Available Online at www.ijcsmc.com

International Journal of Computer Science and Mobile Computing

A Monthly Journal of Computer Science and Information Technology

ISSN 2320–088X

IJCSMC, Vol. 3, Issue. 6, June 2014, pg.24 – 27

RESEARCH ARTICLE

Ranked Keyword Search in Cloud Computing: An Innovative Approach

Nikita Sangwan[#1], Sunita Sangwan[*2]

[1]Shri Baba Mastnath Engineering College, MDU, Rohtak
Haryana (India)
ernikitasangwan@gmail.com

[2]Shri Baba Mastnath Engineering College, MDU, Rohtak, Haryana
guide email id@gmail.com

Abstract:
Cloud computing refers to a computing hardware machine or group of computing hardware machines commonly referred as a server or servers connected through a communication network such as the Internet, an intranet, a local area network (LAN) or wide area network (WAN). Any individual user who has permission to access the server can use the server's processing power to run an application, store data, or perform any other computing task. Therefore, instead of using a personal computer every-time to run the application, the individual can now run the application from anywhere in the world, as the server provides the processing power to the application and the server is also connected to a network via Internet or other connection platforms to be accessed from anywhere. While searching the data in the cloud the attackers prefer the keyword which is not secured properly. The existing technique resolves the optimization complexities in ranked keyword search and its effective utilization of remotely stored encrypted cloud data. But it limits the further optimizations of the search results by preventing cloud server to interact with cloud users to maintain the integrity of actual owner's keyword and the data associated with it. The aim is to define a framework which enhances the accuracy of the ranked keyword search by secured machine learning, which does not affect the data integrity

Keywords:
Search in cloud, secured search engines, Ranked Keyword Search

1. Introduction:-

Cloud computing is the delivery of computing services over the Internet. Cloud services allow individuals and businesses to use software and hardware that are managed by third parties at remote locations. Examples of cloud services include online file storage, social networking sites, webmail, and online business applications. The cloud computing model allows access to information and computer resources from anywhere that a network connection is available. Cloud computing provides a shared

pool of resources, including data storage space, networks, computer processing power, and specialized corporate and user applications.

Cloud computing is a model for enabling convenient, on-demand network access to a shared pool of configurable computing resources (e.g., networks, servers, storage, applications, and services) that can be rapidly provisioned and released with minimal management effort or service provider interaction. This cloud model promotes availability and is composed of five essential characteristics, three service models, and four deployment models.

2. Characteristics:-

The characteristics of cloud computing include on-demand self service, broad network access, resource pooling, rapid elasticity and measured service. On-demand self service means that customers (usually organizations) can request and manage their own computing resources. Broad network access allows services to be offered over the Internet or private networks. Pooled resources means that customers draw from a pool of computing resources, usually in remote data centre. Services can be scaled larger or smaller; and use of a service is measured and customers are billed accordingly.

3. Cloud Models:-

- Delivery Models
  - SaaS
  - PaaS
  - IaaS
- Deployment Models
  - Private cloud
  - Community cloud
  - Public cloud
  - Hybrid cloud
- We propose one more Model: Management Models (trust and tenancy issues)
  - Self-managed
  - 3rd party managed (e.g. public clouds and VPC)

1. Infrastructure as a Service (IaaS) also referred to as *Resource Clouds*, provide resources as services to the user – in other words, they basically provide enhanced virtualization capabilities.
   Examples: Amazon S3, SQL Azure.
2. Platform as a Service (PaaS), provide computational resources via a *platform* upon which applications and services can be developed and hosted. PaaS typically makes use of dedicated APIs to control the behavior of a server hosting engine which executes and replicates the execution according to user requests (e.g. access rate).
   Examples: Force.com, Google App Engine.
3. Software as a Service (SaaS), also sometimes referred to as *Service or Application Clouds* are offering implementations of specific business functions and business processes that are provided with specific cloud capabilities.
   Examples: Google Docs, SAP Business by Design.

Public clouds are offered over the Internet and are owned and operated by a cloud provider. Some examples include services aimed at the general public, such as online photo storage services, e-mail services, or social networking sites.

In a **private cloud**, the cloud infrastructure is operated solely for a specific organization, and is managed by the organization or a third party.

In a **community cloud**, the service is shared by several organizations and made available only to those groups. The infrastructure may be owned and operated by the organizations or by a cloud service provider.

A **hybrid cloud** is a combination of different methods of resource pooling (for example, combining public and community clouds.

4. Problem Statement:-

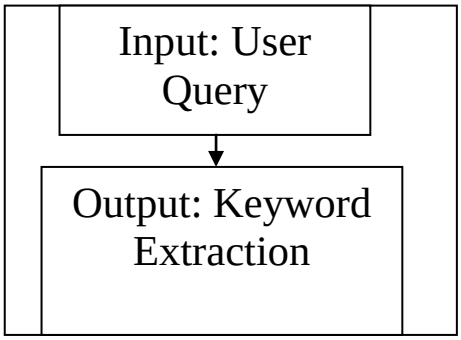
Cloud is one of the major part of Web architecture. The cloud architecture is widely spread in the form of public cloud and available to all the users. There are number of cloud service providers. When a user pass a query for a specific cloud service with well specified requirements then a search engine based work is suggested in this research proposal. In this work at first the user query will be analyzed respective to user requirement specification. Now a parametric match will be performed according to the user interest and other parameters. On the basis of this parametric match the ranking will be assigned to all the eligible cloud services. In this presented work, the ranking criteria are specified on the basis of implicit and explicit properties such as content relevancy, user interest and response time etc.

