Privacy-Respecting Access Control in Collaborative Workspaces

Stefanie Pötzsch, Katrin Borcea-Pfitzmann

Privacy-Respecting Access Control in Collaborative Workspaces

Stefanie Pötzsch and Katrin Borcea-Pfitzmann

Technische Universität Dresden, Faculty of Computer Science
D-01062 Dresden, Germany
{stefanie.poetzsch,katrin.borcea}@tu-dresden.de

Abstract. In these days' information society, people share their life with others not only in their direct, personal environment, but also on the Internet by using social software such as collaborative workspaces. In this context, an important issue is maintaining control over personal data, i.e., who is able to access which information. In this paper, we argue why traditional access control mechanisms are inappropriate for collaborative workspaces in general and present a concept for privacy-respecting access control in a web forum as an instance of collaborative workspaces.

Key words: Access Control, Collaborative Workspaces, Personal Data, Privacy, Web Forum

1 Introduction

The social life of these days' information society of the twenty-first century is altered in fundamental ways by technological developments. This includes the possibilities of social software to support users in sharing their life with others. In this context, two important issues are the promotion of a greater awareness for privacy among users of social software and enabling them to maintain control over personal data, i.e., to determine and enforce who is able to access which information. While there is research going on addressing the first problem [Pöt09a], [Pöt09b], in this paper we focus on the second point. As emphasized by [RI07], sharing of information is an important feature of social software, however not everything is intended to be shared with everyone. Typically, collaborative workspaces, which are in the focus of this paper, implement access control mechanisms that impart users only minimal freedom of decision in this regard. Users can only choose from a small set of "user groups", which are predefined by the administrator of the system, at the best. Hence, in order to enhance users' privacy in collaborative workspaces, there is a need for user-controlled and fine-grained access control to the content and meta-data generated by users during communication.

The paper is organised as follows: In the next section we introduce the basics of collaborative workspaces and highlight the role of personal data in this type of application. In Section 3 we argue why traditional access control mechanisms

are inappropriate for collaborative workspaces in general. Section 4 presents a concept for privacy-respecting access control in a web forum whereby the web forum represents one instance of collaborative workspaces. We conclude the paper with a brief summary and the indication of interesting points for further theoretical and practical research.

2 Collaborative Workspaces and Personal Data

Collaborative workspaces are infrastructures and platforms that enable users to work together, e.g., gathering information or creating contents in a collaborative manner or simply sharing data between each other, e.g., in a wiki, web forum or chat. The main feature of this type of social software is the collaborative creation and modification of content. Thus, the focus is on artifacts produced by a number of users. In addition, another type of social software are social networking sites, where the main focus is on user profiles and traversable connections between these profiles [PP09]. Figure 1 shows a graphical representation of the classification.

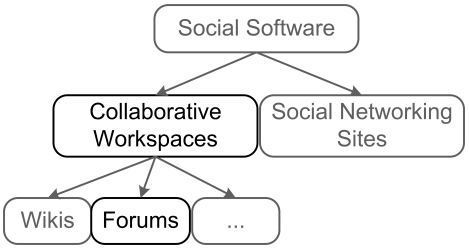

Fig. 1. Classification of social software

For this working paper we focus on collaborative workspaces, and, to be more specific, on web forums as a particular instance of collaborative workspaces. Web forums allow users to discuss particular topics by posting their opinions, experiences, or questions via a web form to a central data storage on the Internet. Several posts that refer to the same subject are grouped into a "thread". Figure 2 provides an overview of the hierarchical structure of content elements in a forum.

