Paper:

Grasp of Disaster Situation and Support Need Inside Affected Area with Social Sensing

– An Analysis of Twitter Data Before and After the 2011 Great East Japan Earthquake Disaster Occurring –

Shosuke Sato, Kazumasa Hanaoka, Makoto Okumura, and Shunichi Koshimura

International Research Institute of Disaster Science (IRIDeS), Tohoku University
Aoba 468-1, Aramaki, Aoba, Sendai 980-0845, Japan
E-mail: ssato@irides.tohoku.ac.jp
[Received October 1, 2015; accepted January 21, 2016]

There are increasing expectations that social sensing, especially the analysis of social media text as a source of information for COP (Common Operational Picture), is useful for decision-making about responses to disasters. This paper reports on a geo-information and content analysis of three million Twitter texts sampled from Japanese Twitter accounts for one month before and after the 2011 Great East Japan Earthquake disaster. The results are as follows. 1) The number of Twitter texts that include geotag (latitude and longitude information) is too small for reliable analysis. However, a method of detecting the tweet's location from the tweet's text using GeoNLP (an automatic technology to tag geo-information from natural language text) is able to identify geo-information, and we have confirmed that many tweets were sent from stricken areas. 2) A comparison of Twitter data distribution before and after the disaster occurred does not identify clearly which areas were significantly affected by the disaster. 3) There were very few Twitter texts that included information about the damage in affected areas and their support needs.

Keywords: social media, Twitter, Common Operational Picture (COP), disaster information system, geo-information, disaster situation

1. Introduction

Expectations about the usefulness of "social sensing" for disaster responses are increasing. This paper defines social sensing as "actions for obtaining information on dynamic individual and social situations and social elements (e.g., logistics, traffic volumes, and the status of natural environments) and their mutual relationships." In the context of a disaster, the term denotes "actions for understanding changes in affected areas, livelihoods, and victims' actions, obtained hourly by estimating victims' situations and affected areas based on location information from mobile phones, probe cars, and social media information."

Many efforts have been made recently to use the social media listed above during a disaster. For example, DISAANA (DISAster-information ANAlyzer) was released experimentally as a system that analyzes disaster-related text in Japanese on Twitter and provides search results in a question and answer format [1]. Rokuse et al. [2] proposed a method for gathering similar tweets in order to construct a system that collects and organizes information from Twitter during a disaster and provides appropriate information according to users' attributes and situations.

Although these approaches make it possible to improve disaster responses through technology innovation, data quality has been insufficiently discussed. There has been no evaluation of social media's effectiveness as an information source for identifying affected areas and understanding the situation during the "information blank period" when public information is lacking immediately after a disaster. In particular, the usefulness of Twitter for understanding the situation in affected areas during a large-scale and widespread disaster such as the Great East Japan Earthquake has not been investigated. Existing research has studied only local disasters [3] and estimated the overall impact of the Great East Japan Earthquake by aggregating and comparing the number of tweets and words [4, 5].

This study is a comprehensive investigation of the hypothesis "Twitter is useful when understanding disaster situations." It tests the hypothesis from the viewpoint of central and local government, NPOs and NGOs, and some private companies that support and relieve affected area. At a minimum, these organizations need data is about time (when), place (where), and situation (what is occurring in what manner, what the important issues are, and what support is required). Since Twitter includes a time stamp with the date and time of transmission, time is not a significant issue for investigation. On the other hand, *place* and *situation* are dependent on the content of each tweet. Therefore, the potential for identifying *place* and *situation* should be investigated in any analysis of the usefulness of Twitter in understanding disaster situations.

This paper uses tweets related to the large-scale,

widespread Great East Japan Earthquake to test Twitter's potential for improving spatial understanding by investigating: 1) whether there are tweets from affected areas (whether tweets from affected areas can be identified) and 2) whether affected areas can be identified by comparing the numbers of tweets before and after the disaster. Twitter's potential for improving situational understanding is tested by investigating 3) whether the place, extent of damage, and support needs can be identified using tweets after the disaster. Questions 1) and 3) are the primary hypotheses for investigation. Question 2) is investigated partly because there may be a significant difference in the number of tweets in inundated and non- inundated areas in the case of a tsunami. During the Great East Japan Earthquake, the tsunami seriously damaged inundated areas. This suggests that rescue and support was urgently needed and there would be frequent tweets that requested assistance. Alternatively, it was possible that victims in inundated areas would be unable to tweet because of the tremendous damage, and this would significantly reduce the number of tweets. Consequently, the authors considered that the number of tweets could be used to identify affected areas, because the number of tweets in areas inundated by the tsunami would either rapidly increase or decrease.

2. Data

Tweets in Japanese from February 11, 2011 to April 11, 2011 were the data for analysis. Tweets before and after the Great East Japan Earthquake were used in order to investigate question 2) whether affected areas can be determined by comparing the number of tweets before and after the disaster. This paper includes data immediately after the disaster because we wished to focus on social media's potential during an "information blank period" when public information is not easily available. A certain amount of public information is likely to be available a week after a disaster. Therefore, data for one month was used to provide a safe margin.

The number of tweets in Japanese from February 11, 2011 to April 11, 2011 was approximately 1.1 billion. The authors believe that each tweet should be read to investigate the effectiveness of using Twitter as a means of understanding the situation in a disaster. However, it is not practical to browse 1.1 billion tweets, and three million tweets were randomly selected from those in the target period to provide the data set. The sampling rate is approximately 0.27%.

Tweets include originals and retweets. The sample of tweets for investigation does not include pairs of original tweets and corresponding retweets, because of the low sample rate. The following analysis therefore does not differentiate between these two kinds of tweets.

The sampled data set consists of tweets from across Japan. However, this paper focuses on those from the Miyagi prefecture, which was significantly affected by the Great East Japan Earthquake, to obtain detailed results.

