# Search System and Search Results Clustering for Researches and Business Computer Projects of Ratchaphruek College

Siriporn **U**mmepean
Ratchaphruek College, Faculty of Business Administration, Business Computer
9 Moo 1, Nakhon-in Road, Bangkhanoon
Bang Kruai, Nonthaburi 11130
E-mail: joknu555@hotmail.com

*Abstract* ---**The objectives of this research were to create a search system and cluster search results for researches and business computer projects of Ratchaphruek College so as to facilitate the searching research data and documents and help users have faster access to required researches. Besides, the users could view another interesting facet data related to required words for searching again.**
**The search system and search results clustering for researches and business computer projects of Ratchaphruek College were developed on operating system, Windows XP; programming language, PHP; and Library. To cluster search results, Carrot2 was applied to help develop programming. It was found that the created search system for researches and business computer projects of Ratchaphruek College comprising 3 parts of evaluation made by 3 experts: part 1, process of system with an average of 3.94; part 2,qualityof service with an average of4.06 ; and part 3, facilities with an average of4.11 .For an analysis of satisfaction of users, the results were categorized into3 parts including: part 1, process of system with an average of 4.09; part 2, quality of service with an average of 4.17; and part 3, facilities with an average of 4.12.The system facilitated faster searching of researches and business computer projects.**

***Keywords-component; Search system: Research***

## I. INTRODUCTION

As Ratchaphruek College is a new college where has been operating approximately 5 years and has goal to make transition from college to university by academic year 2013, a research conducting by personnel in the college is encouraged. In addition, there is large number of projects about business computer of students studying in Faculty of Business Administration, Business Computer Program every academic year.

After a program of Master of Business Administration is managed, its students must conduct thesis. Hence, Ratchaphruek College can obtain increasing researches every year. Significantly, the detail of all researches with a large number is difficult to remember and record in form of documents only; it causes delay in those methods and easily takes a mistake when the researches are required to use. Besides, the search results of the researches needed from the search system is demonstrated in huge number consisting of Title, Abstract and URL, respectively. However, the search results sometimes do not meet the users' required result at first, a repeated searching from that search results is needed or a research required is shown among enormous search results and is not demonstrated on the first page, so the users will waste time searching the required one after the search results are shown.

Apropos of mentioned problems, a concept applying computer technology for helping in development of search system and clustering search results for researches and business computer projects of Ratchaphruek College is presented in order to store and arrange academic data and researches into categories. It will facilitate searching, discovering and disseminating. To cluster the search results, the title and abstract are used as main data. Each cluster is called Label which is shown in Thai phrase, so the users can understand it. An algorithm for clustering and finding Thai phrase is called Suffix Tree Clustering: STC which is mostly utilized in clustering the search results and has flexibility and high speed in clustering.

## II. LITERATURE REVIEW

The main objectives are to develop a search system and cluster search results for researches and business computer projects of Ratchaphruek College so as to facilitate the searching research data and documents and help users have faster access to required researches. Thus, theory, technology and related researches are studied as following.

### *A. Web Service Technology*

Web Service is open standard which the users can publish what services provided and another web application can search or look up the services as a component of software which can work specially. It includes interoperability and operation of data and process. It is able to interact with other systems using technical standard to exchange data, utilize program from

another system which has difference in both hardware and software. The standard is used as mutual agreement is as below.

-Coordination of operation of data through XML

-Coordination of operation of process using standard or SOAP )Simple Object Access Protocol)

-Publishing method and provided services searching: UDDI )Universal Description, Discovery and Integration)

-Web Service Description Language is virtually as manual for the system using in learning instruction for utilizing required Web Services. WSDL is developed based on XML standard.

