# Integrity in Open Collaborative Authoring Systems

Christian Damsgaard Jensen

Informatics and Mathematical Modelling
Technical University of Denmark
Christian.Jensen@imm.dtu.dk

**Abstract.** Open collaborative authoring systems have become increasingly popular within the past decade. The benefits of such systems is best demonstrated by the Wiki and some of the tremendously popular applications build on Wiki technology, in particular the Wikipedia, which is a free encyclopaedia collaboratively edited by Internet users with a minimum of administration.
One of the most serious problems that have emerged in open collaborative authoring systems relates to the quality, especially completeness and correctness of information. Inaccuracies in the Wikipedia have been rumoured to cause students to fail courses, innocent people have been associated with the killing of John F. Kennedy, etc. Improving the correctness, completeness and integrity of information in collaboratively authored documents is therefore of vital importance to the continued success of such systems. In this paper we propose an integrity mechanism for open collaborative authoring systems based on a combination of classic integrity mechanisms from computer security and reputation systems. While the mechanism provides a reputation based assessment of the trustworthiness of the information in a document, the primary purpose is to prevent untrustworthy authors from compromising the integrity of the document.

## 1 Collaborative Authoring Systems

Collaborative authoring systems which support an open and dynamic population of authors, such as the Wiki [1], have become increasingly popular over the past couple of years. Large pieces of documentation, such as the Wikipedia [2], have been compiled using this type of technology and the Wiki technology has become an indispensable part of many computer supported collaborative work (CSCW) tools that support a distributed user base. The Wikipedia project has demonstrated the benefits of this approach by compiling a comprehensive and largely accurate encyclopaedia from the contributions of individual people located around the world. However, the Wikipedia has also exposed one of the weaknesses of collaborative authoring, which is that malicious or incompetent users may compromise the integrity of the document by introducing erroneous entries or corrupting existing entries, e.g., a public figure has found that the entry describing them in the Wikipedia had been modified to defame him [3].

The quality of a collaboratively authored document is determined by a few simple properties, such as whether the document is complete, correct and unbiased. Some of these properties correspond to the properties ensured by existing integrity mechanisms

in computer security, so we intend to leverage this work when designing an integrity mechanism for open collaborative authoring systems. Classic integrity mechanisms [4, 5] associate an integrity level with every author (subject) and document (object), so that authors are assigned the integrity level of the documents that they work on and authors with low integrity are prevented from updating documents with higher integrity levels. Data protected by an integrity mechanism, however, normally have well defined syntax and semantics, whereas the syntax and semantics of collaboratively authored documents are difficult to define. This means that existing integrity mechanisms cannot be used directly. The obvious answer to this problem is to rely on feedback from the users, i.e., some reputation system similar to the ones used by Amazon [6], which corresponds to the approach that is already used in a Wiki. Reputation systems have previously been proposed as an effective means to assess the quality of information from uncertain sources [7, 8], but they only help automate detection of undesirable content and are generally unable to prevent undesirable content from being introduced into the document.

## 2 A Reputation-based Integrity Mechanism

We propose a combination of reputation systems to assess the quality of collaboratively authored documents and traditional integrity mechanisms to prevent unknown or untrusted users from modifying the documents in the collaborative authoring system. The mechanism automatically assigns a "quality rating" to every document in the system, based on the reputation of the last user who updated the document. In order to enforce integrity, we want that only users with a similar or higher reputation than the past user will be able to modify the entry. This means that users with a poor reputation will be unable to update most of the documents in the systems, but more importantly that documents that have a high quality rating may only be modified by the few users who have an equally high reputation. The integrity mechanism is based on two fundamental integrity models: the static integrity model and the dynamic integrity model, which capture respectively the static and dynamic properties of integrity control.

### 2.1 Static Integrity Model

All authors must have an identifier (possibly a pseudonym) which will allow the system to recognise authors and associate them with a quality confidence value (QCV), which indicates the normal level of correctness, completeness and lack of bias in documents by that author, i.e., it encodes the reputation of that author.