Contributions to a forum can contain *personal data* in terms of personal information, expression of thoughts and feelings of the writer. As pointed out by Adams [Ada99], it is important what is deemed sensitive or intimate in the perception of the individual rather than if it can be evaluated by third parties (e.g., lawyers, computer specialists). This argument emphasizes the need for user control with regard to the access control mechanisms implemented in collaborative workspaces. From a privacy perspective, the disclosure of personal data in collaborative workspaces is not target-aimed. Besides all the positive aspects that

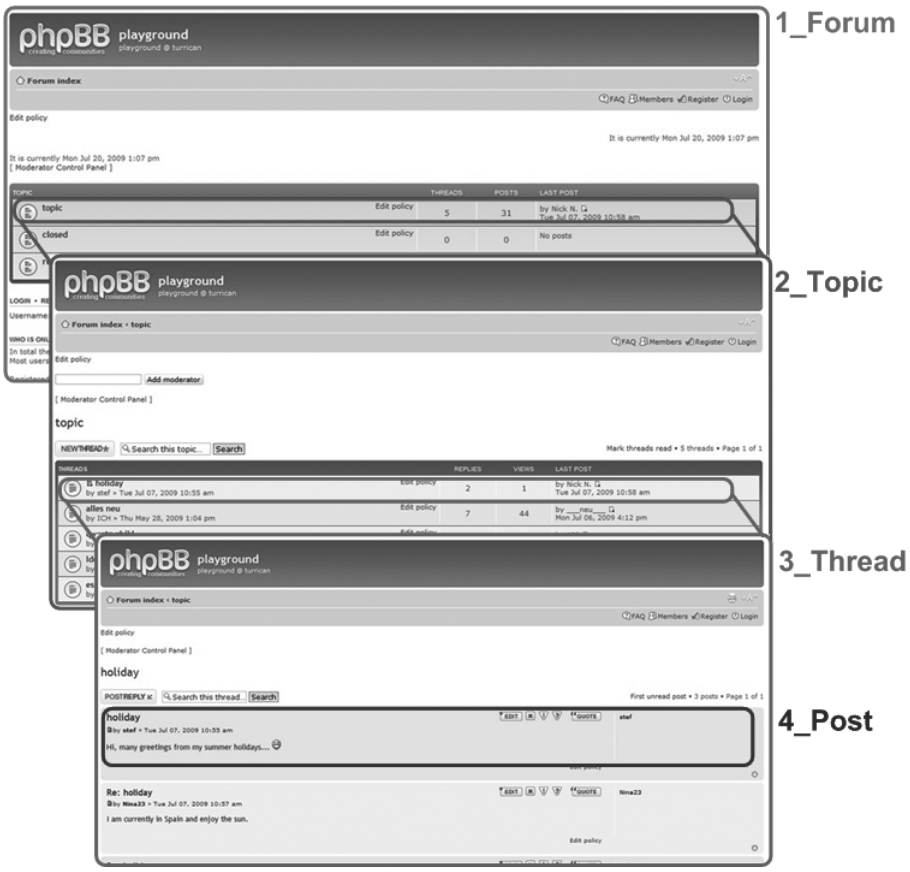

Fig. 2. Hierarchy of forum elements

users of web forums experience, sharing personal data with possibly millions of strangers on the Internet may result also in negative consequences, e.g., identity theft, surveillance, harassment, bullying or cyber stalking [PP09], [KWM+08], [Wor09]. Yet, from a social perspective, revelation to an intended audience is necessary for two reasons: First, the exchange of information, both personal and non-personal, is the major feature of the application and the motivation why people use it. Second, the exchange of implicit and explicit personal data allows users of collaborative workspaces to get an impression of the potential interaction partners and their situations. In this sense, the disclosure of personal data contributes to the success of social interactions and the forming of communities [Cut95].

Our approach for privacy-respecting access control works for all kinds of forums: for those that require a registration with an e-mail address and password or that allow users by other means to contribute under a unique pseudonym as

well as for forums that let users post completely anonymously. Usually, the users of a forum do not possess extensive member profiles stored with the forum service. However, different posts from the same user are still linkable by a nickname. Especially in the case of linkability between different posts from the same writer, privacy-aware users have an interest to restrict access to their contributions in order to limit the risk of being continually monitored by any third party.

