Inferring Intentions of Twitter Users to Visit Places

Chengcheng Zhang, Haruhiko Sato, Satoshi Oyama, Masahito Kurihara

***Abstract*—A method is proposed that uses machine learning to infer the intentions of Twitter users to visit places or attend events. It distinguishes similar but different intentions such as “want to go” and “plan to go.” We discuss several research challenges in preparing the training data, extracting useful features, and building accurate classification models. We also give experiment results indicating that our method works well, and we can use it to infer the popularity and crowdedness of the places to help users make travel plans.**

***Index Terms*—Twitter, intention, event, spatio-temporal data, machine learning.**

I. Introduction

Users of the Twitter micro-blogging service tweet about various aspects of their real-world experiences such as visiting places and attending events. Extracting such information from Twitter is important for estimating the popularity of places and events for use in navigation systems like that shown in Fig 1 and so on.

Fig 1. Mockup of navigation system

Several studies have focused on extracting information from Twitter and other blogs. Chakrabarti and Punera investigated event summarization using tweets [1] as it is impossible for a person to read all of the tweets related to a topic of interest. The method they proposed for summarizing tweets gives the user a better visualization. Recent research has shown that Twitter can be used to detect event occurrences such as earthquakes because Twitter users tend to report on events that they are experiencing [2]. Among the various Twitter and blog-related studies, ones focused on extracting real-world experiences have been one of the most common. Tezuka et al. developed a system that extracts verbs that refer to actions and eliminates the ones indicating movement in order to create a blog map of experiences [3]. Kurashima et al. developed a local information search system that enables the user to specify a location, a time period, or a type of experience in a search query and find relevant Web content [4].

These studies mainly focused on past information such as reports from people who had actually visited certain places or attended certain events. Such information is useful for estimating the popularity of places or events in the past or present. However, future information is more useful for making predictions or plans. Baeza-Yates’s addressed the problem of “future retrieval” and the challenges it faces [5]. Jatowt et al. speculated on the analysis of future-related information in the news and the use of such information for event prediction [6]. Other researchers have used sentiment analysis to predict product sales [7,8] and have analyzed emotional perception of the future [9]. However estimating the popularity of places and events also needs future information, which can be gleaned from tweets and other blog postings. Such information would be useful for making travel plans and for preparing for visitors. Twitter users frequently tweet about their future plans as well as their ongoing activities. We investigated the extraction of intentions to visit places and attend events from tweets.

We focused on detecting three types of intentions related to visits: "Want to go," "Plan to go," and "Have been there before." Tweets related to a certain place or event were collected from Twitter, and each tweet was categorized as belonging to one of the three types by using SVM-based classifiers [10,11].

[1]Graduate School of Information Science and Technology Hokkaido University

{zhangchengcheng, haru, oyama, kurihara}@complex.ist.hokudai.ac.jp

II. Proposed Method

In our proposed method, tweets are first collected from Twitter using the Twitter4j and then labeled by crowdsourcing. Useful features are then extracted from the text, and SVM-light (http://www.svmlight.joachims.org) is used to categorize the tweets by type of intentions.

A. *Collecting tweets from Twitter*

The function search of Twitter4j is used to extract data that match the input query, which is the name of a specific place or event, such as “Sapporo Snow Festival.” The use of Twitter4j enables not only the collection of tweets but also the collection of re-tweets, the tweet creation times, and so on. We collected the tweets posted from the day before the event to the day after the event because it enables us to more clearly see the trends of different types of intentions over the course of the relevant period.

ISBN: 978-988-19252-5-1
ISSN: 2078-0958 (Print); ISSN: 2078-0966 (Online)

B. Labeling data by crowdsourcing

The need to obtain objective and reliable labels has led to increasing attention being paid to crowdsourcing. Crowdsourcing services can collect a large amount of labeled data.

In our research, we used the data for "Sapporo Snow Festival" on February 4 as our dataset. Because there was much data and possibly ambiguous expressions in the tweets, we used the Lancers (http://www.lancers.jp) crowdsourcing service to label the data.

Among the crowd workers, some are highly skilled while some are not. The highly skilled ones give more reliable labels while the lesser skilled ones give variable quality labels. To get the gold standard, we used majority voting. We first posted the task to five workers and obtained labeled data. Then we got the gold standard by majority voting, which means that if three or more workers give the tweet the same label, we treated the label as the gold standard. We calculated the distribution of data labeled the same by three, four, and five workers. The percentages were 0.37, 0.30, and 0.26.

