Journal of Theoretical and Applied Information Technology
30th November 2016. Vol.93. No.2

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

SCALE TO COMPOSITION FAULT INCLINED (SCFI) &SCALE TO COMPOSITION HALENESS (SCH): DESIGN OF HEURISTIC METRICS TO ASSESS SERVICE COMPOSITION IS FAULT INCLINED OR HALE

V.SUJATHA[1], Dr.G.APPA RAO[2]

[1]Asst. Professor, Dept. of CSE, GITAM School of Technology, Hyderabad

[2]Professor, Dept. of CSE, GITAM University, Visakhapatnam

[1]Email:sujatha.varadi@yahoo.com

ABSTRACT

One of the considerable research objective of the current decade is developing applications by composing individual services for web and cloud environment. The majority of research in this context is aimed to optimize the service composition, which is in the raise quality. Many of existing models are heuristic scales defined by machine learning approaches. The constraint of these existing models is that they are centric to one or two de-facto factors of the quality of service such as response time and reliability. This practice of optimizing service composition is not robust and scalable in order to achieve the optimal service composition towards the raise of quality. The other significant issue of research is many of existing models are not linear towards computational complexity, which is due to the compliment of evolutions against the raise in available services count. In this con-text here we devised a two QoS metrics called Scale to Composition Fault Inclined and Scale to Composition Haleness, which enables to assess the services based on multiple number of QoS metrics and also should stabilize the computational complexity to linear. The experiment results are indicating the significance of the proposed model towards scalable and robust QoS-aware service composition.

Keywords: *Web Service Compositions, Composition Support, Service Composition Impact Scale, Service Descriptor Impact Scale, Web Service Composition Fault Proneness.*

1 INTRODUCTION

Web service composition in a service oriented computation discovers and connects several web services for a QoS based optimal solution of the services. The web services are web based software belonging to a service oriented architecture which may be immediately executed for a composite service delivery over the web. A service composition involves the function of first determining the tasks for the client provided business process. Next in the service composition suitable web services are searched for implementing the allotted tasks. The selected services are then combined into a loosely joined single component and the order of priority for executing the services is determined. Finally the web services based on the established priority are executed according to the service requirements.

The tasks implemented in the business process are first identified and these complex tasks are minimized of their complexities. The tasks are executed with a strategy of task ordering [1] to obtain a desired solution from the service. The selection of a service offering a specific task solution is difficult as there are multiple services available from different providers which offer for the same tasks similar or same solutions. An optimal search criteria used to find better web services should reduce the search space with the selection of specific services of optimal or near optimal solutions from a repository of services. The selection of a service over another is in terms of a strategy of ranking the services considering the services QoS parameters, and the service composition context which impacts the ranking of the services based on the priority given to a specific QoS factor. In the composition the available suitable independent web services are adapted and combined in a work flow and based on an order of priority are implements the tasks in terms of the

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

factors of QoS and the context of the composition for an optimal solution.

The optimality achievable from a service composition is defined in terms of Availability, Accessibility, cost, integrity, throughput, round-trip time, reliability, regulatory and Security factors. The optimality achieved differs for different service compositions with different contexts of compositions. For example a service independently may show good throughput but in a composition of services the same service could be a total failure in accommodating and performing in the composition [1][2][3][4][5]. Another example is a context of a service composition may place a higher demand for a QoS factor over another such as a cost optimized service composition where reliability is not the main criteria or a context where a balance of the QoS factors like cost, reliability may be required without compromising security.

The services involved in the previously implemented web service compositions would assist in the fault prone impact assessment before implementing newer web services. The existing State-of-the-art composition techniques [2] [3] [4] [5] [6] [7] [8] are error prone and lacking maturity for guaranteeing totally fault free operations. In the strategies if we may find error prone behavior subsequent to the deployment of these strategies it may result in unimportant solutions, further cost escalations, and severe vulnerabilities. All this makes estimation of the fault proneness in the composition scope necessary. Our approach devised in this paper is a statistical approach introduced for estimating the service composition to determine if it is fault inclined or hale. This proposed strategy is capable of assessing any current approach of web service composition.

2 ASSOCIATED WORK

The contemporary research work has given literature substantially in the strategies for service composition devised using the metaheuristic algorithm and based on QoS awareness. In these research studies towards devising services based on composite offerings the method by Yu et al. [4] uses a greedy strategy integrated to the features of QoS together with a strategy of applied adaptiveness. In least search time the approach effectively determines a solution. An approach for web service composition called WSMO (web service modeling ontology) in [5] by Xiangbing et al. is devised for a composition with QoS service offering using a genetic algorithm to determine an effective optimal solution in the least search time. A chaos PSO based approach introduced in [11] by Li et al is devised for selecting optimality based web services. An approach devised in [12] by Xiangwei et al. is introduced for composing web services with QoS using two strategies of, algorithm of discrete PSO strategy and CPN (color Petri nets). The CPN scheme models among candidate services the relations of multi-attribute, and the multi-constraint, and next the discrete PSO algorithmic approach implements sequentially legal firing to determine an optimal solution [13]. An approach which is a performance efficient composition of web services in [14] by Mao et al. is devised with three metaheuristics based algorithms (PSO, EDA, GA). In the method devised by Zhao et al. [15] has a QoS awareness devised web service composition based on enhancing the improved discrete PSO based immune optimization approach. In this devised approach the local best strategy is enhanced with the IO (immune optimization) algorithm whereas the PSO algorithm is used to determine the global optimization value and to minimize the search capability for a scalable model. The approach based on a hybrid strategy of GRASP and path re-linking algorithm in [16] by Parejo et al. is devised for web service composition offering QoS aware runtime quality optimized services.The above proposed models have considerable importance for their service quality delivered however the limitation of these approaches is that they restrict themselves to only a single or a couple of QoS factors. This hinders the accuracy of determining the scalable factor of the composition. Also in the service composition the associated complexity of computation is O(n2) where for every task the services associated increase in increments and in evaluating the services composition the related complexity becomes higher.

