A Multilateral Secure Reputation Framework for Wikis

Benjamin Kellermann, Stefanie Pötzsch and Sandra Steinbrecher

Technische Universität Dresden
Faculty of Computer Science
D-01062 Dresden, Germany
{Benjamin.Kellermann,Stefanie.Poetzsch,
Sandra.Steinbrecher}@tu-dresden.de

Abstract. Wikis offer users the possibility to share content with each other and thereby create a common information basis. Other users can use this content with or without contributing actively to the content. Content in wikis usually is changed frequently by several authors. Thereby it becomes more and more difficult for users consuming information in wikis to decide which version of content and which author can be trusted. This paper deals with these problems arising in wikis and presents a multilateral secure reputation framework for wikis that we tested with MediaWiki.

Keywords: reputation, wiki, TODO

1 Introduction

Wikis collect and publish information users generate. The probably most popular example is Wikipedia[1]. The English version contains more than 2.6 million articles at the beginning of 2009 and covers nearly the same – and sometimes even more – encyclopaedias in book-form cover. While printing an encyclopaedia is expensive and the review process of articles is usually long, generating content in a wiki is cheap and easy. It needs neither technical nor other specialised know-how from the authors. This leads to the drawback that it becomes very difficult for readers to decide which content or which author can be trusted.

Reputation systems can be used to help users in deciding how much to trust in the quality of content and the credibility of authors. The credibility of authors can be an indicator about the quality new content they generate.

A *content reputation system* allows readers of content to judge on it's quality and give a *rating* to it. The reputation systems collects all ratings given, aggregates them to a reputation of the content and shows it together with the content to possible future readers. The problem with wikis here is that information changes frequently.

For the credibility of authors, an *author reputation system* assigns *author reputation* to authors. This can be done by aggregating ratings or certifications

[1] http://www.wikipedia.org/

given to authors based on their proven expertise, maybe inside or ourside the wiki. If the author reputation is influenced by content he edited inside the wiki the problem is that content is often generated together with other authors. Here the question arises how the reputation of the commonly generated content influences each author's reputation.

An author reputation system can also influence the content reputation by using the author reputation for aggregating the reputation of all content, not only new content.

This paper deals with the problems of determining author and content reputation in wikis and presents a multilateral secure reputation framework trying to help to solve these problems. In Section 2 we describe the possible design issues and the requirements arising in author and content reputation systems for wikis. Based on this discussion we present a content reputation system for the platform MediaWiki in Section 3 and design an author reputation system in Section 4.

2 Scenario

2.1 Content Reputation System

Reputation assigned to content in a wiki can help the content's users, i. e., the readers, to estimate its quality (e. g., truth or usefulness). Therefore, users who are already able to estimate the content can become raters and make use of a *rating algorithm* to give a rating to the content.

The reputation of the content is then calculated from these ratings with the help of a *content reputation algorithm*. There exist countless models to design rating and reputation algorithms [1].

Ben:Den Satz versteh ich gar nicht. Müsste es nicht heißen: Reputation assigned to authors, can help a content reputation algorithm to calculate more precise values.?

The problem with content in wikis is, that it might change after the ratings have been given to it. Here, a *reputation updating algorithm* is needed that updates the reputation depending on the changes made to the content. To best of our knowledge, no such algorithm exists so far. This is the first issue this paper deals with.

2.2 Author Reputation System

Reputation assigned to authors in a wiki can help other users of the wiki to estimate which quality content they edit might have. As for content, the reputation of the authors is calculated with the help of an *author reputation algorithm*. In difference to the reputation algorithm for content, three aspects may be taken into account when calculating the reputation of an author,

Ben: ratings from other users **and** autogenerated ratings (number of edits, produced lines of code...?)

- ratings to the authors,
- certificates, authors may get from internal or external authorities with some *certification algorithm*, as well as
- ratings to the content of the authors.

A problem with the usage of content ratings in the reputation algorithm is that content usually has been created by several authors and it is not easy to determine, which author's edit has the most worthfull influence.

