IJCSE International Journal of Computer Sciences and Engineering Open Access
Research Paper Volume-3, Issue-11 E-ISSN: 2347-2693

Enhancing Security and Efficiency In Cloud Computing Users Using Multi-keyword Ranking Model

S.M.Manimalathi[1*], A.Senthilkumar[2]

[1]*M.Phil Research Scholar, Department of Computer Science, Tamil University, Thanjavur.*
[2]*Asst. Professor, Department of Computer Science, Tamil University, Thanjavur.*

www.ijcseonline.org

Received: Oct/23/2015 Revised: Nov/06/2015 Accepted: Nov/18/2015 Published: Nov/27/2015

Abstract— The venture defines and understand the issue of multi-catchphrase positioned seek over scrambled cloud Information (MRSE) while saving strict framework insightful assurance in the cloud processing paradigm. Information proprietors are motivated to outsource their complex Information administration frameworks from neighborhood sites to the business open cloud for awesome adaptability and monetary savings. But for ensuring Information privacy, delicate Information have to be scrambled before outsourcing, which obsoletes customary Information usage based on plain content catchphrase search. Thus, enabling a scrambled cloud Information seek administration is importance. Considering the vast number of Information clients and archives in the cloud, it is vital to allow numerous catchphrases in the seek demand and return archives in the request of their significance to these keywords. Related works on searchable encryption center on single catchphrase seek or Boolean catchphrase search, and rarely sort the seek results. Among diverse multi-catchphrase semantics, choosing the proficient likeness measure of "coordinate matching," i.e., as numerous matches as possible, to catch the significance of Information archives to the seek query. Specifically, here use the "internal item similarity" i.e., the number of inquiry catchphrases appearing in a document, to quantitatively assess such likeness measure of that archive to the seek query. Amid the file construction, each archive is related with a twofold vector as a sub file where each bit represents whether relating catchphrase is contained in the document. The seek inquiry is too depicted as a twofold vector where each bit implies whether relating catchphrase appears in this seek request, so the likeness could be precisely measured by the internal item of the inquiry vector with the Information vector. However, straightforwardly outsourcing the Information vector or the inquiry vector will damage the file assurance or the seek privacy. The vector space model helps to give sufficient seek accuracy, and the DES encryption empowers clients to include in the positioning while the majority of processing work is done on the server side by operations just on figure text. As a result, Information spillage can be eliminated and Information security is ensured. Intensive security and execution investigation appear that the proposed scheme ensures high security and practical efficiency.

Keywords— MRSE, OTP, Cloud, Item similarity

I. INTRODUCTION

Cloud processing is a colloquial expression utilized to describe a assortment of diverse types of processing concepts that include vast number of PCs that are joined through a real-time correspondence Framework (typically the Internet). Cloud processing is a jargon term without a commjust accepted non-ambiguous scientific or technical definition. In science, cloud processing is a synonym for distributed processing over a Framework and implies the ability to run a program on numerous joined PCs at the same time. The popularity of the term can be attributed to its use in promoting to sell hosted administrations in the sense of application administration provisioning that run customer server software on a remote location. Cloud processing relies on sharing of assets to accomplish coherence and economies of scale similar to a utility (like the electricity grid) over a network. At the foundation of cloud processing is the broader concept of converged foundation and shared services.

Fig. 1. Architecture of the seek over scrambled cloud data

The cloud too focuses on maximizing the effectiveness of the shared resources. Cloud assets are usually not just shared by numerous clients but as well as dynamically re-allocated as per demand. This can work for allocating assets to clients in diverse time zones. For

example, a cloud computer facility which serves European clients amid European business hours with a specific application (e.g. email) while the same assets are getting reallocated and serve North American clients amid North America's business hours with another application (e.g. web server). This approach should maximize the use of processing powers hence reducing environmental damage as well, since less power, air conditioning, Rackspace, and so on, is required for the same functions.

