Knowledge Management for Registered Nurse's Concepts and Clustering with Common Mahalanobis Distances

Jui-Chiao Tseng[1]
[1]Department of Long-Term Care and Health Promotion Education, Min-Hwei Junior College of Health Care Management, Tai-Nan 736, Taiwan.

Abstract- Knowledge management of Registered Nurse's concepts was essential in educational environment. The purpose of this study is to provide an integrated method of fuzzy theory basis for individualized concept structure analysis. This method integrates Fuzzy Logic Model of Perception (FLMP) and Interpretive Structural Modeling (ISM). The combined algorithm could analyze individualized concepts structure based on the comparisons with concept structure of expert. Apply interpretive structural modeling to construct knowledge structure of Registered Nurse's. The results show that there are six clusters and each cluster has its own cognitive characteristics. The methodology can improve knowledge management in classroom more feasible.
Keywords –Mahalanobis Distances, Fuzzy Logic Model of Perception, ISM, Concept Structure.

I. INTRODUCTION

Zadeh developed fuzzy theory and it flourishes methodologies in many fields [1]. One of these fields is cognition diagnosis and it help represent knowledge structure [2, 4]. In addition, cognitive diagnosis is essential for educational environment. As to cognitive diagnosis, clustering technique is useful to classify students and then features of concept structures from each cluster could reveal constructive information for cognitive diagnosis. Therefore, remedial instruction will be more feasible [5]. For the feasibility of remedial instruction based on the cognition diagnosis, clustering method is needed so that students within the same cluster own similar knowledge structures and students among different clusters have the most variance on knowledge structures [6-9].It is a common viewpoint that human knowledge is stored in the form of structural relationship among concepts and their subordinate relationship is fuzzy, not crisp. There are some methodologies for concept structure analysis but little is known about methodologies of individualized concept structure [3]. Therefore, the development for methodology of individualized concept structure is an important issue and it is essential for cognition diagnosis and pedagogy. In this study, the integrated method of individualized concept structure based on fuzzy logic model of perception (FLMP) and interpretive structural modeling (ISM) will be developed. An example of empirical test data of linear algebra concept for students of learning deficiencies will also be analyzed and discussed. For the feasibility of remedial instruction based on the cognition diagnosis, clustering method is needed so that students within the same cluster own similar knowledge structures and students among different clusters have the most variance on knowledge structures [10-14].

We know Gustafson-Kessel clustering algorithm and Gath-Geva clustering algorithm, were developed to detect non-spherical structural clusters, but both of them based on semi-supervised Mahalanobis distance, these two algorithms fail to consider the relationships between cluster centers in the objective function, needing additional prior information. Added a regulating factor of covariance matrix, , to each class in objective function, the fuzzy covariance matrices in the Mahalanobis distance can be directly derived by minimizing the objective function, but the clustering results of this algorithm is still not stable enough. For improving the stability of the clustering results, we replace all of the covariance matrices with the same common covariance matrix in the objective function in the FCM-M algorithm, and then, an improve fuzzy clustering method, called the Fuzzy C-Means algorithm based on common Mahalanobis distance (FCM-CM), is proposed[17-22].

