Proceedings of the Sixth International AAAI Conference on Weblogs and Social Media

Temporal Motifs Reveal the Dynamics of Editor Interactions in Wikipedia

David Jurgens[1,2]
[1]University of California, Los Angeles
Los Angeles, California, USA
jurgens@cs.ucla.edu

Tsai-Ching Lu[2]
[2]HRL Laboratories, LLC
Malibu, California, USA
tlu@hrl.com

Abstract

Wikipedia is a collaborative setting with both combative and cooperative editing. We propose a new method for investigating the types of editor interactions using a novel representation of Wikipedia's revision history as a temporal, bipartite network with multiple node and edge types for users and revisions. From this representation we identify significant author interactions as network motifs and show how the motif types capture important, diverse editing behaviors. Two experiments demonstrate the further benefit of motifs. First, we demonstrate significant performance improvement over a purely revision-based analysis in classifying pages as combative or cooperative page by using motifs; and second we use motifs as a basis for analyzing trends in the dynamics of editor behavior to explain Wikipedia's content growth.

1 Introduction

The open access policy of Wikipedia has made it possible for anyone to modify Wikipedia, e.g., adding, editing, and deleting existing content, or creating and removing pages altogether. While early speculation suggested that this model was unsustainable, Wikipedia has continued to grow and improve in quality (Arazy and Nov 2010). Underlying this growth are the cooperative –and sometimes combative– interactions between editors working on the same content. Because these interactions drive the content creation process, understanding and quantifying their effects is essential for accurately modeling Wikipedia as a system, as well for analyzing the characteristics of editors, pages, and the interactions themselves. In this work, we present a data-driven model of editor interactions using network motifs. Using motifs as the foundation for analysis, we demonstrate how the micro-level dynamics of motifs characterize larger scale phenomena via two tasks: classifying page collaboration characteristics and explaining Wikipedia's continued content growth.

Much of the current analysis of Wikipedia's authors has focused on high-level features such as the number of inter-editor reverts (Kittur et al. 2007) or interactions on Talk pages (Laniado et al. 2011). Recent studies have begun assessing the impact of specific editing behavior by looking at how editors revise each others work (Brandes et al. 2009; Laniado and Tasso 2011). However, the focus has been on identifying contentious behavior between editors, and accordingly almost all models are limited in their ability to represent and analyze cooperative behaviors.

To achieve a full analysis of all types of editing behavior, including both cooperative and combative, we propose a new method of pattern analysis on Wikipedia's revision history. The revision history can be viewed as a bipartite graph from editors to pages. Enriching this graph with temporal information of both who edited the article and how the article was changed enables discovering meaningful editing behavior in the form of network motifs. These temporal motifs are repeated subgraphs of the editing graph which correspond to significant patterns of collaborative interactions. Furthermore, the discovered motifs represent multiple types of editor behavior in a single framework, which allows us to compare different types of interactions and their relations to page properties, such as quality and contentiousness.

The contributions of this paper are as follows. In Sec. 2, we define a new network representation of Wikipedia's revision history as a dynamically evolving bipartite graph with multiple node and edge types. We propose a new method for identifying core editor interactions through a motif-based analysis of the graph. From these discovered motifs, we further propose a new method in Sec. 3 for classifying pages as cooperative and combative based on motif patterns. Last, in Sec. 4, we derive a new generative model for Wikipedia's growth from motifs-based interactions in order to identify the changes in editor behaviors responsible for its sustained growth rate.

2 Network Representation

At a basic level, Wikipedia may be viewed as a bipartite graph from authors to the pages to which they contribute. We expand this representation to encode three additional features: (1) the type of author who made the change, (2) the time at which the change was made, and (3) the magnitude and effect of the change to the page. Together, these three features create a dynamically evolving bipartite graph that has node attributes corresponding to author type and edge

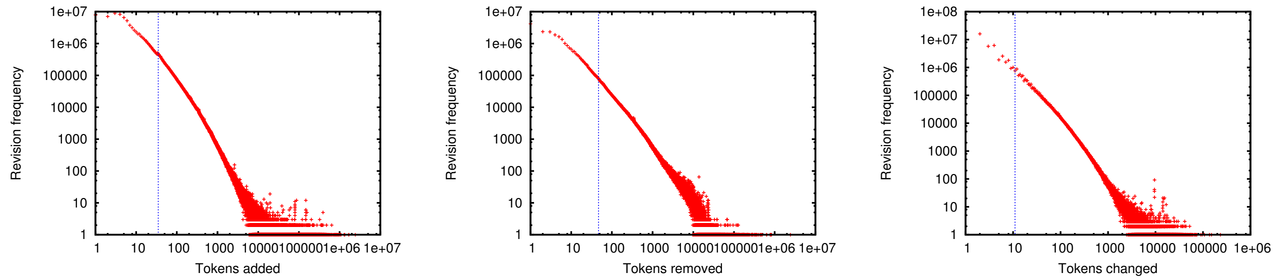

Figure 1: Distributions of the effects for `ADD`, `DELETE`, and `EDIT` types. Vertical lines indicate the division between *major* and *minor* revisions based on the 80/20 rule, where 80% of a type's cumulative effects are due to those to the left of the line.

