Experimental Study of Discovering Essential Information from Customer Inquiry

Keiko Shimazu
Information Media Laboratory
Corporate Research Group
Fuji Xerox Co., Ltd.
430 Sakai Nakai-machi
Ashigarakami-gun
Kanagawa 259-0157 JAPAN
+81-465-80-2327
keiko.shimazu@fujixerox.co.jp

Atsuhito Momma
Information Media Laboratory
Corporate Research Group
Fuji Xerox Co., Ltd.
430 Sakai Nakai-machi
Ashigarakami-gun
Kanagawa 259-0157 JAPAN
+81-465-80-2321
atsuhito.momma@fujixerox.co.jp

Koichi Furukawa
Graduate School of Media and Governance
Keio University
5322 Endo Fujisawa-shi
Kanagawa 252-8520 JAPAN
+81-466-47-5111
furukawa@sfc.keio.ac.jp

ABSTRACT

This paper reports the result of our experimental study on a new method of applying an association rule miner to discover useful information from customer inquiry database in a call center of a company. It has been claimed that association rule mining is not suited for text mining. To overcome this problem, we propose (1) to generate sequential data set of words with dependency structure from the Japanese text database, and (2) to employ a new method for extracting meaningful association rules by applying a new rule selection criterion. Each inquiry in the sequential data was represented as a list of word pairs, each of which consists of a verb and its dependent noun. The association rules were induced regarding each pair of words as an item. The rule selection criterion comes from our principle that we put heavier weights to co-occurrence of multiple items more than single item occurrence. We regarded a rule important if the existence of the items in the rule body significantly affects the occurrence of the item in the rule head. The selected rules were then categorized to form meaningful information classes. With this method, we succeeded in extracting useful information classes from the text database, which were not acquired by only simple keyword retrieval. Also, inquiries with multiple aspects were properly classified into corresponding multiple categories.

Categories and Subject Descriptors

H.2.8 [**Database Management**]: Database Applications - *Data Mining.*

J.1 [**Computer Applications**]: Administrative Data Processing – *Business, Marketing.*

General Terms

Experimentation

Keywords

data mining, text mining, association rule, prior confidence, posterior confidence, syntactic dependency

1. INTRODUCTION

Recent studies on text mining, a research area in data mining, are drawing attention among researchers [15]. It is due to the recent rapid increase in digital documents on the Internet [3, 6].

Our purpose is the extraction of important information from call center data. Call centers are now regarded as the most important interface for companies to communicate with their customers. Call center operators have to respond to various requests from their customers without offending them. In addition, the record is said to contain precious information for understanding the trends of the customers. However, the operators often overlook such hidden information in analyzing the records because they tend to rely solely on static clustering measures and empirical keywords. Also, information that the call center operators deal with in their daily work is quite different from that for the top management with global and long-term viewpoints. Consequently, call center records are not fully utilized for discovering business chances or avoiding risks. This fact is called the bottleneck of the electronic Customer Relationship Management (eCRM). In our experiments, we aimed to discover important information that could not have been extracted by a set of keywords specified by a domain specialist.

In our experiment, call center inquiry data were first converted into sets of items that consist of important words with syntactic dependencies. Then, meaningful item sequences were generated by selecting items necessary for capturing the meaning of the original sentences and sorting them accordingly. The three rule mining algorithms, (1) typical association rule mining algorithm [1], (2) association rule mining with a novel rule selection threshold, the difference between prior and posterior confidence, and (3) the exception rule discovery method [16] were applied to the data and the results were compared.

The remainder of this paper is structured as follows: Section 2 summarizes related work. Section 3 explains the framework that we adopted for important information discovery from text data.

Section 4 reports the results of the experiments on the call center records, the analysis for which is given in Section 5. Section 6 gives the conclusion of the paper.

2. RELATED WORK

2.1 Techniques for Keyword Discovery

For discovery of important information and knowledge from large amount of data, identification of important words that represent the content of each document is the key [13]. While many systems that have been put into practice calculate word importance based on its frequency [14], methods that fulfill its shortcomings also exist.

