Visualization of Dense Tweets based on Locations for Web and Twitter User Communication

Yuanyuan Wang
Nagoya University
yuanw@db.ss.is.nagoya-u.ac.jp

Yukiko Kawai
Kyoto Sangyo University
kawai@cc.kyoto-su.ac.jp

Toyokazu Akiyama
Kyoto Sangyo University
akiyama@cse.kyoto-su.ac.jp

Yoshiharu Ishikawa
Nagoya University
ishikawa@is.nagoya-u.ac.jp

Kazutoshi Sumiya
Kwansei Gakunin University
sumiya@kwansei.ac.jp

ABSTRACT

This paper presents a tweet visualization system to support simultaneous communication between web and Twitter users through both the contents of tweets and web pages based on locations. The system provides a question answering interface attached to web pages, which allows web users to chat with Twitter users while presenting tweets that are associated with web pages. In order to map tweets to web pages, the system matches location names retrieved from geo-tagged tweets and web pages. Furthermore, we detect floor information from tweets, and categorizes tweets based on category names of floor information from web pages. Thus, our system can effectively present most related tweets and their summary information to help web users easily gain more detailed current situation in real time, and it also can effectively present messages from web users to help Twitter users immediately obtain useful information or knowledge.

Keywords

geo-tagged tweets; Twitter; web pages; locations; cross-media communication

1. INTRODUCTION

The advent of Twitter has recently attracted attention and gained popularity of personal update sharing from all over the world. Twitter users can broadcast and share information about their activities, opinions and statuses in short posts, up to 140 characters, using smartphones at anytime and anywhere. Despite the useful information on Twitter, there still exists a lack of Twitter users' requirements. That is, tweet senders are difficult to obtain useful information (e.g., bus timetables, sightseeing maps) about their current locations. Meanwhile, previous works usually focus on detecting a wide range of events based on geographical areas or location mention. For example, GeSoDeck [2] detected geo-social events based on geographical pattern mining and content analysis. Another application of a probabilistic framework [1] for estimating a Twitter user's city-level location based purely on the content of the user's tweets. However, these works did not detect dense tweets by considering with floor or height information of landmarks (e.g., composite facilities), because locations include only latitude and longitude information. Actually, there are many small events such as crowded restaurants, shop sells, and seasonal events based on floors in composite facilities at urban areas; users are difficult to immediately obtain most recent information there from tweets, whilst they browse web pages of composite facilities since these pages are not updated in real time. Furthermore, those events depend on the time of day, and located users on other floors or facilities. Therefore, it is important to visualize tweets through web pages, to facilitate simultaneous communication for web and Twitter users.

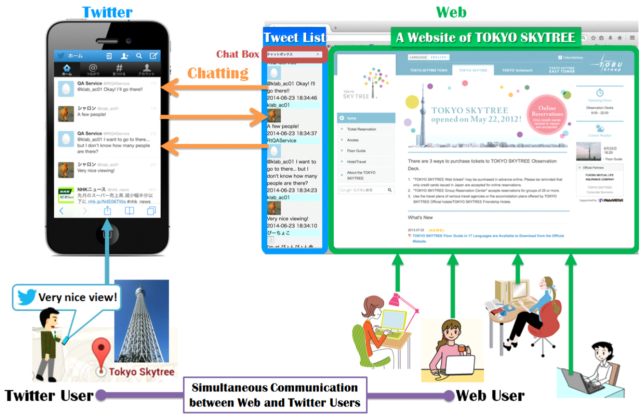

Figure 1: Web and Twitter user communication.

In this research, we have proposed a novel tweet visualization system to associate web pages with their most related dense tweets support for web and Twitter user communication. To achieve this, we first acquire geo-tagged tweets based on content analysis and region selection. Therefore, our method can detect tweets if they are related to, or nearby target locations, even though they do not include location names, or it can also detect tweets posted from Twitter users who are not in target locations. Furthermore, our method can filter out tweets from target locations for which the content is not related to target locations. The system then maps acquired tweets to web pages by matching location names detected from acquired tweets and web

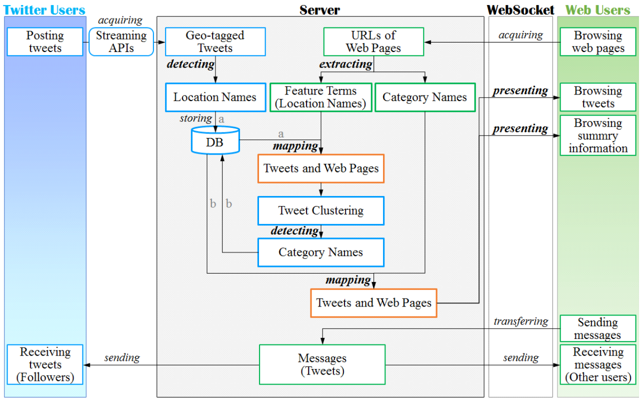

Figure 2: System configuration diagram.

