Emerald Insight

International Journal of Web Information Systems

Tip information from social media based on topic detection

Yuki Hattori Akiyo Nadamoto

Article information:

To cite this document:
Yuki Hattori Akiyo Nadamoto, (2013),"Tip information from social media based on topic detection", International Journal of Web Information Systems, Vol. 9 Iss 1 pp. 83 - 94
Permanent link to this document:
http://dx.doi.org/10.1108/17440081311316406

Downloaded on: 03 April 2015, At: 20:42 (PT)

References: this document contains references to 17 other documents.

To copy this document: permissions@emeraldinsight.com

The fulltext of this document has been downloaded 796 times since 2013*

Users who downloaded this article also downloaded:

Alton Y.K Chua, Snehasish Banerjee, (2013),"Customer knowledge management via social media: the case of Starbucks", Journal of Knowledge Management, Vol. 17 Iss 2 pp. 237-249 http://dx.doi.org/10.1108/13673271311315196

Jaroslav Pokorny, (2013),"NoSQL databases: a step to database scalability in web environment", International Journal of Web Information Systems, Vol. 9 Iss 1 pp. 69-82 http://dx.doi.org/10.1108/17440081311316398

Hao Han, Hidekazu Nakawatase, Keizo Oyama, (2014),"Evaluating credibility of interest reflection on Twitter", International Journal of Web Information Systems, Vol. 10 Iss 4 pp. 343-362 http://dx.doi.org/10.1108/IJWIS-04-2014-0019

Access to this document was granted through an Emerald subscription provided by All users group

For Authors

If you would like to write for this, or any other Emerald publication, then please use our Emerald for Authors service information about how to choose which publication to write for and submission guidelines are available for all. Please visit www.emeraldinsight.com/authors for more information.

About Emerald www.emeraldinsight.com

Emerald is a global publisher linking research and practice to the benefit of society. The company manages a portfolio of more than 290 journals and over 2,350 books and book series volumes, as well as providing an extensive range of online products and additional customer resources and services.

Emerald is both COUNTER 4 and TRANSFER compliant. The organization is a partner of the Committee on Publication Ethics (COPE) and also works with Portico and the LOCKSS initiative for digital archive preservation.

*Related content and download information correct at time of download.

Downloaded by 67.78.117.81 At 20:42 03 April 2015 (PT)

The current issue and full text archive of this journal is available at
www.emeraldinsight.com/1744-0084.htm

Tip information from social media based on topic detection

Yuki Hattori
Graduate School of Natural Science, Konan University, Hyogo, Japan, and
Akiyo Nadamoto
Department of Intelligence and Informatics, Konan University, Hyogo, Japan

Abstract

Purpose – The information of social media is not often written in ordinary web pages. Nevertheless, it is difficult to extract such information from social media because such services include so much information. Furthermore, various topics are mixed in social media communities. The authors designate such important and unique information related to social media as tip information. In this paper, they aim to propose a method to extract tip information that has been classified by topic from social networking services as a first step in extracting tip information from social media.

Design/methodology/approach – Themes of many kinds exist in a social media community because users write contents freely. Then the authors first detect the topics from the community and cluster the comment based on the topics. Subsequently, they extract tip information from each cluster. In this time, the tip information is include a user's experience and it has common important words.

Findings – The authors used an experiment to confirm that their proposed method can extract appropriate tip information from a community that a user specifies. The average precision is 69 per cent. A comparison of the authors' proposed method and baseline which is without detection of topic and clustering, the average precision obtained using the authors' proposed method is 18 per cent greater than the baseline.

Originality/value – The authors have three points to extract tip information from social media: topic detection and clustering from the social media using LDA method; extracting an author's actual experiences; and creation of a tip keyword dictionary from user experiments.

Keywords Social media, Extracting information, Experience mining, Topic detection, Information

Paper type Research paper

Downloaded by 67.78.117.81 At 20:42 03 April 2015 (PT)

1. Introduction

Recently, social media of many kinds that exist on the internet have created numerous and diverse communities. Using social media, users who are members of a social media community post and exchange information that is related to personal behaviour, experimentation, and their own sentiments. Sometimes this information is not included in the contents of ordinary web pages. For example, some festival communities in an SNS present experimental information that is not described on official web pages, such as "When you go to the festival by car, you should exit the expressway one exit early. If you exit at the nearest exit, you will hit a terrible traffic jam". That information is important for users who are not only community members, but also for people who are outside the community. Nevertheless, it is difficult to extract important information from social media because so much information exists and mixed multiple topics exist simultaneously in the communities. We designate such important and unique information related to social media as "tip information".

