Privacy Protection Based Privacy Conflict Detection and Solution in Online Social Networks

Arunee Ratikan[1] and Mikifumi Shikida[2]

[1] School of Information Science
[2] Research Center for Advanced Computing Infrastructure,
Japan Advanced Institute of Science and Technology, Ishikawa, Japan 923-1211
{a.ratikan,shikida}@jaist.ac.jp

Abstract. Online Social Networks (OSNs) such as Facebook, Twitter, and so on recently are major impact in communication and social interaction. Users can share any information with others. However, they have concerns about losing privacy due to lack of an adequate privacy protection provided by the OSNs. The information posted by the user (owner) might leak to unwanted target users. Especially, when collaborative information (e.g. text, photo, video, link), which has associated with the owner and multiple users (co-owners) in the real world, is posted into the OSNs, the co-owners do not have permission to control and might not be aware their information that is being managed by others. To overcome, collective privacy protection (CPP) is proposed to balance between the collaborative information sharing and the privacy protection for the owner and co-owners by majority vote. It enables the owner to create the privacy policy and the co-owners to make a decision in the privacy policy by vote. It additionally identifies and solves the privacy conflicts because at least one co-owner intends to keep private.

Keywords: Online Social Networks, Information sharing, Privacy protection.

1 Introduction

OSNs such as Facebook, Google+, Twitter refer to "online communities whose main goal is to make available an information space, where each social network participant can publish and share information" defined by [1]. This leads to communication and social interaction each other. For example, when the user can share personal stories, interests, activities, services and so on, other users can comment or press like button on this information.

The user in the OSNs can be both reader and creator according to the information consuming and the information sharing. The creator can additionally refer to an owner and a co-owner. The owner creates and posts the collaborative information to the OSNs. The co-owner has participated in creating the collaborative information or can referred by the owner such as tagging or mention. Nonetheless, the co-owner does not post it to the OSNs.

T. Tryfonas and I. Askoxylakis (Eds.): HAS 2014, LNCS 8533, pp. 433–445, 2014.

In case of the collaborative information, the owner can tag and mention the co-owners on the information. Moreover, the OSNs allow the user, who is not the owner of the information to share the information. If the collaborative information leaks to unwanted target users (referring to whom the owner and the co-owner are not willing to share with), it is hard to solve. This is because the owner and co-owners cannot command those users to stop spread the information via mobile phone, word of mouth and so on. Therefore, this causes the owner and co-owners, who have associated with the collaborative information, lose privacy thanks to lack of an adequate privacy protection.

Our goal is to balance between the collaborative information sharing and the privacy protection for the owner and co-owners. Therefore, this research proposes the CPP by applying majority vote concept. The proposed CPP allows the owner to create the privacy policy and the co-owners to make a decision in the privacy policy by vote whether or not the collaborative information should be posted. This is because the owner or co-owners have difficulty with setting the privacy policy [2]. This makes the proposed CPP differ from other research works [3] [4]. The proposed CPP additionally identifies and solves the privacy conflicts because at least one co-owner intends to keep private. The privacy conflict comes from different privacy concerns over the collaborative information by the owner and the co-owners. Furthermore, when the co-owners want to share the information, they also can perform themselves as the owner. This is because the owner and co-owners have right in the collaborative information and each owner might have different privacy preference.

2 Background and Related Works

2.1 Cause of Losing Privacy

The losing privacy in the OSNs can generally cause from four possible ways as follows:

1. The information is shared by the user or other users with poor privacy setting or no privacy setting. Several research works indicated that the users have difficulty with the privacy setting or do not use it [5] [6].
2. The information is tagged or mentioned by other users. These actions are meaningless in privacy protection because the tagged or mentioned users do not have permission the control the information before the information was spread in the OSNs.
3. When the user posted the information via own space provided by the OSNs, it might be shared or re-shared by other users without permission such as retweet in Twitter or share in Facebook.
4. Privacy setting provided by the OSN is not adequate for privacy protection because it allows only the user, who posts the information, to privilege in control the information. This means the other users, who have associated with this information, cannot do anything for their information.

