PACIS 2004 Proceedings

December 2004

# Development of an Automatic Customer Service System on Computer Networks

Judy Tseng
*Chung Hua University*

Yi-Shiang Huang
*Chung Hua University*

Li-Chen Cheng
*Ta Hwa Institute of Technology*

Gwo-Jen Hwang
*National Chi Nan University*

# Development of an Automatic Customer Service System on Computer Networks

Judy C.R. Tseng and Yi-Shiang Huang
Department of Computer Science and Information Engineering
Chung Hua University
Hsinchu, 300, Taiwan, R.O.C.
E-mail: judycrt@chu.edu.tw


Li-Chen Cheng
Department of Information Management
Ta Hwa Institute of Technology
Hsinchu
Taiwan 307, R.O.C.
E-mail: cheng_lichen@yahoo.com.tw

Gwo-Jen Hwang
Information Management Department
National Chi Nan University
Pu-Li, Nan-Tou,
Taiwan 545, R.O.C.
FAX: 886-49-2915205
TEL: 886-915396558
E-mail: gjhwang@ncnu.edu.tw

## Abstract

*While the Internet is used by enterprises for contacting with customers, network-based customer services with e-mail or web-based interfaces have been developed to replace the traditional call centers with telephone interfaces. However, most existing network-based customer services heavily rely on manpower to reply e-mail or on-line requests submitted by customers, which not only increases the cost of providing customer services, but also delay the response time to complete the service procedure. To cope with these problems, this paper proposes an automatic customer service system, which can automatically handle customer requests by analyzing the contents of the requests and finding the most feasible answers from the Frequently-Asked Question (FAQ) database. If a customer is not satisfied with the reply, the system will forward the enquiry to the proper service personnel for further processing. An assistance model is also developed to help the service personnel to find suitable answers from the FAQ database and save new answer to it. Experimental results on practical applications showed that over 87.3% of users were satisfied with the replies given by the system; therefore, we conclude that the system can significantly reduce the service cost and provide more efficient and effective customer service.*

**Keyword:** Customer relationship management, Customer service system, Document matching, Call center, Internet applications


## 1. Introduction

Peppers and Rogers (1993) indicated that most businesses loose their customers by 25% rate per year in average. The cost of finding new ones, however, is five times that of keeping original customers. In addition, Reichheld and Sasser (1990) showed that enterprise would increase over sixty percent profit if they can build a strong relationship with customers. Therefore, one of the most important issues for increasing the competitive advantage of an enterprise is how to provide good service to customers.

Customer Relationship Management, CRM, is an integrated solutions designed to reduce costs and increase profitability by solidifying customer loyalty. A successful CRM includes four key elements, consumer commutation management, customer relationship management, decision support management and integration of financial or logistic system. WorldTalk Corporation 1999 estimates that over sixty million people prefer to deal with their work by e-mail. The investigation of Department of Industrial Technique in Taiwan also observed that 96% of companies have used e-mail to serve their customers (http://news.sina.com.tw/). It seems that e-mail will become a major communication management tool. However, an investigation in 1998 has pointed out that enterprises gave no response to 65% of customer's enquiry. The report also indicated that 63% of responded customers spend 5 days or longer to wait the response. From these observations, it seems to be valuable to build up an efficient system for dealing with customer requests.

E-mail based customer service system has been discussed and presented in several researches (Cohen, 1996; Chang & Hwang, 2001). In most existing systems, automatic e-mail reply functions and customer request classification functions are not taken into consideration. In (Tseng 2002), although the proposed system can automatically handle a customer request by analyzing its contents, the domain experts must assign each key word a weight value in advance, which will spend lots of human resources and the analysis results might be incorrect.

To cope with these problems, this study proposes a new system, ACCS (Automatic Customer Service System), which can automatically reply the requests from customers by invoking a knowledge base containing a set of Frequently Asked Questions (FAQ's) and the corresponding Character Vector (CV) of each FAQ. While receiving a request, ACCS will generate a CV for the request, and compares it with those of the FAQ's by employing the space vector concept to find the most feasible answer for the customer. If no feasible answer can be found, ACCS will forward the enquiry to the proper service staff. After the service personnel provide an answer, the new request and its answer will be recorded in the FAQ database.

Experimental results on practical applications showed that ACCS can correctly answer over 87.3% requests; therefore, the service cost can be significantly reduced and more efficient and effective customer service can be provided.


