Detection, Classification and Visualization of Place-triggerd Geotagged Tweets

Shinya Hiruta[1], **Takuro Yonezawa**[1], **Marko Jurmu**[1,2], **Hideyuki Tokuda**[1]

[1]Keio University
5322 Endo, Fujisawa, Kanagawa, Japan
{hiru, takuro, hxt}@ht.sfc.keio.ac.jp

[2]University of Oulu
Erkki Koiso-Kanttilan katu 3, 90570 Oulu, Finland
marko.jurmu@ee.oulu.fi

ABSTRACT

This paper proposes and evaluates a method to detect and classify tweets that are triggered by places where users locate. Recently, many related works address to detect real world events from social media such as Twitter. However, geotagged tweets often contain noise, which means tweets which are not content-wise related to users' location. This noise is problem for detecting real world events. To address and solve the problem, we define the *Place-Triggered Geotagged Tweet*, meaning tweets which have both geotag and content-based relation to users' location. We designed and implemented a keyword-based matching technique to detect and classify place-triggered geotagged tweets. We evaluated the performance of our method against a ground truth provided by 18 human classifiers, and achieved 82% accuracy. Additionally, we also present two example applications for visualizing place-triggered geotagged tweets.

General Terms

Design, Experimentation

Author Keywords

Microblogs, Location-based Services, Place-triggered Geotagged Tweets

ACM Classification Keywords

H.3.5 Web-based services.

INTRODUCTION

Real world event can be formally structured as a collection of descriptive attributes. Frequently, these attributes are dynamic in nature which means that static *a priori* descriptions cannot be used. Furthermore, a real world event can manifest itself without any a priori static description, in which case dynamic information is the only source for communicating information regarding the event. Example of former is a baseball game that gets postponed because of rain, and an example of the latter is a traffic accident occurring on a motorway and causing significant traffic congestion.

From this framing of real world events, we can conclude that systems which *extract, classify and provide real-time dynamic attributes of the event* are needed. In this paper, we focus on location as a key attribute of both aforementioned event types. This is because location is the most common denominator for a wide variety of events, and in many cases the single most important one. Especially in events describing accident or catastrophe information, location is the first attribute to be resolved.

In designing aforementioned systems, we need to consider the possible data sources and their suitability for this purpose, the actual extraction and classification methods, and finally the APIs that the system can offer for third party applications. Since we classify two types of real world events, those that have a static a priori description and those that don't, we can envision the API providing a subscription-based service for receiving dynamic attributes of an event, given a static unique identifier of the event. More challenging is the API for dynamically occurring events, which can rely on more advanced subscription mechanisms.

Considering the data sources, social networking services are suitable for extraction of dynamic real world data. In this paper, we especially focus on Twitter, due to its public and agile nature as a communication medium. Regarding Twitter, there are several strategies to employ in extraction and analysis. One strategy is to use metadata-only, for example the geotags encoded into tweets. This is however insufficient, since there are a lot of noisy tweets whose contents are not related to the location. Another strategy is to combine the metadata extraction with content-based analysis, meaning first to filter the tweets based on metadata, and subsequently analyze them based on their content. This allows us to identify precise location of tweets using only geotagged data, and to detect meaningful tweets which have a content-based relation to their location.

We define *Place-triggered Geotagged Tweets* as those tweets that have both the geotag metadata as well as content relevant to the associated location. In this paper, we present our first approach towards extraction, analysis and provision of dynamic event attributes by designing, implementing and evaluating a keyword-based classifier for place-triggered

UbiComp '12, Sep 5-Sep 8, 2012, Pittsburgh, USA.