Posting the collaborative information might lead to a crime problem because the OSNs allow the creator to use "Check-in" feature. This feature can reveal actual location where the activities are being performed or were done by many users, therefore a criminal can take advantage from this information within little time to find victim's location available on the SNP.

2.2 Access Control Models and Other Solutions for Privacy Protection

In order to protect the user's privacy, most of research works have been proposed the access control model. Gollu et al. [7] presented a social-networking-based access control mechanism for the information sharing. Identities between the users were viewed as key pair and social relationship. They provided access control list to determine who can access the information. Carminati et al. [8] proposed a rule-based access control mechanism for the OSNs. Type, depth and trust level of existing relationship between the users were used for expressing the complex privacy policy. Hart et al. [9] used the relationship information, which had exist in the OSNs, in a content-based access control model. This model could authenticate the user for accessing the information. Hu et al. [3] proposed a mechanism that detected and resolved the privacy conflict among the users, who had shared ownership of the collaborative information. Their research works enabled these users to provide the policy then calculated the privacy risk and sharing loss. Hu et al. [4] presented collaborative privacy management for shared data in Google+. This research introduced the concept of circle and trust to their model. Squicciarini et al [10] considered that the information might not belong to only one user in some cases, therefore they made a mechanism that supported the information sharing in the OSNs based on the notion of content ownership.

Besides the access control models, there are other solutions for privacy protection. Dinh et al. [11] attempted to construct a circle of trust by proposing the hybrid algorithm from investigating the maximum circle of trust problem. Thus, the user can safely share the information with others or the information will not be leaked to unwanted target users. Li et al. [2] used machine learning techniques and structured semantic knowledge in the ONS to learn the users' privacy setting pattern in the past and users' profiles. Then, this research made recommendation for the privacy setting to the user. Adu-Oppong et al. [12] applied automatically extracted network communities to make privacy policies easier by grouping friends into lists.

Although many access control models and other solutions have been proposed for privacy protection, they allow only the owner to control the privacy setting. Only few research works realized losing privacy of the co-owners, who has associated with the collaborative information. In some research works [3] [4], the owner and co-owners can create the privacy policy; but it cannot satisfy everyone. Possibility of violating privacy remains if at least one co-owner intends to keep private.

3 Research Methodology

Figure 1 depicts the proposed CPP for the privacy protection of owner and co-owners. It composes of five main components: social graph, privacy policy, co-owner invitation, majority vote, and conflict identification and provided solution. The work-flow of the proposed CPP begins at the owner creates the privacy policy. Then, the co-owners are detected in order to send the invitation that they are owned a part of the information and this information is being posted to the owner in the OSNs. The co-owners can vote on the privacy policy whether or not this collaborative information should be posted on the OSNs. The co-owner can be one of three statuses: acceptance, rejection and no response. No response status presents that the co-owner does not accept or reject in time. Nevertheless, when the time is over, the co-owner with no response will be moved to rejection status because the privacy is considered as a high priority. Next, the proposed CPP finds the privacy conflict among the owner and the co-owners, and provides the solution for each conflict. A list of the target users, who can see this information, is suggested to the owner to re-check before uploading it to the OSNs.

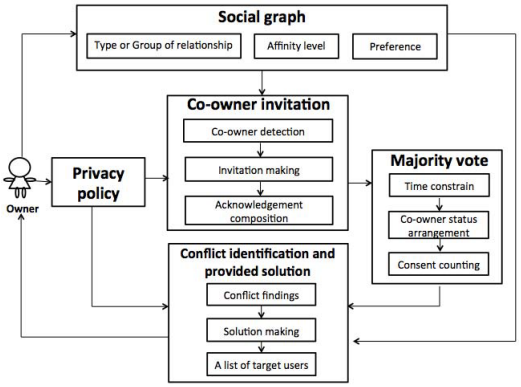

Fig. 1. The proposed collective privacy protection for the owner and co-owners in the information sharing

4 Proposed Collective Privacy Protection

The proposed CPP composes of five main components, which are social graph, privacy policy, co-owner invitation, majority vote, and conflict identification and provided solution. More details are explained as follows.

