AIC を用いた重要語抽出手法の提案・評価

Proposal and Evaluation of Significant Words Selection Method based on AIC

大平 茂輝　　白井 克彦

Shigeki Ohira　　Katsuhiko Shirai

早稲田大学理工学部

School of Science and Engineering, Waseda University,

Tokyo, 169-8555

{ohira, shirai}@shirai.info.waseda.ac.jp

Abstract

We have participated in "ad hoc IR task". Possibly it was more suitable for participation in "automatic term recognition and role analysis task" as a result. However, this is the first time that we actually make a information retrieval system, and so we participated in more standard task, ad hoc IR task, for the time being.

In this research, we present an AIC based method for the selection of significant words in information retrieval, and the comparison experiment with the conventional technique using the χ^2-test was performed. Therefore, it turns out that the high search accuracy is obtained, as compared with χ^2-test in which setting up a suitable level of significance is difficult. Moreover, it shows that by using only term of 1/7 to 1/9 of the whole selected by this AIC based methods, the search accuracy is close to that of the baseline (all terms in use).

Keyword

Significant Words Selection, AIC(Akaike's Information-theoretic Criterion), χ^2-(chi-square) test

1 Introduction

In recent years, research of large vocabulary continuous speech recognition is progressing greatly. Outstanding speech-recognition software such as "Via Voice(IBM)" (the number of vocabularies about 44,000)[1] has appeared.

As a practical toolkit which attains high recognition accuracy and a real time response, there is "Japanese Dictation ToolKit"[2] of IPA (Information-technology Promotion Agency, Japan), in which a word dictionary of 20,000 vocabularies is prepared.

However, few practical speech dialogue systems have been presented, such as ATIS(Air Travel Information System)[3] and the speech translation system[4] of ATR which task is travel guidance. Moreover, such systems only have vocabularies of several thousand words.

To appear in the NTCIR Workshop, 1999, Tokyo, Japan.

If the task of the spoken dialogue system is resticted, such as an simple guidance system, robot operation, etc., such a small vocabulary may be sufficient. However in tasks like information retrieval, the number of words in the vocabulary used there increases according to the number and type. Therefore, it is necessary to use several sets of dictionary.

Thus, from the viewpoint of speech recognition technology, it is important to select significant words, and narrow down the number of words to use, based on the task or the user's requirements.

Moreover, also in the viewpoint of information retrieval technology, it is thought that it is very desirable to extract beforehand an effective vocabulary of significant words as keywords for retrieval, since using an unsuitable vocabulary as a keyword reduces search accuracy.

In this research, we will propose a significant words selection method suitable for the information retrieval by speech.

2 The Significant Words Selection Method Using χ^2-test

As a technique of determining significant words, the method of verifying the correlation by χ^2-test[5], between the word which appears in a document and the category which is assigned to a document is proposed.

Given the condition that each document in the data collection has at least one category code, and each word has a unique ID, the significance of each word is calculated by the following process.

(1) When considering w as the event an object word appears in a document, and c as the event in which the object category is assigned to a document, the 4 numbers of the following events are counted.

n_{11} : number of docs in which $w \cap c$ is realized
n_{12} : number of docs in which $w \cap \neg c$ is realized
n_{21} : number of docs in which $\neg w \cap c$ is realized
n_{22} : number of docs in which $\neg w \cap \neg c$ is realized

At this time, the correlation between the appearance of a word and the assignment of category is expressed in the following contingency table (Table 1) in which term ID and category codes are used as a key.

	c	$\neg c$
w	n_{11}	n_{12}
$\neg w$	n_{21}	n_{22}

Table 1: Contingency Table

(2) Using the contingency table, χ^2-test (test of independence) is made. It is that the following statistic ${\chi_0}^2(w,c)$ is calculated under null hypothesis in which "w and c are independent mutually".

$$
\begin{aligned}
{\chi_0}^2(w,c) &= \frac{\left(n_{11} - c_1 \cdot \frac{w_1}{N}\right)^2}{c_1 \cdot \frac{w_1}{N}} + \frac{\left(n_{12} - c_2 \cdot \frac{w_1}{N}\right)^2}{c_2 \cdot \frac{w_1}{N}} \\
&+ \frac{\left(n_{21} - c_1 \cdot \frac{w_2}{N}\right)^2}{c_1 \cdot \frac{w_2}{N}} + \frac{\left(n_{22} - c_2 \cdot \frac{w_2}{N}\right)^2}{c_2 \cdot \frac{w_2}{N}}
\end{aligned}
$$

$$
\begin{aligned}
(\quad & N = n_{11} + n_{12} + n_{21} + n_{22}, \\
& w_1 = n_{11} + n_{12}, \; w_2 = n_{21} + n_{22}, \\
& c_1 = n_{11} + n_{21}, \; c_2 = n_{12} + n_{22} \quad)
\end{aligned}
$$

(3) In the contingency table, ${\chi_0}^2(w,c)$ is followd by χ^2-distribution of which degrees of freedom $f = (2-1)(2-1) = 1$.

