



## Cascading Guided Search Cloud Service Search Engine

Ms Thamilvaani Arvaree@ Alvar<sup>1</sup>, Assoc. Prof Dr. Rodziah Atan<sup>2</sup>

<sup>1</sup>Assistant Professor, Foundation in Science, Faculty of Science,  
University of Nottingham, Malaysia  
[thamilvaani@yahoo.com](mailto:thamilvaani@yahoo.com)

<sup>2</sup>Associate Professor, Department of Information System,  
Faculty of Computer Science and Informataion Technology,  
Univeristy Putra Malaysia.  
[rodziah@fsktm.upm.edu.my](mailto:rodziah@fsktm.upm.edu.my)

### ABSTRACT

Cloud computing is one contemporary technology in which the research community has recently embarked. This paradigm shifts the location of the infrastructure to the network to reduce the costs associated with the management of hardware and software resources. Developers with innovative ideas for new Internet services no longer require the large capital outlays in hardware to deploy their service or the human expense to operate it. Organization adopts cloud computing services through service provider via Internet. In recent years, numbers of cloud service providers are increase. However, there is no study that focuses on search engine and web portal for cloud computing for users who want to find cloud service. At the same time, "vendor lock in" issues and the lack of common cloud standards delayed the interoperability across these providers. Thus, lead the cloud customer to face challenges and problems in selecting the right service provider who meets their needs. This research focus on meeting the user requeriment in cloud environment whereby, user requirements reflect to search query that entered by the end users and how these search queries are exactly matched with accurate cloud service. Therefore, end of the research cloud service search engine was developed as a proposed tool for meeting the user requirements.

### Indexing terms/Keywords

Cloud Computing, Search Query, Cloud Service, Matching, Search Engine.

---

# Council for Innovative Research

Peer Review Research Publishing System

**Journal:** INTERNATIONAL JOURNAL OF COMPUTERS & TECHNOLOGY

Vol 10, No 9

[editor@cirworld.com](mailto:editor@cirworld.com)

[www.cirworld.com](http://www.cirworld.com), [member.cirworld.com](http://member.cirworld.com)



## INTRODUCTION

Cloud computing is an emerging technology, which has attracted a lot of attention in recent years. While cloud computing is currently a term without a single consensus meaning in the marketplace, it describes a broad movement toward the use of wide area networks, such as the Internet, to enable interaction between IT service providers of many types and consumers. Service providers are expanding their available offerings to include the entire traditional IT stack, from hardware and platforms to application components, software services, and whole applications.

There are many definitions of Cloud Computing, but they all seem to focus on just certain aspects of the technology. After review all the definitions by the experts Luis M. Vaquero, 2009 proposed cloud definitions as “Clouds are a large pool of easily usable and accessible virtualized resources (such as hardware, development platforms and/or services)”. These resources can be dynamically reconfigured to adjust to a variable load (scale), allowing also for an optimum resource utilization. This pool of resources is typically exploited by a pay-per-use model in which guarantees are offered by the Infrastructure Provider by means of customized SLAs. From the summarized definition, seeking the accurate user requirements in the pool of resources is very essential and always an issues among the end users.

## RELATED WORKS

In cloud environment, providers can potentially publish various clouds or cloud services through the Internet, and consumers can access these services provided by cloud application layer through web-portals (Jaeyong Kang, 2010). Even though there are many existing generic search engines that consumers can use for finding cloud services, these engines may return URLs containing not relevant web-pages to meet the original service requirements of consumers. Probability of retrieving the accurate search result is only 20% or some time even 0% at the first page of search results. Intuitively, visiting all the web-page can be time-consuming. Generic search engines (e.g., Google, MSN, etc) are very effective tools for searching URLs for generic user queries. However, they are not designed to reason about the relations among the different types of cloud services and determining which service(s) would be the best or most appropriate service for meeting consumer’s service requirements (Taekgyeong Han, 2010).