3. Analysis: Potential for Understanding Spatial Distribution of Affected Area Using Tweets

This chapter considers the questions about Twitter's potential for improving spatial understanding: whether there are tweets in affected areas (whether tweets from affected areas can be identified), and whether affected areas can be identified by comparing the numbers of tweets before and after the disaster.

In order to examine whether there are tweets in affected areas (whether tweets from affected areas can be identified), information about the sender's location is processed and analyzed.

Twitter allows users to add a geotag that gives the coordinates (latitude and longitude) of each tweet. However, most users tweet without using geotags.

Information about the sender's location can also be obtained from items such as the names of places and landmarks in the text of the tweet. GeoNLP is a fundamental technology that obtains location information from natural sentences in the text [6]. GeoNLP is able to resolve ambiguities such as determining which part of text is a place name and the place name's location by combining GIS (Geographic Information System) and NLP (Natural Language Processing). Our research added location information to each tweet using information produced by GeoNLP. It should be noted that GeoNLP estimates location using the information in the text of the tweet, and that location might not be the sender's physical location. For example, senders in areas other than Sendai can tweet "It seems something terrible is occurring in Sendai city," and GeoNLP will identify "Sendai city" as the location. Although this is the case for the analysis in this chapter, Chapter 4 analyzes only those tweets that have been verified as originating in the Miyagi prefecture.

Table 1 shows the number of tweets with location information, either from geotags or from GeoNLP analysis. **Fig. 1** plots the tweets that have been identified as originating in the Miyagi prefecture in the week following the disaster. **Fig. 1** also plots the tweets identified as originating outside the Miyagi prefecture for reference.

The complete data set contains three million tweets. There were 1,619,370 tweets in Japanese in the month following the disaster (**Table 1**). Only 3,384 tweets (0.21%) included a geotag, and 57 of these were from the Miyagi prefecture. It is not possible to determine the true number of tweets from the Miyagi prefecture out of the three million tweets in the data set, but it is clear that tweets with latitude and longitude information based on the GPS function are very rare.

However, GeoNLP identified location information in 221,780 (13.7%) of tweets in the data set that were sent in the month after the disaster (**Table 1**). GeoNLP identified 5,903 tweets as originating in the Miyagi prefecture. GeoNLP identified more tweets with location information than achieved by geotags by a factor of approximately 66 for Japan and 104 for the Miyagi prefecture.

Table 1. Number of tweets with information about location.

	Area	1st week: Mar. 11-17	2nd week: Mar. 18-24	3rd week: Mar. 25-31	4th week: Apr. 1-7	5th week: Apr. 8-11	Total
Raw Data	Japan	317,693	355,661	365,242	376,266	204,508	1,619,370
Geo tag	Japan	678	637	633	824	612	3,384
	Miyagi	17	9	13	8	10	57
GeoNLP	Japan	64,102	40,171	43,759	45,210	28,538	221,780
	Miyagi	2,627	918	845	869	644	5,903

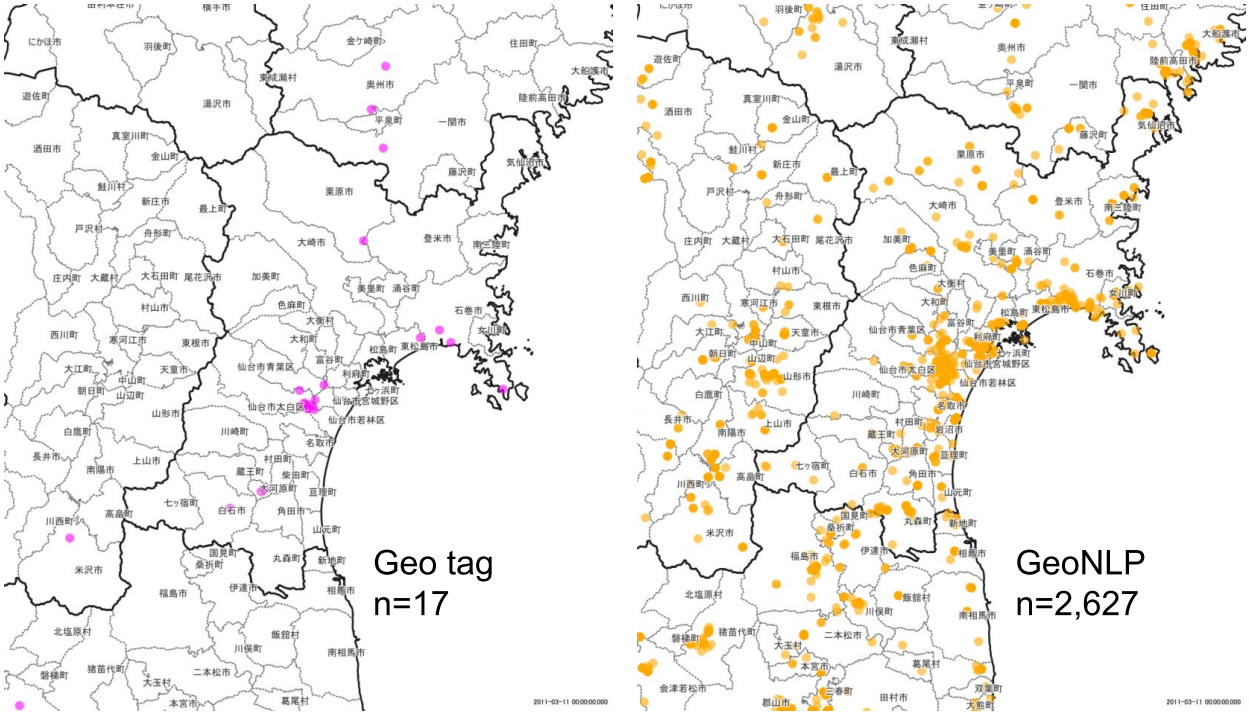

Fig. 1. Tweet location information in the Miyagi prefecture, one week after the disaster: March 11 to March 17, 2011.