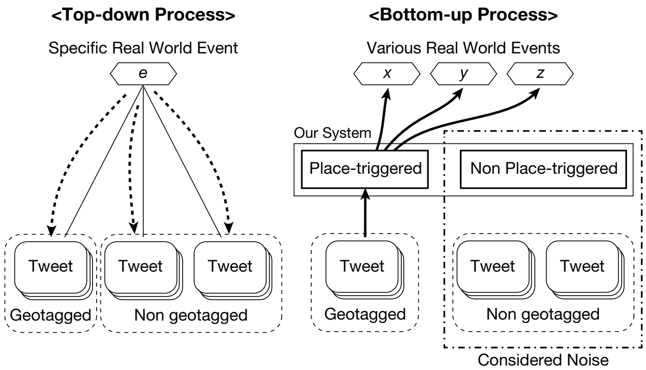

Figure 1. Comparison of approaches to detect events from tweets

geotagged tweets. First, in order to classify the type of place-triggered geotagged tweets, we surveyed geotagged tweets which were collected from Twitter. As a result, we classified the place-triggered geotagged tweets into 5 types: *report of whereabouts*, *food*, *weather*, *back at home* and *earthquake*. Based on these results, we designed filters to classify geotagged tweets into these 5 types by naive keyword matching method. We asked third parties of 18 people to create ground truth. Based on it, we evaluated relevance of our classification and determination accuracy of the filter.

The key contributions of this paper are:

- We survey and classify current usage of geotagged tweets.
- We present the design and implementation of a prototype system which detects and classifies *Place-triggered Geotagged Tweets*.
- We present results from an evaluation study that compares ground truth created by third parties of 18 people with our outcome filtered by the system.

DETECTING EVENTS FROM TWEETS

We define two ways to detect events from social media. One is the top-down process, which first specifies events and then filters the tweets for detection of these events. The other is the bottom-up process, where tweets converge into certain topics of interest, possibly by using certain ground truth as evidence. We take the latter approach, and introduce a new concept of *Place-triggered geotagged tweet*.

Methodology

Top-down Process

We define Top-down Process as a method which is intended for detecting specific events. The objective is to first detect certain events, and then construct a method for detection of these events. The left part of Figure 1 describes an approach of Top-down process. First, a specific event *e* is set, and then tweets are crawled for event detection.

This approach is suitable for events which have obvious characteristics that help to recognize the fact that the event has occurred. For example, forest fire events [3] and live news events [5] are being detected through this process.

Bottom-up Process

In contrast to the needs-oriented process, we define the bottom-up process that classifies the tweets as new point of view. There are two aspects revealing whether a tweet is related to real world events or not. One is the "geotag relation" which means whether a geotag is appended to or not. The other is "content-based relation", meaning whether contents of the tweet refer to the current location or not. These are described in Figure 1.

We define the geotagged tweets whose contents are motivated by events or situations of current locations as "Place-triggered Geotagged Tweets". We consider this bottom-up augmentation of Twitter data as important enabler to detect real world events by mining social media.

Objectives

Our goal is to investigate whether a keyword based classifier, aided by ground truth notion, can effectively classify and augment twitter data for the purpose of real world event detection. We also see this mechanism as a necessary pre-processing and de-noising step for systems mining data from social media.

Our research goal is two-fold: First aim is to *detect* place-triggered geotagged tweets, meaning discovering whether a tweet contains both of the abovementioned location relations. Second aim is to *classify* these place-triggered geotagged tweets by using a method of sequential filtering based on keywords and regular expressions. Our method is based on external classification knowledge, and we focus on measurement of accuracy as well as the relevance of used filters.

Related Work

Recently, many researchers are working for detecting events from social media. When detecting the real world events by leveraging social media, it is important to estimate the place from which the user tweeted.

Sakaki T, et al. detect occurrence of earthquakes in real time by analyzing Twitter stream [8]. In the research, they mainly make use of static locations which are set on user's profile in order to estimate the location where they tweeted. However, users are not necessarily tweeting on that location, because they can tweet from everywhere by their mobile devices. Therefore, inaccurate event locations would be estimated by analyzing inaccurate locations of tweets. In fact, they estimate location error an average of about 300 km from actual source of earthquakes.

Geotagged tweets have the potential to be utilized for real world event detection more accurately. Lee R, et al. detect local events by measuring geographical regularities of geotagged tweets [6]. However, geotagged tweets contain many noises which affect analysis. The research shows low precision rate, in other words, the results contain many unexpected events.

On the other hand, Yin X, et al. propose a method to discover objective information from multiple conflicting data sources on the web [13]. In a similar fashion, Wang D, et al. estimate the objectiveness of information by using Bayesian and maximum-likelihood methods [10], [11]. Schlieder, C et al. addressed a central problem in the field of social reporting [9]. They proposed an approach to the quality problem that is based on the reciprocal confirmation of reports by other reports. These researches consider only the authenticity of information source.