## 2. Relevant Works

On-line FAQ database are frequently adopted in traditional customer service systems, especially for web site users to search for the answers of their problems. Although providing FAQ database can reduce service cost, the web site users must spend lots of time on liking to the database and search for the answers they need.

Since1985, enterprises started to build up call centers to provide customer services. Most of the call centers only serve consumers at regular office hours. Moreover, the service personals must receive a series of training courses before they can offer proper customer service, which implies the requirement of a large amount of training cost. It can be seen that such traditional service systems not only provide inefficient and ineffective service, but also increase the service cost; therefore, the development of automatic customer service systems has become an important issue for enterprises.

To more accurately identify the requests of customers, researchers have proposed several methods in recent years. For example, Hoch (1994) presented the statistical methods of

information retrieval used at INFOCLAS, which is capable of classifying print business letters according to message types such as order, offer, enclosure, etc. In 1996, Cohen (1996) proposed the "key word spotting rule" approach, which can efficiently classify e-mails and has been applied to the development of e-mail management systems. In the meantime, Cooper (1996) reported the FAQfinder, which employed a set of weighted parameters to determine the similarity between the customer's request and the FAQ. Later, Li and Tseng (2001) proposed an Intelligent Network-based Customer Service System (INCSS), which assigns a weighted keyword set to each FAQ, and then compares the keyword set of the customer's request and that of each FAQ to find the most feasible answer. INCSS can automatically reply the requests submitted from the customers; however, the keyword set of each FAQ need to be assigned manually, which is time-consuming.

Other relevant studies include the issues of keyword retrieval and sentence similarity comparison. There are several ways to retrieval keywords, e.g., term extraction method (Arppe, 1995; Salton & Buckley, 1998), phrase extraction method (Krulwich, 1995), and Statistic analysis method (Jones & Gassie, 1990). Term extraction method can detect important terms from classified input text; phrase extraction method is used to detect phrases in the text; statistic analysis method can identify possible keywords from a large amount of unclassified input text by computing the occurrences of each keyword.

## 3. Structure of the Automatic Customer Service System

Figure 1 shows the structure of the Automatic Customer Service System (ACSS), which consists of four databases and six modules. The extracted keywords from consumer requests are kept in the Keyword Database. In the Characteristic Database, each question is represented as a CV. Questions and the corresponding answers are kept in FAQ databases. The workflow of ACSS is given as follows:

(1) Receives a user's e-mail.
(2) Invoke the Question Identification Module to decompose the e-mail into several terms defined in the Keyword Database. According to the formula this study is proposed, each term is transferred to a CV and kept in Characteristic Database.
(3) Invoke the Answer Judgment Module and Answer Picking-up to select the best answer in Characteristic Database and FAQ Database, and reply the answer to the user.
(4) If the customer is not satisfied by the answer provided, the system will distribute the enquiry to the proper service staff and an assistance system is also developed to help the service staff find a suitable answer from the FAQ database.
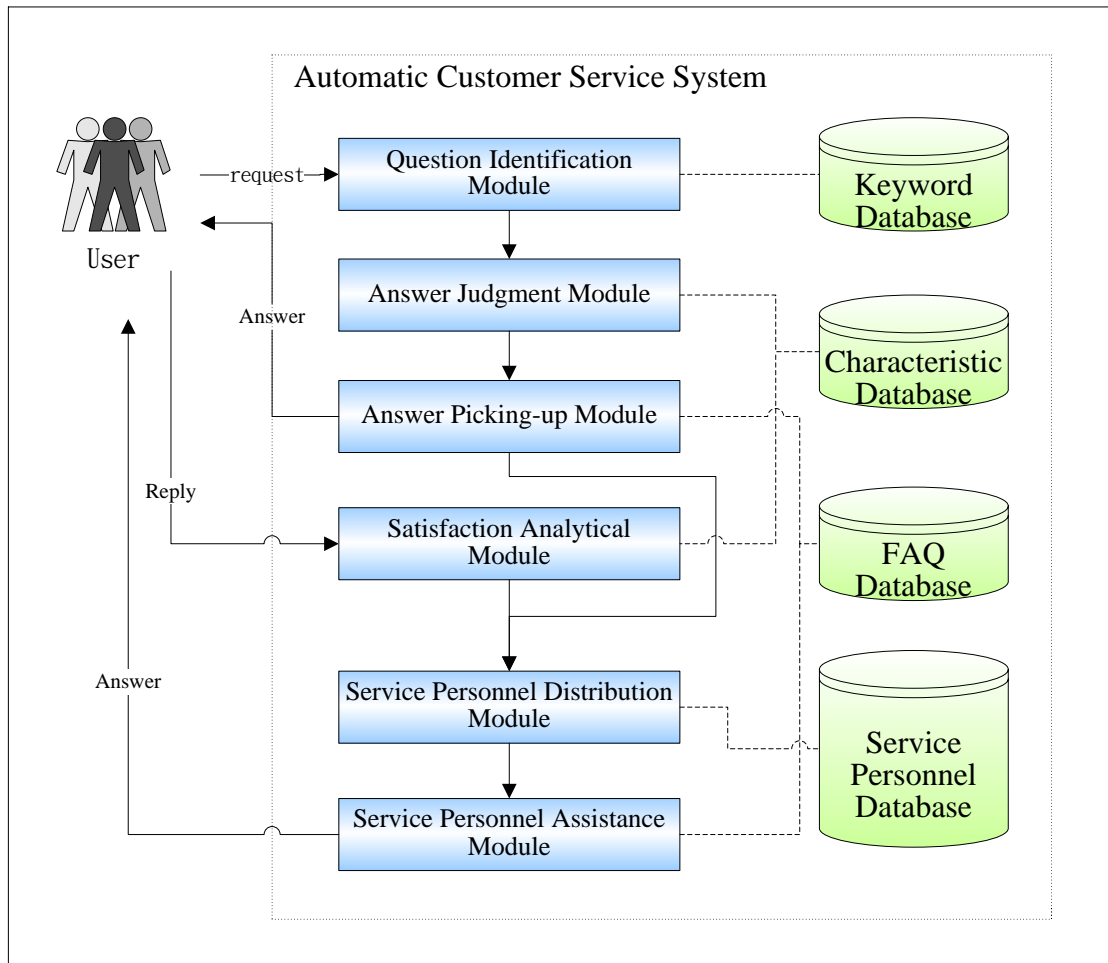
Fig. 1: Structure of the Automatic Customer Service System