If test is made on level of significance α, rejection region of null hypothesis is $\chi^2(f,\alpha)$. For example, when degrees of freedom is 1 and level of significance is 0.05(5%), rejection region $\chi^2(1, 0.05)$ is 3.84 refering to χ^2-distribution table (Table 2).

degrees of freedom	level of significance					
	0.500	0.100	0.050	0.025	0.010	0.005
1	0.45	2.71	3.84	5.02	6.63	7.88
2	1.39	4.61	5.99	7.38	9.21	10.60
3	2.37	6.25	7.81	9.35	11.34	12.84
4	3.36	7.78	9.49	11.14	13.28	14.86
5	4.35	9.24	11.07	12.83	15.09	16.75

Table 2: χ^2-distribution table

Therefore, if

$$ {\chi_0}^2(w,c) > \chi^2(f,\alpha) $$

is fulfilled, null hypothesis is rejected and w and c have correlation. $P(c)$, appearance probability of c, is considered as the weight of a category.
And $I(W)$, the significance of each word W, is defined with the following formula.

$$ I(W) = \sum_{c \in \{x | {\chi_0}^2(w,x) > \chi^2(f,\alpha)\}} P(c) $$

However, it is difficult to set up suitably level of significance required for χ^2-test. Moreover, in the task of the similar document retrieval which searches few similar documents from a lot of documents, since an extreme difference is in both numbers of documents, it is thought that the method using χ^2-test which is weak in the fluctuation with the delicate number of events is not effective.

Then, we will propose a significant words selection method based on Akaike's Information-theoretic Criterion (AIC)[6], as a statistical method in which level of significance is not needed and which is robust for the fluctuation of data.

3 The Significant Words Selection Method based on AIC

3.1 AIC (Akaike's Information-theoretic Criterion)

"Logarithmic Likelihood", i.e., the accuracy of a model, and the number of parameters used for logarithmic likelihood calculation, have a trade-off relationship. AIC is a method which is capable to solve this trade-off well. The general formula of AIC is as follows.

$$
\begin{aligned}
AIC &= -2 \times (\texttt{logarithmic likelihood}) \\
&\quad +2 \times (\texttt{the number of parameters of a model})
\end{aligned}
$$

AIC value is calculated for every model. The model which has the minimum AIC value is considered the optimum model.

3.2 Significant Word Selection Method

We explain about significant word selection method based on AIC. In the same way as the above χ^2-test based method, given the condition that each document in the data collection has at least one category code, and each word has a unique ID, the significance of each word is calculated by the following process.

(1) The logarithmic likelihood MLL and AIC value are calculated about the model assumed that c and w occur independently (Independent Model, IM) and the model assumed that a dependency is amond both (Dependent Model, DM)[7] in the above contingency table in which term ID and category codes are used as a key.

$$
\begin{aligned}
MLL_{IM}(w,c) &= (n_{11} + n_{12}) \log(n_{11} + n_{12}) \\
&+ (n_{11} + n_{21}) \log(n_{11} + n_{21}) \\
&+ (n_{21} + n_{22}) \log(n_{21} + n_{22}) \\
&+ (n_{12} + n_{22}) \log(n_{12} + n_{22}) \\
&- 2N \log N \\
AIC_{IM}(w,c) &= -2 \times MLL_{IM}(w,c) + 2 \times 2 \\
MLL_{DM}(w,c) &= \sum_{i,j} n_{ij} \log n_{ij} - N \log N \\
AIC_{DM}(w,c) &= -2 \times MLL_{DM}(w,c) + 2 \times 3 \\
&\qquad (N = n_{11} + n_{12} + n_{21} + n_{22})
\end{aligned}
$$

(2) Based on the conditions that the model with a smaller AIC value is more optimal, the significance of a word is judged by the following formula.

$$ AIC_{DM}(w,c) < AIC_{IM}(w,c) $$

(3) If (2) is fulfilled, it consider that an appearance of object word W and category c have correlation. Furthermore, the weight of word at this time is considered as consisting of the weight of a category and the weight based on AIC value of a model. Total likelihood P with a higher weight is defined as the following formula using $P(c)$, appearance probability of c, and $P(DM)$ and $P(IM)$, occurence probability of each model.

$$ P = \frac{P(DM)}{P(IM)} \cdot P(c) $$

Moreover, multiplying the logarithm of both side of this equation by 2 and transforming it, the following formulas

$$\begin{aligned} 2\log P &= 2\log P(DM) - 2\log P(IM) + 2\log P(c) \\ 2\log P - 2 &= -(-2\times MLL_{DM} + 2\times 3) \\ &\quad +(-2\times MLL_{IM} + 2\times 2) \\ &\quad +2\log P(c) \\ &= AIC_{IM} - AIC_{DM} + 2\log P(c) \end{aligned}$$

are obtained. Then, it considers that the right side of this equation should have a bigger value in order to increase total likelihood.