II. PROPOSED ALGORITHM

2.1 Method of Fuzzy Approach on Concept Structure Analysis –

The integrated algorithms consist of three steps of algorithms, AMC, ASC and AFISM. Suppose there be M ($m=1,2,\cdots,M$) items which measures A ($a=1,2,\cdots,A$) concepts. There are N task-takers ($n=1,2,\cdots,N$) who take this test. The response data matrix is denoted by $\mathbf{X}=(x_{nm})_{N\times M}$. $x_{nm}=1$ means task-taker n gives correct answer on item m and $x_{nm}=0$ means it doesn't. The item attribute matrix is denoted by $\mathbf{Y}=(y_{ma})_{M\times A}$. $y_{ma}=1$ means item m exactly measures concept a while $y_{ma}=0$ means it doesn't. These A concepts construct 2^A ideal concept vectors, which are denoted by $\mathbf{z}_i=(z_{ia})_{1\times A}=(z_{i1},\ z_{i2},\ \cdots,\ z_{iA})$, $i=1,2,\cdots,I$ and $I=2^A$. Each ideal concept vector expresses one certain kind of concept structure. $z_{ia}=1$ means the ideal concept vector $\mathbf{z}_i$ contains concept a while $z_{ia}=0$ means it doesn't. These ideal concept vectors construct a matrix $\mathbf{Z}=(z_{ia})_{I\times A}$ .Only the two matrices $\mathbf{X}=(x_{nm})_{N\times M}$ and $\mathbf{Y}=(y_{ma})_{M\times A}$ are known already. The following subsections of algorithms, which are AMC, ASC and AFISM, describe the steps to analyze individualized concept structures [15-16].

An example of empirical test data of Registered Nurse's concepts concept for students of learning deficiencies will also be analyzed and discussed. For the feasibility of remedial instruction based on the cognition diagnosis, clustering method is needed so that students within the same cluster own similar knowledge structures and students among different clusters have the most variance on concept structures [23-33].

2.2. Fuzzy Logic Model of Perception –

Suppose there be a combination of two factors . There are levels and levels for factor and respectively. The fuzzy true values are expressed as Fuzzy truth value and express the degree that the combination of and will support prototype [20] [21]. The probability that the combination of could be viewed as prototype can be expressed as follows [13] [22].

$$p(c_i\ ,\ o_j)=(c_i\, o_j)[c_i o_j+(1-c_i)(1-o_j)]^{-1} \tag{1}$$

Integrated method of individualized concept structure based on fuzzy logic model of perception (FLMP) and interpretive structural modeling (ISM) will be developed.

2.3. New Proposed Algorithm Based on Mahalanobis Distances–

For improving the stability of the clustering results, we replace all of the covariance matrices with the same common covariance matrix in the objective function in the FCM-M algorithm, and then, an improve fuzzy clustering method, called the Fuzzy C-Means algorithm based on common Mahalanobis distance (FCM-CM) is proposed. We can obtain the objective function of FCM-CM as following :

$$J_{FCM-CM}^{m}\left(U,A,\Sigma,X\right)=\sum_{i=1}^{c}\sum_{j=1}^{n}\mu_{ij}^{m}d^{2}\left(\underline{x}_j,\underline{a}_i\right) \tag{2}$$

Conditions for FCM-CM are

$$m\in[1,\infty);U=\left[\mu_{ij}\right]_{c\times n};\mu_{ij}\in\left[0,1\right],\ i=1,2,...,c,\ j=1,2,...,n$$

$$\sum_{i=1}^{c}\mu_{ij}=1,\ j=1,2,...,n,\ 0<\sum_{j=1}^{n}\mu_{ij}<n,\ i=1,2,...,c \tag{3}$$

$$d^2\left(\underline{x}_j,\underline{a}_i\right)=\begin{cases}\left(\underline{x}_j-\underline{a}_i\right)'\Sigma^{-1}\left(\underline{x}_j-\underline{a}_i\right)-\ln\left|\Sigma^{-1}\right| & if\left(\underline{x}_j-\underline{a}_i\right)'\Sigma^{-1}\left(\underline{x}_j-\underline{a}_i\right)-\ln\left|\Sigma^{-1}\right|\geq 0\\ 0 & if\left(\underline{x}_j-\underline{a}_i\right)'\Sigma^{-1}\left(\underline{x}_j-\underline{a}_i\right)-\ln\left|\Sigma^{-1}\right|<0\end{cases} \tag{4}$$

Using the Lagrange's multipliers, we can rewrite the objective function of FCM-CM as following

$$\bar{J}=J_{FCM-CM}^{m}\left(U,A,\Sigma,X\right)=\sum_{i=1}^{c}\sum_{j=1}^{n}\mu_{ij}^{m}\left[\left(\underline{x}_j-\underline{a}_i\right)'\Sigma^{-1}\left(\underline{x}_j-\underline{a}_i\right)-\ln\left|\Sigma^{-1}\right|\right]+\sum_{j=1}^{n}\alpha_j\left(1-\sum_{i=1}^{c}\mu_{ij}\right) \tag{5}$$

