AN ACCESS CONTROL MODEL FOR MASSIVE COLLABORATIVE EDITION

Juliana de Melo Bezerra, Celso Massaki Hirata and Edna Maria dos Santos
Computer Science Department, Instituto Tecnologico de Aeronautica, S. J. Campos, Brazil

Keywords: Access control, RBAC, Collaborative editing, Massive collaborative editing.

Abstract: The Web has enabled the elaboration of documents, through editing systems, by a large number of users collaboratively. An access control model is essential to discipline the user access in these systems. We propose an access control model for massive collaborative edition based on RBAC, which takes into account workflow, document structure and organizational structure. Workflow is used to coordinate the collaboration of participants. The document structure allows the parallel edition of parts of the document; and organizational structure allows easier management of users. We designed and implemented an editing system based on the model. We show that the model is useful to specify collaborative editing tools and can be used to categorize and identify collaborative editing systems.

1 INTRODUCTION

Collaborative editing (CE) systems aim to support a distributed group of people editing a shared document collaboratively over a computer network. The major benefits of CE include reduced task completion time by taking advantage of parallelism and improved solution quality by leveraging collective intelligence. Over the past decades, a large number of CE systems have been developed in academia as well as in industry, but most of them were designed for single group and single shared document.

The concept of Massive Collaborative Edition (MCE) is not new. Massive Collaborative Edition is seen as the task of editing a document by a large number of users (ECOO); the number of users may vary from tens to millions. The most well-known example of a system that employs MCE is Wikipedia.

The definition of MCE is starting to be accepted, and we argue that its access control model should be more elaborated. In our view, the document can be structured in a specific manner so that dependencies of edition between the items of documents may arise. Similarly, users can be organized in various groups that can have different roles in the edition of documents. In a single group, users in general have a single role related to the edition of the document (editor role). In some systems users may have other complementary roles such as reviewers and coordinators, but they all in general belong to a single group. Therefore a more complex organization of users is required. In general, groups are part of an organization and their relations form the structure.

Besides, MCE may require a process that can be defined using workflow. Workflow is the automation of a business process during which documents, information or tasks are passed from one participant to another (WfMC, 2003). So, the access control model for MCE shall also include workflow.

RBAC (role-based access control) (NIST, 2001) has been successfully employed in many systems. We propose an extended RBAC model to take into account workflow, document structure and users organizational structure in order to address Massive Collaborative Edition. Our goal is to provide a comprehensive model to aid in the construction of massive collaborative edition tools.

The article is organized as follows. Section 2 discusses MCE, RBAC, workflow, document structure and users organizational structures. Section 3 presents the model for Massive Collaborative Editing, and discusses its possible usages. Section 4 presents some related work and Section 5 concludes our work and discusses future work.

2 FOUNDATIONS

In this section, some concepts are described to

de Melo Bezerra J., Hirata C. and dos Santos E. (2010).
AN ACCESS CONTROL MODEL FOR MASSIVE COLLABORATIVE EDITION.
In *Proceedings of the 12th International Conference on Enterprise Information Systems - Information Systems Analysis and Specification*, pages 135-140
DOI: 10.5220/0002868401350140

support the proposed access control model for MCE. Most of the concepts are originated from well established subjects such as RBAC, Workflow, and Organization Structure.

2.1 MCE

The most prominent example of MCE tool is wiki and, to some extent, Content Management System. A wiki is a website that allows the easy creation and edition of any number of interlinked Web pages, using a WYSIWYG (What You See Is What You Get) text editor, within the browser. Most wikis serve a specific purpose and material is promptly updated by the user community. Such is the case of the collaborative encyclopaedia Wikipedia. Open purpose wikis accept all sorts of content without rigid rules as to how the content should be organized and who can create and modify the pages.

A content management system (CMS) is a collection of procedures used to manage work process in a collaborative environment. The procedures allow for a large number of people to access shared data according to user roles. User roles are used to define what information users can view or edit.

Wiki can be seen as a web content management system (WCMS) software, implemented as a Web application, for creating and managing HTML content. The software provides authoring (and other) tools designed to allow users with little knowledge of programming languages or markup languages to easily create and manage content.