Now $I(W)$, the significance of each word W, is defined with the following formula, and the way of this weighting is named AIC_th.

$$I(W) = \sum_{\dagger}\Big(AIC_{IM}(W,c) - AIC_{DM}(W,c) + 2\log P(c)\Big)$$

$$\dagger : c \in \{x | AIC_{DM}(W,x) < AIC_{IM}(W,x)\}$$

In contrast to such as theoretical word weighting, the way of weighting based on actual numeric value is named AIC_pr and defined as the following.

$AIC_{IM}(w,c) - AIC_{DM}(w,c) > 0$, the weight based on the difference of AIC value of a model, has the distribution as shown in Fig. 1 (number of 1 - 59 on the x axis corresponds to each category, and each plot point represents word).

It shows that the words differ widely in significance according to category. Therfore, it cannot be said that the significant word obtained is comprehensive.

Then, the logarithm is calculated after adding 1 to this value. The reason of calculating the logarithm is to hold down extremely high value, and of adding 1 is to make it value bigger than 0.

Moreover, the reason that their values were not devide by maximum for a difference of AIC value, and not make into one or less value is that we think the weight based on AIC value of a model is more important than the weight of category. In this way, the weight obtained has a distribution as shown in Fig. 2.

That which added these two weights is defined as $I(W)$, the significance of each word W, and it is expressed with the following formula.

$$I(W) = \sum_{\dagger}\Big(\log\big(1 + AIC_{IM}(W,c) - AIC_{DM}(W,c)\big) + P(c)\Big)$$

$$\dagger : c \in \{x | AIC_{DM}(W,x) < AIC_{IM}(W,x)\}$$

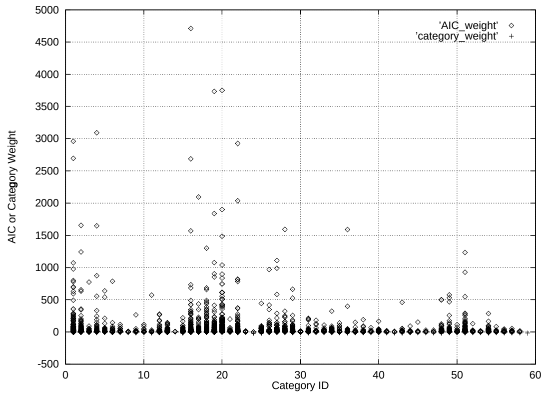

Figure 1: The distribution for a difference of AIC value (AIC_th)

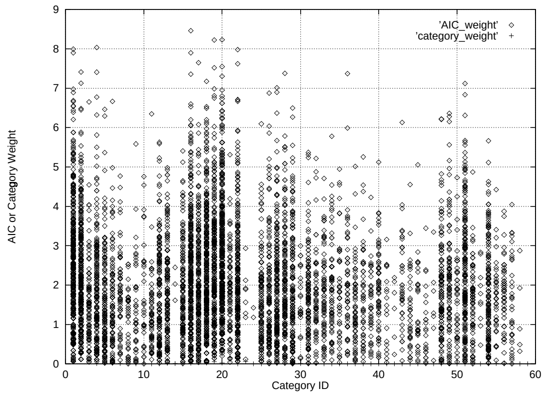

Figure 2: The distribution for a difference of AIC value (AIC_pr)

4 Outline of Information Retrieval System

4.1 Retrieval Model

Similarity $Sim(Q, D_d)$ between query Q and document D_d is calculated by the following formula based on the well known vector space model (w is the term weight).

Since the main object of this research is the evaluation of significant word selection, retrieval techniques such as relevance feedback and query expansion were not applied in this experiment.

$$Sim(Q, Dd) = cos(Q, D_d) = \frac{\sum_{t=1}^{n} w_{q,t}\cdot w_{d,t}}{\sqrt{\sum_{t=1}^{n} w_{q,t}^2}\cdot\sqrt{\sum_{t=1}^{n} w_{d,t}^2}}$$

4.2 Data Preprocessing

The following processes were made on the abstract data for preparation.

- All the half-size (8bit character) symbols in abstract data were removed.
- Chasen ver1.51[14] was used for the morphological analysis of each data. No changes were applied to the program environment or to the dictionary.

  Nouns (common, proper, verbal, time, name, place), numerals, adjectives, noun prefixes, noun nature noun suffixes, and undefined words were extracted for indexing.
- A stopword list consisted of 550 words (299 English, 251 Japanese) was prepared (see appendix).

5 Experiments

In this experiment, the abbreviation of the society was used as the category code for each document. For example, "The Institute of Electronics, Information and Communication Engineers" becomes "IEICE". Based on this categorization, the contingency table was made for all 352K words which occured in 339K docs and 59 categories.