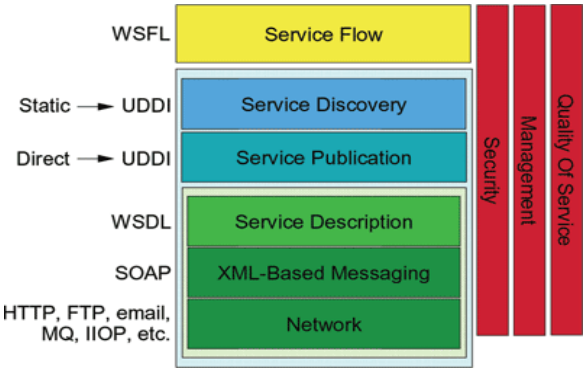


Figure 1. shows the process of web service.

### B. *Search Result Clusteing*

Search Result Clustering is a technique in query expansion by selecting some statement from search results to cluster.

### C. *Suffix Tree Clustering*

Suffix Tree Clustering: STC is a linear time (O(n)) clustering algorithm based on a suffix tree which efficiently identifies sets of documents sharing common phrases or the suffix of a phrase into one cluster and then clusters according to these phrases are created. The original STC algorithm, however, cannot provide an effective evaluation method to assess the quality of clusters. STC algorithm has three logical steps: ((1 document cleaning, (2)identifying base clusters, and ((3 combining base clusters.

### D. *Related Research*

Apropos of other researches, there are many researches developing information technology system as follows.

According to research of Winai,[1] it was the development of information database management system in central region volunteer's club via the Internet by applying computer to facilitate its management and linking data in the Internet network. Hence, it was search system development and its members' data improvement in order to decrease repeated data and also provide faster searching for members' data; the database could be updated with accuracy. However, this developed system stored data in form of database so as to solve problem occurred to the old data directory and was be able to meet the increasing needs of units along with providing convenience for users in efficiently utilizing data.

According to research of Krisada, [2] it was the development of research searching system on local area network of graduate school of Chandrakasem Rajabhat University created by the researcher containing quality of searching via the Internet network with average of 3.79 and generic qualification of search system with average of 3.99; those were higher than the criteria set. In addition, it could facilitate data searching for the users in efficient way.

Yang et al. [3] implemented and proposed a novel algorithm called sentence-based suffix tree clustering algorithm (SSTC) for web documents. The algorithm was implemented based on the Carrot2 by extending it with structure weights of nodes.

## III. METHODOLOGY

### A. *Data Collection*

Data collection of old system operation was made for analyzing the system to cover all contents needed by inquiring people who collected data in each unit.

### B. *System Framework Design*

The design of the search system and search results clustering for researches and business computer projects of Ratchaphruek College was presented comprising 2parts: (1a part of searching for researches and business computer projects and 2) a part of clustering the search results.

- A Part of Searching for Researches and Business Computer Projects .
  It was the part of searching for researches and business computer projects from the database which was be able to search for the title, author, and abstract.
- A Part of Clustering the Search Results
  It was the part of clustering the search results by leading the title and abstract searched to a method of text processing which were word segmentation, stop word removal, and numeric and special character removal. After text processing, those were clustered and created in Thai phrase applying algorithm based on suffix

tree from Library, Carrot2. Then the defined label was returned to the system.

For the process of the system, the users were firstly enter the required words and then the system searched data for researches and projects stored in the database by using those entered. The search results were presented including title and abstract. Then clustering the search results was conducted by using the first 100results of titles and abstracts to cluster in order to create defined label. Lastly, the completed results were shown for the users.