Adler and Alfaro [2] present an author reputation system that aggregates author reputation by observing whether subsequent authors preserve the changes they made.

To the best of our knowledge no reputation algorithm exists so far that considers all three aspects. This is the second issue this paper deals with.

2.3 Rater Reputation System

Raters usually give subjective ratings that are influenced by their personal estimation of the truth or usefulness of the content.

Thus, for the evaluation of a content's or an author's reputation, users do not only have to trust in the rater's honesty but also need some means to map the rater's subjective rating to their own view. We do not further elaborate the numerous existing trust models to implement trust. We demand that the raters and the context of some rated information are weighted according to the trustworthiness they have for the evaluator of a reputation. An example for such a technical trust model that makes use of interpersonal context-specific trust was developed by Abdul-Rahman and Hailes [3]. Unfortunately, this model is by far too complex for practical applications with a large number of users.

To help users to estimate the trustworthiness of raters, a *rater reputation system* is needed. The rater reputation should have an influence to the reputation algorithm for authors and content. The rater reputation system can be similar to the author reputation system and depend on the same aspects (direct ratings, certificates, ratings to raters content). Furthermore, gaining a good reputation as an author should influence the reputation somebody has for rating content or other persons.

All reputation stored can only be evaluated by a user of the reputation system if there is an information flow in the reputation network towards him. The reputation selection for evaluation can be:

- *global:* This means the information flow within the reputation network is complete and every evaluator gets the same reputation of a reputation object.
- *individual:* This means an evaluator only gets a partial view on the reputation available.

Ben: ab hier passts nicht mehr so richtig zum vorhergehenden. Mir ist auch nicht so richtig klar, was du mit dem Rest des Abschnitts ausdrücken willst.

2.4 Requirements

To reach the functional goal of estimating a given content's quality one has to consider several security requirements for the reputation systems assisting this goal. The security requirements we list in the following overlap with the generic security requirements of a reputation system stated by Carrara et al. [4] and follow the requirements derived from Steinbrecher [5]:

update bibtex of Stei_08

Ben: irgendwie sind das so viele Requirements, dass es mir als Leser schwer fälllt zu übersehen ob alles drin wäre. Kann man

Availability of reputation: As a functional requirement, each user of the reputation system wants to access reputations to estimate the quality of content at the time he reads the content. This needs the content reputation to be stored with the content.

Integrity of content and ratings: Users want content and ratings used to calculate the content's reputation to be preserved from manipulations, both in propagation and in storage.

Accountability of authors, raters and certifiers: Users want authors to be accountable for the content they edited and raters to be accountable for the ratings they gave to authors/raters. The same holds for accountability of certifiers.

Completeness of reputation: Users want the aggregated reputation to consider all ratings given. It should not be possible to omit certain ratings.

Pseudonymity of raters and authors: Users want to rate and provide content under a pseudonym to not necessarily allow others to link this rating to their real name. In the real world there are also authors who write under a pseudonym and many services in the Internet also allow the use of pseudonyms instead of real names following EC Directive 95/46 [6]. Nevertheless all activities done under this pseudonym should be linkable.

Convertibility of reputation/ratings: Authors want to transfer reputation/ratings from their real identity to and between their various pseudonyms.

Unlinkability of ratings and content: Users want to rate and provide different content without being linkable. Otherwise behaviour profiles of pseudonyms (e. g., time and frequency of web site visits, valuation of and interest in specific items) could be built. If the pseudonym can be linked to a real name the profile can be related to this real name as well.

Anonymity of users: Users want to evaluate reputation anonymously to prevent others from building personal behaviour profiles of their possible interests.

Confidentiality of ratings: Although a reputation system's functional requirement is to collect and provide information about a reputation object, raters might prefer to provide only a subset of their ratings to a specific group of other users while keeping it confidential to all others.