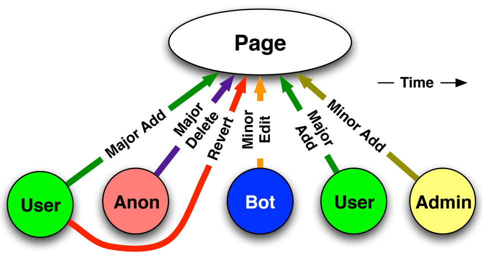

Figure 2: The temporal sequence of edits reveals repeated patterns of editing and author interactions

attributes for the type of revision. We first formally define this graph and then in Section 2.2 describe how the network is derived from Wikipedia's history.

2.1 Definition

The bipartite graph of Wikipedia revisions, $G = (V, E)$, is defined as follows

- The set of nodes V are the union of the sets of all unique Wikipedia author identifiers $A = \{a_1, \ldots, a_m\}$, consisting of registered usernames and IP addresses, and all unique Article pages $P = \{p_1, \ldots, p_n\}$, i.e., $V = A \cup P$.
  - Each author a_i is associated with an attribute defining the class of authors of which it is a member, $c(a_i) = \{registered, anonymous, admin, bot\}$. An author's class is fixed throughout the graph's lifetime.
- The set E contains directed edges from author $a_i \in A$ to page $p_j \in P$, with one edge per revision. Each edge $e_i \in E$ is associated with two attributes:
  - the time t_i at which its revision was made, which is a discrete time value in the range of the network's lifetime $t_0, \ldots, t_n$. Due to Wikipedia revision constraints, two edges to the same page may not have the same time, i.e., no simultaneous revision are allowed.
  - the class describing the effect of the revision, $c(e_i) = \{minor\_add, major\_add, minor\_edit, major\_edit, minor\_delete, major\_delete, revert\}$.

2.2 Network Derivation

The network is derived from a complete revision history of Wikipedia, ending on April 05, 2011. Because the primary focus of this work is on editor interaction on articles, we restrict our analysis to article pages that have at least 10 revisions in their history. The resulting data set contains 2,715,123 articles and 227,034,806 revisions. From this, we derive the node and edge properties of our graph, which encode the effect of the revision.

Author Classes Wikipedia provides a hierarchy of user types, each having its own level of permissions, with bureaucrats and system operators at the top and anonymous users at the bottom. Users near the top of the hierarchy have increased editing and administrative capabilities (e.g., banning users), and are often very engaged in Wikipedia. We divide the hierarchy into four categories: administrator, bot, registered, and anonymous.

The administrator category contains all users with at least `administrator` rights, which also includes bureaucrats and stewards. The bot category contains all user names associated with automated programs that autonomously edit the content of Wikipedia. Bots perform a variety of editing tasks from routine changes such as spelling correction or ensuring consistent Wiki formatting, to high impact changes such as importing large sections of text from other sites into Wikipedia's articles. While users may run bot program under their user accounts, our category contains only those bots who have been approved by Wikipedia admins to operate independently on their own account.

We label the authors as follows. All revisions with an IP address for the author are labeled as anonymous. We use the `Wikipedia:Bots/Status` page to identify the 2271 usernames registered to bots, and `Special:ListUsers/sysop` page to identify 1514 administrators, and label their revisions accordingly. All other authors are classified as registered users.

Revision Classes Wikipedia revisions vary considerably in their effect on a page's content, from small copyediting, to adding new sections. While understanding the author's intent behind these actions is not always possible, we can still categorize their effect into broad classes in order to discover patterns in editing behavior. Specifically, we selected in four high-level categories for revisions: adding, deleting, editing, and reverting. Adding and deleting behaviors correspond to changes in the size of a page's content, whereas editing behavior revises the existing content (e.g., fix typos, merge and split subsections, etc.) but does not significantly affect the page size. Reversion is treated as a special type of behavior where one editor undoes one or more previous revisions, returning the page to an earlier state.

To classify revisions into adding, deleting, and editing behaviors, we analyze revisions using two parameters: (1) the number of whitespace-delimited tokens added or removed from the page, δ, i.e., its change in size, and (2) the number of tokens whose content was changed, θ. Token changes include character insertions, removals, and additions as well as adding and removing tokens. These two parameters capture many characteristic types of user behavior. For example, a revision that fixes grammatical mistakes will have a larger number of token changes but is unlikely to increase the page size; conversely, adding new content to a page will likely have proportional values for δ and θ. In addition, these parameters are largely independent of the page size itself, with Pearson correlations of $\rho(\delta, size) = 0.147$ and $\rho(\theta, size) = 0.142$, which makes them suitable for classification across all revisions.

Article reversions are not directly identifiable from the revision logs and so two processes are used to annotate which revisions constitute a revert. First, the revising author's comment in the logs is searched for a fixed set of phrases that are known to be made by editors when reverting a page's content, e.g., "reverting," "rv," or "rvv." While these capture the majority of reversions, some editors choose not to leave comments, or do not indicate the revert in their comments. Therefore, we adopt a second process used by Kittur et al. (2007) where an MD5 hash is computed for the full text (including wiki markup) of each revision. Then, the hash is matched against the hashes of all previous revisions of the page to check whether the current content is equivalent; if so, the change is marked as a revert.