## IV. RESULTS OF RESEARCH

### *A. Results of System Development*

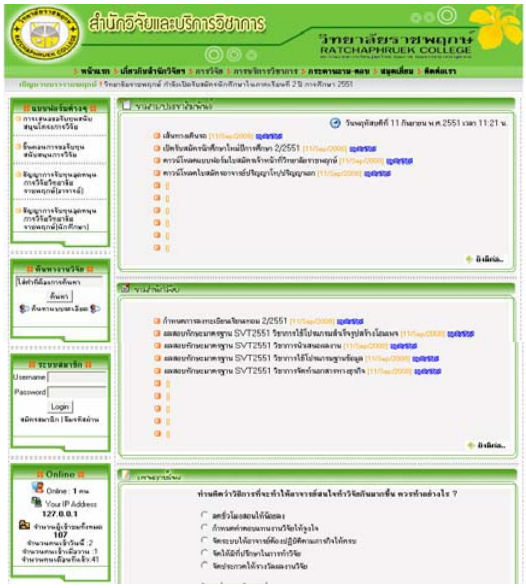


Figure 2. shows a screen of home page of the system.

Figure 2 demonstrated the structure of home page which the users logging in for utilizing the service. The data were shown relating to the topics which were home page, research office, research, academic service, web board, guest book, contact us, forms, search for research, and news and public relations notified according to its topic.

- A screen for searching researches was for research searching by adding required words or clicking on Advance Search as shown in Figure 3.

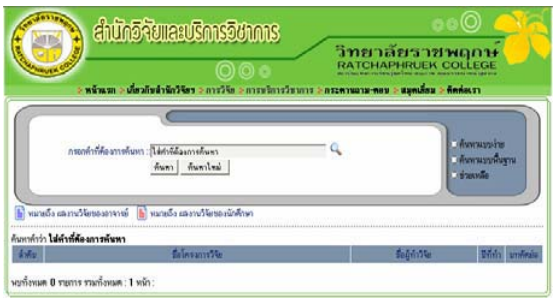


Figure 3 demonstrates a screen for searching researches and business computer projects

- Structure of home page for a system administrator was shown in Figure 4which is for entering username and password to login for system administration. This page contained lists for administrator to manage the system including research, research type, faculty, program, researcher, grant, advisor, member, research data management, news/public relations, research office, poll, guestbook, and log out.


Figure 4. shows screen for system administrator.

- Structure of news/public relations was shown in Figure6 which demonstrated a screen of news/public relations for members. The notified information would be appeared on the home page of members. Items listed in this page were for adding, editing, searching and publishing news/ public relations.


Figure 5. Structure of News/Public Relations

*B. Efficiency Analysis of Search System for Researches and Business Computer Projects.*

TABLE I . EFFICIENCY ANALYSIS MADE BY EXPERTS

| Evaluation Items | Experts | | | Average |
|---|---|---|---|---|
| | 1 | 2 | 3 | |
| Process of System | 3.83 | 4.17 | 3.83 | 3.94 |
| 1. Convenience of system use | 4 | 4 | 3 | 3.67 |
| 2. High speed in operation | 4 | 4 | 4 | 4.00 |
| 3. Easily learn use of system | 4 | 5 | 4 | 4.33 |
| 4. Easy to view and beautiful web page alignment | 4 | 4 | 4 | 4.00 |
| 5. Suitability of linking other web pages | 4 | 5 | 4 | 4.33 |
| 6. Suitability of operation of holistic system | 4 | 4 | 4 | 4.00 |
| Quality of Service | 4.17 | 4.00 | 4.00 | 4.06 |
| 1. Accuracy of searching | 5 | 4 | 4 | 4.33 |
| 2. Security system for system using | 3 | 3 | 3 | 3.00 |
| 3. Accuracy can be monitor at all times. | 4 | 5 | 4 | 4.33 |
| 4. Accuracy in recording, editing, or cancel | 4 | 4 | 4 | 4.00 |
| 5. Errors clearly shown and easy for editing | 4 | 4 | 4 | 4.00 |
| 6. Clear sorting of search | 5 | 4 | 5 | 4.67 |
| Facilities | 4.33 | 4.00 | 4.00 | 4.11 |
| 1. Clear and suitable graphic design | 4 | 4 | 4 | 4.00 |
| 2. Clear letters on web page both in size and color | 5 | 4 | 4 | 4.33 |
| 3. Suitable method and help | 5 | 5 | 4 | 4.67 |
| 4. Contact address of administrator is clearly shown. | 4 | 4 | 4 | 4.00 |
| 5. Manual can be learnt and understood. | 4 | 4 | 4 | 4.00 |
| 6. Error protection from users is created. | 4 | 3 | 4 | 3.67 |

