

# Mediatory Summary Generation: Summary-Passage Extraction for Information Credibility on the Web <sup>\*</sup>

Koichi Kaneko, Hideyuki Shibuki, Masahiro Nakano, Rintaro Miyazaki, Madoka Ishioroshi, and Tatsunori Mori

Graduate School of Environment and Information Sciences, Yokohama National University  
79-7 Tokiwadai, Hodogaya-ku, Yokohama 240-8501, Japan  
{kaneko, shib, nakano, rintaro, ishioroshi, mori}@forest.eis.ynu.ac.jp

**Abstract.** In this paper, we discuss the summarization for supporting a user's judgment on the credibility of information on the Web. In general, if a statement contradicts another statement, the credibility of either of the statements decreases. However, these opposing statements may coexist under certain situations, and presenting such situations is helpful for a user's judgment. Therefore, we focus on the coexistence of these opposing statements, and attempt to develop a system to generate survey reports that contain *mediatory summaries*, which are defined as passages extracted from Web documents in order to present situations in which these opposing statements can coexist. We describe the outline of the summarization system and describe how to improve the TextRank algorithm from the viewpoint of passage extraction for the system. From experimental results, we confirmed that the methods based on the improved TextRank algorithm can extract significant passages, which are actually considered as significant by human assessors, with higher precision than baseline methods.

**Keywords:** Mediatory Summary, Information Credibility, Passage Extraction, Text Summarization

## 1 Introduction

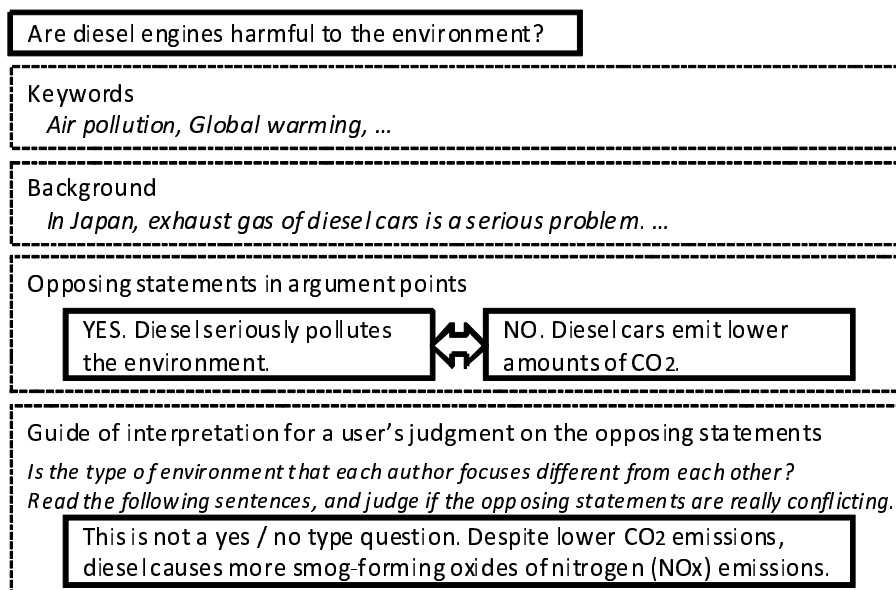
Many pages on the Web contain incorrect or unverifiable information. Therefore, there is a growing demand for technologies that enable us to obtain reliable information. However, it would be almost impossible to automatically judge the accuracy of information presented on the Web. In this case, the second-best approach is to develop a supporting method that judges the credibility of information on the Web.

At present, when we would like to judge the credibility of information on the Web, we often read some relevant Web documents retrieved via Web search engines. However, if the content described in some of the documents conflicts with the content in some other documents, Web search engines do not give users any suggestions on how to interpret the conflict. Furthermore, the amount of retrieved documents is too large to read, and the documents may not be ranked in the order of the credibility of information. Therefore, information retrieval is not sufficient to support a user's judgment on the credibility of information, and therefore, additional techniques are required.

Such techniques for supporting a user's judgment on the credibility of information are classified into three types. The first type involves the extraction of significant descriptions such as statements that may be relevant to a user and the sender of the statements. In this paper, a *statement* is defined as text such as an opinion, evaluation, or objective fact. The term *sender* refers to a person or an

---

<sup>\*</sup> This research is partially supported by National Institute of Information and Communications Technology, Japan and a research project of Graduate School of Environment and Information Sciences, Yokohama National University.



