An Alert System Based on Shared Score for Online Social Networks

T.Shanmughapriya
Department of Information Science and Technology
Anna University,Chennai,India
shanmugapriyat@ssn.edu.in

S.Swamynathan
Department of Information Science and Technology
Anna University,Chennai,India
swamyns@annauniv.edu

ABSTRACT

In recent years, the usage of Online Social Network's(OSN's) has increased in an unprecedented rate. OSN integrates an enormous number of Third Party Application (TPA)services to provide enhanced service set to the users. When users uses the application for the first time, they have to agree to share their profile content requested by the service providers otherwise, the service would be denied to the user. The users are not aware of the attributes that are really required for the application to function and extent of the personal data that they share with the third party applications. In the virtue of accessing the service, most of the time, the user simply agrees to share the attributes requested by the third party applications. The main focus of the paper is to mark the gap between the user's attitude and user's actual behavior of sharing. We estimate the measure of the sharing score of the user, based on what they actually share and compare it with the users preferred privacy score collected from the user through a questionnaire and intimate the user about the gap in the actual and user preferred share score. The list of applications and a score for each application are presented to the user. The score is calculated based on extent of the data the application accesses. The applications are classified into good, satisfactory and bad. Based on the classification emojis are displayed along with the score to help the users in understanding the reason for their over share.

CCS Concepts

•**Human-centered computing** → **Social networking sites;** •**Security and privacy** → *Privacy-preserving protocols;*

Keywords

Online Social Networks,Privacy,Third Party Applications.

ICTCS '16, March 04-05, 2016, Udaipur, India

DOI: http://dx.doi.org/10.1145/2905055.2905324

1. INTRODUCTION

The service provided by OSN, and the number of OSN users are increasing in an unprecedented rate. The OSN allows the users to form virtual relationships and share content among them. The OSN in advent to provide an extended list of services incorporates Third Party Application Services(TPA) offered by different vendors. Personal information of users is required for the TPAs to provide customized services to the users. The users have to solicit the data share by accepting the request raised by the TPA during the first time of application access. Most of the TPAs are based on 'All or Nothing' strategy, we have to either accept to share the contents marked as required in order to use the application or abstain from using the service. The TPAs access the user's data through the APIs provided by the OSN. Once the TPA gets control of the user's information the user's control over the data is fully lost. There are a number of privacy concerns related to the data being shared, (i.e.), The TPA could leak the data intentionally or unknowingly to fourth party agencies like advertisement agencies, data aggregation brokers, analytical engines...etc. The user faces an enormous risk for using the service provided by the TPA.

Application developers can see user's data. Their relationship is not symmetric as in the case of user's friend and also not transparent [4]. The user does not necessarily know who the application owner is. The current 'ALL OR NOTHING' privacy situation forces users to give away access to data that may not even be needed for the application to run. Social networking sites have a responsibility to protect the user data that has been handed over to them. The standard approach is to display a warning screen every time a user adds an application. Since this agreement is present in every application, the warning becomes meaningless. The OSN site has no way of monitoring the flow of information once it has been released.

A recent BBC report [8] noted that these applications do not always provide the same level of privacy protections that Facebook does. Poorly written or malicious programming can reveal personal information about anyone who utilizes the application (Kelly, 2008).The potential risk is due to the fact that the users are not aware of the extent of the personal data they share and lack control mechanisms. we do not tend to address any control mechanism and the paper is only concerned about ensuring the users awareness about their sharing. A survey conducted by felt et al.[6] indicates that 91% of the top 150 Facebook applications have unnecessary access to user's private data violating the principle of least privilege, Apart from the user having elevated sharing

behavior without necessary awareness the users also have selective behavior i.e., the users are only interested in sharing their personal attributes to selected audience depending on certain context.There exist a number of solutions to intimate and control the information shared with the TPA.