In the paper [9] the approach describes a collection of factors of QoS for predicting the possible services for a practical solution. The contemporary strategies for selecting services based on quality-awareness are mostly based on selecting only the best service from the various available services. An approach for finding service compositions which are optimal for a business process is devised in [8] is based on linear programming and on the linear combination of various QoS factors like, availability, execution rate success, response time, execution cost and reputation based solutions. A service composition based on the service factors associated temporal validity is proposed in. [6]. An approach based on the constraints locally as well as

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

globally in [10] uses a model of mixed integers based on linear programming. A services selection strategy models the complex rucksack problem of multi-choice multi-dimension in [7] by assigning to the services various levels of quality for the selection of a best service.

These approaches are mostly based on the user rating given to the models and its various parameters. The QoS solutions are defined by a ranking mechanism/ However the possibility of a scalable ordering of the different models with the ranking mechanism is however rather impossible. The strategies devised for a QoS web service composition are many where the presence of fault proneness is also the same in the composition of the services. Hence there is need for devising a model with a mechanism which evaluates the fault proliferation and resurgence especially in service compositions based on highly complex dynamically coupled architectures.

The models for benchmark service quality assessment are mostly either based on the attributes or dependent on the rating given by the users or further have both the strategies included in their approach. If we consider the attributes specific composition, the relevance of a composition to a specific set of attributes would differ when compared with the attributes associated with another composition. Also in a user ratings based composition the rating importance varies with contextual factors. Further most of the models of benchmarking assess the services by their performance independently. In reality however the performance of a service may influence the functionality of another service. The approach proposed by us is a statistics based impact scale estimation towards service composition under fault proneness. The approach is devised with a metric introduced and termed by us as composition support of service compositions and service descriptors. The heuristic metrics developed in this paper are based on our previously devised statistical assessment models which are strategies based on predictive accuracy and scalable factor called Web Service Composition Impact Scale towards Fault Proneness [12] and QoS metrics for robust service composition [13]. The heuristic metrics devised here in this paper are termed as, Scale to Composition Fault Inclined (SCFI) and Scale to Composition Haleness (SCH). The devised metaheuristic model is not dependent on specific QoS factors or quantity of factors. The model makes impact assessment of every service based on a measurement of their Composition Fault Inclined Scope in terms of multiple factors of QoS.

3 DEFINING A HEURISTIC SCALE TO COMPOSITION FAULT INCLINED (SCFI) AND SCALE TO COMPOSITION HALENESS (SCH)

3.1 Dataset Preprocessing

In the experiments performed the Dataset used here consists of 14 attributes (see table 1). Here every record consists of 14 attribute values which are continuous type and categorical values for a QOS based composition. In the earlier contributions [17] [18], we have provided detailed description of the attributes. For applying a process for optimizing the attribute values a mechanism introduced by us requires for a given dataset numeric and categorical attributes values. For this in this scheme first all the continuous values are transformed to categorical values.

Journal of Theoretical and Applied Information Technology
30th November 2016. Vol.93. No.2

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

Table1: Description Of Dataset Attributes

Attribute ID	Attribute of Complete Record	Description
1	Associability	Services from the single WSDL
2	Cyclic	Number of services in cyclic execution
3	Dependent	No of services dependent of other services
4	Parallel	No of services involved in parallel execution
5	Repetitive	No of execution attempts due to failure
6	Uptime	Service uninterrupted available time
7	Services count	No of services involved in composition
8	Diversity	Services from number of service providers
9	Roundtrip time	Composition execution time
10	Cost	Sum of the services cost involved in composition
11	Reliability	Service accuracy towards response
12	response time	Composition response time
13	versioning ratio	No of times composition is updated
14	Status	Label that indicates the composition is successful or failed

3.2 Attribute Optimization for Defining Scale to Composition Fault Inclined

Let partition the preprocessed set of records based on their labels, such that the records labeled as hale is one set, records labeled as fault inclined is other set. Consider the unique values of each attribute values set $f_i v(NRS)$ in the resultant records-set (NRS) with records labeled as hale and their coverage percentage as $f_i v = \{f_i(v_1, c_1), f_i(v_2, c_2), f_i(v_3, c_3), f_i(v_4, c_4), ..., f_i(v_j, c_j)\}$. Further the attribute optimization for fault inclined records is done as follows:

- Let consider the records set $rs(NRS)$ contains records those labeled as normal.
- Let $f_i(FPS)$ be the attribute f_i of FPS and $f_i(FPS)_{vs}$ be the set of values assigned to that attribute in FPS
- Create an empty set $\overline{f_i(NRS)_{vs}}$ of size $|f_i(FPS)_{vs}|$, then fill it with values from $f_i v(NRS)$ according to their coverage percentage such that $|f_i(FPS)_{vs}| \cong |\overline{f_i(NRS)_{vs}}|$.
- This process is opted to prepare the attribute values vector $\overline{f_i(NRS)_{vs}}$ of each attribute f_i of the NRS,
- This process should be applied for all attributes of the record-set and refer that resultant attributes with values as a set $\overline{NRS}$.
- The canonical correlation (see section 3.4) will be done further, which is between each attribute values set $f_i(FPS)_{vs}$ and $\overline{f_i(NRS)_{vs}}$ of FPS and $\overline{NRS}$ respectively.
- Further, the attributes of the FPS can be considered as optimal, which are having canonical correlation is less than given threshold or zero. Further we form a record set $OFPS$, which is having records with values of only attributes that are assessed as optimal through canonical correlation, and this record set $OFPS$ is used further to define the scale to Composition Fault Inclined (SCFI).

