International Journal of Scientific Research in Computer Science, Engineering and Information Technology
© 2019 IJSRCSEIT | Volume 5 | Issue 4 | ISSN : 2456-3307
DOI : https://doi.org/10.32628/CSEIT195432

Improving Efficiency of Quality Correlation in Query Using Spark

P. Suma Priya*[1], Dr. P. Radhika raju[2], Prof. P. Ananda Rao[3]

*[1]M.Tech Scholar, Department of Computer Science and Engineering, JNTUA College of Engineering, Ananthapuramu, Andhra Pradesh, India

[2]Ad-hoc Assistant Professor, Department of Computer Science and Engineering, JNTUA College of Engineering, Ananthapuramu, Andhra Pradesh, India

[3]Professor, Department of Computer Science and Engineering, JNTUA College of Engineering, Ananthapuramu, Andhra Pradesh, India

ABSTRACT

Service oriented models to easily development of the service-based systems, many services has been developed. Based on the service composition it creates the enterprise services. The main aim of the service-based system is to fulfil the specific quality constraints. Quality aware service composition cannot consider the quality correlations among services, to solve this problem develop an approach efficient query of quality correlations(Q^2C). To consider these type of quality correlations further increasing search space, to reducing search space and provide efficient correlations develop quality correlation index graph(QCIG). In the QCIG quality constraints in the quality dimensions increased of pre-processing time and search time by the quality correlation. When raising this issue which affect the efficiency of Q^2C. To improve efficiency of Q^2C and get fast, better results use spark distributed technique.

Keywords : Apache Spark, Service Composition, Quality of Service, Query of Quality Correlation (Q^2C), Spark Distributed Technique, Quality Correlation Algorithms

I. INTRODUCTION

Service oriented architecture has become a major framework for building complex distributed software systems by discovering web services. Demand of building service-based system in SOA applications with the usage of composed available web services had interested. Most important step in service-based system development is to recognize the suitable services to the service-based system-engineering process. In the service oriented architecture, the service composition has not consider the quality correlations so to consider the proper correlations query of the quality correlation is introduced. For improving efficiency of quality correlation using the spark distributed technique.

Service composition has not totally considered the quality correlations among the services. It does not have proper checking so it will affect the quality of SBSs. Consider the quality correlations among the services an approach has been proposed that approach have poor efficiency. To overcome all these problems query of quality correlation (Q^2C) has proposed. Calculate the quality correlations among the services with proper aggregation functions. Calculate the correlations using the quality correlation model in this quality model three types of correlation models are there, those are adjacent, non-

adjacent quality, hybrid quality correlations. Adjacent quality correlation is a correlation between the two nearest services, that means the nearest distance between the two quality services, each service have price. Non-adjacent quality correlation is a correlation between the two far distance services for calculating the efficient quality correlations. Finally hybrid quality correlation in this model consider the combination of the both nearest and far distance services, at a time taking two are more services. In adjacent quality correlation model accept all quality dimensions (i.e. availability, response time) in non-adjacent quality correlation it can accept only system independent dimensions (i.e.; cost, availability). Construct the QCIG for providing efficient queries to the Q^2C and reduce the search space problem .In future consider the lack of inspection quality correlations it will affect the search space problem. Main purpose of this graph is providing efficient queries without any failures. In the QCIG aggregate, the quality correlations it does not get proper solutions so after that apply all the correlations in the QCIG it is the worst-case n very inefficient. To improve efficiency of Q^2C aggregates the some of the correlations. Based on the QCM taking quality correlations, aggregate the two compatible services generate the new services. Compatible services mean it belongs to the same category of the candidate services. Whenever two services are compatible aggregate those two services, generate new services, like that checking the correlations, and generate the new services. After aggregate, the correlations construct the quality correlation index graph. In the index graph, each services containing relation by checking the relation to the services construct the graph. Compare the three parameters in three stages or constraints to improve efficiency of Q^2C. Those parameters are the number of candidate services, the quality correlation rate, the number of quality dimensions. Moreover, constraints are simple, medium, and severe. While improve efficiency of the quality correlations in the index graph parameters take more time to complete the process and cannot give proper solution. So that it will affect the efficiency of the quality, correlations because of the pre-processing time.

