# Privacy Preserving Interceptor for Online Social Media Applications

T. Shanmughapria(✉) and S. Swamynathan

Department of Information Science and Technology, Anna University, Chennai, India
priyanethiran@gmail.com, swamyns@annauniv.edu

**Abstract.** In the current scenario, online social network (OSN) plays a vital role in user's day to day life. User's store lots of their personal information in OSN, and they share their day to day experiences with their connections. There are enormous number of third party application (TPA) services allied with OSN to provide extended services to the users. The OSN manages the identity verification by authenticating the user and granting access to allied TPAs. The TPA requests permissions to access personal attributes about the user when accessed by the user for the first time. The personal attributes marked as required by the TPA has to be shared to avail the service. The privacy risk increases exponentially with the users TPA usage. The users not only leak their information to TPA but also end up unlocking a new type of threat from correlation with auxiliary information, through the data available from alternative sources. In this paper, we focus on reducing the level of sensitive data exposed to the external parties. The feasibility of providing such a service by restricting the data flow through access control policies is not feasible with the current All or Nothing approach. Hence, in this paper, we propose, the Privacy Preserving Interceptor (PPI) that acts as an interceptor between OSN and TPA to provide the required utility and yet preserves user's privacy. PPI identifies the sensitive attributes shared by the user and transforms the original data into a less sensitive form that still meets the utility goals. Standard Differential Privacy in combination with other perturbation mechanism of replacing with random values is used in PPI. The users privacy remains more or less the same both before and after the data share to TPA and still meets the utility needs of the user.

**Keywords:** Online social networks · Third party applications · Privacy · Data perturbation · Differential privacy

## 1 Introduction

To capture the interactions between the user, OSN, and TPA, we have chosen the Facebook platform [1] for observation. When the users access a TPA for the first time, the users are presented with a list of attributes to be shared with the TPA. The users are allowed to access the TPA only if they agree to share


R. Prasath and A. Gelbukh (Eds.): MIKE 2016, LNAI 10089, pp. 49–57, 2017.
DOI: 10.1007/978-3-319-58130-9_5

the attributes. Otherwise, the service is restricted. Though, such data could be used constructively for sharing users experience and use that knowledge to improve the experience of other users availing the same service; the users are being exposed to more and more privacy threats in such situations. The very act of sharing data is not considered as a breach of privacy. The privacy breach occurs whenever the contextual integrity is wrecked, in other words whenever the information not intended to be disclosed in that particular context gets exposed.

**Risk Analysis.** The risk involved in sharing attributes may lead to attacks [3,5,10,13] like privacy violations, de-anonymization, fake profile creation, information leakage attacks, cyber bullying, identity theft, etc., Women and children are even more vulnerable to these kinds of attacks. [2,15,17]. In an online survey conducted by BullyingUK, it is recorded that 87% of teens of age between 11 and 16, who reported cyber abuse said they were targeted on Facebook, and 20% blamed Twitter [14]. There were reported incidents where social media applications violated the accepted terms of service [9,11]. Consider a scenario where the user shares their attributes (sensitive or insensitive) to many applications from the same application developer. The application developer has access to all the information about a user from all the applications developed by them. There is a possibility to correlate that information and other information available through online resources, to infer the information that was not intended to be shared, thus violating the contextual integrity.

**Motivation.** Applying access control does not solve the current issue of sharing user attributes. Restricting sharing of a required attribute to a TPA, restricts the user accessing the service. Hence a solution is needed to enable the user to share all the required variables in a privacy-preserving way. PPI applies data perturbation techniques like standard differential privacy and randomization to transform the data. As the Privacy Preserving Interceptor resides in between OSN and TPA, users will benefit by using more number of TPA without much privacy concern and on the other hand, OSN and TPA will have an increased number of active users participating. Considering all these benefits, there should be some mutual consensus set up between OSN, PPI, and TPA for such interactions to be feasible. The rest of the paper is organized as follows. Section 2 presents the design of privacy preserving interceptor. Section 3 presents a survey of the existing work in the literature and Sect. 4 presents conclusion and future extensions of the paper.