Figure 1 demonstrates the fact that tweets whose location was identified by GeoNLP are more widespread geographically than those identified by geotags. These tweets do not always border areas affected by the Great East Japan Earthquake in the Miyagi prefecture, and there are tweets from inland and Pacific coastal areas that were seriously damaged by the tsunami.

The analysis summarized above confirms that there were tweets from affected areas in the Miyagi prefecture during the Great East Japan Earthquake and that a number of the tweets from affected areas can be identified by analyzing location information in the tweet's text with GeoNLP. Tweet data with location information identified by GeoNLP are used in the following analyses.

The second question about the potential for improving spatial understanding through Twitter, whether affected areas can be identified by comparing the numbers of tweets before and after the disaster, is discussed in the next section.

Figure 2 maps tweets from the Miyagi prefecture with GeoNLP location information for one month before and after the disaster. **Fig. 2** includes tweets identified as originating from outside the Miyagi prefecture for reference. The number of tweets from the Miyagi prefecture increased threefold, from 1,921 to 5,903, after the disaster. No clear difference between the increase or decrease of tweets from particular places can be observed in the spatial comparison of tweet distribution before and after the disaster because tweets increased after the disaster in all areas.

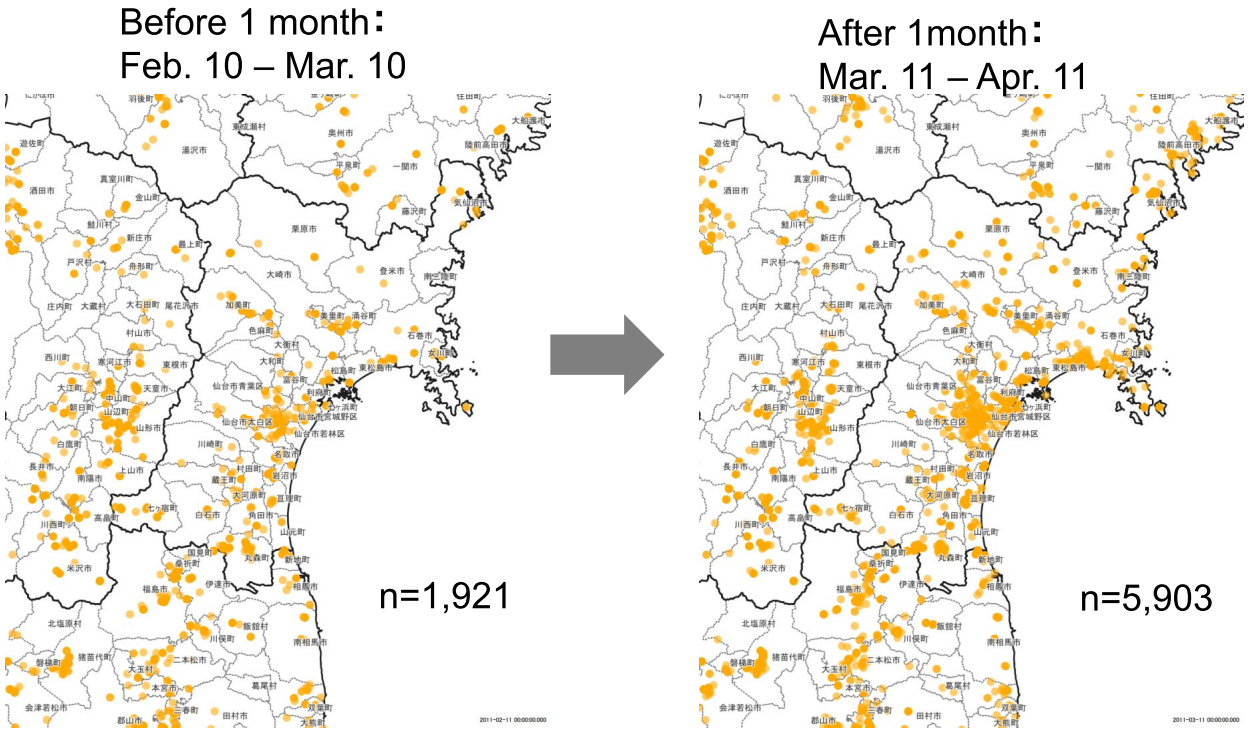

Fig. 2. Comparison of tweet distribution for one month before and after the disaster.

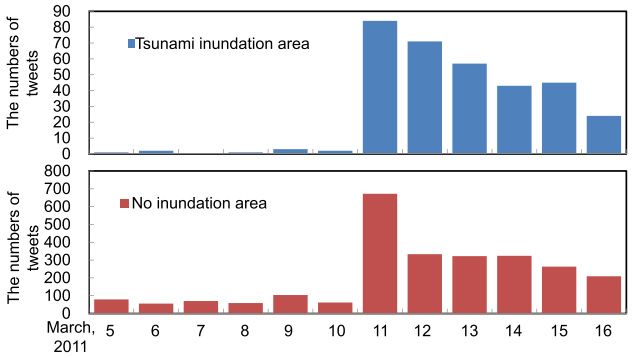

Fig. 3. Changes in the number of tweets in areas inundated and not inundated by the tsunami (Miyagi prefecture).

Figure 3 overlays tweet data with GeoNLP location information on a tsunami inundation map [7] to compare the number of tweets for each day of the week immediately prior to and post the disaster from inundated and non- inundated areas. The number of tweets increased in both inundated and non- inundated areas, suggesting that it is not possible to identify inundated areas solely on the increase or decrease in the number of tweets.

This study originally proposed the hypothesis that "the number of tweets in seriously damaged areas rapidly increases (because rapid environmental changes worth tweeting about frequently occur) or decrease (because it becomes difficult to post information using social media owing to rapid environmental changes and infrastructure outages)." However, the data shows that, in the case of the Great East Japan Earthquake, it is not possible to identify inundated areas based on an increase or decrease in the number of tweets. Therefore, a comparison of tweet distribution before and after a disaster is unlikely to identify affected areas.

