Supporting Judgment of Fact Trustworthiness considering Temporal and Sentimental Aspects

Yusuke Yamamoto[1], Taro Tezuka[2], Adam Jatowt[1], and Katsumi Tanaka[1]

[1] Graduate School of Informatics
Kyoto University
Kyoto, Japan
{yamamoto, adam, tanaka}@dl.kuis.kyoto-u.ac.jp

[2] College of Information Science and Engineering
Institute of Science and Engineering
Ritsumeikan University
Kyoto, Japan
tezuka@media.ritsumei.ac.jp

Abstract. We have developed a system for helping users to determine the trustworthiness of uncertain facts based on sentiment and temporal viewpoints by aggregating information from the Web. Our goal is not to determine whether uncertain facts are true or false, but to provide users with additional data on which the trustworthiness of the information can be determined. The system shows with what sentiment and in what context facts are mentioned on the Web and displays any temporal change in the fact's popularity. Furthermore, the system extracts counter facts and analyzes them in the same way. The majority of 1000 users who evaluated our system found it to be a useful tool for helping to determine the trustworthiness of facts from various viewpoints.

1 Introduction

Nowadays, much information is available from various types of media, such as newspapers, magazines, and television. Furthermore, as the Web has grown and become very popular, people have become more easily and freely able to obtain information. However, along with this facilitated access to large amounts of data, users encounter more and more untrustworthy information. It is especially evident in uncontrolled environments like the Web, where information is often created by anonymous authors. Users often encounter uncertain statements on the Web, such as "soybeans are effective for weight loss." or "Pluto is a planet." Therefore, efficient methods for checking the trustworthiness of information on the Web are becoming necessary.

Perhaps the most common and easiest way to check the information trustworthiness is to use Web search engines such as Google [3]. By inputting questionable

[3] Google, http://www.google.com/

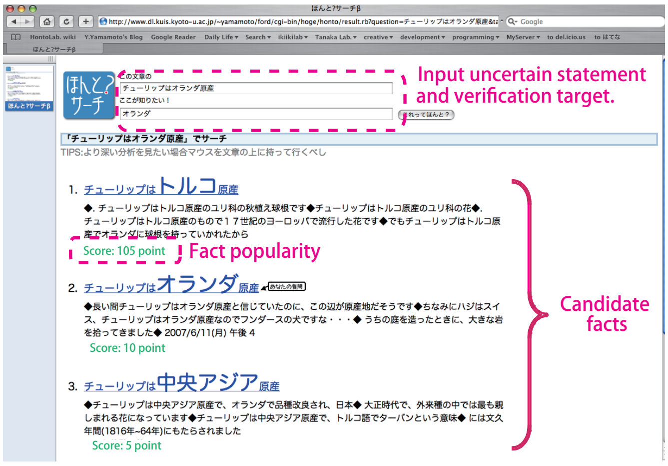

Fig. 1. System overview.

statements into search engines, users can examine their popularity on the Web and, at the same time, find out any contradictory information. However, this process is usually time consuming. To accommodate this need to check information, we have developed a system called "Honto? Search" ("Honto" means "is it true?" in Japanese) that helps users to judge the trustworthiness of statements whose reliability seems disputable [1]. When using our system, users can determine the popularity of questionable statements and their alternatives or counter examples on the Web (Fig.1).

However, in some cases a straightforward approach for estimating fact popularity produces incorrect results. This is because the context of information is not properly taken into account. For example, some pages may claim "soybeans are effective for weight loss; it is a great food!" while other pages may contain the following text "a TV station said that soybeans are effective for weight loss, however, this is a lie..." Obviously, these are counter opinions and should not be agglomerated in a straightforward way. Therefore, we extended our method in order to consider the context of encountered statements and, thus, to be able to more accurately evaluate the results. This extension involves sentiment analysis of information contained in Web pages in order to classify it as supporting or rejecting given questionable facts input by users. Except for sentiment-type context, there is also another important context of information that needs to be taken into account. While some facts are valid at any time point (i.e., "Albert Einstein was born in 1879"), many statements may be true only in certain time frames. For example, "the president of China is Hu Jintao" is a time-sensitive statement. In order to be able to precisely estimate the trustworthiness of such statements, one needs to consider temporal aspect of Web content.