5. DEVELOPMENT SCHEME FOR RANKING:-

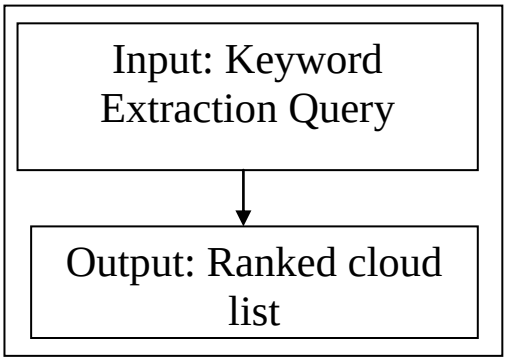
The proposed work is about to optimize the topic based Cloud Service crawling process with the concept of exclusion of duplicate pages. For this a new architecture is proposed, this architecture will use the rank based service selection approach. In this work the ranking is performed respective main criteria's called User Interest Analysis.

The user will interact to the Cloud Service with his topic based query to retrieve the Cloud Service pages. As the page is query performed it will perform request to the Cloud Service and generate the basic URL list. Now it will retrieve the data from the Cloud Service. For the URL collection it will use some concepts like indexing and the ranking. The indexing will provide a fast access to the Cloud Service page where as ranking will arrange the list according to the priority.

Now as a Cloud Service page is fetched, the proposed approach will retrieve the keywords form the document and perform the relevancy match by performing the match of service keywords with user query. Now as a new page is retrieved it will generate the suffix tree and perform a suffix tree based comparison to analyze the relevancy ratio. Based on this factor the initial ranking is assigned to the cloud service.



Now this extracted keyword will work as input to the cloud search architecture and based on the algorithmic approach it will return the effective URL list along with ranking.



Basic concept of the Ranked Keyword Search

The proposed work will be implemented in any Cloud Service based environment. We will implement it in java or .net environment.

6. **Major Challenges:-**

- Most of the work in Cloud area is done on the top layer where the cloud servers exist. But any kind of architectural change in this layer requires complete reconfiguration of cloud network.
- Most of the available clouds have their independent architectures and the resources because of this no such sharing are performed between different clouds.
- Security having lot of flaws in cloud architecture as it is available globally over the net. It requires both the authentication and the encoding mechanism to perform the same work.
- Lot of attacks are there that affects the cloud reliability and efficiency by increasing the unnecessary load on the cloud or to reveal the cloud information.

Conclusion:-

Here we motivate and solve the problem of supporting efficient ranked keyword search for achieving effective utilization of remotely stored outsourced data in a cloud.

We give problem statement and development scheme for our ranked search mechanism, including the efficient support of relevance score dynamics. Clouds offer the opportunity to build data observatories with data, software and expertise together to solve problems such as those associated with economic modeling, climate change, terrorism, Health care and epidemics etc. Clouds could assist greatly in the e-government agenda by providing information in one place to the citizen, together with software to manipulate the data. By enabling a search result authentication mechanism that can detect unexpected behaviors of cloud server like saving cost when handling large number of search requests, software bugs and internal/external attacks.

Repository model is used to make the selection dynamic. Further efficiency can be increased by making slight change in repository model by introducing cache concept.

References:-

1. Akassh A. Mishra, Chinmay Kamat (April 2011),"Migration of search engine process into the cloud"

2. Armbrust.M.Fox. A, Griffith. R, A. D. Joseph, R. H. Katz, A. Konwinski, G. Lee, D. A. Patterson, A. Rabkin, I. Stoica, and M. Zaharia, (Feb 2009) "Above the clouds: A berkeley view of cloud computing," University of California, Berkeley.

3. Boneh .D. G. D. Crescenzo, R. Ostrovsky, and G. Persiano, (2004), "Public key encryption with keyword search," in Proc. of EUROCRYP'04, volume 3027 of LNCS.

4. Y.-C. Chang and M. Mitzenmacher, (ACNS 2005), "Privacy preserving keyword searches on remote encrypted data".

5. Chien Chin Chen and Meng Chang Chen, (SIGIR 2012) "TSCAN: A Content Anatomy Approach to Temporal Topic Summarization".

6. Cong Wang, Ning Cao, Kui Ren, Wenjing Lou "Enabling Secure and Efficient Ranked Keyword Search over Outsourced Cloud Data.

7. Song.D,Agner and Perrig. A, (IEEE 2000), "Practical techniques for searches on encrypted data," in Proc. of IEEE Symposium on Security and Privacy'00.

8. http://en.wikipedia.org/wiki/Web_crawling.

9. AAMEEK SINGH," Search-as-a-Service: Outsourced Search over Outsourced Storage", ACM Transactions on the Web, Vol. 3, No. 4, Article 13, Publication date: September 2009. ACM 1559-1131/2009/09 Anshul Saxena ," A Scalable and Extensible Cloud Service Spider"

10. Bernd Grobauer, Tobias Walloschek, Elmar Stöcker, "Understanding Cloud Computing Vulnerabilities," IEEE Security & Privacy, vol. 9, no. 2, pp. 50-57, March-April 2011, doi:10.1109/MSP.2010.115

11. Cengiz Orencik," Efficient and Secure Ranked Multi-Keyword Search on Encrypted Cloud Data", PAIS 2012, March 30, 2012, Berlin, Germany. ACM 978-1-4503-1143-4/12/03