**Figure 1:** Example of survey report

organization providing a description of certain information on the Web. The second type involves the analysis of semantic relations such as AGREEMENT, CONFLICT, or EVIDENCE between statements (Murakami *et al.*, 2009). The third type involves the summarization and presentation of the extracted and analyzed information to enable the users to easily judge the credibility of statements. In this study, we attempt to develop a system that automatically generates a survey report for verifying the credibility of information (Ando *et al.*, 2008).

The rest of this paper is organized as follows. In section 2, we discuss the different points of the summarization for information credibility and the summarization for general purpose, and describe a survey report for verifying the credibility of information. In section 3, we explain mediatory summarization that is one of the most significant concepts in our summarization. In section 4, we describe the outline of the summarization system for supporting the user's judgment on the credibility of information on the Web. In section 5, we conduct examinations on elementary techniques of our summarization, and discuss the results. In section 6, we provide our conclusion.

## 2 Summarization for Information Credibility

The summarization for the information credibility is a kind of informative summarization of multiple documents, while it is different from the summarization for general purpose as follows. Although a set of documents is given as the input for the summarization for general purpose, a statement such as "Are diesel engines harmful to the environment?" is given for the summarization for the information credibility. In other words, we assume that a user enters a statement, and he/she would like to verify its credibility, such as a query in information retrieval. Therefore, finding a set of documents relevant to the given statement is involved in the process of the summarization for the information credibility.

In general, there are roughly two types of techniques for automatic summarization: sentence extraction and sentence compression. However, the main technique in the summarization for the information credibility is passage extraction. In case a user judges the credibility of information, the sensitive expression of the document, which conveys nuances, is one of the indicators for the judgment that is as significant as semantic relations between sentences. Unfortunately, current techniques for automatic summarization cannot always and completely reflect such sensitive ex-

pression into a summary. Therefore, passage extraction is appropriate to preserve the expression of source documents as much as possible.

Although informing a user of contents described in source documents is the principal purpose of summarization for general purpose (Haghighi and Vanderwende, 2009), it is not sufficient for supporting a user's judgment on the credibility of information. For example, if there are two opposing statements such as "Diesel engines are harmful to the environment" and "Diesel engines are not harmful to the environment," simply showing these statements may not contribute to making a user's judgment on the credibility of the statements easy. In the first place, a user cannot judge which is the case: whether the statement is wrong or whether the statements can coexist under certain situations. Furthermore, while a user's judgment is supported by presenting the reasons of each statement in the former case, it is supported by presenting the situations of the coexistence in the latter case. In other words, in addition to informing a user of the contents described in source documents, presenting how to interpret the contents, which may be also included in a passage extracted from Web documents, is significant for the summarization for the information credibility. Further, a survey report for verifying the credibility of information should be generated as a structured document for a user to be able to easily distinguish between contents described in source documents and guidance passages that present a user with how to interpret the contents.

Figure 1 shows an example of a survey report for verifying the credibility of information. A survey report roughly consists of the following four parts. The first and second parts are a set of keywords relevant to a query statement and a background in a topic of the query statement, respectively. The third part is a list of argument points, especially ones in which opposing statements exist. The fourth part is a list of guides for a user's judgment on the argument points described in the third part, and each guide is generated in different manners according to the system's judgment on whether opposing statements in the argument point can coexist. The first three parts are for making a user understand what issues and opinions exist relevant to the query statement, while the fourth part is for providing information for a user to understand how to interpret the issues and opinions.

