Distributed Wikis and Social Networks: a Good Fit

Babak Esfandiari
babak@sce.carleton.ca

Alan Davoust
davoust@sce.carleton.ca

Department of Systems and Computer Engineering
Carleton University
Ottawa, Ontario, Canada

Abstract

Social networks can play an important role in the process of decentralizing authority in distributed systems. We will focus on distributed wiki systems, and we show how, in the special case of a peer-to-peer wiki, there is a rational incentive for users to self-organize and form a meaningful social network. We discuss to that effect the basic metrics that can be derived from the topology of the social network to help assess the subjective quality of wiki entries. Demos and experimental results will illustrate and support our discussion. We finally speculate as to how these results may also translate to discussion forums or recommender systems.

1. INTRODUCTION

Distributed wiki systems have been developed for many purposes, such as redundancy, scalability, resilience to a central point of technical failure, or the ability to work offline [5]. In such cases, the main technical challenge and concern has been to immediately or eventually reconcile the content of the entries on the different nodes. Peer-to-peer wikis such as Ward Cunningham's Smallest Federated Wiki [2] or our own P2Pedia system [1] take a different angle, one in which each node (i.e., the user and owner of a given peer) acts independently, and where the system allows, or even encourages, complete divergence between the different versions of a wiki entry. The form of collaboration that results from such systems is a passive one, where each new version of an entry is free to build on an existing one. The challenge then becomes one of choosing the most *relevant* entry to consume or to use as the starting point for a new one. This short abstract introduces and motivates the topics of the presentation.

2. THE WIKI ENTRY RANKING PROBLEM

First, let us note that in the case of a traditional wiki such as Wikipedia, multiple versions of an entry do also exist, and the system simply chooses to present the user with the most recent one. This choice may or may not always be the most relevant one, and "edit wars" can be viewed as a way to game the system by exploiting this basic ranking method. In the absence of a central authority, there is no single entry ranking algorithm to be used or enforced. This is both a problem and an advantage: autonomous nodes now have to resort to finding their own personal metrics and ranking algorithm, but this also makes the system more resilient to *trust attacks* (such as social exploitation or Sybil attack [6]), since it would be highly costly (in terms of time, memory, number of messages or accounts that need to be created) and difficult for attacking nodes to game the system, not knowing how exactly each user is planning on ranking the entries.

WWW'16 Companion, April 11–15, 2016, Montréal, Québec, Canada.
ACM 978-1-4503-4144-8/16/04.
http://dx.doi.org/10.1145/2872518.2890584

3. RANKING METRICS

Going back to the problem of choosing ranking metrics, what can be used to help determine the relevance of a document to a peer? The first observation is that in a peer-to-peer wiki like P2Pedia, to view a wiki entry (which is essentially a file) requires downloading it, and thus in turn sharing it with others. Therefore the most popular entries correspond to the most shared ones, and deciding to keep sharing an entry is a way of endorsing it. This link between a peer and the entries she endorses is a first element toward possible metrics.

In P2Pedia, queries for entries propagate via gossip. It is then essential for a peer to make sure there is a direct or indirect connection to the peer that may contain the most relevant entry she is seeking. However, connecting to too many peers means having to sort through a potentially long and irrelevant list of entries. Peers must therefore be selective in the peers they connect to, and they can use the topology of the resulting social network as another element for ranking entries.

P2Pedia makes use of the above elements to come up with rankings such as *most popular entries*, *entries from closest peers*, *entries from most popular peers*, *entries from most similar peers*, etc. We believe that the rankings are more resilient to trust attacks thanks to their diversity, in particular the fact that some rely on crowd-sourcing while others are subjective and dependent on the given peer's sharing and connection behaviour. To benefit from these rankings, peers need to do due diligence on the peers they connect to and the entries they decide to share. The resulting social network of peers is thus a result of the joint incentives of the users in selfishly optimizing the quality of their own rankings.

We have empirically shown that some of these metrics are indeed effective [3] and that their properties also apply to other types of social sharing environments [4], not just for wikis and not just for decentralized environments..

In the talk we will go into more detail as to how these metrics and incentives work and how we validated our claims. We will also give demos that illustrate these points. Finally, we will examine a more general formalization and framework for reasoning about and simulating these systems.

Speaker Biography

Babak Esfandiari is a Professor in the Department of Systems and Computer Engineering at Carleton University (Ottawa, Canada), which he joined in 1999, and where he runs the Network Management and Artificial Intelligence Laboratory. Before that, he worked at Mitel Corporation in the Strategic Technology group. He obtained a Ph.D. from the University of Montpellier II, France in 1997. He is the co-author of about 100 articles and patents in the fields of agent-based systems, peer-to-peer networks, trust, case-based reasoning, web services, and network management. He is on the editorial boards of Computational Intelligence and the Journal of Trust Management.

4. REFERENCES

[1] A. Craig, A. Davoust, and B. Esfandiari. A distributed wiki system based on peer-to-peer file sharing principles. In O. Boissier, B. Benatallah, M. P. Papazoglou, Z. W. Ras, and M. Hacid, editors, *Proceedings of the 2011 IEEE/WIC/ACM International Conference on Web Intelligence, WI 2011, Campus Scientifique de la Doua, Lyon, France, August 22-27, 2011*, pages 364–371. IEEE Computer Society, 2011.

[2] W. Cunningham. Smallest federated wiki project. https://github.com/WardCunningham/Smallest-Federated-Wiki/.

[3] A. Davoust, A. Craig, B. Esfandiari, and V. Kazmierski. P2pedia: a peer-to-peer wiki for decentralized collaboration. *Concurrency and Computation: Practice and Experience*, 27(11):2778–2795, 2015.

[4] A. Davoust and B. Esfandiari. User participation and honesty in online rating systems: What a social network can do. *to appear in Proceedings of the Fifth AAAI Workshop on Incentives and Trust in E-communities (WIT-EC'16)*, 2016.

[5] A. Davoust, H. Skaf-Molli, P. Molli, B. Esfandiari, and K. Aslan. Distributed wikis: a survey. *Concurrency and Computation: Practice and Experience*, 27(11):2751–2777, 2015.

[6] J. R. Douceur. The sybil attack. In P. Druschel, M. F. Kaashoek, and A. I. T. Rowstron, editors, *Peer-to-Peer Systems, First International Workshop, IPTPS 2002, Cambridge, MA, USA, March 7-8, 2002, Revised Papers*, volume 2429 of *Lecture Notes in Computer Science*, pages 251–260. Springer, 2002.