with $\mathbf{x}$ and $\mathbf{w}$ being the vectors of the membership grades. Furthermore by admitting this form of generalization, we allow more than two information granules to form the respective mechanism of generalization or specialization (note that our previous construct was quite restrictive, in this regard).

— A logic-based transformation of the membership grades involving more advanced constructs such as compensative operators, weighted means, a family of OWAs operators and the like.

This study can indicate a useful possibility of experimenting with fuzzy sets. It is quite evident that the experimental studies concerning the use of real-world data is scarce. One can number a very few studies along this line, e.g., [14] and [15]. Practically, there are no experimental data sets available to experiment with and this situation leads to the evident shortage of the ensuing experiment-oriented research. The proposed methodology of exploiting information granulation through fuzzy clustering and developing information granules of different size may be of some help by providing synthetic membership data to experiment with. Moreover, it could be used in synthetic description of clustering results. This may pertain either to the same clustering algorithm and results obtained for different numbers of clusters; in this sense we are interested in learning how these information granules generalize (or specialize) some other elements in the family of the granules. One can envision another scenario where the results of clustering are generated by different clustering methods and one is looking for the relationship of generalization or specialization between the results produced by these methods.

References

[1] C. Blake, E. Keogh, and C. J. Merz, *UCI Repository of Machine Learning Databases*. Irvine, CA: Univ. Calif., Dept. Inform. Comput. Sci..

[2] J. C. Bezdek, *Pattern Recognition with Fuzzy Objective Function Algorithms*. New York: Plenum, 1981.

[3] D. Butnariu and E. P. Klement, *Triangular Norm—Based Measures and Games with Fuzzy Coalitions*. Norwell, MA: Kluwer, 1993.

[4] D. Dubois and H. Prade, "Unfair coins and necessity measures: Toward a possibilistic interpretation of histograms," *Fuzzy Sets Syst.*, vol. 10, pp. 15–20, 1983.

[5] F. Höppner, F. Klawonn, R. Kruse, and T. Runkler, *Fuzzy Cluster Analysis*. New York: Wiley, 1999.

[6] G. J. Klir and B. Yuan, *Fuzzy Sets and Fuzzy Logic: Theory and Applications*. Englewood Cliffs, NJ: Prentice-Hall, 1995.

[7] S. Medasani, J. Kim, and R. Krishnapuram, "Estimation of membership functions for pattern recognition and computer vision," in *Fuzzy Logic and Its Applications to Engineering, Information Sciences, and Intelligent Systems*, K. C. Min and Z. Bien, Eds. Norwell, MA: Kluwer, 1995, pp. 45–54.

[8] W. Pedrycz and F. Gomide, *An Introduction to Fuzzy Sets. Analysis and Design*. Cambridge, MA: MIT Press, 1998.

[9] W. Pedrycz, "Fuzzy equalization in the construction of fuzzy sets," *Fuzzy Sets Syst.*, submitted for publication.

[10] W. Pedrycz and G. Vukovich, "Data-based design of fuzzy sets," *J. Fuzzy Log. Intell. Syst.*, vol. 9, no. 3, 1999.

[11] T. L. Saaty, *The Analytic Hierarchy Process*. New York: McGraw-Hill, 1975.

[12] B. Schweizer and A. Sklar, *Probabilistic Metric Spaces*. Amsterdam, The Netherlands: North-Holland, 1983.

[13] S. Weber, "A general concept of fuzzy connectives, negations and implications based on t-norms," *Fuzzy Sets Syst.*, vol. 11, pp. 115–134, 1983.

[14] H. J. Zimmermann, *Fuzzy Set Theory and Its Applications*, 2nd ed. Norwell, MA: Kluwer, 1991.

[15] H. J. Zimmermann and P. Zysno, "Latent connectives in human decision making," *Fuzzy Sets Syst.*, vol. 4, pp. 37–51, 1980.

Document Retrieval Using Fuzzy-Valued Concept Networks

Shyi-Ming Chen, Yih-Jen Horng, and Chia-Hoang Lee

***Abstract*—This paper presents a new method for document retrieval using fuzzy-valued concept networks, where the relevant degrees between the concepts in a fuzzy-valued concept network are represented by arbitrary shapes of fuzzy numbers. There are two kinds of relevant relationships between any two concepts in a fuzzy-valued concept network, i.e., fuzzy positive association and fuzzy negative association. The relevant matrices and the relationship matrices are used to model the fuzzy-valued concept network. The elements in a relevant matrix represent the relevant degrees between concepts. The elements in a relationship matrix represent the relevant relationships between concepts. Furthermore, we also allow users' queries to be represented by arbitrary shapes of fuzzy numbers and to use fuzzy positive association relationship and fuzzy negative association relationship for formulating their queries for increasing the flexibility of fuzzy information retrieval systems. We also present an information retrieval method in the Internet environment based on the network-type fuzzy-valued concept network architecture.**

***Index Terms*—Document retrieval, fuzzy information retrieval, fuzzy numbers, fuzzy-valued concept networks, network-type fuzzy-valued concept networks, relationship matrices, relevant matrices.**

I. Introduction

Most of the existing information retrieval systems are based on the traditional Boolean logic model [19]. The information retrieval systems based on the Boolean logic model all assume that the documents and the users' queries should be represented by precise index terms. This makes these systems restricted in practical applications especially in the circumstance where the information has uncertainty or fuzziness. In order to overcome the drawbacks of the traditional Boolean logic model, some models like the probability model, the fuzzy set model, and the vector space model are proposed [19]. Since the fuzzy set model can properly represent the inexact and uncertain knowledge of human beings, many researches are devoted to use the fuzzy set theory in the design of fuzzy information retrieval systems. Moreover, many fuzzy information retrieval techniques have been presented such as [1], [3]–[5], [8], [9], [11], [14]–[18], and [20].