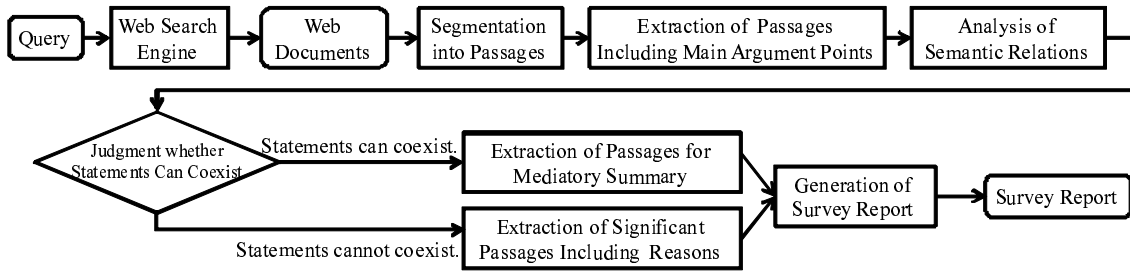
### 3 Mediatory Summarization

In general, if a statement is opposed to another statement, the credibility of either statement will decrease because we tend to think that either of these statements is wrong. However, some of the statements that appeared to contradict each other at first glance are not logically contradictory and can coexist under a certain situation.

For example, let us consider the following two pairs of statements:

- (a1) Diesel engines are harmful to the environment.
- (a2) Diesel engines are not harmful to the environment.
- (b1) Diesel engines are harmful to the environment because of more smog-forming oxides of nitrogen emissions.
- (b2) Diesel engines are not harmful to the environment because of lower carbon dioxide emissions.

The pair of statements (a1) and (a2) appear contradictory. On the other hand, the pair of statements (b1) and (b2) are not logically contradictory because the statements are described in different contexts, namely, air pollution and global warming. If all statements on the Web were described in detail such as the latter pair, a user's judgment on the credibility of the statements would be easier. Unfortunately, the statements on the Web are often abbreviated as the former pair, and such abbreviated statements may make a user to misunderstand these statements as contradictory. Therefore, it is helpful for a user to present the hidden contexts such as "more smog-forming



**Figure 2:** Outline of the system to generate survey reports

oxides of nitrogen emissions” and “lower carbon dioxide emissions” in the above example, and it may assist the user in judging if opposing statements in a conflict can coexist.

We define *pseudo-conflicts* as relations between statements that appear to be contradictory at first glance but can actually coexist, and define *true conflicts* as opposing relations other than pseudo-conflicts. We also define *mediatory summaries* as passages that represent hidden contexts of statements in pseudo-conflicts.

Mediatory summaries are classified into two subtypes: namely direct mediatory summaries and indirect mediatory summaries. A direct mediatory summary is a passage extracted from Web documents, which include both of hidden contexts of statements in a pseudo-conflict and situations. Although direct mediatory summaries make a user’s judgment on the coexistence very easy, passages suitable to direct mediatory summaries may not be always found. In such cases, the system will show indirect mediatory summaries, which are combinations of passages that represent the hidden context of each statement in a pseudo-conflict and situations that the statements can coexist, which is inferred by the system.

#### 4 Outline of the Summarization System

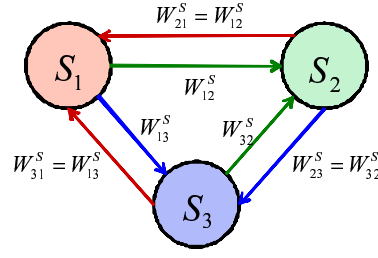
Figure 2 shows the outline of the system to generate survey reports including mediatory summaries. First, a user inputs a query statement such as “Are diesel engines harmful to the environment?” The system retrieves Web documents relevant to the query via TSUBAKI<sup>1</sup> (Shinzato *et al.*, 2008), which is an open-search engine infrastructure based on deep natural language processing and can accept a statement as a query.

We plan to cooperate with the STATEMENT MAP generator (Murakami *et al.*, 2009) in order to analyze semantic relations between statements in Web documents. Since analyzing semantic relations in all retrieved Web documents is expected to take a long time, we must narrow down the total amount of text to be analyzed for semantic relations. Therefore, the retrieved documents are segmented into passages, and some passages including various argument points are selected and inputted to the STATEMENT MAP generator. Passages at this point differ from the ones outputted as a part of the survey report, and are simply recognized by the HTML tag structure. The way to extract passages including various argument points is described in section 5.2.

Using semantic relations in the STATEMENT MAP, the system searches for passages representing situations that opposing statements can coexist according to the contrast structure between statements. If the passages are found, the system extracts and presents a passage to mediate the opposing statements. Otherwise, a pair of passages including reasons and situations of each statement is presented in order to help a user judge which statement is agreed. The extraction of significant passages including reasons is described in section 5.3.

