Regular Paper

DBSJ Journal, Vol.13, No.1
March 2015

Cross-lingual Investigation of User Evaluations for Global Restaurants

Jiawen LE ♥ Hayato YAMANA♦

Twitter, as one of the popular social network services, is now widely used to query public opinions. In this paper, tweets, along with the reviews collected from review websites are used to carry out sentimental analysis, so as to figure out the language-based and location-based effects on user evaluations for six global restaurants. The language expansion is carried out that 34 languages are taken into account. By using a range of new and standard features, a series of classifiers are trained and applied in the later steps of sentiment analysis. Our experimental results show that the location and language effects on user evaluations for restaurants actually exist.

1. Introduction

In recent years, social network service (SNS), a newcomer in the field of social media, has drawn much attention all around the world. Twitter, one of the popular social network services, owns a range of special characteristics that contribute to its huge success. It allows users to express their opinions and share the updated information with each other. Because of the limited length of 140 words per tweet, users feel more free and lighthearted to send a new message, which consequently leads to the tremendous amount of information and the rapid speed of distribution.

Also, in recent years, there has been remarkable progress of globalization. With the increase of the number of transnational enterprises, people from all over the world can use the same product, get the same service, and savor the same meal. However, it is quite common that people from different countries may have totally different feelings about these products, service or meals, probably mostly due to their diverse backgrounds. So, if we are able to analyze such differences by using tweets, we may advance such social science researches.

In this paper, we pick up global restaurants as an example, such as Burger King, Mcdonald's, KFC, Pizza Hut, Subway, and Starbucks to investigate how we can analyze such difference among the world's people. Our contributions consist of the following three points:

- A language expansion has been carried out that tweets written in 34 languages are treated as the research subject.
- Modifications have been made to increase the accuracy of sentiment analysis.
- We have shown that the location and language effects on user evaluations for restaurants actually exist.

The rest of this paper is organized as follows. Section 2 lists prior works. Section 3 discusses the concrete methods and algorithms adopted in this paper. Section 4 describes the experiment and results followed by conclusions in Section 5.

2. Related Work

The sentiment analysis for tweets has been focused by many researchers, and there exist a range of significant works.

In the aspect of opinion mining, a noted work is presented by Pang and Lee [1], which gives a broad view of existing approaches for sentiment analysis and opinion retrieval. Moreover, Liu [2] reviews the methods and works in the field of sentiment analysis of recent years.

Go et al. [3] shows that SVM outperforms other models, such as Naive Bayes and MaxEnt, when classifying tweets into negative ones and positive ones according to the words related to sentiment. They also concluded that unigram feature model had the best performance, which could not be gained by using bigrams and parts-of-speech feature models. Pak and Paroubek [4] characterizes tweets such that tweets have a tendency to include several sentences in the popular newspapers that have no special sentiment polarity. In contrast with the conclusion of Go et al., Pak and Paroubek reported that n-gram and POS strategies both made contributions to increase precision. On the other side, the research of Barbosa and Feng [5] mainly focused on the syntax features such as hashtags, URL links, and exclamations, and then made a combination with the POS model. All the above-mentioned papers only used the common English written tweets into consideration, and did not touch upon the cross-lingual investigations.

As for the cross-lingual sentiment analysis, "Oasys opinion analysis system" presented by Cesarano et al. [6] allows the user to observe the difference of the intensity of opinion over countries. Guo et al. [7] constructed a text mining system to detect the different sentiment in the web texts written in different languages. Cui et al. [8] used emotion tokens to solve the problem of cross-lingual sentiment analysis. Gao et al. [9] analyzed Twitter and the Chinese version of Twitter - Sina Weibo, and compared several different aspects, such as the characteristics of user behaviors and the content of messages by adopting simple statistical model.

♥ Non Member Graduate School of Fundamental Sci. and Eng., Waseda University
lejiawen818@yama.info.waseda.ac.jp

♦ Member Faculty of Sci. and Eng., Waseda University
yamana@waseda.jp

Compared to all these works, this paper focuses on the analysis of cross-lingual user evaluations, which has not been done until now.

