Rescue Activity for the Great East Japan Earthquake Based on a Website that Extracts Rescue Requests from the Net

Shin Aida
Toyohashi University of
Technology
Aichi, Japan

aida@cs.tut.ac.jp

Yasutaka Shindo
Free
ring@quruli.ivory.ne.jp

Masao Utiyama
National Institute of
Information and
Communications Technology
Kyoto, Japan

mutiyama@nict.go.jp

Abstract

At the early phase of the Great East Japan Earthquake a vast number of tweets were made on Twitter. Even though many of them were calling for emergency rescue, they were not found timely due to the vast number of tweets including well-intentioned tweets to support those emergency rescues. In order to deal with the situation, the authors developed and launched a website on March 16, 2011, which automatically extracts rescue requests, categorizes similar statements into several statements and then lists them. This paper covers in detail not only the technology of the system but also how it has already collaborated and been applied to #99japan, a project to support delivering emergency rescue requests. Note that #99japan is an activity to monitor the process of the rescue based on Twitter and coming from the thread started by temporary volunteers who organized on a Japanese textboard 2 ちゃんねる "*2channel*."

1 Introduction

The Great East Japan Earthquake occurred on March 11, 2011 has obviously caused a wide range of damages. At the early phase of the earthquake, a large number of transmissions of information were made not only through mass-media in existence such as TV broadcast, newspaper and magazines but also social media such as Twitter. While the leading mass-media companies focused on information about the seriousness and damages due to the earthquake at affected areas for general public, the information which is useful for disaster survivor were also absolutely imperative. Under this situation, local radio broadcast, local newspaper and social media contributed to satisfy the needs.

From the early phase of the earthquake, a vast number of the tweets with the hashtag '#j_j_helpme', meaning requesting rescues, were made on Twitter, one of the most famous social media in Japan. These kinds of tweets included 【拡散希望】 "[Want to spread the information]", so that well-intentioned people who just wanted to contribute to help people highly tended to retweet those tweets unconditionally. As the result, similar rescue requests were flowed on Twitter and this made it very difficult to trace whether the rescue requests were actually reported to relevant authorities, which is the most important process.

In order to deal with the situation, we developed a website on March 15, 2011, which automatically extracts rescue requests, categorizes similar statements into several statements and then lists them and launched the website on the next day (Aida *et al.*, 2011; Aida *et al.*, 2013). This paper covers in detail not only the technology of the system but also how it has already collaborated and been applied to #99japan, a project to support delivering emergency rescue requests. Note that #99japan is an activity to monitor the process of the rescue based on Twitter and coming from the thread started by temporary volunteers who organized on a Japanese textboard 2 ちゃんねる "*2channel*".

2 Rescue Requests

2.1 Identification of Rescue Requests

We analysed the vast number of rescue requests in the early stages of the earthquake on Twitter; as a result, we identified the following four cases:

1. *Primary rescue request information*;

2. *Secondary rescue request information*, including redundant information;

3. *Non-rescue request information*; and

4. *Rescue completion report*.

We call information categorized in the case 1 or 2 *rescue requests* and one in the case 3 *a non-rescue request* respectively. Information categorized in the case 2 also includes a statement that contains the string 【拡散希望】 "[Want to spread the information]." Volunteers might not know whether or not each request was reported to authorities at each information; however, they still understood the current situation in the disaster area from a remote location by spread requests. Primary rescue requests on Twitter correspond to original tweets by some survivor and each secondary ones correspond to retweet, reply or mention.

On Twitter, it should be noted that statement is preceded by a time series back from the beginning of a sentence, where each quoted statement is postfix, and the primary rescue request is shown at the last place of the tweet in many cases:

```
user1:     a mention for user2; RT
@user2:  a mention for user3; RT
@user3 ... RT @usern:  a primary
rescue request.
```
posted at **MM/DD/YYYY hh:mm:ss**

2.2 Rescue Requests

In order to extract rescue requests on Twitter, we tried to find the words below that tended to be included in rescue requests, categorized them, and then summarized them as regular expressions in heuristics:

1. Rescue requests should include the following:

   - Words included in street addresses in order to identify the address of the rescue request (5 words).

   - Words related to the safety confirmation and to circumstances of life lines (21 words).
   Example: 消息 "whereabouts," 深刻 "serious," 要請 "request," 避難 "evacuation."