The process to generate survey reports contains two different types of extractions of passages. The first one is for passages to contain various argument points. The second one is for significant

<sup>1</sup> <http://tsubaki.ixnlp.nii.ac.jp/>



**Figure 3:** Example of sentences and weight of links between the sentences

passages to include reasons. Therefore, if these two processes are performed in a framework, it will be very efficient. We adopted the TextRank algorithm (Mihalcea and Tarau, 2004), which is a graph-based ranking algorithm for sentence extraction, because of the following reasons: (a) it realizes a united framework and (b) statements and semantic relations are represented as a network structure in the STATEMENT MAP. We describe the improvements in the next section.

## 5 TextRank for Passage Extraction

### 5.1 The TextRank Algorithm

Before we introduce the improved TextRank algorithm for passage extraction, we describe the original TextRank algorithm (Mihalcea and Tarau, 2004). The TextRank algorithm is an application of the PageRank algorithm (Brin and Page, 1998) to natural language documents. It is a graph-based ranking algorithm, which determines the significance of vertexes not using the local information of each vertex but using the global information computed from the entire graph. The idea is based on the fact that a significant vertex is defined as a vertex linked by many other vertexes. In the TextRank algorithm, sentences are regarded as vertexes and the weight of link  $W_{ij}^S$ , which is based on the similarity between sentences  $S_i$  and  $S_j$ , is calculated by the following equation (1).

$$W_{ij}^S = \text{Similarity}(S_i, S_j) = \frac{|\{w_k \mid w_k \in S_i \& w_k \in S_j\}|}{\log(|S_i|) + \log(|S_j|)} \quad (1)$$

Here,  $w_k$  is a noun in the sentences and  $|S_i|$  is the number of nouns in sentence  $S_i$ . Figure 3 shows the example of sentences and weight of links between the sentences. The significance  $TR(S_i)$  of the sentence  $S_i$  is calculated by the following equation (2).

$$TR(S_i) = (1 - d) + d * \sum_{S_j \in In(S_i)} \frac{W_{ij}^S}{\sum_{S_k \in Out(S_j)} W_{jk}^S} TR(S_j) \quad (2)$$

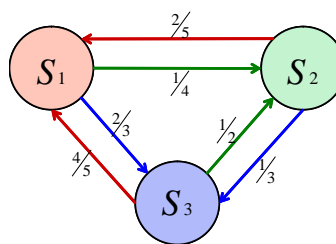
Here,  $In(S_i)$  is a set of sentences linked to  $S_i$ ,  $Out(S_j)$  is a set of sentences linked from  $S_j$ , and  $d$  is the factor that can be set between 0 and 1.  $d$  is set to 0.85, which was the same value adopted by Brin and Page (Brin and Page, 1998), in our experiment. The iteration of calculation with equation (2) is terminated when the following condition is satisfied.

$$\sum_{S_i \in G} (TR^{t+1}(S_i) - TR^t(S_i)) < \varepsilon \quad (3)$$

Here,  $G$  is a set of sentences in the graph and  $TR^t(S_i)$  is the significance of the sentence  $S_i$  at the  $t$ -th iteration.  $\varepsilon$  is a threshold and is set to 0.0001, which was the same value adopted by Mihalcea and Tarau (Mihalcea and Tarau, 2004), in our experiment.

**Table 1:** Example of sentences and words that appeared in the sentences

	$S_1$	$S_2$	$S_3$
$w_1$	✓	✓	✓
$w_2$	✓	✓	
$w_3$	✓	✓	
$w_4$	✓		
$w_5$	✓		
$w_6$		✓	✓
$w_7$			✓



**Figure 4:** Example of the calculation of uniqueness

## 5.2 Passage Extraction by Taking Account of Variety of Argument Points

Mihalcea and Tarau (Mihalcea and Tarau, 2004) used the similarity between sentences based on their content overlap as a weight of link in the calculation of significance. However, such similarity may be unsuitable for the extraction of passages including various argument points because argument points of sentences extracted by similarity depend on majority. Therefore, we introduce *uniqueness* that is a new measure to guarantee more argument points as possible in a limited number of sentences, and assign a value of uniqueness as a weight to each link between sentences.