Searching cloud services over in Internet from service provider are becoming difficult for end users. This is because, although the variety of services and resources are offered in the clouds, there is no portal site and search mechanism that are specialized for cloud computing (Jaeyong Kang and Kwang Mong Sim, 2011). Users need to perform the searching based on their own knowledge by key in the keyword they think is relevant manually. Philip Cy Sheu, 2009 in his statement mentioned that there are no standards, open protocols and search mechanisms for discovering different kinds of clouds and consumers have to perform the searching by entering the keyword manually

Search engines service millions of queries per day, and search collections that contain billions of documents. As the growth in the number of documents that are available in such collections continues, the task of finding documents that are relevant to user queries becomes increasingly costly (Al-Rawi 2010).

The common searching techniques are either by keyword search or image search. Besides, guided search and semantic searching also considered as a common searching techniques as well.

Guided search refer to input the query stage which assist the user to find the suitable and corrected words. The advantage of guided concept is to reduce the probability of inserting wrong words in query. We can see example of this in “Google Suggest”, which analysis web history to display the relevant words. (Abhishek, 2009),(Abhishek,2011). According to the Google blog, suggest feature will help user in following ways:

1. To reduce spelling mistakes.
2. To save time and keystrokes.
3. To help in formulating user search i.e. automatically make search queries for user

Suggestions are very helpful for the novice users because of several reasons like may reduce spelling mistakes and minimize effort in typing long sentences when input the query to search box.

Guided search (conversational interaction will prompt the users for information and guide them towards a solution in the most efficient manner. The question order is dynamically generated based on the content in the case base and the answers the user has provided so far. Users are able to skip questions ( e.g irrelevant questions) and the system will dynamically re compute the next most important question to arrive at a solution as efficiency as possible. Users do not perceive much difference between the clarification questions and guided questions. (Stefania Montani 2010)

In a guided search engine the user does not have the classic text box in which he can write the question. The user interface is composed of various list box menus, automatically generated. The web page that contains the user interface is dynamic, and changes its state as the user generates the query. The idea behind this method consists of helping the user in the formulation of the queries. The program will support the user in the construction of all the queries, reducing the set of all the possible answers to the known one. (Mauro Gaspar, 2003)



With so much of advantages of guided search, this research focus on searching the user requirements with guided environment rather than in open environment which may ignore all the those benefits. Moreover, the end users prefer to be use a simple and basic interface. They do prefer the secure answer which may provide in guided environment.

Cascading refers to series of steps or stages which, once initiated, continues to the final step by virtue of each step being triggered by the preceding one, sometimes with cumulative effect. In this research we adopt this technique in developing the guided search environment in the proposed tool so that the user can easily answer from the easier question to complex questions in order to complete the searching. Asking a well-formulated question is the fastest and the most natural way to express the search goal.

However, unlike the full-sized documents, questions are much shorter and hence require less time and effort from the users for reading and relevance judgment. In addition to questions, users may also be presented with short answers to them, when they point to a particular question. Users can also click on the question and be redirected to the document, containing the answer, for further information. In this sense, questions can be considered as shortcuts to specific answers (Alexander Kotov, 2010). In this research, the cascading method is used to ease the end user of the proposed solution.

## METHODOLOGY

Pre experiments are setup and analysis is drawn. There are two phases of experiment was caaried out. First phase consists of two sub experiments. First is to indentify the user experience and second is to identify the prominent problem faced by user in getting services from the cloud. The second phase of experiement also consists of two sub experiments. First experiment was aimed to identify the common keywords and categories that users are accessing and expect to get results and the second experiement was aimed to identify number of layers that the users navigate to get the services. The results of experiments are treated as an input for the development stage of the research.

Keyword that gathered during the experiement help us to formulate the further steps in this reseach. Keyword searching is identified as a best solution for service searching over in cloud. Therefore, appropriate keyword searching algorithm was rewived and studied. Strike a match similarity ranking algorithm was identified as the best alorithms to perform the keyword searching.