## 2 Privacy Preserving Interceptor

### 2.1 Overview

The goal of the proposed mechanism is to preserve the privacy of the user and enable users to access multiple TPA without much compromise to utility. PPI applies Standard Differential Privacy concepts, Where the disclosure risk of the

user remains the same as before sharing. Utility level of the data is managed as we share all the required attributes as perturbed variables. The perturbed variable resembles the original variable statistically to provide the necessary level of utility. The degree of perturbation depends on the user's privacy requirement. Hence, our design provides a customized solution to the user that strikes a balance between privacy and utility matching the user's need.

The overall interaction between the components OSN, TPA and PPI is shown in the Fig. 1. The users request to OSN is always intercepted by PPI, it identifies the attributes and applies perturbation matching the user's level of privacy. The user's level of privacy is captured based on the user's sharing behavior captured via PPI. Initially, for a new user the privacy level is measured based on a short survey conducted when the user starts to use the application. The PPI eventually learns about the user's actual sharing level based on the attributes shared by users.

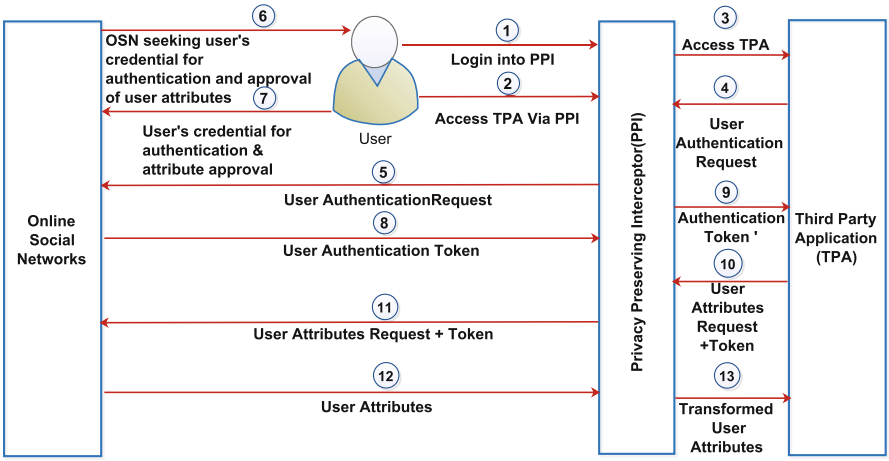


**Fig. 1.** OSN, PPI, TPA interaction pattern

## 2.2 Design of PPI

**Differential Privacy.** A randomized function $f$ gives $\in$ differential privacy if for all data sets $D_1$ and $D_2$ differing on a single user and all S $\subseteq$ Range(f).

$$Pr(f(D_1) \in S) \leq e^{\in} \quad Pr(f(D_2) \in S) \tag{1}$$

where $f(D_1)$ is function applied to data set including the user attributes and $f(D_2)$ is function applied to data set excluding the user attributes and $\in$ gives the required privacy level for all user attributes in the set S. The value of $\in$ is chosen based on the factors deciding the privacy requirement like user's privacy choice and attributes sensitivity.

Laplacian noise as in Eq. 2 is generated with chosen $\in$ and added to original user attributes to generate the perturbed user attributes.

$$\delta_{noise} = \frac{1}{2b} \quad e^{\frac{-(x-\mu)}{b}} \tag{2}$$

where $\mu$ is the mean of the noise signal and b represents the spread of the noise. The spread parameter is based on the global sensitivity parameter $\Delta F(x)$ and the privacy parameter $\in$ as shown in Eq. 3.

$$b = \frac{\Delta F(x)}{\in} \tag{3}$$

The global sensitivity is the difference between the maximum value and the minimum value that could be assigned to an attribute as in Eq. 4.

$$\Delta F(x) = max - min \tag{4}$$

The attribute requested by TPA is replaced with modified value $Attr_{perturbed}$ as in Eq. 5 arrived by adding the laplace noise $\delta_{noise}$ to original attribute (Fig. 2).

