He Says, She Says: Conflict and Coordination in Wikipedia

Aniket Kittur[+], Bongwon Suh[*], Bryan A. Pendleton[*], Ed H. Chi[*]

[+]University of California, Los Angeles
1285 Franz Hall
Los Angeles, CA, 90095 USA
nkittur@ucla.edu

[*] Palo Alto Research Center Inc.
3333 Coyote Hill Road
Palo Alto, CA 94304 USA
{suh, bp, echi}@parc.com

ABSTRACT

Wikipedia, a wiki-based encyclopedia, has become one of the most successful experiments in collaborative knowledge building on the Internet. As Wikipedia continues to grow, the potential for conflict and the need for coordination increase as well. This article examines the growth of such non-direct work and describes the development of tools to characterize conflict and coordination costs in Wikipedia. The results may inform the design of new collaborative knowledge systems.

Author Keywords

Wikipedia, wiki, collaboration, conflict, user model, Web-based interaction, visualization.

ACM Classification Keywords

H.5.3 [Information Interfaces]: Group and Organization Interfaces – Collaborative computing, Computer-supported cooperative work, Web-based interaction, H.3.5 [Information Storage and Retrieval]: Online Information Systems, K.4.3 [Computers and Society]: Organizational Impacts – Computer-supported collaborative work

INTRODUCTION

Collaborative information environments on the web are currently undergoing a fairly extreme revolution. On digg.com, the ranking of news items is determined by the aggregation of user interactions with the site. On del.icio.us, indexing of bookmarked web items is determined by the popularity of tags assigned across many users. A surprisingly successful collaboration environment is the Wikipedia project, an online encyclopedia in which any reader can also be a contributor. The content of virtually every page can be edited by anyone, with those changes immediately visible to subsequent visitors. This participation model has resulted in a highly popular site with a large amount of content (over 1 million articles in the

English Wikipedia alone). Furthermore, much of the content is of surprisingly high quality [12], although vandalism, inaccuracies, user disputes, and other quality issues do continue to plague the site [29].

Wikipedia has shown tremendous and continuing growth, with exponentially increasing numbers of users, articles, and bytes since 2002 [5, 24]. However, the rise of conflict and the costs of coordination are unavoidable in a distributed collaboration system such as Wikipedia, and manifest in scenarios such as conflicts between users, communication costs between users, and the development of procedures and rules for coordination and resolution. Researchers have seen similar costs in other computer mediated communication (CMC) systems such as MOOs and MUDs [8, 9]. Even though researchers have documented the growth of Wikipedia [3, 24, 31], the impact of coordination costs for adding content and users has largely been ignored.

Conflict in online communities is a complex phenomenon. Though often viewed in a negative context, it can also lead to positive benefits such as resolving disagreements, establishing consensus, clarifying issues, and strengthening common values [11]. Here we try to understand the conflict and coordination costs through the concept of indirect work. Viewed from the goal of trying to create high quality content for a collaborative encyclopedia, we define "indirect work" or "conflict and coordination costs" as *excess work in the system that does not directly lead to new article content.* This allows us to develop quantitative measures of coordination costs, and also has broader implications for systems in which maintenance and consolidation occur, such as group work systems [10, 13].

In this paper, we present an overall characterization of conflict and coordination in the development of Wikipedia. We present three novel contributions:

First we demonstrate that at the global level, conflict and coordination costs in Wikipedia are growing. Specifically, direct work (on articles) is decreasing, while indirect work such as discussion, procedure, user coordination, and maintenance activity (such as reverts and anti-vandalism) is increasing.

Second, we build a characterization model for conflict at the article level. Using human-labeled controversy tags as ground truth, we show that a machine learner has high

accuracy in identifying the amount of conflict in an article. A validation survey confirms the generality of the model even on articles that have never been tagged as controversial. The model also identifies a number of interesting conflict-relevant metrics.

Third, we build a user conflict model to investigate the motivation and sources of conflicts through a visualization tool, Revert Graph. A number of test cases show that the tool has great potential to discover and investigate disputes between users.

INTRODUCTION TO PAGE TYPES IN WIKIPEDIA

The main surface contents for people browsing Wikipedia are the article pages. Work done on article pages is immediately viewable to any visitor of the site. In addition to the article pages, there are a number of other pages in which editors can discuss and resolve conflicts, debate about procedures, and contact other users. These pages include:

Article talk page: Used to discuss and build consensus on changes to the article page.

User page and user talk page: Each user on the site has a personal page and a related talk page for discussion of issues ranging from article conflicts to procedure points.

Other "behind-the-scenes" Wikipedia pages: Used to discuss and articulate procedures such as conflict resolution or other Wikipedia policies, and other purposes such as indexing and linking information.

RELATED WORK

Understanding the development of Wikipedia has implications far beyond the development of an online encyclopedia to designers of other collaborative environments, such as computer mediated communication (CMC) systems, and collaborative writing systems.

For example, characterization of Wikipedia's development provides a quantitative analysis of how a Wiki solution to collaborative writing scales to large content and user bases [19]. As for CMC systems, the construction of user-generated content and strong community interaction in Wikipedia are highly reminiscent of MOOs [8]. Mechanisms for dealing with conflict that began to arise in LambdaMOO [9] have, in Wikipedia, evolved to highly sophisticated, formal dispute resolution processes. Indeed, Wikipedia has responded to rising coordination costs in ways not unlike the growth of MOOs and MUDs and other CMC social systems, that is by developing sophisticated policies, procedures, and user classes that fulfill a similar role as a legal system. A Wikipedia editor we surveyed succinctly characterized the process thus:

"The degree of success that one meets in dealing with conflicts (especially conflicts with experience[d] editors) often depends on the efficiency with which one can quote policy and precedent."