As illustrated above, when detecting real world events, noise in geotagged tweets affects the accuracy of results negatively. Many previous works take Top-down process for detecting events, but there are also some works proposing the bottom-up approach. Becker H, et al. focus on distinguishing between messages about real world events and non-event messages from the stream of Twitter messages [2]. Muhammad et al. propose a method for automatic tagging of untagged tweets [1]. Both of them use machine learning method to realize classification from large training set of tweets. Similar to these approaches, our research takes the bottom-up process. We attempt to verify whether the geotagged tweets are motivated by events or situations of current locations.

ESTABLISHING THE GROUND TRUTHS

This section describes the method of preliminary survey to classify the geotagged tweets. Then, we classify the type of place-triggered geotagged tweets based on the survey.

Preliminary Survey

In order to investigate how users tweet with geotag, we conducted a survey to classify types of geotagged tweets based on the content of tweet. For this objective, we crawled geotagged tweets in Twitter around Japan from 2011-11-21 00:00:00 to 2011-12-31 23:59:59. We sampled 2,000 tweets from the total amount of 1,977,531, then classified these tweets to certain types based on their content.

The result of classification is shown in Figure 2. We classified 11 place-triggered types based on the content of tweets in total. Most of the tweets (42.5%) were classified as noise, for example, just replying to other follower as “@someone Good morning !”. Rest of the tweets were classified as place-triggered tweets. In place-triggered tweets, report of whereabouts type accounted for 74.7%, and the other types were under 8%, each. Report of whereabouts is the content such as “I’m at Keio University now”.

Classification of the Place-triggered Geotagged Tweets

From the preliminary survey, many of the place-triggered tweets were confirmed as report of whereabouts. The second most popular type was food, followed by weather information and back at home notifications. The other types were less than 1% of all tweets, so we decided to take account of these four types. Furthermore, we chose the earthquake type in particular, since earthquakes can potentially cause great disaster and thus need to be monitored. Eventually, we classified place-triggered geotagged tweets to five types below:

- Report of whereabouts: A tweet that user refers to his/her current location.

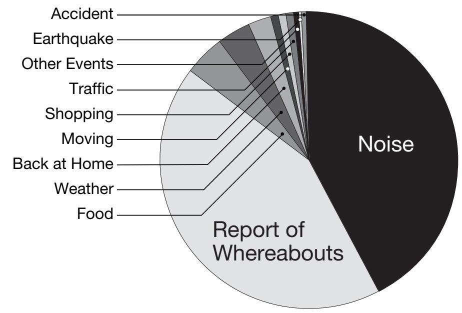

Figure 2. Result of survey of geotagged tweets

- Food: A tweet where user shares information regarding current food or drink.
- Weather: A tweet about weather of the location.
- Back at home: A tweet where user reports the fact that he/she is back at home.
- Earthquake: A tweet in which user reports the feeling of the earthquake.

DETECTION OF PLACE-TRIGGERD GEOTAGGED TWEETS

In this section, we show the approach to detect the place-triggered geotagged tweets. Then, we explain the design and implementation of the proposed system.

Approach

We design filter modules that detect each selected type of place-triggered geotagged tweets. Each filter uses naive keyword matching method to detect the type of tweets, because we assume that people tend to classify tweets mainly by distinctive keywords.

Approaches of each filter module are described below. First, we consider that the tweets from check-in services are the report of whereabouts. From the result of preliminary survey, most of the tweets which classified as the report of whereabouts were made using the check-in services. Many users link their check-in service account to Twitter’s, so that we can acquire check-in activity from crawling Twitter. The filter estimates tweets to be a report of whereabouts, if the *source* information of tweets contain Foursquare, Loctouch [7], or Imakoko-now [4] site URL.

The other filters: *food*, *weather*, *back at home* and *earthquake* use a keyword matching method. As an example of keyword matching, we describe the *food* filter. First, we created a list of synonyms of the keyword “food” using Weblio english thesaurus [12]. If text of a tweet contains words in the synonyms list of food, the tweet is considered to be *food* type. In the same way, we make lists of synonyms of the other keywords. For the first step, we prepare synonyms of verb, noun and adjective for each filter.

The results of classification are returned after the tweets are filtered by above filters. Each filter is independent of each other, one tweet may be classified as more than one type.