In current scenario, the spark distributed technique API is useful to reduce the computation time of quality correlation. When using this technique give thousands time better results compare to normal approach or process, within Nano seconds it will give results. Spark is one of the part in big data, compared to Hadoop and map reduce spark have speed performance.

The remaining paper follows. In section II overview about the various approaches is presented. Section III, presents about problem definition Section IV spark and working of spark distributed API. V implementation VI Shows results and computation time reduction. Finally, section VII includes some enhancements and future work.

II. RELATED WORK

Qiang He [1]. This approach supports the service based system services for structured and successful services .CASS properly captures the quality between the two or more services. This result shows the high efficiency and effectiveness. Qiang He [2] at different maturity levels of multi-tenancy it reduce the multi tenant SBSs, improving efficiency and gets the better and fast response. It is used to reducing the search space for optimization problem. Jong Myoung Kim [3] discovers the correlation between search quality and quality popularity, TF-IDF is use to improve search quality. Mohammad Alrifai [4] at run time it determines the best quality levels based on the available QoS information. Shuiguang Deng [5] by taking the quality correlations CASP can improve the Quality of service composite services. It is focus on Quality aware service composition and takes QoS correlations between services. T. Yu, Y. Zhang [6]

service selection problem with multiple QOS constraints. Mohammad Alrifai [7] Skyline services are use to improve efficiency and solve the problem of QoS based web service composition. Lina Barakat [9] quality aware service selection handles QOS among the services and improving the composition quality. Quality aware pruning techniques are introduce to reduce the cost and results shows that best utility and efficient. Zhenqiu Huang [10] efficient pruning, filtering algorithms are used to handle the large number of services, reducing the search space. Jong Myoung Kim [11] investigated correlation between search quality and quality popularity, Shows a better quality results.

III. PROBLEM DEFINITION

Service based systems do not consider the efficient quality correlations in between the services. Impact the quality of SBSs because of does not have a proper inspection. Approach has been proposed to consider the efficient quality correlations it have poor efficiency. To consider this type of quality correlation further will affect the search space problem. To reduce the search space problem and improve efficiency quality correlation algorithms are used. Those are the quality correlation aggregation and QCIG construction based on those two algorithms provide the efficient queries to Q^2C and reduce the search space problem. In QCA provide the efficient queries and improve efficiency it is compatible the two quality correlations after that create the new correlations until how many correlations are generate. It is Base on the QCA constructs the QCIG in this graph checking the every two services having the relationship. By checking that process constructs the graph based on the relation graph will be construct. After applying, those algorithms cannot find the quality correlation that matches to the services. After completion of the graph construction, improve the efficiency by the graphical representation. Three parameters are use in three difficulty levels for improving efficiency. It is taking more time to complete the total process so because of the computation time efficiency will be reduced. To reduce the computation time and improve the efficiency of quality correlations using the spark distributed technique. It is read the data by using the multithreading processing and take less time to complete the process when compare to the normal processes. This is the challenging problem considered in this paper.

IV. PROPOSED FRAMEWORK

In this framework, explain the stepwise overview of proposed system architecture. There are the five steps including the results.

A. Input Data Set (list of services):

In this step taking the services as input, those services are s1, s2, s3 until s6. Each service has cost and id. In this each service based on the services cost assigns the discount to every service is one of non functional quality constraints. After giving the services as input, read the data internally and generate the service details with tabular format. The data sets are taking in the existed system; it is taking those data related to example for online shopping like Amazon, flipchart, etc. In the Online shopping it set of services are present in that system. It is all services are correlation between one services to another services than calculate the correlation among their service cost.

.