## 4. Algorithms Used in ACCS

Base on the concept of vector space, this study propose a new method to category free text. In the followings, the request analysis algorithm and the similarity comparison algorithm of ACCS are addressed.

### 4.1 Request Analysis Algorithm

In ACCS, several sets of keywords have been defined by domain experts and were saved in the keyword database. A Character Vector (CV) of question $Q_i$ is defined as

$$CV(Q_i) = \left\{ (K_1, W_{i1}), (K_2, W_{i2})..(K_j, W_{ij})..(K_n, W_{in}) \right\} \quad \text{............................................(1)}$$

$$W_{ij} = \frac{f_{ij}}{\sum_{k=1}^{m} f_{kj}} \quad \text{.................................................................................................(2)}$$

where $K_j$ represents $j$th keyword and $f_{ij}$ is the number of occurrences of $K_j$ in question $Q_i$; $\sum_{k=1}^{m} f_{kj}$ is the total number of occurrences of $K_j$ in the FAQ database.

**[Example 1: Question part and answer part of the FAQ database]**
Assume that there are four FAQ's in the database. Table 1 shows the question part of a Chinese FAQ database and Table 2 shows the answer part. Table 3 depicts the keyword set and the number of occurrences of each keyword.

Table 1: Question part of the FAQ's

| serial number | Content of the questions |
|---|---|
| $Q_1$ | 電子報訂報或辦報要費用嗎？ |
| $Q_2$ | 取消訂閱電子報的流程。 |
| $Q_3$ | 使用奇摩分類網站排序優先服務的費用？ |
| $Q_4$ | 網站類目後面加個"@"代表什麼意義？ |

Table 2: Answer part of the FAQ's

| serial number | Answers to question in Table 4.1 |
|---|---|
| $Q_1$ | Yahoo!奇摩電子報為服務性質，不論訂報或辦報皆免費。 |
| $Q_2$ | 在您的個人電子報管理介面，也就是「我的電子報」，在您想要取消訂閱的電子報前勾選「取消訂閱」。您也可以在每封電子報最下方找到我要取消訂閱的連結，點下去後輸入您的 Yahoo!奇摩帳號和密碼，這樣系統也會自動幫您取消訂閱。 |
| $Q_3$ | 分類網站排序優先服務的月費費率依分類目錄之不同而訂定，價格表在您送出您要申請排序優先網站之網址之後，會有線上說明您網站所屬類別下網站排序優先的價格，價格大約在每月新台幣 3,000~20,000 元(含稅)之間。 |
| $Q_4$ | 當您在瀏覽 Yahoo!奇摩分類時，如果看到"@"就表示這個類目會同時現在不只一個地方。比方說：您可以同時在生活資訊和休閒天地的目錄下，找到時尚流行這個類目。又比方說：如果您要找查詢系統，不管是在網路指南或生活資訊下都找得到。 |

