International Journal of Modern Trends in Engineering and Research
www.ijmter.com
e-ISSN No.:2349-9745, Date: 2-4 July, 2015

Personlized Collaborative Tagging in Social Network Using Tag Filter

Pradnya Deshmane[1], Prof.N.R.Wankhade[2]

[1]*Department of Computer Engg, Late G.N.Sapkal COE,Nashik,deshmane.pradnya@gmail.com*
[2]*Department of Computer Engg, Late G.N.Sapkal COE, Nashik, nileshrw_2000@yahoo.com*

Abstract—Collaborative tagging technique is a very famous technique used in online system or social net working system. Our system works on the bottle neck area of some previous tagging method. This system contains the module that extends the tagging functionality capacity and features. Our system is having one policy layer that analyzes the collaborative tagging before it is into action. Our system required this layer to consider user preferences, deliberately defining sources on the basis of tag categories associated with them, and, some other parameters considering their trust wor- thiness. Along with this our system tries to preserve the user's sensitive information which is predictable from his feedback and comments. Hence our system also provides one privacy protection layer. Proposed system avoids misuse of profile by using technology namely tag suppression. Hence our system first filter tags using our policy layer and then for user interest it uses our privacy layer and finally from user interest respective tag is shared. We contribute one phase using which user can view social network status / reputation of particular user. Based on this online so cial network status /reputation of particular user we can take decision whether to allow such user or not. It is completely preventive approach that helps user to take decision whether to block particular user or not.

Keywords- Tagging, Collaborative Tagging, Policy Layer, Privacy Layer

I.INTRODUCTION

Tagging is one of the important features that facilitates user to share some data with important notes etc. This collaborative tagging is nearly supported by any kind of social web based systems. Tagging can be of anything like pages, documents, URLs, media files like video and images etc. Allowing such public tagging on our wall, helps other users belongs to us to view. From one point of view it is great facility for flow of information. If we consider another aspect then it is crucial and it must be focused. Another aspect considers the privacy of particular user. While working in social networking / bookmarking sites user makes friends, manages his profile and share some content with them. Users also tag various content to another user whether he is interested or not. Such unnecessary tagged content on user's wall increase overhead of end user. If that tagged content is malicious or which may spoil user's social status then user's online reputation may get affected. Also user tag cloud helps attacker to predict his behavior. User profile is also source of sensitive information. It also helps attacker to predict user's interest. Hence user privacy is important issue as far as user's profile and his tagging behavior is concern. Also there must be policy layer that helps user to block particular user and show interest in particular content so that such he will not get unwanted and malicious content from unwanted users.

There must be system that helps user to manage their profile and indicate them whether they are open to risk or not. Also there must be system that should support collaborative tagging using which user's interest is hide and tag perturbation is done. Hence there must be system that takes care of policy layer as well as privacy layer.