In [15], Lucarella *et al.* presented an information retrieval method that uses fuzzy concept networks for knowledge representation. A fuzzy concept network consists of nodes and links. Each node in a fuzzy concept network represents a document or a concept, i.e., an index item or a topic of documents. Each link in a fuzzy concept network connects two concepts and is associated with a real value between 0 and 1 which represents the relevant degree between two concepts. By means of the fuzzy inference through fuzzy concept networks, the information retrieval systems are developed. Since the fuzzy inference through the fuzzy concept network is time consuming, in [3] we used concept matrices to model fuzzy concept networks and perform fuzzy inference through concept matrices instead of fuzzy concept network. Since the fuzzy inference through concept matrices

Manuscript received May 10, 1999; revised September 23, 2000. This work was supported by the National Science Council, R.O.C., under Grant NSC 88-2213-E-011-084. This paper was recommended by Associate Editor Teodorescu.

S.-M. Chen is with the Department of Electronic Engineering, National Taiwan University of Science and Technology, Taipei, Taiwan, R.O.C. (e-mail: smchen@et.ntust.edu.tw).

Y.-J. Horng and C.-H. Lee are with the Department of Computer and Information Science, National Chiao Tung University, Hsinchu, Taiwan, R.O.C.

Publisher Item Identifier S 1083-4419(01)00092-9.

can be done more quickly, the fuzzy information retrieval systems we developed in [3] can be more efficient.

However, the fuzzy concept networks presented in [3] and [15] are restricted because the relevant degree between concepts must be a real value between 0 and 1 and the concepts must be linked with the fuzzy positive association relationship. If we can allow the relevant degree between concepts in a fuzzy concept network to be represented by arbitrary shapes of fuzzy numbers and allow the concepts in a fuzzy concept network to be linked with the fuzzy positive association relationship or fuzzy negative association relationship, then there is room for more flexibility. In [5], we presented a method for fuzzy query processing for document retrieval based on extended fuzzy concept networks. However, the method presented in [5] also only allows the relevant degrees between concepts to be represented by real values between 0 and 1. Furthermore, the method presented in [5] is also restricted because users only can use real values between 0 and 1 rather than fuzzy numbers to formulate their queries. In [4], we presented a method for fuzzy query processing for document retrieval, where the relevant degrees between concepts are restricted to be represented by trapezoidal fuzzy numbers, and the relevant relationships between concepts are also restricted to be represented by the fuzzy positive relationship.

In this paper, we use fuzzy-valued concept networks to properly represent fuzzy knowledge for fuzzy information retrieval. A fuzzy-valued concept network consists of nodes and links, each node represents a document or a concept, and each link between two nodes associated with a tuple $(\tilde{\mu}, FR)$ represents the relevance between two nodes, where $\tilde{\mu}$ is a fuzzy number with arbitrary shape representing the relevant degree between two nodes and FR represents the relevant relationship between two nodes, respectively. The values of the relevant degree between any two nodes not only can be real values between 0 and 1, but also can be arbitrary shapes of fuzzy numbers. Moreover, the relevant relationship between any two concepts not only can be a fuzzy positive association relationship, but also can be a fuzzy negative association relationship. In order to reduce the time of fuzzy inference, we use relevant matrices and relationship matrices to model fuzzy-valued concept networks. The elements in a relevant matrix represent the relevant degrees between concepts. The elements in a relationship matrix represent the relevant relationships between concepts. Furthermore, we also allow users' queries to be represented by arbitrary shapes of fuzzy numbers and to use fuzzy positive association relationship and fuzzy negative association relationship for formulating their queries for increasing the flexibility of fuzzy information retrieval systems.

Furthermore, because of the Internet, the documents required by the users should not be bound to a single-host computer. An intelligent information retrieval system must have the capability to help the users to get the documents on different computers through the Internet when the required documents cannot be found on the computers where the users submit their query expressions. Thus, in this paper we also extend the proposed fuzzy-valued concept network architecture to the network-type fuzzy-valued concept network architecture and present a fuzzy information retrieval method based on the network-type fuzzy-valued concept networks in the Internet environment.

The rest of this paper is organized as follows. In Section II, we present the fuzzy-valued concept network architecture for knowledge representation. In Section III, we present a method to model fuzzy-valued concept networks using relevant matrices and relationship matrices. In Section IV, we present an information retrieval method based on the fuzzy-valued concept networks. In Section V, we present a network-type fuzzy-valued concept network architecture for knowledge representation and present a fuzzy information retrieval method in the Internet environment based on the network-type fuzzy-valued concept network architecture. The conclusions are discussed in Section VI.

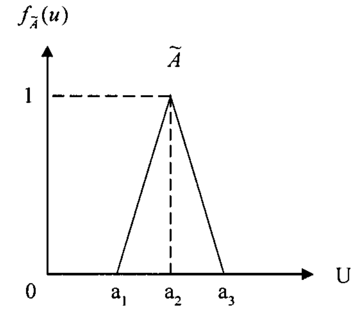

Fig. 1. Triangular fuzzy number.

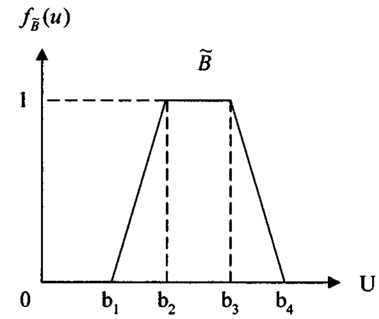

Fig. 2. Trapezoidal fuzzy number.

II. Fuzzy-Valued Concept Networks

In this section, we briefly review the definition of fuzzy numbers [7] and the concepts of fuzzy positive association relationship and fuzzy negative association relationship from [13].