This means that 93% of the data were given the same label three or more times, indicating that data labeled by crowdsourcing are reliable. Because many expressions in the tweets were ambiguous, the more workers who gave the same label, the more reliable the data were. To obtain more reliable data, we also extracted the data that four workers give the same label and that five workers give the same label.

C. Extracting useful features to construct feature vectors

Tweets have three distinctive characteristics: short text, lots of mistakes, unknown new words. These characteristics make it difficult to infer the tweeter's intention simply by using the term frequency of the words in the text. We thus extract five features from the text and use the various combinations of the features and the term frequency of the words in the text to construct feature vectors.

We found that several things can be used to infer the intentions: the sentence tense, the sentence pattern, the time expression, and a turning conjunction (which indicates a turning point in the sentence) such as "けど (but)."

The tense of the sentence can reflect the writer's intention. In English, the focus is on the form of the verb while in Japanese the focus is on the auxiliary words. When the text contains more than one auxiliary word such as "た," we can assume that the tweeter has most likely been to the place because the sentence is in the past tense.

People always use similar expressions or patterns to show their intention of "want to go" or "plan to go". We have analyzed lots of tweet and conclude some expressions of the users shown in Table I.

There is a method that extracts all tweets with this kind of expressions using the string matching. However there will be some problems occurs. First is that it is almost impossible for us to conclude all the expressions. In English, there may be only one phrase that expresses the intention of "want to go," while in Japanese, there are many such phrases. The other is that the format of the expressions can be more than one kind. For example, "行きたい" and "いきたい", they are the same in the reading while not same in the writing.

In our research, I did not use the expressions themselves, but concluded some patterns that can be instead of the real ones. The method is shown in Table II. If a sentence has one of these patterns, we can infer that the user want to go to a place or will go to the place.

Table I. Example patterns

	Want to go	Plan to go
Patterns	行きたい 行ってみたい 行って見よう 見たい 見て見たい 行きてー (Want to go)	行ってきます 行きます 見てきます 見てこよう 見にいく 行ってくる 見てくる 行こう (Plan to go)

Table II. Patterns of "want to go" and "plan to go"

	Want to go	Plan to go
Patterns	Verb + "たい"	Verb + "ます"
	Verb + "う"	Verb basic form

However, some patterns can reflect both "want to go" and "plan to go." In such cases, a time expression in the tweet can help clarify the intention. If the tweeter is planning to go, he/she may include a time-related expression like "tomorrow" or "next Monday." Moreover, time expressions like "yesterday" can help clarify the tense of the tweet and thereby help clarify the intention of "Have been there before." The Table III showed to you the time expressions that often occurred in the tweets.

The same challenge as the one of the first feature, it is difficult to get all kinds of time expressions, so I used a method to extract the time expression by using the speech of the word. In English, the time expressions are almost nouns. In Japanese, the time expression is not only the nouns but also can be adverbs. So after the morphological analysis the word of time expression will have two kinds of speech, which I can use to distinguish the time expression from normal nouns.

At last, it is also important to focus on the turning conjunctions, which can change the meaning of the sentence easily. I want to judge whether the sentence have the turning conjunctions or not to realize the real meaning of the tweets.

Table III. Time expressions

	Past	Future
Time Expressions	昨日 (yesterday) おととい (the day before yesterday) 先週 (last week) 先月 (last month) 去年 (last year)	明日 (tomorrow) 明後日 (the day after tomorrow) 来週 (next week) 来月(next month) 来年 (next year)

The resulting feature vector is in the form

$$V_i = (Tense, Exp1, Exp2, Time, Conj, (tf_1, \ldots, tf_n)),$$

where i is the ID of the tweet. If the tweet includes the auxiliary word "た," *Tense* equals 1. If not, it equals 0. If the tweet includes a pattern that shows the intention of "want to go," *Exp1* equals 1, and if it includes a pattern that shows the intention of "plan to go," *Exp2* equals 1. Otherwise, they equal 0. *Time* is the time expression in the tweet. If the time is a future time, *Time* equals 1. If it is a past time, *Time* equals −1. If there is no time expression, *Time* equals 0. If there is a turning conjunction in the tweet, *Conj* equals 1. If not, it equals 0. The tf_n is the term frequency of words in the text.

D. Using SVMs to classify intentions

We can confider two types of classification models: place/event-dependent classification models and place/event-independent classification models. The former type are used to infer the intention to visit known places or attend known events, while the latter type are used to infer the intention to attend new and/or unknown events.

Here we focus on the place/event-dependent models. A place/event-dependent model is built by training an SVM for each place/event using the training data collected for the place/event. To build a place/event-independent model, a common SVM is trained using training data collected for different places and events.