International Journal of Web Information Systems
Vol. 9 No. 1, 2013
pp. 83-94
© Emerald Group Publishing Limited
1744-0084
DOI 10.1108/17440081311316406

As described in this paper, we propose a means to extract such tip information from social media. The definition of our proposed tip information is:

(1) the information is credible and important; and
(2) a user does not know the information.

In this paper, we target definition (1) as a first step of the research. We consider that information based on a user's experience is more credible than information that is without a user's experience. Furthermore, many common important words exist in the information that a user thinks is important. We regard the tip information as including a user's experience and it has common important words.

We designate the common important word as "a tip keyword".

Themes of many kinds exist in a social media community because users write contents freely. Such topics might be a traffic topic, restaurant topic, best seat topic, or a children topic, each of which might include the fireworks festival community information described above.

Then we detect the topics from the community and cluster the comment based on the topics. Subsequently, we extract tip information from each cluster.

Many methods can be used to detect the topic. Tsutsumida (Tsutsumida *et al.*, 2012) proposed that, when detecting the topic from sparse data, the latent Dirichlet allocation (LDA) (Blei *et al.*, 2003; Griffiths and Steyvers, 2004) and graph-based methods are better than the other methods. As described in this paper, we use LDA methods to detect the topic from the social media content, which is sparse data.

The technical points of this paper are presented below:

(1) Means to detect topic and clustering from the social media using LDA method.
(2) Means to extract tip information:
   - extracting an author's actual experiences; and
   - creation of a tip keyword dictionary from user experiments.

As described herein, we target an SNS as social media. Our target users are not only community members but also outside people. The flow of extracting tip information progresses as follows and as shown in Figure 1:

- The user inputs a query along with tip information that the user wants to know.
- The system extracts communities from an SNS and browses the list of communities.
- The user selects a community from the list.
- The system extracts comments from the 20 threads of the community.
- It detects topics from the threads and clustering comments using LDA.
- It extracts actual experience sentences from each cluster.
- It extracts tip information from the actual experience sentences using a tip keyword dictionary.
- It browses the tip information based on each cluster.

This paper is organized as follows. Section 2 offers a discussion of related work. Section 3 explains tip information extraction. Section 4 presents a prototype system

Downloaded by 67.78.117.81 At 20:42 03 April 2015 (PT)

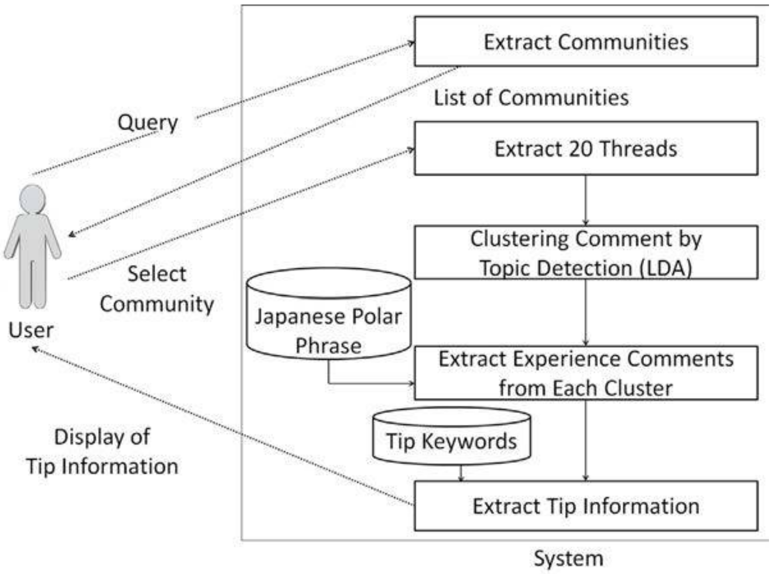


Figure 1.
Flow of extracting tip information

Downloaded by 67.78.117.81 At 20:42 03 April 2015 (PT)

and results of experiments conducted using our system. Section 5 presents the salient conclusions of our study.

2. Related work

Recently, social networking has become a popular application on the internet. Therefore, it is important to obtain information from social networking. Ting *et al.* (Ting, 2008; Ting *et al.*, 2009) introduce a method to collect and analyze multi-source social information and, in doing so, to extract social networks from the data.