Table 3: Number of occurrences of each Keyword

| ID | Keyword | Number of occurrences | ID | Keyword | Number of occurrences |
|---|---|---|---|---|---|
| $K_1$ | 電子報 | 3 | $K_{12}$ | 奇摩 | 4 |
| $K_2$ | 訂報 | 2 | $K_{13}$ | 奇摩分類 | 2 |
| $K_3$ | 辦報 | 1 | $K_{14}$ | 排序 | 1 |
| $K_4$ | 訂閱 | 2 | $K_{15}$ | 搜尋引擎 | 0 |
| $K_5$ | 取消訂閱 | 1 | $K_{16}$ | 設定 | 0 |
| $K_6$ | 試閱 | 0 | $K_{17}$ | 大摩域 | 1 |
| $K_7$ | 費用 | 2 | $K_{18}$ | 網站類目 | 2 |
| $K_8$ | 流程 | 1 | $K_{19}$ | @ | 1 |
| $K_9$ | 即時 | 0 | $K_{20}$ | 會員檔案 | 0 |
| $K_{10}$ | 登入 | 0 | $K_{21}$ | 帳號 | 1 |
| $K_{11}$ | Yahoo | 6 | | | |

Each question in Table 1 can be represent as the following concept vectors:

$$CV(Q_1) = \left\{ \left(k_1, \frac{1}{3}\right), \left(k_2, \frac{1}{2}\right), \left(k_3, \frac{1}{1}\right), \left(k_7, \frac{1}{2}\right) \right\} \quad \text{.................(3)}$$

$$CV(Q_2) = \left\{ \left(k_1, \frac{1}{3}\right), \left(k_4, \frac{1}{2}\right), \left(k_5, \frac{1}{1}\right), \left(k_8, \frac{1}{1}\right) \right\} \quad \text{.................(4)}$$

$$CV(Q_3) = \left\{ \left(k_{12}, \frac{1}{4}\right), \left(k_{13}, \frac{1}{2}\right), \left(k_{14}, \frac{1}{1}\right) \right\} \quad \text{.................(5)}$$

$$CV(Q_4) = \left\{ \left(k_{18}, \frac{1}{2}\right), \left(k_{19}, \frac{1}{1}\right) \right\} \quad \text{.................(6)}$$

Similarly, the CV of customer's request $Q$ can be expressed as follows:

$$CV(Q) = \left\{ (K_1, W_1), (K_2, W_2)..(K_j, W_j)..(K_n, W_n) \right\} \quad \text{.................(7)}$$

$$W_j = \frac{f_j}{\sum_{k=1}^{m} f_{kj}} \quad \text{.................(8)}$$

**[Example 2: Customer's request]**
Table 4 demonstrates a customer's request in Chinese. The keywords and the corresponding weights of the request are depicted in Table 5.

Table 4: Illustrative example of a customer request

| Question |
|---|
| 客服人員您好。我先前訂閱了一份星座分析的電子報,但是現在已經不需要了。但是我在網頁上沒有找到取消訂閱或試閱的連結,請您告知連結位置或是步驟流程。謝謝! |

Table 5: Keywords and the corresponding weights of the illustrative example

| 關鍵詞 | 出現次數 | $W_j$ |
|---|---|---|
| 訂閱 | 1 | 2/3 |
| 電子報 | 2 | 1/2 |
| 取消訂閱 | 1 | 1/1 |
| 試閱 | 1 | 1/2 |
| 流程 | 1 | 1/1 |

From Table 5, CV of the request is represented as follows:

$$CV(Q) = \left\{ \left(k_1, \frac{2}{3}\right), \left(k_4, \frac{1}{2}\right), \left(k_5, \frac{1}{1}\right), \left(k_6, \frac{1}{2}\right), \left(k_8, \frac{1}{1}\right) \right\} \quad \text{.................(9)}$$

*4.2 Similarity Comparison Algorithm*
As mentioned previously, $CV(Q)$ and $CV(Q_i)$ are determined from the proposed question analysis algorithm. Base on the concept of vector space, an algorithm is used to compare the

similarity between customer's request and question part of each FAQ in the database. In our approach, Inner Product and Euclidean Distance methods are employed to compute the degree of similarity.