III. Evaluation

We conducted three experiments to evaluate the performance of our method. We used different models to classify the different kinds of information. We used SVM-Light to do the classification. It provides four kernel functions to the user: a linear function, a polynomial function, a radial basis function, and a sigmoid function. Because we did not know the spatial distribution of the data, we used all four and the linear kernel gave the best result. Thus we describe the results by the linear kernel in the followings.

The labeled data came from the data of Sapporo Snow Festival at February 4th and February 5th, 2013. We used 2-cross-validation. Our objective was to infer three kinds of information: "want to go," "plan to go," and "have been there before." Table IV shows the distribution of information for data at February 4th. The "4 or 5" means that 4 or 5 workers gave the same information label to the tweet. Table V shows the distribution of information for data at February 4th.

Table IV. Distribution for data at February 4th

	Want to go	Plan to go	Have been there before	Else
3 or more workers	417	215	189	508
4 or more workers	311	156	119	318
5 workers	222	89	64	150

The first experiment was on inferring the information of "want to go." Three models were used to evaluate the performance of our proposed method. Model 1 used the tweets for which three or more workers gave the same label as the positive data and the others as the negative data. Model 2 used the tweets for which four or more workers gave the same label as the positive data and the others as the negative data. Model 3 used the tweets for which five workers gave the same label as the positive data and the others as the negative data. We also used three kinds of test data, three or more workers gave the same label (Data 1), four or more workers gave the same label (Data 2) and five workers gave the same label (Data 3). The precision and recall are plotted in Fig 2

Table V. Distribution for data at February 5th

	Want to go	Plan to go	Have been there before	Else
3 or more workers	472	125	195	883
4 or more workers	403	89	137	693
5 workers	313	51	68	382

The second experiment was on inferring the information of "plan to go" by again using three models and the procedure described above. The precision and recall are plotted in Fig 3.

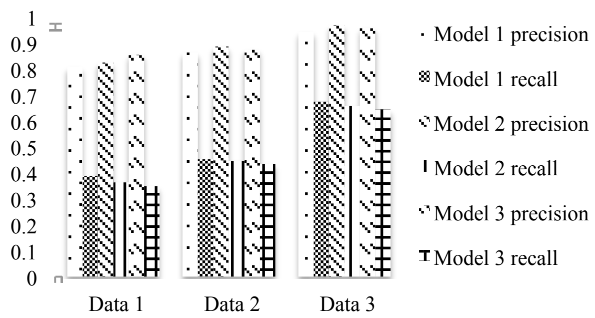

Fig 2. Comparison of results for "want to go"

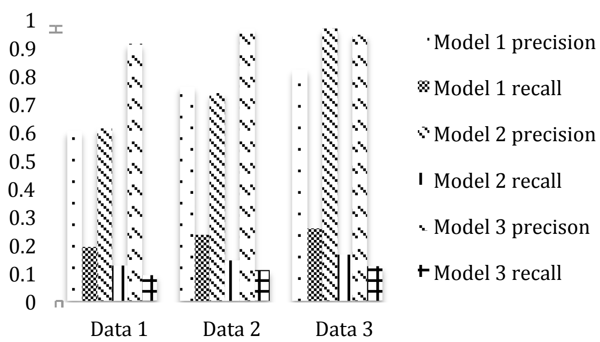

Fig 3. Comparison of results for "plan to go"

The final experiment was on inferring the information of "Have been there before," which was performed as described above. The precision and recall are plotted in Fig 4.

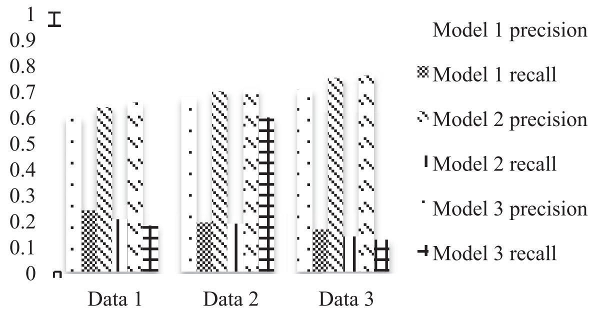

Fig 4. Comparison of results for “have been there before”

As shown in Fig 2, the precision for “want to go” was greater than 80% and the recall was greater than 30% for all combinations of training datasets and test datasets. First, comparing the results by training dataset, we see that recall was the highest and precision was the lowest with the Model 1. In contrast, precision was the highest and recall was the lowest with the Model 3. Next, comparing the results by test dataset, we see that both precision and recall were the highest with Data 3. We attribute this to the data of 5 workers being the most reliable and also to the amount of data being the least. Nevertheless, the precision was still as high as 97% with recall of 69%.