As shown in table I, it was found that the efficiency analysis of search system for researches and business computer projects made by 3 experts with the results comprising 3 parts: part 1, process of system with an average of 3.94; part 2, quality of service with an average of 4.06; and part 3, facilities with an average of 4.11 Those 3 experts had recommendations about the system as following.

- In use of system, the system should provide security system or be utilized only in a network in the specific college.
- A number of logged in users should be calculated.

TABLE II .EFFICIENCY ANALYSIS MADE BY SAMPLING GROUP

| Evaluation Items | Average |
|---|---|
| Process of System | 4.09 |
| 1. Convenience of system use | 4.07 |
| 2. High speed in operation | 4.13 |
| 3. Easily learn use of system | 4.23 |
| 4. Easy to view and beautiful web page alignment | 4.00 |
| 5. Suitability of linking other web pages | 4.13 |
| 6. Suitability of operation of holistic system | 4.00 |
| Quality of Service | 4.17 |
| 1. Accuracy of searching | 4.33 |
| 2. Security system for system using | 3.93 |
| 3. Accuracy can be monitor at all times. | 4.33 |
| 4. Accuracy in recording, editing, or cancel | 4.00 |
| 5.Errors clearly shown and easy for editing | 4.00 |
| 6. Clear sorting of search | 4.43 |
| Facilities | 4.12 |
| 1. Clear and suitable graphic design | 4.00 |
| 2. Clear letters on webpage both in size and color | 4.33 |
| 3. Suitable method and help | 4.33 |
| 4. Contact address of administrator is clearly shown. | 4.13 |
| 5. Manual can be learnt and understood. | 4.00 |
| 6. Error protection from users is created. | 3.93 |

From table II, it was stated that the efficiency analysis of search system for researches and business computer projects made by sampling group of 30 people. The results were categorized into3 parts including : part 1, process of system with an average of 4.09; part 2, quality of service with an average of 4.17; and part 3, facilities with an average of 4.12. The sampling group had recommendations about the system as following.

- Photo of researcher should be placed on web page to honor their biography.
- The tactics in the system should be improved in order to increase users' attention.

## V. CONCLUSION

According to this research, the search system for researches and business computer projects of Ratchaphruek College was studied, developed, and tested. The system could store content of researches and business computer projects in electronic from n efficient way. From evaluation made by experts, an average was

4.04which obtained high level of suitability. The created search system for researches and business computer projects possessed optimum efficiency and could be usable. However, it should be developed to have more space for storing data. If Ratchaphruek College had graduate school, this system with more space would help store data of thesis and mini-thesis.

For recommendation on development of the system, it is recommended that new techniques in data searching should be used in order to search data from content of research and search data with high speed and in more efficient way.

## REFERENCES

[1] Winai Purikasem. The development of information Database management system in Central Region Volunteer's Club Via the Internet. Master Thesis, Master of Science Program in Information Technology, Faculty of Information Technology.2007.

[2] Krisada Ploysri. A Development of Research Searching System on Local Area Network of Graduate School Chandrakasem Rajabhat University. Office of Academic Resources and Information Technology, Chandrakasem Rajabhat University, 2006.

[3] Sanchai Yongkulwanich. Search Engine for Scholars. Office of Academic Resources and Information Technology, Ubon Ratchathani Rajabhat University.1999.

[4] Srijit Rattanakaewkarn. Database Systems. Bangkok: Ramkhamhaeng University, 2000.

[5] SupattraAuewong. Research for Learning Development. [Online].Retrieved July 5, 1998, from www.vattaka.com/reslearn.htm