In general, we believe that naive frequency-based approaches are insufficient and may judge correct facts extracted from the Web as wrong and wrong ones as correct. This drawback is especially problematic when contradicting facts have similar levels of popularity on the Web. In such a case, their sentiment and temporal analysis would provide more insight into the characteristics and trustworthiness of these facts. It would be possible not only to more accurately select correct answers but also to provide additional contextual information related to answers, sentiment to answers, and their dynamics in society.

We have thus developed an extended search system to help users more accurately determine the trustworthiness of facts on the Web. Honto? Search system has been expanded so that it considers information from sentimental and temporal viewpoints. The system has three key factors: counter examples extraction, sentiment distribution analysis, and popularity evolution analysis of facts. Honto? Search proposes counter examples to the input fact and provides a framework for their temporal and sentimental analysis. Sentiment analysis is carried out in order to categorize Web pages containing information about a doubtful fact as positive or negative about the fact and to present final sentiment distribution. This approach is augmented with the prior construction of a large-scale sentiment term dictionary from the Web. A temporal approach is also applied to analyze changes in popularity of facts over time and to visualize them to users.

The remainder of the paper is as follows. In the next section, we discuss related work. Section 3 provides the background description of Honto? Search system. Sections 4 and 5 discuss sentiment and temporal analysis of information on the Web, respectively. In the next section, we demonstrate the experimental results conducted using our system. The last section concludes the paper and outlines our future research directions.

2 Related Work

2.1 Trustworthiness of Web information

The problem that users have in the evaluation of Web resources from the viewpoint of trustworthiness has been recently studied by Nakamura et al. [2]. The authors have undertaken a large-scale analysis of user behaviors and expectations by conducting an online survey of Internet users. The results indicated that average Web users have difficulty in correctly estimating trust levels of encountered information. In addition, users often tend to trust the information without further analysis of its credibility. A study aiming at similar objective was made by Fogg et al. [3].

Many researchers have proposed effective ways of evaluating the trustworthiness of Web pages. PageRank is a well-known algorithm for estimating the trustworthiness of Web documents by considering their relative popularities [4]. Although the PageRank algorithm was quite effective in the past, its efficiency has been recently undermined due to the increase in link spam on the Web. Therefore, several approaches have been proposed to propagate trust among

Web pages with the purpose of combating Web spam [5]. In addition, with the proliferation of user-generated content, some researchers aimed at evaluating the trustworthiness of social content. For example, by focusing on editors' activities, trust and controversy levels of Wikipedia [4] articles were estimated [6]. In this research, we approach the problem of information trustworthiness in a more general way by agglomerating it from the Web and by considering its temporal and sentiment aspects.

2.2 Sentiment analysis

Sentiment analysis of documents is an attractive, yet, at the same time, quite challenging task [7, 8]. The potential benefits that would result from an effective opinion mining of large text collections like the Web cannot be underestimated. For example, companies would be very interested in the feedback from users of their products or services if it was automatically collected from large collections of blog data. In [8], the sentiment of content related to objects is determined using approaches borrowed from natural language processing. In [7], the authors calculate beforehand the probability that any term appears in a given sentiment class based on mining a large news corpus. By using the calculated information, sentiment levels of whole documents are then determined. Although our approach is similar to their method, it is however more general from a theoretical perspective and, also, more accurate. The target of [7, 8] is documents or objects, while our system analyzes sentiments about facts.

2.3 Term Dynamics

Our system uses temporal changes in the occurrence frequency of phrases to filter out obsolete information. Kleinberg's "word burst" is a well known method for examining changes in word frequencies over time [9]. It is a state-based approach that measures term dynamics characterized by transitions between two states: low- and high-frequency. That work, however, was intended to detect significant bursts of terms in text streams, whereas in our system, we compare differences between the frequencies of phrases and their duration over time on the Web. In the area of blog mining, Mei et al. [10] proposed using a Topic-Sentiment Mixture Model to model generation of Web users' positive and negative opinions on a certain subject over time.