4. Analysis: The Potential for Using Tweets to Understand Disaster Situations

This chapter discusses the third question about Twitter's usefulness in a disaster, whether the situation (what is occurring in what manner and what the important issues are) and what support is required) can be determined using tweets sent after the disaster. The following process was used for each tweet sent in the week following the disaster that GeoNLP had identified as being from the Miyagi prefecture.

1) Create a card for each tweet sent in the week following the disaster.
2) Create structured groups based on tweets that contain the same issue.
3) Create separate cards if a tweet contains multiple issues, as one tweet can be included in multiple groups.
4) Separate tweets about direct experience from those about indirect experience.
5) Label each group.

Actions 2) and 3) were performed by multiple evaluators. Action 4) was performed by three evaluators. Evaluator A was a technical staff member in a research institute, and evaluators B and C were involved in disaster research. Actions 3) *grouping* and 5) *labeling* were performed as

Table 2. Results of tweet analysis.

No.	Tweets	Direct		Indirect		Direct/ indirect	Total		Understanding damage in affected areas	Understanding needs for assistance in affected areas
		No. of tweets	Ratio	No. of tweets	Ratio		No. of tweets	Ratio		
1	Comments on status of affected areas	466	10.6%	224	5.1%	2.1	690	15.7%		
2	Comments on safety confirmation of family members and friends	40	0.9%	265	6.0%	0.2	305	6.9%		
3	Concern about safety of family members and friends	161	3.7%	125	2.8%	1.3	286	6.5%		
4	Concern about affected areas	121	2.7%	11	0.2%	11.0	132	3.0%		
5	Plan to go to affected areas	66	1.5%	39	0.9%	1.7	105	2.4%		
6	Requests for safety information on family members and friends	57	1.3%	47	1.1%	1.2	104	2.4%		
7	Requests for information on affected areas	89	2.0%	10	0.2%	8.9	99	2.2%		
8	Plan to provide supplies (food and water, etc.)	39	0.9%	50	1.1%	0.8	89	2.0%		
9	Comments on the government's report	71	1.6%	9	0.2%	7.9	80	1.8%		
10	Announcement on supermarkets and convenience stores	40	0.9%	39	0.9%	1.0	79	1.8%		
11	Comments on the intensities and epicenter	5	0.1%	69	1.6%	0.1	74	1.7%		
12	Requests for rescue	9	0.2%	58	1.3%	0.2	67	1.5%		○
13	Comments on tsunami	9	0.2%	58	1.3%	0.2	67	1.5%		
14	Comments on deaths	0	0.0%	66	1.5%	0.0	66	1.5%		
15	Comments on damage and restoration of the airport	0	0.0%	60	1.4%	0.0	60	1.4%		
16	Comments on nuclear plant accidents and radiation contamination	12	0.3%	48	1.1%	0.3	60	1.4%		
17	Comments on volunteer activities	21	0.5%	34	0.8%	0.6	55	1.2%		
18	Comments on affected areas	15	0.3%	38	0.9%	0.4	53	1.2%	○	
19	Comments on electricity recovery	20	0.5%	32	0.7%	0.6	52	1.2%		
20	Encouragement for victims	44	1.0%	8	0.2%	5.5	52	1.2%		
21	Comments on shelters	3	0.1%	45	1.0%	0.1	48	1.1%		
22	Introduction of information services (radio, SNS account, disaster message bo	1	0.0%	45	1.0%	0.0	46	1.0%		
23	Pray	35	0.8%	9	0.2%	3.9	44	1.0%		
24	Comments on electricity outages	25	0.6%	19	0.4%	1.3	44	1.0%		
25	Comments on planned outages	10	0.2%	30	0.7%	0.3	40	0.9%		
26	Thanks for encouraging and providing information to fiends	15	0.3%	24	0.5%	0.6	39	0.9%		
27	Announcement on gas stations	21	0.5%	14	0.3%	1.5	35	0.8%		
28	Comments on the fire	1	0.0%	34	0.8%	0.0	35	0.8%	○	
29	Announcement of own safety	34	0.8%	1	0.0%	34.0	35	0.8%		
30	Announcement on road damage	12	0.3%	17	0.4%	0.7	29	0.7%	○	
31	Comments on the occurrence of earthquake	1	0.0%	26	0.6%	0.0	27	0.6%		
32	Announcement on supplies (food, water, etc.) provision sites	4	0.1%	22	0.5%	0.2	26	0.6%		
33	Announcement on short supplies (food, water, etc.)	4	0.1%	22	0.5%	0.2	26	0.6%		○
34	Comments on water outages	11	0.2%	14	0.3%	0.8	25	0.6%	○	
35	Comments on railway (subway)	0	0.0%	23	0.5%	0.0	23	0.5%		
36	Comments on intensities and epicenter	3	0.1%	19	0.4%	0.2	22	0.5%		
37	Comments on donations	10	0.2%	10	0.2%	1.0	20	0.5%		
38	Comments on the availability of mobile and fixed phones	13	0.3%	6	0.1%	2.2	19	0.4%	○	
39	Announcement on missing people	0	0.0%	19	0.4%	0.0	19	0.4%		
40	Announcement on isolation	0	0.0%	19	0.4%	0.0	19	0.4%	○	
41	Propose on announcement using Twitter	1	0.0%	17	0.4%	0.1	18	0.4%		