The term "moving to cloud" too refers to an organization moving away from a customary CAPEX model (buy the dedicated hardware and depreciate it over a period of time) to the OPEX model (use a shared cloud foundation and pay as you use it).

Advocates claim that cloud processing permits companies to avoid upfront foundation costs, and center on projects that differentiate their businesses instead of infrastructure. Advocates too claim that cloud processing permits endeavors to get their applications up and running faster, with improved manageability and less maintenance, and empowers IT to more rapidly adjust assets to meet fluctuating and unpredictcapable business demand.

II. MULTI-CATCHPHRASE POSITIONED SEEK OVER SCRAMBLED (MRSE)

CLOUD processing is the long dreamed vision of processing as a utility, where cloud customers can remotely store their Information into the cloud so as to enjoy the on-interest high-quality applications and administrations from a shared pool of configurcapable processing resources. Its awesome adaptability and monetary savings are motivating both individuals and endeavors to outsource their neighborhood complex Information administration framework into the cloud. To protect Information assurance and combat unsolicited accesses in the cloud and beyond, delicate data, for example, e-mails, personal health records, photo albums, tax documents, financial transactions, and so on, might have to be scrambled by Information proprietors before outsourcing to the business open cloud; this, however, obsoletes the customary Information usage administration based on plaincontent catchphrase search.

The trivial solution of downloading all the Information and decrypting locally is clearly impractical, due to the immense sum of bandwidth cost in cloud scale systems. Moreover, aside from eliminating the neighborhood Capacity management, storing Information into the cloud serves no purpose unless they can be easily searched and utilized. Thus, exploring assurance saving and compelling seek administration over scrambled cloud Information is of parse importance. Considering the potentially vast number of on-interest Information clients and immense sum of outsourced Information archives in the cloud, this issue is particularly testing as it is extremely difficult to meet too the necessities of performance, framework usability, and scalability.

On the one hand, to meet the compelling Information recovery need, the vast sum of archives demand the cloud server to perform result significance ranking, instead of returning undifferentiated results. Such positioned seek framework empowers Information clients to find the most pertinent Information quickly, maybe than burdensomely sorting through each match in the content collection. Positioned seek can too completely eliminate vital Framework movement by sending back just the most pertinent data, which is exceedingly desirable in the “pay-as-you-use” cloud paradigm. For assurance protection, such positioning operation, however, should not spill any catchphrase related information. On the other hand, to improve the seek result exactness as well as to upgrade the customer looking experience, it is too vital for such positioning framework to support numerous catchphrases search, as single catchphrase seek regularly yields far too coarse results. As a common practice indicated by today’s web seek engines (e.g., Google search), Information clients might tend to give a set of catchphrases instead of just one as the indicator of their seek interest to recover the most pertinent data. And each catchphrase in the seek demand is capable to help narrow down the seek result further. “Coordinate matching”, as numerous matches as possible, is a proficient likeness measure among such multi-catchphrase semantics to refine the result relevance, and has been widely utilized in the plain content Information recovery (IR) community. However, how to apply it in the scrambled cloud Information seek framework remains a very testing task because of inherent security and assurance obstacles, counting diverse strict necessities like the Information privacy, the file privacy, the catchphrase privacy, and numerous others.

Encryption is a helpful procedure that treats scrambled Information as archives and permits a customer to securely seek through a single catchphrase and recover archives of interest. However, direct application of these approaches to the secure vast scale cloud Information usage framework would not be necessarily suitable, as they are developed as crypto primitives and can’t accommodate such high service-level necessities like framework usability, customer looking experience, and easy Information discovery. Although some recent designs have been proposed to support Boolean catchphrase seek as an attempt to enrich the seek flexibility, they are still not adequate to give clients with acceptable result positioning functionality. Our early works have been aware of this problem, and give solutions to the secure positioned seek over scrambled Information issue but just for queries consisting of a single keyword. How to outline an proficient scrambled Information seek instrument that supports multi-catchphrase semantics without assurance ruptures still remains a testing open problem.