For example, Matsuo proposed an algorithm that extracts important words by excluding general words [8]. It first identifies frequent words within a document and then calculates co-occurrence frequency of each frequent word and other words and finally extracts words that have higher concurrence. This is an extension of KeyGraph [11] and obtains better results in some cases than its origin. In addition, another study reports that nouns are more important than verbs [10] and that semantic relationship among phrases is more useful [4]. Another study claims that the objective of a sentence can be identified by analyzing expression at the end of it [17].

Besides identifying important words, some researchers introduce structure to represent relationship among these words. For example, Zaki proposed an efficient algorithm to induce frequent trees in a forest consisting of ordered labeled rooted trees [19]. He also proposed an algorithm for inducing frequent sequential patterns in [18], where he reported experimental results of word sequential pattern induction.

Apriori 4.03 [2] proposes new methods that effectively select a subset of a large amount of association rules by comparing prior confidence and posterior confidence. It can be regarded as an application of Matsuo's method into association rule mining. In our experiment, we employed this method to extract interesting classes.

2.2 Text Mining from Call Center Information

Representative applications of text mining methods into call center data include one being performed by Nasukawa's system that is based on natural language processing techniques [9]. It allows users to analyze call center data from various viewpoints such as categories of inquiries with similar contents and characteristics of inquiries that require longer time to deal with. It has already been put into practice with practical features such as one that seamlessly analyze inquiries regardless of their media types such as voices through telephone and electronic mail contents. However, to the best of our knowledge, reports that discover clues for brand new business chances and ones that allow users to proactively avoid risks have not been published.

In our experiment, we aimed to identify important information that cannot be discovered by simply utilizing keywords that call center operators customarily use in their daily operation.

3. FRAMEWORK FOR MEANINGFUL INFORMATION DISCOVERY FROM TEXT

With the purpose of identifying important information that cannot be obtained by methods such as keyword search and conventional text data classification techniques, we adopted the overall framework shown in Figure 1. The framework is based on propositions made by professionals who deal with a large amount of text data (inquiries from customers) in their daily work. They claim that important information can be intuitively and effectively discovered not by reading whole documents but by simply skimming sequences of important words within them.

3.1 Conversion to Sequential Data

3.1.1 Parsing and Dependency Information Attachment

In our experiments, we induced association rules by regarding words in the inquiries as basic components, and adopting minimum difference between the prior confidence and the posterior confidence as the rule selection criterion, rather than the conventional threshold with minimum support and the minimum confidence. We expected that our method enable to identify important information classes that cannot be obtained with conventional methods such as keyword search. As the first step, each sentence in inquiries was segmented into words with a dictionary developed solely for the inquiries[1]. Then, sequential data were generated with dependency structure (step1 in Figure 2)[2]. We believe that it does not only eliminate redundancy in interpretation, but also contribute to accurate meaning identification.

Then, we converted inquiry records into sequences of combinations of two words, between which syntactical dependencies exist. For example, the sentence "私は用紙入れの中に紙を入れた" , which means "I put papers in a tray.", is converted to "(私→入れる), (用紙入れ→中), (中→入れる), (紙→入れる)", which is like "(I → put), (a tray → in), (in → put), (paper → put).

Similarly, sentences "紙を用紙入れの中に入れた", which means "Papers was put in a tray", and "私は紙を入れた", which means "I put paper in.", are individually converted to "(紙→入れる), (用紙入れ→入れる)", which is like " (paper → put), (a tray → put)", and "(私→入れる), (紙→入れる)", which is like " (I → put), (papers → put)". One can determine that these three sentences have a same meaning because they posses the same combination "(紙→入れる)", which is like " (papers → put)".

We believe that this method makes it possible to identify meaning of a sentence definitely by excluding ambiguity in interpretation, even if various expressions for a single meaning exist.

[1] Morphological Analyzer "Chasen" [7] was employed for word segmentation.

[2] During word segmentation, meanings of auxiliary verbs were interpreted and were incorporated to the sequential data in order to accurately capture meanings of the original sentences.