II. RELATED WORK

Collaborative tagging is currently an extremely popular online service. Collaborative tagging mainly focus on to task to loosely classify resources based on end user's feedback expressed in the form of free-text labels. Collaborative tagging may be the basis for a semantic network connecting online resources. Meaningless tags are difficult to semantic analysis hence research focuses on effective use of tags collections [1] [2] [3]. P.Mika [1] focused on analysis of data shared between the users and proved that semantic between them can be calculated. This gives us the idea of analysis of tags and its titles. X. Wu et al. [2] cannot focused on tagging but it worked on analysis of annotation in web based social networking which inspired us to go for tagging analysis. This system focuses only on bookmarking web based system. B. Markines et al [3] mainly analyzed the tags and try to find similarities between users by some similarity measures. Here they build an evaluation framework to compare various general folksonomy-based similarity measures, which are derived from several established information-theoretic, statistical, and practical measures. C.Marlow et al. [4] provide a short description of the academic related work to date. They offer a model of tagging systems, specifically in the context of web-based systems, to help us illustrate the possible benefits of these tools. Since many such systems already exist, we provide taxonomy of tagging systems to help inform their analysis and design, and thus enable researchers to frame and compare evidence for the sustainability of such systems. This is important to clear basic ideas about tagging.
B. Carminati et al. [5] discussed about the issue that how to assess the trustworthiness of Web metadata? They discuss how such issue can be addressed through the use of collaborative and Semantic Web technologies. . R. Gross and A. Acquisti [6] study the patterns of information revelation in online social networks and their privacy implications. We analyze the online behavior of more than 4,000 Carnegie Mellon University students who have joined a popular social networking site catered to colleges. They evaluate the amount of information they disclose and study their usage of the site's privacy settings. S.B. Barnes [7] discuss the uproar over privacy issues in social networks by describing a privacy paradox; private versus public space; and, social networking privacy issues. J. Parra-Arnau et al. [8] discussed about analysis of tags and its metadata. Tag analysis can tends to breach the privacy of particular user is proved here.
J. Voß [9] gives an overview of current trends in manual indexing on the Web. Along with a general rise of user generated content; there are more and more tagging systems that allow users to annotate digital resources with tags (keywords) and share their annotations with other users. This paper also shows that tagging should better be seen as a popular form of manual indexing on the Web. G. Adomavicius and A. Tuzhilin [10] give us an idea about recommender systems which analyze the metadata and recommend user suitable options to him. Z. Yun and F. Boqin [11] work on prediction of user profile from tags shared by him or tags shared with him. All the tags used and resources collected by a user constitute the userpsilas personal tag space, which contains valuable information that can be used for building and enhancing the user model. Shepitsen et al. [12] consider the collaborative tagging. Data mining techniques, such as clustering, provide a means to remedy these problems by identifying trends and reducing noise. Tag clusters can also be used as the basis for effective personalized reommendation assisting users in navigation.

III. PROPOSED SYSTEM

Fig.1 shows proposed system. It uses category datasets. . By refereeing category dataset if required tag suppression logic is executed and alternative description / name / term is suggested to user. It helps to make his comment / tags / profile more general and because of which user identity is remain

secured.

Our system is an enhanced collaborative tagging system that consists of a "traditional" bookmarking service with two main additional services. Such services address two main issues. The former allows end users to specify policies that can be used either to explicitly denote resources of interests or to enforce blocking conditions on the browsed data. The latter features a specific PET, namely, tag suppression, to preserve the privacy of registered users by hiding the specific characteristics of their profiles.