3.3 Attribute Optimization for Defining Scale to Composition Haleness

As like as the process explored in section 3.2, consider the unique values of each attribute values set $f_i v(FPS)$ in the resultant records-set (FPS) with records labeled as normal and their coverage percentage as

$$f_i v(FPS) = \{f_i(v_1, c_1), f_i(v_2, c_2), f_i(v_3, c_3), f_i(v_4, c_4), \\ ......, f_i(v_j, c_j)\}$$

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

. Further, the attribute optimization for Fault Inclined records is done as follows:

- Let consider the records set $rs(FPS)$ contains records those labeled as fault inclined.
- Let $f_i(NRS)$ be the attribute f_i of NRS and $f_i(NRS)_{vs}$ be the set of values assigned to that attribute in NRS
- Create an empty set $\overline{f_i(FPS)_{vs}}$ of size $|f_i(NRS)_{vs}|$, then fill it with values from $f_i v(FPS)$ according to their coverage percentage such that $|f_i(NRS)_{vs}| \cong |\overline{f_i(FPS)_{vs}}|$.
- This process is opted to prepare the attribute values vector $\overline{f_i(FPS)_{vs}}$ of each attribute f_i the FPS,
- This process should be applied for all attributes of the record-set and refer that resultant attributes with values as a set $\overline{FPS}$.
- The canonical correlation analysis (see section 3.4) will be done further, which is between each attribute values set $f_i(NRS)_{vs}$ and $\overline{f_i(FPS)_{vs}}$ of NRS and $\overline{FPS}$ respectively.

Further, the attributes of the NRS can be considered as optimal, which are having canonical correlation is less than given threshold or zero. Further we form a record set $ONRS$, which is having records with values of only attributes that are assessed as optimal through canonical correlation, and this record set $ONRS$ is used further to define the Scale to Composition Haleness (SCH).

3.4 Canonical Correlation Analysis

The multidimensional datasets X and Y are two data sets considered and the in-between linear relationships are established with the auto covariance's and cross-covariance's matrices of second-order with standard statistical technique CCA and the results offered are creditable and comprehendible results. The technique is based on finding 2 bases one each for the datasets 'X'and 'Y', where the 'X'and 'Y'datasets matrix of cross-correlation becomes diagonal whereas, correlations of the diagonal is maximized.

The parameters for implementing the canonical correlations in CCA are studied in the paper [19], [20] where, X and Y data vectors should be of equal number; however the data vectors $x \in X$ and $y \in Y$ may have varying dimensions assuming the mean is zero. The canonical correlations computation is solved using the equations of eigenvector.

$$C_{xx}^{-1}C_{xy}C_{yy}^{-1}C_{yx}w_x = \rho^2 w_x$$

$$C_{yy}^{-1}C_{yx}C_{xx}^{-1}C_{xy}w_y = \rho^2 w_y \quad (1)$$

Here $C_{yx} = E\{yx^T\}$ where ρ^2 or Eigen values are square of canonical correlations and 'w_x'and 'w_x' or the Eigen vectors are normalized CCA basis vectors. The solutions to the equations which are considered are those equivalents to non-zero value whose number is equivalent to x and y vectors lesser dimensional value.

The method followed in various ICA and BSS techniques is also used here where, x and y data vectors if prewhitened the solution (1) could be simplified [21]. Following the process of prewhitening, the canonical correlations C_{xx} and C_{yy} are both converted to unit matrices. As $C_{yx} = C_{xy}^T$, Eq (1) is converted to,

$$C_{xy}C_{xy}^T w_x = \rho^2 w_x$$

$$C_{yx}C_{yx}^T w_y = \rho^2 w_y \quad (2)$$

As these equations are however really equations depicting the singular value decomposition (SVD) [22] of the cross-covariance matrix C_{xy}:

$$C_{xy} = U\Sigma V^T = \sum_{i=1}^{L} \rho_i u_i v_i^T \quad (3)$$

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

Here U and V represent orthogonal square matrices $(U^T U = I,\ V^T V = I)$ comprising of u_i and v_i representing singular vectors. In our approach, the singular vectors considered above are w_{xi} and w_{yi} representing basis vectors delivering canonical correlations. The matrices U and V, and the subsequent u_i and v_i singular vectors dimensionalities usually vary according to the varied dimensions x and y data vectors. The pseudo diagonal matrix

$$\Sigma = \begin{bmatrix} D & 0 \\ 0 & 0 \end{bmatrix} \tag{4}$$

includes a diagonal matrix D comprising of singular values equal to non-zero and attached with zero matrices which makes the matrix Σ to be compatible with various dimensions of x and y. The non-zero singular values are basically the nonzero canonical correlations whose number is lesser than any of x and y data vectors dimensions if C_{xy} or the cross-covariance matrix has full rank.

3.5 Defining the Scale to Composition Fault Inclined (SCFI)

Let consider the compositions set $OFPS$ that formed due to canonical correlation analysis (see section 3.2).

