

Density-based Spatiotemporal Analysis System with Photo Image Classifier Using the BoF Model

Tatsuhiro Sakai *, Keiichi Tamura *, Hajime Kitakami *

Abstract

Recently, people have begun to diligently post situational updates, particularly during natural disasters, such as an earthquake, typhoon, heavy storm, and snowfall, on social media; therefore, the enhancement of situation awareness in the real world using social data is one of the most attractive research subjects. In our previous work, we developed a density-based spatiotemporal system to identify topic-related areas in which there are a huge number of geo-tagged tweets related to a topic are posted. In this paper, we propose a novel density-based spatiotemporal system with a photo image classifier in order to enhance situation awareness by showing accurate topic-related photos. The photo image classifier using a support vector machine (SVM) based on the Bag-of-Features (BoF) model is integrated into the conventional density-based spatiotemporal system. To evaluate the proposed system, we used actual data sets related to weather topics, “heavy rain” and “heavy snow,” in Japan. The experimental results indicate that the proposed system can extract photo images related to these weather topics with high accuracy and recall levels.

Keywords: density-based spatiotemporal analysis system, situation awareness, Bag-of-Features, SVM

1 Introduction

In recent years, people have begun to transmit and receive diversified information, such as weather, natural disasters and diseases, via social media. Posting data on the social media sites, such as Twitter and Facebook is called microblogging and it has formed a new and attractive media alternative to conventional media [1, 2]. The messages and photos about an earthquake, typhoon, heavy storm, and snowfall are posted by the people who have just witnessed and heard these [3]; therefore, analyzing them, and developing crisis management systems can lead to an enhanced situation awareness.

Geo-annotated social data (e.g., geo-tagged tweets and photos) are expected to allow us to detect emergency topics and manage information about them spatiotemporally [4]. When people are face a crisis, they actively post geo-tagged messages and photos to update others about the situation. In our previous paper, we proposed a density-based spatiotemporal analysis system to detect emergency topics and analyze the changes in them [5]. We

* Graduate School of Information Sciences, Hiroshima City University, Hiroshima, Japan

implemented the system using geo-tagged tweets to analyze weather topics in Japan and confirmed that it can analyze changes in situations in local areas. In the system, two main techniques are applied: a Naive Bayes classifier and the (ϵ, τ) -density-based adaptive spatiotemporal clustering method. The Naive Bayes classifier extracts geo-tagged tweets, including those related to an observed emergency topic. The (ϵ, τ) -density-based adaptive spatiotemporal clustering method can extract spatiotemporal density local areas from geo-tagged tweets as (ϵ, τ) -density-based adaptive spatiotemporal clusters, which helps us to analyze the location and timing of emergency topics by updating the spatiotemporal clusters.

The density-based spatiotemporal analysis system also provides a map-based Web interface to visualize bursty local areas and browse messages. Moreover, the map-based Web interface visualizes photos that are attached to geo-tagged tweets to enhance situation awareness. However, some of these photos are not related to the observed emergency topic. For example, people post messages related to an observed emergency, a message says “It is raining heavily,” but the attached photo shows the lunch that the user is eating. To present situations correctly, in this case we have to remove these unrelated photos.

To address the issue mentioned above, we propose a novel density-based spatiotemporal analysis system with a photo image classifier in order to enhance situation awareness during a crisis by showing accurate topic-related photos. The main contributions of this study are as follows.

- To create training data sets for observed topics, we utilize a semi-automatic method in which we label photo images attached to topic-related geo-tagged tweets. First, the Naive Bayes classifier extracts candidate photo images to create a training data set. Second, each candidate photo image is labeled manually.
- To classify photo images, feature vectors of photo images are created by using the Bag-of-Features (BoF) model [6] and a support vector machine (SVM) is utilized.
- To evaluate the proposed method, we used an actual data set related to two weather topics in Japan.

The rest of this paper is organized as follows. In Section 2, related work is reviewed. In Section 3, we explain the density-based spatiotemporal analysis system briefly. In Section 4, the new density-based spatiotemporal analysis system with a photo image classification method is described. In Section 5, experimental results are reported, and the paper is concluded in Section 6.