In the present study, we propose a method of extracting credible and useful information from a social networking service. Inui *et al.* (2008) proposed experience mining, which is aimed at collecting instances of personal experiences automatically along with opinions from an extremely large number of user-generated contents such as weblog and forum posts and storing them in an experience database with semantically rich indices. As described herein, we define credible information as that which has been written based on actual experience. Therefore, we extract credible information based on experience mining.

Some studies have extracted useful information as comments written based on the evaluation. Dave *et al.* (2003) develop a method for distinguishing between positive and negative reviews automatically. Liu *et al.* (2005) propose a novel framework for analysis and comparison of consumer opinions of competing products. The system is such that, with a single glance of its visualization, the user can ascertain the strengths and weaknesses of each product in the minds of consumers in terms of various product features. Hu and Liu (2004) sought to mine and to summarize all customer reviews of a product. This summarization task differs from traditional text summarization. In this approach, one mines only the features of the product on which the customers have expressed their opinions and whether the opinions are positive or negative. Yu and Hatzivassiloglou (2003) present a model for classifying opinion sentences as positive or negative in terms of the main perspective being expressed in the opinion. Jansen *et al.* (2009) report research results investigating micro blogging as a form of electronic word-of-mouth for sharing consumer opinions related to brands. They investigated the overall structure of these micro blog posts, the expression types, and the movement in positive or negative sentiment. Morinaga *et al.* (2002) presents a new framework for mining product reputation on the internet. It collects people's opinions related to target products automatically from web pages. Then it uses text mining techniques to ascertain reputations of those products. For this study, we define useful information as comments that are written based on personal experience, this information has not been granted a reputation identifying it as either positive or negative.

We think that useful information contains a common keyword. Turney (2002) has been using a semantic phrase in the review. This approach presents a simple unsupervised learning algorithm for classifying reviews as recommended or not recommended. In this approach, the classification of a review is predicted by the average semantic orientation of the phrases in the review that contain adjectives or adverbs. The semantic orientation of a phrase is calculated as the mutual information between the given phrase and the word "excellent" minus the mutual information between the given phrase and the word "poor".

Some studies have been made identifying the topic using LDA.

Koike *et al.* (2012) identified topics using LDA for collection of news articles and blog articles of the same topic. They analyzed correlation of the topic between news and blogs. Furthermore, they analyzed the change of topic time series. The original LDA assigns a topic for the word. Therefore, a problem exists in that we cannot take account of relations between words. In contrast, Kitajima and Kobayashi (2011) proposed a method for extracting a latent topic that assigns a topic for an event instead of a word. For this study, we identify topics using LDA from comments of SNS. Moreover, we extract tip information for each topic.

3. Extracting tip information

3.1 Latent Dirichlet allocation

Various topics are mixed in the SNS community. Therefore, we classify comments on every topic using LDA for that information. Figure 2 shows a graphical model of LDA. The document set is represented by a collection of word w. K is the number of topics.

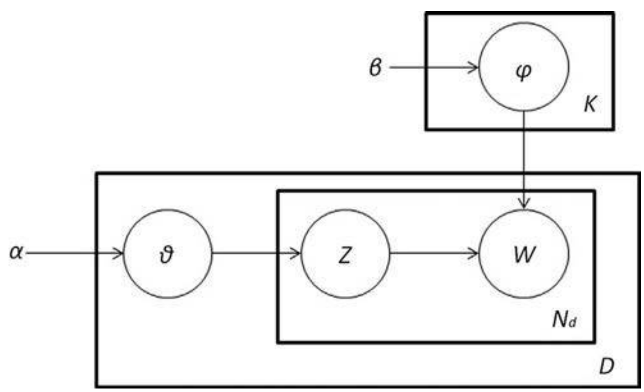


Figure 2.
LDA graphical model

Downloaded by 67.78.117.81 At 20:42 03 April 2015 (PT)

The probability distribution $P(w|zk)$ of word w for each topic zk ($k = 1, \ldots, K$) and the probability distribution $P(zk|d)$ ($d = 1, \ldots, D$) of the topic zk of the document d are generated by inputting the document set and K. Each document has a topic distribution θ, and each word in a document is chosen for topic z according to θ. According to the word distribution φ, topic z will be generated with word w. Also, D is the number of documents, and Nd represents the number of occurrences of words in document d. α and β are hyper-parameters.