3 Related Work on Access Control

Access control mechanisms allow to restrict the access to specified resources. A fair amount of currently available web forums are open to the public, at least with regard to read access. Thus, users of these forums disclose personal data potentially to the "whole world". Even if options for more selective access control are provided by the forum software, still only the technical administrator has the possibility to predefine different access rights on behalf of all users without being able to know about their individual preferences. The most common access control mechanisms used in social and collaborative software are:

- **Access Matrix Model** [Lam71]. The access control matrix is a table, which lists all users of the system in rows and all resources in columns. Each element $e_{u,r}$ of the table specifies the access rights that user u has on resource r. Reading the access control matrix column-by-column provides tuples of rights and user names for each resource, called access control lists (ACLs). Reading the table line-by-line results in a capability list (CL), i.e., for each user it is indicated what access rights she is granted to which resources.

- **Role-Based Access Control** [SCFY96]. Role-based access control mechanisms are similar to ACLs, with this difference: user names are assigned to one or more fixed roles and for each resource it is defined which role is allowed to perform which action.

- **Team-Based or Group-Based Access Control** [Tho97]. For this approach, user names are grouped in teams or groups, e.g., according to their current context of work. Access rights to resources are assigned to these teams or groups, respectively.

A detailed comparison of advantages and disadvantages of these mechanisms with regard to their applicability in the area of social software can be found in [TAPH05], [FWBBP06] and [RI07].

All of the mechanisms indicated above are based on the idea of the existence of an administrative party (e.g. the provider) that defines lists, roles, or groups and assigns the names of all users of the system to these lists, roles, or groups in order to enable the management of access to resources. Even if this management task would not depend on a single administrative party, but could be set by each user for her own contributions as for instance suggested by Razavi and Iversion [RI08] in their approach for social networking sites, another problem remains. In

order to assign a user name to a list, role, or group, it would still be necessary to know about the existence of the user name. The author of a contribution and the user who is to be granted access need at least to meet once - in the physical world or virtually. However, in a public forum for instance, the author of a post is potentially looking for new contacts, who fulfill specified requirements, e.g., live in the same city or are a member of the same fitness centre. This is, the author is not able in any case to specify the names of all other users who should have access to the contribution. Both requirements, namely (i) the existence of an administrative party who decides about access control settings and (ii) that user names are known, are our strongest points of criticism and the reason why we do not consider the introduced approaches as applicable for user-controlled and privacy-respecting access control in collaborative workspaces.

4 Concept For User-Controlled and Privacy-Respecting Access Control

We suggest to enhance the access control features available in the forum software by a finer grained and privacy-respecting approach. This implies that access control policies should be possible to specify not only for the whole forum or for topics, but additionally also for threads and particular posts. The policies need to be set by the user being the owner of the personal data instead of by an administrative party. Furthermore, to respect the privacy of readers, our access control concept for collaborative workspaces must not rely on user names. Instead, according access rules should indicate which properties or certificates someone has to prove to get access to the corresponding resource. Forum platforms typically provide the roles “adminstrator” for addressing technical issues and “moderator” for content-related moderation of topics. Our approach should allow to keep both roles. Hence, we have to consider the following requirements for user-controlled and privacy-respecting access control in a forum whereby these are easily generalisable to further kinds of collaborative workspaces:

- No administrative party, but each user should be able to define and modify access rules to her contributions, i.e., personal information, expression of personal thoughts and feelings.
- Other persons, who should or should not be able to access the personal data are not necessarily known by the user.
- These other persons also have an interest to protect their privacy.
- User-controlled and privacy-respecting access control can be applied to different levels of content granularity (e.g. forum, topic, thread, post).
- An administrator of the forum should be able to address technical issues of the platform, but should not necessarily have access to content data.
- Moderators should be able to moderate particular topics.
- The owner of a resource is always able to have access on it.

