

---

# Fuzzy Similarity Clustering for Consumer-Centric QoS-aware Selection of Web Services

Wei-Li Lin<sup>1</sup>, Chi-Chun Lo<sup>1</sup>, Kuo-Ming Chao<sup>2,3</sup>, Nick Godwin<sup>3</sup>

<sup>1</sup>*Institute of Information Management, National Chiao-Tung University, Hsin-Chu, 300, Taiwan*  
weililin.iim92@nctu.edu.tw, cclo@faculty.nctu.edu.tw  
tel: +886-3-5712121 ext.57401

fax: +996-3-5723792

<sup>2</sup>*Software School, Fudan University, China*

<sup>3</sup>*DSM Research Group, Department of Computing and the Digital Environment, Coventry University,*

*Coventry, CV1 5FB, UK*

[k.chao@coventry.ac.uk](mailto:k.chao@coventry.ac.uk), [a.n.godwin@coventry.ac.uk](mailto:a.n.godwin@coventry.ac.uk)

## ABSTRACT

*Appropriate use of group consensus on service consumers' QoS opinions can improve web service discovery. Web service participants with different backgrounds or preferences may not be easy to reach consensus on their Web service QoS opinions, so they should be treated as multi-groups. Also, to prevent to possible useful opinions from being omitted unintentionally, fuzzy clustering criteria for the "multi-groups" framework should be adopted. In this paper, we have proposed a FMGSAM (Fuzzy Multi-Groups based SAM) to improve the accuracy in group opinion similarity analysis and the efficiency in generating group consensus.*

## 1. Introduction

QoS (Quality of Service) has been considered as a significant criterion in the selection of web services recently [1][2][3][4][5][6]. In our previous work [1], we have designed a model of consumer-centric QoS-aware selection, QCMA (QoS Consensus Moderation Approach), to analyze the group consensus based on their fuzzy opinion similarity and QoS preference with a number of QoS (defined by W3C [1][2][7]). However, the challenges to the single group based opinion similarity and preference analysis for web service selection include:

1. If the fuzzy QoS opinions were collected from web service participants with very different backgrounds which could lead to diverse perceptions, the obtained consensus may not be effective.
2. For some outliers identified via our previous proposed approach which groups all users into one single group can be re-classified into other appropriate groups if multi-groups is adopted, as these opinions may have meaningful correlation. The omission of those outliers without further examination may not be appropriate.

In this paper, an attempt to use Multi-groups based consensus for web service selection to address the above

challenges is reported. If the groups contain highly similar fuzzy QoS opinions, the analysis on the preference order among QoS attributes will be more credible.

However, multi-attributes based clustering is much complicated than single-attribute based clustering. Haixun Wang et al [8] proposed pCluster model in 2003 for pattern similarity clustering in large data sets. The pCluster model can be used for clustering similar objects which were connected by shifting or scaling their relationships in multi-dimensional space. Their work considers that all attributes are equally important, so the weight distribution over the attributes was not discussed. In our applications, the attributes cannot be treated as "equal weight" due to different perceptions.

In 2005, M Fazeli et al [9] proposed a parallel algorithm to tackle multi-features data clustering problems in multi-computer with star topology. The raw data is depicted with a feature vector  $v$ , as a set of measurements  $(v_1, v_2, \dots, v_M)$  that map to properties of a collection of data into a Euclidean space of dimension  $M$ . The squared-error algorithm is taken for the multi-features data clustering and divides  $N$  multi-features data into  $K$  clusters. However, for some disqualified data which is very close to the "boundary" (lower than the required similarity threshold but still can be considered as "certain degree of similarity" in fuzzy concept) could be omitted and this leads to possible distorted classifications.

Therefore, the challenges for building multi-groups based QoS-aware selection of web service including:

1. Associated weight on each QoS attribute should be introduced due to different preference orders given by the users. Therefore, a weighted multi-attributes QoS similarity should be defined.
2. To prevent from not intentionally removing possible meaningful data which just falls outside of the pre-defined group boundaries, the multi-attributes based clustering criteria should be formulated with fuzzy evaluation.

FMG-SAM, an extension of SAM used in QCMA framework, attempts to provide an effective scheme for multi-groups based QoS clustering for web service selection. All the incoming multi-QoSs opinions will be fuzzily clustered into different QoS group. FMGSAM is the proposed mechanism to improve the insufficiency of SAM in the problem domains associated with multi-groups clustering with multi-attributes.