$$Attr_{perturbed} = Attr_{original} + \delta_{noise} \tag{5}$$

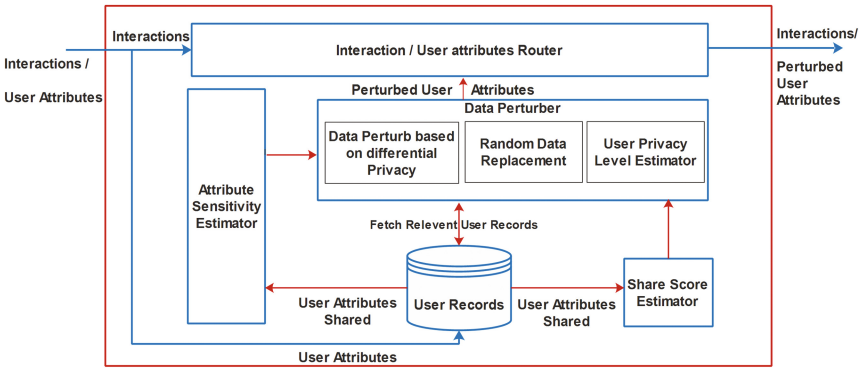


**Fig. 2.** PPI block diagram

**Choosing the Privacy Parameter.** Individual's choice of privacy varies based on multiple constraints like their geographic location, community, age, gender, etc. Therefore, there cannot be a single fixed level of privacy for all users. To address this issue, the module User's privacy level estimator updates the sharing level of the user based on the user's sharing behavior. Initially, a survey was conducted from users to capture their desired sharing level. The desired sharing level varies with the actual sharing level based on the context in which it is shared. So, the share score estimator calculates the actual sharing level of the user from the attributes shared in OSN. The privacy level of the user is measured on a scale of 3 based on the survey questions and the attributes shared by the user in OSN. The attribute table shown in Fig. 3 is used to calculate the share score as in Eq. 6.

$$Sharescore(actual) = \frac{w_1 \Sigma att_1 + w_2 \Sigma att_2 + ... + w_n \Sigma att_n}{n * m} \tag{6}$$

where $w_1, w_2 ... w_n$ are the weights assigned to the attributes, $\Sigma att_1, \Sigma att_2 ...$ $\Sigma att_n$ are attributes, n is the number of attributes and m is the number of applications. There may be variations between actual and desired share score. So, an average of actual and desired recorded from survey questions is used to measure the Share score. The attribute sensitivity estimator finds the sensitivity of data. The sensitivity of the data attributes is measured by finding how frequently the attribute has been shared (the most sensitive attribute is the least shared attribute) and the default classification given by Facebook. The attribute accessed by users are stored in structure as shown in Fig. 3.

| Attribute/ Application | Attribute1 | Attribute2 | Attribute3 | ... | Attribute N |
|---|---|---|---|---|---|
| Application 1 | 1 | 0 | 1 | | 1 |
| Application 2 | 1 | 0 | 0 | | 1 |
| Application 3 | 1 | 0 | 1 | | 0 |
| ... | | | | | |
| Application M | 0 | 1 | 1 | | 1 |

**Fig. 3.** Attribute table

Value '1' stored in the table denotes the attribute being shared. Facebook's classification is shown in Fig. 4, the attribute is assigned weight factor 1 for basic profile elements, weight factor 2 for extended profile elements and a weight factor of 3 for extended permissions. Let N be the number of applications accessed by the user. The sensitivity value calculation for attributes is done as in Eq. 7.

$$sensitivityscore_{att} = \frac{\sum_{i=1}^{N} app_i * wt_{classification}}{\sum_{i=1}^{N} app_i} \tag{7}$$

finally privacy level $\in$ is calculated as in Eq. 8.

$$\in = Sensitivityscore_{att} + Sharescore \tag{8}$$

**Replacement with Random Values.** Certain attributes like name, email id when shared with TPA, the chance is high that it could be correlated with auxiliary information revealing much more information, than currently available. Instead of adding noise to it, PPI uses the approach of distorting the original values by replacing with random attributes. Based on the user's sharing choice the attributes are either shared or replaced with random values picked from the database.

| Permissions / Attributes | 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 | | | * |

**Fig. 4.** Attribute classification

### 2.3 Discussion

The optional attributes can be opted not to share and the required attributes are perturbed thereby balancing the privacy and utility requirement of the user. Let us consider a case for App - Livestream, the attributes requested are public profile elements (name, age, gender, picture), email. Sensitivity value calculated for public profile elements is 1, email is 2. privacy level $\in$ is calculated as 1.22 and laplacian noise (0, 8.19) i.e., 0 mean and spread of 8.19 is added to original value. Assuming data ('AAA',23,'f',pic1.jpg,xyz@gmail.com) is transformed as data' ('XYZ',28,'f',pic1.jpg,xyz@gmail.com). The user can also opt to change the profile photo and email to be replaced with some random elements. PPI provides an attempt to perturb the required attributes. The mechanism could be applied to birthdate, location parameter, timezone, age etc., and random replacement has been applied for all other attributes. In our future extention we wish to apply other techniques to identify the similarity and replace the attributes like likes, action.music, action.books etc., with similar items to preserve the utility. The trust worthiness of the TPA could also be included as a parameter in deciding how much to disclose. Privacy Level Vs Perturbation % is plotted in Fig. 5. value of $\in = 1$ provides the highest perturbation level with 23% noise and value of $\in = 6$ provides the lowest perturbation with 4.5% noise.