Minimizing the objective function respect to all parameters in Equation (4) with the constraint (5), we can obtain the updating equation as follows

(i) $$\frac{\partial \bar{J}}{\partial \underline{a}_i} \overset{\Delta}{=} 0 \Rightarrow \underline{a}_i = \frac{\sum_{j=1}^{n} \mu_{ij}^{m} \underline{x}_j}{\sum_{j=1}^{n} \mu_{ij}^{m}} \quad (6)$$

(ii) $$\frac{\partial \bar{J}}{\partial \alpha_j} \overset{\Delta}{=} 0 \Rightarrow \sum_{i=1}^{c} \mu_{ij} = 1, \quad j = 1, 2, \ldots, n \quad (7)$$

(iii) $$\frac{\partial \bar{J}}{\partial \mu_{ij}} \overset{\Delta}{=} 0, \sum_{i=1}^{c} \mu_{ij} = 1 \Rightarrow \mu_{ij} = \frac{\left[\left(\underline{x}_j - \underline{a}_i\right)' \Sigma^{-1} \left(\underline{x}_j - \underline{a}_i\right) - \ln\left|\Sigma^{-1}\right|\right]^{\frac{-1}{m-1}}}{\sum_{s=1}^{c} \left[\left(\underline{x}_j - \underline{a}_s\right)' \Sigma^{-1} \left(\underline{x}_j - \underline{a}_s\right) - \ln\left|\Sigma^{-1}\right|\right]^{\frac{-1}{m-1}}} \quad (8)$$

III. EXPERIMENT AND RESULT

The performances of clustering Algorithm FCM and FCM-CM all with the fuzzifier m=2, are compared in these experiments. The results of FCM and FCM-CM are obtained by applying the Matlab toolbox developed by [6]. The Mean clustering Accuracies of 100 different initial value sets of FCM and FCM-CM for the Dataset was shown in TABLE 4. From this table, we can find that the Accuracies of FCM is worse than the FCM-CM in the dataset.

Table 1 The Accuracies of four Algorithms

Algorithms	Accuracies
FCM	0.898
FCM-CM	0.927

The performances of our proposed FCM-CM algorithms, FCM-CM is simultaneously better than which of FCM algorithm in the datasets.

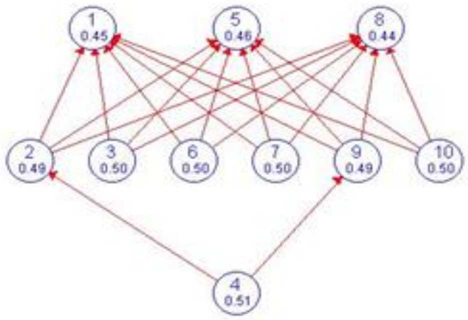
The test includes 19 items with 380 task-takers of Junior College students. The data set comes from the Min-Hwei Junior College of Health Care Management used in the empirical study with learning Registered Nurse's concept. There are 10 concept attributes within each item and they are depicted in Table3.

Table 2. The Details of 10 concept

Classes	Characters about Concepts
1	Basic Nursing 1
2	Basic Nursing 2
3	Basic Medical 1
4	Basic Medical 2
5	Surgical Nursing 1
6	Surgical Nursing 2
7	Production Pediatric Nursing 1
8	Production Pediatric Nursing 2
9	Community spirit Nursing 1
10	Community spirit Nursing 2

Registered Nurse's test for Junior College students is designed by author. The instrument consists of 19 dichotomous items which measure 10 concepts. The data set used in the experimental study is an educational data from Junior College students in Taiwan. There are 380 Junior College students from Taiwan in this test.According the instrument consists of 19 dichotomous items which measure 10 concepts, we use the the Fuzzy C-Means algorithm based on common Mahalanobis distance (called FCM-CM) to cluster 380 Junior College students. We can obtain the six clusters in the dataset. The Accuracies is 92.7% which is better than FCM.