The following section in the paper including: Section 2 briefly describes the SAM in QCMA from our previous work [1]. Section 3 presents the proposed FMG-SAM. Section 4 reports on experimental results with a case study of multi-QoS attributes clustering for a number of hotel booking web services. Finally, Section 5 concludes this work and remarks on the future research direction.

## 2. Similarity Aggregation Method in QCMA

QCMA is employed to obtain and moderate group consensus on QoS in web services [1]. SAM was developed for resolving conflicts that arise from different opinions [10][11]. In SAM the different fuzzy opinions will be aggregated into consensus classes so that they can be measured by their similarities to each other. The procedure to perform SAM is organized into 7 steps described as below [12].

1. Each participant represents his/her subjective fuzzy preference on each specific criterion with a positive trapezoidal fuzzy number shown as Figure 1.

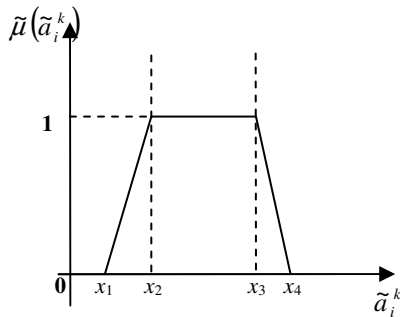


Figure 1: A Trapezoidal Fuzzy Number

Each fuzzy QoS opinion donated as  $wsa_{a_i}^k$ , which is a user's opinion  $k$  in the set of all opinions  $K$  on QoS attribute  $a_i$  which follows the definition in Figure 1.  $WSA_{a_i}$  is represented as a collection of  $wsa_{a_i}^k$  which is formally defined as follows.

$$wsa_{a_i}^k = ((x_1)_{a_i}^k, (x_2)_{a_i}^k, (x_3)_{a_i}^k, (x_4)_{a_i}^k) \quad (1)$$

$$0 \leq (x_1)_{a_i}^k \leq (x_2)_{a_i}^k \leq (x_3)_{a_i}^k \leq (x_4)_{a_i}^k \leq 10$$

$$WSA_{a_i} = \{ wsa_{a_i}^k \mid k \in K, a_i \in S_Q \}$$

$$S_Q = \{ a_1, a_2, a_3, \dots, a_{13} \}$$

where  $a_1 =$  "Performance",  $a_2 =$  "Reliability",  $a_3 =$  "Scalability",  $a_4 =$  "Capacity",  $a_5 =$  "Robustness",  $a_6 =$  "Exception Handling",  $a_7 =$  "Accuracy",  $a_8 =$  "Integrity",  $a_9 =$  "Accessibility",  $a_{10} =$  "Availability",  $a_{11} =$  "Interoperability",  $a_{12} =$  "Security" and  $a_{13} =$  "Friendliness".

2. The opinion similarity between  $wsa_{a_i}^j$  and  $wsa_{a_i}^k$ , which is denoted as  $Sim_{a_i}^{jk}$ , can be obtained via the following equation:

$$Sim_{a_i}^{jk} = \frac{\int (\min \{ \tilde{\mu}(wsa_{a_i}^j), \tilde{\mu}(wsa_{a_i}^k) \}) dx}{\int (\max \{ \tilde{\mu}(wsa_{a_i}^j), \tilde{\mu}(wsa_{a_i}^k) \}) dx} \quad (2)$$

3. To build an Agreement Matrix (AM), which can be represented as equation (3):

$$AM_{n \times n} = \begin{bmatrix} 1 & Sim_{12} & \dots & Sim_{1j} & \dots & Sim_{1n} \\ Sim_{21} & 1 & \vdots & \vdots & \vdots & \vdots \\ \vdots & \dots & 1 & \vdots & \vdots & \vdots \\ Sim_{i1} & \dots & \dots & 1 & \vdots & Sim_{in} \\ \vdots & \dots & \dots & \dots & 1 & \vdots \\ Sim_{n1} & Sim_{n2} & \dots & Sim_{nj} & \dots & 1 \end{bmatrix}_{n \times n} \quad (3)$$