Definition 2.1: A fuzzy number $\tilde{A}$ is a fuzzy set defined in the universe of discourse of U that is both convex and normal. A fuzzy set $\tilde{A}$ is convex if and only if for all u_1, u_2 in U

$$f_{\tilde{A}}(\lambda u_1 + (1-\lambda)u_2) \geq \text{Min}(f_{\tilde{A}}(u_1), f_{\tilde{A}}(u_2)) \tag{1}$$

where $f_{\tilde{A}}$ is the membership function of the fuzzy set $\tilde{A}$, $f_{\tilde{A}}: U \rightarrow [0,1]$ and $\lambda \in [0,1]$. A fuzzy set $\tilde{A}$ is normal if there exists $u_i \in U$, such that $f_{\tilde{A}}(u_i) = 1$, where $f_{\tilde{A}}$ is the membership function of fuzzy set $\tilde{A}$, $f_{\tilde{A}}: U \rightarrow [0,1]$.

From Definition 2.1, we can see that a fuzzy number can be represented by arbitrary shapes. For example, the triangle fuzzy number shown in Fig. 1 and the trapezoidal fuzzy number shown in Fig. 2 are the most often used fuzzy numbers. From Fig. 1, we can see that a triangular fuzzy number $\tilde{A}$ can be represented by a triplet (a_1, a_2, a_3), i.e., $\tilde{A} = (a_1, a_2, a_3)$. From Fig. 2, we can see that a trapezoidal fuzzy number $\tilde{B}$ can be represented by a quadruple (b_1, b_2, b_3, b_4), i.e., $\tilde{B} = (b_1, b_2, b_3, b_4)$.

In this paper, for the convenience of explanations, we assume that the fuzzy numbers used in the fuzzy-valued concept network are all represented by the "close to" shape. However, the fuzzy numbers of arbitrary shapes are also allowed in the fuzzy-valued concept network. According to [7], a "close to γ" fuzzy number is shown in Fig. 3, where γ is a real number or an integer.

The membership function of the fuzzy number "close to γ" is defined by

$$f_{\text{close to } \gamma}(u) = \frac{1}{1 + (\frac{u-\gamma}{\beta})^2} \tag{2}$$

where the crossover points are at $u = \gamma \pm \beta$, and the value of β is the "half-width" of the curve at the crossover points. The larger the value

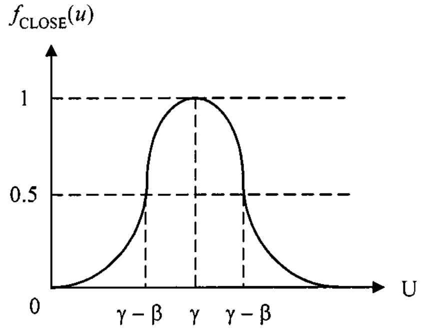

Fig. 3. "Close to γ" fuzzy number.

of β, the wider the curve is. In this paper, we assume that the value of β is 0.1.

According to [12], a fuzzy number $\tilde{A}$ can be decomposed into its level sets or α-cuts), i.e.,

$$\tilde{A} = \int_0^1 \alpha \tilde{A}_\alpha \tag{3}$$

where $\tilde{A}_\alpha = [\alpha_1^{(\alpha)}, a_2^{(\alpha)}]$ is the α-cut of $\tilde{A}$ and $\alpha \in [0,1]$. Assume that there is another fuzzy number $\tilde{B}$

$$\tilde{B} = \int_0^1 \alpha \tilde{B}_\alpha \tag{4}$$

where $\tilde{B}_\alpha = [b_1^{(\alpha)}, b_2^{(\alpha)}]$ is the α-cut of $\tilde{B}$ and $\alpha \in [0,1]$. Then, according to [12], the "OR" operation and the "AND" operation of the fuzzy numbers $\tilde{A}$ and $\tilde{B}$ are defined by

$$\tilde{A} \circledvee \tilde{B} = \int_0^1 \alpha \left[a_1^{(\alpha)} \vee b_1^{(\alpha)}, a_2^{(\alpha)} \vee b_2^{(\alpha)} \right] \tag{5}$$

$$\tilde{A} \circledwedge \tilde{B} = \int_0^1 \alpha \left[a_1^{(\alpha)} \wedge b_1^{(\alpha)}, a_2^{(\alpha)} \wedge b_2^{(\alpha)} \right] \tag{6}$$

where "$\circledvee$" and "$\circledwedge$" are the "OR" operator and the "AND" operator of the fuzzy numbers, respectively, where "$\wedge$" is the minimum operator, "$\vee$" is the maximum operator, and $\alpha \in [0,1]$.

Definition 2.2: Let $\tilde{A}$ and $\tilde{B}$ be two fuzzy numbers of the universe of discourse U with member functions $f_{\tilde{A}}$ and $f_{\tilde{B}}$, respectively, where $f_{\tilde{A}}: U \to [0,1]$ and $f_{\tilde{B}}: U \to [0,1]$. If $\forall\, u_i \in U, f_{\tilde{A}}(u_i) = f_{\tilde{B}}(u_i)$, then the fuzzy numbers $\tilde{A}$ and $\tilde{B}$ are called equal, i.e., $\tilde{A} = \tilde{B}$.

In the following, we briefly review the concepts of fuzzy positive association relationship and fuzzy negative association relationship from [13]

1) *Fuzzy positive association:* It relates concepts that have, in some contexts, a similar meaning, e.g., person $\leftrightarrow$ individual or which are typically used in the same context, e.g., person $\leftrightarrow$ address.
2) *Fuzzy negative association:* It relates concepts which are complementary, e.g. male $\leftrightarrow$ female; incompatible, e.g., unemployed $\leftrightarrow$ freelance; or antonyms, e.g., small $\leftrightarrow$ large.