Persistence of reputation objects: Persistence [7] of members as reputation objects has to be realised resp. the binding of reputation to them. This can be done pseudonymously. The same holds for content with a good reputation that it cannot be easily destroyed by other authors.

Absolute linkability of an author's registration with a reputation system: To prevent a user from leaving a reputation system with a bad reputation and re-entering it with a neutral reputation membership actions of the same user in the same context have to be absolutely linkable.

Certainly some of these requirements are contradicting. The goal of multilateral security [8] is to find a compromise between different design options realising a compromise of security requirements. Our special goal with the the reputation framework is to let this decision to the concrete system designer which security requirements he favours.

2.5 Infrastructure

We assume a user-controlled privacy-enhancing identity management like PRIME to be used [9]. We further assume all communication to be secured by encryption to reach confidentiality of all ratings and actions performed. All actions and ratings have to be secured by digital signatures given under a pseudonym for integrity reasons. By the use of an identity provider accountability of the pseudonym can be given.

Ben: ab hier geht es auch schon nicht mehr ganz um Infrastructure zu gehen, sondern schon um Registrierung. Vielleicht splitten wir das in eine "Prelimenaries" section?

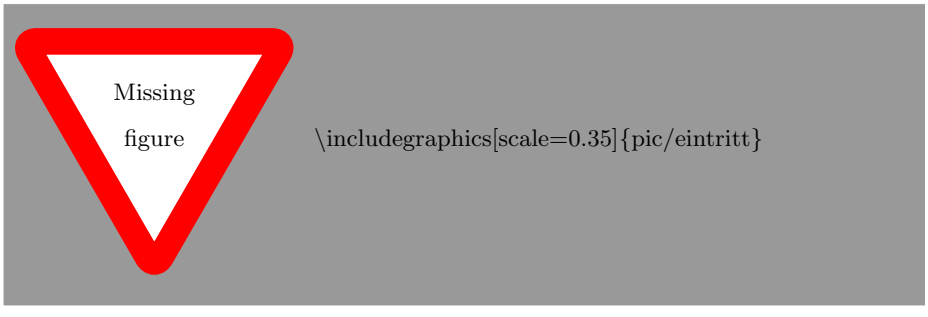

Fig. 1. Registration process enabling unlinkability of a user and his pseudonyms

For the identity management a user registers a basic pseudonym with an identity provider by declaration of her identity data (step 1 in Figure 1). After verifying the data the identity provider issues a basic credential (step 2 in Figure 1).

When the user gains some expertise in a field that can be certified by an independent certifier (e. g., like an university) it may show this basic credential (step 3a in Figure 1) to the certifier.

When the user wants to register in a wiki she sends the reputation provider her basic credential (step 3 in Figure 1). This guarantees no user is able to build up reputation under multiple pseudonyms and every user can be identified in the case of misbehaviour. The reputation provider creates a reputation pseudonym based on the basic pseudonym and sends it back to Alice (step 4 in Figure 1).

The reputation credential contains the pseudonym and its initial reputation. The credential is a pseudonymous convertible credential [10] the user can convert to another pseudonym within the wiki whenever he wants to reach unlinkability of actions. The credentials contain an attribute for the context, $l > 0$ attributes for the last l ratings and an attribute for the expiration date.

After the conversion of the reputation credential to a community pseudonym Alice can register this pseudonym with a community $\mathcal{C}_h$ by showing the converted credential (step 5 in Figure 1). Thereby she agrees that she will collect reputation for her interactions in the community with the reputation network she

registered with. Based on this she gets a community credential to her community pseudonym and becomes a member of the community (step 6 in Figure 1).

By the use of these distinct pseudonyms, unlinkability of the actions performed under these pseudonyms is given initially.

3 Content Reputation System

Ben: Section 3.1 und Section 3.2 hab ich mir noch nicht vorgenommen. Eigentlich müsste das ja Standardvorgehen sein (nicht Wikispezifisch), gibts da schon entsprechenden Text oder Verweis?