3 Honto? Search

In this section, we describe WebQA system called Honto? Search [1] that we developed to help users judge the trustworthiness of given information. Users input a phrase representing a fact whose trustworthiness they doubt into the system and indicate the suspicious part of the phrase. The system then provides users with a popularity estimation for the fact and for alternative or counter

[4] http://www.wikipedia.org/

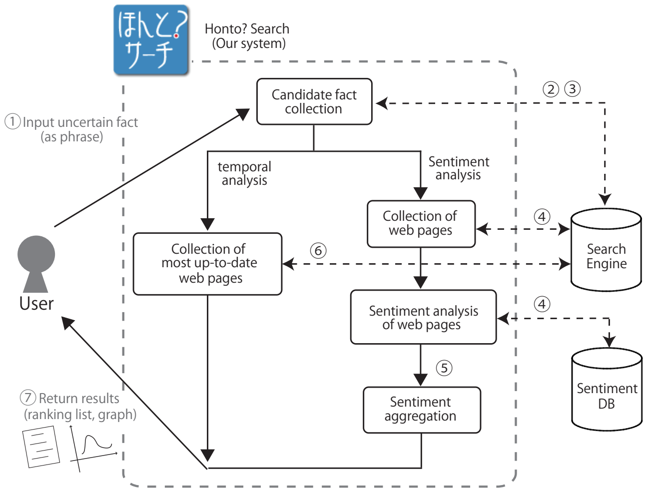

Fig. 2. System flow.

examples to it that occur on the Web. For example, if the user inputs "Tulips are native to the Netherlands" as an uncertain fact and "the Netherlands" as a verification target into Honto? Search, our system returns several facts. The basic idea is to extract the patterns which match "*tulips are native to *(wildcard)*" in snippets returned from a Web search engine. From the result ranking, the user can find that the most popular fact is "tulips are native to Turkey (this is the correct answer)" [5].

Fig.2 summarizes the approach in which extended Honto? Search estimates the trustworthiness of facts. The system works as follows.

1. The user inputs a questionable fact (as a phrase) and its doubtful part into the system.
2. The system divides the phrase given by the user into parts, and constructs a query that is then sent to a Web search engine.
3. The system extracts alternative or counter facts to the original phrase from the search results.
4. For sentiment analysis, the system inputs each fact into a Web search engine and collects search results.
5. The distribution of sentiment on each fact is estimated by aggregating the search results of each fact (Section 4).
6. The system sends the original fact and the counter facts to the Web search engine in order to collect time information (Section 5).
7. Using the time information, the system evaluates the original fact and candidate facts from a temporal viewpoint (Section 5).

[5] Some people have the wrong idea that tulips are native to the Netherlands according to the English Wikipedia article on tulips (http://en.wikipedia.org/wiki/Tulip)

"There are 15 countries in the European Union."	
counter terms	fact popularity
25	187
15	156
10	141

"The President of China is Hu Jintao."	
counter terms	fact popularity
Hu Jintao	589
Jiang Zemin	574

Table 1. Estimation of simple fact popularity.

If the user wants to know how popular a fact is, he or she can easily check the total number of Web pages that include it as a phrase by using a Web search engine. Our system does so automatically by sending all candidate facts as a phrase to a search engine and obtaining the total number of Web pages that include the fact. We call this number the *fact popularity*. Each fact popularity estimation is then presented to the user. By comparing the fact popularity of the original fact with the popularity of alternative facts, the user can get an idea of how strongly the fact is supported on the Web.

Table 1 shows the results of simple fact popularity analysis. In these examples, we input two facts, "there are 15 countries in the European Union" (Example 1) and "the President of China is Hu Jintao" (Example 2) to Honto? Search. Verification targets were "15" and "Hu Jintao", respectively.

Table 1 lists the frequencies of the original and counter facts in the Web search results. For example, for the fact "there are 15 countries in the European Union", we got two counter terms,"25" and "10". The most frequent one was "25", which is the correct answer. The counter term "15" also produced many results, since it was true until 2004. Additionally, the counter term "10", was also frequently reported on the Web, which must have come from expressions such as "10 new countries in the European Union". The user can judge that the original fact may not be trustworthy, since it is not the most frequent one. For the fact, "the President of China is Hu Jintao", we got an counter fact "the President of China is Jiang Zemin", which was actually true until 2003. From the table, the user can judge that the fact "the President of China is Hu Jintao" is reliable, since it is the most frequent one. These are simple estimations which do not consider the temporal aspect.

In addition to calculating simple fact popularity, Honto? Search estimates the sentiment behind the facts by analyzing context in the aggregated Web pages. The system shows the ratio of positive to negative sentiment to users. Fig.3 illustrates an example of sentiment aggregation. In this example, the fact "soybeans are effective for weight loss" was input into the system. The system extracted multiple candidate facts from the Web and performed a sentiment aggregation on them. Two interesting results are illustrated in the figure concerning two facts: "walking is effective for weight loss" and "soybeans are effective for weight loss". From the result, we can find that (1) the fact "soybeans are effective for weight loss" has more negative factors than the fact "walking is effective for

Fig. 3. Example of sentiment aggregation. Result no. 2 is "walking is effective for weight loss" and result no. 7 is "soybean is effective for weight loss". Green and orange colors in bar graphs denote positive and negative sentiments, respectively, while purple color means "unspecified".

weight loss", and (2) opinions on the fact about soybeans are not consistent and so we should not accept the fact.