1. **Social graph**
   It is to create a graph that represents the social relationship among the users in the OSNs as demonstrated in Fig. 2. A node refers to a user in the OSNs. An edge presents relationship between two nodes. Label between

nodes indicates type or group of relationship and affinity level. In this research, preference of user is added in the social graph. This graph supports the notion about importance of relationship quality [6] because relationship between users influences making privacy decision.

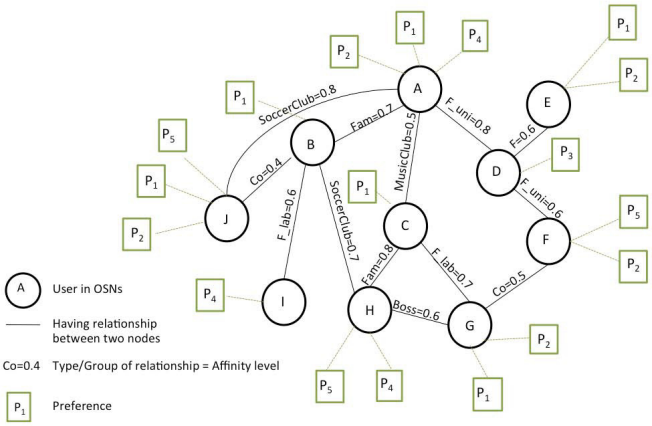

Fig. 2. The simple social graph in the OSNs

2. **Privacy policy**
The privacy policy is designed for limit the number of the users, who can see the collaborative information. Idea of this policy is that the owner tries to match this information to the target users, who might be interested in it. Nonetheless, when the co-owner wants to share the collaborative information, they can change the position to the owner position then create the privacy policy. The privacy policy helps alleviate the information overload by reducing amount of information and the owner can control the distance of the information, which is spread in the OSNs. The owner constructs the privacy policy by using four useful factors obtained from the social graph: type or group of relationship, affinity level, preference and distance for information distribution.

   – **Type or group of relationship** (T/G_Rela)
   In the OSNs, the users have ability to create a group for different purposes. It is fact that members in a contact list cannot have the same role in both the OSNs world and the real-world. Therefore, the T/G_Rela is subjected to each user such as friend, family, boss, co-worker and so on.
   – **Affinity level** (AL)
   This factor refers to how familiar two users are. Generally, the user is not able to give everyone in the contact list with same closeness level. The AL can range from 0.1 to 1.0 (denoting 0.1 is very unfriendly, 1.0 is very familiar)

- **Preference** (Pref)
  The user's preference presents the interests of the user such as music, movie, sport and so on. Using the user's preference is a good way because the information will not leak to other users, who are not interested in that information and does not cause the users annoyance.
- **Distance for information distribution** (Dist)
  It indicates how far the information can be spread to the other users. Controlling the distance can limit the number of the users, who can see the information. Nevertheless, it relies on the purpose of the owner. If the owner intends to spread the information as much as possible without privacy setting, it is possible that the information will be consumed by large number of users.