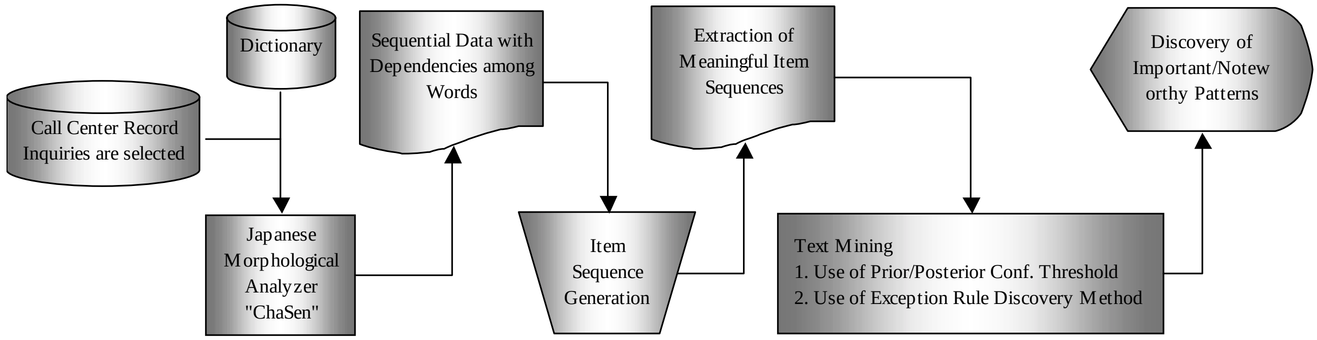

Figure 1. Framework for Important Informaiton Discovery from Text Data

3.1.2 Extraction of Meaningful Sequential Data

Next, meaningful sequences of items (pairs of words where dependency relationship exists) were generated by selecting items necessary for capturing the meaning of the original sentences and sorting them accordingly. Then, following our novel finding that sentence meaning can be obtained solely by browsing its association rules, association rules with bodies consisting of the items and heads representing the classification class for the sequence (step 2 in Figure 2). For example, while a rule "(こと 使用する), (当社 OS 環境), (変更する こと) → ClassN", which contains "(since, use), (change, use), (our company, common OS environment) → ClassN", can be generated, it is excluded in this step since the items in the body are meaningless. An overview of the generated information classes is shown in Table 1. For example, inquiries belonging to Class22 are questions regarding machine operations and/or functional specifications. In Table 1, classes are organized according to the targets and the objectives. This arrangement makes it possible to access all inquiries belonging to "question" by specifying Class20, for example.

3.2 Important Pattern Discovery from Sequential Data

3.2.1 Important Information Acquisition with Prior and Posterior Confidence

In text mining, it is not always the case that frequent words are important, particularly when focusing on the contents [11]. Rather, we regarded a rule important if the existence of the items in the rule body significantly affects the occurrence of the item in the rule head and applied this principle to rule selection. This criterion can be seen as a simplification of Matsuo's proposal [8], and it has already been incorporated in Apriori 4.03 [2]. A rule is selected if the difference between its prior and posterior confidences exceeds a given threshold. The prior confidence of an association rule is the confidence of a rule with the head of the rule and an empty body. The posterior confidence is the confidence of the rule itself. For example, given an association rule "{cheese, tomato} ⇒ {bread}", its prior confidence is the confidence of a rule "{ }⇒{bread}" and its posterior confidence is the confidence of the rule itself (i.e. "{cheese, tomato} ⇒ {bread}").

Our assumption is that word co-occurrence is useful in extracting meaning of sentences. Thus, we expected that meaningful and useful rules can be extracted, while different from common trends in majority of data, by selecting rules whose difference between prior and posterior confidence is large.

3.2.2 Exception Rule Discovery by Default Rules

Machine learning algorithms, often employed as data mining engine, characteristics of target data are induced by explicitly dividing the data into positive examples and negative examples. Meanwhile, sometimes it is difficult to decide whether an example is a positive example or a negative example. Inoue et al. [5] proposed a novel-learning algorithm that is able to deal with incomplete information by means of extended logic programming. They claim that their method enables to learn default rules [12] including exceptions. Suzuki proposed a method to discover default rules and exception rules simultaneously, by regarding rules with high support and confidence as the defaults [16]. When a rule Y → X is captured as a default rule, a related rule Z /→ X' is identified and an exception rule X,Z → X' is induced. Here, X' refers to an atom with the same attributes as X and different attribute values, and /→ indicates that the left part is insufficient