42	Announcement on damage in stations	0	0.0%	18	0.4%	0.0	18	0.4%		
43	Announcement on organizations accepting donations	5	0.1%	13	0.3%	0.4	18	0.4%		
44	Comments on falling snow	15	0.3%	3	0.1%	5.0	18	0.4%		
45	Comments on the availability of parcel delivery services	7	0.2%	10	0.2%	0.7	17	0.4%		
46	Comments on disconnected cell phone cervices	11	0.2%	6	0.1%	1.8	17	0.4%		
47	Comments on gas outages	2	0.0%	14	0.3%	0.1	16	0.4%	○	
48	Comments on activities of on-screen talents	1	0.0%	15	0.3%	0.1	16	0.4%		
49	Comments on the availability of hospitals and facilities	3	0.1%	13	0.3%	0.2	16	0.4%		
50	Announcement on bus operation	4	0.1%	11	0.2%	0.4	15	0.3%		
51	Comments on activities of the Self-Defense Forces	6	0.1%	9	0.2%	0.7	15	0.3%		
52	Comments on surrounding areas	0	0.0%	15	0.3%	0.0	15	0.3%		
53	Requests for information on gas stands	12	0.3%	2	0.0%	6.0	14	0.3%		
54	Comments on fear and anxieties	10	0.2%	4	0.1%	2.5	14	0.3%		
55	Announcement on issued tsunami alerts	2	0.0%	12	0.3%	0.2	14	0.3%		
56	Announcement on railways	13	0.3%	0	0.0%	100.0	13	0.3%		
57	Announcement on canceled or delayed events	3	0.1%	9	0.2%	0.3	12	0.3%		
58	Alerts for affected areas (security and health)	8	0.2%	4	0.1%	2.0	12	0.3%		
59	Comments on shakes	8	0.2%	4	0.1%	2.0	12	0.3%		
60	Comments on gas recovery	3	0.1%	8	0.2%	0.4	11	0.2%		
61	Alerts for aftershocks	2	0.0%	9	0.2%	0.2	11	0.2%		
62	Comments on indoor mess	5	0.1%	5	0.1%	1.0	10	0.2%		
63	Encouragement for family members and friends	4	0.1%	6	0.1%	0.7	10	0.2%		
64	Comments on water recovery	3	0.1%	7	0.2%	0.4	10	0.2%		
65	Alerts for tsunami	1	0.0%	9	0.2%	0.1	10	0.2%		
66	Comments on blockage and recovery of reads	0	0.0%	10	0.2%	0.0	10	0.2%		
67	Comments on how to use Twitter and tips	1	0.0%	8	0.2%	0.1	9	0.2%		
68	Question on safety confirmation methods	6	0.1%	3	0.1%	2.0	9	0.2%		
69	Comments on purchases of local food from affected areas	6	0.1%	3	0.1%	2.0	9	0.2%		
70	Announcement of evacuees list	2	0.0%	7	0.2%	0.3	9	0.2%		
71	Announcement on bathing facilities	4	0.1%	4	0.1%	1.0	8	0.2%		
72	Announcement on aftershock	5	0.1%	3	0.1%	1.7	8	0.2%		
73	Comments on going home	6	0.1%	1	0.0%	6.0	7	0.2%		
74	Comments on mobile phone batteries that were running out	1	0.0%	6	0.1%	0.2	7	0.2%		
75	Comments on tsunami arrival and inundation	1	0.0%	6	0.1%	0.2	7	0.2%	○	
76	Requests for shelter information	5	0.1%	2	0.0%	2.5	7	0.2%		
77	Call for refraining from using Twitter	3	0.1%	3	0.1%	1.0	6	0.1%		
78	Announcement on power sources for mobile phones	2	0.0%	4	0.1%	0.5	6	0.1%		
79	Announcement on predicted tsunami	0	0.0%	6	0.1%	0.0	6	0.1%		
80	Announcement on short supplies (food and water, etc.)	1	0.0%	4	0.1%	0.3	5	0.1%		○
81	Announcement on cold weather	4	0.1%	0	0.0%	100.0	4	0.1%		
82	Comments on damage in the Tokyo Metropolitan area	2	0.0%	2	0.0%	1.0	4	0.1%		
83	Call for evacuation from tsunami	0	0.0%	4	0.1%	0.0	4	0.1%		
84	Comments on communication measures	3	0.1%	1	0.0%	3.0	4	0.1%		
86	Concern about non-affected areas	4	0.1%	0	0.0%	100.0	4	0.1%		
87	Requests for information on provision sites for supplies (food and water, etc.)	3	0.1%	1	0.0%	3.0	4	0.1%		
88	Requests for information on supermarkets and convenient stores	2	0.0%	1	0.0%	2.0	3	0.1%		
89	Call for provision of medical devices and materials	0	0.0%	3	0.1%	0.0	3	0.1%		
90	Call for electricity savings	2	0.0%	1	0.0%	2.0	3	0.1%		
91	Comments on issued evacuation orders and advisories	0	0.0%	3	0.1%	0.0	3	0.1%		
92	Announcement on mortuaries	0	0.0%	2	0.0%	0.0	2	0.0%		
93	Alerts for the fire	0	0.0%	2	0.0%	0.0	2	0.0%		
94	Comments on building damage	0	0.0%	2	0.0%	0.0	2	0.0%		
95	Announcement on sender's evacuation	2	0.0%	0	0.0%	100.0	2	0.0%		
96	Prediction of intensities and epicenter	0	0.0%	2	0.0%	0.0	2	0.0%		
97	Announcement on sites accepting victims	1	0.0%	1	0.0%	1.0	2	0.0%		
98	Announcement on the number of evacuees	0	0.0%	2	0.0%	0.0	2	0.0%		
99	Announcement on the all clear of tsunami alerts	0	0.0%	1	0.0%	0.0	1	0.0%		
100	Prediction of tsunami arrival time and height	0	0.0%	1	0.0%	0.0	1	0.0%		
101	Irrelevant to the disaster	0	0.0%	112	2.5%	0.0	112			
102	Others	8	0.2%	569	12.9%	0.0	577			

Table 3. Tweets likely to be useful for understanding support needs in affected areas.