In this section we want to investigate, how a content reputation system for wikis may be build. Therefore, we want to start considering static content. We describe, how static content can be rated (Section 3.1) and how an aggregated reputation value can be calculated from this single ratings (Section 3.2). Section 3.3 on the next page handles the special case for wikis, where we consider that content might change over time.

3.1 Rating Algorithm

3.2 Reputation Algorithm

An example, how ratings for a specific wikipage may be distributed is shown in Table 1. In this table, a specific rater is linkable over revisions. Table 2 shows

Table 1. Example of given Ratings to a specific article. The Rater is linkable over the Revisions and only the current Reputation of an author is shown.

	Inga (200)	pseudonym (180)	Josha (70)	unknown (730)
rev 42		★★★☆☆		★★★★☆
rev 23	★★☆☆☆		★☆☆☆☆	★★☆☆☆
rev 5	★★★☆☆	★☆☆☆☆		

Table 2. Example of given Ratings to a specific article. The Rater is unlinkable over the Revisions.

Revision	Ratings
42	★★★★☆(730); ★★★☆☆(180)
23	★★☆☆☆(748); ★★☆☆☆(200); ★☆☆☆☆(70)
5	★☆☆☆☆(196); ★★★☆☆(191)

a view to the same ratings, but this time, the rater is not linkable between

multiple ratings. Possible Implementations not only include a linkability within one Wikipage but also a linkability within the whole Wiki. Precondition: Rater reputation is connected to Users master secret, which is issued only once from the government and which he does not want to give away. The next sections are ordered from less potential abuse and less anonymity to most potential abuse and most anonymity.

Linkable within the Whole Wiki Rater fetches a Wiki login by showing his governments Identity to a trusted third party. The trusted third party checks, that every User gets only one login credential. Login credential is re-encrypted by the user to be unlinkable to the real-world identity. This re-encryption is possible only once! Rater shows this login-credential every time, he wants to rate. The server can verify, if a revision is already rated by an author and update old ratings if necessary.

Linkable within One Page Same as in the case "Linkable within the whole Wiki" but the third party issues one credential per user and page. The third party learns from issuing the credentials the pages a user is interested in (except a user fetches credentials for all pages by default, maybe some dummy-credentials also fit).

Unlinkable Rater gets x one-show credentials in a specific time-period. With these credentials, he may rate x Page-revisions in the specific period, proving his Rater-reputation anonymously.

Unlinkable as the User Wants it Combined version of "Linkable within the whole Wiki" and "Unlinkable". A Rater may fetch more than one login credential from the trusted third party (but still only few to prevent abuse) and one may get some one-show credentials for anonymous ratings additionally.

3.3 Reputation Updating Algorithm

As already discussed, content may change in a wiki. Some author may edit a page and store a new revision with his changes. These changes might contain new paragraphs and lots of new content or just some small spelling corrections. Here it becomes difficult to decide whether a rating of an old revision still is valid for the current one. According to our terminology, we call the value of an edit the *revisiondiff reputation*. The knowledge about the revisiondiff reputation is not only important for the calculation of a new content reputation value. Additionally it is needed to calculate some rating for the author. We propose different solutions depending on the amount of raters and their expertise.

Ben: gibts hier vielleicht nen besseren Begriff?

Evtl. noch: Je weniger feingranular die Bewertung (5 Stufen?) desto mehr Bewertungen müssen abgegeben werden um eine Differenz festzustellen.

Democratic System Within the *democratic system*, we consider significant more raters than authors in the wiki system. This means, that each page gets many more ratings, that edits. As there are enough ratings, only ratings given to the last revision are evaluated for the content's reputation value. When there are enough ratings for one revision, the revisiondiff reputation can be calculated from the difference of the old revisions reputation and the new reputation.

Expert System In difference to the democratic system, there are only few raters in an expert system (the experts). It is assumed, that only the few experts are in a position to value the content correctly. In contrast to the quality of the content, it needs less experts skills to decide if a revisiondiff has high impact on the quality of the content or not. Therefore, there still should be many people, who can decide how much the quality changed with a revisiondiff. This means, that in an expert system, the revisiondiff should be rateable directly.