2. Non rescue requests should include the following:

   - Proper names that have been included in the past false rumours (14 words).
   Example: 花山村 "Hanayama village," which became Kurihara city merged with other surrounding municipalities in 2005, so it did not exist in 2011.

   - Official Twitter account names of the news media, because the media has almost already reported to relevant authorities (20 words).
   Examples: radio_rfc_japan, fct_staff, [Aa]sahi, FKSminpo, [Nn][Hh][Kk], nhk_seikatsu, i_jijicom_eqa, kahoku_shimpo, akt_akita_tv,   NTV, telebee_tnc,   NISHINIPPON, zakdesk, 781fm.

   - Specific person's names, such as celebrities and politicians, countries, party names and organization names, because there is no possibility of rescue requests including the thought and creed (9 words).
   Examples: 民主党 "Democratic Party of Japan," 自民党 "Liberal Democratic Party of Japan," 社民党 "Social Democratic Party," 共産党 "Communist Party,"

   - The names of countries and international organizations, because they are almost secondary information of the news media (6 words).
   Examples: アメリカ "USA," フランス "France," 国連 "United Nations," ユニセフ "UNICEF."

   - Words related to the nuclear accident, because rescue requests are not include scientific technical terminology (10 words).
   Examples: セシウム "cesium," ヨウ素 "iodine," ウラン "uranium," プルトニウ

ム "plutonium," 放射線 "radiation," 放射能 "radioactivity."

- Words unused in rescue requests (8 words).
  Examples: 笑 "laughing," 批判 "criticism," テロ "terrorism."

- Words included in tweets using too many hashtags, because they are independent of the original meaning of the tags (6 words).
  Examples: 予測市場 "forecasts market," リスクマネジメント情報 "risk management information."

3 Listing Policy for Rescue Requests

In the early stages of the earthquake, because we needed to release immediately our site extracting rescue requests, we decided to dare to volume display their requests on a single page in order for users not to look over information that they needed due to the info glut. Our site has been displayed 300 requests initially but now displayed 1000 requests at one time.

There were two reasons why we adopted the policy as listed below:

1. Extracted requests might include noise information. If we tightened filtering rules by our regular expression, it would be possible that the serious rescue requests would not be displayed.

2. By displaying rescue requests on the same page, volunteers could search requests in default feature of any web browser and, furthermore, avoid from checking multiple pages due to pagination.

Based on this policy, we manufacture a system extracting rescue requests by way of trial March 15, 2011, and launched the website on the next day (Aida *et al.*, 2011).

3.1 Extraction Algorithm of Rescue Request Information

Overview of the method of extraction algorithm of rescue request information on Twitter is as follows:

1. Obtain HTMLs on the tweet information including each search word listed in Figure 1 which is related to the earthquake disaster.

2. Perform the following process for all information obtained:

   (a) Merge tweets included in the HTMLs into the existing log file of the site and store the merged log file.

   (b) Preclude what appears to be non-rescue request information from the extracted log file by filtering according to the rules described in Section 2.2

   (c) Produce a *similar tweet key* by the following procedure based on the result of (b):

      i. Remove the longest string of up to "@" from the beginning of a tweet sentence.

      ii. Convert a similar tweet key into 15 letters obtained by removing the Japanese syllabary and ASCII characters.

   (d) Classify tweets into equivalent tweet classes, where we call the oldest tweet in equivalent class *represented tweet*, by using an associative array of similar tweet keys.

3. Make a list of latest 1000 items including the following lines and update the site:

   - Sentence of a represented rescue request tweet;

   - Tweet latest date and time of the tweet;

   - Tweet oldest date and time of the tweet:

   - The number of tweets in class including the tweet; and

   - Estimated source URL in the form "http://twitter.com/ screen_name/statuses/ tweet_ID").

4 Rescue Activity

In web-based rescue activities, volunteers needed to share situations about rescue activities. After our site opened, we scouted out for such activities to cooperate the site and really participated in a rescue activity #99japan organized by Kaichi Tamiya.

```
#j_j_helpme #j_i_helpme #hinan
#jishin #jisin #tunami
#311sppt #311care #311sien
#itaisousaku 99japan #anpi
#aitai #Funbaro #hope4japan
#prayforjapan #ganbappe
#save_busshi #save_volunteer
#save_gienkin #save_kids
#saigai #shinsai #tasukeai
#fukkou #fukko #save_miyagi
#save_fukushima #save_iwate
#save_aomori #save_ibaraki
#save_chiba #save_nagano
#save_sendai #save_ishinomaki
#save_iwaki #ishinomaki
#shiogama #rikuzentakata 緊
急地震 余震 火事 怪我 負傷者自宅避難
避難所 孤立 餓死 緊急+救助 食料+不足
物資+不足 食糧+不足 救援 支援 安否
消息 栄村 陸前高田釜石 大船渡 気仙沼
南三陸 歌津 志津川 石巻 松島 亘理 山
元 相馬 いわき 飯舘
```