4. To calculate an average agreement degree (AAD), denoted as  $A(wsa_{a_i}^k)$  is from equation (4):

$$A(wsa_{a_i}^k) = \frac{1}{n-1} \sum_{\substack{j=1 \\ k \neq j}}^n Sim_{kj} \quad (4)$$

5. To obtain the RAD (Relative Agreement Degree) for each individual opinion uses the following formula.

$$RAD(wsa_{a_i}^k) = \frac{A(wsa_{a_i}^k)}{\sum_{j=1}^n A(wsa_{a_i}^j)} \quad (5)$$

6. Assign a weighting variable,  $w_k$ , to each opinion.

7. Obtain the CDC (Consensus Degree Coefficient) for each participant:

$$CDC(wsa_{a_i}^k) = \beta \times w_k + (1 - \beta) \times RAD(wsa_{a_i}^k) \quad (6)$$

where  $\beta$  is a control variable to indicate the relation between CDC and RAD. To simplify the operation of CDC, we set  $\beta$  as zero so that CDC is equal to RAD.

#### 4. The Proposed Approach - FMGSAM

The proposed FMGSAM, derived from SAM, is designed for similarity analysis under multi-groups based QoS-aware selection of web service. The challenges on clustering multi-attributes based QoS are much higher than single dimensional clustering schemes due to the difficulty of identifying appropriate the number of groups and their sizes. In addition, the weight distribution among all QoS attributes must take into account. In order to increase grouping precision and to eliminate unnecessary data loss (outliers), a fuzzy group boundary is introduced.

The system scenario for multi-groups based similarity analysis can be depicted as Figure 2. A number of opinion groups will be formed and these groups have overlapping boundaries.

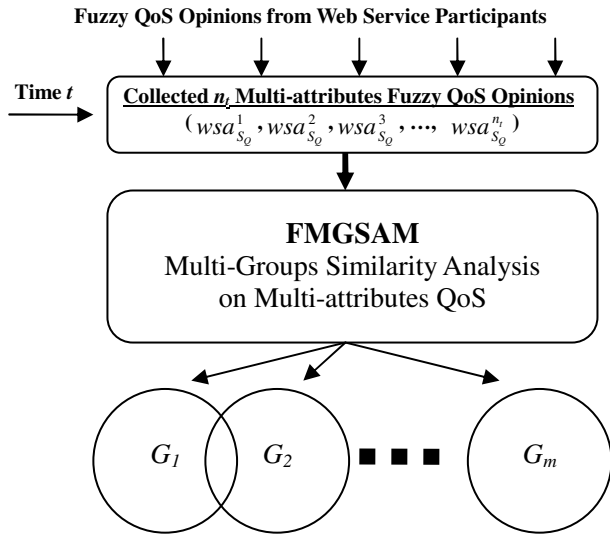


Figure 2: System Scenario for Multi-Groups Similarity

Before FMGSAM,  $n_t$  multi-attributes based fuzzy QoS collected at time  $t$  will be used for a preliminary analysis form a number of opinion groups. Then, the detailed processes of the FMGSAM can be carried which are illustrated as the following steps:

- (1) Represent Fuzzy QoS,  $wsa_{S_Q}^k$ , as a multi-dimensional fuzzy trapezoidal number as  $(wsa_{a_1}^k, \dots, wsa_{a_{13}}^k)$ .  $WSA_{S_Q}$  is a set of  $wsa_{S_Q}^k$  as below:

$$WSA_{S_Q} = \{wsa_{S_Q}^k | k \in K, a_i \in S_Q\} \quad (7)$$

- (2) Perform QoS fuzzy clustering which allocates all  $wsa_{S_Q}^k$  into appropriate groups  $(G_1, G_2, \dots, G_m)$ . For each pair of  $wsa_{S_Q}^j$  and  $wsa_{S_Q}^k$  in clustered  $G_p$ ,  $Sim_{S_Q}^{jk}$  can be obtained via following formulas:

$$Sim_{S_Q}^{jk} = (so_{a_1}^{jk} \times Sim_{a_1}^{jk}, \dots, so_{a_{13}}^{jk} \times Sim_{a_{13}}^{jk}) \quad (8)$$