No.	Tweets	Direct		Indirect		Direct/ indirect	Total		Understanding damage in affected areas	Understanding needs for assistance in affected areas
		No. of tweets	Ratio	No. of tweets	Ratio		No. of tweets	Ratio		
12	Requests for rescue	9	0.2%	58	1.3%	0.2	67	1.5%		○
14	Comments on deaths	0	0.0%	66	1.5%	0.0	66	1.5%		
15	Comments on damage and restoration of the airport	0	0.0%	60	1.4%	0.0	60	1.4%		
18	Comments on affected areas	15	0.3%	38	0.9%	0.4	53	1.2%	○	
28	Comments on the fire	1	0.0%	34	0.8%	0.0	35	0.8%	○	
30	Announcement on road damage	12	0.3%	17	0.4%	0.7	29	0.7%	○	
33	Announcement on short supplies (food, water, etc.)	5	0.1%	22	0.5%	0.2	26	0.6%		○
34	Comments on water outages	11	0.2%	14	0.3%	0.8	25	0.6%	○	
38	Comments on the availability of mobile and fixed phones	13	0.3%	6	0.1%	2.2	19	0.4%	○	
40	Announcement on isolation	0	0.0%	19	0.4%	0.0	19	0.4%		
47	Comments on gas outages	2	0.0%	14	0.3%	0.1	16	0.4%	○	
75	Comments on tsunami arrival and inundation	1	0.0%	6	0.1%	0.2	7	0.2%	○	
89	Call for provision of medical devices and materials	0	0.0%	3	0.0%	0.5	3	0.1%		
94	Comments on building damage	0	0.0%	2	0.0%	0.0	2	0.0%		

follows: if, for example, "a TV program reported a shortage of medical equipment" was tweeted, the "shortage of medical equipment" was focused on as specific information, not the "TV program report." As described in Chapter 3, GeoNLP does not always identify the sender's location, so process 4) was used to assist in this.

Table 2 presents the results of the analysis above, along with the number of tweets in each labeled group. One hundred and two groups were created. Each tweet was marked with a circle if it was useful for "understanding the situation in affected areas" as well as "understanding support needs in affected areas." Tweets that were useful for either "understanding the situation in affected areas" or "understanding support needs in support in affected areas" were marked with a shading in the table. "Useful" labels identified the groups of tweets that contained useful information for understanding the situation in affected areas as well as support needs. On the other hand, "indirect" labels identified the groups of tweets that were useful for either "understanding the situation in affected areas" or "understanding support needs in affected areas." These cards are shaded lightly and not marked with a circle. **Table 3** extracts relevant cells from **Table 2** with the same shading and circles.

Six hundred and ninety tweets (15.7%) were grouped into "comments on the situation in affected areas," which is the largest category (**Table 2**). Examples include "I felt an earthquake in my car" and "terrible... a large-scale fire occurred in Kesennuma city." Many tweets concerned people's safety, such as "safety confirmation for family members and friends" (305 tweets, 6.9%) and "safety concern for family members and friends" (286 tweets, 6.5%).

Groups of tweets likely to be useful for "understanding the situation in affected areas" and "understanding support needs in affected areas" are discussed below.

1) Requests for rescue (nine tweets)

These are useful for understanding rescue needs. "Rescue needed! In a house in Hebita, Ishinomaki City, Miyagi prefecture. Water is reaching the second floor. I cannot move to the roof because I have two children." "Please diffuse!! [Urgent] Parents of my employee are trapped in an immersed car. Could somebody help them? x-x-x, Sakae, Tagajo City. Please diffuse!"

2) Comments on affected area situations (15 tweets)

These are useful for understanding the situation in affected areas. "Inconvenience because of outages. Our town is not as affected as reported in TV: Tomizawa, Taihku district. I just saw that the ground is not safe and buildings have collapsed in Nagamachi-Minami and Nagamachi." "[Please diffuse] The lifeline was completely disabled. Food and information unavailable. I manage to use my mobile phone."

3) Comments on the fire (one tweet)

This is useful for understanding the situation in affected areas, especially fires. "Buncho seems to be all right. Some buildings lost power. No fire. The fire occurred in Miyagino district."

4) Comments on damage to roads (12 tweets)

These are useful for understanding the situation in affected areas, especially damage to roads. "I did not go that way. What I can say is that cars could not get from Ishinomaki and Minatomachi to Watanoha. So I drove from Inai." "If you want to go to Sendai, fill up your car because roads are heavily jammed. There are many cracks in the road and larger gaps of 30 cm. Gaps of about 10 cm are everywhere."

5) Announcements of shortages of supplies (food and water, etc.) (four tweets)

These are useful for understanding supply needs. "Seriously short supplies in Watari elementary school in Miyagi prefecture. One rice ball is shared by four persons. This is the only food for the day." "Clothing needed in Sendai. Especially men's and women's extra-large clothing and women's underwear."

6) Comments on water outages (11 tweets)

These are useful for understanding damage in affected areas, especially water outages. "I found a water outage in my house." "But electricity, water, gas to my apartment have not been reconnected."

7) Comments on the availability of fixed and mobile phones (13 tweets)

These are useful for understanding damage in affected areas, especially to communications. "Finally mobile phone service recovered in central Sendai." "[A friend] called me!!! The phone service is reconnected in Sendai!!!"

8) Comments on gas outages (two tweets)

These are useful for understanding damage in affected areas, especially gas outages. "But electricity, water, gas to my apartment have not been reconnected." "Electricity and water reconnected at about five. Kooriyama, Taihaku district, Sendai city. Gas has not been supplied."

9) Comments on tsunami and inundation (one tweet)

This is useful for understanding damage in affected areas, especially tsunami and inundation. "Tsunami arrived in neighboring districts. I am worried about [personal name] and [personal name]. [Personal name] and Tagajo. Information needed."

Among the 102 groups of tweets, only nine groups were obtained directly from senders and were useful for understanding damage or support needs in affected areas. Note that these nine groups include 213 indirect tweets. Only 69 tweets (1.6%) were included in the nine groups, out of 2,627 tweets. This is likely to be because Twitter is a tool for daily use among communities of individuals, and it is not used for transmitting information about the situation in affected areas or support needs during the disaster.