Figure 1: Search Words to Extract Rescue Requests

4.1 Situation on the Net of In the Early Stages of the Great East Japan Earthquake

In the early stages of the earthquake, many activities have been launched in the Net. Disaster information has been made public in portal sites such as Yahoo! and Google.

In particular, "Google Person Finder" was well-known safety as confirmation system, which was used from the Haiti earthquake of January 2010 (Google, 2011). "ANPI NLP" was a project launched by voluntary researchers to augment the safety information of the system from the Net in a method of natural language processing (Murakami *et al.*, 2011; Neubig *et al.*, 2011).

As a system including rescue operations, "sinsai.info" was also well-known (Seki *et al.*, 2011). The site was constructed using a crowdsourcing tool *Ushahidi*, which was used in many disasters since 2007 and was famous as a system built on the day of the earthquake. Also, a variety of systems were published such as sites displaying a time line of tweets related to the earthquake tags on Twitter,

radiation dose maps and so on.

However, countless volunteers and systems across the country could not quickly cooperate necessarily in the early stage; rather, utilizing existing systems, volunteers had decided the rules by trial and error and was involved in rescue activities.

In textboard *2channel*, the thread on rescue activities (*2channel*, 2011) was opened immediately in the day of the earthquake, Some anonymous volunteers participated in the thread launched a Wiki (*2channel* ID:nx64KwTT, 2011) and a rescue map (#99japan, 2011a) on the earthquake.

4.2 Rescue Assistance Project for The Great East Japan Earthquake: #99japan

March 15, 2011, Kaichi Tamiya (Twitter ID: @ktamiya) was managing a report activity history by using the comments section of his blog (Tamiya, 2011a). March 18, inviting members our rescue support activities taking advantage followers on Twitter, he organized *'the rescue assistance project for The Great East Japan Earthquake #99japan"* (Tamiya, 2011b).

Overview of this project activity was as follows:

- 目的: Twitter 等での救助要請の声を適正機関に伝達する支援を行い，被災者を救う.

  Purpose: We notify agencies about the rescue requests obtained from the Net, such as Twitter, to save the survivors.

- 活動内容: 主に被災者の情報の整理，内容確認，アドバイス，救助要請の代行. 期間は物資が行き渡り，復興段階に入る頃までを予定.

  Activities: We report of rescue requests, verify the accuracy of request information and advice for survivors, until request supplies prevail in disaster-stricken areas and the areas enter the first stages of recovery.

This project adopted an existing editing rule for the rescue request map (#99japan, 2011a) and a supply request map (#99japan, 2011b), which were based on Google maps. In addition administrators of the maps also participated in #99japan, so it was notable that #99japan was one of the rescue projects for survivors in the early phase of the Great East Japan Earthquake.

#99japan was also a *hashtag on Twitter* as well as a project name. Any member of

`#99japan` could tweet many rescue requests and related activities appended with `#99japan` on Twitter, so their tweets shared among members. Of course, their tweets were also visible in the non-members. It was the biggest feature to communicate among users by the hashtag `#99japan` in the Twitter-based activities.

We proposed using our site to volunteers of `#99japan` to find rescue requests on the Net efficiently.

4.3 Activity Flow

We show below the activity flow of information sharing, map editing and source verification in `#99japan` (`@ma_ chiman`, 2011):

- Information sharing and map editing:

  1. Member of the project tweets rescue request information with a street address obtained by our site or other information sources, to which are appended the string "`#99japan`" as a hashtag in order to facilitate to share information to the members in a short period of time.

  2. According to of the shared information, member selects a point of the street address on the map and reports to relevant organizations such as the police, where the point is classified into
     (a) not reported;
     (b) reported but unknown whether resolved;
     (c) resolved already; or
     (d) other.

  3. Member fills in the content of report to the pop-up point in the map in a description rules.

- Source verification work:

  1. To update the local information that is constant number of days elapsed from the time of the rescue request, member follow the information providers and ask the change of environment.