As we assume that a noun is an argument point, the uniqueness of sentence  $S_i$  against sentence  $S_j$  is defined as the number of tokens of nouns that are included in  $S_i$  but not included in  $S_j$ . A value of uniqueness  $W_{ij}^U$  of the link from sentence  $S_i$  to sentence  $S_j$  is calculated by the following equation (4).

$$W_{ij}^U = \frac{|\{w_k \mid w_k \in S_i \& w_k \notin S_j\}|}{|S_j|} \quad (4)$$

Note that the links of uniqueness are directed while the ones of similarity are undirected. For example, in the case of sentences that include nouns, as shown in Table 1, the values of uniqueness between the sentences are shown in Figure 4.

Although using similarity guarantees that a value flowed out from a vertex is equal to the one flowed into the vertex, using uniqueness does not guarantee it. Therefore, scores of vertexes are normalized by dividing them by the sum of scores of all vertexes in a graph.

We conduct an experiment to ascertain whether using uniqueness is effective. In this study, we use the corpus constructed by Nakano et al. (Nakano *et al.*, 2008). They collected Web documents in several topics and manually created dozens of extractive summaries for the information credibility (Nakano *et al.*, 2008). In this experiment, we used the document sets and the summaries of the following five topics: women-only passenger car, metabolic syndrome, minus ion, electrolyzed-reduced water, and jury system in Japan. Table 2 shows the number of all sentences in each topic. Three or four human subjects are engaged in preparing each topic. When a human subject created a summary from Web documents collected in terms of a topic, we recorded sentences that the subject judged as they were going to be useful to create a summary. The sentences that all subjects judged are used as the answer sentences in this experiment. Table 2 shows the number of the answer sentences in each topic.

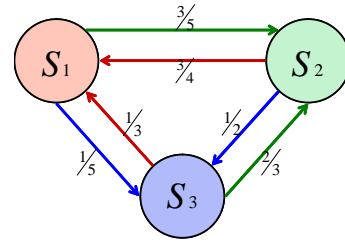
We prepare the following two baseline methods. The first one is the original TextRank method using similarity. The second one is a method using *fungibility*, which is the opposite concept of uniqueness and is calculated by the following equation (5) instead of equation (4).

$$W_{ij}^F = \frac{|\{w_k \mid w_k \in S_i \& w_k \in S_j\}|}{|S_i|} \quad (5)$$

Figure 5 shows an example of fungibility between sentences, as shown in Table 1.

**Table 2:** The number of all sentences and the number of answer sentences in each topic

Topic	All	Answer
Women-only passenger car	780	170
Metabolic syndrome	392	69
Minus ion	743	198
Electrolyzed-reduced water	651	104
Jury system in Japan	504	166



**Figure 5:** Example of the calculation of fungibility

**Table 3:** R-precision of sentence extraction with each method

Topic	Similarity	Fungibility	Uniqueness
Women-only passenger car	0.529	0.488	0.541
Metabolic syndrome	0.594	0.652	0.580
Minus ion	0.465	0.409	0.556
Electrolyzed-reduced water	0.356	0.288	0.356
Jury system in Japan	0.452	0.349	0.530

We evaluate the proposed method and the baselines by the R-precision. R-precision is the precision of the set of sentences up to rank R, where R is the number of answer sentences. R-precision is one of measures that evaluate the effectiveness of ranking mechanism.

Table 3 shows the result. The method using uniqueness is comparable or more effective than other methods in all topics except the topic of metabolic syndrome. The method using uniqueness could not improve the ranking of sentences in the topic of metabolic syndrome, because expressions in an argument point of the topic are not diverse. In other words, the number of sentences, which are not similar to other sentences and are significant, in the topic of metabolic syndrome is fewer than the ones in other topics.