where  $Sim_{a_i}^{jk}$  is similarity between  $wsa_{a_i}^j$  and  $wsa_{a_i}^k$ ,  $so_{a_i}^{jk}$  is the similarity in preference order between  $wsa_{a_i}^j$  and  $wsa_{a_i}^k$ ,  $q$  indicates the number of QoS attributes. By definition in W3C [2],  $q = 13$ .  $Sim_{a_i}^{jk}$  and  $so_{a_i}^{jk}$  can be obtained by following equations:

$$Sim_{a_i}^{jk} = \frac{\min\left\{\left(x_1\right)_{a_i}^j + \int \tilde{\mu}(wsa_{a_i}^j)dx, \left(x_1\right)_{a_i}^k + \int \tilde{\mu}(wsa_{a_i}^k)dx\right\}}{\max\left\{\left(x_1\right)_{a_i}^j + \int \tilde{\mu}(wsa_{a_i}^j)dx, \left(x_1\right)_{a_i}^k + \int \tilde{\mu}(wsa_{a_i}^k)dx\right\}} \quad (9)$$

$$so_{a_i}^{jk} = \frac{q - |o_{a_i}^j - o_{a_i}^k|}{q}$$

- (3) Determine Agreement Matrixes  $(AM_p)_{n_p \times n_p}$  for each clustered fuzzy QoS group  $G_p$ . The corresponding agreement matrixes can be represented as follows:

$$(AM_1) = \begin{bmatrix} 1 & Sim_{S_Q}^{12} & \dots & Sim_{S_Q}^{1j} & \dots & Sim_{S_Q}^{1n_1} \\ Sim_{S_Q}^{21} & 1 & \vdots & \vdots & \vdots & \vdots \\ \vdots & \dots & 1 & \vdots & \vdots & \vdots \\ Sim_{S_Q}^{j1} & \dots & \dots & 1 & \vdots & Sim_{S_Q}^{jn_1} \\ \vdots & \dots & \dots & \dots & 1 & \vdots \\ Sim_{S_Q}^{n_1 1} & Sim_{S_Q}^{n_1 2} & \dots & Sim_{S_Q}^{n_1 j} & \dots & 1 \end{bmatrix}_{n_1 \times n_1}$$

$$(AM_2) = \begin{bmatrix} 1 & Sim_{S_Q}^{12} & \dots & Sim_{S_Q}^{1j} & \dots & Sim_{S_Q}^{1n_2} \\ Sim_{S_Q}^{21} & 1 & \vdots & \vdots & \vdots & \vdots \\ \vdots & \dots & 1 & \vdots & \vdots & \vdots \\ Sim_{S_Q}^{j1} & \dots & \dots & 1 & \vdots & Sim_{S_Q}^{jn_2} \\ \vdots & \dots & \dots & \dots & 1 & \vdots \\ Sim_{S_Q}^{n_2 1} & Sim_{S_Q}^{n_2 2} & \dots & Sim_{S_Q}^{n_2 j} & \dots & 1 \end{bmatrix}_{n_2 \times n_2} \quad (10)$$

.....

$$(AM_m) = \begin{bmatrix} 1 & Sim_{S_Q}^{12} & \dots & Sim_{S_Q}^{1j} & \dots & Sim_{S_Q}^{1n_m} \\ Sim_{S_Q}^{21} & 1 & \vdots & \vdots & \vdots & \vdots \\ \vdots & \dots & 1 & \vdots & \vdots & \vdots \\ Sim_{S_Q}^{j1} & \dots & \dots & 1 & \vdots & Sim_{S_Q}^{jn_m} \\ \vdots & \dots & \dots & \dots & 1 & \vdots \\ Sim_{S_Q}^{n_m 1} & Sim_{S_Q}^{n_m 2} & \dots & Sim_{S_Q}^{n_m j} & \dots & 1 \end{bmatrix}_{n_m \times n_m}$$

(4) Determine the Average Agreement Degree for each  $wsa_{S_Q}^{G_p(k)}$  (the  $k^{\text{th}}$  opinion in  $G_p$ ) as below:

$$A(wsa_{S_Q}^{G_p(k)}) = \frac{1}{n_p - 1} \sum_{\substack{j=1 \\ j \neq G_p(k)}}^{n_p} Sim_{S_Q}^{j, G_p(k)} \quad (11)$$