### 4.2.1 Inner Product Method

Inner Product method is a matrix multiplication method. The similarity between customer's request $Q$ and FAQ $i$ is given as

$$D_i = \sum_{k=1}^{n} (W_k * W_{ik}) \quad\quad (10)$$

For [Example 1] and [Example 2], we have

$$D_1 = \left(\frac{1}{3} * \frac{1}{3}\right) + \left(0 * \frac{1}{2}\right) + \left(0 * \frac{1}{1}\right) + \left(\frac{1}{2} * 0\right) + \left(\frac{1}{1} * 0\right) + \left(0 * \frac{1}{2}\right) + \left(\frac{1}{1} * 0\right) = \frac{1}{9}$$
$$D_2 = \left(\frac{1}{3} * \frac{1}{3}\right) + \left(\frac{1}{2} * \frac{1}{2}\right) + \left(\frac{1}{1} * \frac{1}{1}\right) + \left(\frac{1}{1} * \frac{1}{1}\right) = \frac{69}{28}$$
$$D_3 = \left(\frac{1}{3} * 0\right) + \left(\frac{1}{2} * 0\right) + \left(\frac{1}{1} * 0\right) + \left(\frac{1}{1} * 0\right) + \left(0 * \frac{1}{4}\right) + \left(0 * \frac{1}{2}\right) + \left(0 * \frac{1}{1}\right) = 0$$
$$D_4 = \left(\frac{1}{3} * 0\right) + \left(\frac{1}{2} * 0\right) + \left(\frac{1}{1} * 0\right) + \left(\frac{1}{1} * 0\right) + \left(0 * \frac{1}{2}\right) + \left(0 * \frac{1}{1}\right) = 0$$
$$\quad\quad\quad (11)$$

In Eq. (11), $D_2$ is the maximum value, and hence the answer part of $Q_2$ will be reply to the customer.

### 4.2.2 Euclidean Distance Method

By applying Euclidean Distance method, the similarity between customer's request $Q$ and FAQ $i$ is given as

$$D_i = \sqrt{\sum_{k=1,n} (W_k - W_{ik})^2} \quad\quad (12)$$

For [Example 1] and [Example 2], Eq. (12) gives

$$D_1 = \sqrt{\left(\frac{1}{3} - \frac{1}{3}\right)^2 + \left(0 - \frac{1}{2}\right)^2 + \left(0 - \frac{1}{1}\right)^2 + \left(\frac{1}{2} - 0\right)^2 + \left(\frac{1}{1} - 0\right)^2 + \left(0 - \frac{1}{2}\right)^2 + \left(\frac{1}{1} - 0\right)^2} = \sqrt{\frac{15}{4}}$$
$$D_2 = \sqrt{\left(\frac{1}{3} - \frac{1}{3}\right)^2 + \left(\frac{1}{2} - \frac{1}{2}\right)^2 + \left(\frac{1}{1} - \frac{1}{1}\right)^2 + \left(\frac{1}{1} - \frac{1}{1}\right)^2} = 0$$
$$D_3 = \sqrt{\left(\frac{1}{3} - 0\right)^2 + \left(\frac{1}{2} - 0\right)^2 + \left(\frac{1}{1} - 0\right)^2 + \left(\frac{1}{1} - 0\right)^2 + \left(0 - \frac{1}{4}\right)^2 + \left(0 - \frac{1}{2}\right)^2 + \left(0 - \frac{1}{1}\right)^2} = \sqrt{\frac{529}{144}}$$
$$D_4 = \sqrt{\left(\frac{1}{3} - 0\right)^2 + \left(\frac{1}{2} - 0\right)^2 + \left(\frac{1}{1} - 0\right)^2 + \left(\frac{1}{1} - 0\right)^2 + \left(0 - \frac{1}{2}\right)^2 + \left(0 - \frac{1}{1}\right)^2} = \frac{130}{36}$$
$$\quad\quad\quad (13)$$

In Eq. (13), $D_2$ is the maximum value, which implies that the customer's question $Q$ is most similar to the question $Q_2$. This result is the same as that of the Inner Product method.