Definition 2.3: A fuzzy-valued concept network can be represented as $EFCN(N, L)$, where N is a set of nodes, and each node stands for a concept or a document and L is a set of directed edges between nodes. If $\ell \in L$ then ℓ is associated with a tuple $(\tilde{\mu}, FR)$, where $\tilde{\mu}$ represents the degree of linking strength between nodes and its value is a fuzzy number. FR is the relationship between two nodes linked by the directed edge ℓ, and $FR \in \{P, N\}$, where P stands for fuzzy positive association relationship and N stands for fuzzy negative association relationship.

Example 2.1: Assume that there is a fuzzy-valued concept network as shown in Fig. 4.

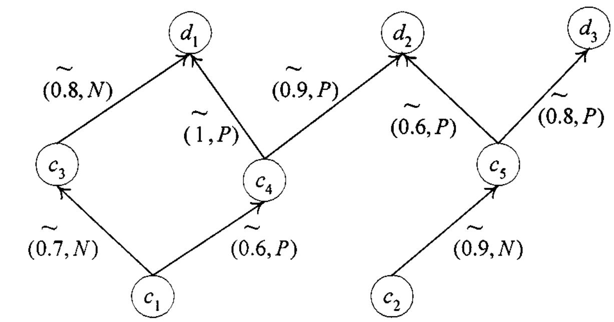

Fig. 4. Example of fuzzy-valued concept network.

From Fig. 4, we can see that the relationship between concept c_4 and concept c_1 is a fuzzy positive association relationship with a relevant degree $\widetilde{0.6}$, i.e., close to 0.6; the relationship between concept c_1 and concept c_3 is a fuzzy negative association relationship with the relevant degree $\widetilde{0.7}$, i.e., close to 0.7; concept c_5 and concept c_2 are linked by a fuzzy positive association relationship with the relevant degree $\widetilde{0.9}$, i.e., close to 0.9. Document d_2 contains concept c_4 with the relevant degree $\widetilde{0.9}$, i.e., close to 0.9 and contains concept c_5 with the relevant degree $\widetilde{0.6}$, i.e., close to 0.6.

III. Relevant Matrices and Relationship Matrices

In this section, we present a method to model fuzzy-valued concept networks using relevant matrices and relationship matrices. The definitions of the transitive closure of relevant matrices and the transitive closure of relationship matrices are also presented in this section.

Definition 3.1: The relevant matrix V is a fuzzy matrix, where the element $\widetilde{v_{ij}}$ represents the relevant degree between concept c_i and concept c_j in a fuzzy-valued concept network, and $\widetilde{v_{ij}}$ is a fuzzy number. If $\widetilde{v_{ij}} = \tilde{0}$, then it means that the relevant degree between concept c_i and concept c_j is not given by the experts in the fuzzy-valued concept network.

Definition 3.2: Let P and Q be two relevant matrices with elements denoted by $\widetilde{p_{ij}}$ and $\widetilde{q_{ij}}$, respectively, where $\widetilde{p_{ij}}$ and $\widetilde{q_{ij}}$ are fuzzy numbers and $1 \leq i \leq j \leq n$. If $\forall i \forall j, \widetilde{p_{ij}} = \widetilde{q_{ij}}$, then the relevant matrices P and Q are called equal, i.e., $P = Q$.

Definition 3.3: Assume that V is a relevant matrix

$$V = \begin{bmatrix} \widetilde{v_{11}} & \widetilde{v_{12}} & \cdots & \widetilde{v_{1n}} \\ \widetilde{v_{21}} & \widetilde{v_{22}} & \cdots & \widetilde{v_{2n}} \\ \vdots & \vdots & \cdots & \vdots \\ \widetilde{v_{n1}} & \widetilde{v_{n2}} & \cdots & \widetilde{v_{nn}} \end{bmatrix}$$

where n is the number of concepts in a fuzzy-valued concept network. See (7), shown at the bottom of the next page, where $\circledvee$ and $\circledwedge$ are the "OR" operator and "AND" operator of fuzzy numbers, respectively. Then there exists a positive integer p, $p \leq n-1$, such that $V^p = V^{p+1} = V^{p+2} = \cdots$. Let $T = V^P$, then T is called the transitive closure of the relevant matrix V.

Definition 3.4: The relationship matrix R is a fuzzy matrix, where the element r_{ij} represents the relationship between concept c_i and concept c_j in a fuzzy-valued concept network and $r_{ij} \in \{P, N, Z\}$, where P stands for the fuzzy positive association relationship, N stands for the fuzzy negative association relationship, and Z stands for the unknown relationship. If $r_{ij} = Z$, then it means that the relationship between concept c_i and concept c_j is not given by the experts in the fuzzy-valued concept network.

Definition 3.5: Let R and S be two relationship matrices with elements represented by r_{ij} and s_{ij}, respectively, where $r_{ij} \in \{P, N, Z\}$

TABLE I
COMBINATION OF FUZZY RELATIONSHIPS

	P	N	Z
P	P	N	Z
N	N	P	Z
Z	Z	Z	Z

and $s_{ij} \in \{P, N, Z\}$, and $1 \le i \le j \le n$. If $r_{ij} = s_{ij}$, then the relationship matrices R and S are called equal, i.e., $R = S$.

Definition 3.6: Assume that R is a relationship matrix and

$$R = \begin{bmatrix} r_{11} & r_{12} & \cdots & r_{1n} \\ r_{21} & r_{22} & \cdots & r_{2n} \\ \vdots & \vdots & \cdots & \vdots \\ r_{n1} & r_{n2} & \cdots & r_{nn} \end{bmatrix}$$

where n is the number of concepts in the fuzzy-valued concept network, $r_{ij} \in \{P, N, Z\}$, $1 \le i \le n$, and $1 \le j \le n$. See (8), shown at the bottom of the page, where "$\boxed{\vee}$" is the operator of choosing the fuzzy relationships whose priority is the highest. In this paper, we give the first priority to the fuzzy negative association relationship (N), the fuzzy positive association relationship (P) gets the second priority, and the relationship (Z) gets the lowest priority, i.e., $N > P > Z$. "$\boxed{\wedge}$" is the operator of choosing the combination of two relationships according to Table I. Then, there exists a positive integer $p, p \le n - 1$ such that $R^p = R^{p+1} = R^{p+2} = \cdots$. Let $L = R^p$, where L is called the transitive closure of the relationship matrix R.