### 5.3 Extraction of Significant Passages by Taking Account of Semantic Relation

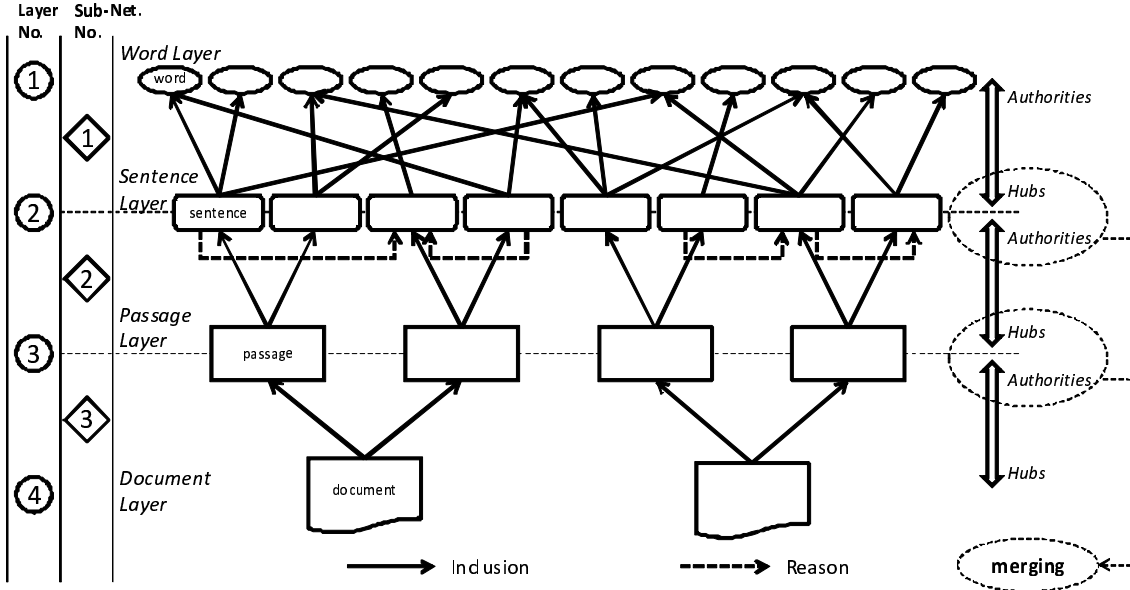
Mihalcea proposed another type of the TextRank algorithm (Mihalcea, 2005), which is based on the HITS algorithm that is one of the algorithms for ranking Web pages (Kleinberg, 1999). Unlike the PageRank algorithm, the HITS algorithm uses two measures, which are authorities and hubs, for scoring significance of vertexes. The values of authority and hub,  $HITS_A^t(V_i)$  and  $HITS_H^t(V_i)$ , of a vertex  $V_i$  at the  $t$ -th iteration in the TextRank algorithm based on the HITS algorithm are calculated by the following equations.

$$HITS_A^{t+1}(V_i) = \sum_{V_j \in In(V_i)} W_{ji} \cdot HITS_H^t(V_j) \quad (6)$$

$$HITS_H^{t+1}(V_i) = \sum_{V_j \in Out(V_i)} W_{ij} \cdot HITS_A^t(V_j) \quad (7)$$

The HITS algorithm is based on the concept that Web pages that are known as hubs lead a user to other authoritative pages. Similarly, statements that are known as reasons support a user’s judgment on the credibility of other statements corresponding to the conclusion. Therefore, we regard reasons and conclusions as hubs and authorities, respectively, and apply the iterative calculation between hubs and authorities to ones between reasons and conclusions.

Next, we assume the influence of significance between passages and sentences as follows. As the passages that include many significant sentences are significant, passages are regarded as hubs of sentences. Similarly, the significance between sentences and words and between documents and



**Figure 6:** Example of the multi-layer network structure

passages are influenced by each other. Therefore, such linguistic units and relations between them are represented by a multi-layer network structure, as shown in Figure 6. Note that reason relations are linked between vertexes in the same layer, while inclusions are linked between vertexes in different layers.

The multi-layer network consists of three sub-networks, as shown in Figure 6, and hubs and authorities of vertexes are calculated in each sub-network. We define the number of layers and sub-networks, as shown in Figure 6, and calculate  $HITS_{A,l}^{t+1}(V_i)$  and  $HITS_{H,l}^{t+1}(V_i)$ , which are the values of authority and hub of a vertex  $V_i$  in a layer  $l$  at the  $t$ -th iteration, respectively, according to the following equations.

$$HITS_{A,l}^{t+1}(V_i) = \sum_{V_j \in In(V_i)} W_{ji} \cdot HITS_{H,l+1}^t(V_j) \quad (8)$$

$$HITS_{H,l}^{t+1}(V_i) = \sum_{V_j \in Out(V_i)} W_{ij} \cdot HITS_{A,l-1}^t(V_j) \quad (9)$$

Note that the authorities of vertexes in a layer  $l$  are calculated in the sub-network numbered to  $l$ , while the hubs of the vertexes are calculated in the sub-network numbered to  $l - 1$ . In order to merge hubs and authorities calculated in different sub-networks, the merged values,  $mHITS_{A,l}^t(V_i)$  and  $mHITS_{H,l}^t(V_i)$ , are calculated by the following equations.