As shown from Figure 1 one student is randomly selected, we see student number 225 in cluster 1, mastery of concept 4 is 0.51. Concept 2 is the basis for concept 2, and 9. From Figure 2, we cannot find any focus item.



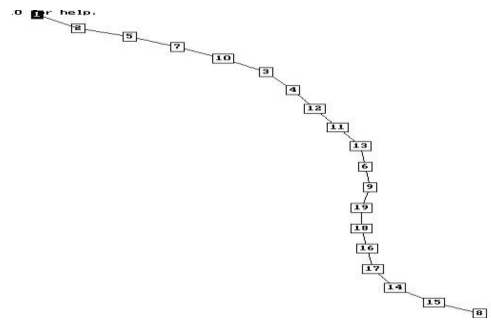
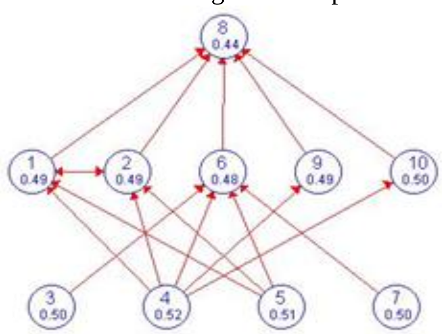


Figure 1. Concept Structure of Student 225 in Cluster 1. Figure 2. Path finder of Student 225 in Cluster 1

As shown from Figure 3 one student is randomly selected, we see student number 138 in cluster 2, mastery of concept 3, 4, 5 and 7 is 0.51. Concept 3, 4, 5 and 7 are the basis for concept 1, 2, 6,9 and 10. From Figure4,we got focus item 6 which belongs to concept 3.



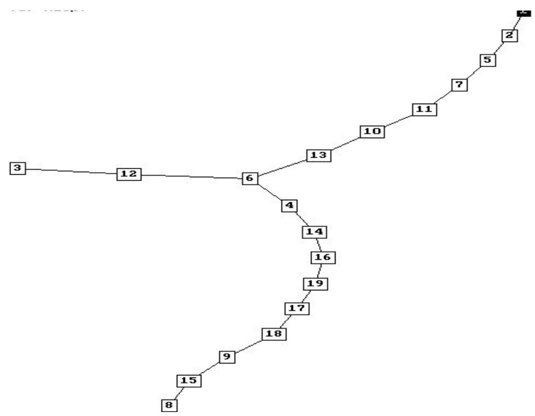
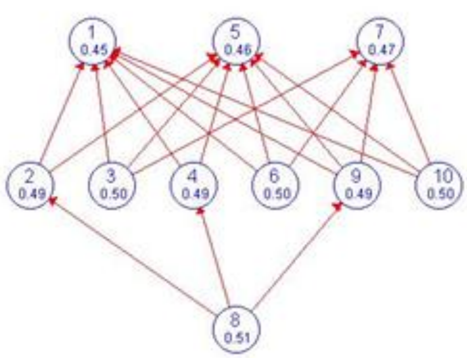


Figure 3. Concept Structure of Student 138 in Cluster 2 Figure 4. Path finder of Student 138 in Cluster 2.

As shown from Figure 5 one student is randomly selected, we see student number 117 in cluster 3, mastery of concept 3 is 0.51. Concept 8 is the basis for concept 2, 4, and 9. From Figure 7, we got focus item 11 which belongs to concept 3.



