ANALYZING HOW EDITORS WRITE ARTICLES IN WIKIPEDIA

S. Itoh[a] , **Y. Yamazaki**[b] and **T. Iba**[c]

[a]Graduate School of Media and Governance, Keio University
5322 Endo, Fujisawa-shi, Kanagawa 252-8520, Japan
si@sfc.keio.ac.jp

[b]Graduate School of Media and Governance, Keio University
5322 Endo, Fujisawa-shi, Kanagawa 252-8520, Japan
yuca@sfc.keio.ac.jp

[c]Faculty of Policy Management, Keio University
5322 Endo, Fujisawa-shi, Kanagawa 252-8520, Japan
iba@sfc.keio.ac.jp

In this research, we analyzed how editors write articles in Japanese Wikipedia. To be more precise, we explored diversity of writer's activity in Wikipedia.

Wikipedia is an encyclopedia on the web. The official version of Japanese Wikipedia was released on September 1st 2002 and is becoming bigger and bigger in numbers of articles and people who write and edit them. Now Japanese Wikipedia has over 500 thousands articles. In Wikipedia, everyone can edit articles they like. In other words, a lot of people, who don't know each other, work together on the Wikipedia platform and co-write articles. We collected history data of all Japanese Wikipedia articles and analyzed them.

We analyzed the collected data in the following viewpoints: (1) numbers of people participate in editing an article, (2) processes of growing articles.

First, to analyze numbers of people participate in editing an article, we plotted annual changes in frequency distributions of editors' number. Then we can see a huge bias, for example, 456336 articles (88% of all articles edited in 2007) are edited by under 10 people while only one article are edited by over 1000 people in 2007. Fig.1 is cumulative distribution graph about number of people editing each article. We found that the structures of frequency distribution are getting stable after 2004.
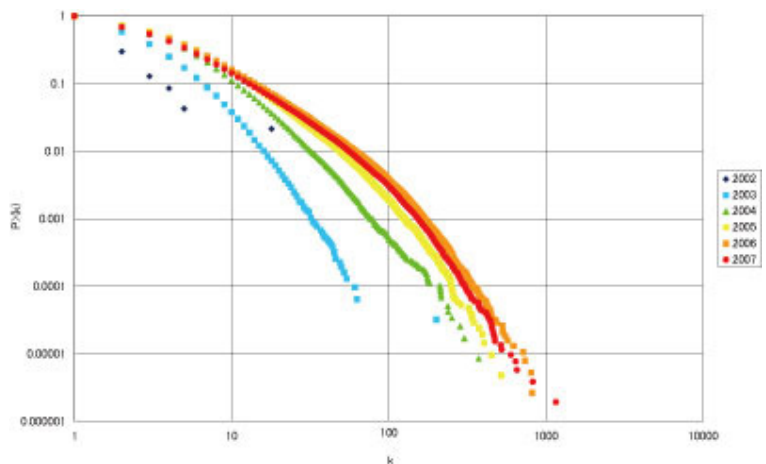

Fig 1: cumulative distribution graph about number of people editing each article

Second, to analyze processes of growing articles, we chose some articles and visualize changes of numbers of bytes in each article and network of editors. Then we can see some characteristic patterns in process of growing articles. For example, Fig.2 is a graph which means a change of number of bytes in "Chromosome theory ", one of featured articles in Japanese Wikipedia. In this article, only one person writes most of article at the start and other people protect the article.

Processes of growing articles are more various than we expected.

Keywords
Wikipedia, time series analysis

References
[1] Y.Yamazaki, T.Ito, T.Iba, and K.Kumasaka, "Growing Patterns of Wikipedia's Networks: Analyzing Laws and Diversity in Japanese Wikipedia", *poster, International Workshop and Conference on Network Science '08*, Norwich, UK, June 2008
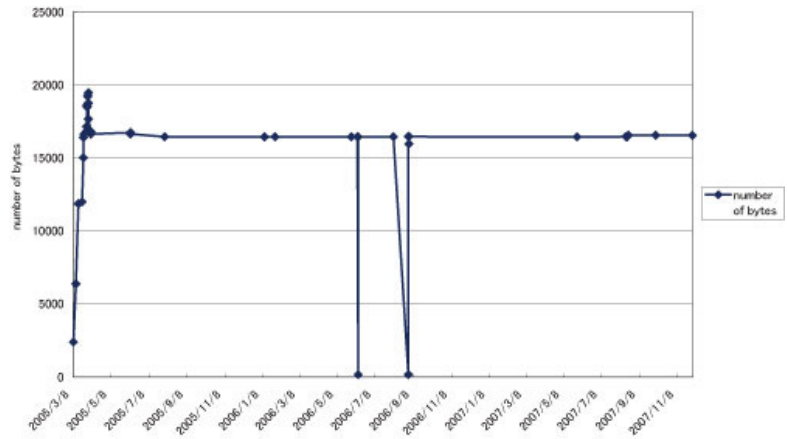[2] http://ja.wikipedia.org/

Fig 2: a change of number of bytes in "Chromosome theory "