As we have mentioned before, the extended system estimates also fact popularity from the temporal aspect. Our system describes the evolution of fact popularity on a graph like Fig.4. In this example, the user input "Chiba Lotte became Pacific League champions" as a doubtful fact and compared the evolution graph of the fact with that of the fact, "Nippon Ham became Pacific League champions" (Chiba Lotte and Nippon Ham are Japanese baseball teams). Both facts are true, but the result graph shows how the popularity of each team differs over time.

4 Sentiment Analysis

In this section, we describe the method to aggregate sentiment about a fact on the Web. To analyze sentiment about a fact, the system operates as follows:

1. The system inputs a fact as a phrase query to a Web search engine, such as Yahoo! [6] and gets top N search results.
2. It analyzes the sentiment contained in each result page and categorizes them as one of the following: "positive", "negative" or "unspecified".
3. It illustrates the ratio of three groups to users as a bar graph.

[6] http://www.yahoo.com/

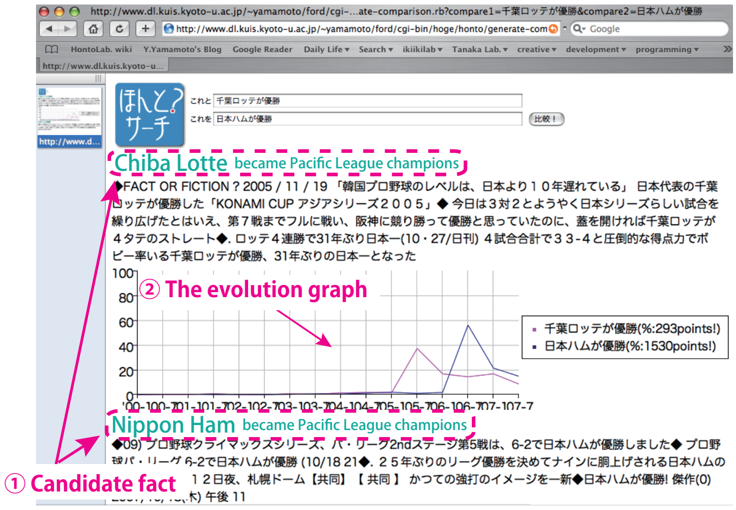

Fig. 4. Example of evolution graph. In the graph, blue line means the popularity change of “Chiba Lotte” and pink line indicates that of “Nippon Ham”.

4.1 Sentiment analysis of pages

Our method uses the Naive Bayes Classifier in order to categorize content as “positive” or “negative”.

Naive Bayes Classifier To statistically analyze the sentiment level of a text, we use a well-known statistical classifier, *naive bayes classifier*. This classifier is based on the Bayes theory and it can be speedily trained and applied. We use it on a multinomial model.

In order to use this classifier, the following assumption is made: all terms that appear in documents are independent of each other. The probability $Pr(d|c)$ that document d belongs to class c is formulated as follows:

$$
\begin{aligned}
Pr(c|d) &= \frac{Pr(c)Pr(d|c)}{\sum_{\gamma} Pr(\gamma)Pr(d|\gamma)} &(1)\\
&= \frac{Pr(c)Pr(L = l_d|c)Pr(d|l_{d,c})}{\sum_{\gamma} Pr(\gamma)Pr(d|\gamma)} &(2)\\
&= \frac{Pr(c)Pr(L = l_d|c)\binom{l_d}{n(d,t)}\prod_{t \in d} \theta_{c,t}^{n(d,t)}}{\sum_{\gamma} Pr(\gamma)Pr(d|\gamma)} &(3)
\end{aligned}
$$

Sentiment	Term seeds
Positive	fair, moderate, appreciate, comfortable
	reliable, happy, enjoyable, honest, useful
Negative	unfair, despite, useless, uncomfortable
	unhappy, trivial, dishonest, unreliable

Table 2. Parts of sentiment term seeds.