In the project, for the first time, characterize and understand the issue of multi-catchphrase positioned seek over scrambled cloud Information (MRSE) while saving strict framework insightful assurance in the cloud processing paradigm. Among diverse multi-catchphrase semantics, pick the proficient likeness measure of "coordinate matching," i.e., as numerous matches as possible, to catch the significance of Information archives to the seek query. Specifically, internal item likeness the number of inquiry catchphrases appearing in a document, to quantitatively assess such likeness measure of that archive to the seek query. Amid the file construction, each archive is related with a twofold vector as a sub-file where each bit represents whether relating catchphrase is contained in the document. The seek inquiry is too depicted as a twofold vector where each bit implies whether relating catchphrase appears in this seek request, so the likeness could be precisely measured by the internal item of the inquiry vector with the Information vector. However, straightforwardly outsourcing the Information vector or the inquiry vector will damage the file assurance or the seek privacy. To meet the challenge of supporting such multi catchphrase semantic without assurance breaches, we propose a fundamental thought for the MRSE utilizing secure internal item computation, which is adapted from a secure k-nearest neighbor (kNN) technique, and then give two significantly improved MRSE plans in a step-by-step manner to accomplish diverse stringent assurance necessities in two risk models with increased attack capabilities. Our contributions are summarized as follows:

A. For the first time, we relook the issue of multi catchphrase positioned seek over scrambled cloud data, and establish a set of strict assurance necessities for such a secure cloud Information usage system.

B. We propose two MRSE plans based on the likeness measure of "coordinate matching" while meeting diverse assurance necessities in two diverse risk models.

C. We relook some further upgrades of our positioned seek instrument to support more seek semantics and dynamic Information operations.

D. Intensive investigation researching assurance and productivity ensures of the proposed plans is given, and tests on the certifiable Information set further appear the proposed plans indeed present low overhead on calculation and communication.

Compared with the preliminary rendition of this paper, this journal rendition proposes two new mechanisms to support more seek semantics. This rendition too studies the support of data/file dynamics in the instrument design. Moreover, we improve the experimental works by adding the investigation and evaluation of two new schemes. In addition to these improvements, we add more investigation on secure internal item and the assurance part.

III. OBJECTIVE

Customer Positioning Guarantee *why* something is mentioned a lot, and that it isn't due to marketing, or self-promotion, maybe than importance. Proposed cloud Capacity frameworks that give confidentiality, respectability and verifiability of customer Information against a UN trusted cloud provider. This OTP utilized to see Information in cloud and it can be utilized once just in a time, when you seek a record and tend to see the record the OTP will send to email and you get the OTP and apply to see the record ensuring data. For this reason, outsourced records must be encrypted. Any kind of Information spillage that would influence Information assurance is respected as unacceptable. To meet the compelling Information recovery need, the vast sum of archives demand the cloud server to perform result significance ranking, instead of returning undifferentiated results. Such positioned seek framework empowers Information clients to find the most pertinent Information quickly, maybe than burdensomely sorting through each match in the content collection. Positioned seek can too carecompletely eliminate unvital Framework movement by sending back just the most pertinent data, which is exceedingly desircapable in the "pay-as-you-use" cloud paradigm. For assurance protection, such positioning operation, however, should not spill any catchphrase related information. On the other hand, to improve the seek result exactness as well as to upgrade the customer looking experience, it is too vital for such positioning framework to support numerous catchphrases search, as single catchphrase seek regularly yields far too coarse results.

IV. PROPOSED SYSTEM

The research, they introduce the concepts of similarity relevance and scheme robustness to formulate the privacy issue in searchable encryption schemes, and then solve the insecurity problem by proposing a Self-Destruction Multi-Keyword searchable encryption (SED-MKSE) scheme. Novel technologies in the cryptography community and information retrieval community are employed, including homomorphic encryption and vector space model.

The proposed scheme, the majority of computing work is done on the cloud while the user takes part in ranking, which guarantees top k multi-keyword retrieval over encrypted cloud data with high security based on Sedas and practical efficiency.