There are three reasons the society name was used for categorization.

1. There is no document without one category.
2. In order to reduce the amount of calculation, it was necessary to set the number of categories to 100 or lower.
3. The correlation with a word and a society name is presumed to be an important key for document similarity.

5.1 Experiment Procedure

First, an experiment to determine the optimum level of significance for χ^2-test based word selection was made.

The level of significance was set to 5.0, 2.5, 1.0, 0.5(%), and the number of selected words was set to 10000, 20000, 30000, 40000, 50000. The precision and recall of the search was calculated for all conditions.

Next, an experiment was made to compare AIC and χ^2-test based word selection and random word extraction. For this experiment, the optimum level of significance for χ^2-test was set based on the results of the previous experiment. The number of selected words are also set as in the previous experiment.

All experiments are made by using the training search topics (30 query). Relevance assesments set used for evaluation of search accuracy is JE Collection 1 (JE-1).

5.2 Experiment Results

5.2.1 Determination of The Optimum Level of Significance in χ^2-test

The 11-point recall-precision and non inter-polated average precision for all relevant documents (average precision) in each number of significant words at the time of changing a level of significance in χ^2-test are shown in Figs.3-7, and Table.3.

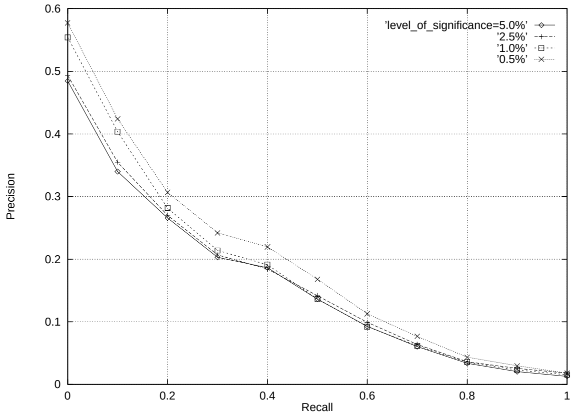

Figure 3: Comparison of the search accuracy by level of significance in χ^2-test
(number of significant words 10000)

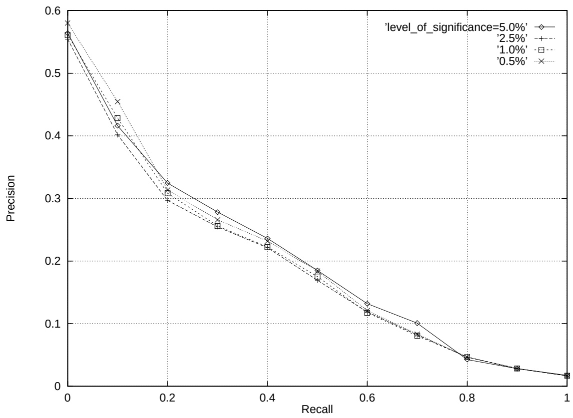

Figure 4: Comparison of the search accuracy by level of significance in χ^2-test
(number of significant words 20000)

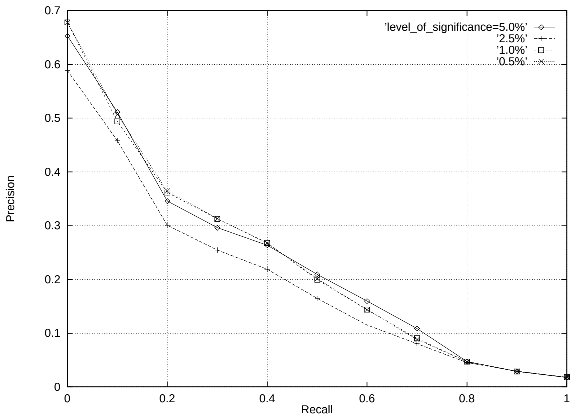

Figure 5: Comparison of the search accuracy by level of significance in χ^2-test
(number of significant words 30000)

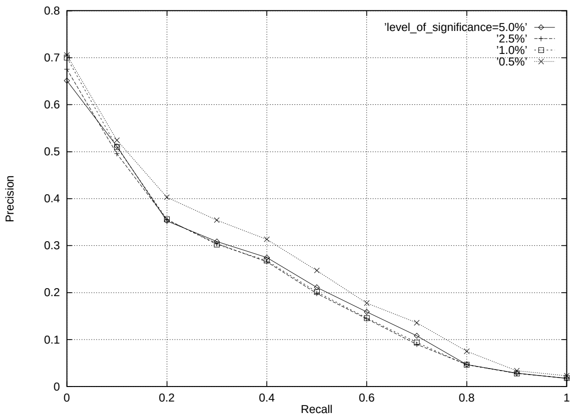

Figure 6: Comparison of the search accuracy by level of significance in χ^2-test
(number of significant words 40000)