2 Related Work

Social media have been forming as a new alternative to conventional media [1, 2]. The exploitation of social data for enhancing situation awareness is just beginning [7]. Twitter, in particular, has become a very interesting research subject, because a huge number of users tweet during a crisis [3]. For example, Miyabe et al. [8] analyzed what is commonly called as use of tweets for trend analysis after the Great East Japan Earthquake in Japan. Moreover Murakami et al. [9] reported the results of mining tweets about the tsunami during the Great East Japan Earthquake. Vieweg et al. [10] analyzed tweets generated during two concurrent emergency events in North America. These case studies emphasize the importance of using social media to analyze situations during a crisis.

Some researchers have attempted to develop information systems that enhance situation awareness during a crisis. Sakaki et al. [11] proposed a method for estimating epicenters using tweets and developed an alert system that sends an e-mail to users to alert them when an earthquake occurs. Aramaki et al. [12] proposed a method for detecting flu epidemics using tweets that mention flu, which they extracted using a term-based SVM. Kreiner et al. [13] developed a classification method in which tweets are classified into seven categories, which is known as the United Nations Development Program Security Model, in order to provide or seek crisis information. In our previous paper, we proposed a density-based spatiotemporal analysis system to obtain information about the location and time of emergency topics. Our system can extract changes in topics from moment to moment during a crisis.

Recently, photo images posted on social media have been playing an important role because visual transmission of information has considerable power [14]. Fruin et al. [15] developed TweetPhoto, which displays photos related to news topics attached to posted tweets. Tweets are clustered and the attached photos are scored according to terms. In our previous paper [5], we proposed a density-based spatiotemporal analysis system that visualizes photos in spatiotemporal areas related to emergency topics; however, the density-based spatiotemporal analysis system visualizes the photos, even if they are not related to the emergency topics.

3 Density-based Spatiotemporal Analysis System

In this section, we present a density-based spatiotemporal analysis system briefly.

3.1 Data Model

In this study, geotagged tweets are defined as a sequence. Let the i -th geotagged tweet in $GTS = \{gt_1, \dots, gt_n\}$ be gt_i ; then, gt_i consists of four items: $gt_i = \langle text_i, pt_i, pl_i, api_i \rangle$, where $text_i$ is a short text message, pt_i is posted time, pl_i is location (e.g., latitude and longitude), and api_i is an attached photo image. If $text_{\phi_t^{(etp)}(k)}$ is related to an emergency topic etp , then $rgt_k^{(etp)} = gt_{\phi_t^{(etp)}(k)}$. A set of relevant geotagged tweets is $RGTS^{(etp)} = \{rgt_1^{(etp)}, \dots, rgt_{num(etp)}^{(etp)}\}$, where $RGTS^{(etp)} \subset GTS$. Function $\phi_t^{(etp)}(k)$ is an injective function.

$$\phi_t^{(etp)}(k) : RGTS^{(etp)} \rightarrow GTS; rgt_k^{(etp)} \mapsto gt_{\phi_t^{(etp)}(k)} \quad (1)$$

3.2 System Overview

Suppose that the geotagged tweets arrive singly and the observed emergency topic is etp . The density-based spatiotemporal system comprises three steps.

1. First, to extract geotagged tweets that include the observed emergency topic related to a natural disaster as relevant geotagged tweets, we utilize the Naive Bayes classifier technique in the Tweet Classifier. The Tweet Classifier classifies the i -th geotagged tweet gt_i . If and only if gt_i is classified as “relevant,” which means that gt_i is a relevant geotagged tweet and $rgt_j^{(etp)} = gt_i$, where $\phi_t^{(etp)}(j) = i$ is related to the observed emergency topic, we go to the next step.