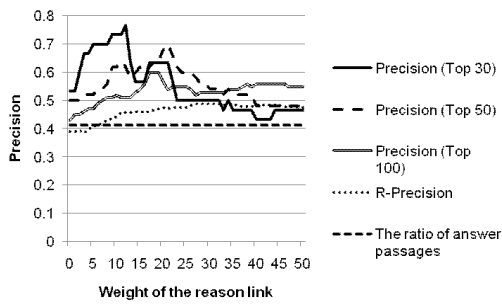
$$mHITS_{A,l}^t(V_i) = \alpha \cdot HITS_{A,l}^t(V_i) + (1 - \alpha) \cdot HITS_{H,l}^t(V_i) \quad (10)$$

$$mHITS_{H,l}^t(V_i) = (1 - \beta) \cdot HITS_{A,l}^t(V_i) + \beta \cdot HITS_{H,l}^t(V_i) \quad (11)$$

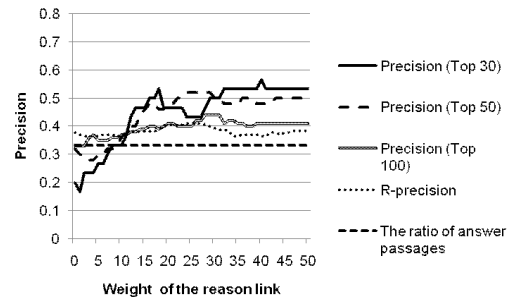
Here,  $\alpha$  and  $\beta$  are coefficients and are set to 0.5 in this study. The merged values are used for new values of the vertexes in the next iteration.

We conduct an experiment to certify whether a method using the multi-layer network structure introduced reason relations is effective. We prepared the answer data by using human subjects. Human subjects created extractive summaries. Four human subjects are engaged in preparing two topics: “Non-washed rice is delicious” and “LASIK operation is painful.” In this experiment, we define a passage as a series of five sentences. We define answer passages as passages including at





**Figure 7:** Precision of the topic “Non-washed rice is delicious”



**Figure 8:** Precision of the topic “LASIK operation is painful”

least two sentences that are actually used in extractive summaries. Note that some of the answer sentences adopted in the experiment in section 5.2 are not used in the extractive summaries. In other words, we adopted a stricter criterion of significant sentence in this experiment. The numbers of the passages and the answer passages in the topic “Non-washed rice is delicious” are 534 and 221 respectively. The numbers in the topic “LASIK operation is painful” are 535 and 178 respectively. Reason relations are also annotated by the human subjects. The weight of link of the reason relations is changed from 0 to 50 in order to investigate how much the reason relations have an effect on the extraction of significant passages. We evaluate the method by using R-precision and precision of the top 30, 50 and 100 passages.

Figures 7 and 8 show the experimental results. In the general tendency of the results, the proposed method using the multi-layer network structure in which reason relations are introduced can extract significant passages with higher precision than the ratio of answer passages that means the precision of extracting random passages. The precisions in the top 30 results are the highest in comparison with the precisions in the top 50 results and the top 100 results, and the precisions in the top 50 results is higher than the precisions in the top 100 results. In other words, significant passages are ranked higher by the proposed method. This means that the proposed method is effective. In the top 30 results, if the weight of the reason link is set to an optimum value, the precision in each topic is considerably improved to 76.7% and 56.7% respectively. However, at weights of reason link lower than 10, the precision in the topic of “LASIK operation is painful” is lower than the ration of answer passages. This is because the TextRank algorithm gives high significance to passages including frequent words, and the number of frequent words in the answer passages of the topic “LASIK operation is painful” is fewer than the one in the other topic. However, increasing the weight of reason link can improve results in such topic. Therefore, the proposed method based on the multi-layer network in which reason relations are introduced is effective for extraction of significant passages that do not include frequent words. One of the future works is a study on the optimum weight of reason link. As shown in Figure 7, the highest value in precision is not in the condition when the weight of the reason link is set to 50. It means that the significance of passage is not always determined by the reason relations. Therefore, introducing other semantic relations than reasons is another future work.