Imrul Kayes et al [7] in A Survey on Privacy and Security in Online Social Networks has coined that the solutions for controlling the data could be provided by either anonymizing or either by enforcing granular privacy policies or by providing platform support to run the applicarion itself, so that the data transfer to external parties could be limited.The Alert system designed fetches the user's data shared from OSN to the TPA. The above data was collected for users in the age group of 19 to 37 and a share score was calculated based on the attributes shared by the user. The user was also asked to attend a simple survey which was used to measure the user's preferred score. The gap between the user's actual and preferred sharing were intimated to the user. The TPA classifier classifies the application based on the attributes acquired by the applications. Emojis are displayed along with the score for the users to understand their sharing nature in a better way. Our solution intimates the user their actual state of privacy so that the necessary control could be made and the user's privacy could be preserved. The paper also focuses on the detailed survey of the work done in the current area. Section II discusses the related work, Section III discusses an overview of the system, data collection, experiment setup and results. Finally, in Section IV, we conclude the paper highlighting potential future work.

2. RELATED WORK

This section provides insight about the commercial and research related work relevant to the proposed work. Based on the literature, all the approaches to computing the sharing score till now are based on the measurement of privacy score in the user-user sharing scenario and none of them were measuring the information shared by the user to third party applications.The AVG privacy fix[1] and MyPermissions[3] are the commercial solutions closely related to the proposed work. AVG estimates the worth of user information stored in OSN and intimates the user in terms of cost(worth of user's data in OSN).The AVG Privacy fix covers up 8 different Facebook privacy settings and intimates the user of how much they are worth in Facebook. The user can choose the option of removing the attribute across the entire range of applications, However if the application has it as a required attribute, it's not removed for such applications. In MyPermissions, the user is intimated about the attributes they share with third-party applications and also supports control mechanism, which assists the user through revoke option which when chosen simply uninstalls the concerned third party application. The solutions lack the capability of intimating the seriousness to users. Common users lack skill in understanding the information presented as normal text and an approach like ours would be helpful to raise the awareness in users about their sharing through a simple score and information about the applications used by the users are presented through emojis describing the TPAs nature of data acquisition.

The solutions for controlling and intimating the information disclosure in literature follows. Kun Liu and Evimaria Terzi[9] have proposed a framework for computing the privacy scores of users in OSNs. The model is very similar to the proposed method and involves only intimation. Mathematical models have been developed to calculate privacy score as an aggregation of combined sensitivity and visibility values of data shared by the user. The method is very closely related to our approach, but it only calculates the privacy score of the user based on the information shared in user to user scenario and does not take into account the information shared to the third party applications. We consider that the information shared with the third party applications have more serious consequences, because we will not be having any control over the data once it is passed to the third party server.

In an effort to increase the awareness of users, visual aid has been considered as a key tool to improve the user's perception. A mechanism based on visual feedback presented by Strater, K. and Lipford, H. R.[11] tends to ease confusing and time-consuming privacy settings. View-centric privacy solutions are built on the intuition that a better interface for setting privacy controls can impact users understanding of privacy settings and thus ensures correctly exercising privacy controls. This interface is a collection of tabbed pages, where each page shows a different view of the profile as seen by a particular audience along with controls for restricting the information shared with that group. Management is tedious for users with many groups and involves more user intervention in the case of controlling the share.
Paul et al.[10] uses the same concept of creating visual impact.The method applies color coding for different privacy visibilities to minimize the cognitive overhead of the authorization task C4PS(Colours for Privacy Settings). This approach applies four color schemes for different groups of users. The color of the buttons shows the visibility of the data. The only drawback is that it does not automate the process of profile setting and depends on the user's input for every single move. Individuals alter their behavior to disclose or not to disclose frequently, hence doing it manually may be cumbersome.

The approaches mentioned below discusses the work done in the area of third party applications and its privacy implications in social networks. P. Anthonysamy et al. has proposed an approach [12] to provide fine-grained access control mechanism for TPA to access the user's private data. Every request from the TPA is intercepted by the Collaborative Privacy Manager (CPM) application. Users share their access control configurations for TPAs with other users in the OSN, who can reuse and rate such configurations. The mechanism to implement such an approach requires the support from the TPA.