3.2 Clustering comment by topic

As described in this paper, we regard a comment made on an SNS as a document. Parameter α is 0.5 and β is 0.1 in LDA. When we fix the number of topic K and the threshold of probability distribution $P(zk|d)(d = 1, \ldots, D)$ of the topic zk in document d, we produce several pre-experiments. In our pre-experiments, we first change the K as 3, 5, 7, 10, 15, 20. The best K is 10, then we fix the K as 10. Next, we change the threshold as 0.1, 0.2, 0.3, 0.4, and 0.5. Then we fix the threshold of P as 0.3.

3.3 Extracting experience sentences

After clustering the comments using LDA, we extract experience sentences from the comments. Our method is based on experience mining (Inui *et al.*, 2008). Experience mining is intended for the automatic collection of instances of personal experiences as well as opinions from social media. It consists of a topic object, experiencer, event expression, event type, factuality, and a source pointer. We assume that tip information is based on the author's event expression. We specifically examine only the event expression in experience mining. Their proposed event expression consists of a sentiment, a happening, and an action. They describe the sentiment as predicative expressions of an emotion or subjective evaluation, a happening as predicative expressions referring to a non-volitional event or state related to the use of a topic object and with sentiment orientation, and the action as predicative expressions referring to the experiencers' volitional actions related to the use of a topic object. Then they propose the *Japanese Polar Phrase Dictionary* (Kobayashi *et al.*, 2005; Higashiyama *et al.*, 2008). It consists of happening words, which consist of a verb, adjective, and noun, and their positive/negative value as a sentiment of the word. We use the *Japanese Polar Phrase Dictionary* to extract sentiment and happening sentences.

However, the action depends on the domain. For example, one action related to shopping is "buy". One action related to food is "eat". Then we create an action word dictionary of the domain dependence from each community of the SNSs. Table I presents an example of

Verb (in Japanese)	Noun (in Japanese)
Go (iku)	Use (riyou)
Lose way (mayou)	Move (idou)
Able to (dekiru)	Cheers (kanpai)
Dance (odoru)	Ensure (kakuho)
Buy (kau)	Participate (sanka)
View (miru)	Excitement (koufun)
Drink (nomu)	Activity (katudou)
Eat (taberu)	Discovery (hakken)

Table I.
Example of festival domain action words

Downloaded by 67.78.117.81 At 20:42 03 April 2015 (PT)

festival domain action words. We first gather 3,000 comments from festival communities. Then we extract the action words manually and count the term frequency (TF) of the nouns and verbs. We then infer the top 50 words of TF as festival domain action words.

After we create an action word dictionary, we extract experience sentences using a *Japanese Polar Phrase Dictionary* and an action word dictionary.

The sentences become tip information candidates.

3.4 Creating a tip keyword dictionary

We extract tip keywords using our experimental method. First, we gather tip information from comments of SNSs, and extract common keywords from it. Specifically in our experiment, we target festival communities. In our experiment, five participants read 2,000 comments and judge the comments as tip information or not. We regard information that is judged as tip information from four of five participants as tip information. Then we extract tip keywords from the sentences that were judged as tip information. Table II presents the quantities of comments that participants judged as tip information. Table III presents examples of tip keywords in festival communities. We regard that a comment is include one or more tip keywords as tip information.

After we create a tip keyword dictionary, we extract tip information from the tip information candidates extracted in Section 3.3.

4. Prototype system and experiments

4.1 Prototype system

We develop our prototype system. For it, we use PHP as the programming language, and MySQL as a DBMS. The target SNS is mixi which is the most popular SNS in Japan. We use JGibbLDA as an LDA tool.