The strike match similarity ranking algorithm was motivated by the following requirements:

1. A true reflection of lexical similarity - strings with small differences should be recognized as being similar. In particular, a significant substring overlap should point to a high level of similarity between the strings.
2. Robustness to changes of word order - two strings which contain the same words, but in a different order, should be recognized as being similar. On the other hand, if one string is just a random anagram of the characters contained in the other, then it should (usually) be recognized as dissimilar.
3. Language Independence - the algorithm should work not only in English, but in many different languages.

Specialized search engine to identify the cloud services were developed as a proposed solution for this research. This section explains about the Cloud Service Search Engine (CSSE) system which consists of service discovery and web interface. Basically there three main actor involves in this system. Firstly, cloud providers. Cloud providers update their cloud services by being registered members into the database file of CSSE. They have the full authorized to update the details of the services and registered users profile. Secondly, cloud users. Cloud users of this system grouped into two categories. There are unregistered users and registered users. Both groups of users have different privileges and access to the system. Unregistered users only can access the website to get the details of available service providers for their request. Therefore, CSSE operates as follows. User send search query to the web interface. The system will identify the possible service providers who match with the search query entered by the users. If the users wish to get the services, they are required to register themselves and update their profile as a registered user. Once become registered users, they are free to access not only to identify possible available service provider whereas, able to retrieve the services from that particular service provider website. The service discovery carried out three main functionalities as follows:-

1. Query processor
2. Matching
3. Ranking

### Query Processor

When a user sends a query to CSSE, the query processor analyses the query and converts the query into an appropriate form which can be processed by the matching procedure.

### Matching Procedure

In the matching procedure, the similarity between search query entered by users and cloud services from cloud provider is determined by referring to the matching procedure. Equivalent method is used to identify the similarity of two strings. Equivalence methods compare two strings and return a value of true or false according to whether the method deems those two strings to be, in some sense, equivalent.

**Ranking**

Similarity ranking methods compare a given string to a set of strings and rank those strings in order of similarity. To produce a ranking, we need a way of saying that one match is better than another. This is done by returning a numeric measure of similarity as the result of each comparison. Figure 1 below illustrates the complete process in CSSE system.

