



Wiki sémantique sur un réseau pair-à-pair

Hala Skaf-Molli, Charbel Rahhal, Pascal Molli

► To cite this version:

Hala Skaf-Molli, Charbel Rahhal, Pascal Molli. Wiki sémantique sur un réseau pair-à-pair. Atelier IC2.0 en association avec le 19èmes journées Francophone d'Ingénierie des Connaissances - IC 2008, INRIA, Jun 2008, Nancy, France. 4p. inria-00432219

HAL Id: inria-00432219

<https://hal.inria.fr/inria-00432219>

Submitted on 14 Nov 2009

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Wiki sémantique sur un réseau pair-à-pair

Hala Skaf-Molli, Charbel Rahhal and Pascal Molli

INRIA Nancy-Grand Est
University Henri Poincaré
BP 239 - 54506 Vandoeuvre-lès-Nancy Cedex
skaf@loria.fr

Résumé :

Un wiki sémantique sur réseau pair-à-pair est un outil d'édition collaborative combinant les avantages des systèmes pair-à-pair et du Web sémantique. L'intégration des technologies du Web sémantique dans un Wiki améliore la structuration des pages wiki, la navigation entre ses pages et la recherche d'information. L'architecture P2P permet un passage à l'échelle, une amélioration des performances, une résistance aux pannes et le travail déconnecté.

Mots-clés : Wiki sémantique, réseau pair-à-pair, CSCW.

1 Introduction

Les Wikis sémantiques sont une nouvelle génération d'outils d'édition collaborative. Ils permettent d'éditer des pages wikis contenant des données sémantiques Völkel *et al.* (2006); Buffa *et al.* (2008). Les utilisateurs peuvent alors collaborer non seulement pour écrire des pages wikis mais aussi pour y mettre des connaissances sémantiques Schaffert (2006). L'aspect sémantique permet de rendre le contenu d'un Wiki "compréhensible" par une machine. Généralement, ceci est réalisé par l'annotation des liens entre les pages. Ainsi, les liens deviennent typés. Par exemple, un lien entre la page *Paris* et la page *France* peut-être annoté par *est capitale de*. Les Wikis sémantiques essaient de préserver les avantages de wikis classiques, c'est-à-dire la simplicité de création et de modification des pages. "Semantic Wikipedia" est un exemple d'édition collaborative autour des wikis sémantiques Völkel *et al.* (2006).

Les wikis sémantiques actuels sont des systèmes centralisés. Ils passent donc mal à l'échelle, ne tolèrent pas les pannes et ne permettent pas le travail déconnecté. Ces problèmes sont résolus par les wikis sur réseau pair-à-pair Weiss *et al.* (2007); Morris (2007).

Un wiki sur réseau P2P est composé d'un réseau de serveurs Wiki. Chaque serveur réplique les mêmes pages wikis. La charge des accès est alors répartie sur l'ensemble des noeuds ; les performances générales du système sont alors grandement améliorée. Une panne d'un serveur n'entraîne pas l'arrêt total du système ; les accès sont alors pris en charge par les autres serveurs. Si un utilisateur héberge un serveur wiki sur sa propre machine, il peut alors travailler en mode déconnecté. Le problème essentiel des wikis

sur réseaux pair-à-pair est d'assurer la cohérence des différentes copies situées sur les différents serveurs et de gérer les conflits d'édition concurrente.

Malheureusement, les wikis sur réseaux P2P ne sont pas capables de gérer des pages contenant à la fois du texte et des données sémantiques. En effet, il faut garantir que la réplification de ce nouveau type de données (texte et données sémantique) assure les propriétés de convergence et de respect des intentions. Si ce n'est pas le cas, alors il n'est pas possible de garantir que tous les serveurs de wikis sémantiques convergeront vers le même état sans perdre de modifications.

Notre objectif est d'adapter les algorithmes de réplification existant afin de prendre en compte ce nouveau type de données où des données sémantiques sont intégrées dans le texte. Il faut alors s'assurer que les propriétés de convergence et de respect des intentions sont respectées.

2 SWOOKI : Wiki sémantique sur réseau P2P et réplification optimiste

SWOOKI est un wiki sémantique sur un réseau pair-à-pair. C'est un système distribué composé d'un réseau de serveurs wikis sémantiques autonomes. Dans ce réseau, chaque noeud héberge une copie des pages wikis et une base de connaissances (figure 1). Les serveurs jouent le même rôle et fournissent les mêmes services.