The development of policies and procedures in Wikipedia is related to research in Online Dispute Resolution (ODR), which looks at resolution processes assisted by information technology for both online and offline disputes. Recent work has pointed to interesting characterizations of how ODR is currently handled [6], and how technology can play an important role in processes such as consensus building. Just as in ODR, the growth of conflict and coordination costs in Wikipedia and the effectiveness of ways of combating it are important to the Wikipedia community as a whole to maintain the continued forward progress of the system.

Few researchers have examined conflict and coordination in Wikipedia, and none have taken a comprehensive approach across the global, article, and user levels. However, there are a number of studies which have separately examined these factors.

There have been many attempts to quantify the growth of Wikipedia and its dynamics as a complex network [3, 24, 31]. These studies suggest that structural properties of Wikipedia are consistent with those found in many common types of networks. For example, the growth and link structure of individual language Wikipedias are very similar to each other and also to networks such as the World Wide Web [3, 24].

Another major topic of research is assessing the quality of content produced in Wikipedia. Lih [16] analyzed the change in quality of Wikipedia articles before and after they had been cited in the press. Stvilia et al. [22] used factor analysis to determine a set of metrics related to article quality, including the number of edits, unique editors, and anonymous edits. Perhaps the most widely cited quality assessment is the comparison by experts of selected articles in Wikipedia and the Encyclopedia Britannica [12]. In a comparison of scientific entries, Wikipedia was shown to contain an average of about four errors to Britannica's three, though reviewers cited readability and structure issues in the Wikipedia content.

Perhaps most relevant to the study of conflict, Viegas et al. examined article creation through "history flow visualizations" [23]. In this technique they visualized how article edit histories changed at a sentence-by-sentence level. In addition, they provide statistics on certain types of conflict, in particular vandalism. Using data from the May 2003 Wikipedia they analyzed vandalism as characterized by mass deletions (90% smaller than a previous maximum).

Buriol et al. [3] also report data relevant to conflict in the Wikipedia. They examined the distribution of article reverts, in which an article is restored to a prior version, often to fight vandalism or to promote one side of a conflict. They found that the fraction of reverts has been increasing over time to approximately 6% of all edits in January 2006, with about 70% of those made within an hour. They showed that the Wikipedia "three revert rule" policy [30] (which states that no user should revert the same page more than 3 times in 24

hours) had an immediate effect of decreasing the percentage of double-reverts (often the mark of an "edit war" [27]).

APPROACH AND SCOPE OF ANALYSIS

The size of the Wikipedia dataset has made conducting full-scale analyses across all articles, revisions, and users extremely difficult, leading most studies to use only a subset of the data. However, this can be problematic for a number of reasons. First, the distribution of revisions to articles is highly skewed and follows a power law [3], making random sampling difficult. Second, trends in historical data are difficult to identify without analyzing the full dataset. Finally, calculating revision-based metrics (such as identifying true reverts or the change in content over time) is impossible to do for all articles without using the entire content base.

In the following analyses, we used a complete history dump of the English Wikipedia that was generated on July, 2 2006. The dump included over 58 million revisions, from more than 4.7 million wiki pages, of which 2.4 million are article-related entries in the encyclopedia, totaling approximately 800 gigabytes of data. To process this data, we imported the raw text into the Hadoop [14] distributing computing environment running on a cluster of commodity machines, while importing the structure into a clone of the Wikipedia's own databases for direct analysis. The Hadoop infrastructure allowed us to quickly explore new full-scale content analysis techniques while minimizing code optimization time. The database allowed us to inspect Wikipedia statistics in their native format.

COORDINATION COSTS AT THE GLOBAL LEVEL

To characterize the global growth of conflict and coordination costs we analyzed the distribution of edits throughout the entire history of Wikipedia. Judging from content alone Wikipedia has maintained robust and remarkable exponential growth [24]. However, a deeper analysis of where growth is occurring shows a slightly different picture.
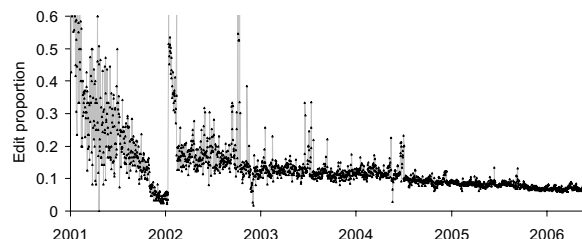
Direct and Indirect Work

The primary evidence of increasing coordination costs in Wikipedia is that the amount of direct work going into articles is decreasing. Edits to article pages continue to be the primary focus of edits on Wikipedia. These edits represent direct work happening on Wikipedia immediately viewable by visiting users. However, despite the overall growth of Wikipedia, the percentage of edits made to article pages has decreased over the years (see Figure 1) from over 90% of all edits in 2001 to roughly 70% in July of 2006.
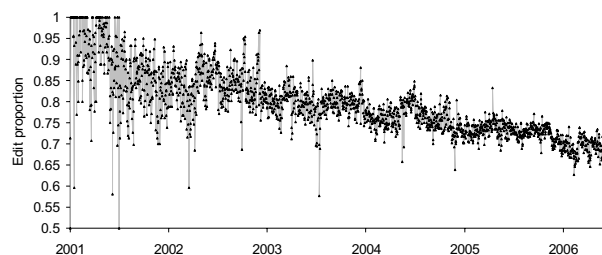

Figure 1. Decrease in proportion of direct work (edits to article pages).

Furthermore, the percentage of edits resulting in the creation of new pages has decreased to less than 10% (see Figure 2), indicating that proportionally less work is going into creating new topics and articles. One explanation is that the maturation of the topic vocabulary in Wikipedia is making it more difficult to find new topics to write about and easier to add or change an existing topic.