Example 3.1: Assume that there is a fuzzy-valued concept network as shown in Fig. 5.

Then, we can use the relevant matrix V and the relationship matrix R shown as follows to model the fuzzy-valued concept network, where

$$V = \begin{array}{c} \\ c_1 \\ c_2 \\ c_3 \\ c_4 \\ c_5 \end{array} \begin{array}{c} \begin{array}{ccccc} c_1 & c_2 & c_3 & c_4 & c_5 \end{array} \\ \begin{bmatrix} \tilde{1} & \tilde{0} & \widetilde{0.7} & \widetilde{0.6} & \tilde{0} \\ \tilde{0} & \tilde{1} & \tilde{0} & \tilde{0} & \widetilde{0.9} \\ \widetilde{0.7} & \tilde{0} & \tilde{1} & \tilde{0} & \tilde{0} \\ \widetilde{0.6} & \tilde{0} & \tilde{0} & \tilde{1} & \tilde{0} \\ \tilde{0} & \widetilde{0.9} & \tilde{0} & \tilde{0} & \tilde{1} \end{bmatrix} \end{array}$$

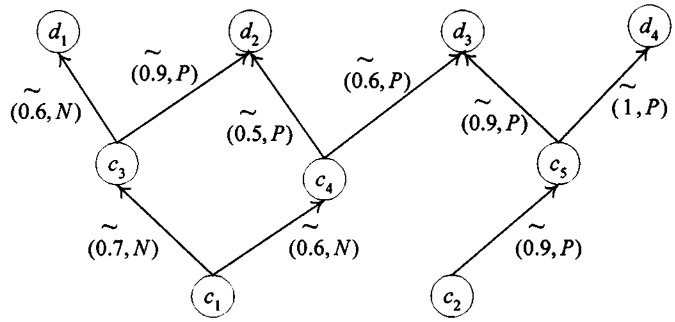

Fig. 5. Fuzzy-valued concept network used in Example 3.1.

$$R = \begin{array}{c} \\ c_1 \\ c_2 \\ c_3 \\ c_4 \\ c_5 \end{array} \begin{array}{c} \begin{array}{ccccc} c_1 & c_2 & c_3 & c_4 & c_5 \end{array} \\ \begin{bmatrix} P & Z & N & N & Z \\ Z & P & Z & Z & P \\ N & Z & P & Z & Z \\ N & Z & Z & P & Z \\ Z & P & Z & Z & P \end{bmatrix} \end{array}.$$

According to Definition 3.3 and Definition 3.6 we can obtain the transitive closure T of the relevant matrix V and the transitive closure L of the relationship matrix R shown as follows:

$$T = \begin{array}{c} \\ c_1 \\ c_2 \\ c_3 \\ c_4 \\ c_5 \end{array} \begin{array}{c} \begin{array}{ccccc} c_1 & c_2 & c_3 & c_4 & c_5 \end{array} \\ \begin{bmatrix} \tilde{1} & \tilde{0} & \widetilde{0.7} & \widetilde{0.6} & \tilde{0} \\ \tilde{0} & \tilde{1} & \tilde{0} & \tilde{0} & \widetilde{0.9} \\ \widetilde{0.7} & \tilde{0} & \tilde{1} & \widetilde{0.6} & \tilde{0} \\ \widetilde{0.6} & \tilde{0} & \widetilde{0.6} & \tilde{1} & \tilde{0} \\ \tilde{0} & \widetilde{0.9} & \tilde{0} & \tilde{0} & \tilde{1} \end{bmatrix} \end{array}$$

$$L = \begin{array}{c} \\ c_1 \\ c_2 \\ c_3 \\ c_4 \\ c_5 \end{array} \begin{array}{c} \begin{array}{ccccc} c_1 & c_2 & c_3 & c_4 & c_5 \end{array} \\ \begin{bmatrix} P & Z & N & N & Z \\ Z & P & Z & Z & P \\ N & Z & P & P & Z \\ N & Z & P & P & Z \\ Z & P & Z & Z & P \end{bmatrix} \end{array}.$$

IV. Fuzzy Query Processing for Document Retrieval Using Fuzzy-Valued Concept Networks

In this section, we present a method for fuzzy query processing for document retrieval using fuzzy-valued concept networks. First, we in-

$$V^2 = V \odot V = \begin{bmatrix} \overset{\circledvee}{\underset{i=1,\ldots,n}{}} (\widetilde{v_{1i}} \circledwedge \widetilde{v_{i1}}) & \overset{\circledvee}{\underset{i=1,\ldots,n}{}} (\widetilde{v_{1i}} \circledwedge \widetilde{v_{i2}}) & \cdots & \overset{\circledvee}{\underset{i=1,\ldots,n}{}} (\widetilde{v_{1i}} \circledwedge \widetilde{v_{in}}) \\ \overset{\circledvee}{\underset{i=1,\ldots,n}{}} (\widetilde{v_{2i}} \circledwedge \widetilde{v_{i1}}) & \overset{\circledvee}{\underset{i=1,\ldots,n}{}} (\widetilde{v_{2i}} \circledwedge \widetilde{v_{i2}}) & \cdots & \overset{\circledvee}{\underset{i=1,\ldots,n}{}} (\widetilde{v_{2i}} \circledwedge \widetilde{v_{in}}) \\ \vdots & \vdots & \vdots & \vdots \\ \overset{\circledvee}{\underset{i=1,\ldots,n}{}} (\widetilde{v_{n1}} \circledwedge \widetilde{v_{i1}}) & \overset{\circledvee}{\underset{i=1,\ldots,n}{}} (\widetilde{v_{ni}} \circledwedge \widetilde{v_{i2}}) & \cdots & \overset{\circledvee}{\underset{i=1,\ldots,n}{}} (\widetilde{v_{ni}} \circledwedge \widetilde{v_{in}}) \end{bmatrix} \quad (7)$$