In our prototype system, the first user inputs a query related to the theme of a community of which the user wants to know tip information using our

Downloaded by 67.78.117.81 At 20:42 03 April 2015 (PT)

Communities (number of comments)	A	B	C	D	E
PL Fireworks Festival (596)	114	155	321	138	98
Autumn Leaves in Kyoto (287)	27	102	137	69	49
a-nation (881)	16	82	338	200	28
Beach in Kansai area (99)	23	33	11	16	24
Kyoto Gion Festival (203)	42	33	19	24	40

Table II.
Number of tip information comments

Communities	Number of all comments	Number of comments that are correctly classified	Precision (%)
PL Fireworks show	569	459	81
Nabana no Sato	125	96	77
Beach in Kansai area	171	125	73
Autumn Leaves in Kyoto	162	124	77
Gathering of clams	227	175	77
Average of precision	–	–	77

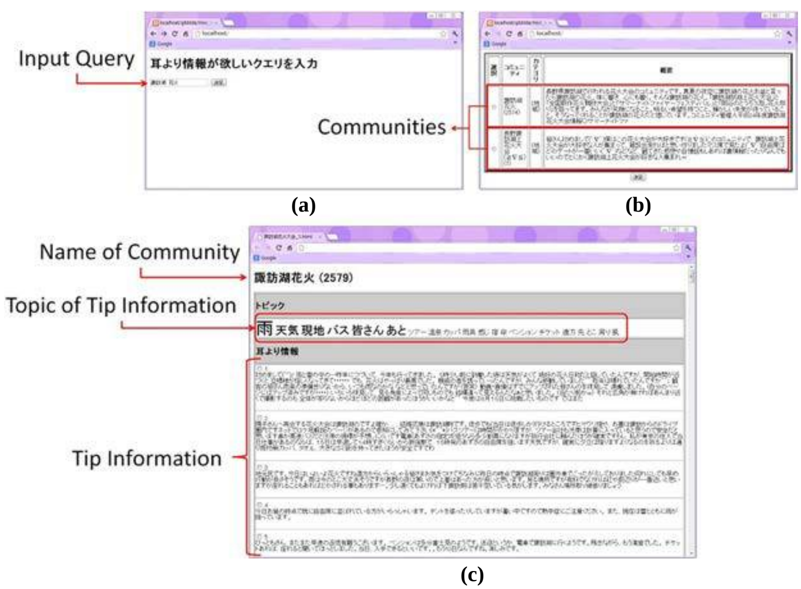
Table III.
Examples of tip keywords

system (Figure 3(a)). Then, the system outputs a list of communities (Figure 3(b)). Next, a user selects the community. Subsequently, the system detects a topic from the comment and clusters them. It extracts comments that include tip information from the cluster. The system displays topics of tip information which is extracted by LDA and also displays the comment which includes tip information in each cluster (Figure 3(c)). At this time, the system displays the top 20 of the probability distribution $P(w|zk)$. The larger values are displayed in larger characters. The character size is three steps. If the P is greater than 0.1, then the size is the largest one, and for P of less than 0.02, the size is the smallest one.

4.2 Experiments

We conducted experiments of two types to confirm the accuracy of our proposed method using our prototype system. In experiment 1, we evaluated the LDA method as good for detecting a topic from the comments. In experiment 2, we used an experiment to confirm that our proposed method can extract appropriate tip information from a community that a user specifies.

Experiment 1: availability of LDA method to detect a topic from comments of SNS. Comments of SNS are usually sparse data and the grammar of the sentence is not good. We first conduct an experiment of the LDA methods as good for such data. In our experiment, we use five communities, which are “PL Fireworks show”, “Nabana no sato”, “Beach in Kansai area”, “Autumn Leaves in Kyoto”, and “Gathering of clams”. In our experiment, we regard a sentence in each community is a document. We use LDA method to detect topics from comments and to calculate the accuracy. At this time, we use the parameter α as 0.5, β as 0.1, the topic number of K as 10, and the threshold of the probability distribution $P(zk|d)$ as 0.3.



Notes: (a) Input; (b) list of communities; (c) output

Figure 3.
Display of the result in our prototype system

Downloaded by 67.78.117.81 At 20:42 03 April 2015 (PT)

The average of the result is 77 per cent (Table IV). Then, LDA is good for comments of SNS, which can be regarded as sparse and dirty grammar data.

Experiment 2: availability of extracting tip information. We conducted an experiment of availability of tip information extraction. In this experiment, we compared our proposed method with the baseline which includes no detection of topics or clustering. Datasets are eight themes of communities (total 4,383 comments) that discuss festivals. The five participants read the comments of all eight themes and judged the comments as tip information or not.

The eight themes of communities were the following:

(1) *PL Fireworks show*. This is the largest fireworks festival in Western Japan. Every year, many people come to the festival, creating terrible traffic jams.

(2) *Nabana no sato*. This illuminated park is a famous tourist destination.

(3) *Beach in Kansai area*. This beach is in Western Japan.