All other non-revert revisions are classified by assessing whether content creation or deletion was responsible for the majority of the token changes. Specifically, a revision r is classified according to

$$class(r) = \begin{cases} \texttt{EDIT} & \text{if } \frac{|\delta|}{\theta} < 0.5 \\ \texttt{ADD} & \text{else if } \delta > 0 \\ \texttt{DELETE} & \text{otherwise} \end{cases}$$

Simply, a revision is classified as EDIT if more existing tokens were changed than were added or deleted, and otherwise classified according to its net effect on the page size.

However, revisions within a single class are not equal in magnitude. Figure 1 shows the parameter value distribution for the non-reversion classes, which follow power-law like distributions. The relative rarity of large changes is consistent with the earlier 2004 observation of Viégas, Wattenberg, and Dave (2004) who noted that among size-reducing edits, only 6% removed more than 50 characters. Therefore, to further distinguish edits based on the magnitude of their effect in addition to the type, we partition each class into *major* and *minor* subcategories, with the exception of REVERT. Based on the shape of the effect distributions, the difference between major and minor was selected using the Pareto principle, or "80/20 rule" (Newman 2005), which in the present case translates as 80% of the editing effects coming from the most frequent 20% of the types of edit instances. That is, the revisions with small effects account for the majority of the *cumulative* effects to the content. Therefore, we divide the classes such that ADD revisions with $\delta > 35$ are major, DELETE revisions with $\delta < -47$ are major, and EDIT revisions with $\theta > 11$ are major.

2.3 Detecting Interaction Motifs

The revision graph contains editor interactions as temporally contiguous subgraphs. Figure 2 illustrates a subset of a page's history as sequence of classified revisions. These revisions can be partitioned by time to identify temporal *motifs*, or repeated network structures that constitute meaningful building blocks of a more complex network (Milo et al. 2002). The distribution of these motifs across all of Wikipedia's history reveals how particular interactions impact the overall growth of its content.

The set of candidate motifs was selected from all subgraphs made of three contiguous edits on a single page, which contains a variety of subgraph configurations: from two vertices (one author making three changes in a row) to four vertices (three different authors). This motif range was selected to capture the reactions of authors when another author has revised their changes. Using four vertex types and seven edge types, 39,788 unique motif types are possible.

Null Model Typically, the importance of motifs is evaluated by comparing the frequency of the motif relative to its frequency in a null model. The choice in null model is key to motif detection; the structural (and temporal) configuration of the networks in the null model must match those of the original in order to properly estimate the expected motif frequencies. In a simple network, with undirected edges and no edge and vertex types, the null model may be computed by randomizing the test network while still preserving its degree sequence, i.e., the number of edges connected to each vertex remains the same while the identity of the connected vertices varies. The expected frequencies of each motif are calculated by repeatedly sampling networks from a null model and computing the frequency distributions of the motifs therein.

We construct our null model based on individual articles' revision histories. Following the null model construction suggested by Kovanen et al. (2011), the revisions for each page are randomized while still preserving both the degree sequence of the authors as well as the vertex and edge types distributions. Essentially, each page must be revised in the same way by the same type of author, but the identity and relative ordering of that revision may change. We randomize the edges of the network 5,000 times to estimate the expected motif frequencies.

Motif Statistics The revision dataset contained 218,888,574 total motifs, 39,034 of which were unique. For each motif, we calculated the Z-score relative to its expected frequency in the null model. As expected, a statistical analysis of motif frequency reveals that the revision process is highly non-random: 9,971 motifs had $z \geq 1$ and 27,697 motifs had $z \leq -1$, while only 1,366 had $|z| < 1$. The large percentage of motifs with $z \leq -1$ suggests that while the

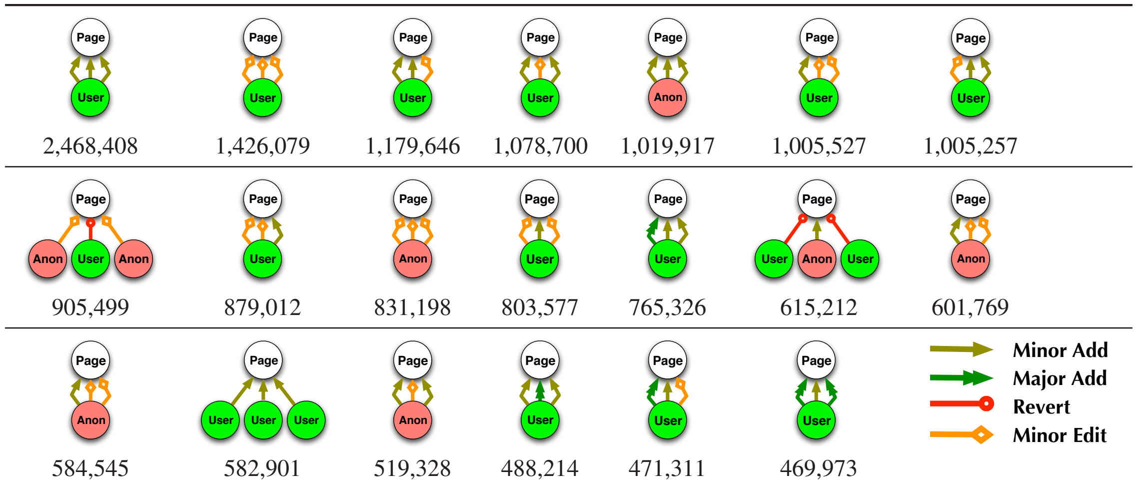

Table 1: The most frequent 20 motifs reveal that motifs involving a single author are the predominant method of revising. The frequency of each motif is shown below.