To address these points, we propose to let the user define access control policies together with her contributions indicating the attributes a reader has to possess

and to prove. In order to protect the privacy also of the other persons, properties or attributes should be presentable in an anonymous way and not linkable when repeatedly used. This requirement can be fulfilled using the concept of *anonymous credentials* proposed by Chaum in 1985 [Cha85] and technically realised in the Identity Mixer (short: Idemix) system [CvH02]. The idea of access control based on anonymous credentials and policies is not new in general and was demonstrated in selected use cases for user - service provider - scenarios in the project PRIME ([ACK+09], [HBPP05]). We build on the results of PRIME and investigate the applicability of the concept and the implementation in collaborative scenarios between a number of users, where all parties have an interest to protect their privacy on the one hand, but to engage in social interaction on the other hand. Whereas [FWBBP06] started from scratch and did a prototypical implementation of an e-Learning application that was specifically designed to work with PRIME technology, this paper devotes to enhancing existing forum software (phpBB). It demonstrates the feasibility of maintaining existing concepts of the platform and integrating new privacy-enhancing functionality at the same time.

All proofs for attributes including the proof for possessing a particular role (that may be required by an access control policy) can be realised by showing the appropriate credential. This implies that the process of creating a new resource includes that the originator of that resource receives the corresponding credential (`cred:Owner-`*`Thread-ID`* or `cred:Owner-`*`Post-ID`*) from the forum platform and stores it on the local device. The roles *administrator* and *moderator* can be realised with help of the credential-based access control approach as well, i.e., the according credentials (`cred:Admin-Forum` and `cred:Moderator-`*`Topic-ID`*) are issued to the corresponding persons. Together with a new resource, default access control policies are created, which ensure that users who show the administrator credential or moderator credential get the required access granted to fulfill their roles. The owner of a resource possessing the owner credential always has access to that resource and can modify the access control policies to, e.g., allow further other users with certain provable properties read and maybe also write access to the resource.

In general, credentials are offered by particular organisations, so called *credential issuers*. Credential issuers need to be known to the public, so that everybody has a chance to get credentials certifying attributes of the user. In the course of this paper, we need to assume an existing infrastructure of credential issuers. Regarding the question which credentials can be used in the access control policies, there are two possibilities: Either a set of all possible credentials needs to be globally defined or a generally accepted standard for defining new credentials is required. Certainly, both alternatives have advantages and disadvantages. The efforts and costs of determining a globally defined set of credentials are comparable with the assumption of knowing all user names. Yet, having knowledge of all credentials instead of all user names offers an improvement in terms of privacy. More flexibility in the definition of credentials and, connected to this, also in the definition of access control policies, can be provided if a general standard

of credential definition would exist. Originators of resources could apply this standard to specify the possession of which attributes a resource requester has to prove. If someone tries to access a resource and that user does not possess the corresponding newly defined credential, which is requested by the access control policy, she needs to be able to get information which credential is required and how to get it. A detailed analysis of possible attacks on privacy and the requested trust structures for both of the sketched alternatives lies beyond the scope of this paper, however it indicates an interesting point for further research.

The following example scenario serves as demonstration how access control based on credentials and access control policies in a web forum should work:

> *Assuming someone – let's call him Hannes – posts a message to the thread "Fit for summer" in a publicly accessible forum. The access control policy of the thread is derived from the parent topic, which is set to be open for reading and writing exclusively for people who have proven to be male. Hannes additionally restricts access to his post to allow only men being member of the same fitness centre, which Hannes attends.*

Table 1. Example of an Access Control Policy

(1) Forum:	[(cred:Admin-Forum) OR (everybody*[default]*)] AND
(2) Topic:	[(cred:Moderator-SportsAndCars) OR (everybody*[default]*)] AND
(3) Thread:	[(cred:Moderator-SportsAndCars) OR (cred:Owner-FitForSummer) OR (cred:male)] AND
(4) Post:	[(cred:Moderator-SportsAndCars) OR (cred:Owner-PostFromHannes) OR (cred:memberOfFitnessCentreXYZ)]