$$R^2 = R * R = \begin{bmatrix} \overset{\boxed{\vee}}{\underset{i=1,\ldots,n}{}} (r_{1i} \; \boxed{\wedge} r_{i1}) & \overset{\boxed{\vee}}{\underset{i=1,\ldots,n}{}} (r_{1i} \; \boxed{\wedge} r_{i2}) & \cdots & \overset{\boxed{\vee}}{\underset{i=1,\ldots,n}{}} (r_{1i} \; \boxed{\wedge} r_{in}) \\ \overset{\boxed{\vee}}{\underset{i=1,\ldots,n}{}} (r_{2i} \; \boxed{\wedge} r_{i1}) & \overset{\boxed{\vee}}{\underset{i=1,\ldots,n}{}} (r_{2i} \; \boxed{\wedge} r_{i2}) & \cdots & \overset{\boxed{\vee}}{\underset{i=1,\ldots,n}{}} (r_{2i} \; \boxed{\wedge} r_{in}) \\ \overset{\boxed{\vee}}{\underset{i=1,\ldots,n}{}} \vdots & \vdots & \vdots & \vdots \\ \overset{\boxed{\vee}}{\underset{i=1,\ldots,n}{}} (r_{ni} \; \boxed{\wedge} r_{i1}) & \overset{\boxed{\vee}}{\underset{i=1,\ldots,n}{}} (r_{ni} \; \boxed{\wedge} r_{i2}) & \cdots & \overset{\boxed{\vee}}{\underset{i=1,\ldots,n}{}} (r_{ni} \; \boxed{\wedge} r_{in}) \end{bmatrix} \quad (8)$$

troduce the definitions of document descriptor relevant matrices and document descriptor relationship matrices.

Definition 4.1: Let D be a set of documents in a fuzzy-valued concept network, $D = \{d_1, d_2, \ldots, d_m\}$ and let C be a set of concepts in a fuzzy-valued concept network, $C = \{c_1, c_2, \ldots, c_n\}$. Then, the document descriptor relevant matrix E is shown as follows:

$$E = \begin{array}{c} \\ d_1 \\ d_2 \\ \vdots \\ d_m \end{array} \begin{array}{c} \begin{array}{cccc} c_1 & c_2 & \cdots & c_n \end{array} \\ \begin{bmatrix} \widetilde{e_{11}} & \widetilde{e_{12}} & \cdots & \widetilde{e_{1n}} \\ \widetilde{e_{21}} & \widetilde{e_{22}} & \cdots & \widetilde{e_{2n}} \\ \vdots & \vdots & \cdots & \vdots \\ \widetilde{e_{m1}} & \widetilde{e_{m2}} & \cdots & \widetilde{e_{mn}} \end{bmatrix} \end{array}$$

where

- m number of documents;
- n number of concepts, $\widetilde{e_{ij}}$ stands for the relevant degree between document d_i and concept c_j;
- $\widetilde{e_{ij}}$ fuzzy number, $1 \le i \le m$, and $1 \le j \le n$.

Definition 4.2: The document descriptor relationship matrix F is shown as follows:

$$F = \begin{array}{c} \\ d_1 \\ d_2 \\ \vdots \\ d_m \end{array} \begin{array}{c} \begin{array}{cccc} c_1 & c_2 & \cdots & c_n \end{array} \\ \begin{bmatrix} f_{11} & f_{12} & \cdots & f_{1n} \\ f_{21} & f_{22} & \cdots & f_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ f_{m1} & f_{m2} & \cdots & f_{mn} \end{bmatrix} \end{array}$$

where f_{ij} stands for the fuzzy relationship between document d_i and concept c_j, $f_{ij} \in \{P, N, Z\}$.

However, the experts may forget to set the relevant degrees and relationships between some documents and some concepts. Since the implicit relevant degrees and relationships between concepts can be obtained from the transitive closure T of the relevant matrix V and the transitive closure L of the relationship matrix R, we can use the transitive closure T of the relevant matrix V and the transitive closure L of the relationship matrix R to get the implicit relevant degrees and relationships between documents and concepts. Let $E^* = E \odot T$, then E^* includes the implicit relevant degrees between documents and concepts. Let $F^* = F * L$, then F^* includes the implicit relevant relationships between documents and concepts. E^* and F^* will then be used as a basis for similarity measures between queries and documents. Each row of E^* can be thought as a document descriptor relevant vector and each row of F^* can be thought as a document descriptor relationship vector.

The user's query Q can be represented by a query descriptor relevant vector $\overline{qv}$ and a query descriptor relationship vector $\overline{qr}$ shown as follows:

$$\overline{qv} = \langle \tilde{x_1}, \tilde{x_2}, \ldots, \tilde{x_n} \rangle$$
$$\overline{qr} = \langle y_1, y_2, \ldots, y_n \rangle$$

where $\tilde{x_i}$ means the relevant degree between desired documents and concept c_i, $\tilde{x_i}$ is a fuzzy number, and $1 \le i \le n$; y_i means the relationship between desired documents and concept c_i and $y_i \in \{P, N\}$. If $y_i = P$, then the desired documents should contain c_i; if $y_i = N$, then the desired documents should not contain concept c_i. Moreover, if the user doesn't set the values of $\tilde{x_i}$ and y_i, then concept c_i is thought as been neglected by the user, and $\tilde{x}_i$ and y_i will be labeled as "-". That is, the users "do not care" whether the retrieved documents contain concept c_i or not.