3. Methodology

3.1 Overview

The overview of our methodology is shown in Figure 1. After collecting tweets and review-data, the location definition step is carried out followed by constructing two dictionaries manually based on the datasets and online dictionaries. Then three main classifiers are trained and used to classify the tweets. Finally, based on the classification results and the data analysis results, the cross-lingual effect on evaluations is clarified out.

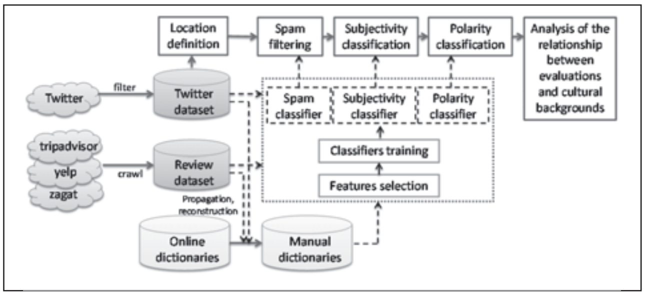

Figure 1: The main flow of our method

3.2 Data Collection

Two kinds of dataset are constructed: tweet-dataset and review-dataset.

As for the tweet-dataset, 9,523,211 restaurant-related tweets were gathered in 4 months (from Sep. 2013 to Dec. 2013), by using the Streaming API and Search API of Twitter. All the tweets are restricted by the names of target restaurants that are translated into multi-languages.

Then, as an auxiliary dataset, the review-dataset is constructed by gathering the English-written reviews from the popular review websites[1,2,3], including the text comments and their corresponding scores. Totally 55,031 reviews were collected.

3.3 Translation and Pre-filtering

In this research, 34 language codes used in Twitter, i.e., en, es, id, ja, fr, pt, tl, ru, tr, zh, ar, th, et, nl, it, de, ko, bg, sv, pl, vi, sk, da, ht, lt, lv, sl, fi, is, no, fa, hu, el, uk, are taken as the target languages. The selection of the target languages is based on tweet amounts, language populations, and whether can be translated by machine translation tools. Due to the complexity of tweet texts, and also the translation performances of the machine translation tools for certain languages, a part of the tweet-dataset could not be correctly translated into English, and were then discarded. Here, validation of correct translation is open question, but we assume that each word itself has been translated correctly.

The remaining data is then filtered by the pre-defined condition of having the relationship with restaurants, which is restricted by a list of restaurant related words shown in Table 1. These words are obtained by calculating the occurrence frequency of each word in the set of text reviews in the review-dataset followed by selecting out them from descending order of their frequencies. Here, 45 most frequent words were selected to filter the original tweets.

Table 1: Restaurant related words

restaurant, restaurants, food, foods, drink, drinks, dinner, dinners, lunch, lunches, breakfast, breakfasts, club, clubs, bar ,bars, pizza, pizzas, burger, burgers, coffee, cafe, cheese, grill, sushi , yelp, taco, steak, fry, fried, bbq, bakery, baked, yummy, yum, tasty, taste, tastes, delicious, eat, ate, eaten, eating, meal, meals

3.4 Location Definition

Location of a tweet is defined by using 'coordinates' item and 'geo' item in the user profile.

First, a manually constructed location name dictionary[4] is used to query the names of counties or cities appeared in the above-mentioned location-related items of tweets. Second, if we cannot identify the location of tweets by the above technique, Yahoo yql API[5] is used to parse these items of the undefined tweets. After these two steps, the ratio of the tweets labeled with location names was 72.8%.

3.5 SPAM Filtering

Because of the fact that the tweet-dataset contains so many undesired spam tweets, a filtering step to discard the obvious spam tweets is necessary. Strictly, whether a tweet is spam or not in this research should depend on whether the content of the tweet contains some useful information to indicate the subjective opinions towards the restaurants. In this research; however, a simple spam filtering technique is applied.

Firstly, advertisements and pure 'check-in' tweets (e.g., *'I'm at Burger King http://t.co/gWSyUMLD'*) are regarded as 'spam'. In addition, tweets posted in a certain short time period which have exactly the same contents are also considered as 'spam'.