B. Quality Correlation Algorithms

In this step, the quality correlation algorithms are used to calculate the quality correlations among the services and provide the efficient queries, improve the efficiency of CQ. Two algorithms are used those are the QCA and QCIG construction.

a. Quality Correlation Aggregation

Quality correlation aggregation is use for aggregating the quality correlations and generates the new correlations. In this aggregation, the correlations are compatible based on the same class of the candidate services. It is aggregate the services until no new correlations are to be creating. Main aim of this approach is creating the new correlations.

b. Quality correlation index graph:

The purpose of the Quality correlation index graph is to construct the graph based on the quality correlations by checking the relation to the two services .two quality services relation is important for construct the graph. It is checking from the entry node until end of the services.

In the quality correlation aggregation, applying all the correlations after applying the results it is worst case and inefficient. To improve the efficiency of QC aggregates the more services. To aggregate the services first compatible the two services, that means two services are belongs to same class of the candidate services. After compatible, the new correlations are generating, it will compatible until no new correlations are create. For efficiency purpose, generate the individual correlations by using the QCA. After that, the QCIG is constructing based on containing the relation among the quality correlations. By checking the relation insert, the correlations from entry node to end .to checking the QC construct this graph. Improve the efficiency of QC using the three parameters in three stages. To completion of the total process, it is taking more time so affect the efficiency of the quality correlation. To reduce the computation time SDT is use. By using this technique, the results will be better compare to the existing system.

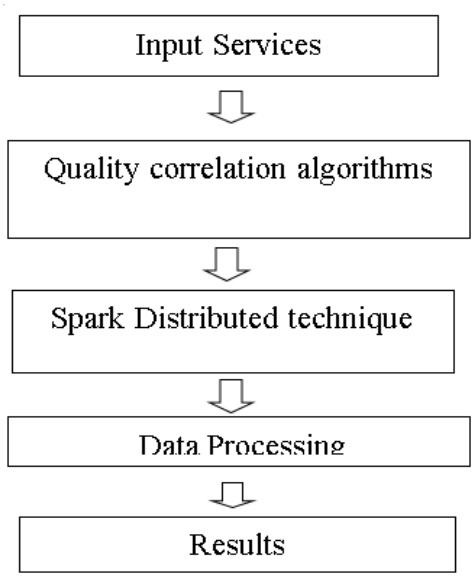

Figure 1. Overview of the proposed system architecture

C. Spark Distributed Technique:

It is one of the best approaches to handle the large number of applications; it can access the multiple data at a time. The main purpose of this technique in this framework is reducing the computation time. Before algorithms cannot give the correct results so to give better results applied this SDT. It's read the data in efficient manner by using the distributed thread processing mechanism. Normal program will read the data in single thread and take more execution time but the SDT read the data in multiple threads take less time to complete the process. This is the main advantage in this technique. it is reducing the computational results and improving the efficiency.

D. Processing Data:

It is used to processing the input data and gives the results without any failures. After applying the distributed technique, the data will be processes and generated the better results.

V. IMPLEMENTATION

Quality aware service composition cannot consider the quality correlations among the services. For calculating the quality correlations in different models, this data set will be use. By using this data set shows the correlation between one service to another service.

Data sets used

Data sets are collect from the SBSs for the experiments. Dataset related to the online shopping for the service based system services. Users loading the datasets it will automatically generate the services based on that services calculate the quality correlations in different stages. in the below figure represents the dataset details, every services have the some cost. For example s1,1 services have 936 cost and s2,2 have 860 like that every services have cost for calculating the QC. Figure 2