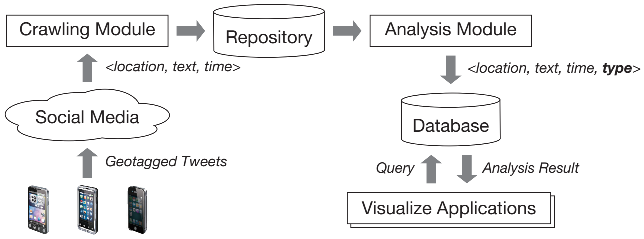

Figure 3. Module configuration

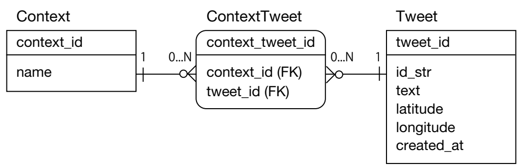

Figure 4. Database structures

As a method of keyword matching, filters use regular expressions. We also examined the possibility of morphological analysis, but decided to apply simple keyword matching method to avoid performance degradation due to object analysis. By increasing the pattern of keywords in the list of synonyms, we have confirmed the accuracy of determination equal to or better than if we use the morphological analysis. Visualization applications are intended to assist in the detection of real world events with place-triggered geotagged tweets. We describe details of implemented applications in Section .

Design and Implementation

We design the system that analyzes and visualizes crawled geotagged tweets. The proposed system is composed of three modules: crawling module, tweets analysis module, and visualization applications. Figure 3 shows the module architecture.

The crawling module collects tweets whose location information is set nearby Japan. Tweets are crawled from Twitter Streaming API in real time. We acquire about 50,000 to 70,000 geotagged tweets per day. In tweets analysis module, tweets are classified into different place-triggered types by using the method described in Section . The classified tweets are saved to the database. Database structure is shown at Figure 4. A tweet can have multiple place-triggered types that are described as Context and ContextTweet tables in the figure. The tweet record has five fields, id_str is unique ID provided by Twitter api, text is contents of tweet, latitude and longitude is location coordnate of the tweet and created_at is the timestamp when user tweeted. The visualization applications are intended to assist in the detection of real world events with place-triggered geotagged tweets. We describe detail of the applications at Section . The system is implemented using Ruby, PHP and MySQL.

API

We define an application programming interface to access the analyzed tweet data stored in the database. The API is currently used for two visualization applications described in Section . Applications invoke the API through an HTTP request, followed by server response containing the result as a JSON dataset. The result contains the tweet list as well as classification information of each included tweet.

We show detail of the API below. Resource URL is:
`http://example.com/api/getTweet`
The abovementioned domain name is tentative, and should be replaced in application implementations. Applications use HTTP GET method to acquire tweets. There are 3 parameters to specify query conditions.

- timestamp
  Returns tweets generated after the given date and duration. Date should be formatted as:
  YYYYMMDDhhmmss:(durationSec)
  Example values: 20120905120000:600
- bounds
  Returns tweets located within a given latitude/longitude pairs of NorthEast and SouthWest rectangle. The parameter is specified by:
  NElatitude,NElongitude,SWlatitude,SWlongitude.
  Example values:
  35.604094,139.585396,35.753852,139.844261
- context (optional)
  Returns tweets which are filterd by the given place-triggered types. Each place-triggered type is identified by integer, which is set in advance by the system. The parameter is specified by comma splitted integer. If more than one ID is specified, filter works through the OR grouping condition. Example values: 1,2,3

Example request is shown below:
`http://example.com/api/getTweet?context=&timestamp=20120905120000:600&bounds=35.604094,139.585396,35.753852,139.844261`
The API returns tweets from 2012-09-05 12:00:00 to 2012-09-05 12:10:00 tweeted on location of the given coordinates:

```
[
    {
        "context": [
            "checkin"
        ],
        "created_at": "Tue Sep 05 12:00:10 +0000 2012",
        "id": 123456789012345678,
        "id_str": "123456789012345678",
        "source":
        "<a href="http://foursquare.com">foursquare</a>",
        "text": "I'm at Tokyo Sta. http://example.com/foo"
    }
]
```

Each hash represents a tweet and multiple tweets can be included in a array. In a hash, the "context" attribute represents place-triggered type, and others are raw tweet data acquired from Twitter API.

VISUALIZATIONS

In this section, we describe details of two application prototypes using the place-triggered geotagged tweets. First, we mention the animation visualizer which is intended to discover the real world events. Then, we describe the web based interactive interface which is more focused on interactive browsing of analyzed tweets.