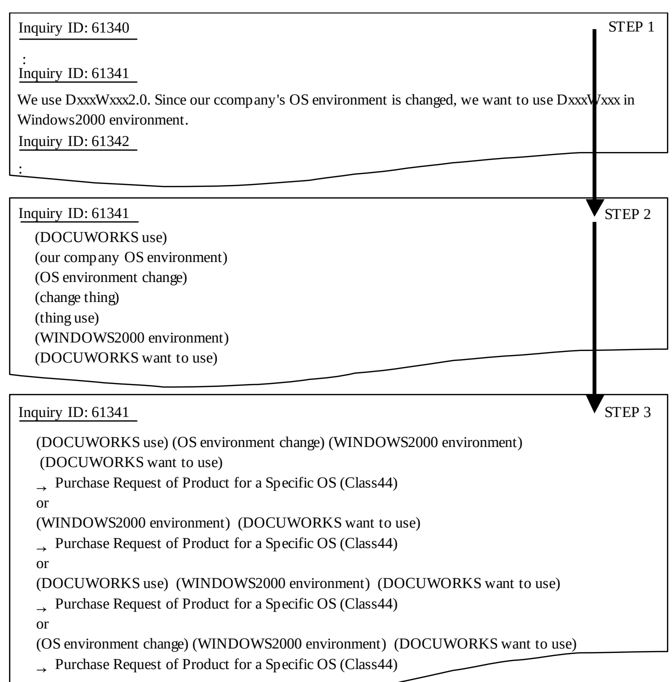

Figure 2. Preparation of Sequential Data

Table 1. Classification Classes

	Class Category	Class01	Class02	Class03	Class04	……
Class Category	Target / Purpose	Pre-purchase Info. Request/Ordering Procedure	Oprations/Functional Specification	Performance	OS-specific Support	……
Class10	Complaint	Class11	Class12	Class13	Class14	……
Class20	Questionnaire	Class21	Class22	Class23	Class24	……
Class30	Acknowledgement	Class31	Class32	Class33	Class34	……
⋮	⋮	⋮	⋮	⋮	⋮	⋮

to explain the right part.

We conducted experiments to verify our assumption that the rules induced with the method described in section 3.2.1 already include those obtained by Suzuki's method. As a result, one meaningful exception rule was acquired by Suzuki's method, which was one of and the 20 rules that were obtained by our method.

4. EXPERIMENTS ON CALL CENTER RECORDS

4.1 Target Data

In this experiment, we used 626 inquiries about a specific product from April 1 to July 31, 2002. The same experiments were conducted on 725 inquiries about the same product from August 1 to October 30 2002, in order to identify differences in the results.

4.2 Meaningful Sequential Data Preparation

Inquiries were converted into sequential data consisting of words with all dependency information, referring to the dictionary dedicated for the data. The average number of items per an inquiry is 14.9. After meaningful item sequences were obtained in section 3.1.2, each inquiry contains 7.1 items in average. Overall data contains 9,598 word occurrences, 1,950 distinctive words, and 8,157 distinctive items.

When dependencies to verbs were solely employed in sequential data generation, each inquiry contains 7.5 items in average.

4.3 Pattern Acquisition from Sequential Data

4.3.1 Meaningful Sequential Patterns Irrelevant to Frequency

In order to obtain important rules that are characterized by word co-occurrence, rules whose difference between the prior confidence and the posterior confidence exceeds 30% were extracted (the minimum support was set to 0.02). Table 2 lists 20 rules with the largest differences. Seven rules are regarding machine operations and/or functional specifications (Class02), which include two questionnaires (Class22). In addition to three rules on purchase operations, inquiries on the compatibility among different operating systems (Class04) and questions on performance (Class03), both of which were not extracted with a conventional rule selection criteria (i.e., the minimum support of 0.6 and the minimum confidence of 40%), were obtained. Four rules appeared in both rule selection criteria. The maximal number and the minimum number of inquiries that match each rule are 10 and 2, whose average is 4.6.

4.3.2 Exception Patterns

Treating rules obtained with the above conventional selection criteria as the default rules, references rules were searched by switching attribute values (classification classes) in the head. Then, two exception rules in Table 3 were obtained as rules whose confidence increases by adding items in reference rule's body into their body. The first rule turned out to be already obtained in section 4.3.1. The confidence factors for acquiring each exception rule are 45.45 and 26.08.