Figure 2. Decrease in proportion of article edits that create new articles.

In contrast, the amount of indirect work spent on activities such as conflict resolution, consensus building, or community management has increased over the lifespan of Wikipedia. As shown in Figure 3, over time the percentage of edits going toward policy, procedure, and other Wikipedia-specific pages has gone from roughly 2% to around 12%.
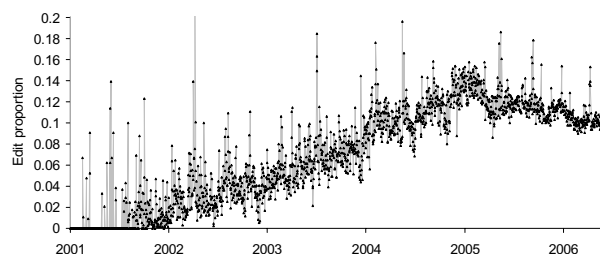
Figure 3. Increase in proportion of edits to procedure and Wikipedia-specific pages (counted as "Other" in Figure 4.)

Thus overall, user, user talk, procedure, and other non-article pages have become a larger percentage of the total edits made in the system. These trends are summarized in Figure 4, which clearly shows the decreasing percentage of edits going to direct work (article edits) and the increasing percentage of edits going to indirect work across different page types.
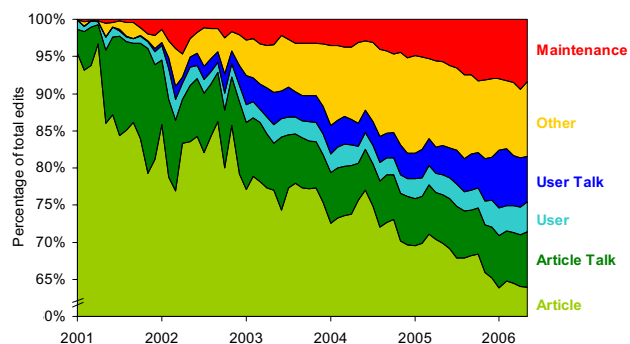
Figure 4. Changing percentage of edits over time showing that decreasing direct work (article) and increasing indirect work (article talk, user, user talk, other, and maintenance).

Maintenance Work

Another type of indirect work is effort spent on maintenance activities. There are two main types of identifiable maintenance work: combating vandalism and making reverts.

One form of maintenance work is reverts. A revert refers to a situation in which a user changes an article back to a previously written version, casting out changes that have been made since. Any work done on the article since the revert (including the revert itself) is lost.

Reverts were measured using two separate methods, a bottom-up data driven method and a top-down user driven method. In the data driven method, we computed a unique identifier of every revision made to every article using the MD5 hashing scheme [20], (commonly used to check that files are identical) and identified when a later revision exactly matched the hash of a previous article, indicating a revert. The advantage of this method is that it does not depend on users to label reverts, which can be inconsistent.

However, the disadvantage of this method is that it does not pick up partial reverts, in which only some of the text in an article is reverted. To capture partial reverts we used a user-dependent metric, counting revisions whose comments included the text "revert" or "rv" (a commonly used abbreviation of revert). The combination of both the data-driven and human-labeled methods described above provide converging evidence on the true change in reverts over time.

Table 1 shows that reverts calculated by the two methods have slightly different characteristics. MD5 (identity) reverts actually capture more revisions than user-labeled (comment) reverts (3.7M vs. 2.4M), suggesting that a substantial number of reverts are not labeled as such. The union of the two methods may provide the most accurate view of reverts, resulting in 3,917,008 reverts marked by comments or MD5 hashes. In other words, approximately 6.7% of the work in Wikipedia goes to restoring articles to previous versions. Figure 5 shows that this number has been

steadily increasing to its current high of about 7% (based on comment-marked reverts).

Type of revert	Edits	Survival (Mean)	Survival (Median)
Comment reverts	2,422,482	0.85 days	7.6 min
MD5 reverts	3,711,638	4.5 days	9.7 min
Vandalism	577,643	2.1 days	11.3 min

**Table 1. Counts and survival times for maintenance work (reverts and vandalism).**
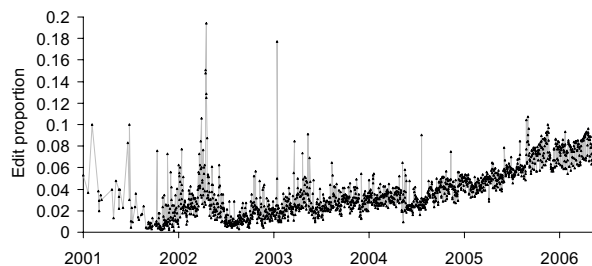
Figure 5. Increase in proportion of edits that are reverts.

Another form of maintenance work in Wikipedia is combating vandalism. Vandalism in Wikipedia refers to a user degrading the quality of an article either by deleting parts of it or intentionally adding inaccurate or inflammatory content, often including swear words. While vandalism has been a particularly visible issue in Wikipedia due to high-profile cases (e.g., [29]), there has been little global characterization of it. The notable exception is [23], which examined ~3500 edits in which mass deletion of content occurred. However, due to the tremendous growth of Wikipedia it has been difficult to get a comprehensive view of vandalism.

We investigated vandalism across all revisions of all articles in Wikipedia. It is difficult to get a perfectly accurate measure of vandalism on Wikipedia since it can take many forms and vandalism to one person may not be considered as such to another. Thus to characterize vandalism we relied on the judgments of users combating it. Specifically, we looked through the edit history for each article for revision comments including any form of the word "vandal" or "rvv" ("revert due to vandalism"), which are put there by users when removing vandalism.