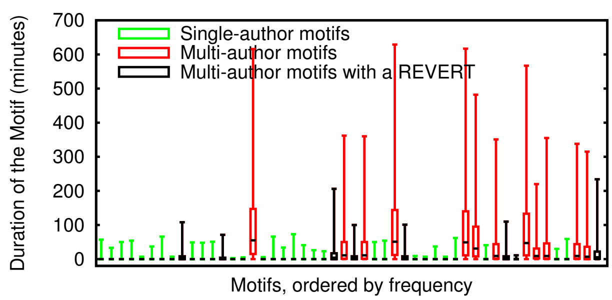

Figure 3: The distribution of 50 most frequent motif durations reveals three distinct types of temporal behaviors based on a motif's authors and edit types. Boxes denote 25th and 75th quartiles, whiskers denote the 2nd and 98th.

Motif type	Frequency	%
Single-author	45,201,434	20.65
Two-authors (all)	70,061,603	32.01
with intra-motif REVERT	13,847,427	(6.33)
Three-authors (all)	103,625,537	47.34
with intra-motif REVERT	28,621,782	(13.08)

Table 2: The frequency distribution of all motif instances

many motif configurations can be created from the revision and author types, relatively few of these motif interactions actually drive the content creation and revision process.

The frequency distribution of the motif instances highlights the differences in single-author editing relative to multi-author collaborations. Among the most frequent motifs, shown in Table 1, single-author motifs dominate. However, multi-author interactions become increasingly common, with cooperative, non-reverting multi-author interactions (such the 16th most-common) accounting for 37 of the top 100 most frequent motifs, and ultimately 59.9% of the total motif instances. Table 2 summarizes motif frequencies with respect to the number of authors, with subtotals for motifs where one author in the motif has reverted another.

The temporal properties of motifs reveal distinct interaction and editing behaviors. Fig. 3 shows the duration distributions for the 50 most frequent motifs. Three temporal categories emerge, which fall sharply based on the participants and editing types of a motif. First, single-author motifs are very quick in duration, with 1052 of the possible 1330 single-author motif types having a median duration of under five minutes, including all those in the most frequent 50 motifs. In addition, the most frequent single-author motif instances account for more than just MINOR revision types, with 10 of the 39 single-author motifs in Fig. 3 including at least one MAJOR_ADD. This suggests that major content addition is not a singular event, but rather the result of multiple contributions done on a short time scale.

Second, multi-author motifs fall into two temporal classes based on whether they contain a REVERT revision. Multi-author motifs without a REVERT occur on much longer time scales than single-author motifs, with only 713 of the 23,594 multi-author motif types having a median duration under five minutes, and only 40% having a median duration under 60 minutes. In contrast, multi-author motifs with a REVERT occur much faster, with 6,558 of all 13,961 such motifs having a median time under 5 minutes, including all those in the most frequent 50 motifs. The strong correspondence between reverting an edit and combating vandalism suggests that such short durations are due to active participation by Wikipedia community members, such as the Counter Vandalism Unit, which actively monitors recent revisions for potential vandalism. In addition, the low median motif duration concurs with the work of Viégas, Wattenberg, and Dave (2004) who noted around half of mass-deletion vandalism was corrected within 3 minutes, as well as a small-scale 2009 study that reported the median time for reverting an incidence of vandalism was four minutes (Cobb 2009).

3 Cooperative and Combative Behavior

Author interactions are likely to differ based on the topic of an article. We consider whether the types of interactions seen on contentious pages differ from those of high quality pages, with the assumption that high quality pages will have more cooperative interactions, while contentious pages will have more combative interactions.

Recent work has assessed cooperative and combative behavior among Wikipedia users (Brandes et al. 2009; Sumi et al. 2011; Halfaker, Kittur, and Riedl 2011), with most analyses focusing on specific combative behaviors related to page reversions, where two or more warring editors periodically revert each other's changes. In contrast, we analyze combative behavior at a micro-level through user motifs in order to (1) quantify the effects of combative behavior beyond reverts, and (2) compare the relative distribution of these interactions with those of high quality, cooperative pages.

The interactions associated with cooperative and combative behavior were determined by using established categories of pages. Combative behavior is estimated using the motifs found on 720 pages listed in `Wikipedia:List of Controversial Articles`; similarly, cooperative behavior is estimated using the motifs found in 10,149 pages in `Wikipedia:Good Articles` and `Wikipedia:Featured articles`.