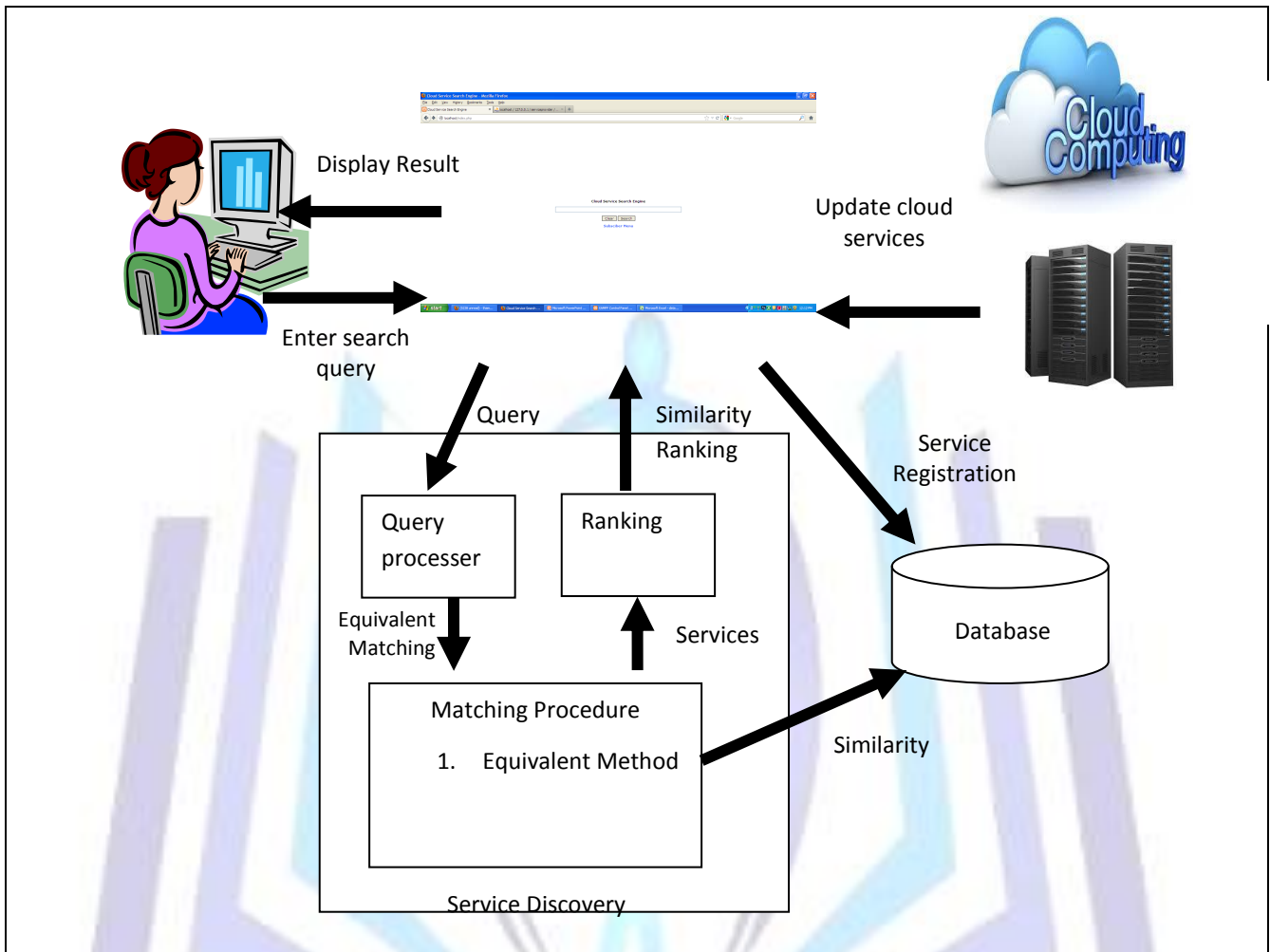
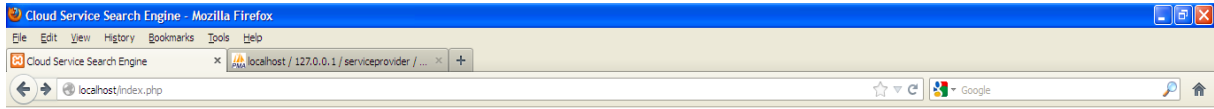


Figure1: CSSE System

**CLOUD SERVICE SEARCH ENGINE (CSSE) INTERFACE**

Cloud Service Search Engine (CSSE) is easily can be access by user via Internet connection. CSSE has web based interface. This, enable the user access the page at anytime and anywhere with Internet connection. The CSSE interface is very simple. Figure 2 and 3 shows the main pages of CSSE. Users need to enter the service name that they expected to retrieve as a keyword. This keyword treated as a search query by the CSSE.

CSSE able to match the user query with relevant search query and services. Besides, its able to shows the percentage matches of the search query with the services found for user convenience. The similarity rate and response time are two measurements that the CSSE meet. CSSE able to retrieve the services fast and accurate for the user needs. By this way, the user meet their required services accurately by knowing the similarity of the services.