The percentage of all edits marked as vandalism is shown in Figure 6. Vandalism appears to be increasing as a proportion of all edits, though it remains at a fairly low level (1-2% of all edits). We also measured the survival time of vandalism edits. For the 577,643 edits marked as vandalism, the mean survival time was 2.1 days, with a median of 11.3 minutes. This suggests that most vandalism is fixed relatively quickly on Wikipedia though, surprisingly, still slower than the typical revert (see Table 1).
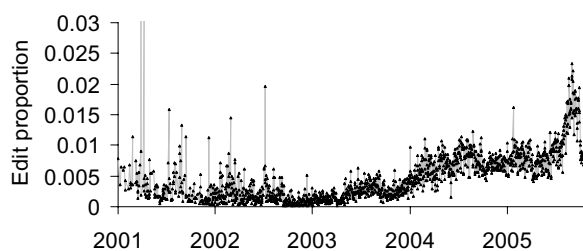
Figure 6. Small increase in proportion of edits marked as vandalism.

Overall, the above global data supports the idea that conflict and coordination costs are rising in Wikipedia, measured as a decrease in direct work and an increase in indirect and maintenance work.

CONFLICT MODEL AT THE ARTICLE LEVEL

While global statistics provide an overview of the growth of conflict in Wikipedia, we also wanted to better understand and characterize article-level conflicts. Our goal was to develop an automated way to identify what properties make an article high in conflict using machine learning techniques and simple, efficiently computable metrics.

Defining Metrics

We extracted a set of 30 page-level objective metrics for each article for comparison (see Table 2). Metrics were selected to be easily computable and scalable to a large set of documents. As the purpose was to see if statistics about the editing history of a document were enough to identify its level of conflict, neither user-dependent measures (e.g., category assignments) nor semantic features (e.g., actual article content) were used.

Metric type	Page
Revisions (#)	Article, talk, article/talk
Page length	Article, talk, article/talk
Unique editors	Article, talk, article/talk
Unique editors / revisions	Article, talk
Links from other articles	Article, talk
Links to other articles	Article, talk
Anonymous edits (#, %)	Article, talk
Administrator edits (#, %)	Article, talk
Minor edits (#, %)	Article, talk
Reverts (#, by unique editors)	Article

Table 2. Page metrics used as inputs into the machine learner. # refers to the raw count, % refers to the percentage for each metric.

Identifying a Marker for Conflict

In addition to the metrics we also required a human-labeled marker for determining the degree of conflict an article is in. Here we were aided by Wikipedia's categorization scheme. In Wikipedia articles can be labeled with multiple tags. Tags

are markers which are replaced with templates in the final version of an article seen by browsers; for example, Figure 7 shows the final browser output for the tag "{{controversial}}". Placing a "controversial" tag on a page also automatically places that page in the "List of controversial topics" category.
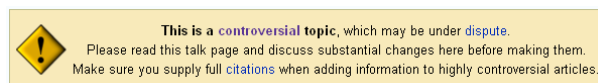
Figure 7. "Controversial" tag

The "controversial" tag provides us with human-labeled conflict data. However, looking at whether the latest revision had the tag or not would be both limited and noisy, as the tag could have been on an article for hundreds of revisions and just happen to have been removed in the very latest revision. It also only provides a coarse, in-or-out decision criterion.

Instead, we developed a measure called the Controversial Revision Count (CRC). The CRC is the count of the total number of revisions in which the "controversial" tag was applied to the article (see Figure 8), providing a parametric measure of conflict for an article. We calculated the CRC for all revisions of every article on Wikipedia (58+ million), ending up with 1343 articles with CRC scores greater than zero (meaning they had at least one "controversial" revision), 272 of which were marked as controversial in their latest revision.
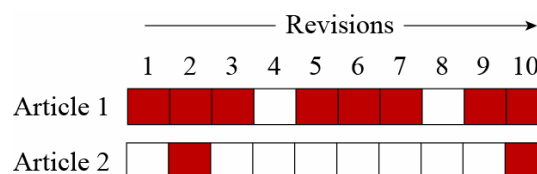

Figure 8. Articles with differing CRC counts. Shaded squares represent revisions that have been tagged as controversial. Article 1 has a CRC of 8, while Article 2 has a CRC of 2.

Machine Learning

Training the Model

The machine learner's goal was to predict CRC scores from the raw page statistics. To do this we used the SMOreg Support Vector Machine (SVM) regression algorithm [21] in the YALE machine learning environment [17].

We computed page metrics and CRC scores for each "controversial"-labeled article. Only pages labeled "controversial" in the latest revision in our dataset were used to train the model. Five-fold cross-validation on this set (training on 4 out of 5 randomly-selected partitions and testing on the remaining partition) gave an R^2 of 0.897 (see Figure 9). This means that the machine learner, by using a combination of page metrics to predict CRC scores, is able to account for about 90% of the variation in the CRC scores. This suggests that the learned model was very effective at predicting CRC from the page metrics.
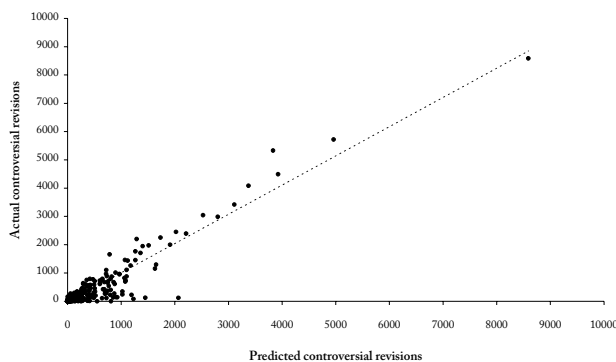
Figure 9. Model performance on articles tagged as controversial. $R^2 = 0.897$.