Further, form a set $F(OFPS)$ such that

$$\begin{aligned} F(OFPS) = \{ & f_1(OFPS) = \{v_{11}, v_{12}, v_{13}, \ldots\ldots v_{1a}\}, \\ & f_2(OFPS) = \{v_{21}, v_{22}, v_{23}, \ldots\ldots v_{2b}\}, \\ & . \\ & . \\ & . \\ & . \\ & ., \\ & f_i(OFPS) = \{v_{i1}, v_{i2}, v_{i3}, \ldots\ldots v_{ic}\}\} \end{aligned}$$

Here in the above description $f_i(OFPS) = \{v_{i1}, v_{i2}, v_{i3}, \ldots\ldots\} \forall i = 1..n$ represents the optimal attribute f_i and the unique values $\{v_{i1}, v_{i2}, v_{i3}, \ldots\ldots\}$ assigned to that attribute of all the records in set $OFPS$.

Rank the each value v_{ij} of optimal attribute f_i, which is based on their coverage in the $f_i(OFPS)$.

Further represent each composition $\{r_i \forall i = 1\ldots|OFPS| \wedge r_i \in OFPS\}$ as a set $rs(r_i)$ with the respective rank of the value of each optimal attribute as follows:

$$\begin{aligned} r_i = \{ & f_1(v_j \forall j \in \{1..|f_1|\}), \\ & f_2(v_j \forall j \in \{1..|f_2|\}), \\ & f_3(v_j \forall j \in \{1..|f_3|\}), \\ & \ldots\ldots, \\ & f_i(v_j \forall j \in \{1..|f_i|\})\} \end{aligned}$$

$$rs(r_i) = \{r(f_1(v_j \forall j \in \{1..|f_1|\})), r(f_2(v_j \forall j \in \{1..|f_2|\})), r(f_3(v_j \forall j \in \{1..|f_3|\})), \ldots\ldots, r(f_i(v_j \forall j \in \{1..|f_i|\}))\}$$

Here in this description r_i is a record that belongs to the $OFPS$, which is representing the set of respective values of the optimal attributes. The representation $f_i(v_j \forall j \in \{1..|f_i|\})$ is the value v_j of optimal attribute f_i, and $|f_i|$ represents the size of all possible values to the attribute f_i. And the set $rs(r_i)$ is representing composition r_i by the respective ranks of the values of the optimal attributes. The representation $r(f_i(v_j \forall j \in \{1..|f_i|\}))$ is the rank of the value v_j of the attribute f_i.

Further, for each $rs(r_i)$, find the aggregate rank $ar(r_i)$ as follows, which is an average of ranks representing the respective values of the optimal attributes of the composition r_i

$$ar(r_i) = \frac{\sum_{i=1}^{n} r(f_i(v_j \forall j \in \{1\ldots|f_i|\})}{n}$$

The standards defined by ANOVA [23],

(i) The measure average reflects the centrality of the distribution, but not significant to consider it alone as representation of the distribution, since it is not considering the uniform distribution.

(ii) The standard deviation of these ranks represents the how they deviated from

Journal of Theoretical and Applied Information Technology
30th November 2016. Vol.93. No.2

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

each other, which is also not confirming the distribution status.

(iii) The kurtosis [23] represents the state of uniform distribution. If kurtosis found to be platy-kurtic (kurtosis value less than three), then it is representing the uniform distribution.

(iv) Henceforth, the distribution with platy-kurtic value is significant to consider as uniform distribution.

Henceforth, we measure the kurtosis of each distribution and order them by their kurtosis from minimal to maximal. The kurtosis of the ranks of each composition t_i is measured as follows:

$$\sigma_{ar(r_i)} = \sqrt{\frac{\sum_{i=1}^{n}(r(f_i(v_j \forall j \in \{1...|f_i|\}) - ar(r_i))^2}{n}}$$

$$m4 = \frac{\sum_{i=1}^{n}(r(f_i(v_j \forall j \in \{1...|f_i|\}) - ar(r_i))^4}{n}$$

$$g_{(r_i)} = \frac{m4}{\sigma_{ar(r_i)}}$$

Here in these equations $\sigma_{ar(r_i)}$ represents the variation observed between ranks of optimal attributes of a composition r_t and $g_{(r_i)}$ represents the kurtosis observed between the ranks of the optimal attributes of the composition r_i

Further we consider the compositions with platy-kurtic distribution of the ranks, and then mean of the ranks of these records will be considered as a scale to assess the composition fault Inclined.

$$\mu(OFPS) = \frac{\sum_{i=1}^{n} ar(r_i)}{n}$$

Here $\mu(OFPS)$ represents the mean of the aggregate ranks of n compositions of $OFPS$

$$SCFI = \frac{\sum_{i=1}^{m} ar(r_i)}{m}$$

Here in the above equation $SCFI$ represents the scale to composition fault Inclined, m represents the number of records with platy-kurtic rank distribution ($g_{(r_i)} < 3$) and having the rank greater than $\mu(OFPS)$.

The lower and upper bounds of the scale will be assessed as follows:

$$stdv_{OFPS} = \sqrt{\frac{\sum_{i=1}^{m}\left(ar(r_i) - SCFI\right)^2}{m-1}}$$

Here in above equation the standard deviation of the aggregate ranks of all record in $ODRS$ is measured

$$SCFI_{low} = SCFI - stdv_{OFPS}$$

$$SCFI_{upr} = SCFI + stdv_{OFPS}$$

3.6 Scale to Composition Haleness (SCH).

The scale that devised here in this section is aimed to assess the Haleness state of the composition. The Strategy that explored on compositions set $OFPS$ to define SCFI (see section 3.5) is also applied on compositions set $ONRS$ to devise Scale to Composition Haleness (SCH). The process applied on $ONRS$ is briefed here:

Form a set $F(ONRS)$ such that

$$\begin{aligned} F(ONRS) = \{ & f_1(ONRS) = \{v_{11}, v_{12}, v_{13}, ...v_{1a}\}, \\ & f_2(ONRS) = \{v_{21}, v_{22}, v_{23}, ...v_{2b}\}, \\ & . \\ & . \\ & . \\ & . \\ & ., \\ & f_i(ONRS) = \{v_{i1}, v_{i2}, v_{i3}, ...v_{ic}\}\} \end{aligned}$$

Here in the above description $f_i(ONRS) = \{v_{i1}, v_{i2}, v_{i3}, ...\} \forall i = 1..n$ represents the optimal attribute f_i and the unique values $\{v_{i1}, v_{i2}, v_{i3}, ...\}$ assigned to that attribute of all the records in set $ONRS$.