Animation Visualizer

We implemented the animation visualizer which plots the place-triggered geotagged tweets on a map. Any range of regions can be visualized, animating images in any time interval. Figure 5 shows the plot of tweet at intervals of 4 hours in April 29, 2012. Column A represents all the geotagged tweets, column B separates place-triggered geotagged tweets, which is shown by o sign, from noise. Column C shows the each type of place-triggered geotagged tweets. Each type of plot shows the report of whereabouts, food, weather, back at home, earthquake and noise.

We discovered some interesting cases by visualizing place-triggered geotagged tweets. On column A, it was detected that cluster surrounded by red circle, A-1 to A-5, is moving northeast at a constant rate. It cannot be estimated whether it is meaningful or not by column A, but it can be determined as noise by column B. It was found to be a Bot that follows a user who tweets with location information. The tweets are not related to the location, but generated randomly. Furthermore, we found that the tweets about the earthquake emerged rapidly in the plot C-1. In fact, a big earthquake occurred at May 28, 2012, 19:28. As explained above, by performing visualization plots for place-triggered geotagged tweets, we have been able to remove noise and detect real world events more easily.

Web-based Interactive Interface

We also implemented the other application which plots place-triggered geotagged tweets on Google Maps. Figure 6 shows a screenshot of the visualization interface. Place-triggered geotagged tweets are plotted on the map, colored by types. By clicking a pin, details of the tweet like texts and pictures can be displayed. Right part of the display shows control panels. You can adjust time windows by sliders, and the tweets are retrieved from the database each time you scroll the map, giving a sense of interactivity. The interface helps users to detect more precise events by displaying detail of each tweets. For example, by enabling filtering by food and picture, popular restaurants often remarked by users can be discovered, and by enabling filtering by weather and picture, guerrilla heavy rain can be visualized.

There are limitations that tweets can only be plotted up to 200 at the same time, from the problem on the performance of the API on Google Maps. Our de-noising method can make it possible to plot more important tweets. We aim to launch this service as a web application, which helps to overcome these limitations.

EVALUATION

In this section, we describe the detail of evaluation study, which was conducted from two perspectives. First, in order to verify whether the classification of place-triggered geotagged tweets is reasonable, we asked 18 people to classify geotagged tweets crawled from Twitter. Next, the results are

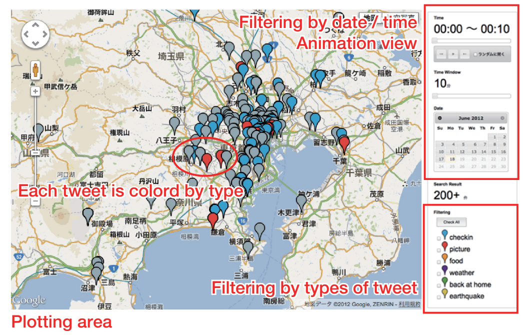

Figure 6. Interactive visualization interface for place-triggered geotagged tweets

compared with the implemented analysis module, in order to evaluate the performance of the module.

Classification of Geotagged Tweets by Third Parties

Methodology

In order to evaluate whether the preliminary classification is reasonable, we asked 18 third party people to classify tweets. Participants were 12 men in their 20s, 5 women in their 20s and 2 men in their 30s. We used different tweets pool from the preliminary classification described at Section . The evaluation targets are geotagged tweet which were crawled from 2012-01-01 00:00:00 to 2012-03-31 23:59:59, total amount is 4,254,257. Each participant reviewed 500 tweets which were randomly sampled from the dataset.

For reduction of effort to review many tweets, we created a web based evaluation tool. Figure 7 shows screenshot of the tool. Before the conduction of evaluation study, we explained the concept of place-triggered geotagged tweets to the participants. The participants see details of each tweets, which contains text, tweeted time and location, and are asked whether the tweet can be classified to any given type. If the tweet cannot be classified to these types, but it is clearly considered to be a place-triggered tweet, the *other* type is selected. On the other hand, if the tweet cannot be considered to be a place-triggered tweet, no types are selected. Multiple selection of types is also allowed.