To first estimate the difference in the interactions in the two sets of articles, we construct a baseline probability distribution over all motifs using the motif frequencies in all pages. From this baseline distribution, we calculate the Kullback-Leibler (KL) divergence to the average motif distribution in the Cooperative and Combative sets. The KL-divergence serving as a directional measure of the deviation of the second distribution from the first, allowing us to quantify the difference in the sets' motif distributions. The resulting analysis reveals that the motif probabilities in Combative articles differs significantly more from the baseline distribution than that of the Cooperative set: the KL-divergence from baselines to the Cooperative distribution is 0.143, while the KL-divergence to the Combative distribution is 0.283.

The KL-divergence of motif distributions suggest clear distinctions can be made between the page classes on the basis of their editor interactions. Therefore, we form a classification problem of predicting a page's class from its motif distribution. Training data was constructed from all pages in the controversial and cooperative sets. In addition, we selected a random 1% sample of 26,746 pages to represent a third, neutral page class, for a total of 37,614 pages.

In order to test that accuracy is due to the information conveyed by motif's configuration rather than its raw author-edit types, we test against three baselines: (1) selecting a random page class, (2) selecting the most frequent page class, neutral, and (3) classifying based on the author-edit type distribution alone. The third baseline enables us to separate out the contribution of the type information from the information in the specific ordering of the types in the motifs. To avoid potentially overfitting due to the high number of motif features, we train multiple motif-based classifiers, using only the k most frequent motif types in all of Wikipedia as features. Motif-based and author-edit type based SVM classifiers were trained using WEKA (Hall et al. 2009) with 10-fold cross-validation to calculate accuracy. Because the page classes are different sizes, we report the micro-averaged F-score, which calculates the average F-score of all classes weighted by class size. Figure 4 shows the F-score for Motif-models with different values of k relative to the baselines, which are independent of k.

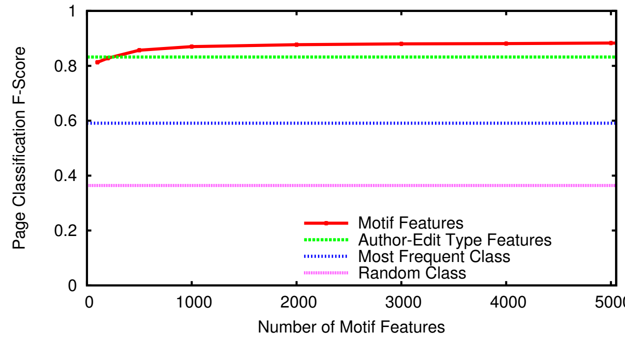

Figure 4: F-scores for classifying a page's type from either the distribution of its author-edit types or the k most frequent motifs, relative to two baselines

Both author-edit type and motif-based classifiers perform well above the random and most frequent label baselines. Notably, the 100-motif feature model achieves a 22.2% absolute improvement in F-score over the most-frequent class baseline while using only 11% of the possible motif instances as features. Increasing the number of available motif features significantly increases performance, with 5,000 motif features achieving a 5.1% absolute F-score improvement over the author-edit type based classifier. The increased improvement as more motif types are added reveals that the more rare motif types still convey significant information to discriminate the quality of a page's interactions.

The impact of using motifs for classification is further emphasized through the performance improvement in each page class type, shown in Table 3. An examination of the motif instances that most contribute to the classifier accuracy at different k values suggests that similar motif configurations serve as distinguishing features between the classes, despite the motifs' overlap in revision types. For example, in the SVM classifier trained with k=500, motifs involving users and anonymous authors reverting each other (i.e., mutual reversion) provided the most positive evidence of a page being in the Contentious set, while most negative evidence of being in the Contentious set was due to motifs where one author reverts another but the third revision in the motif is some form of content creation, e.g., a major addition. Analysis of the author-edit type distribution alone cannot capture these ordering distinctions in how the revision process occurs between authors.

Model	Combative	Cooperative	Neutral
Author-Edit Type	.238	.667	.910
Motif (k=100)	.142	.630	.900
Motif (k=500)	.470	.722	.918
Motif (k=5000)	**.477**	**.784**	**.931**

Table 3: Classifier F-scores for each page class

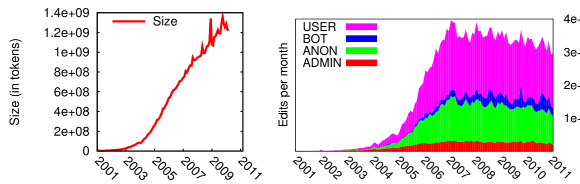

Figure 5: Content growth rate (left) has remained relatively constant despite a highly varied editing frequency (right)

4 Analyzing Content Growth

The dynamics of Wikipedia's editing behavior had changed significantly since its inception. As Fig. 5 shows, relatively few revisions resulted in rapid initial growth, while a corresponding, seemingly exponential rise in the number of edits from 2002-2007 did little to change the content growth rate. Furthermore, despite a slow decline in the number of edits, content growth remains steady. Suh et al. (2009) note a similar decline editing and article creation rate since 2007. Surprisingly, the relative amounts of the different revision types (not shown) has remained stable during Wikipedia's lifetime, with the exception of a growing proportion of user and admin REVERT and anonymous MINOR_EDIT revision. The stability of revision types and content growth rate coupled with large shifts in editing rates points to an unexplained change in the content creation process.