## 3 Related Work-Comparison

There has been considerable work done to resolve the privacy issues in OSN-TPA scenario. To understand the seriousness of the privacy risk involved in

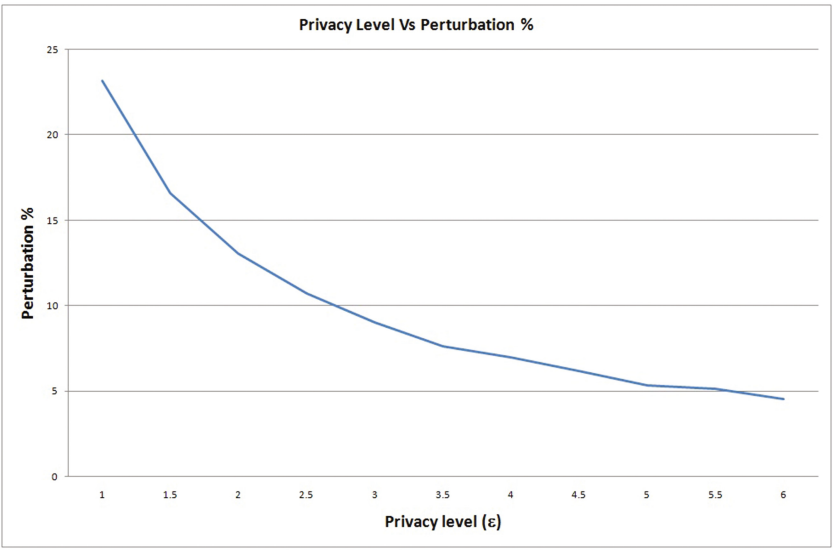


**Fig. 5.** Privacy level vs perturbation %

sharing user attributes with TPA Chaabane et al. [6] conducted an experiment to study the interaction between OSN applications and other external parties. The research revealed shocking results that the Facebook and RenRen applications interacted with hundreds of different fourth-party tracking entities. A similar study was conducted by Aldhafferi et al. [3] showed that the Personal data collected through TPA could be used for data matching to reveal sensitive information posing serious risks to privacy. [12] Kong et al. proposed a framework that utilizes the structure feature learning model to capture the relations among the permission requests and its textual descriptions and functionalities. The work provides insights for users to be aware of potential risks of permission requests.

Defining the privacy setting in the right manner is the most important and cumbersome task for a naive user. Hence Anthonysamy et al. [4] have proposed CPM framework that helps the users to gain control over their data shared with TPAs by utilizing the social construct of friends to identify the best configurations and make use of the same.

Cheng et al. [7] have proposed a framework based on access control mechanism, wherein the applications are split into an internal and external component. Allowing the internal components to access private information but restricting the access to external parties. Implementing such a solution may help to limit the optional attributes, whereas limiting the required attributes leads to the user being denied to access the service. Another framework based on PBAC (Permission-Based Access Control) proposed by [16] Tomy et al., has been designed to give users complete control over their data and to decide on the

information disclosure with third party applications. The work aims to provide the necessary awareness for the user in understanding the privacy risk involved in sharing sensitive data.

An approach similar to ours proposed by Egele et al. [8] have designed a browser plugin to intercept the data flow between OSN and TPA and through control mechanisms users can protect their profile data from malicious applications. The control mechanism of restricting the data once again stops the user from accessing the service. Whereas our approach does not limit the attribute, rather it grants access to the required attributes by generalizing them.