5. Conclusion

In order to test conclusively the hypothesis that "Twitter is useful for understanding the situation during disasters," this paper has analyzed information about location and the subject matter of tweets, using a data set of sampled tweets about the Great East Japan Earthquake. The results are summarized as follows.

Question 1) whether there are tweets from affected areas (whether tweets from affected areas can be identified); because few users add geotags to their tweets, few tweets were identified as being sent from affected areas. More tweets sent from affected areas were identified using means other than geotags.

Question 2) whether affected areas can be identified by comparing the numbers of tweets before and after the disaster; the number of tweets tended to increase after the disaster in all areas. Affected areas cannot at present be identified by comparing the numbers of tweets before and after the disaster.

Question 3) whether the situation in affected areas and support needs can be identified using tweets after the disaster; very few tweets were useful for understanding the situation in affected areas or support needs.

The analysis above demonstrates that it is extremely difficult to identify affected areas and accurately understand the situation in affected areas during an "information blank period" with insufficient public information after a large-scale and widespread disaster such as the Great East Japan Earthquake. It is difficult to identify which tweets are from affected areas because people outside badly affected areas also tweeted frequently. Tweets with information about the sender's location were extremely scarce because most tweets did not include geotags or descriptions of locations. In order to improve this situation, the following ideas are proposed: 1) the development of an algorithm to identify tweets from affected areas accurately; 2) the enhancement of users' literacy to encourage them to provide information that assists in identifying their location and support needs; and 3) the establishment of a special SNS for disasters.

Regarding the first proposal, it is possible to develop an algorithm by analyzing the characteristics of direct tweets that are useful for understanding damage and support needs. However, this paper analyzed very few tweets (69) in this category and more data are required to create an algorithm. The second proposal may be a burden on affected users during a disaster. The third proposal is useful to a certain extent because tweets relevant to a disaster can be identified by establishing a disaster mode interface different from the usual Twitter interface. Uchida et al. [8] developed a system for tweets that automatically inserts location information and conducted demonstration experiments. However, there are few tweets useful for understanding damage and support needs in a large-scale disaster, as shown in this analysis, and therefore the cost-effectiveness and necessity of the second and third proposals should be examined.

Acknowledgements

This study was funded from several sources. They are JSPS KAKENHI Grant Number 24710180: (Principal Investigator, Shosuke Sato); JSPS KAKENHI Grant Number 25242035 (Principal Investigator, Shunichi Koshimura); and JST (Japan Science and Technology Agency) Strategic Basic Research Programs (CREST) (Principal Investigator, Shunichi Koshimura). NTT DATA Corporation and Nightley Inc. supported Twitter data processing. Ms. Sanae Amita (assistant technician at the International Research Institute of Disaster Science, Tohoku University) helped handle materials. The authors wish to thank them.

References:

[1] National Institute of Information and Communications Technology, "DISAANA," 2014, http://www.nict.go.jp/press/2014/11/05-1.html [accessed October 1, 2015]

[2] T. Rokuse, S. Nagashima, O. Uchida, and F. Toriumi, "A Proposal on Information Service System at the Time of Large-Scale Disaster Using Twitter," Forum on Information Technology, Vol.4, No.12, pp. 651-652, August, 2013.

[3] T. Ishikara, A. Kawasaki, and K. Meguro, "Investigation of information sharing by Twitter users during the 2011 heavy snow disaster in San-in region of western Japan," Journal of Social Safety Science, No.17, pp. 135-143, 2012.

[4] T. Sakaki, J. Marui, F. Toriumi, K. Shinoda, K. Kazama, S. Kurihara, and I. Noda, "Changes in social media use in large-scale dis-

asters," The Association for Natural Language Processing, No.18, pp. 1228-1231, 2012.
[5] T. Sakaki, M. Okazaki, and Y. Matsuo, "Earthquake shakes Twitter users: real-time event detection by social sensors," Proceedings of the 19th international conference on World wide web, pp. 851-860, 2010.
[6] National Institute of Informatics, GeoNLP, https://geonlp.ex.nii.ac.jp/ [accessed October 1, 2015]
[7] Ministry of Land, Infrastructure, Transport and Tourism: Tsunami inundation area data in the 2011 Great East Japan Earthquake disaster, http://fukkou.csis.u-tokyo.ac.jp/ [accessed October 1, 2015]
[8] O. Uchida, M. Kosugi, G. Endo, T. Funayama, K. Utsu, S. Tajima, M. Tomita, Y. Kajita, and Y. Yamamoto, "A Disaster Related Information Sharing System Based on the Utilization of Twitter," Proceedings of the IEICE Engineering Sciences Society/NOLTA Society Conference, 2015.

Name:
Shosuke Sato

Affiliation:
Assistant Professor, Ph.D. Informatics, International Research Institute of Disaster Science (IRIDeS), Tohoku University

Address:
Aoba 468-1, Aramaki, Aoba, Sendai 980-0845, Japan

Brief Career:
2009- JSPS Research Fellows (DC2)
2011- Assistant Professor, Disaster Control Research Center (DCRC), Graduate School of Engineering, Tohoku University
2012- Assistant Professor, International Research Institute of Disaster Science (IRIDeS), Tohoku University

Selected Publications:
• S. Sato et al., "Grasp of utilization of social networking services in restoration process – Interview survey for N city-related citizens of the Great East Japan Earthquake –," Journal of Disaster Research, Vol.10, No.5, pp. 948-955, 2015.
• S. Sato et al., "A Behavior Experiment and Analysis of Tsunami Evacuation Focused on Visitors and Navigation – Case of the Central Area of Ishinomaki City –," Journal of Japan Society Civil Engineering, Ser.2, Vol.71, No.2, pp. I_1639-I_1644, 2015.
• S. Sato et al., "A Survey on the Audible Quality of Outdoor Public Address Speakers for the Disaster Reduction Broadcasting System in the Central Area of Ishinomaki City," IEICE Transactions on Communications, Special Section on Information and Communication Systems for Safe and Secure Life, Vol.E98-A, No.8, pp. 1671-1673, 2015.
• S. Sato and F. Imamura, "A Survey and Analysis on Publicity Activity of Recovery Progress Visualization in the 2011 Great East Japan Earthquake Disaster – Focus on Affected Local Governments in Miyagi Prefecture –," Journal of Social Safety Science, No.24, pp. 171-181, 2014.