where $\begin{pmatrix} l_d \\ n(d,t) \end{pmatrix} = \frac{l_d!}{n(d,t_1!)n(d,t_2)!\ldots}$ is the multinomial coefficient. $\theta_{c,t}$ means the probability that term t appears in a document in class c. Let term t occur $n(d,t)$ times in document d, which is said to have *length* $l_d = \sum_t n(d,t)$. Using this formulation, we have to estimate $Pr(d|c)$ from training sets and prior probability. Here, class c is either "positive" or "negative". In order to estimate $Pr(d|c)$, we use the training data sets that are collected using the method which is described later. We then smooth the parameter of $\theta_{c,t}$ in order to estimate $Pr(d|c)$ as accurately as possible [11]. Finally, in order to classify a document into the two classes, we evaluate the logarithmic likelihood ratio LR. LR is defined as

$$LR = log\frac{Pr(pos|d)}{Pr(neg|d)} \tag{4}$$

$$= log\frac{Pr(pos)}{Pr(neg)} + \frac{\sum_{t \in d} n(d,t) log\theta_{pos,t}}{\sum_{t \in d} n(d,t) log\theta_{neg,t}} \tag{5}$$

If LR is much greater than 0, document d has positive sentiment, whereas if LR is below 0, document d has negative sentiment. When the value of LR is around 0, it means that the sentiment contained in document d cannot be specified. When our system classifies documents, the threshold $\alpha_{pos/neg}$ is used, and if $|LR| >= \alpha_{pos/neg}$, our system classifies documents into positive/negative sentiment.

Training data To make training data sets, we use the following approach. First we prepare positive term seeds and negative term seeds using 31 terms (see Table 2). These seeds are subjectively constructed. Next, we input the terms from each sentiment term seed as a query into a Web search engine and obtain search results. Then, the snippets collected with positive seeds are regarded as positive sentiment training sets and the snippets collected with negative seeds are regarded as negative sentiment training sets. This approach is based on the assumption that the prepared positive (or negative) seeds frequently appear in positive (or negative) documents. This method is a simple way to make training sets, and the Naive Bayes classifier, which is based on this approach, works well, as described in Section 4.1. Therefore, we use the approach above.

After collecting the training data sets, we count the terms' frequencies in positive and negative training data sets. In this step, only nouns, verbs, adjectives and adverbs are extracted. Finally, we calculate the probability $\theta_{pos,t}$ ($\theta_{neg,t}$) that each extracted term t appears as positive sentiment (or negative sentiment). Through the above procedure, we calculated $\theta_{pos,t}$ ($\theta_{neg,t}$) of 85,776 terms using 62,000 Web search results as training data sets.

5 Visualization of temporal evolution

Since a simple fact popularity does not take temporal factors into account, it is often not appropriate to use it for trustworthiness judgement. For example, consider the fact, "the Chinese president is Mr. Jiang Zemin". This fact was correct only until 2003. This example shows that the trustworthiness of a fact is strongly dependent on time. Our system helps users to temporally judge the trustworthiness of a fact by portraying its temporal evolution. Here, we define the popularity of fact A, FP_A, at time period t (FP is fact popularity).

Fact popularity of fact A during time period t

$FP_A(t)$: the number of Web pages that refer to fact A and were last modified during time period t

Using FP, we can estimate which fact out of several alternative facts is the most reliable for a certain period of time. That is, if we want to estimate which fact is more reliable, fact A or fact B in time period t, we only have to compare $FP_A(t)$ with $FP_B(t)$. If $FP_A(t)$ is greater than $FP_B(t)$, we can estimate that fact A was truer in period t than fact B. We calculate popularity for the fact, which is the fact the user inputs into our system, and $FP(t)$ for alternative facts, which our system generates from the original fact. We then identify the fact for which $FP(t)$ has the greatest score to be the most reliable fact during period t.

To collect Web pages with timestamps, we use the Ask.jp[7] search engine. This search engine allows the use of a temporally structured query such as "keywords between: date1, date2" for which relevant Web pages modified between date1 and date2 are returned. The time periods for which the support of facts is to be estimated can have an arbitrary length.

Fig.5 shows that although FP for the fact "the President of China is Jiang Zemin" is higher than that of "the President of China is Hu Jintao" at the beginning, they reversed around 2003. In fact, Jiang Zemin was the President of China until March 2003 and was later substituted by Hu Jintao.