Rank the each value v_{ij} of optimal attribute f_i, which is based on their coverage in the $f_i(ONRS)$.

Further represent each composition $\{r_i \forall i = 1...|ONRS| \wedge r_i \in ONRS\}$ as a set $rs(r_i)$ with the respective rank of the value of each optimal attribute as follows:

Journal of Theoretical and Applied Information Technology
30th November 2016. Vol.93. No.2

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

$$\begin{aligned} r_i = \{ & f_1(v_j \forall j \in \{1..|f_1|\}), \\ & f_2(v_j \forall j \in \{1..|f_2|\}), \\ & f_3(v_j \forall j \in \{1..|f_3|\}), \\ & ...., \\ & f_i(v_j \forall j \in \{1..|f_i|\})\} \end{aligned}$$

$$\begin{aligned} rs(r_i) = \{ & r(f_1(v_j \forall j \in \{1..|f_1|\})), \\ & r(f_2(v_j \forall j \in \{1..|f_2|\})), \\ & r(f_3(v_j \forall j \in \{1..|f_3|\})), \\ & ...., \\ & r(f_i(v_j \forall j \in \{1..|f_i|\}))\} \end{aligned}$$

For each $rs(r_i)$, find the aggregate rank $ar(r_i)$ as follows, which is an average of ranks representing the respective values of the optimal attributes of the composition r_i

$$ar(r_i) = \frac{\sum_{i=1}^{n} r(f_i(v_j \forall j \in \{1...|f_i|\})}{n}$$

According to the ANOVA [23] standards (explored in section 3.5), we measure the kurtosis of each distribution and order them by their kurtosis from minimal to maximal. The kurtosis of the ranks of each composition t_i is measured as follows:

$$\sigma_{ar(r_i)} = \sqrt{\frac{\sum_{i=1}^{n} (r(f_i(v_j \forall j \in \{1...|f_i|\}) - ar(r_i))^2}{n}}$$

$$m4 = \frac{\sum_{i=1}^{n} (r(f_i(v_j \forall j \in \{1...|f_i|\}) - ar(r_i))^4}{n}$$

$$g_{(r_i)} = \frac{m4}{\sigma_{ar(r_i)}}$$

Here in these equations $\sigma_{ar(r_i)}$ represents the variation observed between ranks of optimal attributes of a composition r_t and $g_{(r_i)}$ represents the kurtosis observed between the ranks of the optimal attributes of the composition r_i

Further we consider the compositions with platy-kurtic distribution of the ranks, and then mean of the ranks of these records will be considered as a scale to assess the Composition Fault Inclined.

$$\mu(ONRS) = \frac{\sum_{i=1}^{n} ar(r_i)}{n}$$

Here $\mu(ONRS)$ represents the mean of the aggregate ranks of n compositions of $ONRS$

$$SCH = \frac{\sum_{i=1}^{m} ar(r_i)}{m}$$

Here in the above equation SCH represents the scale to Composition haleness, m represents the number of records with platy-kurtic rank distribution ($g_{(r_i)} < 3$) and having the aggregate rank greater than $\mu(ONRS)$.

The lower and upper bounds of the scale will be assessed as follows:

$$\overline{\mu} = \frac{\sum_{i=1}^{m} \{ar(r_i) \exists g_{(r_i)} < 3\}}{m}$$

The above equation is finding the mean $\overline{\mu}$ of the aggregate rank of the records with platy-kurtic feature rank distribution.

$$stdv_{ONRS} = \sqrt{\frac{\sum_{i=1}^{m} \left(ar(r_i) - \overline{\mu}\right)^2}{m-1}}$$

Here in above equation the standard deviation of the aggregate ranks of all record in $ONRS$ is measured

$$SCH_{low} = SCH - stdv_{ONRS}$$

$$SCH_{upr} = SCH + stdv_{ONRS}$$

The Scale to Composition Fault Inclined (SDP) and Scale to Composition Haleness (SCH) that are assessed from the given compositions set for training will be used further to assess the scope of Composition Fault Inclined or haleness of a given composition.

For a given composition mr to be tested, as explored in section 3.5,

Journal of Theoretical and Applied Information Technology
30th November 2016. Vol.93. No.2

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

- Preprocess the given composition and reform it as record mr_{SCFI} with only values of optimal attributes of the $SCFI$ and similarly reform it as record mr_{SCH} with only values of optimal attributes of the SCH.
- To assess the Fault Inclined of the composition mr_{SCFI}, form $rs(mr_{SCFI})$ with the respective rank of the value of each optimal attribute of $SCFI$.
- To assess the Composition Haleness of the composition mr_{SCH}, form $rs(mr_{SCH})$ with the respective rank of the value of each optimal attribute of SCH.
- Find the aggregate rank of the mr_{SCFI} as $ar(mr_{SCFI})$ and also find the aggregate rank of the mr_{SCH} as $ar(mr_{SCH})$