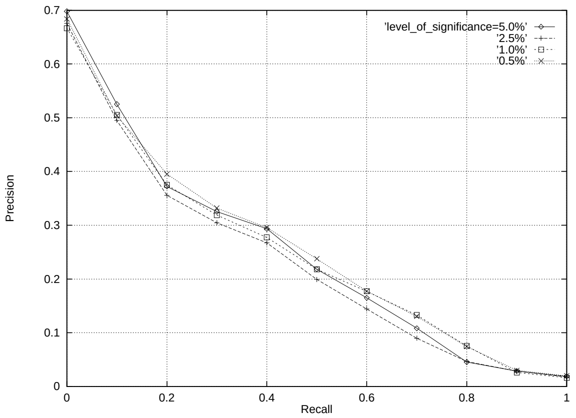

Figure 7: Comparison of the search accuracy by level of significance in χ^2-test
(number of significant words 50000)

LoS	number of significant words				
(%)	10,000	20,000	30,000	40,000	50,000
5.0	.1431	.1886	.2152	.2184	.2260
2.5	.1492	.1736	.1812	.2080	.2083
1.0	.1565	.1790	.2099	.2128	.2279
0.5	.1785	.1856	.2135	.2458	.2368

(LoS : Level of Significance)

Table 3: Relationship between level of significance and average precision
(using search topics for training)

From these results, it is clear that the significance level of 0.5% has achieved the best performance. However, a correlation between the significance level and search accuracy can not be observed from these results, since the search accuracy does not correspond with change of significance level. Theoretically, the accuracy should decline as the level of significance increases, but the results show that the accuracy for significance level 2.5% is generally worse than other levels.

Similar experiments were made on the evaluation topics. The average precision for these experiments are shown in Table 4. For this experiment, the accuracy of significance level 1.0% was the best, and the accuracy of significance level 2.5% was the worst.

These results prove that the optimization of significance level for χ^2-test based word selection is difficult.

LoS	number of significant words				
(%)	10,000	20,000	30,000	40,000	50,000
5.0	.1455	.1933	.1930	.1889	.1914
2.5	.1443	.1923	.1907	.1914	.1914
1.0	.1591	.1868	.1933	.1917	.1975
0.5	.1458	.1835	.1917	.1958	.1956

(LoS : Level of Significance)

Table 4: Relationship between level of significance and average precision
(using search topics for evaluation)

5.2.2 Comparison of three Methods, AIC, χ^2-test, and Random Extraction

The result when performing significant words selection and search experiment based on three methods, AIC, χ^2-test (level of significance 0.5%), and random extraction is shown below.

The 11-point recall-precision graph for every method is shown in Fig 8-11.

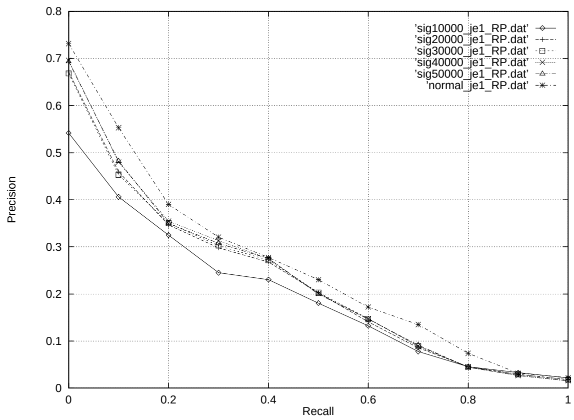

Figure 8: AIC_th

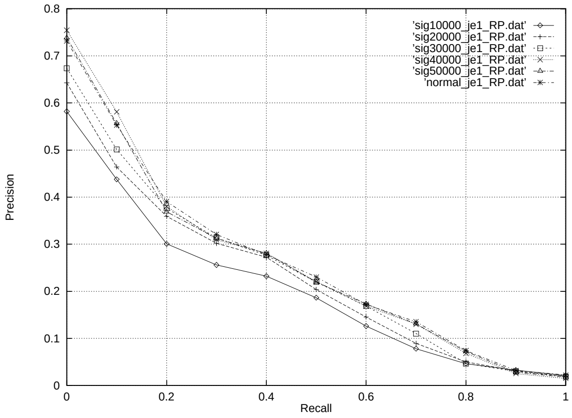

Figure 9: AIC_pr

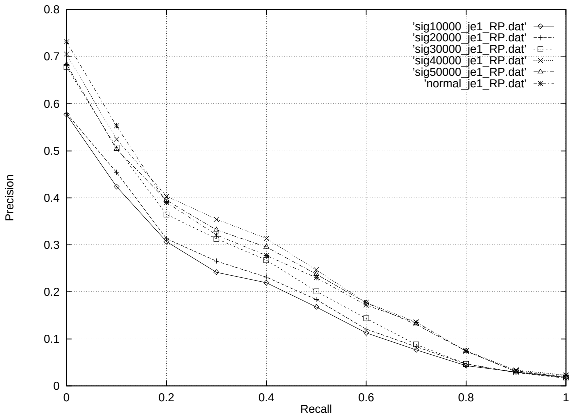

Figure 10: χ^2-test (level of significance 0.5%)