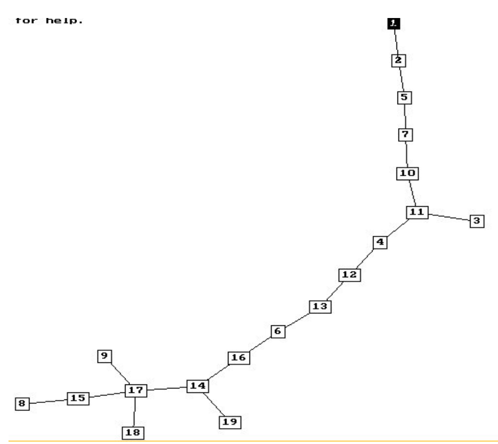
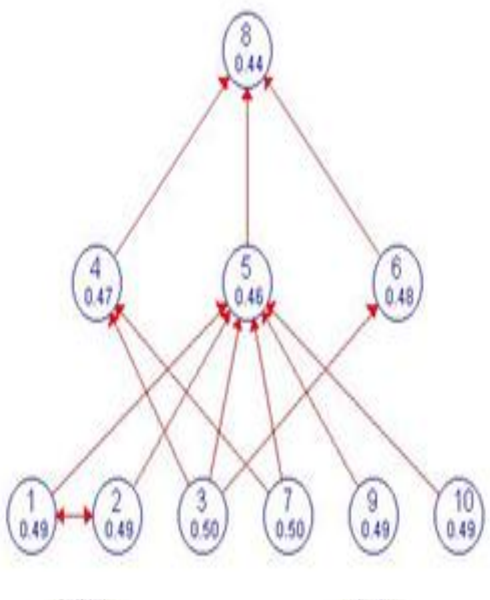


Figure 5. Concept Structure of Student 117 in Cluster 3. Figure.6. Path finder of Student 117 in Cluster 3.

As shown from Figure 7 one student is randomly selected, we see student number 19 in cluster 4, mastery of concept 1 and 2 is 0.49. Concept 1and 2 are the basis for concept 4. From Figure8, we got focus item 13 which belongs to concept 1 and focus item 17 which belongs to concept 2.



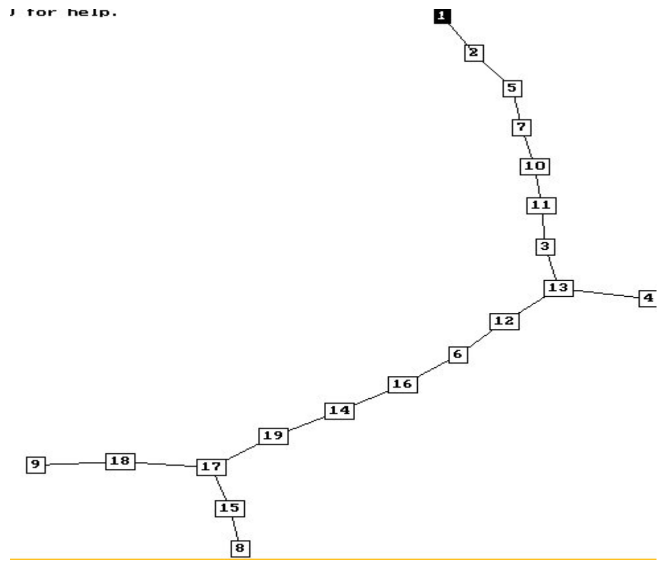
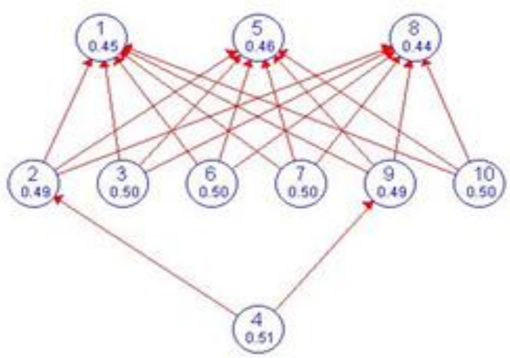


Figure 7. Concept Structure of Student 19 in Cluster 4 Figure 8. Path finder of Student 19 in Cluster 4

As shown from Figure 9 one student is randomly selected, we see student number 356 in cluster 5, mastery of concept 4 is 0.51. Concept 4 are the basis for concept 9 and 2. From Figure 10, we got focus item 16 and 11 which all belong to concept 4.