A Bayesian classifier is used here, because Bayesian classification is usually robust to the noisy information. The training features include the number of the followers and friends of the user, the ratio of the number of followers and friends, the date of the registration, average number of new friends and followers per day, the latest 20 posted tweets, and also some syntax characteristics. 1000 and 200 manual labeled tweets are taken as the training-dataset and test-dataset respectively, and the performance of the 'spam' classifier turns out to be of an accuracy of

[1] http://www.tripadvisor.com/
[2] http://www.yelp.com/
[3] http://www.zagat.com/

[4] The country names and codes in the dictionary are collected from http://en.wikipedia.org/wiki/ISO_3166-1/.
[5] http://developer.yahoo.com/yql/

97.8%. This trained classifier is then applied to the whole dataset to filter out the 'spam' tweets.

3.6 Features for Sentiment Classification

3.6.1 Dictionaries Construction

Before the features selection step, two dictionaries are constructed: total-word-dictionary (*tw_total_dict*) and initial-polarity-dictionary (*pol_dict_ini*).

First, total-word-dictionary records all the words appeared in the tweet-dataset, with their occurrence frequencies. Here, in order to discard noisy words, restriction of frequency no less than 3 is applied. The size of total-word-dictionary became 58,625 entries.

Then, for the later sentiment analysis, we prepared initial-polarity-dictionary by combining existing popular authoritative polarity dictionaries on the Internet shown in Table 2. The size of initial-polarity-dictionary became 125,277 entries.

Table 2: The structure of the initial polarity dictionary

Label in the dictionary	Source
Positive	Positive Score > 0.75, or Positive Score – Negative Score > 0.5 (SentiWordNet[6]), Strong Positive (MPQA[7]), Positive category (the General Inquirer[8])
Negative	Negative Score > 0.75, or Negative Score – Positive Score > 0.5 (SentiWordNet), Strong Negative (MPQA), Negative category (the General Inquirer)
Neutral	Positive Score = 0 and Negative Score = 0 (SentiWordNet)

3.6.2 Syntax Features

A big difference between tweets and common texts lies in the syntax characteristics of them. Tweets own many unique syntax characteristics that other sentences do not have, including '@', the retweet mark, the URL link, and the hashtag. These characteristics bring some inconvenience while preprocessing the tweet texts; however, they are quite informative in the task of sentiment analysis.

In this research, totally 10 syntax characteristics are taken into consideration. They are exclamation marks (!), question marks (?), upper-case words, capitalized words, hashtags (#), at marks (@), retweet marks (RT), URL links, emoticons, and slang words. All these characteristics are counted by their occurrences in one tweet, and this 10-dimension vector is regarded as '*syn*' feature. Here, a manually built emoticon dictionary with 300 entries and slang dictionary with 200 entries are referred to during the counting process.

3.6.3 Enhanced Polarity Dictionary

Since *pol_dict_ini* cannot cover all the words in *tw_total_dict*, *pol_dict_ini* is enhanced by the following method.

First, for each word in *pol_dict_ini*, we set the polarity score as 2, -2, and 0 if it is labeled as Positive, Negative, and Neutral in *pol_dict_ini* respectively. Second, for each word NOT appears in *pol_dict_ini*, but appears in *tw_total_dict*, we put the polarity score as 1 for 'positive inclined' word, as -1 for 'negative inclined' word, and as 0 for all other words.

During the second step, we adopt PMI (Pointwise Mutual Information) calculation for each word w_1 which does not appears in *pol_dict_ini*, but appears in *tw_total_dict*. Here, w_2 is the word which appears in *tw_total_dict*.

$$PMI(w_1, w_2) = \log \frac{p(w_1, w_2)}{p(w_1) \cdot p(w_2)}$$

where, $p(w_1, w_2)$ is the co-occurrence probability of both word w_1 and w_2 in one tweet, and $p(w_1)$ and $p(w_2)$ are the occurrence probabilities of word w_1 and w_2 in one tweet respectively.

Then, for each word w_1, a set of PMI values is calculated followed by sorting them in descending order. After that, majority voting among a set of w_2 polarities whose $PMI(w_1, w_2)$ have the top-10 ranked probability. Based on the majority voting, we put polarity score to each word w_1. For example, the result of majority voting shows 'positive' wins, we put 'positive inclined' to the word.