The two tools, wiki and WCMS, illustrate why the web is becoming a space where hundreds or thousands of people share their knowledge and resources. One of the reasons of the success of these collaborative tools is that they do not require any specific skills for publishing and editing. However, the tools provide limited functionalities for collaborative authoring of shared documents.

Based on the previous experience in research of collaborative editing, the ECOO project-team has been concerned with various issues specific to the mass collaboration such as consistency maintenance of structured data, consistency maintenance in peer-to-peer environments and awareness aspects in large groups. A research direction adopted by ECOO team is to use P2P techniques to distribute collaborative documents. The P2P techniques raise the issues of supporting collaborative edits, and of maintaining consistency, over a massive population of users, shared documents, and sites. Ignat and Norrie (2008) studied a number of alternative P2P, decentralized approaches, applied to collaborative wiki editing, contrasted with current centralized systems. However no access control model is provided.

2.2 RBAC and Workflow

The RBAC (Sandhu et al., 2000), role-based access control (also called role-based security), has become the predominant model for advanced access control because it reduces the complexity and its related cost. Within an organization, roles are created for various job functions. The permissions to perform certain operations are assigned to specific roles. Users are assigned to particular roles, and through those role assignments, acquire the permissions to perform particular tasks. Since users are not assigned to permissions directly, but only acquire them through their role (or roles), management of individual user rights becomes a matter of simply assigning appropriate roles to the user.

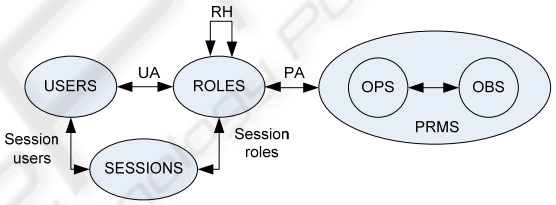

Figure 1: RBAC model.

When defining a hierarchical RBAC model (Figure 1) the following conventions are useful (Sandhu et al., 2000):

- OBS: the set of objects. An object is a resource used in some operation.
- OPS: the set of operations. An operations is an activity that can be performed (e.g. edit, save, and to include a picture).
- USERS: the set of users.
- ROLES: the set of roles. A role is a job function or title which defines an authority level.
- PRMS: the set of permission. A permission is an approval of a mode of access to an object.

$$PRMS = 2^{(OPS \times OBS)}$$

- PA (permission assignment): a many-to-may permission-to-role assignment relation.

$$PA \subseteq PRMS \times ROLES$$

- UA (user assignment): a many-to-many user-to-role assignment relation.

$$UA \subseteq USERS \times ROLES$$

- SESSIONS: the set of sessions. A session is a mapping between a user and a possible role.
- RH (role hierarchy): a partial order on the ROLES called inheritance relation, written as $\geq$.

For r1, r2 ∈ ROLES, it is true that r1 ≥ r2, only if all permissions of r2 are also permissions of r1.

The RBAC simplifies authorization management, because authorization can be administrated as a whole for all users belonging to a role, rather than at the level of individual users. The hierarchical characteristic of RBAC is granted due to the possibility to define a role hierarchy. The use of RBAC to manage user privileges within a single system or application is widely accepted as a best practice.

A workflow is a pattern of activity enabled by a systematic organization of resources, defined roles and information flows, into a work process that can be documented and learned. Workflows are designed to achieve processing intents of some sort, such as service provision or information processing. In collaborative editing, a workflow consists of a sequence of connected editing tasks over a document, i.e. a partial order of tasks. With respect to RBAC, each task is seen as a subset of permissions (operations over objects). Although users have the permissions stated by their roles, only the permissions specified by the current task in the workflow can be performed by the users.

Workflow is required to MCE systems since users are expected to perform different tasks through different roles at different situations or states.

2.3 Document Structure

In a Massive Collaborative Edition, a document is edited by a large number of users. The document can be classified into basic or structured.

Basic documents are artifacts that can be independently published. They are self contained and they do not rely or rely very little on other contents. In other words, their edition can be independent from the edition of other artifacts. Examples of basic documents are research articles and newspaper articles in proceeding and newspaper, respectively.