Framework Architecture

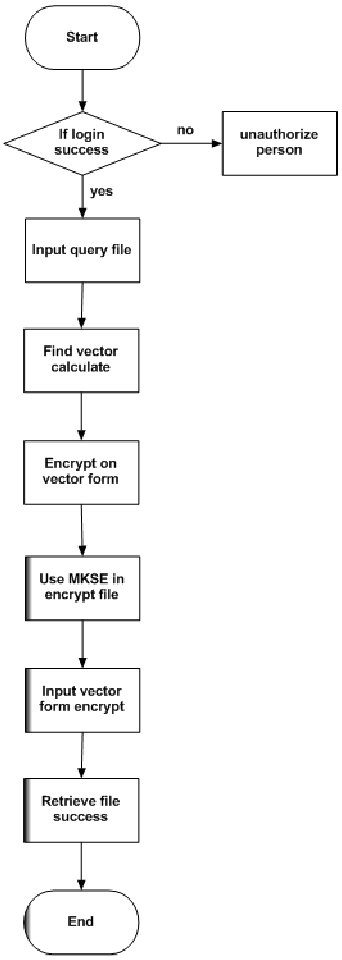


Fig.2 Architecture of the MRSE Implementation.

Advantages

- Secured Multi-keyword Ranked Search: To design search schemes which allow multi-keyword query and provide result similarity ranking for valuable data retrieval, instead of returning undifferentiated results.
- Privacy: To prevent cloud server from learning additional information from dataset and index, and to meet privacy requirements.
- Effectiveness with high performance: Above goals on functionality and privacy should be achieved with low communication and computation overhead.
- Thorough analysis on security demonstrates the proposed scheme guarantees high data privacy. Furthermore, performance analysis and experimental results show that the scheme is efficient for practical utilization.

V. EXISTING SYSTEM

In the existing system, To improve feasibility and save on the expense in the cloud paradigm, it is preferred to get the retrieval result with the most relevant files that match users' interest instead of all the files, which indicates that the files should be ranked in the order of relevance by users' interest and only the files with the highest relevance's are sent back to users. A series of searchable symmetric encryption schemes have been proposed to enable search on cipher text. Traditional SSE schemes enable users to securely retrieve the cipher text, but these schemes support only Boolean keyword search, i.e., whether a keyword exists in a file or not, without considering the difference of relevance with the queried keyword of these files in the result.

Preventing the cloud from involving in ranking and entrusting all the work to the user is a natural way to avoid information leakage. However, the limited computational power on the user side and the high computational overhead precludes information security.

Disadvantages

- To improve security without sacrificing efficiency, schemes presented in show that they support top-k multi keyword retrieval under various scenarios.
- Authors of made attempts to solve the problem of top-k multi-keyword over encrypted cloud data.
- These schemes, however, suffer from two problems - Boolean representation and how to strike a balance between security and efficiency.
- In the former, files are ranked only by the number of retrieved keywords, which impairs search accuracy. In the latter, security is implicitly compromised to tradeoff for efficiency, which is particularly undesirable in security-oriented applications.

Module Description:

- User Creation For SSE
- Vector Calculation
- MKSE design
- User Login and Retrieve data
- User Query for SSE

a. User Creation for SSE

The user create their own login credential when they want entering in cloud. The keys are automatically generated when the user register in cloud. Considering the large number of data users and documents in the cloud, it is necessary to allow multi-keyword in the search query and return documents in the order of their relevancy with the queried keywords. Scoring is a natural way to weight the

relevance. Based on the relevance score, files can then be ranked in either ascending or descending. Several models have been proposed to score and rank files in information retrieval (IR) community.

b. Vector Calculation

Although all data files, indices and requests are in encrypted form before being outsourced onto cloud, the cloud server can still obtain additional information through statistical analysis. They denote the possible information leakage with statistic leakage. There are two possible statistic leakages, including term distribution and inter distribution. The term distribution of term t is t's frequency distribution of scores on each file .The inter distribution of file f is file f's frequency distribution of scores of each term. Term distribution and inter distribution are specific. They can be deduced either directly from cipher text or indirectly via statistical analysis over access and search pattern.