Important Metrics

A key question we aimed to address was to understand what features of a page characterize its level of conflict. The machine learner provides insight to this in the weights it assigns to various page metrics. These weights are determined by the utility of a metric in predicting CRC scores, and are shown in order of importance in Table 3.

⬆	1. Revisions (article talk)
⬆	2. Minor edits (article talk)
⬇	3. Unique editors (article talk)
⬆	4. Revisions (article)
⬇	5. Unique editors (article)
⬆	6. Anonymous edits (article talk)
⬇	7. Anonymous edits (article)

Table 3. Highly weighted metrics, rank ordered. Up arrows indicate positive correlation with conflict; down arrows indicate negative correlation with conflict

By far the most important metric to the model was the number of revisions made to an article talk page (#1 above). This is not unexpected, as article talk pages are intended as places to discuss and resolve conflicts and coordinate changes. Some of the metrics are more surprising; for example, one might expect that the more points of view are involved, the more likely conflicts will arise. However, the number of unique editors involved in an article *negatively* correlates with conflict (#5 above), suggesting that having more points of view can defuse conflict.

Another interesting finding is that while anonymous edits to the article talk page correlate with increased conflict (#6), they correlate with reduced conflict when made to the main article page (#7). This suggests that anonymous editors may be valuable contributors to Wikipedia on the article page where they are adding or refining article content. However, anonymity on the article talk page, where heated discussions often occur, seems to fan the flames. This suggests that anonymity may be a two-edged sword, useful in lowering participation costs for content but less so in conflict resolution situations.

Such findings may have more general implications for strategies of dealing with conflict. For example, one potential method for reducing conflict, if desired, might be to increase the number of people involved. They also raise the question whether anonymity on the article talk page may hurt more than it helps, though providing the ability for serious anonymous contributors to discuss articles may outweigh the risks. Balancing the desire to reduce controversies and the need to provide protection for minority opinions expressed through anonymity is an interesting design consideration for online collaboration spaces.

Generalization and Validation

Given that the model is quite successful in predicting conflict for articles that have been tagged as "controversial", our next goal was to generalize it to non-tagged articles as well. This is especially important because only a tiny percentage of articles have been tagged as "controversial" at any point in their lifespan. Generalizing to articles that had never been tagged as "controversial" would be a strong indication of success for the model.

To validate the model we asked Wikipedia administrators to provide a baseline to compare to. We first applied the model to all articles in Wikipedia with over 100 edits, generating CRC predictions for each. A small set of 28 articles were then sampled to represent a range of predicted CRC values (limited by the extensive time needed for human evaluation; even these articles took some users multiple hours to rate). We developed an online survey using these 28 articles, with separate 7-point Likert scale ratings for conflict, quality, and vandalism.

Thirteen administrators completed the survey, providing a metric of comparison for the model. The predicted CRC scores for each article were correlated with the mean ratings made by users (see Figure 10). The results supported the validity of the model, showing significant agreement between the model's predicted scores and users' conflict ratings (by Pearson correlation, $r(27) = .47, p = .012$). There were also significant correlations between vandalism and predicted CRC ($r(27) = .430, p = .022$) and vandalism and user conflict ratings ($r(27) = .393, p = .039$). There were no significant correlations between quality and any of the other metrics.

The above results suggest that the model succeeds in identifying conflict even for articles that have never been tagged as such. However, there are limitations to the model and the rating method we used. For example, we had users rate conflict separately from vandalism, while the two may actually be related (as suggested by the significant correlation between them). Predicted CRC appears to be affected by both, as correlating it with the average of the conflict and vandalism scores raises the correlation coefficient to .576 ($p < .001$). Also, the model may be affected by edit patterns that superficially look like conflict but instead reflect frequent updating, such as the documenting of current events.
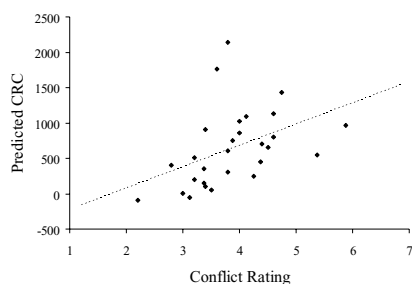
Figure 10. Model-predicted CRC scores vs. user conflict ratings for a set of 28 sampled articles. The top two outliers are articles with high vandalism scores but low conflict scores.

Despite these limitations, this analysis demonstrates the potential for automated prediction of conflict from metrics. Using efficiently computable metrics and standard machine learning techniques, we were able to successfully predict the degree of conflict of an article. Further applications of the model are proposed in the Discussion.

REVERTS AS CONFLICT AT THE USER LEVEL

The characterization of conflicts between users is crucial to understanding the motivation of users and the sources of conflicts. The goals are to 1) identify users involved in conflicts; 2) characterize ongoing conflicts; and 3) develop a tool that can help in understanding the conflicts.

User Conflict Model

To characterize user conflicts, we needed a measure of how much conflict a user is engaged in. One metric to capture this is the history of reverts between users. Reverts are often used to block other editors' contributions and to promote one's own viewpoints, and thus include information about whom a person is engaged in conflict with, how much, and on which pages. As shown in Table 4, they represent a very rich dataset as a huge number of reverts have been made in Wikipedia.

Users Total	3,769,347
Users who made at least one revert	402,454
Reverts (MD5 hash method. see Table 1)	3,711,638
Self-reverts	582,373
Pages that have revert records	721,866
Pages with 50 reverts or more	9,973
Reverts made to pages with 50 reverts or more	1,542,701

Table 4. User and revert statistics

It is important to note that there are many other ways in which conflict is expressed, of which reverts are only one form. Furthermore, reverts are complex actions that are an inherent part of the suggestion and negotiation that happens on discussion pages. This process of settling on an accepted change is an important part of the natural evolution of content on Wikipedia. However, reverts remain a useful proxy for understanding conflicts between users, especially in reflecting extra work put into the system.