(4) *Kyoto Gion Festival*. This is the oldest and most famous festival in Kyoto.

(5) *Autumn Leaves in Kyoto*. In autumn in Kyoto, many tree leaves turn red or yellow. Many people come to Kyoto to see them.

(6) *Kobe Luminarie*. This illumination for Christmas in Kobe is a famous event related to illumination displays. Many people come to Kobe to see it at Christmas-time.

(7) *Lake Suwa Fireworks show*. This is the largest fireworks festival in Eastern Japan. Many people come to Lake Suwa to see it.

(8) *Gathering of clams*. In Japan, many people come to the seaside to gather clams.

Results and discussion.

Accuracy of the proposed method. Experimental results are shown in Figure 4. The average precision is 69 per cent. The results of "PL Fireworks show", "Nabana no sato", "Autumn Leaves in Kyoto", and "Gathering of clams" are greater than 70 per cent. However, the results of "Kobe Luminarie" are less than 50 per cent. In some unfortunate cases, many comments discuss another theme in the communities. For example, few people discuss Kobe Luminarie itself, but they discuss another night view in Kobe because Kobe is famous for its beautiful night view. That theme differs from Kobe Luminarie itself, but the comments are included because they include many tip keywords such as "go", "come", and "met". We should consider that the discussion theme is appropriate for the community theme. Moreover, the results differ among participants. For example, for the Gathering of clams, the respective precision measures of participants A and B are greater than 90 per cent, but the precision of participant D is less than 50 per cent because the tip

Communities	Number of all comments	Number of comments that are correctly classified	Precision (%)
PL Fireworks show	569	459	81
Nabana no Sato	125	96	77
Beach in Kansai area	171	125	73
Autumn Leaves in Kyoto	162	124	77
Gathering of clams	227	175	77
Average of precision	–	–	77

Table IV.
Precision of clustering by LDA

Downloaded by 67.78.117.81 At 20:42 03 April 2015 (PT)

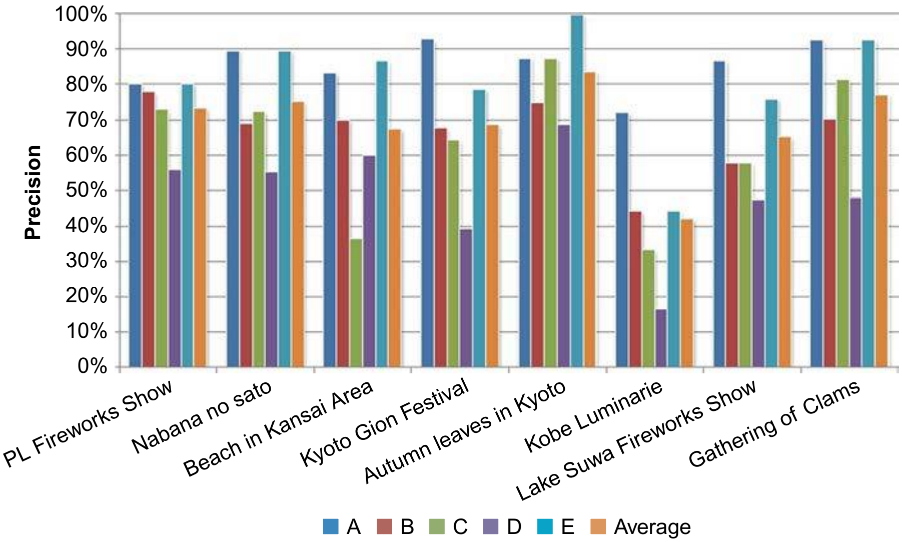


Figure 4.
Accuracy of proposed method

Downloaded by 67.78.117.81 At 20:42 03 April 2015 (PT)

information depends on the user's knowledge. A user who knows Gathering of clams well judged that the comment is not tip information. However, a user who does not know it well judged the comment as tip information. In this way, judgments related to tip information depend on the person making them, which is a point that must be considered in the near future. Table V presents an example of the results.

Comparison with the baseline. The results of comparing our proposed method with baseline which is without detection of topic and clustering in Figure 5. The average precision obtained using our proposed method is 18 per cent greater than the baseline. The result shows that extraction of tip information based on topic detection and clustering prove useful.