3. **Co-owner invitation**
   This component is proposed to inform the co-owners that they are a part of the collaborative information. It differs from previous works [8] [11] because this component makes the co-owners know that their information is being managed by others. In many cases, the co-owners lose the privacy that causes from the owner share the collaborative information without permission. This component additionally is important since it helps the co-owners realize whether or not posting this collaborative information might cause them trouble in the future. Then, the co-owner can make a decision by vote on the privacy policy created by the owner. After that the vote result will be collected and transferred to the majority vote component. The vote result of each co-owner is considered as the co-owner's status. The co-owner can be one of three statuses: acceptance, rejection and no response. Acceptance status means the co-owner agrees with the privacy policy. Rejection status indicates the co-owner denies the privacy policy or he/she needs to keep this information private. No response status presents the co-owner does not accept or reject in time.
4. **Majority vote**
   It has a duty to seek the consent of all co-owners as much as possible because allowing all of owner and co-owners to create the privacy policy is difficult to meet all of desires in one times. This component starts with gathering all of co-owners' status from the co-owner invitation component. Nonetheless, it takes time for collecting the vote results; so it needs to specific time. When the time is over, the co-owner with no response status will be moved to rejection status to protect the privacy. After status arrangement, the vote results will be counted if the number of acceptances is more than half of the vote results, this means that the collaborative information can be posted to the OSNs. The advantage of the majority vote is that if there is one co-owner rejects this privacy policy, he/she is still provided the privacy protection.
5. **Conflict identification and provided solution**
   It is designed for finding the cause of conflicts among the co-owners, who accept and reject the privacy policy, and making solution for those conflicts. Then, a list of target users is recommended to the owner. The owner can verify it before uploading the information. The conflicts are also identified

when sharing and re-share are occurred. In order to find the conflicts, the social graph is required because it can indicate how each user connects or has associated with. Moreover, it can represent the mutual friends as depicted in Fig. 2. User C is a mutual friend of user A and user B, while user H is a mutual friend of user B and user C. The mutual friend is necessary for detecting the conflict between the co-owners, who accept and reject the privacy policy. The proposed CPP shows the co-owners the list of users, who can see the collaborative information. These users have direct relationship with the owner and pass a condition of the privacy policy.

5 Experiment and Results

The experiment aims to analyze the factor and combination of factors, which help the information not leak to unwanted target users and to investigate the opinion of co-ownership by using the proposed architecture as shown in Fig. 1 and a questionnaire. The analysis results in this experiment will be used in the privacy policy, which is a part of the proposed CPP.

5.1 Experimental Setup

In order to study the factor and the combination of factors, which have influential on sharing sensitive information, a virtual social graph was built that helped the respondents imagined the flow of the information when it was shared in the OSNs. In this experiment, it was not created by the real data due to permission requirement. The virtual social graph consisted of 88 nodes, 201 edges as shown in Fig. 4. Each node referred to a user in the OSNs and one user had few preferences e.g. sport, music, game, food, travel. The edge presented a relationship and 10 affinity levels (ranging from 0.1 (very unfriendly) to 1.0 (very familiar)) between two nodes. In this experiment, the collaborative information additionally was assumed that it has associated with one owners and five co-owners. Three co-owners accepted the privacy policy created by the owner so that the vote results had a majority. There are 15 types for investigation, which compose of four groups according to the number of factors as follows in Table 1.

Each respondent were shown many scenarios as denoted in Fig. 3 according to types of information e.g. text, photo, video and link. The respondents then imagined that they were the owner and created the privacy policy that composed of the one factor and combination of factors. Also, they observed the flow of information with consideration of co-owners, who rejected the privacy policy. At the same privacy policy, the respondent swapped a position to co-owners and observed the flow of information again. Moreover, the co-owner could change a position from the co-owner to the owner in order to post this collaborative information.

Figure 4 indicates that the owner 0 created the privacy policy by using the combination of T/G_Rela and Dist (setting as 1 hop) factors. The co-owner 1 and 4 rejected this privacy policy. Therefore, the users, who have relationship

with the owner 0 and co-owner 1, and the owner 0 and the co-owner 4, could not see this information due to conflicts (user 20 and 13). Figure 5 shows the position change from the co-owner 2 in Fig. 4 to the owner 2. The owner 2 also created the privacy policy policy by using the same combination of factors. Nonetheless, the co-owner 4 and 6 rejected this policy. By consideration of results, it can be divided into two main groups. Firstly, users could not see the information because of rejection of the owner 4 (user 3) and not consistency with this policy (user 8 and 43), Secondly, the user can see the information due to acceptance of the co-owner 1 (user 20 and 42) and consistency with this policy (user 9, 10, 11, 22, 23, 24, 39, 45, 47).