For an individual user, using reverts to understand conflicts is difficult because 1) multiple users are often involved in chains of reverts; 2) the edit history is typically long and tedious to browse; and 3) there exist various types of reverts such as the "edit war" (repeated reverts between two users), the "wheel war" (repeated revert between two administrators), self-reverts, and so on. To address these challenges, we developed a user conflict model based on the following principles.

- Self-reverts are disregarded.

- The amount of dispute between two users is related to the number of reverts between them (their "revert relationship").

- When both users A and B revert edits of user C, users A and B are assumed to be in the same user group unless A and B have a revert relationship also.

- When a page is reverted to an older version, we only count the revert relationship between the reverter and the user who made the immediate last edit.

Revert Graph – Visualizing User Conflict

Based on the above user conflict model, we built a tool called Revert Graph to visualize user conflict on a particular article. Revert Graph retrieves all users who have participated in reverts and visualizes a graph based on revert relationships between the users (Figure 11 and Figure 12).
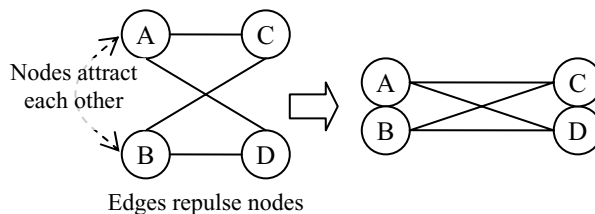
Figure 11. Force directed layout structure employed in Revert Graph. Users (represented as nodes) attract each other unless they have a revert relationship. A revert is represented as an edge. When there are reverts between users, they push against each other. Left figure: Nodes are evenly distributed as an initial layout. Right figure: When forces are deployed, nodes are rearranged in two user groups.

User Cluster: Using the above principles, we can identify user clusters based on the assumption that a group of users have closer views on a topic the more they revert users in another user group. Once users are laid out on the screen, the above user conflict model is simulated by the force directed layout [15] as shown in Figure 11. The force directed layout rearranges users based on their link structure. As forces in the graph become stabilized, social structures between users emerge as shown in Figure 12. The size of a node is proportional to the log of the number of reverts. Nodes are also color coded based on user status: green for administrators, gray for registered users, and white for anonymous contributors.
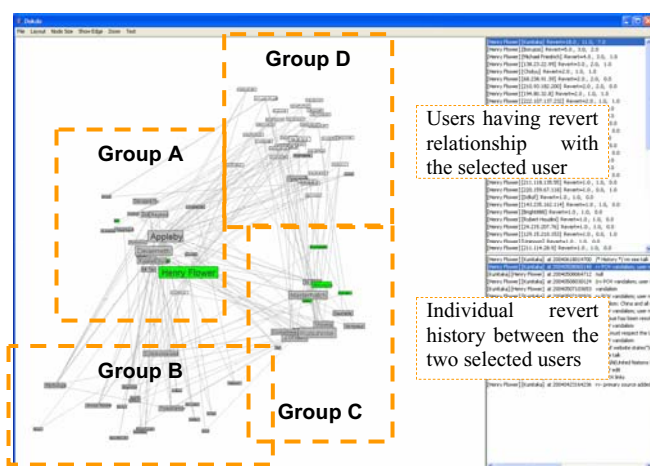
Figure 12. Revert Graph for the Wikipedia page on *Dokdo* [26]. Revert Graph uses force directed layout to simulate revert relationship between users. The tool also allows its users to drill down into revert relationships, which enables them to investigate the nature of the conflicts.

Revert Graph allows easy identification of user groups representing opinion groups, the motivation of edits, and the conflict detail. The tool provides intuitive user clusters and interactive revert history browsing. Revert Graph also enables the investigation of revert relationships at the level of individual reverts. When a user node is selected by clicking, the upper right panel displays the list of users that have revert relationships with the user. Selecting a second user in the list, the bottom right panel shows revert records between the two users. Clicking an item in the bottom right list launches a web browser showing the revert record.

Case Study
We performed a pilot case study to investigate the effectiveness of our user conflict model. A set of test cases were loaded into Revert Graph. When user nodes form a set of user clusters, we evaluated the characteristics of each user cluster by manually looking up the users' edit log.

The Wikipedia page on *Dokdo* (Figure 12) is one example where we were able to find interesting user clusters. *Dokdo* is a disputed islet in the Sea of Japan (East Sea) currently controlled by South Korea, but also claimed by Japan as *Takeshima* [26]. Figure 12 shows user groups discovered on the *Dokdo* article. We manually labeled each user based on his/her position on the issue and summarized the result as shown in Table 5. Group D, where 31 out of 34 users are non-registered users, is not considered in this analysis because they mainly have very short edit histories.

Number of users in group	A	B	C	Total
Users who support the Korean claims	10	6	0	16
Users who support the Japanese claims	1	8	7	16
Neutral or Unidentified	7	3	6	17

Table 5. User Groups on the *Dokdo* article.

The result in Table 5 shows that the identified user groups indeed represent opinion groups. Users in each group

showed cohesiveness toward the issue. The majority of users in Group A supports the Korean claims while users in Group C show the opposite pattern. Users in Group B don't have enough revert history to have been classified accurately.