Each section of the document has an integrity level (IL) associated with it, which corresponds to the QCV of the latest author. This means that it is the integrity level (QCV) of the author that determines the current integrity level (IL) of the document, the integrity label of a document is modified to reflect the integrity label of the author, which is the opposite of the low watermark policy [5]. Moreover, authors are only allowed to modify a document if their QCV is higher than the IL of the document, so authors with a poor reputation cannot modify high integrity documents. We believe

that it is reasonable to assume that authors who have a history of writing complete, correct and unbiased documents are likely to continue in that style, so new documents edited by such authors will benefit from their involvement, i.e., the document will be raised to the higher level of the author.

### 2.2 Dynamic Integrity Model

New accounts are created with a QCV which is lower than any other integrity label in the system. This means that newly created accounts are only able to modify documents by other new authors and create new documents at the lowest integrity level. Authors who create enough quality documents will eventually be promoted, so they will be allowed to contribute to documents with higher integrity levels.

Authors are promoted as a consequence of the promotion of documents that they have authored. This requires a mechanism to assess the quality of their work, which can be done automatically, e.g., at regular intervals using some of the existing information rating techniques [7, 8], or manually using feedback from the users (readers and/or other authors). While the automatic techniques look promising, we believe that they are not yet sufficiently mature for general use, so we propose a simpler and more transparent manual approach, which is initiated by the author who wishes to promote one of his documents.[1] The documents are forwarded to a number of higher integrity authors who use a voting mechanism to decide on the promotion. The group of higher integrity authors who are asked to vote on the promotion, should include a decreasing number of members with an increasing QCV in order to prevent a hoard of vandals from promoting each other. If one of the documents are promoted, then the QCV of the author is updated to the new IL of the document.

Finally, we need a mechanism to deal with any documents that are mistakenly promoted. Any author whose QCV dominates the IL of the document may request that the document is demoted, using essentially the same mechanism as promotion. If a document is demoted, so are the integrity labels of the latest author to modify the document and the author who initiated the promotion.

## 3 Discussion

In this paper, we proposed a mechanism which combines existing assessment techniques with integrity control mechanisms from computer security, in order to provide quality information to the reader and prevent untrustworthy users from corrupting high quality documents.

Documents are internally labelled with an integrity label, which provides the reader with an idea about the provenance of the document and whether the content should be trusted. The system also associates integrity labels with authors, which allows the system to prevent authors who have primarily authored low quality documents from

[1] The complete details of the protocols and mechanisms needed to promote and demote authors and documents are too lengthy to be included here.

modifying documents with a high integrity (quality) label. The integrity mechanism is designed to ensure that the editing process does not lower the integrity of documents.

The proposed integrity mechanism for open collaborative authoring systems has the following integrity properties:

1. Unknown authors can only modify the documents of other unknown authors
2. Normal authoring procedures will never decrease the integrity label of documents
3. Collaborative filtering techniques are used to promote documents that are complete, correct and unbiased to a higher integrity level

We believe that these properties will help raise the quality of documents produced and stored in open collaborative authoring systems. The system is currently being developed at the Technical University of Denmark and we expect that the implementation of the mechanism in an existing Wiki system will allow us to experiment more with such policies in a real world environment.

## References

1. What is Wiki. `http://www.wiki.org/wiki.cgi?WhatIsWiki`, visited 28 December 2006
2. Wikipedia, the free encyclopedia. `http://en.wikipedia.org/wiki/Wiki`, visited 28 December 2006
3. John Seigenthaler (2005) A false Wikipedia 'biography'. Editorial in USA TODAY, 29 November 2005
4. K. J. Biba (1977) Integrity Considerations for Secure Computer Systems. Technical Report MTR-3153, The MITRE Corporation, Bedford, Massachusetts, U.S.A.
5. Timothy Fraser (2000) LOMAC: LowWater-Mark Integrity Protection for COTS Environments. In Proceedings of the IEEE Symposium on Security and Privacy, Oakland, California, U.S.A.
6. Amazon website. `http://www.amazon.com`, visited 28 December 2006
7. Pierpaolo Dondio, Stephen Barrett, Stefan Weber and Jean-Marc Seigneur (2006) Extracting Trust from Domain Analysis: a Case Study on Wikipedia Project. In Proceedings of the 3rd International Conference on Autonomic and Trusted Computing, IEEE, 2006
8. Ilya Zaihrayeu, Paulo Pinheiro da Silva and Deborah L. McGuinness (2005) IWTrust: Improving User Trust in Answers from the Web. In Proceedings of 3rd International Conference on Trust Management, Rocquencourt, France, 2005