

## Analyzing Information Flow and Context for Facebook Fan Pages\*

Kwanho KIM<sup>†</sup>, Josué OBREGON<sup>††</sup>, Nonmembers, and Jae-Yoon JUNG<sup>††a)</sup>, Member

**SUMMARY** As the recent growth of online social network services such as Facebook and Twitter, people are able to easily share information with each other by writing posts or commenting for another's posts. In this paper, we firstly suggest a method of discovering information flows of posts on Facebook and their underlying contexts by incorporating process mining and text mining techniques. Based on comments collected from Facebook, the experiment results illustrate how the proposed method can be applied to analyze information flows and contexts of posts on social network services.

**key words:** information flow, context analysis, process mining, text mining, Facebook

## 1. Introduction

Recently, social network services such as Facebook and Twitter are increasingly growing as an online platform where individuals and organizations post and share information such as news, opinions and advertisements in a fast and easy manner. The original posts are notified to the subscribers of the sites, and the comments on the post let the information be delivered to the friends of the subscriber again. Since comments on a post reflect how information is actually propagated from and to people, discovering hidden information flows and their underlying contexts from commenting behaviors of people is a significant issue of information diffusion analysis on social networks. Sophisticated understanding about information propagation mechanisms that people actually possess can be achieved by information flows and contexts, which accelerates many social services such as viral marketing [1], social suggestion [2], and influence analysis [3].

Specifically, this research focuses on *Facebook fan pages* (FFPs) which are to connect people with the same interest. A FFP includes many posts each of which is composed of a textual description and comments, as shown in Fig. 1. The description of a post represents its specific *context* while its comments contain the *information flow* of the context based on people's interactions in chronological order. Here, the sequence of ordered comments on a specific

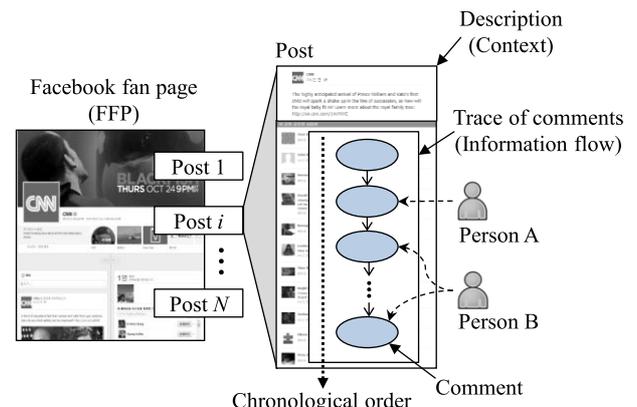


Fig. 1 Information flow and context of a post on a Facebook fan page.

post is called the *trace* of comments on the post.

There are many studies on information flows based on social interactions such as calculation of influences [4], analysis on social network structures [5], [6], and discovery of information diffusion processes [7]. However, their major drawback is the lack of ability to suggest the corresponding context with an information flow. Moreover, the previous methods cannot capture the dynamic nature of information flows since they heavily rely on the static links between people, called friendship, rather than focusing on the dynamic information propagation established by people for a context.

In this paper, we present an integrated method of analyzing the context from comments on a FFP, as well as discovering meaningful information flows from the FFP. While the previous method deals with only information flow discovery from re-posting behaviors of people [8], the integrated method proposed in this paper first suggests application of trace clustering to finding several meaningful information flows and then reveals the underlying context of each flow based on frequent keywords. In summary, by using the people's comments on posts, the proposed method can discover information flows that imply the actual relations among people, as well as their underlying contexts simultaneously. Specifically, process mining and text mining techniques are incorporated to achieve the goal of our method.

The experiment results based on comments collected in a FFP demonstrate that our method is able to suggest a yet another viewpoint to understand information propagation among people through revealing hidden information flows and their contexts.

Manuscript received July 4, 2013.

Manuscript revised October 29, 2013.

<sup>†</sup>The author is with Incheon National University, Incheon, 406-772, Republic of Korea.