## 6 Conclusion

In this paper, we discussed the summarization for supporting a user’s judgment on the credibility of information on the Web. We focus on the coexistence of a statement and its opposing statement, and attempt to develop a system to generate survey reports that contain *mediatory summaries*, which are defined as passages extracted from Web documents in order to present situations in which these opposing statements can coexist. We described the outline of the summarization

system. As elementary techniques for the system, we improved the TextRank algorithm from the following two viewpoints: (a) passage extraction by taking account of a variety of argument points and (b) passage extraction by taking account of reason relation. From the experimental result of passage extraction in the former viewpoint, we confirmed that the method using uniqueness can extract significant passages, which are actually judged as significant by human assessors, with higher precision than the one using similarity and the one using fungibility. From the experimental result of passage extraction in the latter viewpoint, the method using a multi-layer network introduced reason relations can extract significant passages with higher precision than the ratio of answer passages in the general tendency of the results. We confirmed that significant passages were ranked higher by the proposed method. Future works are as follows: (1) to investigate a method based on the combination of uniqueness and fungibility, (2) to investigate the optimum weight of the reason link, and (3) to examine the effect of semantic relations besides reasons.

## References

- Ando, S., K. Inui, M. Ishioroshi, Y. Matsumoto, S. Matsuyoshi, R. Miyazaki, T. Mori, K. Murakami, M. Nakano, S. Nakazawa, Y. Okajima, H. Shibuki and T. Suzuki. 2008. Information Credibility Survey Reporting: A Prototype System and Project Roadmap. *Proceedings of 2nd International Symposium on Universal Communication (ISUC 2008)*.
- Brin, S. and L. Page. 1998. The anatomy of a large-scale hypertextual Web search engine. *Computer Networks and ISDN Systems*, 30(1-7), 107-117
- Haghighi, A. and L. Vanderwende. 2009. Exploring Content Models for Multi-Document Summarization. *Proceedings of The 2009 Annual Conference of the North American Chapter of the Association for Computational Linguistics on Human Language Technology (HLT-NAACL 2009)*, pp.362-370.
- Keinberg, J. 1999. Authoritative sources in a hyperlinked environment. *Journal of the ACM*, 46(5), 604-632.
- Mihalcea, R. 2005. Language Independent Extractive Summarization. *Proceedings of the 43th Annual Meeting of the Association for Computational Linguistics (ACL 2005)*, pp.25-30.
- Mihalcea, R. and P. Tarau. 2004. TextRank: Bringing order into texts. In L. Dekang and W. Dekai, editors, *Proceeding of the 2004 Conference on Empirical Methods in Natural Language Processing (EMNLP 2004)*, pp.404-411.
- Murakami, K., E. Nichols, S. Matsuyoshi, A. Sumida, S. Masuda, K. Inui and Y. Matsumoto. 2009. Statement Map: Assisting Information Credibility Analysis by Visualizing Arguments. *Proceedings of 3rd Workshop on Information Credibility on the Web (WICOW 2009)*, pp.43-50.
- Nakano, M., H. Shibuki, R. Miyazaki, M. Ishioroshi and T. Mori. 2008. Experiment and analysis of manual summarization aiming at automated summarization that supports the judgment of the information credibility. Information Processing Society of Japan (IPSJ) SIG Technical Report 2008-NL-187, 107-114 (In Japanese).
- Okazaki, N., Y. Matsuo, N. Matsumura and M. Ishizuka. 2003. Sentence Extraction by Spreading Activation with Similarity Measure. *IEICE Transactions on Information and Systems (Special Issue on Text Processing for Information Access)*, E86-D(9), 915-926.
- Shinzato, K., T. Shibata, D. Kawahara, C. Hashimoto and S. Kurohashi. 2008. TSUBAKI: An Open Search Engine Infrastructure for Developing New Information Access Methodology. *Proceedings of Third International Joint Conference on Natural Language Processing (IJCNLP 2008)*, pp.189-196.