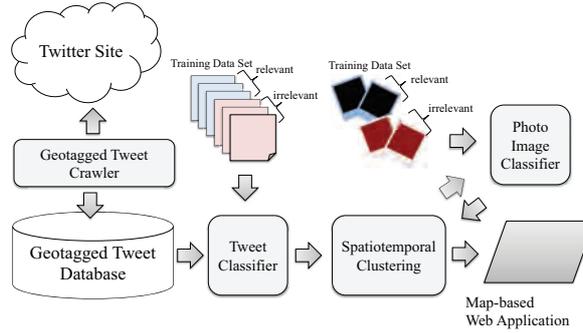


Figure 1: System overview of new density-based spatiotemporal analysis system

2. Next, the Spatiotemporal Clustering Manager executes an incremental algorithm for extracting (ϵ, τ) -density-based adaptive spatiotemporal clusters. The Spatiotemporal Clustering Manager is based on the (ϵ, τ) -density-based adaptive spatiotemporal clustering algorithm that was proposed in [16]. The incremental algorithm updates the states of the extracted (ϵ, τ) -density-based adaptive spatiotemporal clusters and extracts new (ϵ, τ) -density-based adaptive spatiotemporal clusters every time a relevant geotagged tweet arrives.
3. Then, the extracted and updated (ϵ, τ) -density-based adaptive spatiotemporal clusters, which constitute bursty local areas, text messages, and attached photos, are visualized through map-based Web interfaces. If users click on an icon on the map, text messages and the attached photos in the spatiotemporal clusters are visualized using markers on the map.

4 Proposed System

In this section, we propose a new density-based spatiotemporal analysis system with a photo image classifier.

4.1 System Overview

Figure 1 shows an overview of our new density-based spatiotemporal analysis system. In the new system, after spatiotemporal clustering, for each (ϵ, τ) -density-based adaptive spatiotemporal cluster, a set of photo images attached to relevant geotagged tweets is extracted and these images are classified using the photo image classifier. Let a training data set be $TPI^{(etp)} = \{(tpi_1^{(etp)}, tc_1), (tpi_2^{(etp)}, tc_2), \dots, (tpi_l^{(etp)}, tc_l)\}$, where $tc_i = \{relevant, irrelevant\} \in CLASS^{(etp)}$. The training data set is created using a semi-automatic method. The photo image classifier is based on an SVM with the BoF model.

4.2 Semi-automatic Method for Creating Training Data Set

It is difficult to create a training data set manually, because topic-related photo images have to be extracted from a huge number of geotagged tweets. The semi-automatic method can reduce the load on the user. We actively use the information provided by the outputs of the Naive Bayes classifier. First, we gather a set of relevant geotagged tweets, which are classified as “relevant” by the Naive Bayes classifier. Then, we label photo images in the set of relevant geotagged tweets manually. Photo images are labeled “relevant” or “irrelevant.”

4.3 Support Vector Machine with Bag-of-Features

In the BoF model, a photo image represents the frequencies of key features, which are used to create a codebook. The BoF model comprises three steps: (1) local feature extraction, (2) key-feature identification, and (3) histogramming. The second step is used only to create the image histograms of the training data set.

In the local feature extraction step, local features of each photo image are extracted using Speeded-up Robust Features (SURF) [17], which is a robust local feature detector first introduced by Bay et al. Let local features of $tpi_i^{(etp)}$ in $TPI^{(etp)}$ be $TLF_i^{(etp)} = \{tlf_{i,1}^{(etp)}, tlf_{i,2}^{(etp)}, \dots, tlf_{i,numl f^{(etp)}(i)}^{(etp)}\}$, where $tlf_{i,j}^{(etp)}$ is a local feature that is extracted using SURF and $numl f^{(etp)}(i)$ is the number of local features in $tpi_i^{(etp)}$.

In the key-feature identification steps, all local features of the training data set are clustered using k-means clustering. After clustering, a set of cluster centers is extracted as a codebook $CB^{(etp)} = \{cbv_1^{(etp)}, cbv_2^{(etp)}, \dots, cbv_k^{(etp)}\}$, where $cbv_i^{(etp)}$ is the center of cluster i and is called a codebook vector. These codebook vectors represents key features.