  2. Member contacts the neighbourhood residents, asks their situation and writes a checking status to the map.

  3. Depending on the content and the presence or absence of reply, member reflects the following information on the pop-up of the point at the map:

(a) resolved;
(b) there is no new information;
(c) not contact; or
(d) other information,

which are added the modified date and Twitter ID of reporting member.

4.4 Cooperative Rescue Activities with `#99japan` and Our Site

We continued to improve the site after the date as March 20 to conform to the needs of members.

We show a line graph about the number of accesses to the site as Figure 2. Analyzing the graph,
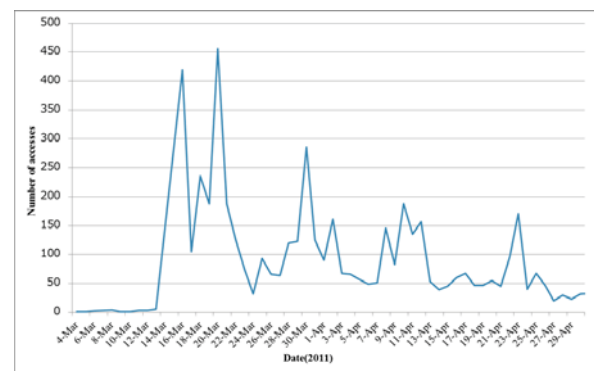

Figure 2: Changes in the Number of Accesses Our Site Extracting Rescue Requests.

we have confirmed the following trends:

- The number of accesses was particularly high activity immediately after the start of March 20 and the release date.

- Early April emerged from an emergency state, the number of accesses was increased temporarily.

- When large afterquakes was invoked in Miyagi April 4 and Fukushima of April 11 and 12 respectively, the number of accesses was increased again.

There was no quantitative data to show whether our site was helpful. Because most of the project members at the time is the anonymous participant, it is too difficult for our site to survey efficiency of rescue. However, we got good evaluated tweets on Twitter by the project `#99japan` chair Tamiya, `@ma_chiman` who opened the official website (`@ma_ chiman`, 2011) of the project, and the rescue map (`#99japan`, 2011a) administrator `@juntaro33` as follows:

- Immediately after organizing `#99japan`, `@ma_chiman` suggested that the project members used our site to find rescue requests. (March 20, 2011.)

- When we implemented several additional functions into our site, `@ma_chiman` and `@juntaro33` admire our implementation. (March 25, 2011.)

- `@juntaro33` said that he was using our site as the most useful now. (April 2, 2011.)

- `@juntaro33` said that most of the information on the map were obtained from our site. (April 6, 2011.)

By the project activity used the rescue map and the supply request map (Figure 3), The project had reporting and supporting activities more than 200 points on the rescue and supply request maps, respectively, in about three weeks until early April.
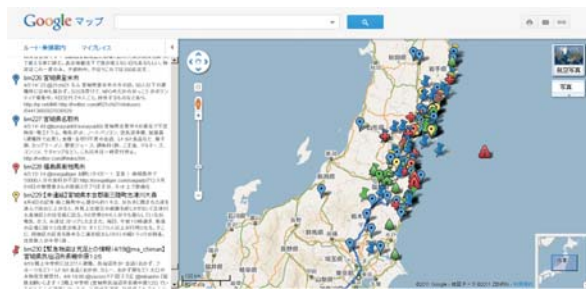

Figure 3: The Supply Request Map of `#99japan`.

5 Consideration for Our Actively

Looking back on `#99japan`, it was important to have *2channel* textboard as a virtual place to gather in case of large earthquake existed. In fact, *2channel* users became volunteers and shared disaster information such as rescue requests on the day of the earthquake on several social media, such as Twitter, mixi which was a famous SNS in Japan, Wiki sites and so on.

Then, Twitter users and others naturally joined the `#99japan` and they could share more rescue requests on Twitter. In particular, utilizing "#99japan" as a hashtag was a very significant as a mechanism that could be shared rescue progress and completion report. Importance of progress and completion report like this has been pointed out by many researchers on Twitter (Yamazaki *et al.*, 2012).

In the rescue activities, "freshness of the rescue request information" was particularly important; our site had contributed to the report activities support efficient, utilized as a source of information sources to `#99japan`.