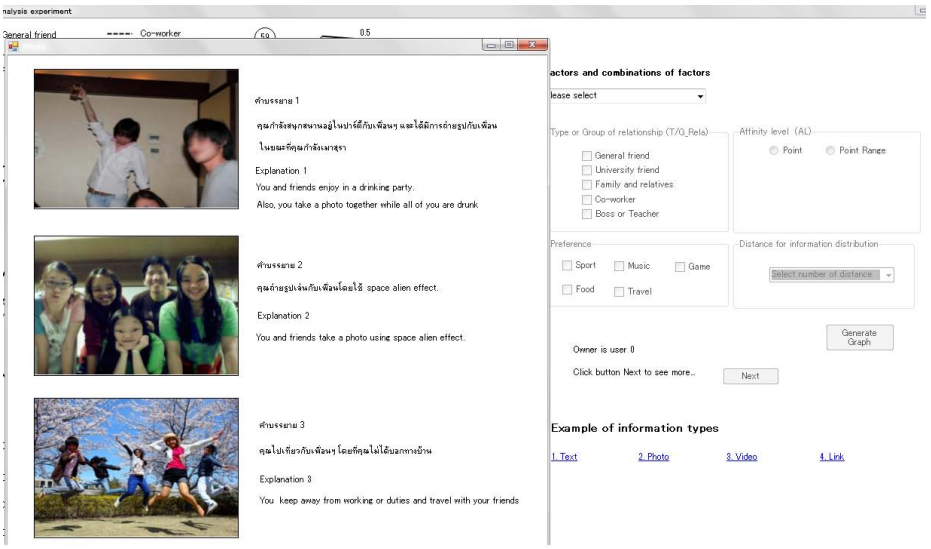

Fig. 3. An example of scenarios

After the experiment, the respondents did the questionnaire to evaluate the performance of each factor and combination of factors. The questions in the questionnaire were answered by 24 respondents: 14 male and 10 female. All respondents were asked about general questions, privacy in the OSNs, opinion of co-ownership as described in next Sect. 5.2

5.2 Results

Each question about the privacy in the OSNs and the opinion of co-ownership in the questionnaire was answered by Yes/No, explanation and a 5-point Likert scale. The performance of each factor and combination of factors was investigated by Mean and Standard deviation.

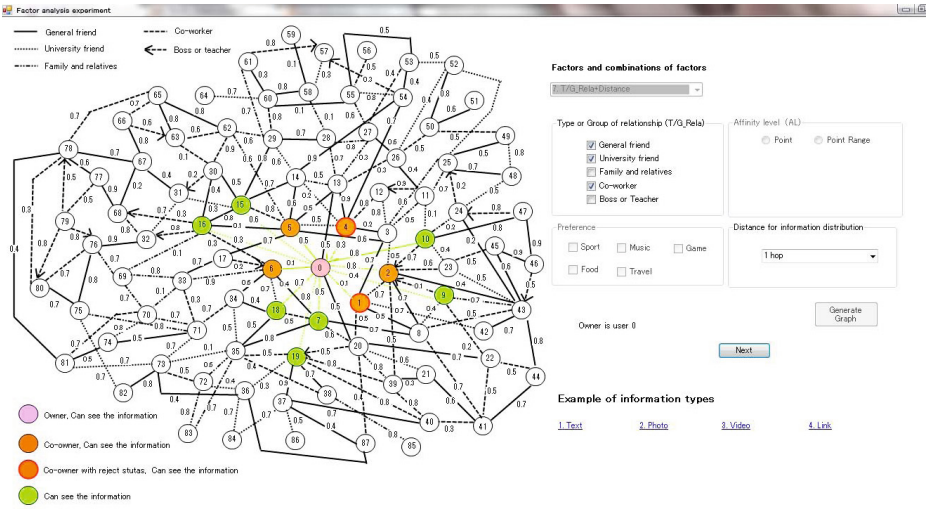

Fig. 4. An example of changing the position from the co-owner to the owner and results (before)