In the histogramming step, for each training photo image $tpi_i^{(etp)}$ in $TPI^{(etp)}$, the local features are classified.

$$clf_{i,j}^{(etp)} = \arg \min_c dist(tlf_{i,j}^{(etp)}, cbv_c^{(etp)}) \quad (2)$$

Let a histogram of $tpi_i^{(etp)}$ be $thist_i^{(etp)}$; the j -th element of $thist_i^{(etp)}$ is denoted by $thist_i^{(etp)}[j]$, where $1 \leq j \leq k$.

$$thist_i^{(etp)}[j] = \sum_{l=1}^{numl f^{(etp)}(i)} f(clf_{i,l}^{(etp)}, j),$$

$$f(c, j) = \begin{cases} 1 & (if\ c = j) \\ 0 & (otherwise) \end{cases} \quad (3)$$

Let a feature vector set $FVT^{(etp)}$ of the training data set be

$$FVT^{(etp)} = \{(thist_1^{(etp)}, c_1), (thist_2^{(etp)}, c_2), \dots, (thist_i^{(etp)}, c_i)\}. \quad (4)$$

In this study, an SVM was trained using this $FVT^{(etp)}$. The codebook $CB^{(etp)}$ is used for creating histograms of attached photo images.

4.4 Algorithm

The photo image classifier classifies attached photo images according to the following steps. First, for each (ϵ, τ) -density-based adaptive spatiotemporal cluster $ASTC^{(etp)}$, the attached photo images are extracted. A set of attached photo images in $ASTC^{(etp)}$ is denoted by

$$SAPI^{(etp)} = \{(sapi_1^{(etp)}, sc_1^{(etp)}), (sapi_2^{(etp)}, sc_2^{(etp)}), \dots, (sapi_{sm}^{(etp)}, sc_{sm}^{(etp)})\}, \quad (5)$$

where sm is the number of attached photo images in $ASTC^{(etp)}$. Moreover, $sapi_i^{(etp)}$ in $SAPI^{(etp)}$ that is not classified is $sc_i^{(etp)} = NULL$.

Second, for each $sapi_i^{(etp)}$ in $SAPI^{(etp)}$, local features are extracted. Let the local features of $sapi_i^{(etp)}$ be $SALF_i^{(etp)} = \{sal f_{i,1}^{(etp)}, sal f_{i,2}^{(etp)}, \dots, sal f_{i,numsa^{(etp)}(i)}^{(etp)}\}$, where $sal f_{i,j}^{(etp)}$ is a

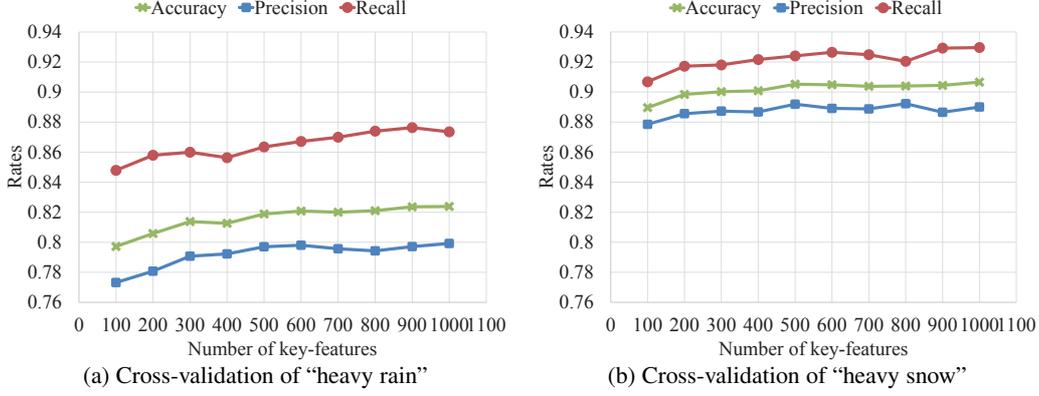


Figure 2: Cross-validation

local feature that is extracted using SURF and $numsa^{(etp)}(i)$ is the number of local features in $sapi_i^{(etp)}$.