6 Conclusion and Future Work

We developed and launched a website which extracted and listed rescue requests among all the information on Twitter on March 16, 2011, which was right after the Great East Japan Earthquake.

Through participating and collaborating with the activities of `#99japan`, a relief project of the Great East Japan Earthquake based on Twitter, it has turned out that exchanging messages to find appropriate information, report and check the status of the process on a timely basis based on the information of rescue requests listed on the website.

According to the requests of the members of `#99japan`, we also made efforts in order to increase the precision of extraction of information of rescue requests and to improve functions of the website for that. It is reported that social media was taken advantage of in backup activities at the early phase after the earthquake.

Further research could be analyzing the needs among the log file of the website which is still processing at present, and creating a well adaptive system for the disaster recovery relief.

It should be also noted that fortunately several important factors got connected by chance in `#99japan`. However it is also important to refine a framework of social system for letting volunteers work effectively and rapidly at disasters and at following restriction activities now in normal time.

References

Shin Aida, Yasutaka Shindo, and Masao Utiyama. 2011. *The automatically listing site extracted similar rescue requests.* http://www.selab.cs.tut.ac.jp/~aida/

Shin Aida, Yasutaka Shindo, and Masao Utiyama. 2013. *Rescue Activity for the Great East Japan Earthquake Based on a Website that Extracts Rescue Requests from the Net. Journal of Natural Language Processing*, 20(3):405–422.

Google. 2011. *More resources for those affected by the Japan earthquake and tsunami. The Official google.org blog.* http://blog.google.org/2011/03/more-resources-for-those-affected-by.html

Hiroaki Kuze. 2010. http://xtter.openlaszlo-ason.com/XTTER/1500ttr/

Koji Murakami *et al.* 2011. *ANPI NLP.* http://trans-aid.jp/ANPI_NLP/

Graham Neubig, Yuichiroh Matsubayashi, Masato Hagiwara, and Koji Murakami. 2011. *Safety Information Mining — What can NLP do in a disaster —. Proceedings of the 5th International Joint Conference on Natural Language Processing (IJCNLP)*, 965–973.

Haruyuki Seki *et al.* 2011. *sinsai.info.* http://www.sinsai.info/

Kaichi Tamiya. 2011. Tamiya's Blog: 東北関東大震災，救助依頼連絡先 *("Rescue Request Contact for The Great East Japan Earthquake").* http://ameblo.jp/ktamiya/entry-10829792004.html

Kaichi Tamiya. 2011. 東北関東大震災の救助支援プロジェクト「#99japan」参加者を募集します *("We Invite the Participants of the Rescue Assistance Project* #99japan *in the The Great East Japan Earthquake").* http://twipla.jp/events/6133

Fumi Yamazaki *et al.*. 2012. #shinsaidata 「東日本大震災ビッグデータワークショップ — *Project 311 —*」ブレスト *("The Big Data Workshop of The Great East Japan Earthquake — Project 311 —: A Brainstorming on Twitter").* http://togetter.com/li/372103

2channel users. 2011. 【私にも】三陸沖地震災害の情報支援【できる】*("Information support of The Great East Japan Earthquake Disaster").* http://logsoku.com/thread/hayabusa.2ch.net/eq/1299829654/

2channel ID:nx64KwTT user. 2011. 東北大震災 まとめ *Wiki("The Great East Japan Earthquake Summary Wiki").* http://www45.atwiki.jp/acuser001

#99japan. 2011. 共同編集*:* 被害リアルマップ東北地方太平洋沖地震 *("Rescue Request Map for the Great East Japan Earthquake Based on Google Maps").* https://maps.google.co.jp/maps/ms?ie=UTF8&hl=ja&msa=0&ll=38.255436,140.998535&spn=10.259815,16.54541&z=6&msid=212756209350684899471.00049f93fb04a48b1dce9

#99japan. 2011. 共同編集*:* 物資 要請・提供マップ東北関東大震災 *("Supply Request Map for the Great East Japan Earthquake Based on Google Maps").* https://maps.google.co.jp/maps/ms?ie=UTF8&hl=ja&msa=0&msid=212756209350684899471.00049ea27cf60c4292136&ll=37.827141,140.306396&spn=2.290855,3.488159

@ma_chiman. 2011. 東日本震災支援 #99japan（救急ジャパン）公式サイト *("Rescue Assistance Project for The Great East Japan Earthquake* #99japan*: Official Site").* https://sites.google.com/site/sharp99japan/