Figure 7. Evaluation tool for tweet classification

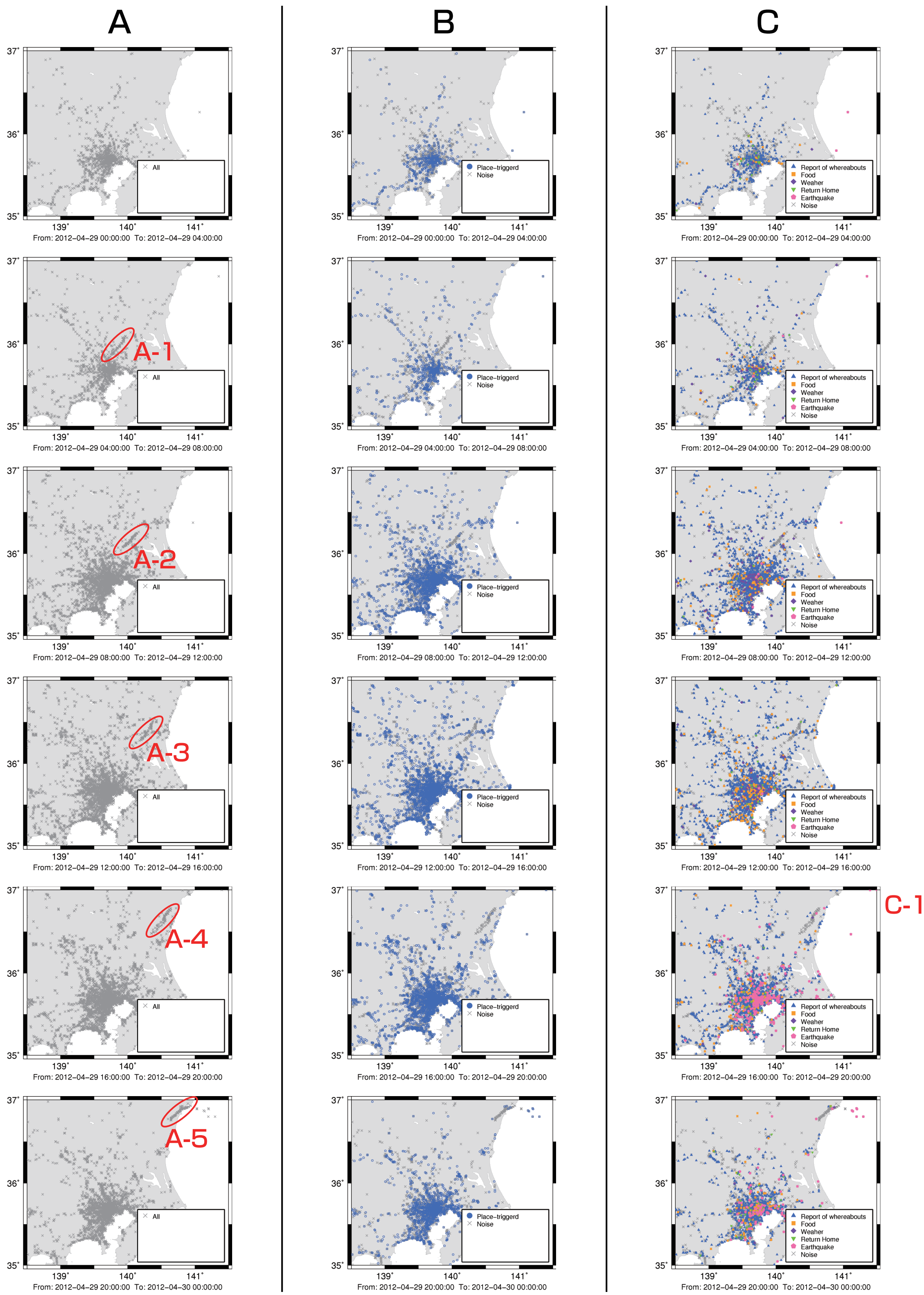

Figure 5. Plotted result of place-triggered geotagged tweets

Table 1. **Result of place-triggered geotagged tweets classification by third parties**

Type of Tweets	Number	Percentage
Report of whereabouts	4,300	76 %
Food	664	12 %
Weather	255	4.5 %
Back at home	61	1.1 %
Earthquake	44	0.77 %
Other	354	6.2 %
Total	5,678	100 %

Results and Discussions

We collected 8,988 tweets for the ground truths in this evaluation study. Twelve results are missed due to an application or network error. The breakdown of the result is as follows: 5,027 tweets are place-triggered (55.93%) , 3,961 tweets are non place-triggered (44.07%). Further breakdown of the place-triggered types is described in Table 1. Total amount of classification is 5,678, since multiple tagging is allowed.