On the other hand, there are artifacts whose contents are dependent on the contents that are placed on other artifacts. Therefore their edition must take into account the edition of the documents which they rely on. They are called structured documents. Example of that artifact is book chapter which rely on other chapters. For some documents, the dependency can be weak and they can be seen as basic documents. An example of that document is pages of Wikipedia.

The document organization enables the parallel work of items of document by various users simultaneously. In order to discipline the access control by various users on the artifacts, some scheme of control is required.

2.4 Organization Structure

Organizations are a variant of clustered users. The structure of an organization determines the modes in which the users collaborate to meet a goal (Daft, 2008). Organizational structure allows the expressed allocation of responsibilities for different functions and processes to different entities such as the branch, department, workgroup and individual.

Users within the functional divisions of an organization tend to perform a specialized set of tasks, for instance the engineering department would be staffed only with engineers. Functional organization leads to operational efficiencies. As examples of organizanial structure, there are hierarchical and matrix organizations.

The hierarchical organization groups each organizational function into a division (or product). Each division contains all the necessary resources and functions within it. An example of divisional structure is the computer manufacturer, with the following divisions: printers, computers, and cameras. Each division has its own personnel (users) and tasks. The organization business objectives are met through the divisions' work.

The matrix organization groups users by both function and division (or product). A matrix organization is a response to the weaknesses of functional structure. Matrix organization, however, may lead to management conflicts since users can be subordinated to two different managers. An example of matrix organization is a university that is structured into departments (Computational, Electrical, Mechanics, etc) and pro-rectories (for instance, graduate and undergraduate).

An organizational structure may group hundreds to thousands of users who have positions in the organization. In MCE, the positions may both be structured, for instance in a hierarchy, and may have specific roles, therefore some the access control model to discipline of the access control of positions should be available.

3 ACCESS CONTROL MODEL FOR MASSIVE COLLABORATIVE EDITING

The access control model for Massive Collaborative

Edition is based on the RBAC model and considers: workflow, document structure, and organization structure.

3.1 Key Components

In our proposed model, the document can be structured in items, which are the objects of RBAC. The workflow is a partial order of tasks (subsets of permissions) in order to accomplish its life cycle. Besides, the organization is represented by a structure of groups that contain users. So, a user can belong to multiple groups; while a group can have multiple roles.

Figure 2 depicts the proposed model. The majority of the conventions are from RBAC, except for the following:

- OS (Object Structure): a partial order on the OBS written as ≥. For o1, o2 ∈ OBS, it is true that o1 ≥ o2, only if o2 is a part of o1, consequently all permissions related to o2 are also permissions to o1.
- Workflow: a partial order on the tasks written as ≤. A task is a subset of permissions. For t1, t2 ∈ 2^{PRMS}, it is true that t1 ≤ t2, only if it is required to perform t1 before t2.
- GROUPS: the set of groups. A group comprises of users.
- GS (Group Structure): a partial order on the GROUPS written as ≥. For g1, g2 ∈ GROUPS, it is true that g1 ≥ g2, only if all roles of g2 are also roles of g1.
- UA (user assignment): a many-to-many user-to-group assignment relation.

$$UA \subseteq USERS \times GROUPS$$

- GA (group assignment): a many-to-many role-to-group assignment relation.

$$GA \subseteq ROLES \times GROUPS$$

- SESSIONS: the set of sessions. A session is a mapping between a user, a group and a possible role.

Items of document (or objects) are assigned to operations. This assignment is called permission. Permissions are organized in workflow tasks. Permissions are also assigned to role. However, a role has permissions only if the current task of the workflow enables those permissions. So, the tasks discipline the permissions during the work process.

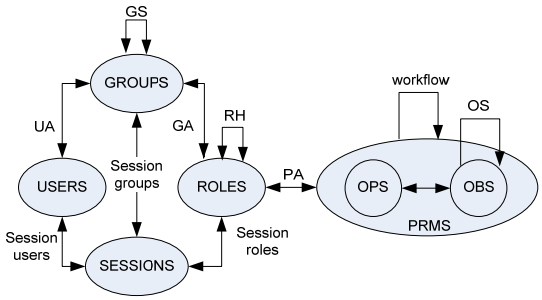

Figure 2: Model for Massive Collaboration Editing.