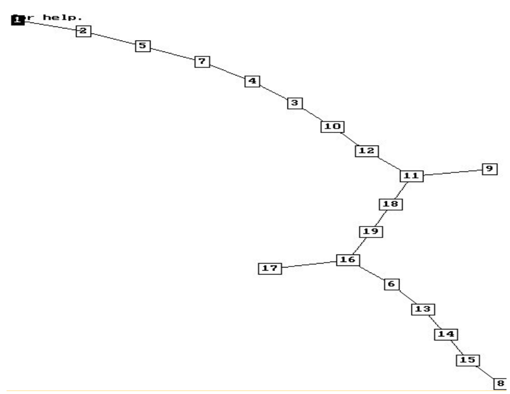


Figure 9. Concept Structure of Student 356 in Cluster 5 Fig. 10. Path finder of Student 356 in Cluster 5

As shown from Figure 11 one student is randomly selected, we see student number 79 in cluster 6, mastery of concept 2 is 0.49. Concept 2 are the basis for concept 1,5 and 8. From Figure 12, we got focus item 17 which belongs to concept 2.

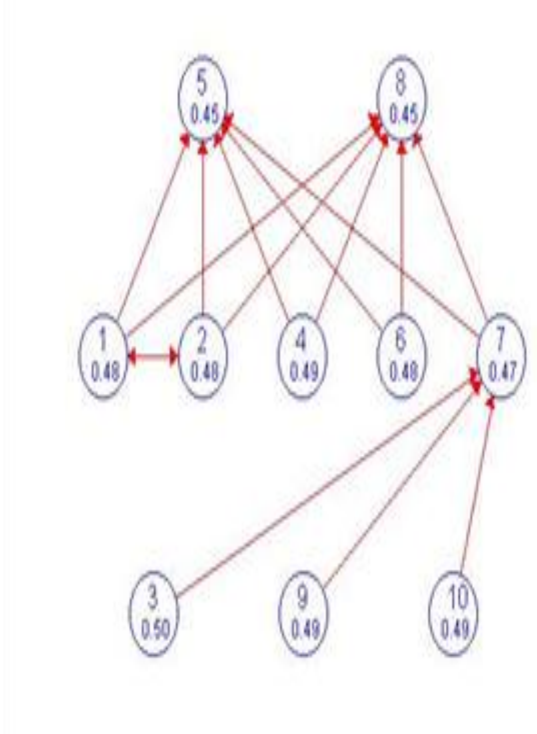


Figure 11. Concept Structure of Student 79 in Cluster 6

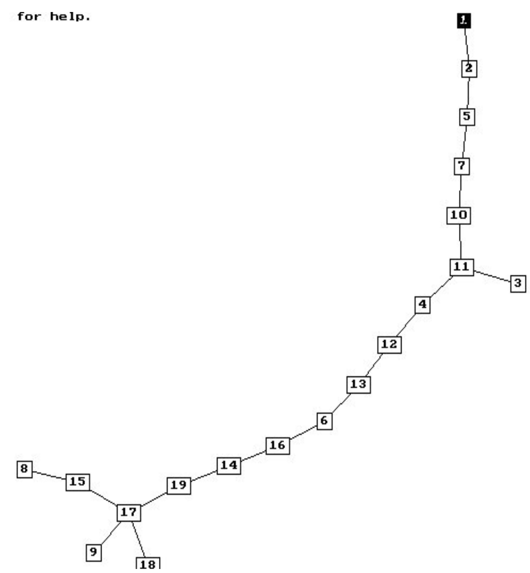


Figure. 12. Path finder of Student 79 in Cluster 6

There are 10 concept attribute within each item. Although the combined algorithm of FLMP and ISM could provide the concept structure of each task-taker respectively, it is unfeasible to display the concept structure of all task-takers in this paper. For the feasibility of remedial instruction based on the cognition diagnosis, clustering method is needed so that students within the same cluster own similar knowledge structures and students among different clusters have the most variance on concept structures. Master of concepts for these six students are different. It is also clearly understood that concept structures of these six students vary a lot.