6 Evaluation

In this section, we present the results of our experiment. The experiment had two main parts. The first was to estimate the efficiency of sentiment analysis by

[7] http://ask.jp/

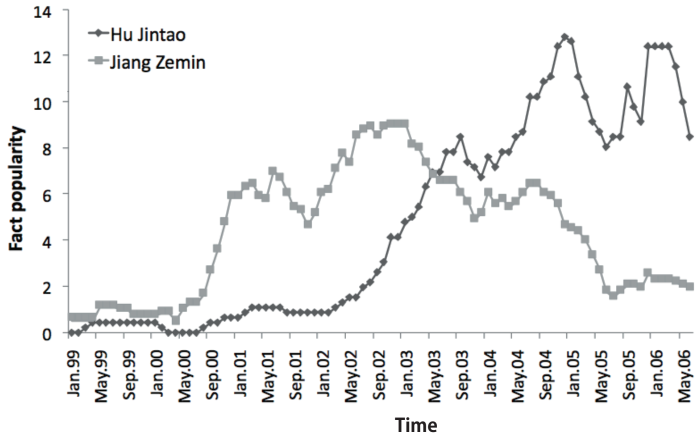

Fig. 5. Example of temporal evolution.

evaluating its precision and recall levels. The second part was a user test. The goal of the user test was to confirm whether or not our system is useful when users use it for checking the trustworthiness of a fact.

6.1 Precision and recall of sentiment analysis

Sentiment aggregation of Web pages strongly depends on the ability to estimate the sentiment contained in a text. In this section, we evaluated the capability to estimate sentiment contained in a text using our system by comparing our sentiment analysis method with the one described in [7], which uses a preconstructed sentiment dictionary. We made test sets from 196 randomly collected blogs. They were manually categorized into three groups: "positive (58)", "negative (59)", and "unspecified (79)" (the numbers in brackets are the numbers of blogs that were categorized into each sentiment group). We evaluated the efficiency of the two sentiment analysis methods using the test sets with different levels of threshold $\alpha_{pos/neg}$. We have found the optimum values of this threshold to be between 8 and 17. For using a Naive Bayes Classifier, we set the prior probability $P(pos)$ as 0.5. That is, we assumed that positive pages and negative pages were uniformly distributed on the Web. The result is given in Table 3. $Prec_{pos}(Prec_{neg})$ means the precision of positive (negative) sentiment decisions. $Recall_{pos}(Recall_{neg})$ means the recall of positive (negative) sentiment decisions. These values are calculated using all the test sets, thus, also the ones whose sentiment cannot be specified (unspecified sentiment category). In order to better gauge the performance of our method, we have also evaluated the efficiency by removing test sets whose sentiment could not be specified. $Prec_{pos'}$, $Prec_{neg'}$, $Recall_{pos'}$ and $Recall_{neg'}$ mean the precision and recall of positive (negative) sen-

$\alpha_{pos/neg}$	$Prec_{pos}$	$Recall_{pos}$	$Prec_{neg}$	$Recall_{neg}$	$Prec_{pos'}$	$Recall_{pos'}$	$Prec_{neg'}$	$Recall_{neg'}$
8	0.464	0.224	0.455	0.612	1.000	0.542	0.732	1.000
9	0.462	0.207	0.452	0.571	1.000	0.545	0.737	1.000
10	0.478	0.190	0.441	0.531	1.000	0.550	0.743	1.000
11	0.524	0.190	0.456	0.531	1.000	0.579	0.765	1.000
12	0.556	0.172	0.434	0.469	1.000	0.556	0.742	1.000
13	0.625	0.172	0.449	0.449	1.000	0.556	0.733	1.000
14	0.615	0.138	0.455	0.408	1.000	0.533	0.741	1.000
15	0.615	0.138	0.462	0.367	1.000	0.571	0.750	1.000
16	0.667	0.138	0.457	0.327	1.000	0.571	0.727	1.000
17	0.727	0.138	0.438	0.286	1.000	0.571	0.700	1.000
Baseline	0.342	0.224	0.364	0.735	0.813	0.406	0.655	0.923

Table 3. Evaluation of sentiment decisions using our approach vs. the baseline method.

timents on the test datasets composed of only positive and negative categories (i.e. blogs whose sentiment is unspecified are not taken into account).

From Table 3 we can see that our method classifies sentiment contained in documents more accurately than the baseline method. In particular, if we do not take unspecified sentiment into account, the performance of the Naive Bayes classifier is very high (the precision is about 100%(positive)/74%(negative) and the recall is about 56%(positive)/100%(negative)). That is, there were few wrong decisions. From these results, we can notice that the problem of our system is to correctly classify ambiguous content. Therefore, we have to (1) make more exact and (2) larger test sets in order to evaluate the performance more precisely. Improving the method to select sentiment term seeds and the algorithm to estimate sentiment is a part of our future work.