5. DISCUSSION

5.1 Effectiveness of Preprocessing

Applied studies on text mining for call center data claim the importance of verbs, rather than nouns [10], and expressions at the end of sentences [17]. When incorporating only dependency information related to verbs generated sequential data, 741 meaningful association rules were identified among 10,333 rules, or about 7%. In this experiment, target data for text mining were generated by selecting rules that preserve meanings of the source data, after preparing item sequences with all available dependencies. With this enhancement, the ratio of meaningless rules, such as “(こと, できる) ⇒ Request for Information/Questionnaire on Purchase Operation”, where “こと” and “できる” mean “matter” and “able”, reduced down to 5%. In experiments on data sets without rule selection, only 75 meaningful rules were obtained among 380 rules. Thus, we suppose that our proposed preprocessing, consisting of dependency information addition and meaningful item sequence selection, certainly contributes to effective rule induction.

5.2 Meaningful Information Acquisition

In the head of association rules induced with the conventional selection criteria, “Machine Operation/Functional Specification” appears in half of the rules. “Pre-purchase Information Request/Purchase Procedure”, “Operational Scheme”, and “Complaint” appear in two rules, respectively. Assuming that this is the overall trend of the inquiries, this corresponds to the trend reported in monthly reports periodically prepared by a call center staff member. In other words, these facts can be obtained by accessing inquiry database on the Intranet. On the other hand, our experiments revealed the unnoticed facts: (1) inquiries on usage of files created by the product and attached to email are noticeable among those in “Machine Operation/Functional Specification” class, and (2) inquiries and claims on the usage and specification of a new version of the product released in a year ago exist. These facts prove that our method is able to capture keywords that tend to be overseen, buried in a pile of information.

Table 2. Meaningful Rules Independent from Frequency

				diff. of conf.	confidence	Inquiries
(where, download)			⇒ Question on Download Site	96.63	100.00	3
(personal computer, OS)			⇒ OS-specific Support	95.28	100.00	3
(support window, teach)			⇒ Question on Operational Scheme	95.01	100.00	4
(HTTP, exist),	(FTP, exist),	(HTTP, FTP)	⇒ Download Site	91.51	100.00	4
(FTP, download),	(HTTP, download),	(where, download)	⇒ Operation Procedure/Functional Spec.	91.51	100.00	3
(telephone number, teach)			⇒ Operational Scheme	89.35	100.00	3
(soft product XYZ, catalog)			⇒ Pre-purchase Info Req./Purchase Proc.	79.38	100.00	5
(license, purchase[demand])			⇒ Pre-purchase Info Req./Purchase Proc.	79.38	100.00	3
(restriction, exist [question]）			⇒ Question on Performance	73.79	75.00	3
(HP, watch)			⇒ Operating on Download Site	71.50	75.00	3
(OS, do)			⇒ OS-specific Support	70.28	75.00	3
(soft product XYZ, purchase[demand]）			⇒ Question on Purchase Procedure	65.16	75.00	3
(operation, listen[demand]）			⇒ Question on Op. Proc./Func. Spec.	61.23	83.33	10
(usage, teach[demand])			⇒ Question on Op. Proc./Func. Spec.	52.90	75.00	3
(soft product XYZ, Ver3.02)			⇒ Notification of Function Currently in Use	52.90	75.00	3
(file, open),	(mail, attach)		⇒ Operation Procedure/Functional Spec.	46.23	100.00	10
(trial version, download),	(soft product XYZ, trial version)		⇒ Operation Procedure/Functional Spec.	41.51	50.00	7
(trial version, use),	(soft product XYZ, trial version)		⇒ Operation Procedure/Functional Spec.	41.51	50.00	7
(X company, scanner)	(scanner, use),	(Ver10.2, use)	⇒ Complaint/Dissatisfaction/Uncertainty	49.60	71.43	2
(scanner, use),	(X company, scanner)		⇒ Operation Procedure/Functional Spec.	33.73	87.50	7

Table 3. Exception Patterns

				confidence	inquiries
(X company, scanner)	(scanner, use),	(Ver10.2, use)	⇒ Complaint/Dissatisfaction/Uncertainty	71.43	2
(this, use),	(matter, able)		⇒ Operation Proc./Func. Specification	87.50	3

Among association rules derived with the novel selection criterion, six rules regarding file download operations from homepages appeared among the 20 rules. By closely observing these rules, the fact that users are perplexed in selecting proper connection protocol (i.e. HTTP or FTP) was obtained. This fact was also unnoticed in daily compiling operations. In addition, the fact that users of a particular scanner tend to issue complaints ("(X 社, スキャナ), (スキャナ, 使う), (Ver10.2, 使う) ⇒ 苦情・不満・不安", or "(manufactured by X, scanner), (scanner, use), (Ver10.2, use) ⇒ Complaint/dissatisfaction /Uncertainty") was also overlooked by professionals, which was also obtained the exception rule discovery method [16].