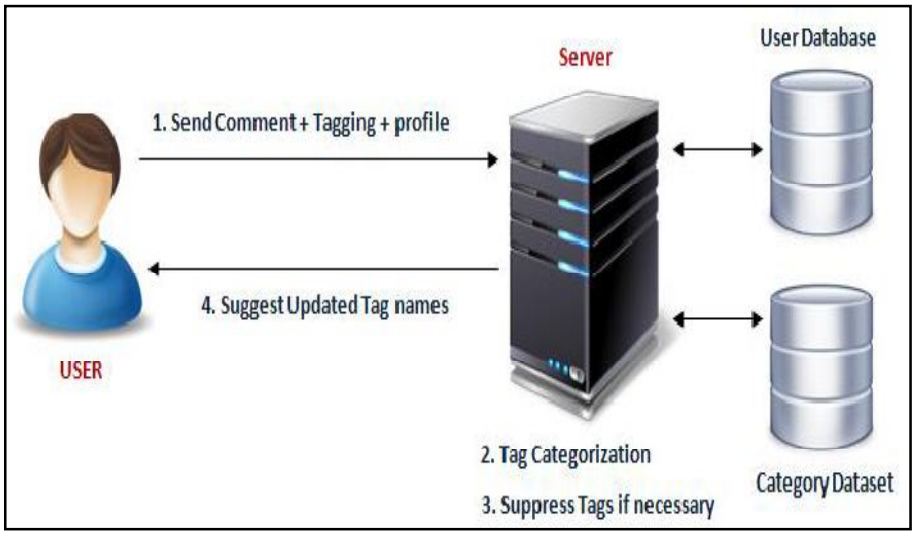

Figure 1: Basic Architecture Diagram

Collaborative tagging features and potential should be utilized with its full capacity. While extending the tagging capacity its core architecture should be as it is so that user can easily make use of it. Its implementation should be very smart and easy since data required is too big. While implementing collaborative tagging we have considered 2 aspects of it .In one aspect user can set some policies using his setting panel provided by the system. This allows user to specify his area of interest i.e. define area of blocking. Another aspect is focused to preserve user privacy by hiding some specific fields from his profile. This feature is called tag suppression. Tag suppression is on data set. Initially to define major categories, we have used delicious dataset. This dataset is having triplet like date, userId and bookmark URL with category. This dataset is too big hence we have used certain records. To categorize these records we have to use K-Means algorithm. By using this algorithm we have got 20categories.Each of these 20categories having sub-categories. The main objective is to find tags / sub-categories close to main category. With respect to user profile we can draw histogram out of it. We can draw this histogram in which Y-axis represents the summation of user proposed tags ratings belongs to particular category and X-axis represents main categories. With the help of histogram we can see the details of frequency of particular category.

3.1. Algorithm

3.1.1. K-Means Algorithm

K-Means algorithm is used for clustering. Here we used this algorithm to extract major categories from mentioned delicious dataset.

Steps

Input: k: the number of clusters

Output: A set of k clusters

Step 1: Choose k numbers of clusters to be determined
Step 2: Choose Ck centroids randomly as the initial centers of the clusters
Step 3: Repeat
3.1: Assign each object to their closest cluster center using Euclidean distance
3.2: Compute new cluster center by calculating mean points
Step 4: Until
4.1: No change in cluster center OR
4.2: No object changes its clusters

3.1.2 Shannon's Entropy:

This is used to find out the privacy level of user because of his apparent profile s as its Shannon's Entropy H(s).
This Shannon's entropy is defined as

$$H(s) = - \sum_{i=1}^{n} S_i \log_2 S_i$$

Here S is apparent profile of user.

IV. RESULTS

To develop this system we used JAVA and delicious dataset is used for analysis. We have extracted the main categories using delicious dataset. After this categorization of tags we have extracted tags for 10 users and we have tested tags suppression techniques for these users. We have drawn histogram based on the tags ratings and categories.
For example:
From data we have got following tags with their ratings in brackets as follows.
1 is considered when certain tag is far from main category and 10 is consider when tag is nearest to main category.
Computers (8), forum (2), dev (8), java (9), and development (9) from this we have finalized one main category named "Programming"
Lyrics (4), Ham radio (8), music (9), videos (9), radio (8), From this we have finalized another main category named "Entertainment"
Asset management (7), Economics (9), content management (7), stats (9), ngos (2). From this we have finalized another category named "Management Study" and we have drawn histograms out of it.

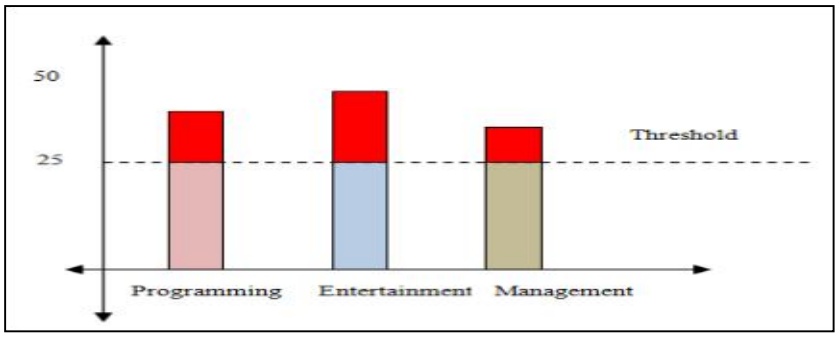

Figure 2.Histogram of Tag Categorization

Here red segments shows critical regions we have to cut off it up to threshold level by assigning "Programming" , "Entertainment"," Management" category names to very specific tags representing user profile uniquely.