3.2 Discussion

In the proposed model, the objects can be structured, for instance into a hierarchical structure. The catalogue of products in virtual shop or courses of university are example of such a structure. In general, the completion of an operation of objects in higher level depends on the completion of operations of lower level in its tree. In order to ensure the upwards completion, workflows should be defined so that completions of operations of lower levels enable the completion of operations in higher level. A similar policy can be defined to start operations of sub-objects. In this case, the downwards initiation can be used. With workflow, objects have a discipline to be worked with.

In a hierarchical organization, users have positions and positions are structured according to a tree hierarchy. In general, positions of higher hierarchies have access or can perform roles of lower hierarchy positions. In order to ensure the downwards authorization, groups should be defined so that users of higher positions must belong to all groups of lower positions in the tree. With groups, the assignment of roles to groups is easier to both make and verify. So, the group structure (GS) allows that a user to have at least the same access of his/her subordinates in a hierarchical organization. The proposed model can also be applied to a matrix organization, where a user can belong to two groups: the functional group (e.g. engineering group) and the technical one (e.g. electrical system of some product).

It is possible to include organization aspects using only RBAC by representing each group as a role. However, it is not easy to adapt this model in case of organization changes. So, our approach separates roles and groups. In our approach the organization is well defined and structured into groups, and any change in this structure can be reflected to the access control model. As examples of organization changes, there are the creation, elimination, and fusion of groups.

The proposed model also allows for user delegation and organizational cooperation. Delegation is to give to a user u_1 the same permissions of another user u_2 during a specific period. In general during this period, u_2 is out of work. The delegation can be represented in the proposed model by adding u_1 into the same group of u_2. Besides, organizational cooperation can be represented too. Cooperation between companies occurs when a company c_1 is contracted by another company c_2 in order to develop together a solution (product or service). In this case company c_2 has its own access control model and needs to give access to c_1 for specific objects. So, an organization structure can be introduced into the existent model to represent c_1 and the relations between groups from c_1 and c_2 can be well established.

The proposed model is quite general and can be used for both specifying collaborative editing tools and determining classes of collaborative tools. The specification of massive collaborative editing systems may require the configuration of web tools, which may be awkward and tiresome. The configuration can require thousands of assignments of users, roles, operations, and objects to provide the functionality with access control if one chooses to use the conventional RBAC model. Using the proposed model, the number of assignments can be reduced significantly and are easier to make.

Our experience with the construction of collaborative editing system of a graduate studies catalogue, with 18 users, 21 groups, 21 document items, 1 workflow, 6 roles, and 5 tasks, showed a significant reduction of assignments of roles to the users. It was possible to reduce from 178 to 74. In the implementation of the catalogue, we also benefited from other reductions. The proposed model allows using one workflow definition and 6 role definitions for all the document items. It is expected that for larger numbers of users and document items, the reductions of assignment numbers are very significant.

The group structure allowed also an easier maintenance of the system. When a user enters or leaves the system (or when she/he is replaced) it suffices to make changes in the groups that she/he belongs to.

We also advocate that the model can be used to classify the collaborative editing systems. For instance, Wikipedia can be seen as an instance of the model with the following definitions:

- USERS: any web user.
- GROUPS: world (not structured).
- OBS: any page regarding some subject.
- OPS: edit.
- PRMS: edit any object.
- ROLES: collaborator, which can edit any page.

Since there is one group in Wikipedia, every user belongs to it. The group has just one role (collaborator), and this a role is assigned to the group. The objects are not structured. Since there is one permission, a workflow is not defined. Therefore, it is possible to use instances of the model to classify collaborative systems. It is not difficult to model other systems such as Google Docs (2008). The model for Google Docs is presented below:

- USERS: any person invited by the document owner.
- GROUPS: world (not structured).
- OBS: a document, including DOC, XLS, ODT, ODS, RTF, CSV, PPT, etc.
- OPS: to create, to read, and to change a document by inserting tables and images, adding comments and formulas, changing letter style, etc. It is also possible to organize the documents in folders, and to publish the document as a web page.
- PRMS: the same as OPS, but applied to a specific document.
- ROLES: author, reader and collaborator. The author is the document owner. Readers and collaborators are chosen by the author. A reader can only read the document, while a collaborator can read and change it.