The TPAs communication with the Facebook API has to be altered to that of the intercepting application. The practical viability of accepting and making such modification is uncertain. One solution based on access control framework was proposed by Yuan Cheng et al[5]. for social networking platforms where TPAs are classified as internal components and external components. Internal components can use all the personal information but cannot transmit further, the external components may use the essential and non-sensitive data attributes. The OSNs release all the non-sensitive information to third party service provider, which may also lead to personal attribute disclosure by correlating with the background information that could be aggregated from external sources. Existing solutions in the field does not provide methods to alert the user using the information of what

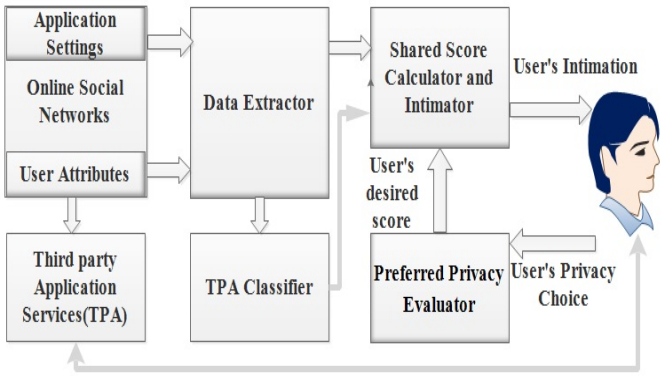

Figure 1: Block Diagram of Shared Score Intimator

is being shared with the third party services and this information changes frequently with the list of applications used by the user. Our solution provides a means to alert the user in case of any difference between their actual and desired share score.

3. SHARE SCORE CALCULATOR

3.1 Overview

The proposed solution was applied for Facebook scenario and results were drawn. Facebook acts as an identity provider for various third party services. First time when the user accesses the service, the user will be requested to share their attributes to third party applications. Once the user grants permission, the access token is handed over to TPA and the TPA could access the attributes by sending the token along with the request, the server checks the authorization and allows the TPA to access the data. Once the data is granted to the third party services, the user has no control over the information. The third party services could intentionally or

Attributes \ Permissions	Basic	Extended Profile Properties	Extended Permissions
Public Profile	*		
FriendsList	*		
Email	*		
Personal Description		*	
Work History		*	
Education History		*	
Home Town		*	
Birthday		*	
Likes		*	
Current City		*	
Friends Birthday			*
Messages			*
Checkins			*
Friends Personal Info			*
Relationships			*
Photos			*
Relationship Status			*

Figure 2: Attribute Category

accidentally leak the data to external services. There exist several solutions and the foremost solution that could be provided is preventing it before the data is actually transferred to such third party services. The share score calculator as shown in Fig 1.has data extractor, preferred privacy evaluator, TPA classifier and shared score estimator as core modules. The data extractor is responsible for extracting the user's Facebook setting, that has the details about the data shared with TPAs. The preferred privacy evaluator estimates the user's desired share level from the questionnaire answered by the user. The TPA classifier classifies the application based on the attributes acquired by the applications. The data attributes requested by applications are collected and a score is assigned to the application based on the extent it collects the user attributes and applications are classified as good, satisfactory and bad. The score of the application along with the emoji is displayed.The display of emoji is visually appealing and aids in creating awareness even for the naive users. The shared score calculator and the intimate module are responsible for measuring the shared score based on the sensitivity of the data shared and intimates when there is a discrepancy between the actual and desired share score.The privacy is context oriented and is entirely decided by the user as to how much of content has to be shared with whom and in which context. The Share score calculator calculates the sharing score of users and intimates the user about the extent the user shares their personal attribute to access the service and thus operates as an alerting service.

3.2 Data Collection

The data collection involved 122 users in the age group of 19 to 37. The data was collected from 170 applications for 17 user attributes. The attributes belongs to one of the categories specified by Facebook as Basic or Extended Profile Properties or Extended Permissions as listed in the Fig 2. We use this categorization later to assign weight for TPAs score computation. The user application setting information depicting what the application accesses from the user's profile is collected and the user's preferred share score is also measured through the questionnaire the user answers during the data collection. The data collected from the user is stored in CSV format file with all the attributes in the columns and the application accessing the attributes in the rows. If the attribute is accessed by the application, a binary value of '1' is placed in the cell intersecting the application and attribute name else a '0' is placed.