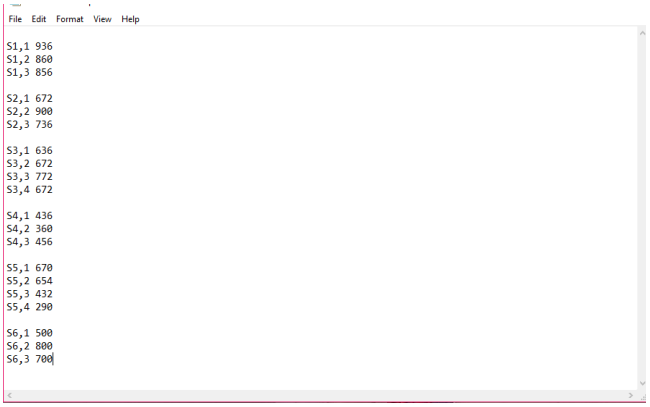

Figure 2. Dataset details

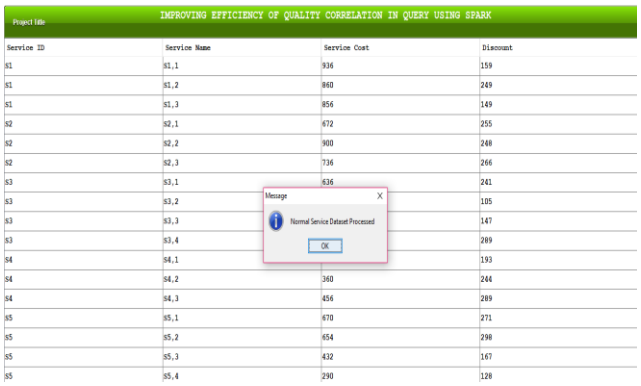

Figure 3. Loaded list of Services

As shown in figure 3 it is evident that it facilitates loading dataset. In this dataset for every service have the id, cost and discount value for that cost. When loading the dataset it will automatically processing and get the data like displayed in the above figure. For example S1 is the service id and S1,1 is the name of that id and cost for that service is 930 and discount value is 224.

VI. EXPERIMENTAL RESULTS

The experimental results show the computation time reduction compare between the different approaches.

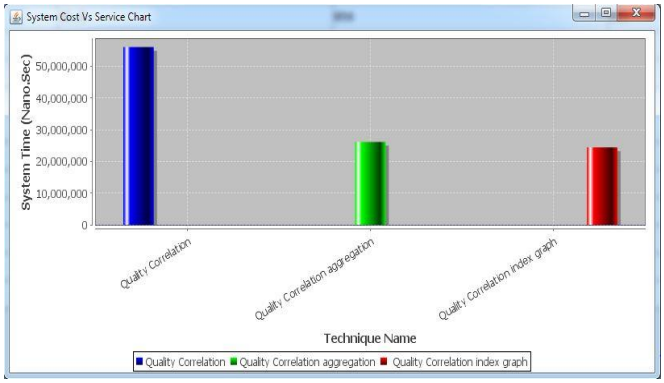

Figure 4. Improving efficiency in between three parameters

As presented in the fig 4, three techniques are represented the quality correlation, quality correlation aggregation and quality correlation index graph in the x-axis while the y-axis showed computational time in nanoseconds. It is evident that the computation time with existing system various time values are required. Evaluate the system cost for these models within Nano seconds, after getting the results the system time will be accuracy and it will take less time to complete the total process.

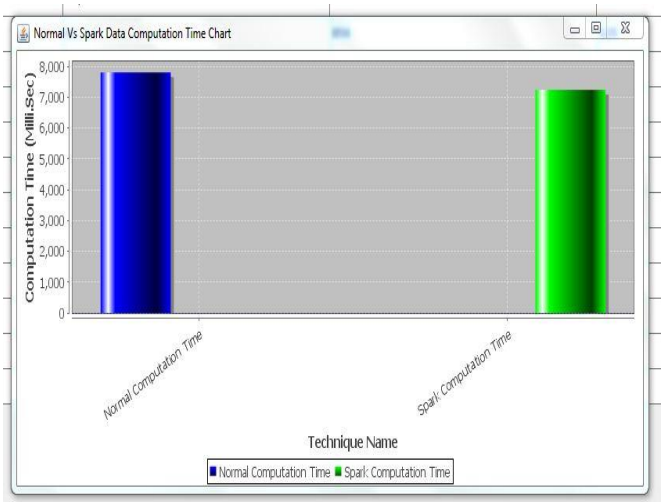

Figure 5. Computation time of the normal system and spark system

As presented in the figure 5, it is evident that the spark computation time against the normal computation time. Normal system and spark system computation time represented in x-axis and while the y-axis showed computational time in milliseconds. Compare to the normal computation system time while more computation time and stark computation system is less time to normal system. The spark system is better computation time.