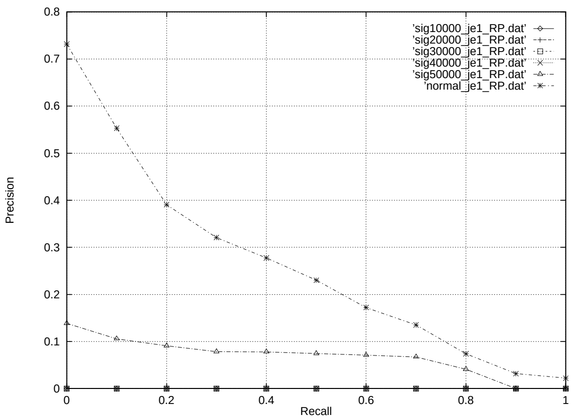

Figure 11: Random Extraction

The 11-point recall-precision graph for every number of significant words is as follows(Fig. 12-16). For comparison, the result at the time of setting up the worst level of significance (2.5%) was also plotted.

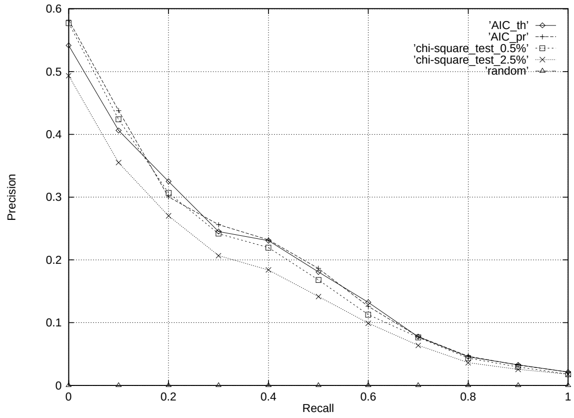

Figure 12: Search result by three methods (number of significant words 10000)

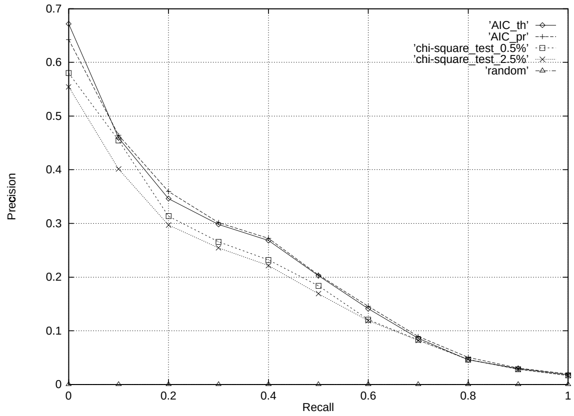

Figure 13: Search result by three methods (number of significant words 20000)

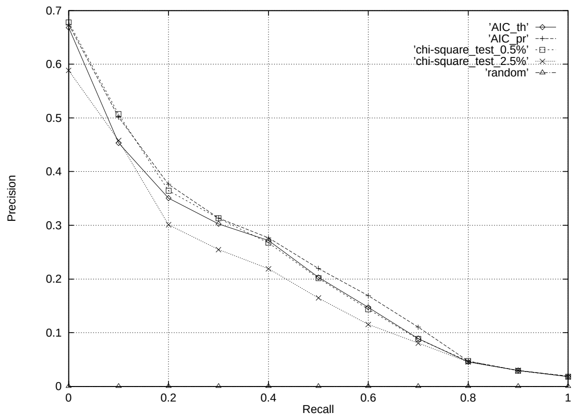

Figure 14: Search result by three methods (number of significant words 30000)

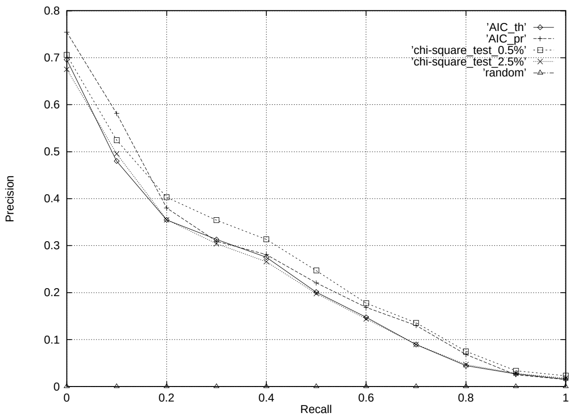

Figure 15: Search result by three methods (number of significant words 40000)