### 4.3 Service Personnel Distribution Algorithm

By assuming that each service personnel is responsible for a particular group of questions. In this system, $CV(P_i)$ is assigned for each personnel. This gives

$$CV(P_i) = \{(K_1, W_{i1}), (K_2, W_{i2})..(K_j, W_{ij})..(K_n, W_{in})\} \quad\quad (14)$$

where Kj represents the jth keyword in one particular group; Wij is the relevance of Kj to service personnel i.

Personnel arrangement algorithm is similar to the similarity comparison algorithm given above. Based on the same computation procedure, each unsolved request can be assigned to the most proper service personnel.

## 5. Implementation and Evaluation

ACSS was implemented on a PC server with Intel Pentium4 1.6G CPU and 512MB RAM. The development environment consisted of Microsoft Windows 2000 operating system, Apache 1.3 web server, MySQL 3.23 database management system and J2SDK 1.4 Java development kit. Figures 2 and 3 show ACSS user interfaces for maintaining the keyword database and the FAQ database.



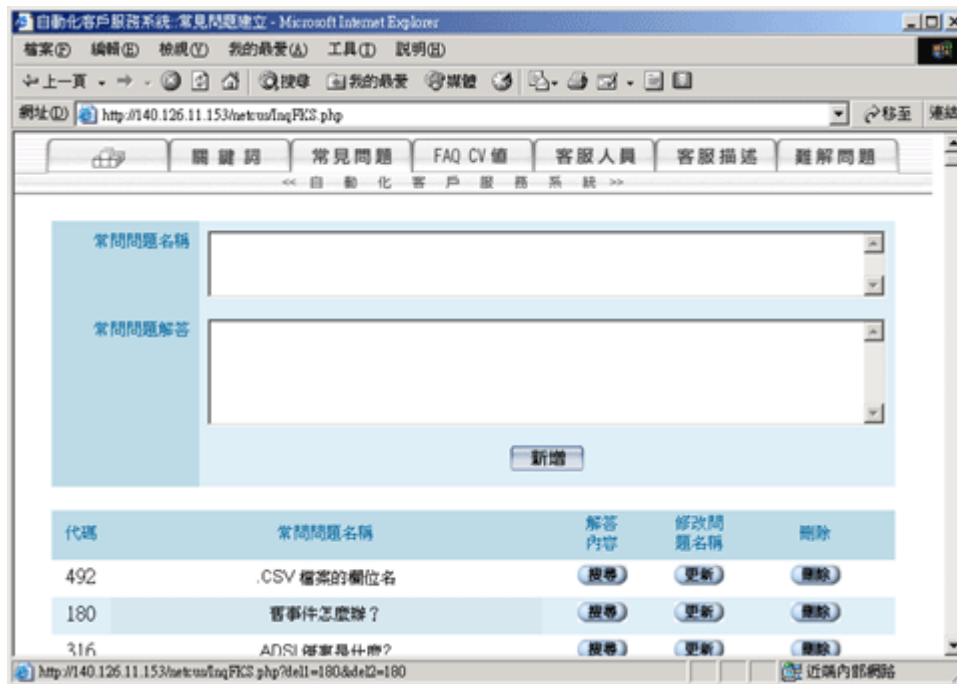Figure 2. ACSS user interfaces for maintaining the keyword database

Figure 3. ACSS user interfaces for maintaining the FAQ database

Several experiments have been conducted to evaluate the performance of ACSS by comparing its results with those of other approaches. Two FAQ databases with different languages were adopted in the experiments: the Yahoo!-Kimo database contained 578 FAQ's and 953 key words in Chinese, while the MiTAC Internation Corp database contained 808 keywords and 569 FAQ's in English.

Two versions of data were generated from each FAQ database: the classified data based on six interrogatives, that is, "who", "what", "where", "when", "why" and "how", and the original data without being classified by the six interrogatives. Since there are several Chinese phrases that can be used to represent each of the six interrogatives, an English-to-Chinese mapping table (see Table 6) is used to deal with Chinese requests and FAQ's.

Table 6: English-to-Chinese mapping table

| English word | Chinese phrase |
|---|---|
| What | 什麼, 那些, 哪些, 何為, 甚麼, 何謂, 那一些, 哪一些, 幾個 |
| Where | 哪裡, 那裡, 何處 |
| Who | 誰, 哪一位, 那一位, 那位, 哪位 |
| Why | 為什麼, 為何, 為甚麼 |
| When | 何時, 什麼時候, 多久, 什麼時間 |
| How | 如何, 怎樣, 有何, 怎麼樣, 怎麼辦, 怎麼, 怎麼會 |

In addition, three types of requests (as shown in Table 7) are used to evaluate the performance of different approaches, including the requests that are similar to the ones being asked before, the requests that are quite different from any request being asked, and the mixed case of the

two types of requests.

