SmartWiki: A reliable and conflict-refrained Wiki model based on reader differentiation and social context analysis

Haifeng Zhao [a,*], William Kallander [a], Henric Johnson [b], Shyhtsun Felix Wu [a,b]

[a] *Computer Science Department, University of California Davis, One Shields Ave., Davis 95616, United States*
[b] *Computer Science Department, Blekinge Institute of Technology, Valhallavägen, 371 41 Karlskrona, Sweden*

ABSTRACT

Wiki systems, such as Wikipedia, provide a multitude of opportunities for large-scale online knowledge collaboration. Despite Wikipedia's successes with the open editing model, dissenting voices give rise to unreliable content due to conflicts amongst contributors. Frequently modified controversial articles by dissent editors hardly present reliable knowledge. Some overheated controversial articles may be locked by Wikipedia administrators who might leave their own bias in the topic. It could undermine both the neutrality and freedom policies of Wikipedia. As Richard Rorty suggested "Take Care of Freedom and Truth Will Take Care of Itself"[1], we present a new open Wiki model in this paper, called TrustWiki, which bridge readers closer to the reliable information while allowing editors to freely contribute. From our perspective, the conflict issue results from presenting the same knowledge to all readers, without regard for the difference of readers and the revealing of the underlying social context, which both causes the bias of contributors and affects the knowledge perception of readers. TrustWiki differentiates two types of readers, "value adherents" who prefer compatible viewpoints and "truth diggers" who crave for the truth. It provides two different knowledge representation models to cater for both types of readers. Social context, including social background and relationship information, is embedded in both knowledge representations to present readers with personalized and credible knowledge. To our knowledge, this is the first paper on knowledge representation combining both psychological acceptance and truth reveal to meet the needs of different readers. Although this new Wiki model focuses on reducing conflicts and reinforcing the neutrality policy of Wikipedia, it also casts light on the other content reliability problems in Wiki systems, such as vandalism and minority opinion suppression.

© 2013 Elsevier B.V. All rights reserved.

1. Introduction

1.1. Background

Wiki systems are widely-used online collaborative applications which allow multiple contributors from diverse backgrounds and dispersed geographic locations to collaborate in creating and editing manuals, books, and other public knowledge bases. Among the numerous Wiki ecosystems, Wikipedia is probably the most widely known. Its essential idea, that a useful encyclopedia of knowledge can be created by allowing anyone (even anonymous users) to create and edit articles, is predicated on the principles of openness and neutrality. Wikipedia has grown to over 3.9 million articles (in English alone) and with millions of contributors (as of May 2012). In the face of such scale, the openness policy has invited

conflicts, or the inclusion of bias, debate, and abuse inside Wikipedia articles covering controversial topics.

Opening unrestricted editing access to everyone makes this lofty goal of maintaining a Neutral Point Of View (NPOV) overly optimistic due to the inevitable biases and idiosyncrasies of human editors. Instead of collaborating harmoniously, some individuals attempt to dominate "their" articles and nullify all previous edits with which they disagree, often forcing administrators to lock down the editing access of those articles in contention. However, this "lock" method is, by itself, also at odds with the NPOV policy since Wikipedia administrators may subjectively choose their own preferred article edits before locking such articles.

Wikipedia assumes that "the articles are agreed on by consensus".[1] This assumption treats unreliable content as a consensus problem, and thereby posits that, while misleading information can and will be contributed, over time the quality will improve as editors reach consensus and the resulting article moves toward a

---

* Corresponding author.
 E-mail address: hfzhao@ucdavis.edu (H. Zhao).
 URL: http://dsl.cs.ucdavis.edu/zhf/homepage/ (H. Zhao).

---

[1] http://en.wikipedia.org/wiki/Wikipedia The version of 18:45, 25 May 2012.

```
Time: 04:48, 2 March 2007          Time: 11:16, 14 March 2007
  From IP: 64.149.176.160            From IP: 62.160.219.253
On May 15 2006, the US State Department    On May 15 2006, the US State Department
announced that it would restore full       announced that it would restore full
diplomatic relations with Libya, even      diplomatic relations with Libya, once
after Gaddafi declared Libya's weapons     Gaddafi declared he was abandoning
of mass destruction programs. The State    Libya's weapons of mass destruction
Department also stated that Libya would    program. The State Department also said
be removed from the list of nations that   that Libya would be removed from the list
support terrorism.                         of nations supporting terrorism. On
On December 19 2006, the Libyan courts     August 31, 2006, however, Gaddafi openly
announced their final verdict in the HIV   called upon his supporters to "kill
trial in Libya. The case had generated     enemies" who asked for political change.
intense interest globally.                 On December 19 2006, the Libyan courts
                                           announced their final verdict in the HIV
                                           trial in Libya. The case had generated
                                           intense interest globally.
```

Fig. 1. Two historical updates on the page of Muammar al-Gaddafi.