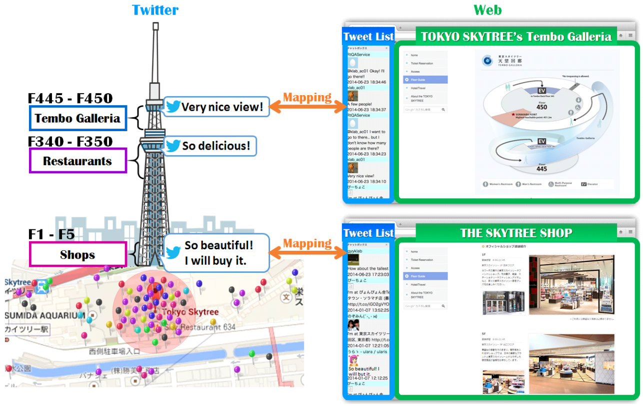

Figure 3: Tweets associated with web pages.

pages while users browse them; and categorizes acquired tweets based on category names of floor information from web pages. As depicted in Figure 3, our visualization system has two features: 1) mapping real time tweets to web pages based on both location and floor information; 2) attaching a chat box to web pages so that web users can communicate with Twitter users who follow an account of our system.

2. SYSTEM OVERVIEW

The processing flow of our system can be constructed as shown in Figure 2. To user our system, which is on the basis of existing Web services, users are first required to simply install a toolbar (a Firefox add-on) when they want to browse web pages, and Twitter users are first required to follow an account[1] as followers of our system. Once a user browses a web page, the system records the information into a server database, which is used for mapping tweets to the web page based on a location name detected from the tweets and the web page, and categorizing the tweets based on category names of floor information from the web page. The functions of our system are described as follows:

- A user selects a web page to browse, the system then returns a tweet list, in which most related real time tweets and their summary information with the web page are presented in Web browser.
- When the user sends a message, the system presents it in the tweet list, users who browse the same web page, or Twitter users who follow our system, can receive it.
- When a Twitter user replies the message of the user through Twitter service; the system presents the reply relating to the Web page in the tweet list in real time.

3. PROTOTYPE SYSTEM

The prototype system has three stages: analysis, mapping, and interface. In the analysis stage, we acquired geo-tagged tweets, without duplicates, from a specified region by using The Streaming APIs version 1.1. The specified region is defined as a rectangular region by a northeast point and a southwest point. In the mapping stage, we mapped acquired tweets and web pages based on their location names, and categorized the acquired tweets based on category names of floor information, when locations are composite facilities. During the interface stage, a browsing interface is developed into a Web browser by using an add-on for Firefox, and users are connected to our system through WebSocket to receive and send messages. Meanwhile, Twitter users can receive and send messages (tweets) through Twitter services.

An example is shown in Figure 3, which depicts a user browsing an official website of Tokyo Skytree in the Web browser of our system. Streaming tweets, e.g., "Very nice view!" located on Tembo Galleria Floor of the Tokyo Skytree, are associated with the web page of **Tembo Galleria** based on a location name, "Tokyo Skytree," and a category name of the floor level, "Tembo Galleria," even though the tweets do not include them; and streaming tweets, e.g., "So beautiful! I will buy it" located on Shop Floor of the Tokyo Skytree, are associated with the web page of **The Skytree Shop**. In the same manner, a tweet, "Very delicious!" located on Restaurant Floor of the Tokyo Skytree can be detected when the user browses the web page of **Sky Restaurant**. This allows the user to gain insight into the congestion level or gain impressions of each floor of the Tokyo Skytree from presented tweets, and he or she also easily knows where and when more people are in the Tokyo Skytree now.

4. DISCUSSION

Mapping of tweets and web pages. Our mapping method needs to consider space structures (planes and heights) from tweets in real world, and hierarchical link structures from web pages to associate traditional web pages with real time tweets by a vertical and horizontal analysis.

Visualization of dense tweets. It should be possible to visualize the summary information of tweets based on generic platforms with visualization tools to portray time perspectives. This should be useful for E-commerce to help users (e.g., customers) to understand the social data easily and efficiently as a store visualization system.

5. REFERENCES

[1] Z. Cheng, J. Caverlee, and K. Lee. You are where you tweet: a content-based approach to geo-locating twitter users. In *Proc. of CIKM 2010*, pages 759–768, 2010.

[2] X. Gao, J. Cao, Z. Jin, X. Li, and J. Li. Gesodeck: a geo-social event detection and tracking system. In *Proc. of MM 2013*, pages 471–472, 2013.

[1]https://Twitter.com/@RtQAService