CONCLUSION

Collaborative tagging is one of the popular services available online. Collaborative tagging is nearly supported by any kind of social web based systems. The proposed architecture of collaborative tagging system consists of bookmarking service and two additional services built on it. Also the privacy-preserving technology namely tag suppression has been implemented. An extensive performance evaluation of this architecture showing its effect tiveness in terms of privacy guarantees, data utility, and filtering capabilities for two key scenarios, for example, parental control and resource recommendation. Since we are not aware of similar experimental studies, we believe that what reported in this paper can be useful to evaluate further future developments in the area. Future work includes the development of a full prototype for the experimented system and its testing and use in further scenarios.

REFERANCES

[1] Javier Parra-Arnau, Andrea Perego, Elena Ferrari, Fellow, IEEE Jordi Forne and David Rebollo-Monedero "Privacy-Preserving Enhanced Collaborative Tagging" IEEE transaction on Knowledge and data engineering, Vol.26, No.1, January 2014

[2] P. Mika, "Ontologies Are Us: A Unified Model of Social Networks and Semantics," Proc. Int'l Semantic Web Conf. (ISWC '05), Y. Gil, E. Motta, V. Benjamins, and M. Musen, eds.,pp. 522- 536, 2005

[3] X. Wu, L. Zhang, and Y. Yu, "Exploring Social Annotations for the Semantic Web," Proc. 15th Int'l World Wide Web Conf. (WWW), pp. 417-426, 2006

[4] B. Markines, C. Cattuto, F. Menczer, D. Benz, A. Hotho, and S.Gerd, "Evaluating Similarity Measures for Emergent Semantics of Social Tagging," Proc. 18th Int'l Conf. World Wide Web (WWW), pp. 641-650, 2009

[5] C. Marlow, M. Naaman, D. Boyd, and M. Davis, "HT06, Tagging Paper, Taxonomy, Flickr, Academic Article, to Read," Proc. 17th Conf. Hypertext and Hypermedia (HYPERTEXT), pp. 31-40, 2006

[6] B. Carminati, E. Ferrari, and A. Perego,"Combining Social Networks and Semantic Web Technologies for Personalizing Web Access," Proc. Fourth Int'l Conf. Collaborative Computing: Networking, Applications and Work sharing, pp. 126-144, 2008

[7] R. Gross and A. Acquisti, "Information Revelation and Privacy in Online Social Networks," Proc. ACM Workshop Privacy Electronic Soc. (WPES), pp. 71-80, 2005

[8] B. Barnes, "A Privacy Paradox: Social Networking in the United States," First Monday, vol. 11, no. 9, Sept. 2006

[9] J. Parra-Arnau, D. Rebollo-Monedero, and J. Forne', "A Privacy-Preserving Architecture for the Semantic Web Based on Tag Suppression," Proc. Seventh Int'l Conf. Trust, Privacy, Security, Digital Business (TrustBus), pp. 58-68, Aug. 2010

[10] J. Voß, "Tagging, Folksonomy and Co - Renaissance of Manual Indexing?" Computer Research Repository, vol.abs/cs/0701072 ,2007

[11] G. Adomavicius and A. Tuzhilin, "Toward the Next Generation of Recommender Systems: A Survey of the State-of-the-Art and Possible Extensions," IEEE Trans. Knowledge Data Eng., vol. 17,no. 6, pp. 734-749, June 2005

[12] Z. Yun and F. Boqin, "Tag-Based User Modeling Using Formal Concept Analysis," Proc. IEEE Eighth Int'l Conf. Computer Information Technology (CIT), pp. 485-490, 2008

[13] Shepitsen, J. Gemmell, B. Mobasher, and R. Burke, "Personalized Recommendation in Social Tagging Systems Using Hierarchical Clustering," Proc. ACM Conf. Recommender Systems (RecSys), pp. 259-266, 2008