"stable" version. This effect works well in mitigating certain types of unreliable content, such as vandalism, as the majority of editors are honest and responsible. However, it remains vulnerable to disputes existing in hundreds of thousands of pages, especially those related to contentious historical, religious and political subjects. Even if an article appears to be "stable", it may still retain bias, as some earlier contributors may have preferred to throw in the towel rather than engage in endless editing wars. With regard to a controversial article, there could be hundreds of editors with considerably diverse evidence, statements and viewpoints to present. For the vulnerable reader who is unfamiliar with a topic presented in an article, how does s/he evaluate the credibility of information from a plethora of unknown or anonymous editors?

Take a locked page "Muammar Gaddafi" (the former leader of Libya) for example. Fig. 1 shows two historical updates which resemble one another but have obviously different sentiments and evidence:

The lefthand update positively expects the US government to restore diplomatic ties with Libya and implies that the country is not supporting terrorism. However, the righthand update suggests that the US government would only restore diplomatic ties contingent upon the cessation of Libya's weapons of mass destruction programs. Meanwhile, the right update also puts forth evidence that denounces Libya's autocracy as further supporting terrorism. Both of the prose contributors are anonymous. Which update should be more trusted? While Wikipedia authorizes administrators to evaluate opinions and evidence on some controversial topics,[2] there are no guarantees that these "experts" will completely avoid personal bias and present universally fair viewpoints to readers, which is is at odds with the NPOV policy.

The current strategy for maintaining NPOV sacrifices reader control in favor of contributors. This strategy exerts effort to ensure complete accuracy for all readers, which is impossible in cases where subject matter experts disagree. The critical issue here is that not all readers have the same goal in seeking knowledge about a topic. This important factor is overlooked by the current NPOV strategy, where the judgement and decision making of readers is ignored. At a high level, we draw a distinction between two different motivations for readers, where the degree of discernment for the truth is starkly contrasted.

In the more general case, we observe that (lenient-type) readers refer to information in Wikipedia without overly concern about absolute truth in depth, especially when the article is irrelevant to their everyday lives. In this article, we name them as "*value adherents*". From the psychological perspective, readers have a tendency to favor information that confirms their beliefs or hypotheses [2]. Moreover, such judgements and deliberations about trust are subject to social influences, since people are social by nature [2]. Social context factors, including social background and interpersonal relationships, play important roles in judgement and knowledge perception, which is absent in the current knowledge representation of Wiki systems. The importance, trustworthiness, and compatibility of the same pieces of information vary significantly to readers with different social contexts. The diversity of social contexts stretches across many variables, such as social relationships, cultural factors, ethnicity, education, and other human factors. Thus, when facing the dilemma of conflicting information from multiple sources, people tend to believe friends or authorities with whom common values and/or similar social backgrounds are shared. These social context factors are absent in the knowledge representation model of traditional Wiki systems.

In contrast to the value adherents described above, there are readers who crave absolute truth in depth. They are skeptical, critical thinking, and open-minded in their willingness to take a holistic view of controversial opinions before allowing themselves to take a stand. In this article, we name them as "*truth diggers*". They refer to Wikipedia seeking solely accurate information and welcome conflicting opinions on controversial topics that deepen their understanding. Unfortunately, the knowledge representation model in traditional Wiki systems cannot efficiently filter and organize conflicting opinions. Worse, controversial topics that have been administratively locked may even undermine the NPOV policy in terms of their needs. Richard Rorty suggested that "If we take care of freedom, truth will take care of itself" [1]. Censoring unfavored editors undermines the discovery of truth for this type of reader. These readers would rather view a clear and well-organized knowledge representation model that reveals conflicting opinions so that the merits of differing opinions may be weighed before adopting a standpoint for themselves.

1.2. Contribution

In this paper, we introduce an alternate Wiki system, TrustWiki, which presents two knowledge representation formats for simulta-