

Radial History Flow: Direct Visualization of Author Dynamics in Wikipedia

Taeil Jeon, Jihyun Lee, Wonho Lee, Wonjong Rhee, and Bongwon Suh

Seoul National University, Seoul, Korea

Abstract

In this paper, we propose Radial History Flow, which enhances History Flow by enabling comparison of author dynamics shown in the revision histories of Wikipedia articles. The key objective of this study is to unravel the complex relationships among the authors by linking two authors when one has a history of revising the other's contribution on the article. Using the revision history dataset of the Wikipedia article on 'electric energy', we perform a case study and demonstrate the feasibility of Radial History Flow.

Categories and Subject Descriptors (according to ACM CCS): H.5.2 [Information Interfaces and Presentation (e.g., HCI)]: General—

1. Introduction

Through information technology, Crowd Capital perspective has introduced a new type of human interaction that accumulates marginal efforts of the internet community into an united and easily accessible commodity [PS13]. Wikipedia is one of the well-known Crowd Capitals, and it has gathered public knowledge into a single and well-structured encyclopedia [LC01]. Although Wikipedia is known to be mostly reliable, the contents built by an anonymous crowd has the fundamental issue of possible conflicts among the contributors. Furthermore, it is often difficult to track the dynamics of revision histories that result in the current Wikipedia page.

History Flow [VWD04] attempts to alleviate this problem by visualizing the revisions of a Wikipedia article by laying out the revision records in a chronological order. History Flow allows tracking of overall collaboration patterns of the edits, and highlights prominent user behaviours that are difficult to detect otherwise. When the collaboration dynamics become sufficiently complicated, however, it is difficult to identify who has revised which part of the article or to determine whose contribution has been revised.

In this paper, we aim to enhance History Flow to allow for a further comparison of contributor dynamics and Wikipedia edit histories. We introduce Radial History Flow which uses the mechanism of Friend Wheel [HSS10] that shows the relationship between different objects by drawing a link between the related nodes in a circular layout. The key objec-

tive of the enhancement is to unravel the complex relationships among the authors by linking two authors when one has a history of revising the other's contribution on the article.

2. Data Source

We use the dataset of the Wikipedia article on 'electric energy' collected from 2012 to 2014 [EE114]. Over the three year period, 150 major revisions had been performed. We use this log data to visualize the general flow of revision history, where we lay out each version sequentially.

3. Radial History Flow

3.1. Modifications

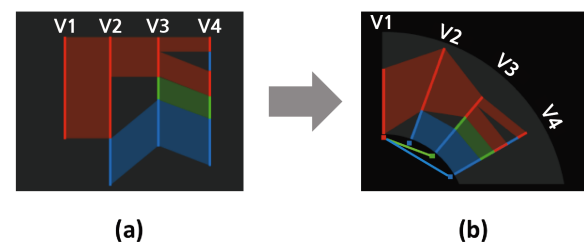


Figure 1: (a) History Flow mechanism (b) Radial History Flow mechanism

Figure 1 (a) shows the well-known History Flow and Figure 1 (b) shows the proposed Radial History Flow. Radial History Flow displays how the article has been edited in a circular, instead of a linear, layout. The vertical lines in History Flow become the radial boundaries in Radial History Flow. Each radial boundary is then used to represent an individual version of the article. Every author is assigned with a unique colour, and a dot positioned below each radial boundary is chosen to have the colour of the author who invoked that version of the article. For example, the colour of the dot underneath the version 4 radial boundary is blue, indicating that the author assigned with the colour blue edited the article of the version 3.

3.2. Main Feature

Radial History Flow enables the user to explore the connections among the contributors. The user can hover the mouse on the flow segment created by an author of the interest, and investigate the cooperating or conflicting revisions relevant to the author. For instance, the lines A1 and A2 in Figure 2 indicate that the contributor of version 22 edited the revisions made in version 14 and 16. Similarly, the line B in Figure 2 shows that the contribution of version 22 was edited by the contributor of the version 24.