Whenever someone requests access to Hannes' post, the access control policy is evaluated according to the hierarchical order of content elements of the forum (cf. Table 1). In our example, step (1) ensures that authorised users are either an administrator of the forum or – since we have chosen a public forum for the example – any regular user. Step (2) specifies that users are allowed to read the topic "Sports and Cars" if they are a moderator of this topic or anybody else. The latter applies since the example does not specify any restriction on topic level as well. Step (3) ensures that only users who are either moderator of the topic "Sports and Cars" or who are owner of the thread or who are male get read access to the thread "Fit for summer". At last, step (4) determines that only users who are either moderator of the topic "Sports and Cars", owner of the post, or a member of the fitness centre XYZ can read the post created by Hannes. Accordingly, read access to Hannes' post is only granted if the whole policy (steps 1 – 4) is evaluated to be "true". Similar to this example for *read access*, further policies need to be defined in order to specify *add, edit* or *delete* rights of a resource. All users who add a post to a particular thread have the opportunity to further restrict access to their own contribution. Obviously, it is

not possible for them to overwrite access control policies of parent elements (or any other element) for which they do not possess the corresponding credentials.

If the presented access control concept is used in a very restrictive way, forum users will experience a high level of privacy but a low amount of interactions. Vice versa, if the access control is handled very open users could lose much of their privacy. Certainly, it would be inappropriate to use the proposed privacy extension for every contribution in a forum. However, having this feature at disposal may encourage privacy-aware users to discuss issues, which they would not address in public forums, or to state unpopular and uncensored opinions to a specified audience. In order to practically test the approach in real-life scenarios and to collect data aiming at determining and study the compromises between privacy and social interaction that different types of forum users make, we currently work on the implementation of the presented concept.

5 Conclusions and Future Work

In the paper, we showed that traditional access control mechanisms are inappropriate for privacy-respecting access control in collaborative workspaces. Further, we elaborated requirements on user-controlled and privacy-respecting access control in a web forum as an instance of collaborative workspaces. We presented a concept of enhancing existing features of a forum with access control mechanism based on anonymous credentials and access control policies. These can be individually specified by each user for her contributions on a fine-grained level. This way, the requirements outlined in Section 4 can be fulfilled.

Hence, specifying individual access control rules on content items represents a useful privacy enhancement of the application, but it also requires additional effort from the users. Further research is needed to investigate whether and how users become able to understand their benefit from this additional effort. The concept for privacy-respecting access control in collaborative workspaces described in this paper is currently being implemented by extending the access control component of the popular forum software *phpBB*. This is, we will be able to conduct experiments with end users in the near future and report about the practical applicability of our approach.

So far, the privacy enhancement is completely based on advances of the access control mechanism of the forum. In the future we want to elaborate on questions related to the management of different predefined access control settings and pseudonyms in collaborative scenarios. Building on technical solutions for scenarios where the user interacts with a single provider, which are already developed within PRIME, we want to point out that interactions between an arbitrary number of users are expected to be more dynamic with regard to access control preferences and the use of pseudonyms. For instance, Alice has used a pseudonym $\mathtt{A}_B$ to communicate with Bob and a pseudonym A_H when talking to Hannes. In case Alice starts an interaction, which involves both Bob and Hannes, she needs to decide which pseudonym to chose for this communication. Thus,

existing mechanisms need to be extended and to proof their suitability for real life in our setting.

Another interesting point that needs detailed discussion and elaboration is the question of credential-issuing and revoking of credentials in case the certified claim is no longer valid, e.g., a person no longer attends the indicated fitness centre or is no longer moderator of a topic in the forum.

Acknowledgments. We thank Rainer Böhme, Stefan Köpsell, and our anonymous reviewers for their helpful comments as well as Hagen Wahrig for working on the practical realisation of the concept. The research leading to these results has received funding from the European Community's Seventh Framework Programme (FP7/2007-2013) under grant agreement no. 216483.

References

[ACK+09] C.A. Ardagna, J. Camenisch, M. Kohlweiss, R. Leenes, G. Neven, B. Priem, P. Samarati, D. Sommer, and M. Verdicchio. Exploiting Cryptography for Privacy-Enhanced Access Control: A result of the PRIME Project. *Journal of Computer Security (JCS)*, 2009. to appear.