IV. CONCLUSION

A Fuzzy C-Means algorithm based on Mahalanobis distance (FCM-M) was proposed to improve those limitations of above two algorithms, but it is not stable enough when some of its covariance matrices are not equal. An improved Fuzzy C-Means algorithm based on Common Mahalanobis distance (FCM-CM) is proposed. The experimental results of two real data sets consistently show that the performance of our proposed FCM-CM algorithm is better than FCM algorithms. In this paper, each cluster of data can easily describe features of knowledge structures. We can manage the knowledge structures of Registered Nurse's Concepts to construct the model of features in the pattern recognition completely. An integrated method of FLMP and ISM for analyzing individualized concept structure is provided. With this integrated algorithm, the graphs of concept structures will display the characteristics of knowledge structure. This result corresponds with foundation of cognition diagnosis in psychometrics. This study investigates an integrated methodology to display knowledge structures based on fuzzy clustering with Mahalanobis Distances. In addition, empirical test data of Registered Nurse's for Junior College students are discussed. It shows that knowledge structures will be feasible for remedial instruction [7]. This procedure will also useful for cognition diagnosis. To sum up, this integrated algorithm could improve the assessment methodology of cognition diagnosis and manage the knowledge structures of Registered Nurse's Concepts easily.

http://dx.doi.org/10.21172/ijiet.104.16

V. REFERENCE

[1] R. Krishnapuram and J. Kim (1999), A note on the Gustafson-Kessel and adaptive fuzzy clustering algorithm, IEEE Transactions on Fuzzy Systems, vol. 7, pp. 453-461.

[2] R. Coppi, P. Giordani and P. D'Urso (2006), Component models for fuzzy data, Psychometrika, vol. 71, pp. 733-761.

[3] R. W. Schvaneveldt (1990), Pathfinder Associative Networks: Studies in Knowledge Organization, Norwood, NJ: Ablex Publishing Corp.

[4] Y. H. Lin, M. W. Bart, and K. J. Huang (2006), Generalized polytomous ordering theory [manual and software], National Taichung University, Taiwan.

[5] Hasanzadeh R. P. R., Moradi M. H. and Sadeghi S. H. H. (2005), Fuzzy clustering to the detection of defects from nondestructive testing, 3rd International Conference: Sciences of Electronic Technologies of Information and Telecommunication, March 27-31, 2005, Tunisia.

[6] B. Feil, B. Balasko and J. Abonyi (2007), Visualization of fuzzy clusters by fuzzy Sammon mapping projection: application to the analysis of phase space trajectories, Soft Computing, vol. 11, pp. 479-488.

[7] K. K. Tatsuoka, and M. M. Tatsuoka (1997), Computerized cognitive diagnostic adaptive testing: effect on remedial instruction as empirical validation, Journal of Educational Measurement, vol. 34, pp. 3-20.Kundur D., Hatzinakos D., Towards robust logo watermarking using multiresolution image fusion, IEEE Transcations on Multimedia 6 (2004) 185-197.

[8] H.-C. Liu, J.-M. Yih, W.-C. Lin and D.-B. Wu: Journal of Multiple Valued Logic & Soft Computing, 15 (2009), pp.581-595.

[9] G. J. Klir, and B. Yuan, Fuzzy Sets and Fuzzy Logic: Theory and Applications, Prentice-Hall, New York, NY, 1995.

[10] L. A. Zadeh, Fuzzy sets, Information and Control, 1965, Vol. 8, pp. 338-353.

[11] R. Coppi, P. Giordani and P. D'Urso, Component Models for Fuzzy Data, Psychometrika, Vol. 71, 2006, pp. 733-761.