Figure 13 shows two examples where distinct user clusters emerged. In the Terri Schiavo case, one group supports her husband's descision (Group A) – the removal of Schiavo's feeding tube, while another group represents those who object to euthanasia (Group B). In the Charles Darwin case, one group represents users who view evolution as a fact (Group F) and other groups classify it as a theory (Groups D, E). Both examples show cohesive user groups.
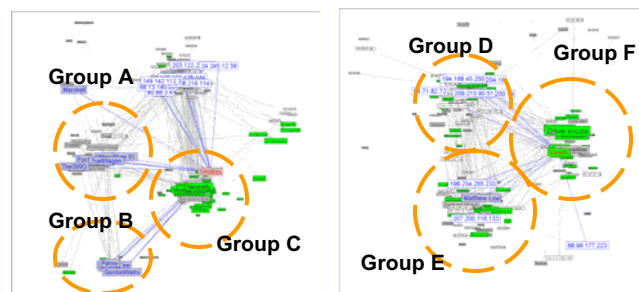
Figure 13. Case study results. Our preliminary investigation revealed some interesting characteristics of user groups. Left: Revert Graph applied to a set of the users who participated in the Terri Schiavo page [28]. Users in Group A appear to be sympathetic to her husband's decision. Users in Group B tend to defend more religious and/or conservative values. Group C is largely composed of admins. Right: Revert Graph for the Charles Darwin page [25]. Users in Group F tend to classify evolution as fact. Group D and E appear to be divided by a variety of disagreements.

As shown in the examples, Revert Graph enables users to intuitively identify and explore user groups. We also found that there are limitations to this tool. The force directed layout does not always produce optimal user groups, requiring sufficient revert relationships to be available. Also, since Revert Graph relies on revert relationships, the tool cannot detect conflicts between users who were not involved in reverts.

DISCUSSION

Conflict and Coordination Costs
As social collaborative knowledge systems grow, so do opportunities for conflict and coordination costs. In the first part of this article we demonstrate a way to quantify these costs at the global level that provides insights into how growth in Wikipedia is occurring. We show that, even though Wikipedia continues to grow exponentially, the rate of creation of new articles and content is decreasing, while levels of maintenance and indirect work are increasing. These results provide the first comprehensive view of this phenomenon, and reflect the entire history of all Wikipedia articles rather than a small sampling of pages.

These data are consistent with the findings from studies of group work systems which suggest that, to keep functioning,

a group must engage in both task-focused and group maintenance activity [4][10]. Continued growth in Wikipedia is not maintained merely by an increase in articles and quality content; the sophisticated procedures developed for coordinating users and dealing with conflict are vital for a community where people may not agree on everything.

This suggests that despite the many unique qualities of Wikipedia – such as its size, low participation costs, and "swarm" intelligence – the results found here may have broad application to other systems in which maintenance activities occur, or where multiple viewpoints interact, including many group work systems [10][13]. Furthermore, the characterization of the growth of coordination costs in Wikipedia provides insights into how large knowledge systems such as collaborative hypertext and organizational memory systems evolve [1][18].

Predicting Conflict From Content
In the second part of this article we showed how conflict can be predicted from article level content. We developed a new measure (the Conflict Revision Count) which was used to train a machine learning model to predict conflict from simple, efficiently computable content metrics. The model was used to predict conflict on untrained articles never marked as being in conflict, and was shown to correlate highly with experts' conflict ratings, corroborating its validity.

This analysis demonstrates the possibility of developing automatic detection and prediction models of complex phenomena such as controversies in large scale social collaborative systems. The process of building such models can also lead to identifying important and sometimes unintuitive metrics correlated with the phenomena investigated (such as the negative correlation found between conflict and the number of unique article editors involved) which can inform the design of new policies and tools. These techniques may have wider applicability to large scale social knowledge systems with a marker for the phenomena of interest (in our case, CRC) and relevant computable metrics (e.g., edits, page views, etc.).

Visualizing Conflict Through Revert Relationships
The third part of this article presents a novel way of visualizing conflict between users. By applying a force-directed layout of the graph describing revert relationships between users we were able to cluster users into groups based on shared points of view.

This visualization technique provides a research tool for modeling conflict in large scale online communities in which relationships between users can be quantified either as conflict behavior (in which case edges repel each other, as demonstrated here) or as collaborative behavior (in which case edges would attract each other). This can also be a practical tool for users in such environments trying to make sense of complex relationships between users, such as in online dispute resolution [6].

APPLICATIONS AND FUTURE WORK

One of the significant advantages of an open-source system is the efficiency of matching users with work relevant to their interests and expertise, and Wikipedia is no exception [2][7]. However, this strength could turn into a weakness if self-selection leads to insufficient diversity in points of view, leading to lower quality articles or increased unproductive conflict.

One way to deal with this problem is to provide alternate methods of matching users with work based on the needs of the community. Such routing methods have been shown to have large effects on user behavior [7]. This is especially relevant to conflict, as only a tiny percentage of conflict articles are tagged as such, and thus do not attract attention until already embroiled in revert wars.

A future application of the conflict model developed here is to identify controversial articles before they have reached a critical conflict point. Our data suggest that an effective way to resolve conflict is to increase the number of users involved in editing the article, rather than have the same few people arguing back and forth. Even if a small percentage of users involved themselves in these pages they could prove vital to defusing conflict before it gets out of hand.

In addition to the "article conflict detector", another application is a "user conflict detector" based on Revert Graph which helps surface editing and revert patterns that could identify high-conflict users. Such a "social dashboard" could help community members make sense of the edit history of a particular article or user, and could provide a way to identify or apply social pressure to high-conflict users.

These applications highlight the idea of channeling attention on Wikipedia to areas that need it most, rather than relying solely on user interest. This approach could play an increasingly relevant role in the continued growth of Wikipedia as conflict and coordination costs continue to rise. An important part of this is surfacing content and statistics relevant to attention-allocation and sensemaking tasks that would not normally be available to a single user.

CONCLUSION

Throughout this paper we have presented methods to characterize conflict in Wikipedia at the global, article, and user levels. First we presented details of the growth of conflict and coordination costs at the global level across Wikipedia's history. We then showed that conflicts at the local article level can be modeled and predicted using a machine learner. Finally, we depicted the conflicts that occur at the user level, demonstrating the use of visualization in making sense of disputes between users.