As shown in Fig 3, the results for “plan to go” were similar. Recall increased with the amount of training data while precision increased with the degree of test data reliability. However, recall was low for every combination of training data and test data. This was due to the lack of training data.

As shown in Fig 4, the results for “Have been there before” were almost the same as for “plan to go.” The precision was good but recall was not. Looking at the amount of data available for the three types of intentions, we see that the amount for “have been there before” is the smallest. Therefore, the precision and recall were the lowest.

In conclusion, our proposed method performed well for inferring “want to go.” The precision was 97% with recall of 69%. For the other two types of information, the precision was as high as 85%, but the recall was around 15%. Higher recall could be obtained by using more training data, and higher precision could be obtained by using more reliable data.

IV. Future Work

There are several research challenges from data collected to implementing the proposed method. So far, we have focused on finding useful features in the texts of tweets. We could also use information about the users taken from their profiles. We plan to evaluate the use of such information.

The final goal of our research is to construct a navigation system. We have collected the data of “Sapporo Snow Festival” from February 4^{th} to February 13^{th}. Our model for inferring “want to go” works well on the data of February 4^{th}, and we can use this model to do the classification of the other days. The change in the percentage of tweeters who “want to go” is easily understood from Fig 5. The number started off relatively high and then decreased during the course of the event. The increase on February 7 is attributed to February 8 being a Friday, which means that more people can go to the festival on Friday evening or over the weekend. This is just like the trend seen in restaurants, i.e., always busy on Friday night and the weekend.

To better model such information, we need to label data for other days such as the day before the end of the event. We then might be able to get more useful data for inferring “plan to go” and “Have been there before.” By obtaining good models for different types of information, we should be able to complete the mockup shown in Fig 1.

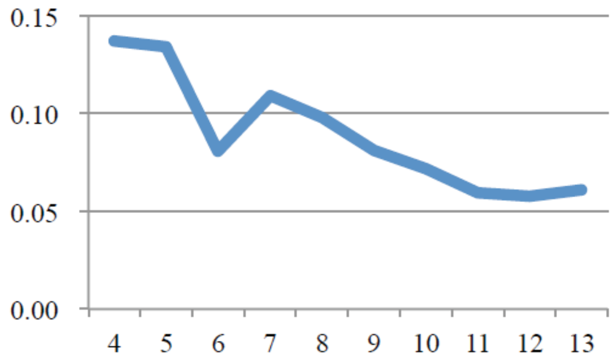

Fig 5. Trend for “want to go” during the Sapporo Snow Festival

V. Conclusion

Our proposed method uses machine learning to infer the intention to visit places or attend events from data collected from Twitter. Our model for inferring “want to go” works well. Our models for “plan to go” and “have been there before” have high precision but low recall due to the lack of training data.

Acknowledgment

This research is supported in part by a grant from the Artificial Intelligence Research Promotion Foundation.

References

[1] Chakrabarti D, Punera K. Event Summarization Using Tweets, ICWSM 2011.

[2] Sakaki T, Okazaki M, Matsuo Y. Earthquake Shakes Twitter Users: Real-time Event Detection by Social sensors, WWW 2010.

[3] Tezuka T, Kurashima T, Tanaka K. Toward Tighter Integration of Web Search with a Geographic Information System, WWW 2006.

[4] Kurashima, T., Tezuka, T. and Tanaka, K. Mining and Visualization of Visitor Experiences from Urban Blogs, DEXA 2006.

[5] Baeza-Yates R. Searching the Future, SIGIR Workshop MF/IR 2005.

[6] Jatowt, Adam, et al. Supporting Analysis of Future-related Information in News Archives and the Web, JCDL 2009.

[7] Mishne G, Glance N S. Predicting Movie Sales from Blogger Sentiment, AAAI Spring Symposium: Computational Approaches to Analyzing Weblogs 2006.

[8] Liu Y, Huang X, An A, et al. ARSA: A Sentiment-aware Model for Predicting Sales Performance Using Blogs, SIGIR 2007.

[9] Pepe A, Bollen J. Between Conjecture and Memento: Shaping A Collective Emotional Perception of the Future, AAAI Spring Symposium: Emotion, Personality, and Social Behavior 2008.

[10] Boser B E, Guyon I M, Vapnik V N. A Training Algorithm for Optimal Margin Classifiers, COLT 1992.