Table 7: Experiments with different types of requests

| Number | Content of the Experiment |
|---|---|
| Case 1 | All of the customer requests are similar to the question part of some FAQ's (i.e., such questions have been asked before). |
| Case 2 | All of the customer requests are quite different to the question part of any FAQ (i.e., such questions have not been asked before); however, the answers of the requests exist in the FAQ database. |
| Case 3 | Customer requests are randomly selected from Case1 and Case2. |

### 5.1 Experiments to Find the Best-Fit Answer

Tables 8 and 9 present the experimental results by respectively applying ACSS with Euclidean distance method ($ACCS_{ED}$), ACSS with Inner Product method ($ACCS_{IP}$) and INCSS to reply 550 requests from customers. It can be seen that the correct reply ratio of ACCS was higher than INCSS no matter which similarity comparison methods were adopted. The experimental results also reveal that ACCS is able to handle those questions that have not been asked before.

Table 8: Experimental results with Yahoo!-Kimo FAQ database

| | Original data | | | Data classified by six interrogatives | | |
|---|---|---|---|---|---|---|
| | $ACSS_{ED}$ | $ACSS_{IP}$ | INCSS | $ACSS_{ED}$ | $ACSS_{IP}$ | INCSS |
| Case 1 | 97.2% | 94.4% | 77.7% | 95.1% | 94.4% | 84.7% |
| Case 2 | 81.1% | 80.8% | 62.5% | 78.8% | 81.6% | 59.6% |
| Case 3 | 88.8% | 88.1% | 70.5% | 88.2% | 87.4% | 70.3% |
| Average | 89.3% | 87.7% | 70.2% | 87.3% | 87.8% | 71.5% |

Table 9: Experimental results with Mitac FAQ database

| | Original data | | | Data classified by six interrogatives | | |
|---|---|---|---|---|---|---|
| | $ACSS_{ED}$ | $ACSS_{IP}$ | INCSS | $ACSS_{ED}$ | $ACSS_{IP}$ | INCSS |
| Case 1 | 98.1% | 92.9% | 73.2% | 94.7% | 92.1% | 80.4% |
| Case 2 | 92.9% | 90.5% | 66.3% | 84.4% | 82.9% | 66.0% |
| Case 3 | 95.7% | 91.8% | 69.3% | 89.0% | 86.7% | 72.8% |
| Average | 95.5% | 91.7% | 69.6% | 89.3% | 83.9% | 73.0% |

### 5.2 Experiments for Finding Five Best-Fit Answers

Tables 10 and 11 illustrate the comparative results between ACSS and INCSS (in getting the "Top 5" possible answers). The customer service personnel can choose the answer among the possible answers to reply directly. From Tables 10 and 11, it is observed that the correctness

of $ACCS_{VC-ED}$ or $ACCS_{VC-IP}$ is higher than INCSS about 2%~5%. In addition, the comparative results show the correctness for Random data and Classified Data in each system are close. In average, $ACCS_{VC-IP}$ provides the best performance in getting top-five possible answers among these three systems.

Table 10: Experiments for finding five best-fit answers from Yahoo!KIMO FAQ database

|  | Original data | | | Data classified by six interrogatives | | |
|---|---|---|---|---|---|---|
|  | $ACSS_{ED}$ | $ACSS_{IP}$ | INCSS | $ACSS_{ED}$ | $ACSS_{IP}$ | INCSS |
| Case 1 | 99.3% | 98.5% | 96.3% | 98.5% | 98.5% | 96.8% |
| Case 2 | 85.2% | 95.7% | 81.0% | 81.3% | 91.0% | 75.9% |
| Case 3 | 92.4% | 97.1% | 88.3% | 90.3% | 94.9% | 84.5% |
| Average | 92.3% | 96.1% | 88.5% | 90.0% | 94.8% | 85.7% |

Table 11: Experiments for finding five best-fit answers from Mitac FAQ database

|  | Original data | | | Data classified by six interrogatives | | |
|---|---|---|---|---|---|---|
|  | $ACSS_{ED}$ | $ACSS_{IP}$ | INCSS | $ACSS_{ED}$ | $ACSS_{IP}$ | INCSS |
| Case 1 | 99.0% | 98.1% | 93.5% | 95.8% | 96.3% | 94.6% |
| Case 2 | 95.8% | 96.6% | 90.2% | 87.9% | 86.1% | 85.0% |
| Case 3 | 97.3% | 97.0% | 91.3% | 91.1% | 91.6% | 89.1% |
| Average | 97.3% | 97.2% | 91.6% | 91.6% | 91.3% | 89.5% |