4 Design of an Author Reputation System

Reputation assigned to authors in a wiki can help other users of the wiki to estimate which quality content they edit might have. As for content users can make use of a rating algorithm to give a rating to the content. Also internal or external authorities who are able to estimate an author's expertise can make use of **a certification algorithm** to assign a certain certificate to an author.

The reputation of the author then can be calculated from the ratings and the certificates given to the author as well as from the ratings given to the content he edited with the help of a **reputation algorithm**.

Adler and Alfaro [2] present an author reputation system that aggregates author reputation by observing whether subsequent authors preserve the changes they made.

But to the best of our knowledge no reputation algorithm exists so far that considers all three aspects (ratings and certificates to authors as well as ratings to content the author edited). This is the second issue this paper deals with. The problem with the usage of content ratings in the reputation algorithm is that content usually has been created by several authors.

4.1 Updating Authors Reputation

day 1 Mariangela (50) rates article Foo with ★★★★☆
day 2 Foo has a reputation value of ★★☆☆☆
day 3 Mariangelas reputation changes to 100

What happens with the article reputation?

Possibility 1 Mariangelas reputation changed because it did not matched her skill before. → the article reputation should be updated.

Possibility 2 Mariangelas reputation changed because it her skill changed. → the article reputation should not be updated.

How to decide, if skill changed or not? → Maybe, it is sufficient to distinguish between reputation changes resulting from new {Author,Rater}-ratings and reputation changes resulting from {Author,Rater}-rating updates?

References

[1] Jøsang, A., Ismail, R., Boyd, C.: A survey of trust and reputation systems for online service provision. Decision Support Systems **43**(2) (2007) 618–644

[2] Adler, B.T., de Alfaro, L.: A content-driven reputation system for the wikipedia. In: WWW '07: Proceedings of the 16th international conference on World Wide Web, New York, NY, USA, ACM (2007) 261–270

[3] Abdul-Rahman, A., Hailes, S.: Supporting trust in virtual communities. In: HICSS '00: Proceedings of the 33rd Hawaii Intern. Conference on System Sciences. Volume 6., Washington, DC, USA, IEEE Computer Society (2000) 6007

[4] ENISA: Position paper. reputation-based systems: a security analysis. available from `http://www.enisa.europa.eu/act/it/oar/reputation-systems/reputation-based-systems-a-security-analysis/at_download/fullReport` (last visited 07/01/09) (2007)

[5] Steinbrecher, S.: Enhancing multilateral security in and by reputation systems. In: to be published in Fourth FIDIS International Summer School 2008, in cooperation with IFIP WG 9.2, 9.6/11.7, 11.6; Springer 2009. (2008)

[6] : Directive 95/46 EC. Official Journal L281,23/11/1995 pp. 31-50 (1995)

[7] Resnick, P., Kuwabara, K., Zeckhauser, R., Friedman, E.: Reputation systems. Communications of the ACM **43**(12) (2000) 45–48

[8] Rannenberg, K., Pfitzmann, A., Müller, G.: IT security and multilateral security. In Müller, G., Rannenberg, K., eds.: Multilateral Security in Communications. Volume 3 (Technology, Infrastructure, Economy)., München, Addison-Wesley (1999) 21–29

[9] Casassa-Mont, M., Crosta, S., Kriegelstein, T., Sommer, D.: Architecture v2. Deliverable D14.2.c, PRIME (March 2007) `https://www.prime-project.eu/prime_products/reports/arch/pub_del_D14.2.c_ec_WP14.2_v1_Final.pdf`.

[10] Chaum, D.: Showing credentials without identification. Signatures transferred between unconditionally unlinkable pseudonyms. In: Advances in Cryptology - EUROCRYPT '85, Workshop on the Theory and Application of of Cryptographic Techniques, New York, NY, USA, Springer-Verlag (1986) 241–244