3.3 Experiment Setup

Net beans IDE with Derby backend was used for building the application. The Application has a upload module which imports the CSV format data into the derby database. The format of the CSV is as in Fig 2. A unique user ID is allotted to Users and all the personally identifiable information (PII) was stripped. A questionnaire as given in Fig 3 was given to the users. The questionnaire answered by the user was used to measure the preferred privacy level of the user.The share score was calculated on a scale of 1 to 10. The preferred share score of 10 means the user's attitude towards sharing attribute is high and has a very low level of privacy preference, whereas a shared score of 1 means the user is very concerned about sharing variables and has a high level of privacy preference. The user's actual privacy was calculated from the details fetched from the application setting of the user's Facebook account. The sensitivity value was

Attribute / Application	Attribute 1	Attribute 2	Attribute 3	...	Attribute17	User ID
Application1	1	1	0		1	UID 1
Application2	1	1	0		1	UID 2
Application3	1	1	1		0	UID 3
...						...
Application 170	1	1	1		1	UID 122

Figure 3: Structure of the shared Attribute data

1. Do you have the practice of inspecting the attributes requested by the application during the first time use of application ?
   a. Always
   b. Never
   c. Rarely

2. Do you have the practice of modifying the privacy settings of what you share with third party applications ?
   a. Always
   b. Never
   c. Rarely
   d. Not Aware

3. Have you ever denied third party services , when it requests to access the attributes ?
   a. Always
   b. Never
   c. Rarely

4. Given a scale of 1 to 5 .How would you rate your sharing level ? _____ .

Figure 4: Quetionaire presented to the users for capturing their share score

assigned to the user attribute based on the user's attitude. The concept behind the calculation of the sensitivity value is that the most sensitive value will be the least visible. For every user the sensitivity of the attribute is measured and is used as a weight factor in calculating the sharing score. The user who shares public profile element will have a less share score than a user who shares more personal information like relationship status. The share score module uses weight of the attribute to calculate the sharing score as given in the equation 1.

$$Sharescore(SC) = w_1 \Sigma att_1 + w_2 \Sigma att_2 + ... + w_n \Sigma att_n \quad (1)$$

Where $w_1, w_2, , w_n$ are the weights assigned to n attributes and $\sum att_i$ represents the aggregated score of any particular attribute across all the applications. As shown in Fig 3 attribute value '1' is assigned if it is shared with that application, else a '0' is assigned.The score for all the selected attributes are computed and combined with the weights assigned based on the sensitivity. The score has been scaled between 1 and 10 and compared with the preferred share score. The user is intimated in the case of discrepancy. This allows the user to be intimated about their difference in actual and preferred share score, creating awareness in their actual sharing levels.

The TPA classifier classifies the application based on the attributes acquired by the applications. The data attributes requested by applications are collected and a score is assigned to the application based on the extent it collects the user attributes. Weightage is assigned to the attributes based on the Facebook's classification as public profile, extended profile properties, open graph permissions and page Permissions. The score is calculated as in equation 2 and applications are classified as good, satisfactory and bad. The score of the application along with the emoji is displayed.

$$TPAScore = \sum w_i att_i \quad (2)$$

Where i=1,2,3..n, n represents the number of attributes, w_i represents weight assigned based on the attribute classification of att_i.

3.4 Results

The user's actual score, desired score and the list of applications accessed by the users along with their score and emoji is displayed to the user as shown in Fig.4. The actual and the desired share score of random samples that represent the original data of the users were plotted as shown in Fig.5. The share score comparison chart was used to estimate the percentage of users sharing above the desired mark and it was estimated that almost 53% of users shared 1.5 times more than that of their desired share level, around 6% have shared almost 4 times more than that of their desired share

Privacy Score

The users Actual Score :6.4
The users Desired Score :5.5
The Application Accessed by User

academia.edu	21	😐
angry birds friends	11	🙂
authorstream	11	🙂
candy crush saga	11	🙂
castle ville	7	🙂
city girl life	16	🙂
depositfiles.com	57	☹
docstoc.com	11	🙂
icims	45	☹
glassdoor	36	😐
reseachgate	11	🙂
slide share	21	😐

Figure 5: Alert System Screen Shot

level and 41% of the users had their preference more or less matching with the actual.