VII. CONCLUSIONS AND FUTURE SCOPE:

The framework which is develop to improve efficiency of query of quality correlation (Q^2C).the problems are involved in quality correlation that process needs user involvement with knowledge on spark distributed technique for improving efficiency of quality correlation. In the literature, these problems are found with the existing system. In the proposed the main framework name SDT for reducing computation time. Compare to the experimental results proposed system showed the better performance compare to the existing system. In future work, implement the different techniques/ approaches to reduce the computation time of quality correlations.

VIII. REFERENCES

1. Qiang He, "Quality-Aware Service Selection for Service-Based Systems Based on Iterative Multi-Attribute Combinatorial Auction" IEEE TRANSACTIONS ON SOFTWARE ENGINEERING, VOL. 40, NO. 2, FEBRUARY 2014.
2. Qiang He, Jun Han, Feifei Chen, Yanchun Wang "QoS-Aware Service Selection for Customizable Multi-Tenant Service-Based Systems: Maturity and Approaches ", 2015 IEEE 8th International Conference on Cloud Computing.
3. Jong Myoung Kim, Hancheol Park, Gahgene Gweon, and Jeong Hur "The Correlation between Search Quality and Query Popularity" 2016 International Conference on Big Data and Smart Computing (Big Comp), 18-20 jan 2016.
4. M. Alrifai and T. Risse, "Combining Global Optimization with Local Selection for Efficient QoS-Aware Service Composition," Proc. the 18th International Conference on World Wide Web (WWW2009), Madrid, Spain, pp. 881-890, 2009 .
5. Deng, H. Wu, D. Hu, and J. L. Zhao, "Service Selection for Composition with QoS Correlations," IEEE Transactions on Services Computing, vol. 9, no. 2, pp. 291-303, 2016.
6. T. Yu, Y. Zhang, and K.-J. Lin, "Efficient Algorithms for Web Services Selection with End-to-End QoS Constraints," ACM Transactions on the Web, vol. 1, no. 1, 2007.
7. M. Alrifai, D. Skoutas, and T. Risse, "Selecting Skyline Services for QoS-based Web Service Composition," Proc. the 19th International Conference on World Wide Web, Raleigh, North Carolina, USA, pp. 11- 20, 2010.
8. Yiwenzhang, guangming cui," Efficient query of quality correlation for service composition" IEEE transactions on service computing.

9. Lina Barakat, Simon Miles, Michael Luck "Efficient Correlation-aware Service Selection" 2012 IEEE 19th International Conference on Web Services.
10. Zhenqiu Huang, Wei Jiang, Songlin Hu, Zhiyong Liu "Effective Pruning Algorithm for QoS-Aware Service Composition" 2009 IEEE Conference on Commerce and Enterprise Computing.
11. Jong Myoung Kim, Hancheol Park, "The Correlation between Search Quality and Query Popularity"2016 IEEE transactions on big data computing.
12. John F. Kolen and Tim Hutcheson "Reducing the Time Complexity of the Fuzzy C-Means Algorithm" IEEE transactions on fuzzy systems, vol 10, n0 2, APRIL 2002.
13. Zhen Ye, Xiaofang, Zhou "Genetic Algorithm Based QoS-Aware Service Compositions in Cloud Computing" 16th international conference, pp. 321-334, 2011.

Cite this article as :

P. Suma Priya, Dr. P. Radhika Raju, Prof. P. Ananda Rao, "Improving Efficiency of Quality Correlation in Query Using Spark", International Journal of Scientific Research in Computer Science, Engineering and Information Technology (IJSRCSEIT), ISSN : 2456-3307, Volume 5 Issue 4, pp. 182-188, July-August 2019. Available at doi : https://doi.org/10.32628/CSEIT195432
Journal URL : http://ijsrcseit.com/CSEIT195432