Further the state of given composition mr is assessed as follows:

Table 2 The Scale Of Composition Fault Inclined And Haleness And Assessing Strategy

$ar(mr_{SCFI}) \le SCFI_{low}$ && $ar(mr_{SCH}) \ge SCH_{upr}$	Haleness
$ar(mr_{SCFI}) \le SCFI$ && $ar(mr_{SCH}) \ge SCH_{upr}$	Haleness
$ar(mr_{SCFI}) \ge SCFI_{upr}$ && $ar(mr_{SCH}) \ge SCH_{upr}$	Inclined
$ar(mr_{SCFI}) \ge SCFI$ && $ar(mr_{SCH}) \ge SCH_{upr}$	Inclined
$ar(mr_{SCFI}) \le SCFI_{low}$ && $ar(mr_{SCH}) \ge SCH$	Haleness
$ar(mr_{SCFI}) \le SCFI$ && $ar(mr_{SCH}) \ge SCH$	Haleness
$ar(mr_{SCFI}) \ge SCFI_{upr}$ && $ar(mr_{SCH}) \ge SCH$	Inclined
$ar(mr_{SCFI}) \ge SCFI$ && $ar(mr_{SCH}) \ge SCH$	Inclined
$ar(mr_{SCFI}) \le SCFI_{low}$ && $ar(mr_{SCH}) \le SCH_{low}$	Inclined
$ar(mr_{SCFI}) \le SCFI$ && $ar(mr_{SCH}) \le SCH_{low}$	Inclined
$ar(mr_{SCFI}) \ge SCFI_{upr}$ && $ar(mr_{SCH}) \le SCH_{low}$	Inclined
$ar(mr_{SCFI}) \ge SCFI$ && $ar(mr_{SCH}) \le SCH_{low}$	Inclined
$ar(mr_{SCFI}) \le SCFI_{low}$ && $ar(mr_{SCH}) < SCH$	Haleness
$ar(mr_{SCFI}) \le SCFI$ && $ar(mr_{SCH}) < SCH$	Inclined
$ar(mr_{SCFI}) \ge SCFI_{upr}$ && $ar(mr_{SCH}) < SCH$	Inclined
$ar(mr_{SCFI}) \ge SCFI$ && $ar(mr_{SCH}) < SCH$	Inclined

Here in the table 2, all possible combinations of SCFI and SCH and the impact of those combinations explored. Regardless of the mr_{SCH}, if mr_{SCFI} is greater than $SCFI$ then the record confirmed to be Composition Fault Inclined. But in contrast, the Composition Haleness is dependent of $SCFI$, which is indicating that though the composition mr_{SCH} is more than the value of SCH, it's mr_{SCFI} must be less than the $SCFI$ to conclude that the given composition mr is scaled as Haleness. This may leads to slight increase in false positives in prediction but strictly avoids false negatives, which is an accuracy measurement of fault Inclined.

4 EXPERIMENTAL RESULTS AND PERFORMANCE ANALYSIS

The experiments were carried out on dataset that explored in section 3.1. Initially partitioned the processed dataset into normal and fault inclined records and then the optimal attributes of fault inclined compositions and normal compositions were traced out, which is by using the process explored (see section 3.2, 3.3, and 3.4). Further the scale to Composition Fault Inclined (SCFI) and Scale to Composition Haleness (SCH) were devised through the process explored (see section 3.5 and 3.6). The exploration of the input data and results were shown in table 3, 4 and 5. The visualization of the optimal scope of attributes for fault inclined and hale compositions can be found in fig 1 and fig 2.

Journal of Theoretical and Applied Information Technology
30th November 2016. Vol.93. No.2

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

Table 1 Statistics Of The Experiment Results

Total Number of Compositions	303
Range of QoS attributes of a composition	13
Compositions used for defining scale	80% (242 records)
Compositions used for performance analysis	20% (61 records)
Scale of Composition Fault Inclined $SCFI$ observed	7.11324
$SCFI_{low}$ observed	5.575204
$SCFI_{upr}$ observed	8.651277
Scale of Composition Haleness SCH observed	2.982372
SCH_{low} observed	2.35923
SCH_{upr} observed	3.605513

Table 2 Selected Qos Attributes Of The Compositions Labeled As Fault Inclined Under Different Canonical Correlation Threshold

<0.04		<0.05		<0.051(0.06) (mean of the CC		<0.1		<0.2	
Attribute ID	CC value	ID	CC	ID	CC	ID	CC	ID	CC
1	0.000398	1	0.000398	1	0.000398	1	0.000398	1	0.000398
7	0.021013	4	0.044147	3	0.050546	2	0.107357	2	0.107357
8	0.029505	7	0.021013	4	0.044147	3	0.050546	3	0.050546
9	0.03851	8	0.029505	5	0.050081	4	0.044147	4	0.044147
12	0.021347	9	0.03851	7	0.021013	5	0.050081	5	0.050081
		11	0.042488	8	0.029505	7	0.021013	6	0.103026
		12	0.021347	9	0.03851	8	0.029505	7	0.021013
				11	0.042488	9	0.03851	8	0.029505
				12	0.021347	10	0.098274	9	0.03851
						11	0.042488	10	0.098274
						12	0.021347	11	0.042488
						13	0.099685	12	0.021347
								13	0.099685