To create histograms of attached photo images, the codebook $CB^{(etp)}$ is used. A histogram of $sapi_i^{(etp)}$ is denoted by $shist_i^{(etp)}$, and the j -th element of $sahist_i^{(etp)}$ is denoted by $sahist_i^{(etp)}[j]$, where $1 \leq j \leq k$:

$$sahist_i^{(etp)}[j] = \sum_{l=1}^{numsa^{(etp)}(i)} f(clf_{i,l}^{(etp)}, j),$$

$$f(c, j) = \begin{cases} 1 & (\text{if } c = j) \\ 0 & (\text{otherwise}) \end{cases} \quad (6)$$

To classify $sapi_i^{(etp)}$, $sahist_i^{(etp)}$ is inputted into the photo image classifier, which outputs "relevant" or "irrelevant."

5 Experimental Results

In this section, we report the results of experiments using the new density-based spatiotemporal analysis system with a photo image classifier.

5.1 Experiment Overview

To evaluate the proposed method, we used crawled geotagged tweets and their attached photo images. We analyzed the natural disasters "heavy rain" and "heavy snow" in Japan. First, we performed cross-validation to evaluate the classification performance of the photo image classifier. Second, we evaluated the new density-based spatiotemporal analysis system over a few days. The training data set $TPI^{(etp)}$ was composed of photo images in the set of relevant geotagged tweets extracted by the Naive Bayes classifier, which were posted in July, January, and February. The number of photo images in the training data set $TPI^{(etp)}$ was 1000, divided equally between the "relevant" and "irrelevant" class. When any of the following conditions were satisfied, a photo image was labeled as the "relevant" class.

- The photo image was taken outdoors and confirmed that it was raining or snowing.

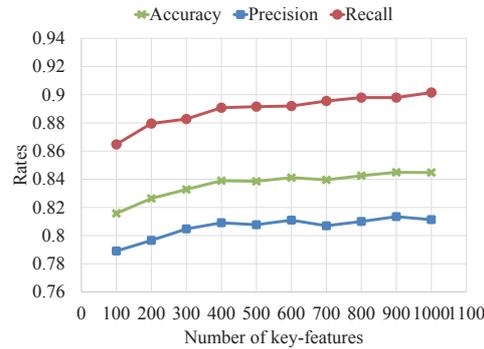


Figure 3: Cross-validation of “heavy rain” using RTDS

- The photo image was taken indoors but rain or snow can be seen, for example, through a window.
- The photo image showed a person wearing rain or snow gear.

A photo image that did not meet any of these conditions was labeled “irrelevant.”

5.2 Cross-validation

To evaluate the classification performance of the photo image classifier, we performed a cross-validation. The number of partitions for the cross-validation were 2, 4, 6, 8, and 10. We compared the results when the number of key features was varied between 100 and 1000. Figures 2(a) and 2(b) show the accuracy, recall, and precision values of the cross-validation. These figures show the average of each partition. It can be seen that the greater the number of key features, the higher the accuracy, precision, and recall. In particular, when the number of key features is 1000, the accuracy, precision and recall values of “heavy rain” are 0.82, 0.80, and 0.87, respectively. Moreover, the accuracy, precision and recall values of “heavy snow” are 0.91, 0.89, and 0.93, respectively.

The experimental results show that the performance for “heavy rain” was lower than that for “heavy snow.” In the classification results of “heavy rain,” photo images including people were misclassified as “relevant.” Then, we recreated the training data set (denoted by RTDS.) First, 67 photo images in the “relevant” class that showed people were removed from the training data set and others were added. Second, 50 photo images in the “irrelevant” class were randomly removed from the training data set and other photo images showing people were added. Figure 3 shows the results of cross-validation using RTDS. The results show higher accuracy, precision, and recall values than those in Figure 2(a). In particular, the results using RTDS show the highest performance when the number of key features was 1000, and the accuracy, precision and recall values are 0.84, 0.81, and 0.90, respectively.