5. Conclusion

As described in this paper, we proposed a method for extracting tip information based on the author's experience. Important tip information uses some tip keywords from SNSs as a first step in extracting tip information from social media. First, we proposed detection of the topic and clustering comments using LDA method. Subsequently, we proposed a means for extracting tip information from SNS. We conducted an experiment demonstrating the accuracy of our proposed methods.

Future work will include the following tasks:

- For some comments of which the contents differ from a community's theme in extracted tip information, we should consider that the contents of comments fit the community theme.
- We do not consider the comment context, but we should.
- We should consider personalization because whether the information should be regarded as tip information is a subjective judgment.

Theme	Topic	Tip information
Autumn Leaves in Kyoto	Tofuku-ji temple Japanese garden Jissouin	It is getting difficult to find the good time to see autumn leaves. I personally recommend these spots: “Komyo-ji Temple” in the Rakusei area for looking up autumn leaves; “Jojakko-ji Temple” in the Arashiyama area for being surrounded by autumn leaves; “Tofuku-ji Temple” for looking down on autumn leaves: “Eikando” and “Enko-ji Temple “ for viewing carpets of autumn leaves; “Hosenin” in the Ohara area and “Genkoan” in the Rakuhoku area for looking at autumn leaves in a frame-like environment. To be honest, I do not want some of them to become very famous
Gathering of clams	Hamana Lake Seaside The spring-tide	I went to an area near the south side of Yu-utouoohashi over Hamana Lake at the time of aspring tide. The result was about a bucket and half of large clams. According to the locals, fisher-men sort clams around here and throw small ones away, so clams are increasing in number. The area seems to have been quite a well-kept secret place for the past several years, but where to park and how to get to the seaside are a bit of problem
PL Fireworks show	Tondabayashi Car Train	I definitely recommend taking a train. The train on the way home is super-crowded, so I advise you to get on at Tondabayashi-nishiguchi Stn. or Kawanishi Stn., even if it means walking one station further. Driving a car makes you miserable. My relatives living in Seika Town, Kyoto once came by car. They say that they left Tondabayashi at 11 p.m. and got home at 4 a.m.

Table V.
Example of the results of experiment

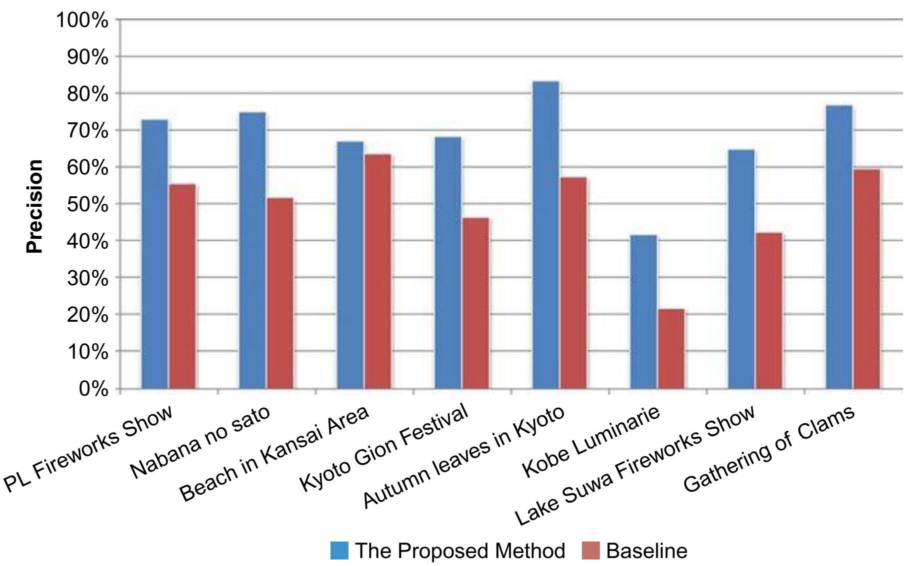


Figure 5.
Comparing with baseline

Downloaded by 67.78.117.81 At 20:42 03 April 2015 (PT)

References

Blei, D.M., Ng, A.Y. and Jordan, M.I. (2003), "Latent Dirichlet allocation", *Journal of Machine Learning Research*, Vol. 3, pp. 993-1022.

Dave, K., Lawrence, S. and Pennock, D.M. (2003), "Mining the peanut gallery: opinion extraction and semantic classification of product reviews", *Proceedings of the 12th International World Wide Web Conference*, pp. 519-28.