In general, relevant keywords are not provided in advance in extracting noteworthy information from a large amount of data. Thus, keywords with low frequency tend to be forgotten and not to be used in text retrieval. Conversely, we believe that our method is effective in discovering noteworthy trends, independent from the overall trends of inquiries.

5.3 Effectiveness of Exception Rules

Among the two exception rules obtained in the experiments, one rule was also obtained by our proposed method and other rule was useless. No exception rule was obtained from another data set.

In the exception rule discovery, a similar rule selection criterion was employed as the criterion based on prior and posterior confidences. In an example on medical data shown in [16], rules whose support and confidence exceed 20% and 75% were selected as the default rules. The maximum confidence of the reference rules is 50%, and rules whose support and confidence exceed 3.6% and 80% were identified as the exception rules. In our experiments, lower thresholds were employed in identifying frequent patterns with the conventional rule selection criterion and the minimum difference between prior and posterior confidences was set to less than 50% in exception rule extraction, because with the thresholds in the example in [16] did not provide successful results. We suppose that this is caused by the difference in target data.

The call center inquiries for our experiments differ from that in [16] in that various expressions for the same status exist. They were processed as different items. For example, "(Ver10.2, 購入する), (xxxx, yyyy) ⇒ ClassN", which means "(Ver10.2, purchase), (xxxx, yyyy) ⇒ ClassN", and "(Ver10.2, 買う), (xxxx, yyyy) ⇒ ClassN", which means "(Ver10.2, buy), (xxxx, yyyy) ⇒ ClassN", were individually induced as distinct rules. Due to this fact, both support and confidence tend to be low.

We believe that the key cause for the low support and confidence is the characteristics of the target data (total word occurrences: 9,598, distinctive words: 1,950, distinct items: 8,157) and that improvement in the dictionary (thesaurus) for preprocessing enables the employment of higher thresholds and leads more useful results. However, we also suppose that the complete elimination of variations in transcription is impossible, as long as we deal with "natural language". In addition, association rules for medical data tend to employ definite situations such as "recovery" or "death" in the rule head, while the default rules for call center data dynamically change over time. We suppose that our method is able to obtain important data without omission in the domains with volatile default rules than the exception rule discovery method.

6. SUMMARY

In this paper, we conducted experiments to identify meaningful information by applying association rule induction algorithm to text mining. The distinctive features of our method are (1) dependencies among words were incorporated into the sequential data, and (2) association rules were selected based on the difference between the prior confidence and the posterior confidence. As a result, meaningful classes covering relatively few data were successfully extracted. In applying the exception rule discovery method based on non-monotonic reasoning to the same data set, we could not obtain effective results because of its strict rule selection threshold. Conversely, our method, with a relatively relaxed selection criterion, was able to acquire useful rules.

In addition, we indicated the following two reasons for the effectiveness of our proposal where sequential data with dependency information are generated in preprocessing: (1) it preserves meanings of the raw data, and (2) it is applicable to a conventional data mining technique (i.e., association rule induction). In particular, reason (1) proves a suggestion posed by a domain professional and we expect that our method will contribute to lighten the arrangement and reporting operation by call center staff members. Meanwhile, Zaki claims that meaning of text can be preserved in text mining by taking word order into consideration in preprocessing [19]. We are afraid that incorporation of word order may cause further diffusion of association rules (i.e., lower support and confidence). We are going to closely examine this issue.

In this experiment, we found a rule that can be seen as the clue for proactive risk avoidance (issues caused by the combination of a specific scanner machine and a software product). However, from another data set, such rule was not obtained. We suppose that a novel method must be developed to identify information that can be used as future prediction.