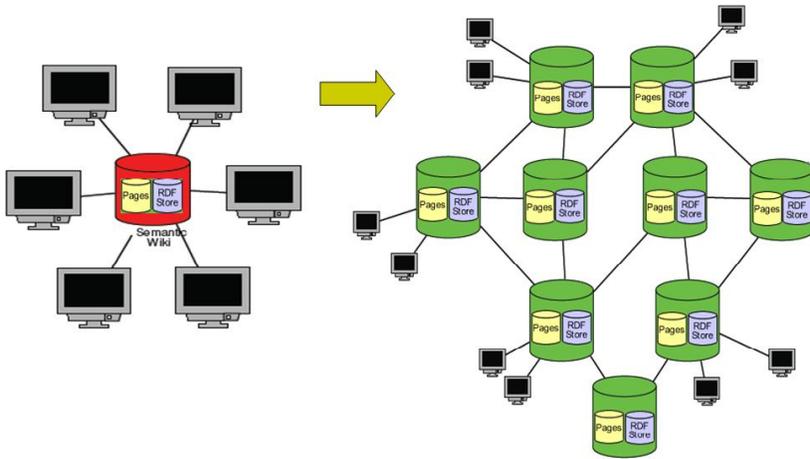


FIG. 1 – Wiki sémantique sur un réseau pair-à-pair

Dans ce réseau, une opération générée sur un serveur est : (1) exécutée localement immédiatement, (2) propagée vers les autres sites, (3) reçue par les autres sites, (4) intégrée et ré-exécutée par chaque site. Une opération est donc exécutée localement sur un état s d'un site et finalement ré-exécutée sur un autre site sur un état s' . s' n'est pas forcément égal à s . En effet, si des opérations concurrentes ont été produites, alors

l'état sur lequel l'opération est ré-exécutée n'est pas le même état sur lequel l'opération a été générée. C'est pour cela qu'un algorithme d'intégration est nécessaire. Une opération reçue est transformée pour tenir compte des opérations concurrentes avant d'être re-exécutée.

SWOOKI ne gère que deux opérations : *insertion d'une ligne dans une page wiki* et *suppression d'une ligne dans une page wiki*. Une mise à jour est traitée comme une suppression suivie d'une insertion. Une ligne peut contenir 0 ou plusieurs données sémantiques. Par exemple, un utilisateur peut ajouter une nouvelle ligne dans la page wiki concernant la France. Cette insertion va être détectée à travers l'opération :

`ins("PAGE_FRANCE",1<"La france a pour capitale [[capitale : :Paris]]")<7)`. Cette opération est exécutée localement quand l'utilisateur sauve sa page wiki. Les données sémantiques sont alors extraites et insérées dans la base de connaissance locale. L'opération est ensuite propagée aux autres sites pour y être intégrée et ré-exécutée.

Un système collaboratif répliqué est correct s'il assure la Causalité, Convergence et préservation des Intentions (modèle CCI Sun *et al.* (1998)) :

- la causalité assure que si une opération op2 a été exécutée après une opération op1 sur un site, alors cet ordre est respecté sur tous les sites.
- la convergence garantit que toutes les copies sont identiques lorsque le système est au repos.
- L'intention d'une opération est l'effet observé lors de la génération de cette opération.
- la préservation de l'intention : Pour toute opération O, les effets observés lors de l'exécution de O sur tous les sites sont identiques à l'intention de O quelques soient les opérations concurrentes.

WOOT Oster *et al.* (2006) est un algorithme de réplication préservant le modèle CCI. Malheureusement, WOOT ne peut pas assurer la convergence des bases de connaissances. En effet, les seules intentions prises en compte sont celles des opérations mettant à jour une structure linéaire. Or, la base de connaissance est un graphe.

Nous avons redéfini les intentions des opérations "insert" et "delete" pour des lignes contenant des données sémantiques et nous avons modifié l'algorithme WOOT en conséquence. SWOOKI est capable d'assurer le modèle CCI sur un type de données nouveau mêlant texte et données sémantiques.

3 Conclusion

Un wiki sémantique sur réseau pair-à-pair combine les avantages des systèmes pair-à-pair et du Web sémantique. C'est un système distribué composé d'un réseau de serveurs wikis sémantiques autonomes. Dans ce réseau, chaque noeud héberge une copie des pages wikis et une base de connaissances. SWOOKI est capable d'assurer le modèle CCI sur le texte et les bases de connaissances. SWOOKI est disponible en Open Source à l'adresse (wooki.sourceforge.net).

Références

- BUFFA M., GANDON F. L., ERETEO G., SANDER P. & FARON C. (2008). Sweetwiki : A semantic wiki. *Journal of Web Semantic*, **6**(1), 84–97.
- MORRIS J. (2007). DistriWiki : a distributed peer-to-peer wiki network. *Proceedings of the 2007 international symposium on Wikis*, p. 69–74.
- OSTER G., URSO P., MOLLI P. & IMINE A. (2006). Data Consistency for P2P Collaborative Editing. In *Proceedings of the ACM Conference on Computer-Supported Cooperative Work - CSCW 2006*, p. 259–267, Banff, Alberta, Canada : ACM Press.
- SCHAFFERT S. (2006). Ikewiki : A semantic wiki for collaborative knowledge management. In *WETICE*, p. 388–396 : IEEE Computer Society.
- SUN C., JIA X., ZHANG Y., YANG Y. & CHEN D. (1998). Achieving convergence, causality preservation, and intention preservation in real-time cooperative editing systems. *ACM Transactions on Computer-Human Interaction (TOCHI)*, **5**(1), 63–108.
- VÖLKEL M., KRÖTZSCH M., VRANDECIC D., HALLER H. & STUDER R. (2006). Semantic Wikipedia. *Proceedings of the 15th international conference on World Wide Web*, p. 585–594.
- WEISS S., URSO P. & MOLLI P. (2007). Wooki : a p2p wiki-based collaborative writing tool. In *Web Information Systems Engineering*, Nancy, France : Springer.