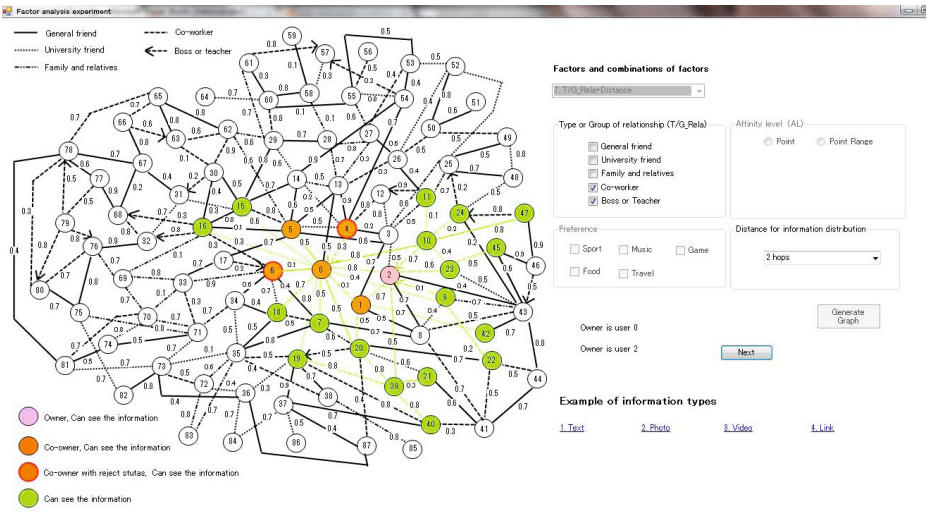

Fig. 5. An example of changing the position from the co-owner to the owner and results (after)

1. **General questions**
   Most of the respondents age range from 21 to 40 years old and have good experience in using the OSNs. They normally have more than one account. 62.5% of them spend time on using the OSNs about 1-4 hours a day. Purposes of using the OSNs generally are entertainment, information sharing, consuming and relationship maintenance.

2. **Privacy in the OSNs**
The results show that most of the respondents concerned the privacy with 3.91±0.99. 62.5% of them made an effort to avoid the sensitive information leak by using a custom setting. The remaining respondents use a default setting provided by the OSNs. However, they still complained that the privacy setting was difficult to understand [2] [13]. Moreover, it had many steps to complete setting. Some of them state that they would not post much information via the OSNs, which little provided the privacy setting such as Line (Timeline) or Instagram.

   Photo and text were considered that they easily leak to unwanted target users. This is because they were easy to recognize by others and using tagging and mention on the information can be done without permission.

   Family and Boss are the group of people, who the respondents did not want the sensitive information leak to. Although the respondents care feeling of them, the respondents still need private. Some opinion was expressed that if some sensitive information leaks to their family, it could lead to misunderstanding, worry, or disputation. Boss influenced career because several decisions relied on the boss. Posing the information sometimes could refer to an image of organization.

   Further interesting analysis, the respondents would not be worried much if the sensitive information leaked to people, who the respondents did not have relationship with [14].

   From Table 1, the number of factors influenced the privacy protection that if the number of factors increased, it helped the sensitive information not leak to unwanted target users. A combination of T/G_Rela, AL, Pref and Dist is considered as the most important for the privacy protection because it could filter unwanted target users. Although the performance of Pref factor or combinations containing this factor dropped, minority of the respondents believed that combinations with the Pref factor would scope the number of users, who had similar point of view.

   From the respondents' opinions, the T/G_Rela and the AL factors were a basis that the privacy setting should have. This was relevant to the results in Table 1. If the combination of factors contains the T/G_Rela and the AL factors, the performance increased.