4 RELATED WORK

The related work proposes extensions of RBAC model that can partially address MCE systems, according to the following characteristics: workflow, document structure and organization structure.

The concept of document structure is proposed by Buegge et al. (2006) within the scope of the management of artefacts produced in distributed software development. In their work, the goal is not MCE, and workflow and organization structure are not mentioned.

Sun and Pan (2005) propose the FRWM model, where the relation between roles and permissions is made via workflow tasks. Although the workflow concept is considered, they do not take into account the document and organization structure.

Wang and Long (2007) present a model associating a workflow task to the users and not to the permissions. In their proposal, document structure is not addressed. Their model also considers organization that contains users, but the organization is not related to roles. In our approach,

a group has roles and a user is associated only with groups.

Li et al. (2008) propose the H-TRBAC which extends the concept of hierarchy to tasks, roles and permissions. The task hierarchy is the own workflow. The role hierarchy is already defined in the RBAC model (Sandhu et al., 2000). The permission hierarchy is a partial order of permissions; therefore if a role grants permission, the role also grants all the permissions downwards in the hierarchy. Our model does not allow this type of hierarchy, but it provides the object hierarchy to represent the document structure. Li et al. (2008) do not consider organization structure.

Zhu and Lv (2008) propose the ACEC model, where the document, being used in a cooperative editing, is defined by a set of sections. Their representation is a type of document structure. Their model also includes the concept of workflow activity, but the organization structure is not mentioned.

5 CONCLUSIONS

We presented an access control for massive collaborative edition that considers workflow, document structure and organization structure. The proposed model is an extension of the RBAC model. The model can be used to ease the specification work of collaborative editing systems and we advocate that the model can used to classify systems. We expect that the model can aid the collaborative editing systems with hundreds to millions of users with significant reductions in the numbers of assignments.

The research is not complete. Now, we are currently investigating more ways of assignments using particular structures. For instance, we would like to have flexible assignments in order to cope with the dynamic changes in organizations. For instance, the delegation of role if a certain condition is met presents itself as interesting research problem. Another problem is how to specify or extend the model for conflict resolution in matrix organization. As mentioned, entities can respond to two different entities, therefore some form of specifying priorities may be required.

REFERENCES

Daft, R. L. (2008). *Organization Theory and Design*. Cengage Learning. 10th ed. pp 88-132.

Cooperation Environment. ECOO project. http://www.loria.fr/equipes/ecoo/english/index.html

Fischer, L. The Workflow Handbook 2003 Published in association with the Workflow Management Coalition (WfMC) 3rd ed.

Google. Google Docs & Spreadsheets. (2008). Create and share your work online. http://docs.google.com.

Ignat, C.L. and Norrie, M.C. (2008). Multi-level editing of hierarchical documents. *Journal of Computer Supported Cooperative Work*, 17(5-6):423-468.

Sandhu, R., Ferraiolo, D.F. and Kuhn, D.R. (2000). The NIST Model for Role Based Access Control: Toward a Unified Standard. In *5th ACM Workshop Role-Based Access Control*. pp 47-63.

Wang W. and Long Y. (2007). Research on extension to role based access control mechanism on workflow platform. In *Third International Conference on Natural Computation* (ICNC).

Sun, Y. And Pan, P. (2005). PRES-A Practical Flexible RBAC Workflow System. In *7th International Conference on Electronic Commerce* (ICEC).

Li, J, LI, X., Xie, S. et al. (2008). Multi-Hierarchy and Fine-Grained Task-role-based Access Control in Collaborative Environments. In *Industrial Engineering and Engineering Management* (IEEM).

Bruegge, B., De Lucia, A., Fasano, F and Tortora, G. (2006). Supporting Distributed Software Development with fine-grained Artefact Management. In *International Conference on Global Software Engineering* (ICGSE).

Zhu, F. and Lv, Qiang. (2008). ACEAC: A Novel Access Control Model for Cooperative Editing with Workflow. In *International Symposium on Electronic Commerce and Security*.