As a result, we have a new polarity dictionary (*pol_dict*) with the same vocabulary of *tw_total_dict* as shown in Table 3, where each word is mapped to a score of 5 scales (i.e. 2, 1, 0, -1, and 2). By using *pol_dict*, each tweet can be projected to a 5-dimension vector, and each dimension records the count of the unigram words in this category. This vector is named as '*5s*' feature.

Table 3: The comparison of polarity word counts

Polarity score	*pol_dict_ini* ∩ *tw_total_dict*	*pol_dict* ∩ *tw_total_dict*
2	17,494	17,494
1	0	7,108
0	19,581	24,332
-1	0	4,957
-2	4,734	4,734
Total	41,809	58,625

3.6.4 Polarity Score based on Review-Dataset

We also prepare another polarity score to enhance the polarity analysis of tweets. Here, we use social review data, that usually consist of review texts and 5 scale score, on popular review websites[1,2,3].

Each entry of review-dataset as mentioned in 3.2 has a tuple structure such as $(text, score)$, where text is review text and score is its polarity score. The total vocabulary of the review-dataset is shown as W_{rv}. For each word w_i in W_{rv}, the polarity score for each word based on review-dataset is calculated by

[6] http://sentiwordnet.isti.cnr.it/
[7] http://mpqa.cs.pitt.edu/
[8] http://www.wjh.harvard.edu/~inquirer/

$$pol_{w_i} = \frac{\sum_{text_j \in TX_{w_i}} score_j}{|TX_{w_i}|}$$

where, TX_{w_i} is the set of review texts, in which the word w_i is included, $text_j$ is a review text, and $score_j$ is the corresponding polarity score of $text_j$.

Then for each tweet tw_i in tweet-dataset, the polarity score based on review-dataset is given by

$$pol_{tw_i} = \frac{\sum_{w_j \in W_{tw_i}} pol_{w_j}}{|W_{tw_i}|}$$

where, W_{tw_i} is the word set of tw_i. This float polarity score for each tweet is named as '*rv*' feature.

3.6.5 Polarity Score based on CCA with Review-Dataset

This feature is constructed by adopting canonical correlation analysis (CCA), which is a classical statistics method to figure out the latent relations among multiple variables, to review-dataset.

Each entry of review-dataset has a tuple structure such as $(text, score)$ as described in 3.6.4. The review-dataset is taken as the condition set, and the first correlated variable parameters are decided by the CCA process. Then, for each tweet in tweet-dataset, the first correlated variable is applied to calculate its polarity score. Finally, a float number is given to each tweet as '*cca*' feature.

3.6.6 Polarity Score based on Co-occurrence in 3 Word Window

This feature is constructed based on the assumption that the relationship between two words is deep if they appear closely. Thus, we restrict the appearance of two words within adjacent three words followed by adopting a propagation algorithm.

Inspired by the research of Brody and Elhadad [10], in which a propagation algorithm is applied to analyze the sentiment of online reviews, a modified graph-based propagation algorithm is adopted here to obtain the polarity score of each word in *tw_total_dict* based on the three-word-window neighboring relationship.

First, a co-occurrence dictionary is constructed by parsing all the tweets in the tweet-dataset. It consists of both word pair $w_i\_w_j$, and the frequency of appearance, noted as $freq(w_i, w_j)$, in the three-word-window. Then, as an initial propagation graph, all the words in *tw_total_dict* are taken as the nodes of the graph. The value of each node is initialized as 1, and -1 for the words in the Positive category and Negative category of *pol_dict_ini* respectively. For other words, the initial node value is set as 0. Then, for each iteration, the polarity of each node is updated by

$$pol'_{n_i} = (1-\alpha) \cdot \frac{\sum_{n_j \in NEI_{n_i}} pol_{n_j} \cdot \left(1 + \log\left(freq(n_i, n_j)\right)\right)}{\sum_{n_j \in NEI_{n_i}} \left(1 + \log\left(freq(n_i, n_j)\right)\right)} + \alpha \cdot pol_{n_i}$$

where, NEI_{n_i} is the set of the nodes neighbored with node n_i, and α is a tuning parameter, which is set as 0.6. After running the iterations to convergence, each node has a float value indicating its polarity, i.e., polarity of the word, followed by obtaining a polarity dictionary. Then, the same formula to calculate the polarity score of each tweet in 3.6.4 is applied. This float score for each tweet is named as '*co3*' feature.