5.3 System Evaluation

We extracted (ϵ, τ) -density-based adaptive spatiotemporal clusters and classified attached photo images. The experimental period of “heavy rain” was from August 1 to August 10, and that of “heavy snow” from February 10 to February 18. During these experimental periods, heavy rain and snow fell in Japan. Figures 4 and 5 show examples of photo images classified as into “relevant” on August 10 and February 14, respectively. The figures show



Figure 4: Examples of photo images classified as into “relevant” on August 10



Figure 5: Examples of photo images classified as into “relevant” on February 14

that the proposed method can classify photo images related to “heavy rain” and “heavy snow.” Moreover, Tables 1 and Table 2 list the number of attached photo images in the spatiotemporal clusters, the number of photo images in the “relevant” class, and the accuracy, precision, and recall values for each day. The range of accuracy, precision, and recall values are from 0.59 to 0.86, from 0.29 to 0.88, and from 0.40 to 1.00, respectively. Although the experimental results show an almost high performance, a low performance can be seen on some days. In the experiments for “heavy rain,” it is conceivable that the landscape and trend of photo images changed between July and August. Therefore, the precision value was low on some days. In the experiments for “heavy snow,” although the experimental results for days when these were a large number of photo images show a high performance, those for days when there were a small number of photo images show a low performance. Precision is the most important quality in them of enhancing situation awareness, because when its level is high the system does not show a photo image not related to the observed emergency topic. Therefore, our future work in which we intend to improve precision is important.

6 Conclusion

In this paper, we proposed a novel density-based spatiotemporal analysis system with a photo image classifier. The aim of this study was to enhance situation awareness during a crisis by showing on the social media photo images that are accurately related to the topic. We developed the photo image classifier using the BoF model and an SVM based on BoF in the density-based spatiotemporal analysis system. To evaluate the proposed system, we used an actual data set related to “heavy rain” and “heavy snow” in Japan. The experimental results indicate that the proposed system can classify photo images related to “heavy rain” and “heavy snow” with high accuracy and recall levels. In our future work, an automatic method for creating a training data set and a new classification method that achieves high precision will be addressed.

Table 1: System evaluation of “heavy rain” from August 1 to August 10

Date	Number of photo images	Number of photo images in “relevant”	Accuracy	Precision	Recall
8/1	130	68	0.70	0.62	0.76
8/2	217	97	0.76	0.55	0.88
8/3	203	88	0.74	0.60	0.75
8/4	63	19	0.86	0.63	0.86
8/5	63	30	0.68	0.43	0.81
8/6	87	41	0.61	0.32	0.68
8/7	53	34	0.62	0.41	1.00
8/8	230	111	0.65	0.36	0.82
8/9	412	178	0.73	0.42	0.89
8/10	572	269	0.76	0.59	0.84

Table 2: System evaluation of “heavy snow” from February 10 to February 18

Date	Number of photo images	Number of photo images in “relevant”	Accuracy	Precision	Recall
2/10	210	108	0.70	0.79	0.67
2/11	206	136	0.72	0.76	0.80
2/12	23	7	0.65	0.29	0.40
2/13	87	42	0.68	0.64	0.68
2/14	2196	1465	0.79	0.84	0.84
2/15	940	582	0.72	0.88	0.73
2/16	306	186	0.70	0.75	0.76
2/17	152	80	0.66	0.66	0.68
2/18	58	35	0.59	0.46	0.76

Acknowledgments

This work was supported by JSPS KAKENHI Grant Number 26330139 and Hiroshima City University Grant for Special Academic Research (General Studies).