Assume that there are two tuples, i.e., $\langle \tilde{A}, B \rangle$ and $\langle \tilde{C}, D \rangle$, where $\tilde{A}$ and $\tilde{C}$ are fuzzy numbers, $B \in \{P, N, Z\}$, and $D \in \{P, N, Z\}$, then the degree of similarity between $\langle \tilde{A}, B \rangle$ and $\langle \tilde{C}, D \rangle$ can be calculated by

$$Y(\langle \tilde{A}, B \rangle, \langle \tilde{C}, D \rangle) = \begin{cases} 0, & \text{if } B \ne D \\ 1 - \sup_\alpha \frac{\left|a_1^{(\alpha)} - c_1^{(\alpha)}\right| + \left|a_2^{(\alpha)} - c_2^{(\alpha)}\right|}{2}, & \text{if } B = D \end{cases} \tag{9}$$

where $\alpha \in [0, 1]$ and $Y(\langle \tilde{A}, B \rangle, \langle \tilde{C}, D \rangle) \in [0, 1]$.

Assume that the document descriptor relevant vector $\overline{dv_i}$ and the document descriptor relationship vector $\overline{dr_i}$ are represented as follows:

$$\overline{dv_i} = \langle \tilde{v}_{i1}, \tilde{v}_{i2}, \ldots \tilde{v}_{in} \rangle$$
$$\overline{dr_i} = \langle r_{i1}, r_{i2}, \ldots, r_{in} \rangle.$$

Then, the degree of satisfaction that document d_i satisfies the user's query Q can be evaluated by (10), shown at the bottom of the page, where

- $\overline{qv(j)}$ jth element of the query descriptor relevant vector;
- $\overline{qv}, \overline{qr(j)}$ jth element of query descriptor relationship vector $\overline{qr}$, $1 \le j \le n$, $RS(d_i) \in [0, 1]$;
- k number of concepts not neglected by the user query.

The information retrieval system would display every document having the degree of satisfaction greater than a threshold value λ, where $\lambda \in [0, 1]$, in a sequential order from the document with the highest degree of satisfaction to that with the lowest one.

Example 4.1: Assume that we have the same fuzzy-valued concept network as shown in Example 3.1, and there are four documents d_1, d_2, d_3, d_4 as shown in Fig. 5. Then, the document descriptor relevant matrix E and the document descriptor relationship matrix F are as follows:

$$E = \begin{array}{c} \\ d_1 \\ d_2 \\ d_3 \\ d_4 \end{array} \begin{array}{c} \begin{array}{ccccc} c_1 & c_2 & c_3 & c_4 & c_5 \end{array} \\ \begin{bmatrix} \tilde{0} & \tilde{0} & \widetilde{0.6} & \tilde{0} & \tilde{0} \\ \tilde{0} & \tilde{0} & \widetilde{0.9} & \widetilde{0.5} & \tilde{0} \\ \tilde{0} & \tilde{0} & \tilde{0} & \widetilde{0.6} & \widetilde{0.9} \\ \tilde{0} & \tilde{0} & \tilde{0} & \tilde{0} & \tilde{1} \end{bmatrix} \end{array}$$

$$F = \begin{array}{c} \\ d_1 \\ d_2 \\ d_3 \\ d_4 \end{array} \begin{array}{c} \begin{array}{ccccc} c_1 & c_2 & c_3 & c_4 & c_5 \end{array} \\ \begin{bmatrix} Z & Z & N & Z & Z \\ Z & Z & P & P & Z \\ Z & Z & Z & P & P \\ Z & Z & Z & Z & P \end{bmatrix} \end{array}.$$

The transitive closure T of the relevant matrix V and the transitive closure L of the relationship matrix R has been obtained as shown in Example 3.1. Because $E^* = E \odot T$ and $F^* = F * L$, we can obtain E^* and F^* shown as follows:

$$E^* = \begin{array}{c} \\ d_1 \\ d_2 \\ d_3 \\ d_4 \end{array} \begin{array}{c} \begin{array}{ccccc} c_1 & c_2 & c_3 & c_4 & c_5 \end{array} \\ \begin{bmatrix} \widetilde{0.6} & \tilde{0} & \widetilde{0.6} & \widetilde{0.6} & \tilde{0} \\ \widetilde{0.7} & \tilde{0} & \widetilde{0.9} & \widetilde{0.6} & \tilde{0} \\ \widetilde{0.6} & \widetilde{0.9} & \widetilde{0.6} & \widetilde{0.6} & \widetilde{0.9} \\ \tilde{0} & \widetilde{0.9} & \tilde{0} & \tilde{0} & \tilde{1} \end{bmatrix} \end{array}$$

$$F^* = \begin{array}{c} \\ d_1 \\ d_2 \\ d_3 \\ d_4 \end{array} \begin{array}{c} \begin{array}{ccccc} c_1 & c_2 & c_3 & c_4 & c_5 \end{array} \\ \begin{bmatrix} P & Z & N & N & Z \\ N & Z & P & P & Z \\ N & P & P & P & P \\ Z & P & Z & Z & P \end{bmatrix} \end{array}.$$

$$RS(d_i) = \frac{\sum_{\overline{qv(j)} \ne \text{“-” and } \overline{qr(j)} \ne \text{“-” and } j=1,\ldots,n} Y(\langle \tilde{v}_{ij}, r_{ij} \rangle, \langle \tilde{x}_j, y_j \rangle)}{k} \tag{10}$$