4. CONCLUSION

The Facebook Platform has been proving itself as one of the top most social networking applications. It has almost

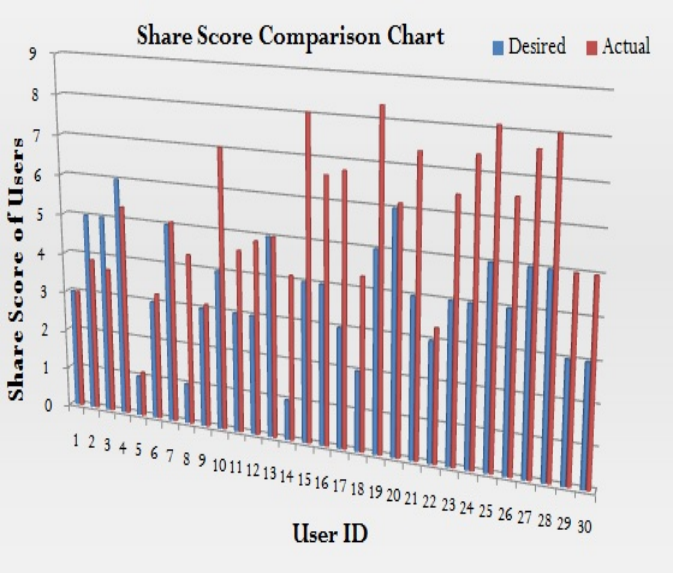

Figure 6: Share score comparison

1.49 billion monthly active users[2] and a myriad number of third party applications are connected with Facebook. The users are more and more interested in accessing applications that provide new features and tend to disclose their attributes in order to access the applications. This raises the privacy concern in such applications. Our solution is based on the concept that user alone is responsible for the dissemination of his privacy-relevant data. Currently, we have not applied any control measures towards controlling the privacy leak. The applications design motive was to insist that the user shares the attribute only with awareness. The application does not measure the indirect leakage of data attributes from third party applications to fourth party applications like applications that do analytics, applications that acts as a data aggregating broker and ad-based applications. The flow of data in that could be done my monitoring the network traffic from third party applications to external applications. The share score calculator intimates the user during over-share and also points out the cause of over-share by displaying emoji based on the classification of Third Party Applications.

5. REFERENCES

[1] AVG Privacy Fix. https://itunes.apple.com/us/app/avg-rivacyfix/id688795921?ls=1&mt=8, 2014. [Online; accessed 19-June-2015].

[2] Number of active facebook users as of 2015. http://www.statista.com/statistics/264810/number-of-monthly-active-facebook-users-worldwide/., 2014. [Online; accessed 19-June-2015].

[3] Mypermissions. https://play.google.com/store/apps/details?id=com.Mypermissions.mypermissions, 2015. [Online; accessed 19-June-2015].

[4] A. Ali-Eldin and J. van den Berg. A self-disclosure framework for social mobile applications. In *New Technologies, Mobility and Security (NTMS), 2014 6th International Conference on*, pages 1–5. IEEE, 2014.

[5] Y. Cheng, J. Park, and R. Sandhu. Preserving user privacy from third-party applications in online social networks. In *Proceedings of the 22nd international conference on World Wide Web companion*, pages 723–728. International World Wide Web Conferences Steering Committee, 2013.

[6] A. Felt and D. Evans. Privacy protection for social networking apis. *Web 2.0 Security and Privacy (W2SP'08)*, May 2008.

[7] I. Kayes and A. Iamnitchi. A survey on privacy and security in online social networks. *arXiv preprint arXiv:1504.03342*, 2015.

[8] Kelly. Identity at risk on Facebook. http://news.bbc.co.uk/2/hi/programmes/click_online/7375772.stm., 2008. [Online; accessed 19-June-2015].

[9] K. Liu and E. Terzi. A framework for computing the privacy scores of users in online social networks. *ACM Transactions on Knowledge Discovery from Data (TKDD)*, 5(1):6, 2010.

[10] T. Paul, M. Stopczynski, D. Puscher, M. Volkamer, and T. Strufe. C4ps-helping facebookers manage their privacy settings. In *Social Informatics*, pages 188–201. Springer, 2012.

[11] K. Strater and H. R. Lipford. Strategies and struggles with privacy in an online social networking community. In *Proceedings of the 22nd British HCI Group Annual Conference on People and Computers: Culture, Creativity, Interaction-Volume 1*, pages 111–119. British Computer Society, 2008.

[12] N. Wang, H. Xu, and J. Grossklags. Third-party apps on facebook: privacy and the illusion of control. In *Proceedings of the 5th ACM symposium on computer human interaction for management of information technology*, page 4. ACM, 2011.