3.6.7 POS-relationship Feature

POS (part-of-speech) information is also usually used in the NLP analysis. It has been reported that part-of-speech pairs are especially sentiment expressive in previous researches. Here, all the tweets are first processed by the Stanford Parser[9] to extract the dependencies trees. Then the typical relationships between POSs, shown below, are extracted. In this research, 10 most common and sentiment expressive POS relationships are chosen manually followed by giving a polarity label to each relationship manually; such that '*negation*' relationship expresses the opposite meaning of the latter POS. The POS relationships include '*adjectival complement*', '*adverbial modifier*', '*adjectival modifier*', '*conjunct*', '*direct object*', '*negation*', '*nominal subject*', '*purpose clause*', '*relative clause modifier*', and '*open clausal complement*'.

Then, to decide the polarity of a tweet, a simple majority voting technique is applied so that the larger number of polarity labels in the tweet wins. This feature is called '*pos*' in the later analysis steps.

4. Experiment and Results

4.1 Preprocessing

Before analyzing tweet-dataset, the preprocesses are adopted: 1) 'RT' and URL deletion; 2) Emoticons conversion; 3) Lower-casing; 4) HTML transcoding; 5) Hashtags conversion; 6) Punctuation deletion; 7) Word segmentation; 8) Non-alphabet words and single alphabet words deletion; 9) Stop words deletion; 10) Repeated alphabets reduction; 11) Chat words conversion; 12) Lemmatization. Here, when Stanford Parser is used, only 1) to 5) are carried out.

As for the review-dataset, the following preprocesses are adopted: 1) Lower-casing; 2) Word segmentation; 3) Non-alphabet words and single alphabet words deletion; 4) Stop words deletion; 5) Chat words conversion; 6) Lemmatization.

4.2 Sentiment Classification

Sentiment classification consists of two steps. The first step, subjectivity classification, is to classify the spam filtered dataset into the subjective dataset and the objective dataset. The second step, polarity classification, is to classify the subjective dataset into the positive dataset and the negative dataset. In each step, a pre-trained classifier is applied to carry

[9] http://nlp.stanford.edu/software/lex-parser.shtml

out the classification as shown below.

Features selection: In 3.6, 6 features are introduced. They are *'syn'*, *'5s'*, *'rv'*, *'cca'*, *'co3'*, and *'pos'* features. All the combinations of them are implemented.

Training method: The SVM (Linear, RBF, and Polynomial) and the Naïve Bayes (Gaussian, Multinomial, and Bernoulli) are used in this experiment.

Training implementation: The total number of implementation variations turns out to be:

$$(2^6 - 1) \cdot 6 = 378$$

Validation method: The standard 10-fold cross-validation is applied here.

Training dataset: For the subjectivity classifier, 1000 manually labeled tweets (500 subjective, 500 objective) are prepared. For the polarity classifier, 1000 manually labeled subjective tweets (500 positive, 500 negative) are prepared.

Part of the results of the subjectivity classifiers and the polarity classifiers are shown in Table 4 and Table 5.

Table 4: Subjectivity classifiers performance

syn	*5s*	*rv*	*co3*	*cca*	*pos*	accuracy
•		•		•		74.7%
	•		•	•		74.9%
	•	•		•	•	75.8%
	•	•	•			76.4%
•	•	•	•		•	76.5%
•		•	•	•	•	77.5%
•		•	•		•	**78.4%**

Table 5: Polarity classifiers performance

syn	*5s*	*rv*	*co3*	*cca*	*pos*	accuracy
•	•					82.2%
			•	•	•	85.3%
•			•	•	•	87.2%
•	•	•		•	•	89.6%
		•	•		•	89.9%
		•	•	•		90.6%
		•	•	•	•	**91.1%**

The black bot indicates that the feature is applied. According to these tables, the best-performed subjectivity classifier is obtained by the features combination of *'syn'*, *'rv'*, *'co3'*, and *'pos'*, with SVM polynomial training method, while the best-performed polarity classifier is obtained by the features combination of *'rv'*, *'co3'*, *'cca'*, and *'pos'*, with the SVM linear training method. These two classifiers are used in the classification step for the whole spam filtered tweet-dataset.