7. ACKNOWLEDGMENTS

We express our deep gratitude to Mr. Takemi Yamazaki, who has given indispensable supports and considerations for our research and experiments. In addition, we thank Mr. Yohei Yamane and Mr. Tetsushi Sakurai for their devoted assistance in experiments.

8. REFERENCES

[1] Agrawal R. Fast Algorithms for Data Mining Applications, in Proceedings of the 20th International Conference on Very Large Databases (Santiago Chile, 1994), 487-489.

[2] Borgel C. Apriori: Finding Association Rules/Hyperedges with the Apriori Algorithm. http://fuzzy.cs.uni-magdeburg.de/~borgelt/apriori/ .

[3] Hearst M. A. Untangling Text Data Mining (invited paper). in Proceedings of the 37th Annual Meeting of the Association for Computational Linguistics (College Park MD, 1999).

[4] Hisamitsu T., Niwa Y., and Tsujii J. A Method of Measuring Term Representativeness - Baseline Method Using Co-occurrence Distribution. in Proceedings of the 18th International Conference on Computational Linguistics (Saabrucken Germany, July 2000), 320-326.

[5] Inoue K., and Kudoh Y. Learning extended logic programs. in Proceedings of the Fifteenth International Joint Conference on Artificial Intelligence (Nagoya Japan, August 1997), 176-181.

[6] Laurence S., and Giles L. Searching the World Wide Web. Science, 280, 5360 (1998), 98-100.

[7] Matsumoto Y., Kitauchi A., Yamashita T., Hirano Y., Matsuda H., Takaoka K., and Asahara M. Morphological Analysis System ChaSen version 2.2.1 Manual. http://chasen.aist-nara.ac.jp/chasen/doc/chasen-2.2.1.pdf .

[8] Matsuo Y., Ohsawa Y., and Ishizuka M. KeyWorld: Extracting Keywords in a Document as a Small World. in Proceedings of the Fourth International Conference on Discovery Science (Washington D.C., 2001), 271-281.

[9] Nasukawa T., and Nagano T. Text analysis and knowledge mining system. IBM Systems Journal 40, 4 (Winter 2001), 967-984.

[10] Nagano T., Takeda K., and Nasukawa T. Information Extraction for Text Mining. in IPSJ SIG Notes FI60-5 (2000), 31-38 (in Japanese).

[11] Ohsawa Y., Benson N. E., and Yachida M. KeyGraph: Automatic Indexing by Co-occurrence Graph based on Building Construction Metaphor. in Proceedings of 5th Advanced Digital Library Conference (Santa Barbara CA, April 1998), 12-18.

[12] Reiter R. A Logic for Default Reasoning. Artificial Intelligence 13, 2 (1980), 81-132.

[13] Sakurai S., Ichimura Y., Suyama A., and Orihara R. Inductive learning of a knowledge dictionary for a text mining system. In Proceedings of the 14th International Conference on Industrial and Engineering Applications of Artificial Intelligence and Expert Systems (Budapest Hungary, June 2001), 247-252.

[14] Segal R., and Kephart J. MailCat: An Intelligent Assistant for Organizing E-Mail. in Proceedings of the 3rd International Conference on Autonomous Agents (Seattle WA, May 1999), 276-282.

[15] Smyth P., Pregibon D., and Faloutsos C. Data-driven evolution of data mining algorithms. Commun. ACM 45, 8 (2002), 33-37.

[16] Suzuki, E. and Tsumoto, S. Evaluating Hypothesis-Driven Exception-Rule Discovery with Medical Data Sets. In Proceedings of the Fourth Pacific-Asia Conference on Knowledge Discovery and Data Mining (Kyoto Japan, April 2000), 208-211.

[17] Tanabe T., Yoshimura K., and Shudo K. Modality Expressions in Japanese and Their Automatic Paraphrasing. in Proceedings of the Sixth Natural Language Processing Pacific Rim Symposium (Tokyo Japan, 2001), 507-512.

[18] Zaki M. J. Efficient Enumeration of Frequent Sequences. in Proceedings of the seventh international conference on Information and knowledge management (Bethesda MD, November 1998), 68-75.

[19] Zaki M. J. Efficiently Mining Frequent Trees in a Forest. in Proceedings of the eighth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (Edmonton Canada, July 2002), 71-80.