We consider that the five types of preliminary classifications are mostly appropriate, since most of tweets were classified to these types by the third party people. However, 354 tweets were classified to the other type, which was place-triggered but cannot be classified to the preliminary types, so we decided to investigate these results more closely. For example, "Going out to buy something ..." and "I'm attending the event of ... ." are classified to place-triggered tweets by third parties. Unfortunately, since the standard of classification differs very much by each person, it was difficult to find any tendency within the type *other* to classify any further types.

Performance of the Tweets Classification Module

Methodology

To evaluate the performance of our tweet analysis module, we compared the module output with the ground truth collected in the previous subsection. We used the same tweets pools of that was used in the classification by third parties.

The evaluation is conducted in two aspects. First we examined the accuracy rate, that whether the system correctly detected place-triggered tweets or not. Second, we inspected the accuracy of each place-triggered type. For each type, we define the number of total tweets that is classified as the type by participants as C, number of tweets that is detected as the type by the system as N, and number of correct answer which the system detected as R. Then, we calculated the following measures: $Precision = \frac{R}{N}$, $Recall = \frac{R}{C}$ and $F\text{-}measure = \frac{2 \cdot Precision \cdot Recall}{Precision + Recall}$

Results and Discussions

First, we show the overall accuracy of the module. The false-positive and true-negative rate is shown in Table 2. As we show in Table 2, the place-triggered geotagged tweets could be detected with an accuracy of 82%. The false-positive rate is relatively low at 2.18%, meaning the system can get rid of most noise. Whereas, the false-negative rate is still 15.84%, meaning that significant amount of geotagged tweets are missed.

Table 2. Accuracy rate of detecting place-triggered geotagged tweets

	Positive	Negative
True	40.09 %	15.84 %
False	2.18 %	41.89 %

Table 3. Result of place-triggered geotagged tweets classification by system

Type of Tweets	Precision	Recall	F-measure
Report of whereabouts	93.18 %	77.16 %	84.42 %
Food	53.6 %	17.8 %	26.7 %
Weather	57 %	21 %	30 %
Back at Home	54 %	23 %	32 %
Earthquake	76 %	66 %	71 %

Second, we describe accuracies of each place-triggered type. The breakdown of detection accuracy in each type of place-triggered tweets is shown in Table 3.

- Report of Whereabouts
  The report of whereabouts has been detected with high accuracy, as the F-measure shows 84.42%. There are still some missed tweets, because there are tweets reporting whereabouts without using check-in service, thus they cannot be detected by the system for now.
- Food, Weather and Back at Home
  On the other hand, food, weather and back at home types could not be detected well. We consider that there are many tweets which mention the types of food, weather and back at home, without containing words within our dictionary. To improve recall, we should apply not only simple keyword matching method, but more detail method such as language analysis.
- Earthquake
  Earthquake type is detected with a relatively high accuracy. We conceive that keyword matching method is suitable for earthquake, since the report of earthquake occasion tend to be short words.

OPEN ISSUES AND FUTURE WORK

We introduced a new concept of place-triggered geotagged tweets, and presented the result of classification accuracy by using simple keyword matching algorithm. However, several limitations and room for improvement still remain. Here, we briefly discuss these remaining issues and provide implications for future research.

Expanding the classification

Although we classified place-triggered geotagged tweets into five categories based on their relative frequencies, other types of tweets that fulfill the criteria of place-triggered geotagged tweets still remain (see Figure 2). Additionally, since we only focused on Japanese tweets for defining the categories, it is necessary to investigate tweets in other countries. More complete categories of place-triggered geotagged tweets can be useful for future systems aiming to reliably detect real world events.

Improving detection accuracy

We leveraged a simple keyword-matching algorithm to detect place-triggered geotagged tweets. Although we presented reasonable accuracy with the proposed method, more efficient detection methods should be investigated. One promising direction is to utilize results from linguistic analysis research (e.g., unsupervised classification, allowing the system to converge to a finite number of types with significant relative frequencies). Furthermore, as real world places and events are often associated with dedicated terminologies and language constructs, the use of slang should be analyzed.