We explain this evolution in editor behavior by analyzing dynamics of author interactions. To analyze these dynamics, we first build a generative model of how content is created as a result of the interactions. While Wikipedia content is the result of multiple editors collaborating, the types of interactions on a page may vary widely according to topic matter. For example, a highly specialized topic may have few authors knowledgeable enough to edit the article and so the editing behavior is primarily single-author; whereas, a controversial page may enjoy wide participation with many authors revising or reverting each other's work. We characterize the types of collaborations seen on a page in terms of editing *behaviors*, which represent probability distributions over the types of interactions authors are likely to have when working on a page. Using behaviors as a starting point, we first formalize the generative model and then analyze how changes in this generative process help explain changes in Wikipedia's dynamics.

4.1 Generative Model

Behaviors are defined as a set of multinomial distributions, $\{\phi_1, \ldots, \phi_B\}$ over the space of interactions represented through unique motif instances $\{m_1, \ldots, m_M\}$. In turn, an article is defined as a distribution over behaviors θ, which reflects the kinds of interactions editors are likely to have for that article. The generative case is summarized as follows:

1. sample $\theta_j \sim Dirichlet(\alpha)$
2. sample $\phi_{1,\ldots,B} \sim Dirichlet(\beta)$
3. for each N_i interaction in article j
   (a) sample a behavior $\phi_{i,j} \sim$ multinomial θ_j
   (b) sample a motif $m_{i,j} \sim$ multinomial $\phi_{i,j}$

Both the distribution of behaviors in a document and the interactions of a behavior are conditioned by Dirichlet priors, which enforces a sparsity constraint where few behaviors are expected per document and each behavior is concentrated on a few motif types. In summary, the motif history of an article is built by repeatedly sampling behaviors for that page, and then sampling motifs for each behavior.

This generative model is analogous to that of Latent Dirichlet allocation (Blei, Ng, and Jordan 2003), where motifs are equivalent to tokens and behaviors to topics. Accordingly, we perform Gibbs sampling to infer the θ and ϕ parameters of the model, with $\alpha = \frac{1}{B}$ and $\beta = 0.01$. The model is implemented using Mallet (McCallum 2002).

4.2 Tracking Changes in Behaviors

To track changes in the behaviors, the generative model was first inferred from the motif present in all article histories up to April, 2011 for articles with at least 10 edits. We set the number of behaviors to B=20 in order to capture high-level patterns in the changes of editor interactions. Next, the temporal motif sequence was partitioned so that all motifs whose last revision ended in the same month were grouped in the same set. Finally, for each month we create a probability mass distribution over the set of behaviors $\{\phi_1, \ldots, \phi_{20}\}$ by summing the posterior probabilities $p(\phi_i|m_j)$ for each motif m_j that ended in that month. Once the distribution is normalized, the relative mass assigned to each behavior reveals shifts in behavior prominence given the observed motifs in each month. Figure 7 visualizes the changes in mass for all 20 behaviors.

Two notable trends are evident in the behavior timeline (Fig. 7). First, aside from the initial year when little data is available, many behaviors are present in the same amount throughout Wikipedia's history. However, several behaviors run counter to the first trend and notably shrink or grow prior to stabilizing in 2007, which matches the stabilization in edit frequency. We highlight four behaviors in Figure 6 whose relative probability mass changed most. During 2002-2007, the probability mass of Behaviors B and C decreased to 24.6% and 32.4% of their earlier respective levels, while A and D more than doubled.

Overall, the trends suggest that the early growth was fueled by content addition from single authors or collaborating between two authors (B) and contributions from administrators. These early behaviors have given way to increases in behaviors associated with editing (A) and maintaining quality or vandalism detection (D). The stability of the remaining behaviors suggests that the core behaviors responsible for editing growth have not changed, explaining the contin-