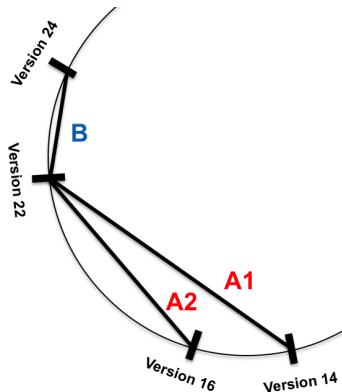


Figure 2: The two lines A1 and A2 indicate the revision histories of an author on other authors' works. The single line B indicates the revision histories of another author on to the contributions made by the author of interest.

3.3. Additional Features

Radial History Flow provides a variety of features to facilitate the exploration of revision history. First, the user can adjust the scale of the visualization to control the heights of the flow segments and examine the details. Radial History Flow also allows the user to reset the starting point and the ending point of the History Flow. Since these two versions are positioned adjacent to each other, the user can directly compare the two versions of the article. These two versions

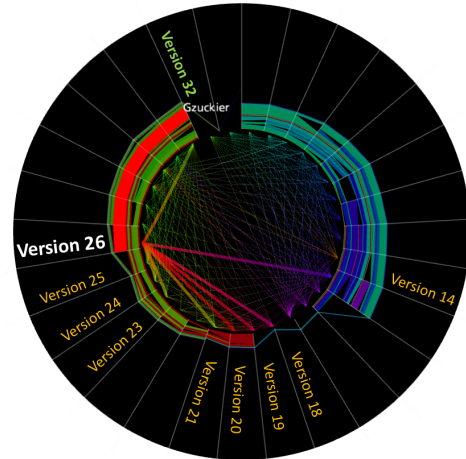


Figure 3: Out of 150 versions of Wikipedia article on 'electric energy', we have selected version 5 as the starting point and version 34 as the ending point.

are positioned at the 12 o'clock position in Figure 3, but in general they can be moved to any other position.

4. Case Study

To examine the feasibility of Radial History Flow, we performed a case study on the revision dynamics of the Wikipedia article on 'electric energy'. Out of 150 versions of Wikipedia article on 'electric energy', we have selected version 5 as the starting point and version 34 as the ending point.

As shown in Figure 3, the part written by Gzuckier is highlighted in red colour and a number of lines are drawn to indicate how the person's works are related to other authors' works. Using the conjugated links in the layout, we can recognize that Gzuckier started participating from version 26 and edited the works of version 14, 18, 19, 20, 21, 23, 24, and 25. Unlike all the other links connected to the version 26, the one particular link that bridges version 26 and version 32 indicates that Gzuckier's contribution in version 26 was edited by another author in version 32.

5. Conclusion and Future Work

In this paper, we propose Radial History Flow, that uses the revision history dataset of Wikipedia. Radial History Flow draws a link between the two dots that represent the two editing authors to indicate that one author has edited the other author's work. Therefore, it enables the user to unveil the complex relationships among the authors, clarifying who has edited whose portion of the article. Our next steps will be to further generalize Radial History Flow in such a way that it can be applied to a wider range of edit history datasets.

References

- [EE114] Electric energy: Revision history, 2014. URL: https://en.wikipedia.org/w/index.php?title=Electric_energy&action=history. 1
- [HSS10] HANSEN D., SHNEIDERMAN B., SMITH M. A.: *Analyzing social media networks with NodeXL: Insights from a connected world*. Morgan Kaufmann, 2010. 1
- [LC01] LEUF B., CUNNINGHAM W.: *The Wiki way: quick collaboration on the Web*, vol. 9. Addison-Wesley Boston, 2001. 1
- [PS13] PRPIC J., SHUKLA P.: The theory of crowd capital. In *System Sciences (HICSS), 2013 46th Hawaii International Conference on* (2013), IEEE, pp. 3505–3514. 1
- [VWD04] VIÉGAS F. B., WATTENBERG M., DAVE K.: Studying cooperation and conflict between authors with history flow visualizations. In *Proceedings of the SIGCHI conference on Human factors in computing systems* (2004), ACM, pp. 575–582. 1