**Cloud Service Search Engine**

[Subscriber Menu](#)

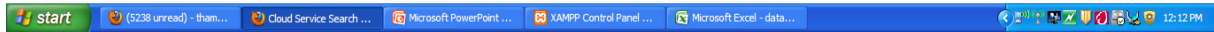
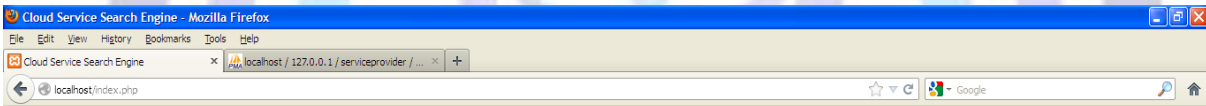


Figure2: CSSE Main Screen



**Cloud Service Search Engine - Result**

Search query: **engineering**

No.	URL	Percentage
1.	wikipedia.org/wiki/Computer_science#Software_engineering	74.07%
2.	http://jobsearch.monster.com	68.97%
3.	http://jobsearch.monster.com	68.97%

Query returned in **0.004074** second(s)

[Subscriber Menu](#)

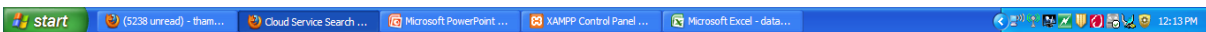


Figure1: CSSE result screen

## CONCLUSION

In conclusion, In conclusion, the technology of search engines is a very dynamic field, always looking for improvements and new ideas in order to satisfy user needs. The ability of the system to find relevant information based on the user's search query to a successful system. This ability can be significantly enhanced by employing an approximate string matching algorithm. This research collect set of search query from the experiment carried out. Strike match algorithm works very well in information retrieval of web applications and tested out by developing CSSE to validate this algorithm with sample prototype which significantly improves effectiveness of service searching over in cloud.

## REFERENCES

- [1] Luis M. Vaquero, LR-M,JCML 2009, 'A Break in the Clouds: Towards a Cloud Definition', *ACM SIGCOMM Computer Communication Review*, vol 39, no. 1, pp. 50-55
- [2] Abhishek 2011, *Let Google Suggests what you Search". Gtricks Sitemap.*, viewed 28 March 2013, <<http://www.gtricks.com>>.
- [3] Abhishek July 19, 2009., *Top 10 Weird Suggestions From Google That Are Really Amusing*, <<http://www.gtricks.com>>.
- [4] Jaeyong Kang and Kwang Mong Sim 2011, 'A Cloud Portal with a Cloud Service Search Engine', *International Conference on Information and Intelligent Computing*, IACSIT Press, Singapore.
- [5] Alexander Kotov, CZ 2010, 'Towards Natural Question-Guided Search', *International World Wide Web Conference Committee (IW3C2)*, ACM, Raleigh, North Carolina, USA.
- [6] Al-Rawi, SSAA-KB 2010, "Comparison the Efficiency of Some Search Engine on Arabic Keywords and Roots", *International Journal of Computational Linguistics Research*, vol 1, no. 1
- [7] PHILLIP C-Y SHEU, SWQWKHARP 2009, 'Semantic Computing, Cloud Computing, and Semantic Search Engine', *IEEE International Conference on Semantic Computing (2009)*.
- [8] Mauro Gaspar, DG 2003, 'Towards an ontology-guided search engine', Technical Report, Department of Computer Science, University of Bologna, Italy.

## Author' biography with Photo



**Thamilvani Arvaree @ Alvar** receive a Bachelor's degree in Computer Science in 2001 from Inti College Malaysia and Master in Science (Software Engineering) in 2010 from Universiti Putra Malaysia. With her interest in research and teaching currently she pursuing her Phd at Universiti Putra Malaysia. Her area of interest are software engineering, software requirements, software metrics and cloud computing. She is now working as a Assistant Profesor in University of Nottigham, Malaysia.



**Rodziah Atan** received a Bachelor's degree in Computer Science in 1996 from Agricultural University of Malaysia and Master Science in 1998 from Universiti Putra Malaysia. With research experience while her M.S. studies, she completed her PhD from the same university in 2005. She has been supported by the government of Malaysia and the University's Young Lecturer Scheme (SLAB). Her field of interest is software process and business process modeling and pursuing for new knowledge in bioinformatics visualization tools. She is now as a Associate Professor cum researcher in Universiti Putra Malaysia