[Ada99] Anne Adams. The implications of users' privacy perception on communication and information privacy policies. In *In Proceedings of Telecommunications Policy Research Conference*, Washington, DC, 1999.

[Cha85] David Chaum. Security without identification: Transaction systems to make big brother obsolete. In *Communications of the ACM*, volume 28, pages 1030–1044, 1985.

[Cut95] R. H. Cutler. Distributed presence and community in cyberspace. *Interpersonal Computer and Technology*, 3(2):12–32, 1995.

[CvH02] Jan Camenisch and Els van Herreweghen. Design and implementation of the idemix anonymous credential system. In *Proceedings of the 9th ACM conference on Computer and communications security*, pages 21 – 30, 2002.

[FWBBP06] Elke Franz, Hagen Wahrig, Alexander Böttcher, and Katrin Borcea-Pfitzmann. Access Control in A Privacy-Aware eLearning Environment. In *First International Conference on Availability, Reliability and Security*, pages 879–886, 2006.

[HBPP05] Marit Hansen, Katrin Borcea-Pfitzmann, and Andreas Pfitzmann. PRIME - Ein europäisches Projekt für nutzerbestimmtes Identitätsmanagement. *it - Information Technology, Oldenbourg*, 6(47):352–359, 2005.

[KWM+08] Da-Yu Kao, Shiuh-Jeng Wang, Kush Mathur, Saransh Jain, and Frank Fu-Yuan Huang. Privacy Concealments: Detective Strategies Unveiling Cyberstalking on Internet. In *APSCC '08: Proceedings of the 2008 IEEE Asia-Pacific Services Computing Conference*, pages 1364–1368, Washington, DC, USA, 2008. IEEE Computer Society.

[Lam71] B. Lampson. Protection. In *In 5th Princeton Symposium on Information Science and Systems*, pages 437–443, 1971.

[Pöt09a] Stefanie Pötzsch. *Privacy Awareness: A Means to Solve the Privacy Paradox?*, volume 298 of *IFIP Advances in Information and Communication Technology*, pages 226–236. Springer, Boston, 2009.

[Pöt09b] Stefanie Pötzsch. Untersuchung des Einflusses von wahrgenommener Privatsphäre und Anonymität auf die Kommunikation in einer Online-Community. In S. Fischer, E. Maehle, and R. Reischuk, editors, *Informatik 2009, Im Fokus das Leben, 28 September - 02 October 2009, Lübeck*, volume 154 of *Lecture Notes in Informatics*, pages 2152 – 2165, Bonn, 2009. Gesellschaft fr Informatik.

[PP09] Martin Pekárek and Stefanie Pötzsch. A comparison of privacy issues in collaborative workspaces and social networks. *Identity in the Information Society*, 2009. Special Issue on Social Web and Identity.

[RI07] M. N. Razavi and L. Iverson. Towards usable privacy for social software. Technical Report LERSSE-TR-2007-03, University of British Columbia, 2007.

[RI08] M. N. Razavi and L. Iverson. Supporting selective information sharing with people-tagging. In *In Proceedings of the ACM CHI '08 Extended Abstracts on Human Factors in Computing Systems*, Florence, Italy, April 2008.

[SCFY96] R. S. Sandhu, E. J. Coyne, H. L. Feinstein, and C. E. Youman. Role-based access control models. *IEEE Computer*, 29(2):38–47, 1996.

[TAPH05] William Tolone, Gail-Joon Ahn, Tanusree Pai, and Seng-Phil Hong. Access control in collaborative systems. *ACM Comput. Surv.*, 37(1):29–41, 2005.

[Tho97] Roshan K. Thomas. Team-based access control (TMAC): a primitive for applying role-based access controls in collaborative environments. In *RBAC '97: Proceedings of the second ACM workshop on Role-based access control*, pages 13–19, New York, NY, USA, 1997. ACM.

[Wor09] Online harassment and cyberstalking cumulative statistics for the years 2000-2008, 2009. http://www.haltabuse.org/resources/stats/Cumulative2000-2008.pdf.