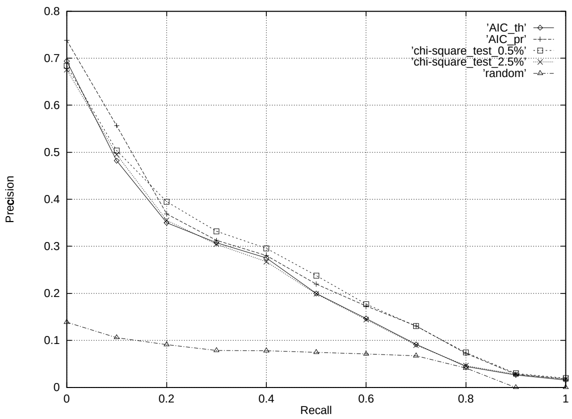

Figure 16: Search result by three methods (number of significant words 50000)

which compared three methods and average precision for every number of significant words is shown in Table 5.

	number of significant words				
	10,000	20,000	30,000	40,000	50,000
AIC_th	.1804 -.0229	.2033 -.0027	.2060 -.0051	.2111 +.0003	.2108 -
AIC_pr	.1820 -.0543	.2047 -.0316	.2202 -.0161	.2389 +.0026	.2363 -
χ^2-test (best:0.5%)	.1785 -.0583	.1856 -.0512	.2135 -.0233	.2458 +.0090	.2368 -
χ^2-test (worst:2.5%)	.1431 -.0652	.1736 -.0347	.1812 -.0271	.2080 -.0003	.2083 -
random	-	-	-	-	.0627

Table 5: Comparison of three methods, AIC, χ^2-test, random extraction

From Figs 12-16, it sis clear that AIC-based word selection is more effective than χ^2-test.

In the method based on AIC, AIC_th is compared with AIC_pr which changed the way of word weighting.

In AIC_th, if the number of significant words becomes 30,000 or more, accuracy is remarkably bad compared with χ^2-test based method which set up the optimum significance level. Although it means that the omission of effective words is large when increasing the number of significant words to extraction, accuracy is better than the case where the worst significance level is set up.

On the one hand, even if, as for the case of AIC_pr, the number of significant words becomes more than 30,000, accuracy is not bad so much. That is, it means that the omission when increasing the number of significant words is small, and can be said that it is stable way of weighting.

Moreover, the decline of accuracy corresponding to the reduction of selected words was lower for AIC compared to that of χ^2-test, as shown in Figs 7-10.

The experiment results show that even by selecting 40,000 words using the AIC-based method, the search accuracy was close to that of the baseline. This means that by using the proposed method, the dictionary can be reduced to 1/7 - 1/9 of the original, while preserving accuracy, proving the effectiveness of our method.

Meanwhile, for the conventional method based on χ^2-test, it is necessary to set an optimum significance level to achieve similar accuracy. However, our experiments could not clarify a theoretical method to do so.

6 Conclusion

In this research, the method of performing significant word selection based on AIC in information retrieval was proposed, and comparison experiment with the method using χ^2-test which is the conventional technique was performed.

Consequently, when the number of significant words is set as 10000, 20000, 30000, 40000, and 50000, collectively, the method using AIC has search accuracy better than the method using χ^2-test, proving that the terms obtained by the method based on AIC contains more effective terms for retrieval.

For the conventional χ^2-test based method, we clarified the difficulty to derive an optimum significance level. Our experiments also showed that the decline of accuracy resulting from the reduction of selected words was higher for χ^2-test compared to AIC.

These results prove the superiority of our method.

Future tasks include the application of IR techniques such as relevance feedback or query expansion, and the implementation of our method to a spoken dialogue IR system with an automatic vocabulary switching function.

Acknowledgements

We appreciate the members of the Knowledge-Based Information Processing Lab of KDD R&D Laboratories for their advice in this research.

Supplementary Explanation

There was a bug in the program used for official submission, so the results in this paper differ from submitted results.

References

[1] Interface, CQ 出版社, pp.100-108, August 1998.

[2] "Japanese Dictation ToolKit", http://www.itakura.nuee.nagoya-u.ac.jp/ takeda/IPA/

[3] Proc. ARPA Spoken Language Systems Technology Workshop, Morgan Kaufmann Publishers, San Francisco(1995).

[4] T. Shimizu, et al: "Fast word-graph generation for spontaneous conversational speech traslation", Proc. ICASSP97, pp.95-98(1997).

[5] Ng, Goh, Low: "Feature Selection, Perceptron Learning, and a Usability Case Study for Text Categorization," Proc of ACM SIGIR'97, 1997.

[6] Akaike, H.: A New Look at Statistical Model Identification. IEEE Transactions on Automatic Control, 19(1974) 716-723.

[7] 橋本, 松本: "テキスト検索のための相関ルールの発見", 信学技報 AI97-47, pp.13-18.

[8] 大平, 帆足, 松本, 橋本, 白井, "AIC を用いた重要語抽出手法の提案", 信学会春季総合大会, pp. 80, 1999.