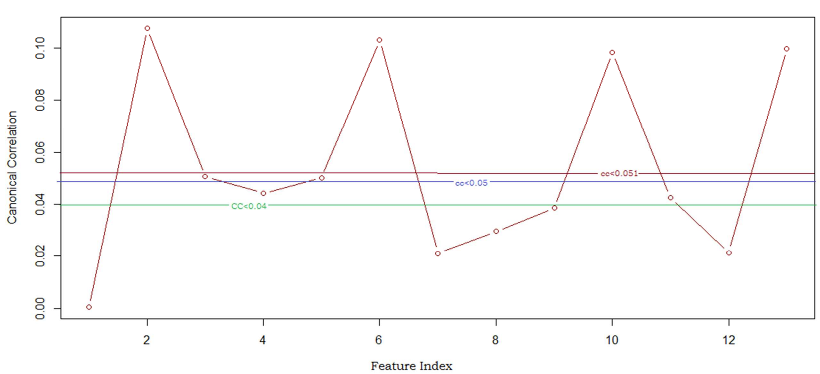

Figure 1: Qos Attributes Of The Compositions Labeled As Fault Inclined And Their Optimality Under Divergent CC Thresholds

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

Table 3 Selected Qos Attributes Of The Compositions Labeled As Hale Under Different Canonical Correlation Thresholds

<0.04		<0.06		<0.07585903 (mean of the cc vales) (0.08)		<0.1		<0.2 (yes, It's all)	
1	0.011762	1	0.011762	1	0.011762	1	0.011762	1	0.011762
5	0.004778	5	0.004778	2	0.069674	2	0.069674	2	0.069674
8	0.039421	8	0.039421	5	0.004778	3	0.082918	3	0.082918
11	0.021706	9	0.043843	8	0.039421	5	0.004778	4	0.190604
		11	0.021706	9	0.043843	8	0.039421	5	0.004778
		13	0.043684	11	0.021706	9	0.043843	6	0.177771
				12	0.066248	10	0.089583	7	0.144176
				13	0.043684	11	0.021706	8	0.039421
						12	0.066248	9	0.043843
						13	0.043684	10	0.089583
								11	0.021706
								12	0.066248
								13	0.043684

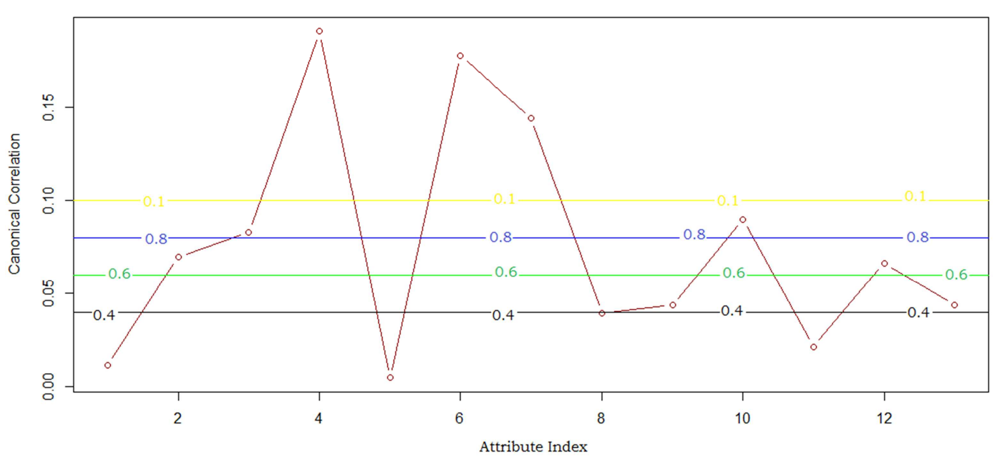

Figure 2: Qos Attributes Of The Compositions Labeled As Hale And Their Optimality Under Divergent CC Thresholds

4.1 Performance Analysis

The robustness and prediction accuracy of the scales *SCFI* and *SCH* are assessed through 62 records, which are of the combination of 40 fault inclined and 22 hale compositions.

The prediction statistics are as follow:

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

The count of true positives are (records predicted as truly fault inclined) 40, the count of true negatives are (records predicted as truly hale) 20, the count of false positives are (records predicted as falsely fault inclined) 2 and the count of false negatives are (the records predicted as falsely hale) 0.

Since the experimental results indicating that fault inclined composition prediction is 40 out of given 40 compositions, hence prediction error towards fault inclined compositions are 0.0. The prediction accuracy of normal compositions is observed as 20 among the given 22 compositions, hence the prediction error ratio is approx. 0.09, which is to be negligible as in the track of sensitive service composition needs, the composition actually fault inclined should not be diagnosed as hale, in contrast a hale composition can be suspected falsely as fault inclined and may recommend to further assessment strategies.

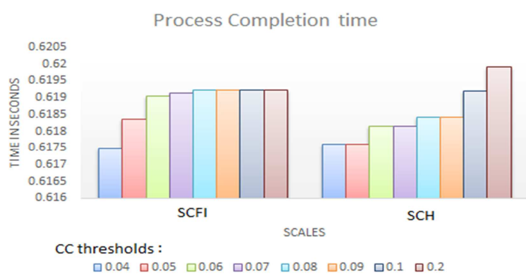

Figure 3: Process completion time of SCFI and SCH under divergent optimal attributes selected through various canonical correlation thresholds

The prediction accuracy of the model devised here in this manuscript is explored through statistical assessment metrics called precision, recall and f-measure (see table 5). The value obtained for metric recall indicating that the devised model is highly robust and scalable towards assessing fault prone scope, and the precision also indicating that the prediction accuracy of the model is high and approximately it is 97%.