Discovering real events

We presented two applications for visualizing and interacting with place-triggered geotagged tweets. We consider that users can discover real world events with more ease by analyzing place-triggered geotagged tweets, compared with analyzing all tweets including noise. However, more automatic or efficient methods, such as temporal-spatial analysis of place-triggered geotagged tweets, should be investigated.

Security & spoofing location

When treating location information, we should consider whether the attached location is genuine or not. If a user spoofs his/her location, system might detect non-real events. Additionally, collaborative spoofing by groups of users can severely harm the detection process. In future work, we plan to take these types of attacks into account in the system design.

CONCLUSIONS

Detecting real world events from geotagged tweets whose location have no association with their actual contents significantly restricts overall performance. In this research, we defined *Place-triggered Geotagged Tweets*: Tweets containing both geotag and content-based relation to your location, and designed a system for their detection and classification. We classified the place-triggered geotagged tweets as 5 types: *report of whereabouts*, *food*, *weather*, *back at home* and *earthquake*, based on ground truth established by a survey as well as a study featuring 18 human classifiers. We conducted evaluation study and showed that the system can detect place-triggered geotagged tweets with an overall accuracy of 82%. Furthermore, we implemented visualization applications to detect real world events from place-triggered geotagged tweets. Our work contributes to current state-of-the-art through a pre-survey, design and implementation of a prototype system, as well as an evaluation against a ground truth notion.

ACKNOWLEDGMENTS

This research is partly supported by National Institute of Information and Communications Technology and Research Laboratories NTT Docomo Inc. Third author would also like to thank the Academy of Finland for financial support.

REFERENCES

1. M. Asif Hossain Khan, M. Iwai, and K. Sezaki. Towards urban phenomenon sensing by automatic tagging of tweets. In *Proceedings of the Ninth International Conference on Networked Sensing Systems*, 2011.
2. H. Becker, M. Naaman, and L. Gravano. Beyond trending topics: Real-world event identification on twitter. In *Proceedings of the Fifth International AAAI Conference on Weblogs and Social Media*, pages 438–441, 2011.
3. B. De Longueville, R. S. Smith, and G. Luraschi. ”omg, from here, i can see the flames!”: a use case of mining location based social networks to acquire spatio-temporal data on forest fires. In *Proceedings of the 2009 International Workshop on Location Based Social Networks*, pages 73–80, 2009.
4. fujita-lab.com. Imakoko-now. `http://imakoko-gps.appspot.com/`, 2012.
5. A. Jackoway, H. Samet, and J. Sankaranarayanan. Identification of live news events using twitter. In *Proceedings of the 3rd ACM SIGSPATIAL International Workshop on Location-Based Social Networks*, pages 25–32, 2011.
6. R. Lee and K. Sumiya. Measuring geographical regularities of crowd behaviors for twitter-based geo-social event detection. In *Proceedings of the 2nd ACM SIGSPATIAL International Workshop on Location Based Social Networks*, pages 1–10, 2010.
7. NHN Japan Corp. Loctouch. `http://tou.ch/`, 2012.
8. T. Sakaki, M. Okazaki, and Y. Matsuo. Earthquake shakes twitter users: Real-time event detection by social sensors. In *Proceedings of the 19th International Conference on World Wide Web*, pages 851–860, 2010.
9. C. Schlieder and O. Yanenko. Spatio-temporal proximity and social distance: a confirmation framework for social reporting. In *Proceedings of the 2nd ACM SIGSPATIAL International Workshop on Location Based Social Networks*, pages 60–67, 2010.
10. D. Wang, T. Abdelzaher, H. Ahmadi, J. Pasternack, D. Roth, M. Gupta, J. Han, O. Fatemieh, H. Le, and C. Aggarwal. On bayesian interpretation of fact-finding in information networks. In *Proceedings of the 14th International Conference on Information Fusion*, pages 1–8, 2011.
11. D. Wang, L. Kaplan, H. Le, and T. Abdelzaher. On truth discovery in social sensing: a maximum likelihood estimation approach. In *Proceedings of the 11th international conference on Information Processing in Sensor Networks*, pages 233–244, 2012.
12. Weblio, Inc. Weblio thesaurus. `http://thesaurus.weblio.jp/`, 2012.
13. X. Yin, J. Han, and P. Yu. Truth discovery with multiple conflicting information providers on the web. *IEEE Transactions on Knowledge and Data Engineering*, 20(6):796–808, June 2008.