$1 \leq G_p(k), j \leq n_p$

(5) Determine the Relative Agreement Degree for each  $wsa_{S_Q}^{G_p(k)}$  in group  $G_p$  as below:

$$RAD(wsa_{S_Q}^{G_p(k)}) = \frac{A(wsa_{S_Q}^{G_p(k)})}{\sum_{j=1}^{n_p} A(wsa_{S_Q}^{G_p(j)})} \quad (12)$$

$G_p(k) \neq G_p(j)$

(6) Determine the Consensus Degree Coefficient for each  $wsa_{S_Q}^{G_p(k)}$  in group  $G_p$  with assigned  $w_{G_p(k)}$  and  $\beta$  as follows:

$$CDC(wsa_{S_Q}^{G_p(k)}) = \beta \times w_{G_p(k)} + (1 - \beta) \times RAD(wsa_{S_Q}^{G_p(k)}) \quad (13)$$

The auxiliary algorithm for do fuzzy comparison between  $Sim_{S_Q}^{jk}$  and  $\tilde{d}_{S_Q}$  is defined in algorithm Fuzzy\_Comparison( $Sim_{S_Q}^{jk}, \tilde{d}_{S_Q}, m\_operator$ ).  $\tilde{d}_{S_Q}$  is set as similarity distance for fuzzy clustering which is represented as  $\tilde{d}_{S_Q} = \{\tilde{d}_{a_1}, \tilde{d}_{a_2}, \dots, \tilde{d}_{a_{13}}\}$ .  $\tilde{d}_{a_i}$  is defined as fuzzy interval which is represented as  $(d_{a_i}^l, d_{a_i}^u)$ ,  $0 \leq d_{a_i}^l < d_{a_i}^u \leq 1$ . The algorithm Fuzzy\_Comparison( $Sim_{S_Q}^{jk}, \tilde{d}_{S_Q}, m\_operator$ ) is illustrated as follows:

**Algorithm Fuzzy\_Comparison( $Sim_{S_Q}^{jk}, \tilde{d}_{S_Q}, m\_operator$ )**

1.  $m\_credit \leftarrow 0$ ; /\*  $m\_credit$ : indicator for the fuzzy comparison
2.  $m\_result \leftarrow \text{null}$ ; /\* indicate if the result of comparison will be E\_Large, E\_Equal or E\_Less
3. for  $i = 1$  to 13 /\* Number of QoS attributes
4. if  $(so_{a_i}^{jk} \times Sim_{a_i}^{jk}) > d_{a_i}^u$
5.  $m\_credit \leftarrow m\_credit + |so_{a_i}^{jk} \times Sim_{a_i}^{jk} - d_{a_i}^u|$

6. end if  $(so_{a_i}^{jk} \times Sim_{a_i}^{jk}) > \tilde{d}_{a_i}$
7. if  $(so_{a_i}^{jk} \times Sim_{a_i}^{jk}) < d_{a_i}^l$
8.  $m\_credit \leftarrow m\_credit - |so_{a_i}^{jk} \times Sim_{a_i}^{jk} - d_{a_i}^l|$
9. end if  $(so_{a_i}^{jk} \times Sim_{a_i}^{jk}) < d_{a_i}^l$
10. end for  $i = 1$  to 13
11. if  $m\_credit > 0.1$  then /\* Fuzzily larger than
12.  $m\_result \leftarrow E\_Larger$
13. elseif  $(-0.1 \leq m\_credit \leq 0.1)$
14.  $m\_result \leftarrow E\_Equal$
15. else /\* Fuzzily less than
16.  $m\_result \leftarrow E\_Less$
17. end if  $m\_credit > 1$
18. if  $m\_operator = "\tilde{\geq}"$  then
19. if  $(m\_result = E\_Large)$  or  $(m\_result = E\_Equal)$  then
20. return (true)
21. else
22. return (false)
23. endif
24. endif  $m\_operator = "\tilde{\geq}"$
25. if  $m\_operator = "\tilde{>}"$  then
26. if  $(m\_result = E\_Large)$  then
27. return (true)
28. else
29. return (false)
30. endif
31. endif  $m\_operator = "\tilde{>}"$
32. End Algorithm Fuzzy\_Comparison

## 5. Experiment

This section describes a simplified experiment using FMGSAM to conduct the similarity analysis of hotel book web services which include 5 QoS opinions on 3 different QoS attributes. The procedures are illustrated as follows:

1. Assume 5 opinions  $wsa_{S_Q}^1, wsa_{S_Q}^2, wsa_{S_Q}^3, wsa_{S_Q}^4$ , and  $wsa_{S_Q}^5$  in the set  $WSA_{S_Q}$ . Each  $wsa_{S_Q}^k$  ( $1 \leq k \leq 5$ ) represents three opinions on different QoS attributes ( $wsa_{a_1}^k, wsa_{a_2}^k$  and  $wsa_{a_3}^k$  and  $S_Q = \{a_1, a_2, a_3\}$ ). Each opinion  $wsa_{a_i}^k$  is represented as fuzzy trapezoidal numbers defined in (1). The 5 opinions collected in  $WSA_{S_Q}$  are defined as below:

$$wsa_{S_Q}^1 = \{(4,5,6,7), (5,6,7,8), (3,4,5,6)\};$$

$$wsa_{S_Q}^2 = \{(6,7,8,9), (7,8,9,10), (6,7,9,10)\};$$

$$wsa_{S_Q}^3 = \{(3,4,5,6), (2,3,4,5), (3,4,5,6)\};$$

$$wsa_{S_Q}^4 = \{(2,3,4,5), (6,7,8,9), (3,4,5,6)\};$$

$$wsa_{s_o}^5 = \{(5,6,7,8), (7,8,9,10), (6,7,8,9)\};$$

2. The relative position of preference order denoted as  $o_{a_i}^k$  for each  $wsa_{s_o}^k$  can be defined as follows:

$$(o_{a_1}^1, o_{a_2}^1, o_{a_{13}}^1) \text{ for } wsa_{s_o}^1 = (1, 3, 2)$$

$$(o_{a_1}^2, o_{a_2}^2, o_{a_{13}}^2) \text{ for } wsa_{s_o}^2 = (3, 1, 2)$$

$$(o_{a_1}^3, o_{a_2}^3, o_{a_{13}}^3) \text{ for } wsa_{s_o}^3 = (1, 2, 3)$$

$$(o_{a_1}^4, o_{a_2}^4, o_{a_{13}}^4) \text{ for } wsa_{s_o}^4 = (2, 1, 3)$$

$$(o_{a_1}^5, o_{a_2}^5, o_{a_{13}}^5) \text{ for } wsa_{s_o}^5 = (3, 2, 1)$$

3. For opinions clustering, the  $\tilde{d}_{s_o}$  is initialized as:

$$\tilde{d}_{s_o} = \{(0.4, 0.6), (0.4, 0.6), (0.4, 0.6)\}$$

4. Let  $wsa_{s_o}^1$  be selected as the first group centre. The distance of  $wsa_{s_o}^1$  to the rest opinions, similarity  $Sim_{s_o}^{1j}$  ( $2 \leq j \leq 5$ ), will be evaluated via equation (8) which obtains the following values:

$$Sim_{s_o}^{12} = (0.75, 0.26, 0.18); \quad Sim_{s_o}^{13} = (0.83, 0.38, 0.67)$$

$$Sim_{s_o}^{14} = (0.45, 0.29, 0.67); \quad Sim_{s_o}^{15} = (0.28, 0.52, 0.42)$$

5. Select  $wsa_{s_o}^j$  as an input for fuzzy comparison  $Sim_{s_o}^{1j}$  and use  $\tilde{d}_{s_o} ((0.4, 0.6), (0.4, 0.6), (0.4, 0.6))$ , which is defined in algorithm Fuzzy\_Comparison as threshold to determine the group membership, if  $Sim_{s_o}^{1j} \gtrsim \tilde{d}_{s_o}$  is satisfied.

$$Sim_{s_o}^{12} \lesssim \tilde{d}_{s_o} \text{ due to } m\_credit = -0.21 < -0.1$$

$$Sim_{s_o}^{13} \gtrsim \tilde{d}_{s_o} \text{ due to } m\_credit = 0.28 > 0.1$$

$$Sim_{s_o}^{14} \cong \tilde{d}_{s_o} \text{ due to } -0.1 \cong m\_credit \cong 0.1$$

$$Sim_{s_o}^{15} \lesssim \tilde{d}_{s_o} \text{ due to } m\_credit = -0.12 < -0.1$$