3. **Opinion of co-ownership**
Almost all respondents have experience in losing privacy because the owner posted the collaborative information without their permission. Around 64% of these respondents faced trouble after the collaborative information, which was sensitive, was posted to the OSNs. They were worried about information leak with 3.86±0.77.

   There are two different opinions when the respondents were asked about asking the co-owner's permission before posting the collaborative information. The respondents imagined that they were the co-owner. 83.33% of them think that asking the permission was necessary with four reasons.

   – The respondents did not want the information leak to others, whom the respondent do not want to share with.

Table 1. Influence of factors on privacy protection

Factor	Personal information	Confidential business information	Freedom of expression	Improper morality	Behavior embarrassing
1. T/G_Rela	3.54±1.47	3.29±1.60	3.25±1.50	3.75±1.92	3.46±2.00
2. AL	3.42±1.50	2.92±1.53	3.29±1.30	3.46±1.50	3.33±1.34
3. Pref	2.63±1.73	2.67±1.61	3.63±1.38	2.75±1.39	2.54±1.25
4. Dist	3.63±1.12	2.88±1.56	2.92±1.05	3.08±1.21	3.04±1.23
5. T/G_Rela+AL	4.04±1.30	3.58±1.47	3.75±0.98	4.13±1.03	3.71±1.12
6. T/G_Rela+Pref	3.46±1.47	3.33±1.43	3.88+1.08	3.50±1.06	3.17±0.96
7. T/G_Rela+Dist	3.88±1.15	3.38±1.50	3.45±1.30	3.42±1.28	3.29±1.33
8. AL+Pref	3.30±1.31	2.98±1.43	3.83±1.12	3.29±1.23	3.13±1.26
9. AL+Dist	3.46±1.22	3.08±1.50	3.29±1.27	3.38±1.20	3.33±1.17
10. Pref+Dist	3.25±1.22	2.92±1.50	3.25±1.29	2.83±1.17	2.79±1.10
11. T/G_Rela+AL+Pref	3.92±1.32	3.63±1.44	4.08±0.97	4.17±1.00	3.96±1.00
12. T/G_Rela+AL+Dist	4.25±0.85	3.71±1.30	3.79±1.14	4.04±0.95	3.95±1.04
13. T/G_Rela+Pref+Dist	3.83+1.12	3.46±1.38	3.92±1.18	3.92±0.88	3.79±0.78
14. AL+Pref+Dist	3.67±0.96	3.50±1.35	3.96±1.20	3.77±0.95	3.58±1.06
15. T/G_Rela+AL+Pref+Dist	**4.50±0.93**	**3.79±1.44**	**4.21±1.22**	**4.38±0.71**	**4.21±0.72**

- They should have right to decide whether or not this information could be posted because the could not know which the information would cause them trouble in the future.
- Sensitivity level of privacy toward each information relied on person. In other words, each person has different privacy concern when seeing the same information.
- They should know their information is being managed by whom because they were worried who would see the information.

Nonetheless, the remaining respondents state that no need to ask their permission when they were the co-owner. Three reasons are explained below.

- They could not expect the owner to use the privacy setting, thus the co-owner should have to be careful the collaborative information by themselves.
- Giving the permission every times was boring task.
- They did not care much the privacy.

6 Discussion

Analysis results can imply that the respondents are worried when the collaborative information, which is sensitive, leak to the users, who the respondents have relationship with especially family and boss. They generally have many roles depending on society. They thus perform different behaviors when are in different societies. Although they want to post the collaborative information to the OSNs, they need private by not revealing some information to others because of negative feedback. As a result, most of the respondents believe that the combination of T/G_Rela, AL, Pref and Dist factors helps protect the privacy for leaking the information. On the other hand, they do not care much if the collaborative information will leak to other users, who the respondents have no

relationship with or not familiar with because they might not meet in the real world. Asking the co-owner's permission is expressed that the owner takes responsibility to the co-owners' privacy and it is suitable way although sometimes waiting for the permission might make the information not fresh or up to date.