Academic Societies & Scientific Organizations:
• Institute of Social Safety Science (ISSS)
• Japan Society for Natural Disaster Science (JSNDS)
• Japan Society of Civil Engineers (JSCE)
• Japan Society for Disaster Information Studies (JASDIS)
• Institute of Electronics, Information and Communication Engineers (IEICE)

Name:
Kazumasa Hanaoka

Affiliation:
Assistant Professor, Ph.D., International Research Institute of Disaster Science (IRIDeS), Tohoku University

Address:
Aoba 468-1, Aramaki, Aoba, Sendai 980-0845, Japan

Brief Career:
2009- Assistant Professor, Department of Geography, Ritsumeikan University
2012- Assistant Professor, International Research Institute of Disaster Science (IRIDeS), Tohoku University

Selected Publications:
• K. Hanaoka, O. Murao, and K. Sugiyasu, "Characteristics of Housing Damage and Human Suffering Caused by Typhoon Haiyan in Philippines: The Relationship between Disaster Damage and Vulnerability," Journal of Social Safety Science, No.24, pp. 1-9, 2014.
• K. Hanaoka, T. Nakaya, and T. Tabuchi, "Generating Large-scale and Health-related Synthetic Population Microdata at a Neighbourhood Level in Japan," Proceedings of GIScience 2014, 2014.
• K. Hanaoka et al., "Network-based Spatial Interpolation of Commuting Trajectories: Application of a University Commuting Management Project in Kyoto, Japan," Journal of Transport Geography, Vol.34, pp. 274-281, 2014.

Academic Societies & Scientific Organizations:
• Institute of Social Safety Science (ISSS)
• GIS Association of Japan (GISA)
• Association of Japanese Geographers (AJG)
• City Planning Institute of Japan (CPIJ)

Name:
Makoto Okumura

Affiliation:
Deputy Director, Professor, Dr. Eng., International Research Institute of Disaster Science (IRIDeS), Tohoku University

Address:
Aoba 468-1, Aramaki, Aoba, Sendai 980-0845, Japan

Brief Career:
1987- Research Associate, Dept. of Civil Engineering, Kyoto University
1995- Assistant Professor, Department of Civil and Environmental Engineering, Hiroshima University
2008- Professor, Center for Northeast Asian Studies, Tohoku University
2012- Professor, International Research Institute of Disaster Science (IRIDeS), Tohoku University

Selected Publications:
• H. Tirtom, M. Okumura, H. Yamaguchi, and R. Das, "Network Fortification Model for Intercity Passenger Transportation," Journal of the Eastern Asia Society for the Transport Studies, Vol.11, 2016 (in press).
• R. Das and M. Okumura, "Effect of Demand Distributions and Declining Urgency of a Perishable Item in Relief Ordering Policy," Journal of the Eastern Asia Society for the Transport Studies, Vol.11, 2016 (in press).
• M. Okumura and M. Tsukai, "Business Service Location with Spatially Stochastic Demands: Agglomeration Economies Generated by the Intersection of Costs and Localized Uncertain Demand" in C. Karlsson, B. Johansson, K. Kobayashi, and R. Stough (Eds.), "Knowledge, Innovation and Space," Edward Elgar Publishing, 2014.
• M. Okumura, M. Tsukai, and N. Takada, "A rank-size rule in a firm, produced from a hierarchical branch office location model," Review of Urban and Regional Development, Vol.22, Issue 2-3, pp. 73-88. 2010.

Academic Societies & Scientific Organizations:
• Japan Society of Civil Engineers (JSCE)
• City Planning Institute of Japan (CPIJ)
• Eastern Asia Society for Transportation Studies (EASTS)
• Applied Regional Science Conference (ARSC)
• Japan Society of Traffic Engineers (JSTE)

Name:
Shunichi Koshimura

Affiliation:
Professor, Dr. Eng., International Research Institute of Disaster Science (IRIDeS), Tohoku University

Address:
Aoba 468-1, Aramaki, Aoba, Sendai 980-0845, Japan

Brief Career:
2000- JSPS Research Fellows (PD)
2002- Research Scientist, Disaster Reduction and Human Renovation Institute
2005- Associate Professor, Graduate School of Engineering, Tohoku University
2012- Professor, International Research Institute of Disaster Science (IRIDeS), Tohoku University

Selected Publications:
• E. Mas, B. Adriano, and S. Koshimura, "An Integrated Simulation of Tsunami Hazard and Human Evacuation in La Punta, Peru," Journal of Disaster Research, Vol.8, No.2, pp. 285-295, 2013.
• S. Koshimura, T. Oie, H. Yanagisawa, and F. Imamura, "Developing fragility functions for tsunami damage estimation using numerical model and post-tsunami data from Banda Aceh, Indonesia," Coastal Engineering Journal, JSCE, Vol.51, No.3, pp. 243-273, 2009.
• S. Koshimura, Y. Namegaya, and H. Yanagisawa, "Tsunami Fragility – A new measure to assess tsunami damage," Journal of Disaster Research, Vol.4, No.6, pp. 479-488, 2009.

Academic Societies & Scientific Organizations:
• Japan Society of Civil Engineers (JSCE)
• Institute of Social Safety Science (ISSS)
• Japan Association for Earthquake Engineering (JAEE)
• Japan Society for Computational Engineering and Science (JSCES)
• American Geophysical Union (AGU)