- In this case,  $wsa_{s_o}^3$  and  $wsa_{s_o}^4$  are clustered into the first group but  $wsa_{s_o}^4$  can also be classified into the other group due to  $Sim_{s_o}^{14} \cong \tilde{d}_{s_o}$  (so that  $wsa_{s_o}^4$  is still reserved in pool of opinions that are to be clustered).

$$\rightarrow G_1 = \{ wsa_{s_o}^1, wsa_{s_o}^3, wsa_{s_o}^4 \}$$

6. Similar to the generation of first group started from

step 4, the  $wsa_{s_o}^2$  is selected as the 2<sup>nd</sup> group centre. The rest opinions will be evaluated using similarity  $Sim_{s_o}^{2j}$  ( $j = 4, 5$ ) that is equation (8) to examine if they could belong to the 2<sup>nd</sup> group:

$$Sim_{s_o}^{24} = (0.34, 0.89, 0.38); \quad Sim_{s_o}^{25} = (0.88, 0.67, 0.60)$$

7. Similar to Step 5, the Fuzzy\_Comparison algorithms is employed and the same threshold values  $\tilde{d}_{s_o}$  are adopted to determine their m\_credit values.

$$Sim_{s_o}^{24} \gtrsim \tilde{d}_{s_o} \text{ due to } m\_credit = 0.21 > 0.1$$

$$Sim_{s_o}^{25} \gtrsim \tilde{d}_{s_o} \text{ due to } m\_credit = 0.35 > 0.1$$

- As result,  $wsa_{s_o}^4$  and  $wsa_{s_o}^5$  are qualified for the 2<sup>nd</sup> group.

- So,  $G_2 = \{ wsa_{s_o}^2, wsa_{s_o}^4, wsa_{s_o}^5 \}$  is obtained.

8. According to the above steps, we can get  $G_1 = \{ wsa_{s_o}^1, wsa_{s_o}^3, wsa_{s_o}^4 \}$  and  $G_2 = \{ wsa_{s_o}^2, wsa_{s_o}^4, wsa_{s_o}^5 \}$ . The following figure is a diagram representation of 2 different opinion groups and their memberships.

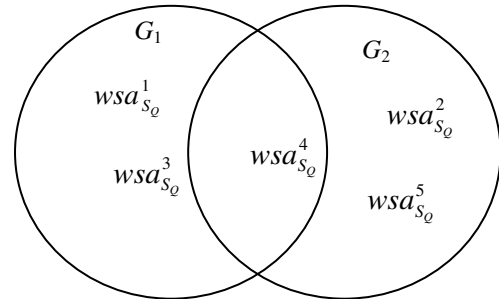


Figure 3: The Clustering of 5 Opinions

According to the above illustration, we can conclude FMGSAM is better than SAM in terms of similarity and efficiency:

1. Regarding similarity, the lowest similarity obtained by the SAM is  $Sim_{s_o}^{35}$  which is  $Sim_{s_o}^{35} = (0.23, 0.44, 0.21)$ . If we use m\_credit in fuzzy comparison algorithms as criterion,  $Sim_{s_o}^{35}$  is -0.36. However, for FMGSAM, the lowest similarity is  $Sim_{s_o}^{45}$  in  $G_2$ ,  $Sim_{s_o}^{45} = (0.38, 0.60, 0.21)$ , and its corresponding m\_credit is -0.21. The improvement of similarity analysis can be obtained via the following calculation:

$$\begin{aligned} Sim_{S_0}^{45} / Sim_{S_0}^{35} &= (0.38, 0.60, 0.21) / (0.23, 0.44, 0.21) \\ &= (1.65, 1.37, 1) \end{aligned}$$

So, the improvement for  $Sim_{S_0}^{45} / Sim_{S_0}^{35}$  can be is  $(1.65 + 1.37 + 1) / 3 = 1.34$ . It means that FMGSAM is more accurate than SAM by 34% in this case.

- Regarding efficiency, the AM generation involves the most complex process in these approaches. For SAM, the complexity of AM generation will be  $O(n^2)$ . However, the complexity of AM generation for FMGSAM for  $G_1$  or  $G_2$ , is  $O(n_1^2 + n_2^2) < O(n^2)$ . We can gain the following efficiency improvement:

$$Efficiency(FMGSAM/SAM) = O(n_1^2 + n_2^2) / O(n^2)$$

$$\therefore n^2 = 5, n_1 = n_2 = 3$$

$$\begin{aligned} \therefore Efficiency(FMGSAM/SAM) &= O(n_1^2 + n_2^2) / O(n^2) = (9 + 9) / 25 = 18/25 \\ &= 0.72 \end{aligned}$$

→ The improvement in the instance is 28%.