The vector calculation is performed to the term frequency user querying file. Here access pattern refers to which keywords and the corresponding files have been retrieved during each search request, and search pattern refers to whether the keywords retrieved between two requests are the same.

c. MKSE design

Existing SSE schemes employ server-side ranking based on order preserving encryption to improve the efficiency of retrieval over encrypted cloud data. However, server-side ranking based on order-preserving encryption violates the privacy of sensitive information, which is considered un comprisable in the security-oriented third-party cloud computing scenario, i.e., security cannot be tradeoff for efficiency. To achieve data privacy, ranking has to be left to the user side.

Traditional user-side schemes, however, load heavy computational burden and high communication overhead on the user side, due to the interaction between the server and the user including searchable index return and ranking score calculation. Thus, the user side ranking schemes are challenged by practical use. A more server-siding scheme might be a better solution to privacy issues. They propose a new searchable encryption scheme, in which novel technologies in cryptography community and IR community are employed, including homomorphic encryption and vector space model.

In the proposed scheme, the data owner encrypts the searchable index with homomorphic encryption. When the cloud server receives query consisting of multi-keyword, it computes the scores from the encrypted index stored on cloud, and then returns the encrypted scores of files to the data user. Next, the data user decrypts the scores and picks out the top-k highest-scoring files' identifiers to request to the cloud server.

The retrieval takes a Multi-Keyword communication between the cloud server and the data user.

d. User Login and Retrieve data

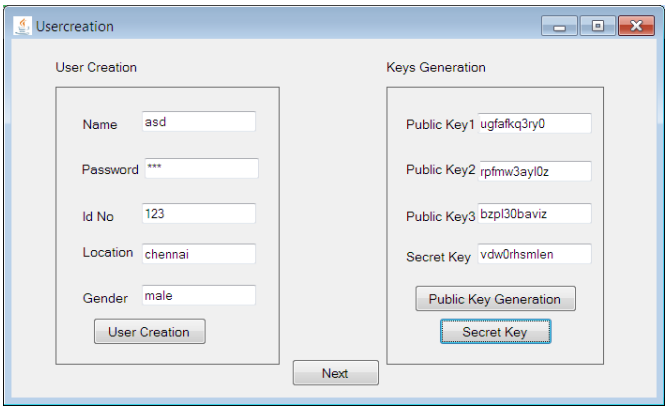
After find the term frequency of the data vector value can be calculated. Then we upload file to client and at same time it can be send to cloud. Because the uploading cost is similar. But the client receive the original file and the encrypted file can send to the cloud.

e. User Query for SSE

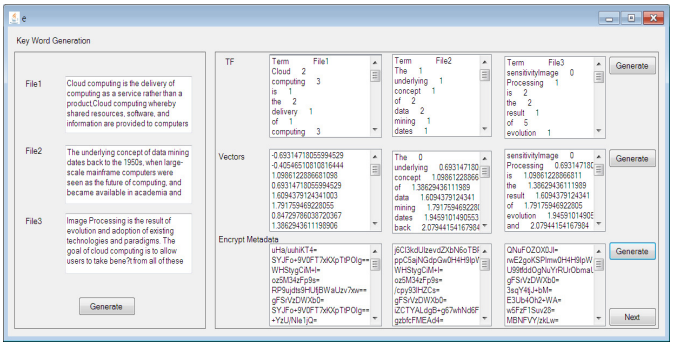
To alleviate the computational burden on user side, computing work should be at the server side, so they need an encryption scheme to guarantee the operability and security at the same time on server side. MKSE allows specific types of computations to be carried out on the corresponding cipher text. The result is the cipher text of the result of the same operations performed on the plaintext. That is, MKSE allows computation of cipher text without knowing anything about the plaintext to get the correct encrypted result. Although it has such a fine property, original fully homomorphic encryption scheme, which employs ideal lattices over a polynomial ring, is too complicated and inefficient for practical utilization.