4.3 Tweet Dataset

Based on the 'list of restaurant chains' on Wikipedia, 6 restaurants, i.e., *Burger King, Mcdonald's, KFC, Pizza Hut, Subway,* and *Starbucks,* that have branches in the worldwide are chosen as the research subject. Also, 33 countries are selected: United states (US), United Kingdom (GB), Australia (AU), Indonesia (ID), Malaysia (MY), Canada (CA), Philippines (PH), Singapore (SG), Brazil (BR), India (IN), South Africa (ZA), Japan (JP), Mexico (MX), France (FR), and Netherlands (NL), Greece (GR), Thailand (TH), China (CN), Russia (RU), Spain (ES), Argentina (AR), Chile (CL), South Korea (KR), Germany (DE), Italy (IT), Ireland (IE), Venezuela (VE), Colombia (CO), Poland (PL), Egypt (EG), Ukraine (UA), New Zealand (NZ), and Viet Nam (VN).

Only tweets from these 33 countries are selected from whole tweet-dataset which consists of 10 million tweets. After applying pre-filtering and spam-filtering, the number of tweets became approximately 3 million. This dataset was used in the later analysis.

Figure 2 shows the distribution of tweets over the 6 restaurants in each country. As shown in Figure 2, the following conclusions can be obtained: 1) In different countries, the distribution of tweets over the 6 restaurants is quite different; 2) These distributions may give information for the popularities of each restaurant in each country.

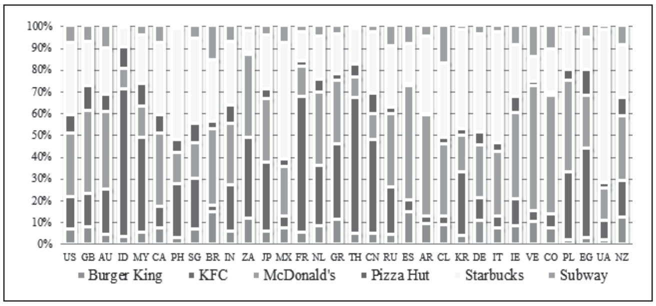

Figure 2: Distribution of tweets related to 6 restaurants

4.4 Sentiment Analysis

Here, we pick-up 15 countries' average sentiment scores for each restaurant and show the difference in Figure 3. For example, we can observe that McDonald's and Pizza Hut have positive sentiments among 15 countries'; however, other restaurants have variety of sentiments country by country.

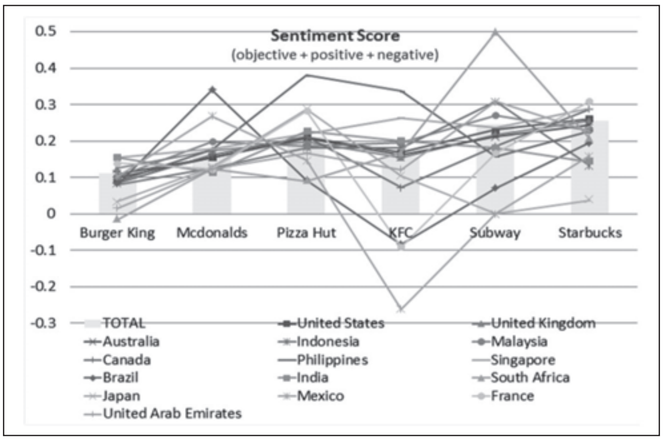

Figure 3: Sentiment difference among 15 countries

4.5 Location and Language based Analysis

We also investigated how 33 countries are similar each other by clustering the sentiment scores of 33 countries. K-means method was applied to cluster these countries into several groups. Here, k was empirically set as 10 after varying k from 2 to 10. The result is shown in Figure 4. The countries in same cluster are colored by the same color.