According to the above evaluation, we can conclude that FMGSAM is better than SAM.

## 7. Conclusion

This paper reported our proposed clustering scheme for web service consumer QoS opinions. The proposed FMGSAM possesses the following points:

- The similarity analysis for multi-attributes based QoS defined by W3C [2] can be performed via multi-attributes based clustering by FMGSAM. The different weight among QoS attributes and similarity for each individual QoS attribute has been considered.
- The FMGSAM can conduct higher precision on multi-groups opinions clustering analysis which includes multi-attributes QoS and the users from different backgrounds.
- The FMGSAM also achieve higher efficiency under multi-groups framework. The improvement in efficiency can be formulated as  $(1 - O((n_1^2 + n_2^2) / n^2))$ .

The future work will focus on development of the whole model of QoS-aware selection of web service, which including multi-groups based fuzzy similarity and preference order analysis.

## Acknowledgement

The research leading to these results has received funding from the European Community's Seventh Framework Programme FP7/2007-2013 under grant agreement n° No224609.

## References

- [1] Wei-Li Lin, Chi-Chun Lo, Kuo-Ming Chao, Muhammad Younas, *Consumer-Centric QoS-aware Selection of Web Service*, Journal of Computer and Systems Science, Elsevier Science, 74 (2008), pp. 211-231
- [2] W3C, *QoS for Web Service: Requirements and Possible Approaches*, W3C Working Group Note 25 November, <http://www.w3c.or.kr/kr-office/TR/2003>, 2003
- [3] Kuo-Ming Chao, Muhammad Younas, Chi-Chun Lo, Tao-Hsin Tan: *Fuzzy Matchmaking for Web Service*, Proceedings of 19 IEEE Conference on Advanced Network and Information Application, IEEE CS, pp 721-726, 2005
- [4] Menasce, 2004 D.A. Menasce', *Composing web services: A QoS view*, IEEE Internet Computing 8 (2004) (6), pp. 88-90.
- [5] Michael C. Jaeger, Gregor Rojec-Goldmann, and Gero Mühl. *Qos aggregation for service composition using workflow patterns*, In Proceedings of the 8th International Enterprise Distributed Object Computing Conference (EDOC 2004), page 149-159, Monterey, California, USA, September. 2004. IEEE CS Press.
- [6] Yutu Liu, Anne H.H. Ngu, Liangzhao Zeng *QoS Computation and Policing in Dynamic Web Service Selection* ACM WWW Conference, May 2004, New York, USA, pp. 66-73
- [7] E Herrera-Viedma, F. Herrera, and F. Chiclana: *A Consensus Model for Multiperson Decision Making With Different Preference Structures*. IEEE Transactions on Systems, Man and Cybernetics, IEEE, Vol.32.394-402, 2002
- [8] Haixun Wang, Wei Wang, Joing Yang, Philip S. Yu, *Clustering by Pattern Similarity in Large Data Sets*, Data Mining, Third IEEE International Conference on, pp. 187-194, 2003
- [9] M. Fazeli, H. Sarbazi-Azad, R. Farivar, *Parallel Clustering on the Star Graph*, Lecture Notes on Computer Science, Vol. 3719/2005, pp 287-292, 2005
- [10] F. Chiclana, F Herrera, and E. Herrera-Viedma: *Integrating Multiplicative Preference Relations in A Multipurpose Decision-making Model Based on Fuzzy Preference Relations*. Fuzzy Sets and Systems, Elsevier Science, Vol. 122. 277-291, 2001
- [11] F. Chiclana, F. Herrera, E. Herrera-Viedma: *A Classification Method of Alternatives for Multiple Preference Ordering Criteria Based on Fuzzy Majority*. Journal of Fuzzy Math., Vol. 34, 224-229, 1996
- [12] C-L Huang, K-M Chao, C-C Lo: *A Moderated Fuzzy Matchmaking for Web Services*, Proceedings of The 5th International Conference on Computer and Information Technology, Shanghai, China, IEEE CS, 2005, pp 1116-1122.