We believe that the characterization on the growth of conflict and coordination costs provides insights into how a Wiki solution to collaboration in hypertext systems can scale to very large sizes, with potential implications for the study of other large groupware and organizational memory

systems. The methods developed for predicting conflict from simple metrics and visualizing user conflict also present novel ways to analyze large scale online collaborative systems in which users interact to produce knowledge. Further research is needed to explore how these findings generalize to other collaborative knowledge systems.

Acknowledgements

REFERENCES

1. Baecker, R. Cooperative hypertext and organizational memory. In *Readings in Groupware and Computer-Supported Cooperative Work* (1992), 519-580.

2. Benkler, Y. Coase's penguin,or Linux and the nature of the firm. *Yale Law Journal 112* (2002), 367-445.

3. Buriol, L., Castillo, C., Donato, D., Leonardi, S., and Millozzi, S.: "Temporal Evolution of the Wikigraph", *Proc. Web Intelligence Conference 2006*. IEEE CS Press (2006), 45-51.

4. Butler, B., Sproull, L., Kiesler, S., and Kraut, R. Community Effort in Online Groups: Who Does the Work and Why? in *Leadership at a Distance*, Erlbaum Associates (2001), 346-362.

5. Capocci, A. Servedio, V.D.P., Colaiori, F., Buriol, L.S., Donato, D. Leonardi, S., and Caldarelli, G., Preferential attachment in the growth of social networks: the case of Wikipedia, arXiv.org:physics/0602026 (2006).

6. Conley-Tyler, M. 115 and counting: The state of ODR 2004. *Proc. 3rd Annual Forum on Online Dispute Resolution* (2004).

7. Cosley, D., Frankowski, D., Terveen, L., Riedl, J. Using Intelligent Task Routing and Contribution Review to Help Communities Build Artifacts of Lasting Value. *Proc. CHI 2006*, ACM Press (2006), 22-27.

8. Curtis, P. *Mudding: Social phenomena in text-based virtual realities.* Palo Alto Research Center, 1992.

9. Dibbell, Julian. A Rape in Cyberspace. *Village Voice, 21* (1993), 26-32.

10. Dourish, P., & Bellotti, V., Awareness and coordination in shared workspaces. *Proc. ACM Conference on Computer Supported Cooperative Work (CSCW '92),* ACM Press (1992), 107-114.

11. Franco, V., Piirto, R., Hu, H. Y., Lewenstein, B. V., Underwood, R., and Vidal, N. K. Anatomy of a flame: conflict and community building on the Internet. *Tech. and Society Magazine, IEEE, 14* (1995) 12-21.

12. Giles, G. Internet encyclopaedias go head to head. *Nature*, *438*, 7070 (2005), 900-901.

13. Grudin, J. Groupware and social dynamics: Eight challenges for developers. *Communications of the ACM 37* (1994), 92-105.

14. Hadoop Project. http://lucene.apache.org/hadoop/.

15. Heer, J., Card, S. K. and Landlay, J. Prefuse: A toolkit for interactive information visualization, *Proc. CHI 2002*, ACM Press (2005), 421-430.

16. Lih, A. Wikipedia as participatory journalism. *Proc. 5th Intl. Sym. Online Journalism,* (2004), 16-17.

17. Mierswa, I. , Klinkberg, R. , Fischer, S., and Ritthoff, O. YALE -- Yet Another Learning Environment. *LLWA 03 - Tagungsband der GI-Workshop-Woche Lernen* (2003).

18. Neuwirth, C. M., Kaufer, D. S., Chandhok, R., and Morris, J. H. Issues in the design of computer support for co-authoring and commenting. *Proc. CSCW 1990*, ACM Press (1990), 183-195.

19. Noel, S., and Robert, J. M. How the Web is used to support collaborative writing. *Behaviour & Information Technology, 22* (2003) 245-262.

20. Rivest, R., The MD5 Message-Digest Algorithm, RFC 1321, MIT and RSA Data Security, Inc., April 1992.

21. Smola, A.J., and Scholkopf, B. A tutorial on support vector regression. *Statistics and Computing, 14* (2004), 199-222.

22. Stvilia, B., Twidale, M. B., Smith, L. C., Gasser, L. Assessing information quality of a community-based encyclopedia. In *Proc. ICIQ* (2005), 442-454.

23. Viégas, F. B., Wattenberg M., Dave K., Studying cooperation and conflict between authors with history flow visualizations, *Proc. CHI 2004*, ACM Press (2004), 575-582.

24. Voss, J. Measuring Wikipedia. *Proc. 10th Intl. Conf. Intl. Soc. for Scientometrics and Informetrics* (2005).

25. Wikipedia.org, Charles Darwin. http://en.wikipedia.org/wiki/Charles_Darwin

26. Wikipedia.org, Dokdo. http://en.wikipedia.org/wiki/Dokdo

27. Wikipedia.org, Edit war, http://en.wikipedia.org/wiki/Wikipedia:Edit_war

28. Wikipedia.org, Terri Schiavo. http://en.wikipedia.org/wiki/Terri_Schiavo

29. Wikipedia.org, John Siegenthaler Sr. controversy. http://en.wikipedia.org/wiki/John_R._Seigenthaler_Sr._Wikipedia_biography_controversy

30. Wikipedia.org, Three-revert-rule. http://en.wikipedia.org/wiki/Wikipedia:Three-revert_rule

31. Zlatic, V., Bozicevic, M., Stefancic, H., Domazet, M. Wikipedias: Collaborative web-based encyclopedias as complex networks. *Physical Review E 74* (2006).