6.2 User test

In this section we report on the results of the user test we did regarding the effectiveness of our system. The user test was online and it was done between 7th and 12th February 2008 by a group of 1000 Internet users in Japan. Subjects were divided into four categories depending on their age: 20-29, 30-39, 40-49 and 50-59 years old. Each category consisted of 250 respondents, where half were males and half females. We asked the users to judge the trustworthiness of facts using Honto? Search and to give their impressions about the system in a free form. The given uncertain fact was "soybean is effective for weight loss". We did not limit the time to make the judgement and asked the users to use the system until they could finally confirm the trustworthiness of the fact. The survey was done in the Japanese language (we show here translated results).

In summary, 44.1% of users had a good impression of Honto? Search, 31.1% had a bad impression, and 24.8% had a neutral impression. Although, more users evaluated our system as good than as bad, the number of users who evaluated our system as good is still not enough. Fig.6 summarizes the aspects that users

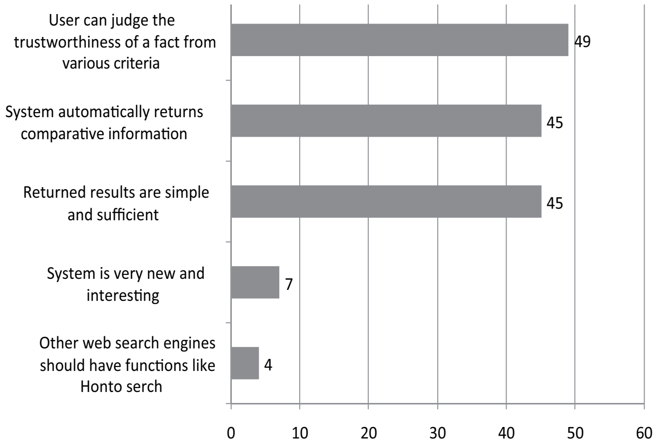

Fig. 6. Main reasons for judging our system as good.

evaluated as good and Fig.7 indicates the aspects that users evaluated as bad in our system. From Fig.6, we can see that some functions that we proposed were useful because users could judge the trustworthiness of facts quickly and concisely and make necessary comparisons. If users do the same thing using conventional Web search engines, it is time consuming. The most negative thing, on the other hand, was that our system works very slowly. It takes about 10 seconds to return results. Thus, we have to improve the response speed of our system. Focusing on the answers related to the trustworthiness, we can see that some users do not understand the results generated by the system and the mechanism in which trustworthiness of facts is estimated. It is, thus, important to provide users with the information that is easy to be understood. We plan to improve the interface to make it easier and more transparent to users.

7 Conclusion

In this paper, we have proposed a method to estimate the trustworthiness of a fact from sentimental and temporal viewpoints. We implemented a trustworthiness judgment support system, Honto? Search, based on the proposed method. Our approach uses (1) candidate fact extraction, which enables users to compare facts, (2) sentiment aggregation, which is the summarization of sentiment about a fact on the Web, and (3) visualization of temporal evolution about a fact, which enables users to know how a fact's popularity has changed over time. With Honto? Search users can compare their query with automatically collected similar facts in order to have more confidence in their judgment. Furthermore,

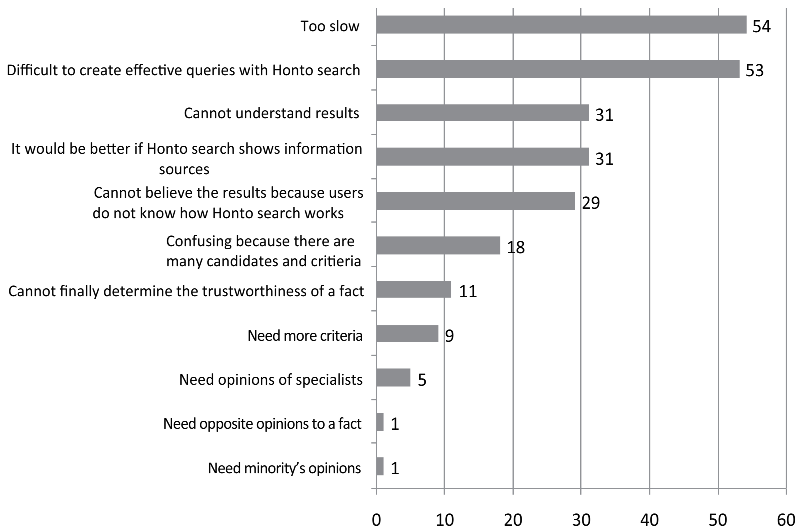

Fig. 7. Main reasons for judging our system as bad.

users can estimate sentimental and temporal characteristics of the fact, which cannot be obtained using conventional naive frequency-based approaches.