### 5.3 Limitations of ACCS
From the experimental results, some factors that might influence the finding of correct FAQ's are observed and listed as follows:

(1) Keyword in different locations of a sentence that leads to inverse meaning of a request. For example, the sentences "How to copy contact information from Yahoo! Address Book to Microsoft Outlook Express?" and "How to copy contact information from Microsoft Outlook Express to Yahoo! Address Book?" contains the same keywords "Address Book", "Yahoo" and "Microsoft Outlook Express"; however, the meanings of these two questions are opposite. Therefore, if the customer is not satisfied with the answer, the system will try another answer with the same weight.
(2) Customer requests are not clear or contain too few keywords. For example: The system might have difficulty to find a good answer for the question "What is My Yahoo!?" since only one keyword "Yahoo" does not provide sufficient information.
(3) Customer requests contain too many keywords that are not relevant to the right answer. For example: the user want to know "How to find the information about on-line game?"; however, the question is "I want to play Yahoo! Game, the on-line help show me to the system announcement of Yahoo! Game. But I cannot find the hyperlink. Please tell me how can I find the system announcement?" Those words with underline are the frequently appeared keywords that might confuse the system in finding the right answer

## 6. Conclusions

This study proposed an Automatic Customer Service System (ACCS) that can automatically reply customer requests by selecting the mostly feasible answers from the FAQ database. Moreover, it can learn from customers' questions to increase the accuracy of replying answers. From several experimental results, it can be seen that ACCS is able to provide high efficiency services and relieve the work burdens of service personals, and hence the 24-hour service can be offered and each customer's waiting time can be significantly reduced, which will not only decrease the service cost, but also increase the competition advantages of enterprises.

The idea of this research can also be applied to other applications, such as the development of an intelligent BBS system that can find past relevant discussions for a new proposed question to avoid duplicated discussion contents, and the development of an intelligent tutoring assistance system to solve the students' problems in learning and to ease the burden of the teachers.

## Acknowledgement

## References

[1] http://news.sina.com.tw/sinaNews/rtn/twFinance/2001/0211/2752555.html

[2] Arppe Antti, "Term Extraction from Unrestricted Text", http://www.lingsoft.fi/doc/ nptool/term-extraction.html, 1995.

[3] Chang, Y.M., and Hwang, G.J. "Development of an Adaptive Online Customer Service System," *National Computer Symposium*, Taiwan, 2001.

[4] Cohen, W. "Learning Rules That Classify E-mail," *AAAI Spring Symposium on Machine Learning in Information Access*, 1996.

[5] Cooper, E. "Improving FAQfinder's Performance: Setting Parameters by Genetic Programming," *AAAI Spring Symposium on MLIA Technical Papers*, 1996.

[6] Hoch, R. "Using IR techniques for text classification in document analysis", *Proceedings of SIGIR-94, 17th ACM International Conference on Research and Development in Information Retrieval*, 1994, pp. 31-40.

[7] Jones, L. P., Gassie, E. W., & Radhakrishnan, S. "INDEX: The Statistical Basis for an Automatic Conceptual Phrase-indexing System," *Journal of American Society for Information Science*, 41(2), 1990, pp. 87-98.

[8] Krulwich, B. "Learning Document Category Descriptions through the Extraction of Semantically Significant Phrase" Workshop on Data Engineering for Inductive Learning, IJCAI-1995, Montreal, Canada, Aug. 20 1995.

[9] Li, M.Y., and Tseng, Judy C.R. "Development of an Intelligent Network-based Customer Service System," *2001 Taiwan Academic Network Conference (TANET2001)*, October 24-26, 2001.

[10] Peppers, D., and Rogers, M. "Building Relationships One Customer at a Time", *The One to One Future*, New York, 1993.

[11] Reichheld, F.F., and Sasser, W.E. "Zero Defections: Quality Comes to Service," *Harvard Business Review*, 1990, pp. 105-111.

[12] Salton, G., and Buckley, C. "Term Weighting Approaches in Automatic Information Retrieval," *Journal of Information Proceeding and Management*, Vol. 24, No. 3, 1998, pp. 513-524.