VI. IMPLEMENTATION

A. User Registeration and Key Generation

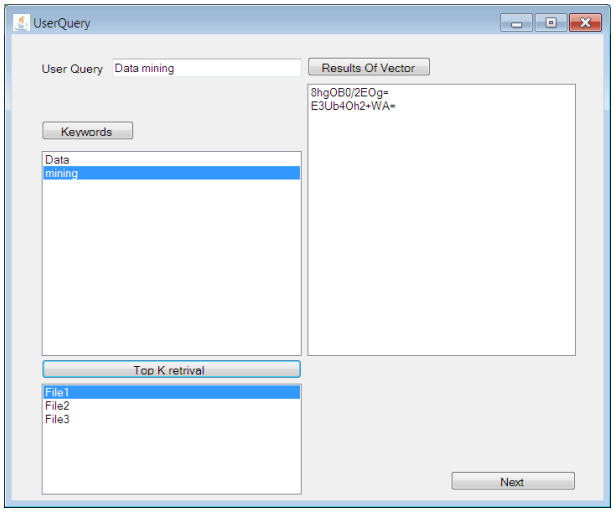


B. File is Computed in to RS Vector



C. Upload File to cloud By Encryption

D. User Query For File Retrieval



VII. CONCLUSION AND FUTURE WORK

This research, for the first time they define and solve the problem of multi-keyword ranked search over encrypted cloud data with Security, and establish a variety of privacy requirements. Among various multi-keyword semantics, they choose the efficient similarity measure of "coordinate matching", i.e., as many matches as possible, to effectively capture the relevance of outsourced documents to the query keywords, and use "inner product similarity" to quantitatively evaluate such similarity measure. For meeting the challenge of supporting multi-keyword semantic in improved security without privacy breaches, they propose a basic idea of SED-MRSE using secure inner product computation.

Then they give two significantly improved MRSE schemes to achieve various stringent privacy requirements in two different threat models. Thorough analysis investigating privacy and efficiency guarantees of proposed schemes is given, and experiments on the real-world dataset show the proposed schemes introduce enhanced security with low overhead on both computation and communication.

REFERENCES

[1] Karapakula, A.; Puramchand, M.; Rafi, G.M."Coordinate matching for effective capturing the similarity between query keywords and outsourced documents", Published in: Sustainable Energy and Intelligent Systems (SEISCON 2012), IET Chennai 3rd International on, Date of Conference:27-29 Dec. 2012Page(s):1- 7.

[2] Laveti, G.; Dept. of E.C.E, Sangivalasa, Visakhapatnam, India; Sasibhushana Rao, G.; Kumar, M.N.V.S.S.; Goswami, R. ," A Novel TOA measurement based global evolutionary search estimator for defence applications", Published in: Electrical, Electronics, Signals, Communication and Optimization (EESCO), 2015 International Conference on, Date of Conference:24-25 Jan. 2015, Page(s):1 – 4

[3] Zhiyong Xu; Shenzhen Inst. of Adv. Technol., Shenzhen, China; Wansheng Kang; Ruixuan Li; Kinchoong Yow ," Efficient Multi-Keyword Ranked Query on Encrypted Data in the Cloud", Published in:Parallel and Distributed Systems (ICPADS), 2012 IEEE 18th International Conference on, Date of Conference:17-19 Dec. 2012,Page(s):244 – 251.

[4] Mantel, C.; Dept. of Photonics Eng., Tech. Univ. of Denmark, Lyngby, Denmark; Ferchiu, S.C.; Forchhammer, S.," Comparing subjective and objective quality assessment of HDR images compressed with JPEG-XT",Published in: Multimedia Signal Processing (MMSP), 2014 IEEE 16th International Workshop on, Date of Conference:22-24 Sept. 2014, Page(s):1 – 6

[5] Woan Yun Hsiao; Inst. of Electron. Eng., Nat. Tsing-Hua Univ., Hsinchu, Taiwan; Chin Yu Mei; Wen Chao

Shen; Tzong Sheng Chang ," A high density Twin-Gate OTP cell in pure 28nm CMOS process ",Published in: VLSI Technology, Systems and Application (VLSI-TSA), Proceedings of Technical Program - 2014 International Symposium on ,Date of Conference:28-30 April 2014 Page(s):1 – 2.