## 4 Conclusion

Preserving privacy in OSN is a challenging task. The very business model of OSN is based on analyzing the data shared by the user and using it as a monetizing fuel to run the business. The data is not only used within OSN but also the complexity increases by sharing it with third party application services to provide an extended service set to the users. Once the data is passed to TPA, the user's control over the data is lost, and they could even give the data to advertising agencies or data aggregating companies. The user's privacy threat spectrum widens with the inclusion of TPA. In this paper, we have collected information about the required attributes requested by the application. We have computed the privacy level of individual users and sensitivity of attributes to define the privacy parameter $\in$. Our proposed system Privacy Preserving Interceptor (PPI), intercepts the user's request and forwards the perturbed data to TPA. PPI perturbs the sensitive attribute by adding Laplacian noise or by replacing with random values. In our future extension, we wish to model the user's privacy level to include other features like the age of the user, gender, geographic location, cultural background. The attribute sharing is contextual. So, the TPA's trustworthiness could also be included as a parameter to decide the degree of disclosure. Improvements to perturbation techniques are being explored for certain attributes in which rather than replacing with random values, we are considering replacement with similar values to provide better utility.

## References

1. Facebook platform. http://developers.facebook.com/
2. Teen cyberbullying (2016). http://www.deccanchronicle.com/lifestyle/health-and-wellbeing/230816/teen-cyberbullying-more-common-among-friends-dating-partners.html. Accessed 7 Aug 2016
3. Aldhafferi, N., Watson, C., Sajeev, A.: Personal information privacy settings of online social networks and their suitability for mobile internet devices. arXiv preprint arXiv:1305.2770 (2013)
4. Anthonysamy, P., Rashid, A., Walkerdine, J., Greenwood, P., Larkou, G.: Collaborative privacy management for third-party applications in online social networks. In: Proceedings of the 1st Workshop on Privacy and Security in Online Social Media, p. 5. ACM (2012)

5. Bilge, L., Strufe, T., Balzarotti, D., Kirda, E.: All your contacts are belong to us: automated identity theft attacks on social networks. In: Proceedings of the 18th International Conference on World Wide Web, pp. 551–560. ACM (2009)
6. Chaabane, A., Ding, Y., Dey, R., Kaafar, M.A., Ross, K.W.: A closer look at third-party OSN applications: are they leaking your personal information? In: Passive and Active Measurement, pp. 235–246. Springer, Cham (2014)
7. Cheng, Y., Park, J., Sandhu, R.: Preserving user privacy from third-party applications in online social networks. In: Proceedings of the 22nd International Conference on World Wide Web Companion, pp. 723–728. International World Wide Web Conferences Steering Committee (2013)
8. Egele, M., Moser, A., Kruegel, C., Kirda, E.: Pox: protecting users from malicious facebook applications. Comput. Commun. **35**(12), 1507–1515 (2012)
9. Mills, E.: Facebook suspends app. that permitted peephole (2008). http://news.cnet.com/8301-10784_3-9977762-7.htm. Accessed 10 May 2016
10. Jernigan, C., Mistree, B.F.: Gaydar: Facebook friendships expose sexual orientation. First Monday **14**(10) (2009)
11. Kelly: identity at risk on Facebook (2008). http://news.bbc.co.uk/2/hi/programmes/click_online/7375772.stm. Accessed 19 June 2015
12. Kong, D., Jin, H.: Towards permission request prediction on mobile apps via structure feature learning. In: Proceedings of SIAM International Conference on Data Mining (SDM 2015). SIAM (2015)
13. Kosinski, M., Stillwell, D., Graepel, T.: Private traits and attributes are predictable from digital records of human behavior. Proc. Nat. Acad. Sci. **110**(15), 5802–5805 (2013)
14. MadhumithaMurgia: cyber bullying (2016). http://www.telegraph.co.uk/technology/2016/06/19/facebook-leads-the-way-in-online-compassion-but-others-need-to-f/. Accessed 7 Aug 2016
15. Selkie, E.M., Fales, J.L., Moreno, M.A.: Cyberbullying prevalence among us middle and high school-aged adolescents: a systematic review and quality assessment. J. Adolesc. Health **58**(2), 125–133 (2016)
16. Tomy, S., Pardede, E., Taniar, D., Pardede, E.: Controlling privacy disclosure of third party applications in online social networks. Int. J. Web Inf. Syst. **12**(2) (2016)
17. Ybarra, M.L., Mitchell, K.J.: How risky are social networking sites? A comparison of places online where youth sexual solicitation and harassment occurs. Pediatrics **121**(2), e350–e357 (2008)