References

- [1] Akshay Java, Xiaodan Song, Tim Finin, and Belle Tseng. Why we twitter: Understanding microblogging usage and communities. In *Proceedings of the 9th WebKDD and 1st SNA-KDD 2007 Workshop on Web Mining and Social Network Analysis*, WebKDD/SNA-KDD '07, pp. 56–65, 2007.
- [2] Bernard J. Jansen, Mimi Zhang, Kate Sobel, and Abdur Chowdury. Twitter power: Tweets as electronic word of mouth. *J. Am. Soc. Inf. Sci. Technol.*, Vol. 60, No. 11, pp. 2169–2188, 2009.
- [3] Marcelo Mendoza, Barbara Poblete, and Carlos Castillo. Twitter under crisis: Can we trust what we rt? In *Proceedings of the First Workshop on Social Media Analytics*, SOMA '10, pp. 71–79, 2010.
- [4] Mor Naaman. Geographic information from georeferenced social media data. *SIGSPATIAL Special*, Vol. 3, No. 2, pp. 54–61, 2011.
- [5] Tatsuhiro Sakai and Keiichi Tamura. Real-time analysis application for identifying bursty localareas related to emergency topics. *SpringerPlus*, Vol. 4, No. 162, 2015.

- [6] Gabriella Csurka, Christopher R. Dance, Lixin Fan, Jutta Willamowski, and Cedric Bray. Visual categorization with bags of keypoints. In *In Workshop on Statistical Learning in Computer Vision, ECCV*, pp. 1–22, 2004.
- [7] Jie Yin, Andrew Lampert, Mark Cameron, Bella Robinson, and Robert Power. Using social media to enhance emergency situation awareness. *IEEE Intelligent Systems*, Vol. 27, No. 6, pp. 52–59, 2012.
- [8] Mai Miyabe, Asako Miura, and Eiji Aramaki. Use trend analysis of twitter after the great east japan earthquake. In *Proceedings of the ACM 2012 Conference on Computer Supported Cooperative Work Companion, CSCW '12*, pp. 175–178, 2012.
- [9] Akiko Murakami and Tetsuya Nasukawa. Tweeting about the tsunami?: Mining twitter for information on the tohoku earthquake and tsunami. In *Proceedings of the 21st International Conference Companion on World Wide Web, WWW '12 Companion*, pp. 709–710, 2012.
- [10] Sarah Vieweg, Amanda L. Hughes, Kate Starbird, and Leysia Palen. Microblogging during two natural hazards events: What twitter may contribute to situational awareness. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, CHI '10*, pp. 1079–1088, 2010.
- [11] Takeshi Sakaki, Makoto Okazaki, and Yutaka Matsuo. Earthquake shakes twitter users: Real-time event detection by social sensors. In *Proceedings of the 19th International Conference on World Wide Web, WWW '10*, pp. 851–860, 2010.
- [12] Eiji Aramaki, Sachiko Maskawa, and Mizuki Morita. Twitter catches the flu: Detecting influenza epidemics using twitter. In *Proceedings of the Conference on Empirical Methods in Natural Language Processing, EMNLP '11*, pp. 1568–1576, 2011.
- [13] Karl Kreiner, Aapo Immonen, and Hanna Suominen. Crisis management knowledge from social media. In *Proceedings of the 18th Australasian Document Computing Symposium, ADCS '13*, pp. 105–108, 2013.
- [14] David J. Crandall, Lars Backstrom, Daniel Huttenlocher, and Jon Kleinberg. Mapping the world's photos. In *Proceedings of the 18th International Conference on World Wide Web, WWW '09*, pp. 761–770, 2009.
- [15] Brendan C. Fruin, Hanan Samet, and Jagan Sankaranarayanan. Tweetphoto: Photos from news tweets. In *Proceedings of the 20th International Conference on Advances in Geographic Information Systems, SIGSPATIAL '12*, pp. 582–585, 2012.
- [16] Tatsuhiko Sakai, Keiichi Tamura, and Hajime Kitakami. Emergency situation awareness during natural disasters using density-based adaptive spatiotemporal clustering. In *Database Systems for Advanced Applications, DASFAA 2015 International Workshops, SeCoP, BDMS, and Posters, Hanoi, Vietnam, April 20-23, 2015*, Vol. 9052, pp. 155–169, 2015.
- [17] Herbert Bay, Andreas Ess, Tinne Tuytelaars, and Luc Van Gool. Speeded-up robust features (SURF). *Computer Vision and Image Understanding (CVIU)*, Vol. 110, No. 3, pp. 346–359, 2008.