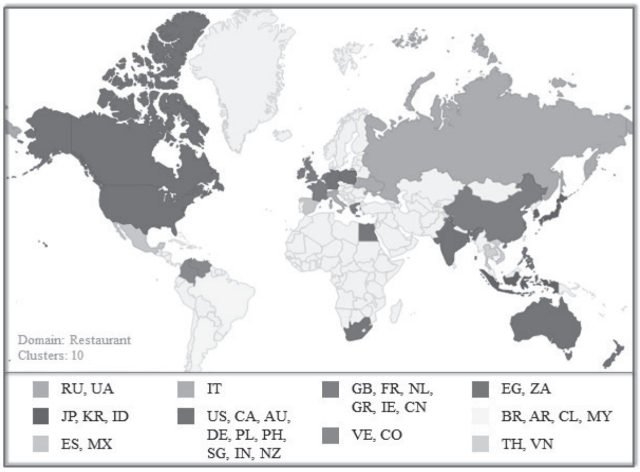

Figure 4: Clustering result of 33 countries

Upon observing Figure 4, some explanations can be led: 1) location-based effects are quite obvious for the cluster of RU and UA, the cluster of TH and VN, the cluster of EG and ZA, and the cluster of most of the Western European countries. We think they are clustered into the same cluster according to their location-based effects; 2) English-speaking countries such as US, CA, AU, SG, IN, PH, and NZ are clustered into the same group. Also Spanish-speaking countries such as ES and MX are clustered into the same group. It suggests that the language-based effects also exist; 3) Comparing to most of the European countries, some countries, such as ES and IT, seem to have quite different opinions for these restaurants, which may suggest that they have some special attitudes considering the food culture.

On the other hand, some confusing results still exist. For example, CN is clustered into the Western European cultural background group, and MY is clustered into the South American group. These confusing results may be led from other effective element except for location and language background, such as the eating patterns, the brand reputation, marketing strategies, and some locally specialized products and services.

5. Conclusions

In this research, the relationship between user evaluations for restaurants and other backgrounds, location and language, was investigated. The investigation is based on 33 countries around the world, and on tweets written in 34 languages. Three key classifiers, i.e., spam classifier, subjectivity classifier, and polarity classifier, were trained to obtain the accuracy of 97.8%, 78.4%, and 91.1% respectively. The results show that the location and language effects on user evaluations for restaurants actually exist.

Our future work includes 1) confirming the reliability of our analysis whether automatic translation has some wrong effects on it or not, and 2) expanding the analysis to other fields to confirm the potential of our method.

[References]

[1] Pang B. and Lee L., “Opinion mining and sentiment analysis,” *J. of Foundations and trends in information retrieval*, 2(1-2), pp.1-135, 2008.
[2] Liu B., “Sentiment analysis and subjectivity,” *Handbook of natural language processing*, 2nd Ed., 2010.
[3] Go A., Bhayani R., and Huang L., “Twitter sentiment classification using distant supervision,” *Project Report*, Stanford, 2009.
[4] Pak A. and Paroubek P., “Twitter as a Corpus for Sentiment Analysis and Opinion Mining,” *Proc. of LREC’10*, 2010.
[5] Barbosa L. and Feng J., “Robust sentiment detection on twitter from biased and noisy data,” *Proc. of COLING’10*, pp.36-44, 2010.
[6] Cesarano C., Picariello A., Recupero D. R., et al., “The OASYS 2.0 Opinion Analysis System,” *Proc. of ICWSM*, pp.313-314, 2007.
[7] Guo H., Zhu H., Guo Z, et al., “OpinionIt: a text mining system for cross-lingual opinion analysis,” *Proc. of CIKM2010*, pp.1199-1208, 2010.
[8] Cui A., Zhang M., Liu Y., et al., “Emotion tokens: Bridging the gap among multilingual twitter sentiment analysis,” *LNCS*, 7097, pp.238-249, 2011.
[9] Gao Q., Abel F., Houben G. J., et al., “A comparative study of users’ microblogging behavior on Sina Weibo and Twitter,” *LNCS*, 7379, pp.88-101, 2012.
[10] Brody S., Elhadad N., “An unsupervised aspect-sentiment model for online reviews,” *Proc. of HLT’10*, pp.804-812, 2010.

Jiawen LE
She received her Master of Engineering degree at Waseda University in Sep. 2014. Her research interests include analysis of cultural difference among countries by using social media data.

Hayato YAMANA
He received his Doctor of Engineering degree at Waseda University in 1993. He began his career at the Electrotechnical Laboratory (ETL) of the former Ministry of International Trade and Industry (MITI). He was subsequently appointed Associate Professor of Computer Science at Waseda University in 2000, and has been a professor in that department since 2005.