Table 4: The precision, recall and the f-measure of the predictions

Precision	0.952381
Recall	1
F-Measure	0.97561

The process time of the application is stable since the increase in number of optimal attributes is not influencing the process complexity (see fig 3)

5 CONCLUSION

This manuscript introduced a novel heuristic scale to assess the fault Inclined scope of the given service composition. In regard to this, two heuristic metrics called Scale to Composition Fault Inclined (SCFI) and Scale to Composition Haleness (SCH) is devised. In contrast to the existing benchmarking models, the proposed metrics are assessing the composition fault Inclined scope and haleness referred as SCFI and SCH respectively. Further the combinations of these SCFI and SCH values of the given service compositions are used to assess the state of that composition. The process opted to devise these metrics is initially finding the optimal attributes of the given fault inclined and hale composition records that represented by QoS attributes, which is done through the canonical correlation analysis. Further the service compositions of fault inclined and hale with optimal attributes are used to assess the metrics SCFI and SCH. The experimental results are optimistic and concluding the prediction accuracy and robustness. The future work can be the definition of fuzzy model to estimate the combination of SCFI and SCH values.

REFERENCES:

[1] R. Anane, K.-M. Chao, Y. Li, Hybrid composition of web services and grid services,in: e-Technology, e-Commerce and e-Service, 2005. EEE '05. Proceedings. The 2005 IEEE Inter-national Conference on, 2005, pp. 426-431.

[2] J. El Hadad, M. Manouvrier, M. Rukoz, Tqos: Transactional and qos-aware selection algo-rithm for automatic web service composition, Services Computing, IEEE Transactions on 3 (1) (2010) 73-85.

[3] Strunk, Qos-aware service composition: A survey, in: Web Services (ECOWS), 2010 IEEE 8th European Conference on, 2010, pp. 67-74.

[4] Y. Yu, H. Ma, M. Zhang, An adaptive genetic programming approach to qos-aware web services composition, in: Evolutionary Computation (CEC), 2013 IEEE Congress on, 2013, pp. 1740-1747.

[5] Z. Xiangbing, M. Hongjiang, M. Fang, An optimal approach to the qos-based wsmo web service composition using genetic algorithm, in: Service-Oriented Computing - ICSOC 2012 Workshops, Vol. 7759, Springer Berlin Heidelberg, 2013, pp. 127-139.

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

[6] Ardagna, D., &Pernici, B. (2007). Adaptive service composition in flexible processes; IEEE Transactions on Software Engineering, 33(6), 369–384.
[7] Zheng, H., Zhao, W., Yang, J., &Bouguettaya, A. (2012). QoS analysis for web service compositions with complex structures. IEEE Transactions on Services Computing, 1939-1374, PP(99), 1. http://dx.doi.org/10.1109/TSC.2012.7
[8] Strunk, A. (2010). QoS-aware service composition: A Survey. In 2010 IEEE 8th European conference on web services (ECOWS) (pp. 67–74).
[9] Bonatti, P. A., Festa, P. (2005). On optimal service selection. In WWW '05: 14th interna-tional conference on world wide web (pp. 530–538)
[10] Ardagna, D., Pernici, B. (2005). Global and local QoS guarantee in web service selection. In BPM Workshops (pp. 32–46)
[11] W. Li, H. Yan-xiang, Web service composition based on qos with chaos particle swarm optimization, in: Wireless Communications Networking and Mobile Computing (WiCOM), 2010 6th International Conference on, 2010, pp. 1-4.
[12] L. Xiangwei, Z. Yin, Web service composition with global constraint based on discrete particle swarm optimization, in: Web Mining and Web-based Application, 2009. Second Pacific-Asia Conference on, 2009, pp. 183-186.
[13] W. Li, H. Yan-xiang, A web service composition algorithm based on global qos optimizing with mocaco, in: Algorithms and Architectures for Parallel Processing, Vol. 6082 of Lec-ture Notes in Computer Science, Springer Berlin Heidelberg, 2010, pp. 218-224.
[14] C. Mao, J. Chen, X. Yu, An empirical study on meta-heuristic search-based web service composition, in: e-Business Engineering (ICEBE), 2012 IEEE Ninth International Confer-ence on, 2012, pp. 117-122.
[15] X. Zhao, B. Song, P. Huang, Z. Wen, J. Weng, Y. Fan, An improved discrete immune opti-mization algorithm based on pso for qos-driven web service composition, Applied Soft Computing 12 (8) (2012) 2208-2216.
[16] Parejo, José Antonio, Sergio Segura, Pablo Fernandez, and Antonio Ruiz-Cortés. "QoS-aware web services composition using GRASP with Path Relinking." Expert Systems with Applications 41, no. 9 (2014): 4211-4223.
[17] Sujatha Varadi, G Appa Rao; "Quality of Service Centric Web Service Composition: Assessing Composition Impact Scale towards Fault Proneness"; Vol 14, No 9-C (2014): Global Journal of Computer Science and Technology
[18] V.Sujatha, Dr.G.Appa Rao and T. Tharun; "QoS Discrepancy Impact (qdi) and Cohesion Between Services (cbs): QoS metrics for robust service composition", ICMEET-2015, Department of Electronics and Communication Engineering, GITAM University, Visakhapatnam, Andhra Pradesh, India, proceedings in Springer LNEE series
[19] A. Rencher, Methods of Multivariate Analysis, 2nd ed., Wiley, 2002.
[20] M. Borga, "Canonical correlation: a tutorial", Linkoping University, Linkoping, Sweden, 2001, 12 pages. Available at http://www.imt.liu.se/~magnus/cca/tutorial/.
[21] A. Hyv¨arinen, J. Karhunen, and E. Oja, Independent Component Analysis. Wiley, 2001
[22] S. Haykin, Modern Filters. MacMillan, 1989.
[23] http://www.statsoft.com/textbook/anova-manova