accessed in the shared information(timeline) with category
2. Convert receiver names to appropriate numerical values
3. Based on the numerical values, consider value as point
4. In each point calculate the Euclidean distance using the formula (Eq.2)
5. Short list the distance value to the ascending order
6. Based on this, take the first K values with its categories

Timelines are accessed user information that are collected from the specific timeline. For example, consider the user A share some information to their specific category or group member then the timeline is identified that User's Timeline identifies user names who are accessed the shared information with specific category, sensitiveness level and accessed date and time. Timeline is identified for transferring information to other user in specific date and time. Implementing these timeline values the dissertation identifies the shared data evaluation period on the network. Using this user information this system processes the next level.

5. RESULTS AND DISCUSSIONS

This process is implemented by using training data from the conversion of sample Social Network's data. In this process user names are converted to the numerical values based on the registration order. This order is based on the user's registration type and registration period. Simply, this order is based on the First Come First served (FCFS) order. The following table lists the users that are shared to more number of users with their category.

Table 1 Comparison analysis

POINT	CATEGORY
1.	Friend

2.	Neighbor
3.	Relation
4.	Relation

Friend	Colleague	Neighbor	Relation
1	0	1	2

The following table illustrates the threshold value for different content can be shared in different users.

Table 2 Threshold value Analysis

User content	Decision voting	Sensitivity voting	KNN algorithm
Content A	0.63	0.52	0.68
Content B	0.54	0.51	0.82
Content C	0.58	0.62	0.74
Content D	0.73	0.68	0.91
Content D	0.51	0.49	0.65

A graph is plotted for these values and the curve moves constantly in the KNN algorithm

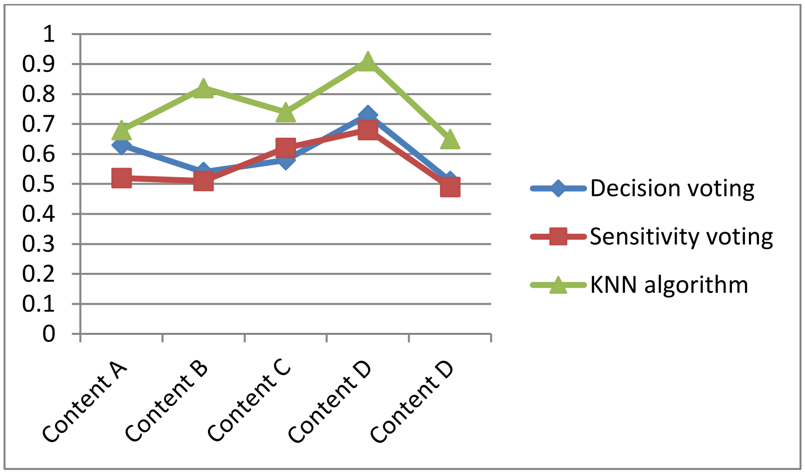

Figure 2 Graphical Analysis

Figure 2 analyzes the graphical representation of three algorithm concepts. The first one is decision voting scheme. Results in this method plots a curve that slightly increases and drops in suddenly so that sensitive levels in the trusted category increases in a small value(for example 0.5625). Sensitive voting schemes results plot the curve in the similar manner of decision voting scheme. Sensitive levels in the trusted category increases in the same manner (for example (for example 0.453125). Finally Nearest Neighbor Approach implementation results plot the curve that increases constantly and sensitive levels in the trusted category increases in a medium value (for example 0.75). Based on the comparison of three results Nearest Neighbor Approach implementation results gives the more number of sensitive values so that trusted user levels increases in this method.

6. CONCLUSION

To implements a framework for identifying the privacy policies in OSN to increase the security. Multiparty policy mechanisms are newly introduced. K-NN approach is used to categorize the friends list and add the limited functionalities when information shared in the Social Network. The framework also includes the information sharing timeline activity and the activity protects that information viewed in unauthorized people.

REFERENCES

1. Andrew Besmer & Heather Richter Lipford "Moving Beyond Untagging: Photo Privacy in a Tagged World" -April 10–15, 2010, Atlanta, GA, USA
2. Bennet Hammer, James Parrish, "Pictures is worth a thousand words, but are they the words that matter? – an analysis of the influence of image comments on social networking sites on the recruiters' evaluation of job candidates" - Nova Southeastern University.
3. Bert Huang, Angelika Kimmig?, Lise Getoor, and Jennifer Golbeck "A Flexible Framework for Probabilistic Modelsof Social Trust"-University of Maryland, College Park, MD 20742
4. David Zejda "From Subjective Trust to Objective Trustworthiness in On-line Social Networks: Overview and Challenges"-University of Hradec Kralove-Faculty of Informatics and Management
5. Ed Novak, Qun Li (2011) "A Survey of Security and Privacy in Online Social Networks" Department of Computer Science, The College of William and Mary.
6. Hongxin Hu and Gail-Joon Ahn-"Multiparty Authorization Framework for Data Sharing in Online Social Networks"-Arizona State University, Tempe, AZ 85287, USA.
7. Mohamed Shehab, Anna Squicciarini(2012), Gail-Joon Ahn, Irini Kokkinou –"Access control for online social networks third party Applications" -ELSEVIER-computers & security 31 897e911.
8. Michelle Madejski, Maritza Johnson, Steven M. Bellovin (2010) "The Failure of Online Social Network Privacy Settings" -CUCS-010-11
9. Pamela Wisniewski, David C. Wilson, Heather Richter-Lipford (2011) "A New Social Order: Mechanisms for Social Network Site Boundary Regulation" - AIS Electronic Library (AISeL).