[9] Witten, Ian H. et al: "Managing Gigabytes: Compressing and Indexing Documents and Images," Van Nostrand Reinhold, New York, pp. 141-148, 1994.

[10] G. Salton, M.J. McGill: "The SMART and SIRE Experimental Retrieval Systems," Readings in Information Retrieval, pp.381-399, 1997.

[11] Signhal, A.: "AT&T at TREC-6," Proceedings of the Sixth Text REtrieval Conference(TREC-6), pp. 215-225, 1998.

[12] Signhal, A. et al: "AT&T at TREC-7," Presentation Papers for Text REtrieval Conference, pp. 141-151, 1999.

[13] 岸田和明, "情報検索の理論と技術", 勁草書房, 1999.

[14] 「日本語形態素解析システム ChaSen」 http://cl.aist-nara.ac.jp/lab/nlt/chasen.html

Appendix : Stopword List

a about above aboveboard across after afterward afterwards again against albeit all almost alone along already also although always among amongst an and another any anyhow anyone anything anywhere are around as at

b be became become becomes becoming becouse been before beforehand behind being below beside besides between beyond both but by

c can cannot co could

d do down during

e each eg either else elsewhere enough etc even ever every everyone everything everywhere except

f few first for former formerly from further

g gradually

h had has have he hence her here hereafter hereby herein hereupon hers herself him himself his how however

i ie if in inc indeed into is it its itself

j

k

l last latter latterly least less ltd

m many may me meanwhile might more moreover most mostly much must my myself

n namely neither never nevertheless next no nobody none noone nor not nothing now nowhere

o of off often on once one only onto or other others otherwise our ours ourselves out over own

p per perhaps periodically

q

r rather

s same seem seemed seeming seems several she should since so some somehow someone something sometime sometimes somewhere still such

t than that the their them themselves then thence there thereafter thereby therefore therein thereupon these they this those though through throughout thru thus to together too toward towords

u under until up upon us usually

v very via

w was we well were what whaterer whatsoever when whence whenever whensoever where whereabouts whereafter whereas whereat whereby wherefrom wherein whereinto whereof whereon whereto whereunto whereupon wherever wherewith whether which whichever whichsoever while whilst whither whithout who whoever whole whom whomever whomsoever whose whosoever why will with within without would

x

y yet you your yours yourself yourselves

z

あらまし, あり方, あるとき, いか, いくつ, いくつか, いるか, いわば, お, おいで, おき, か, かち, かつ, から, がら, きっかけ, きを, こう, ここ, こと, このうち, このまま, この間, この際, この方, ごと, ごろ, さいこ, さま, しあい, しか, しそ, した, しゅん, じ, す, すい, すぎ, すなわち, すみずみ, ずつ, せんと, そう, そく, そこ, その, そのため, そのまま, そのもの, その後, その他, たい, たこ, ただ, ただし, たる, だれ, ち,っ, つき, つど, でどこ, ときの, ところ, としこ, とも, どうし, どの, どの事, どれだけ, ない, なすこと, など, により, はじめ, はや, ばく, ふ, べき, ほしい, ほんや, ま, ます, また, まとめ, み, みて, め, もつ, もと, ゆえ, よりこ, よる, ら, らば, りと, りや, をの

以後, 以前, 影響, 何等, 課題, 我, 我々, 解決, 解消, 解説, 改善, 概括, 概観, 概説, 概要, 概略, 各, 確認, 学問, 勘案, 観察, 寄稿, 期待, 機会, 疑問, 議論, 急務, 協議, 近年, 具体例, 経緯, 軽視, 決断, 結局, 結論, 月, 懸念, 検証, 検討, 研究, 研究開発, 見地, 見通し, 現行, 現在, 後半, 工夫, 考え, 考え方, 考案, 考究, 考察, 考慮, 行く末, 貢献, 克服, 今まで, 今回, 今後, 今日, 最近, 採用, 在り方, 参考, 仕方, 指摘, 氏, 私見, 試み, 試験, 時々刻々, 時代, 自ら, 自分, 実験, 実現, 実際, 実施, 実証, 若干, 若年, 従来, 重要視, 初め, 初めて, 初頭, 将来, 人々, 人達, 推察, 遂行, 数々, 数年, 世の中, 成果, 設け, 設置, 説明, 前半, 想定, 対処, 大半, 着目, 注目, 長い間, 長年, 直観, 追求, 通り, 提案, 提起, 提唱, 展望, 当社, 当初, 到来, 動向, 動揺, 導入, 特色, 日々, 年, 把握, 発表, 反映, 反省, 彼, 筆者, 普通, 分野, 文献, 平成, 報告, 暴露, 本, 本文, 本来, 未来, 魅力, 命題, 目標, 問題解決, 問題点, 約, 夕方, 余裕, 様子, 欲しい, 利用, 理解, 理由, 論文