<sup>††</sup>The authors are with Kyung Hee University, Yongin, Gyeonggi-do, 446-701, Republic of Korea.

\*This research was supported by the National Research Foundation of Korea (NRF) grant funded by the Korea government (MISP) and the Ministry of Education (Nos. 2013R1A2A2A03014718 and 2012R1A6A3A01038452).

a) E-mail: jyjung@khu.ac.kr (Corresponding author)

DOI: 10.1587/transinf.E97.D.811





combining similar traces but also to analyze their underlying contexts. The experiment results illustrated how the proposed method could be applied to the understanding of the information propagation of posts on FFPs. It is believed that the method allows Web analysts to figure out who are interested in the posts on the site and how the information of the post are propagated on the social networks based on the comments on FFPs.

The proposed method can also be widely adopted for most social network services just as illustrated for Facebook. It is because the method depends only on information related to comments such as user sequences and textual descriptions which are provided by the social network services.

For the future work, we plan to further extend the proposed method to accommodate additional information of social networks such as friend relations and users' profiles.

#### References

- [1] P. Domingos, "Mining social networks for viral marketing," *IEEE Intell. Syst.*, vol.20, no.1, pp.80–82, Jan. 2005.
  - [2] Q. Li, J. Wang, Y.P. Chen, and Z. Lin, "User comments for news recommendation in forum-based social media," *Inf. Sci.*, vol.180, no.24, pp.4929–4939, Dec. 2010.
  - [3] M. Cha, H. Haddadi, F. Benevenuto, and K.P. Gummadi, "Measuring user influence on twitter: The million follower fallacy," *Proc. 4th Int. Conf. on Weblogs and Social Media*, pp.10–17, Washington, DC, May 2010.
  - [4] X. Jin, C. Wang, J. Luo, X. Yu, and J. Han, "LikeMiner: A system for mining the power of 'Like' in social media networks," *Proc. 17th Int. Conf. on Knowledge Discovery and Data Mining*, pp.753–756, San Diego, CA, Aug. 2011.
  - [5] W.M.P. van der Aalst, T. Weijters, and L. Maruster, "Workflow mining: Discovering process models from event logs," *IEEE Trans. Knowl. Data Eng.*, vol.16, no.9, pp.1128–1142, Sept. 2004.
  - [6] W.M.P. van der Aalst and M. Song, "Mining social networks: Uncovering interaction patterns in business processes," *Lect. Notes Comput. Sci.*, vol.3080, pp.244–260, June 2004.
  - [7] P. Sriprasertsuk and W. Kameyama, "Information distribution analysis based on human's behavior state model and the small-world network," *IEICE Trans. Inf. & Syst.*, vol.E92-D, no.4, pp.608–619, April 2009.
  - [8] K. Kim, J.-Y. Jung, and J. Park, "Discovery of information diffusion process in social networks," *IEICE Trans. Inf. & Syst.*, vol.E95-D, no.5, pp.1539–1542, May 2012.
  - [9] J.-Y. Jung, "PROCL: A process log clustering system," *J. Society for e-Business Studies*, vol.13, no.2, pp.181–194, May 2008.
  - [10] C. Günther and W.M.P. van der Aalst, "Fuzzy mining – adaptive process simplification based on multi-perspective metrics," *Lect. Notes Comput. Sci.*, vol.4714, pp.328–343, Sept. 2007.
  - [11] N. Zhong, L. Yuefeng, and W. Sheng-Tang, "Effective pattern discovery for text mining," *IEEE Trans. Knowl. Data Eng.*, vol.24, no.1, pp.30–44, Jan. 2012.
  - [12] W.M.P. van der Aalst, T. Weijters, A.J. Weijters, B.F. van Dongen, A.K. Medeiros, M. Song, and H.M.W. Verbeek, "Business process mining: an industrial application," *Inf. Syst.*, vol.32, no.5, pp.713–732, Sept. 2006.
-