7 Conclusion and Future Works

The CPP is proposed to balance between the collaborative information sharing and the privacy protection for the owner and co-owners by majority vote. It enables the owner to create the privacy policy and the co-owners to make a decision in the privacy policy by vote. It additionally identifies and solves the privacy conflicts because at least one co-owner intends to keep private. We have analyzed the factors, which help protect the privacy, the privacy in the OSNs, and opinion of co-ownership via the survey. Asking permission from the co-owners is necessary because it helps the collaborative information not leak to unwanted target users. For the future work, we plan to classify the sensitive information in order to help remind the owner before posting.

References

[1] Dhia, I.B.: Access control in social networks: A reachability-based approach. In: Proceedings of the 2012 Joint EDBT/ICDT Workshops, EDBT-ICDT 2012, pp. 227–232. ACM, New York (2012)
[2] Li, Q., Li, J., Wang, H., Ginjala, A.: Semantics-enhanced privacy recommendation for social networking sites. In: 2011 IEEE 10th International Conference on Trust, Security and Privacy in Computing and Communications (TrustCom), pp. 226–233 (2011)
[3] Hu, H., Ahn, G.J., Jorgensen, J.: Detecting and resolving privacy conflicts for collaborative data sharing in online social networks. In: Proceedings of the 27th Annual Computer Security Applications Conference, ACSAC 2011, pp. 103–112. ACM, New York (2011)
[4] Hu, H., Ahn, G.J., Jorgensen, J.: Enabling collaborative data sharing in google+. In: 2012 IEEE Global Communications Conference (GLOBECOM), pp. 720–725 (2012)
[5] Acquisti, A., Grossklags, J.: Privacy and rationality in individual decision making. IEEE Security Privacy 3(1), 26–33 (2005)
[6] Banks, L., Wu, S.: All friends are not created equal: An interaction intensity based approach to privacy in online social networks. In: International Conference on Computational Science and Engineering, CSE 2009, vol. 4, pp. 970–974 (2009)
[7] Gollu, K., Saroiu, S., Wolman, A.: A social networking-based access control scheme for personal content. In: Proceeding of the 21st ACM Symposium on Operating Systems Principles, SOSP 2007 (2007) (Work-in-Progress Session)
[8] Carminati, B., Ferrari, E., Perego, A.: Rule-based access control for social networks. In: Meersman, R., Tari, Z., Herrero, P. (eds.) OTM 2006 Workshops. LNCS, vol. 4278, pp. 1734–1744. Springer, Heidelberg (2006)
[9] Hart, M., Johnson, R., Stent, A.: More content- less control: Access control in the web 2.0. In: IEEE Web 2.0 Privacy and Security Workshop (2007)

[10] Squicciarini, A.C., Shehab, M., Wede, J.: Privacy policies for shared content in social network sites. The VLDB Journal 19(6), 777–796 (2010)
[11] Dinh, T.N., Shen, Y., Thai, M.T.: The walls have ears: optimize sharing for visibility and privacy in online social networks. In: CIKM, pp. 1452–1461 (2012)
[12] Adu-Oppong, F., Gardiner, C., Kapadia, A., Tsang, P.: Social circles: Tackling privacy in social networks. In: Symposium on Usable Privacy and Security, SOSP (2008)
[13] Church, L., Anderson, J., Bonneau, J., Stajano, F.: Privacy stories: Confidence in privacy behaviors through end user programming. In: Proceedings of the 5th Symposium on Usable Privacy and Security, SOUPS 2009. ACM, New York (2009)
[14] Wellman, B., Salaff, J., Dimitrova, D., Garton, L., Gulia, M., Haythornthwaite, C.: Computer networks as social networks: Collaborative work, telework, and virtual community. Annual Review of Sociology 22(1), 213–238 (1996)