[6] Chang-Lung Tsai; Dept. of Comput. Sci. & Inf. Eng., Chinese Culture Univ., Taipei, Taiwan; Chun-Jung Chen; Deng-Jie Zhuang," Secure OTP and Biometric Verification Scheme for Mobile Banking",Published in:Mobile, Ubiquitous, and Intelligent Computing (MUSIC), 2012 Third FTRA International Conference on ,Date of Conference:26-28 June 2012 ,Page(s):138 - 141

[7] Chao-Hsi Huang; Inst. of Comput. Sci. & Inf. Eng., Nat. Ilan Univ., Ilan, Taiwan; Shih-Chih Huang," RFID systems integrated OTP security authentication design",Published in:Signal and Information Processing Association Annual Summit and Conference (APSIPA), 2013 Asia-Pacific, Date of Conference:Oct. 29 2013-Nov. 1 2013, Page(s):1 – 8.

[8] Wenge, O.; Multimedia Commun. Lab. (KOM) Tech., Univ. Darmstadt, Darmstadt, Germany; Schuller, D.; Steinmetz, R.," Towards Establishing Security-Aware Cloud Markets", Published in:Cloud Computing Technology and Science (CloudCom), 2014 IEEE 6th International Conference on, Date of Conference:15-18 Dec. 2014, Page(s):1027 – 1032

[9] Nagalakshmi, N.; Rajalakshmi, S.," Enabled security based on elliptic curve cryptography with optimal resource allocation schema in cloud computing environment", Published in:Computing, Communication and Information Systems (NCCCIS), 2015 IEEE Seventh National Conference on, Date of Conference:13-14 Feb. 2015,Page(s):17 – 22.

[10] Khanna, P.; Inst. of Eng. & Technol., JK Lakshmipat Univ., Jaipur, India; Jain, S.; Babu, B.V.," Cloud Broker: Working in federated structures",Published in:Advances in Computing, Communications and Informatics (ICACCI, 2014 International Conference on, Date of Conference:24-27 Sept. 2014,Page(s):1273 – 1278.

[11] Yefei Zha; Sch. of Comput. Sci. & Eng., Southeast Univ., Nanjing, China; Yuqing Zhai," An Improved Collaborative Filtering Model Considering Item Similarity",Published in:Information Science and Cloud Computing Companion (ISCC-C), 2013 International Conference on, Date of Conference:7-8 Dec. 2013,Page(s):428 – 434.

[12] Wang Weijie; Dept. of Comput. Sci. & Technol., East China Normal Univ., Shanghai, China; Yang Jing; He Liang," An Improved Collaborative Filtering Based on Item Similarity Modified and Common Ratings", Published in: Cyberworlds (CW), 2012 International Conference on, Date of Conference:25-27 Sept. 2012,Page(s):231 – 235.

[13] Bai Juan; Dept. of Inf. Eng., North China Univ. of Water Conservancy & Electron. Power, Zhengzhou, China," Collaborative filtering recommendation algorithm based on semantic similarity of item", Published in:Advanced Computational Intelligence (ICACI), 2012 IEEE Fifth International Conference on, Date of Conference:18-20 Oct. 2012, Page(s):452 – 454.

[14] Feng Xie; Dept. of Autom., Tsinghua Univ., Beijing, China; Zhen Chen; Jiaxing Shang; Wenliang Huang ," Item Similarity Learning Methods for Collaborative Filtering Recommender Systems", Published in:Advanced Information Networking and Applications (AINA), 2015 IEEE 29th International Conference on, Date of Conference:24-27 March 2015, Page(s):896 – 903.