Griffiths, T. and Steyvers, M. (2004), "Finding scientific topics", *Proceedings of the National Academy of Sciences*, Vol. 101, pp. 5228-35.

Higashiyama, M., Inui, K. and Matsumoto, Y. (2008), "Acquiring noun polarity knowledge using selectional preferences (in Japanese)", *Proceedings of the 14th Annual Meeting of the Association for Natural Language Processing*, pp. 584-7.

Hu, M. and Liu, B. (2004), "Mining and summarizing customer reviews", *Proceedings of ACM-KDD*, pp. 168-77.

Inui, K., Abe, S., Morita, H., Eguchi, M., Sumida, A., Sao, C., Hara, K., Murakami, K. and Matsuyoshi, S. (2008), "Experience mining: building a large-scale database of personal experiences and opinions from web documents", *Proceedings of 49th 2008 IEEE/WIC/ACM International Conference on Web Intelligence*, pp. 314-21.

Jansen, B., Zhang, M., Sobel, K. and Chowdury, A. (2009), "Twitter power: tweets as electronic word of mouth", *Journal of the American Society for Information Science and Technology*, Vol. 60, pp. 2169-88.

Kitajima, R. and Kobayashi, I. (2011), "Summarization using latent Dirichlet allocation based on events in a document (in Japanese)", *Proceedings of the 3th Forum on Data Engineering and Information Management, Japan*.

Kobayashi, N., Inui, K., Matsumoto, Y., Tateishi, K. and Fukushima, T. (2005), "Collecting evaluative expressions for opinion extraction", *Proceedings of the 1st International Joint Conference on Natural Language Processing* (*IJCNLP-04*), pp. 596-605.

Koike, D., Yokomoto, D., Makita, K., Suzuki, H., Utsuro, T., Kawada, Y., Yoshioka, M., Kando, N., Fukuhara, T., Nakagawa, H., Kiyota, Y. and Seki, Y. (2012), "Analyzing correlation of topics in news and blogs and their changes: a case study of topics on earthquake disaster (in Japanese)", *Proceedings of the 4th Forum on Data Engineering and Information Management, Japan*.

Liu, B., Hu, M. and Cheng, J. (2005), "Opinion observer: analyzing and comparing opinions on the web", *Proceedings of the 14th International Conference on World Wide Web*, pp. 342-51.

Morinaga, S., Yamanishi, K., Tateishi, K. and Fukushima, T. (2002), "Mining product reputations on the web", *Proceedings of the 8th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, Canada*, pp. 341-9.

Ting, I.H. (2008), "Web mining techniques for on-line social networks analysis", *Proceedings of the 5th International Conference on Service Systems and Service Management, Melbourne, Australia*, pp. 696-700.

Ting, I.H., Wu, H.J. and Chang, P.S. (2009), "Analyzing multi-source social data for extracting and mining social networks", *Proceedings of International Conference on Computational Science and Engineering*, pp. 815-20.

Tsutsumida, K., Nakatsuji, M., Uchiyama, T., Toda, H. and Uchiyama, T. (2012), "Cross-domain recommendation using web access logs (in Japanese)", *Proceedings of the Information Fundamentals and Access Technologies, Japan*, Vol. 4, pp. 1-8.

Downloaded by 67.78.117.81 At 20:42 03 April 2015 (PT)

Turney, P.D. (2002), "Thumbs up or thumbs down? Semantic orientation applied to unsupervised classification of reviews", *Proceedings of the 40th Annual Meeting of the Association for Computational Linguistics*, pp. 417-24.

Yu, H. and Hatzivassiloglou, V. (2003), "Towards answering opinion questions: separating facts from opinions and identifying the polarity of opinion sentences", *Proceedings of the 2003 Conference on Empirical Methods in Natural Language Processing*, pp. 129-36.

Corresponding author
Akiyo Nadamoto can be contacted at: nadamoto@konan-u.ac.jp

To purchase reprints of this article please e-mail: **reprints@emeraldinsight.com**
Or visit our web site for further details: **www.emeraldinsight.com/reprints**

Downloaded by 67.78.117.81 At 20:42 03 April 2015 (PT)

This article has been cited by:

1. Bojan Božić, Werner Winiwarter. 2013. A showcase of semantic time series processing. *International Journal of Web Information Systems* **9**:2, 117-141. [Abstract] [Full Text] [PDF]

Downloaded by 67.78.117.81 At 20:42 03 April 2015 (PT)