Nevertheless, Honto? Search system has some problems. In the process of extracting candidate facts, a rather naive natural language technique is employed; therefore, if a fact's meaning is the same as another fact's meaning and their spellings are different, the system returns different results. Moreover, the interface and the time required to generate the results are not yet satisfactory and therefore should be optimized. In the future, we plan to work on improving the above drawbacks and we would like also to estimate reliability of Web pages and Web sites based on a sentiment and temporal analysis.

Acknowledgement

This work was supported in part by a MEXT Global COE Program "Informatics Education and Research Center for Knowledge-Circulating Society" (Project Leader: Katsumi Tanaka), a MEXT Grant for "Development of Fundamental Software Technologies for Digital Archives", Software Technologies for Search and Integration across Heterogeneous-Media Archives (Project Leader: Katsumi Tanaka), a MEXT Grant-in-Aid for Scientific Research on Priority Areas "Cyber Infrastructure for the Information-explosion Era", Planning Research "Contents Fusion and Seamless Search for Information Explosion" (Project Leader: Katsumi Tanaka, A01-00-02, Grant#: 18049041), the National Institute of Information and Communications Technology, a MEXT Grant-in-Aid for Young

Scientists (B) "Trust Decision Assistance for Web Utilization based on Information Integration" (Project Leader: Taro Tezuka, Grant#: 18700086), and a MEXT Grant-in-Aid for Young Scientists(B) "Information Discovery Using Web Archives" (Project Leader: Adam Jatowt, Grant#: 18700111).

References

1. Yamamoto, Y., Tezuka, T., Jatowt, A., Tanaka, K.: Honto? Search: Estimating Trustworthiness of Web Information by Search Results Aggregation and Temporal Analysis. Proc. the 9th Asia-Pacific Web Conference (APWeb2007), LNCS **4505** (2007) 253–264
2. Nakamura, S., Konishi, S., Jatowt, A., Ohshima, H., Kondo, H., Tezuka, T., Oyama, S., Tanaka, K.: Trustworthiness Analysis of Web Search Results. Proc. ECDL2007, LNCS **4675** (2007) 38–49
3. Fogg, B., Swani, P., Treinen, M., Marshall, J., Laraki, O., Osipovich, A., Varma, C., Fang, N., Paul, J., Rangnekar, A., et al.: What makes Web sites credible?: a report on a large quantitative study. Proceedings of the SIGCHI conference on Human factors in computing systems (CHI2001) (2001) 61–68
4. Page, L., Brin, S., Motwani, R., Winograd, T.: The pagerank citation ranking: Bringing order to the web (1998)
5. Gyongyi, Z., Garcia-Molina, H., Pedersen, J.: Combating web spam with trustrank. Proceedings of the Thirtieth international conference on Very large data bases-Volume 30 (VLDB2004) (2004) 576–587
6. Vuong, B., Lim, E., Sun, A., Le, M., Lauw, H.: On ranking controversies in wikipedia: models and evaluation. Proceedings of the international conference on Web search and web data mining (WSDM2008) (2008) 171–182
7. Kumamoto, T., Nadamoto, A., Tanaka, K.: Design and Evaluation of News Reader "wEE" Expressing Emotional Aspects of News Articles. Trans. IEICE of Japan (2007) vol.J90–D, no.2, pp. 185–195
8. Yi, J., Niblack, W.: Sentiment Mining in WebFountain. Proceedings of the 21st International Conference on Data Engineering (ICDE2005) (2005)
9. Kleinberg, J.: Temporal Dynamics of On-Line Information Streams. Data Stream Management: Processing High-Speed Data Streams, Springer (2005)
10. Mei, Q., Ling, X., Wondra, M., Su, H., Zhai, C.: Topic sentiment mixture: modeling facets and opinions in weblogs. Proceedings of the 16th international conference on World Wide Web (WWW2007) (2007) 171–180
11. Chakrabarti, S.: Mining the Web: Discovering Knowledge from Hypertext Data. Morgan Kaufmann (2003)