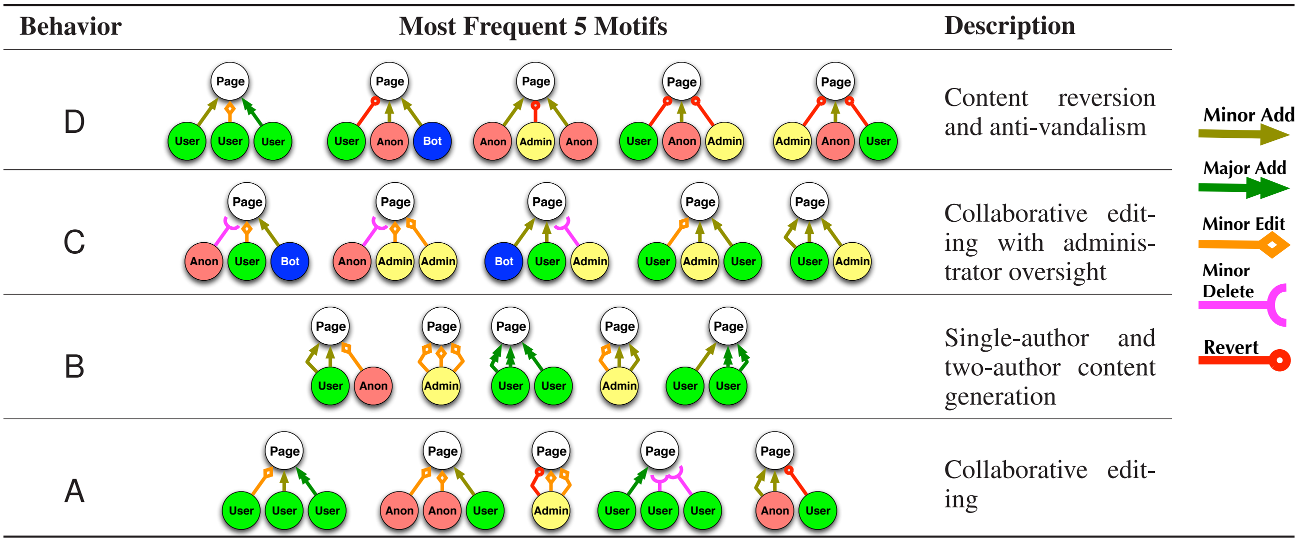

Figure 6: The most probable motifs associated with the labeled behaviors in Fig. 7 show distinct editing characteristics. The relative probability mass of behaviors B and C decreased, while A and D increased.

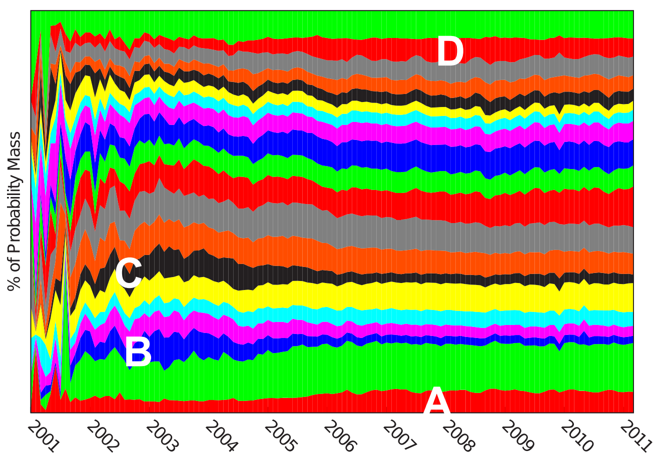

Figure 7: Changes in the relative amounts of probability mass assigned to all 20 behaviors show evolving editor dynamics consistent with the changes in editing frequency. Labeled behaviors are further illustrated in Fig. 6

ued growth. Instead, the increased revision volume without a matching spike in growth rate is due to behaviors focusing on non-augmenting revisions.

5 Related Work

The present work joins both network and temporal analysis to capture the dynamics of author interactions in a single model. Many related approaches have built simpler models for similar analyses. To our knowledge, Wu, Harrigan, and Cunningham (2011) performed the only other motif analysis of Wikipedia, using a small-scale motif analysis on a static, untyped, bipartite graph of author-article and article-article edges for a subset of Wikipedia's articles. Their analysis revealed that motif distributions in topical subcategories varied significantly, suggesting different patterns of contributor behavior based on the content topic. Our model provides a richer motif representation of author behavior and is scalable for analyzing all of Wikipedia as a whole. Vuong et al. (2008) also use a bipartite author-article graph to measure the disputedness of an article's content by differentiating revisions according to the age of the page and author's length of time as an editor. Our analysis indicates that this type of revision-only analysis can be significantly improved by using higher-order motif representations.

Other studies have used author-author interaction graphs to measure article quality or controversy. Laniado et al. (2011) analyze the patterns of editor interactions on author Talk pages. Laniado and Tasso (2011) build a co-authorship network from authors interactions on the same page in order to identify high quality editors on the basis of network properties. Kittur et al. (2007) analyze conflict and coordination in pages both by training a classifier over page-related features, and by building a co-author network for identifying communities of conflicting authors. Brandes et al. (2009) added information to directed edges in their author-author network in order to capture the overall effect of one editor on another's content, thereby enabling the detection of different editor roles or combative interaction patterns. Of these works, only Kittur et al. (2007) addresses multiple types of editing behaviors, but still use separate frameworks for cooperation and conflict; in contrast, our representation accurately models multiple behaviors, allowing for a richer analysis of author dynamics and page characteristics.

Last, several works have begun to analyze the temporal aspects of editing behavior. Roth, Taraborelli, and Gilbert (2008) assess the impact differences of anonymous, regular, and administrative users on Wikipedia's growth. Their study revealed that the frequency of user edits is a strong indicator of page growth, while high concentrations of users or admins on the same page suggested slowed growth. Sumi et al. (2011) analyze the burstiness of an article's revision and Talk page histories to assess whether the article is undergoing a period of high editor conflict or has been vandalized. In contrast, the author and edit features used by these works are incorporated by both our motif model and the behavior-based generative model to capture the temporal changes in higher-level properties.

6 Conclusion and Future Work

We have proposed a novel representation of Wikipedia as a temporal, bipartite graph and shown how motif instances in that graph capture important aspects of author dynamics using a single framework. Furthermore, we demonstrated the utility of the motif representation in two tasks. First, we showed that the temporal information captured in motifs resulted in increased classifier accuracy over the simple revisions types alone for determining a page's editor behavior. Second, we demonstrated how previously unaddressed changes in Wikipedia author behavior and content growth can be explained through a motif-based generative model.