[12] Y. H. Lin, W. M. Bart and K. J. Huang: WPIRS Software [Manual and Sftware for Gneralized Soring of Item Relational Structure] (2006).

[13] Y. H. Lin, M. W. Bart, and K. J. Huang, Generalized Polytomous Ordering Theory [manual and software], National Taichung University, Taiwan, 2006.

[14] T. Sato, Introduction to S-P Curve Theory Analysis and Evaluation, Tokyo, Meiji Tosho, 1985.

[15] L. F. Blixt, and T.E. Dinero, An Initial Book Look at The Validity of Diagnoses Based on Sato's Caution Index, Education and Psychological Measurement, Vol. 45, 1985, pp. 55-61.T.

[16] R. Coppi, P. Giordani and P. D'Urso (2006), Component models for fuzzy data, Psychometrika, vol. 71, pp. 733-761.

[17] T. Sato, The S-P Chart and The Caution Index, NEC Educational Information Bulletin 80-1, C&C Systems Research Laboratories, Nippon Electic Co., Ltd,. Tokyo, Japan, 1980.

[18] U. Bodenhofer, A Similarity-Based Generalization of Fuzzy Orderings Preserving the Classical Axioms, Fuzziness Knowledge- Based Systems, Vol. 8, 2000, pp. 593-610.

[19] T. Sato, Introduction to S-P Curve Theory Analysis and Evaluation, Tokyo, Meiji Tosho, 1985.

[20] K. K. Tatsuoka, and M. M. Tatsuoka, Computerized Cognitive Diagnostic Adaptive Testing: Effect on Remedial Instruction as Empirical Validation, Journal of Educational Measurement, Vol. 34, 1997, pp. 3-20.

[21] Y. H. Lin, and S. M. Chen, The Integrated Analysis of S-P Chart and Ordering Theory on Equality Axiom Concepts Test for Sixth Graders, WSEAS Transactions on Mathematics, Vol. 5, 2006, pp. 1303-1308.

[22] K. Tatsuka, and M. Tatsuoka, Spotting Erroneous Rules of Operation by The Individual Consistency Index, Journal of Education Measurement, Vol. 20, 1983, pp. 221-230.

[23] D. J. Chen, A. F. Lai, and I. C. Liu, The Design and Implementation of a Diagnostic Test System Based on the Enhanced S-P Model, Journal of Information Science and Engineering, Vol. 21, 2005, pp. 1007-1030.

[24] L. A. Zadeh, Information and Control (1965). Vol. 8, p.338.

[25] Y. H. Lin, M. W. Bart, and K. J. Huang, Generalized Polytomous Ordering Theory(2006). [manual and software], National Taichung University, Taiwan.

[26] G. Klir and B. Yuan, Fuzzy Sets and Fuzzy Logic, Theory and Applications(1995). Prentice Hall.

[27] D. W. Massaro and D. Friedman, Psychological Review(1990). Vol. 97, p.225 .

[28] C. S. Crowther, W. H. Batchelder and X. Hu, Psychological Review (1995). Vol.102, p.396.

[29] J. N. Warfield, Crossing Theory and Hierarchy Mapping(1977).Vol.7, p. 505.

[30] R. Krishnapuram and J. Kim, A note on the Gustafson-Kessel and adaptive fuzzy clustering algorithm, IEEE Transactions on Fuzzy Systems(1999). Vol. 7, no. 4 August.

[31] Gath, and A. B. Geva, Unsupervised optimal fuzzy clustering, IEEE Trans. Pattern Anal. Machine Intell(1989). Vol.11, pp.773-781.

[32] Hasanzadeh R. P. R., Moradi M. H. and Sadeghi S. H. H., Fuzzy clustering to the detection of defects from nondestructive testing, 3rd International Conference: Sciences of Electronic Technologies of Information and Telecommunication(2005). March 27-31, Tunisia.

[33] J. C. Dunn, A fuzzy relative of the isodata process and its use in detecting compact, well-separated clusters J. Cybern(1973). Vol.3, vol.3, pp. 32-57.