The insights from the new Wikipedia model raise several possible future works. First, the edge types were driven by coarse-grained features aimed at capturing what was done to the page. However, the Wikipedia revision process lends itself to a variety of other classifications, such as by examining the revision comment or considering edit longevity. Further, the network representation could be expanded to include editor interaction on the Talk pages, which might reveal collaborative sequences such as Talk page discussion followed by article revision. As a second future work, we plan use our motif framework as a way to analyze other evolving collaborative systems, such as non-Wikimedia Wikis, such as Wikia and Conservapedia, which have very different editing policies and user bases. Last, we plan to analyze the causal relationships between motifs: do specific interactions change the probabilities for other interactions occurring later in time? This work will include both considering more sophisticated null models that take causality into account (Holme and Saramäki 2012) and investigating the evolutionary dynamics of the system as a whole.

References

Arazy, O., and Nov, O. 2010. Determinants of wikipedia quality: the roles of global and local contribution inequality. In *Proceedings of the 2010 ACM conference on Computer supported cooperative work*, 233–236. ACM.

Blei, D.; Ng, A.; and Jordan, M. 2003. Latent dirichlet allocation. *The Journal of Machine Learning Research* 3:993–1022.

Brandes, U.; Kenis, P.; Lerner, J.; and Van Raaij, D. 2009. Network analysis of collaboration structure in wikipedia. In *Proceedings of the 18th international conference on World Wide Web (WWW)*, 731–740. ACM.

Cobb, L. 2009. Study of vandalism survival times. http://en.wikipedia.org/wiki/Wikipedia:Wikipedia˙Signpost/2009-06-22/Vandalism.

Halfaker, A.; Kittur, A.; and Riedl, J. 2011. Don't bite the newbies: how reverts affect the quantity and quality of wikipedia work. In *Proceedings of the 7th International Symposium on Wikis and Open Collaboration*, 163–172. ACM.

Hall, M.; Frank, E.; Holmes, G.; Pfahringer, B.; Reutemann, P.; and Witten, I. 2009. The WEKA data mining software: an update. *ACM SIGKDD Explorations Newsletter* 11(1):10–18.

Holme, P., and Saramäki, J. 2012. Temporal networks. *Physical Reports. to appear*.

Kittur, A.; Suh, B.; Pendleton, B.; and Chi, E. 2007. He says, she says: Conflict and coordination in wikipedia. In *Proceedings of the SIGCHI conference on Human factors in computing systems*, 453–462. ACM.

Kovanen, L.; Karsai, M.; Kaski, K.; Kertsz, J.; and Saramki, J. 2011. Temporal motifs in time-dependent networks. *Journal of Statistical Mechanics: Theory and Experiment* (11):P11005.

Laniado, D., and Tasso, R. 2011. Co-authorship 2.0: Patterns of collaboration in wikipedia. In *Proceedings of the 22nd ACM conference on Hypertext and hypermedia*, 201–210. ACM.

Laniado, D.; Tasso, R.; Volkovich, Y.; and Kaltenbrunner, A. 2011. When the wikipedians talk: network and tree structure of wikipedia discussion pages. In *Proceedings of the International AAAI Conference on Weblogs and Social Media*.

McCallum, A. K. 2002. MALLET: A Machine Learning for Language Toolkit. http://mallet.cs.umass.edu.

Milo, R.; Shen-Orr, S.; Itzkovitz, S.; Kashtan, N.; Chklovskii, D.; and Alon, U. 2002. Network motifs: simple building blocks of complex networks. *Science* 298(5594):824.

Newman, M. 2005. Power laws, pareto distributions and zipf's law. *Contemporary physics* 46(5):323–351.

Roth, C.; Taraborelli, D.; and Gilbert, N. 2008. Measuring wiki viability: An empirical assessment of the social dynamics of a large sample of wikis. In *Proceedings of the 4th International Symposium on Wikis*, 27. ACM.

Suh, B.; Convertino, G.; Chi, E.; and Pirolli, P. 2009. The singularity is not near: slowing growth of wikipedia. In *Proceedings of the 5th International Symposium on Wikis and Open Collaboration*, 8. ACM.

Sumi, R.; Yasseri, T.; Rung, A.; Kornai, A.; and Kertész, J. 2011. Edit wars in wikipedia. In *IEEE Third International Confernece on Social Computing*.

Viégas, F.; Wattenberg, M.; and Dave, K. 2004. Studying cooperation and conflict between authors with history flow visualizations. In *Proceedings of the SIGCHI conference on Human factors in computing systems*, 575–582. ACM.

Vuong, B.; Lim, E.; Sun, A.; Le, M.; and Lauw, H. 2008. On ranking controversies in wikipedia: models and evaluation. In *Proceedings of the international conference on Web search and web data mining*, 171–182. ACM.

Wu, G.; Harrigan, M.; and Cunningham, P. 2011. Characterizing wikipedia pages using edit network motif profiles. In *Proceedings of the 3rd international workshop on Search and mining user-generated contents*, 45–52. ACM.