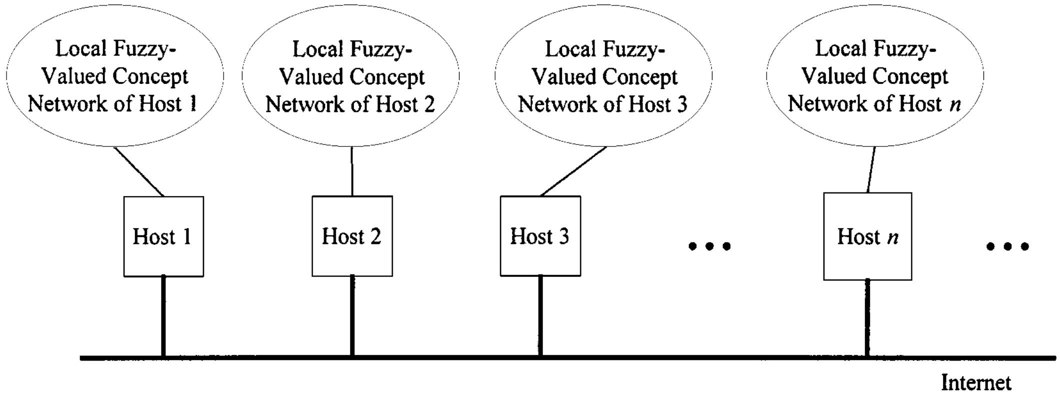

Fig. 6. Architecture of the network-type fuzzy-value concept network.

Assume that the user's query Q is represented by the query descriptor relevant vector $\overline{qv}$ and the query descriptor relationship vector $\overline{qr}$ shown as follows:

$$\overline{qv} = \langle \widetilde{0.6}, \text{-}, \widetilde{0.8}, \widetilde{0.6}, \widetilde{0.7} \rangle$$
$$\overline{qr} = \langle P, \text{-}, P, N, P \rangle.$$

Then, based on (10), we can get

$$RS(d_1) = \frac{1+0+1+0}{4} = \frac{2}{4} = 0.5$$
$$RS(d_2) = \frac{0+0.9+0+0}{4} = \frac{0.9}{4} = 0.225$$
$$RS(d_3) = \frac{0+0.8+0+0.8}{4} = \frac{1.6}{4} = 0.4$$
$$RS(d_4) = \frac{0+0+0+0.7}{4} = \frac{0.7}{4} = 0.175.$$

Assume that the information retrieval threshold value $\alpha = 0.2$, then the sequential order from highest retrieval status value to that with the lowest retrieval status value is $d_1 > d_3 > d_2$. In this case, document d_1 is the best choice for the user's query, and document d_4 will not be retrieved in this example due to the fact that its degree of satisfaction is smaller than 0.2.

V. Fuzzy Query Processing Using Fuzzy-Valued Concept Networks in the Internet Environment

Since the Internet became prevalent [6], [21], the information about the documents needed by the user should not be bound on a single host computer. When the users' queries cannot be satisfied on the local computer, the information retrieval system should expand its searching capability to other computers on the Internet until the required documents are either found or they do not exist.

In this section, we present the network-type fuzzy-valued concept networks architecture as the basis for fuzzy information retrieval in the Internet environment. The architecture of the network-type fuzzy-value concept network is shown in Fig. 6.

From Fig. 6, we can see that each host links to the Internet by the bold black lines. Each host has its local fuzzy-valued concept network as the knowledge base of the documents and concepts. Substantially, the local fuzzy-valued concept networks inside these hosts are the same as the ones presented in the previous sections. That is, the fuzzy-valued concept networks inside these hosts allow the values of the relevant degrees between concepts to be arbitrary shapes of fuzzy numbers, and the relevant relationships between nodes to be not only fuzzy positive association relationship, but also fuzzy negative association relationship.

Since the local fuzzy-valued concept networks inside these hosts are the same as the ones presented in the previous sections, we can also model these local fuzzy-valued concept networks by using relevant matrices and relationship matrices. Furthermore, we can get the transitive closures of the relevant matrices and the transitive closures of the relationship matrices when the relevant matrices and relationship matrices are known. The implicit relevant degrees and implicit relationships between concepts then can be found in the transitive closures of the relevant matrices and the transitive closures of the relationship matrices, respectively.

The document descriptor relevant matrices and document descriptor relationship matrices can model the relevant degrees and fuzzy relationships between documents and concepts in each local fuzzy-valued concept network inside each host in the Internet environment. However, the experts may forget to set the relevant degrees or fuzzy relationships between some documents and concepts. Because all associate concepts are linked together, we can get the implicit relevant degrees and fuzzy relationships between documents and concepts by the transitive closures of the relevant matrices and the transitive closures of the relationship matrices. Assume the document descriptor relevant matrix is E, and the transitive closure of the relevant matrix is T, let $E^* = E \odot T$, then E^* includes all the implicit relevant degrees between documents and concepts. Assume that the document descriptor relation matrix is F, and assume that the transitive closure of the relationship matrix is L, let $F^* = F * L$, then F^* includes all the implicit relationships between documents and concepts.

By the previous discussions, we know that the fuzzy-valued concept networks contain nodes and links. These nodes stand for either documents or concepts. In the network-type fuzzy-valued concept network architecture, we assume that each local fuzzy-valued concept network may have an identical number of concept nodes and a different number of document nodes. Therefore, the relevant matrices and relationship matrices used to model the local fuzzy-valued concept networks on each host are identical. But the document descriptor relevant matrices and the document descriptor relationship matrices are different on different hosts.

Example 5.1: Assume that Figs. 7 and 8 are two local fuzzy-valued concept networks on host 1 and host 2, respectively, which are linked by the Internet. From Figs. 7 and 8, we can see that concepts c_1, c_2, c_3, c_4, c_5, and documents d_1, d_2, d_3 are located on host 1, and that concepts c_2, c_3, c_5, c_6, c_7 and documents d_4, d_5, d_6 are located on host 2.

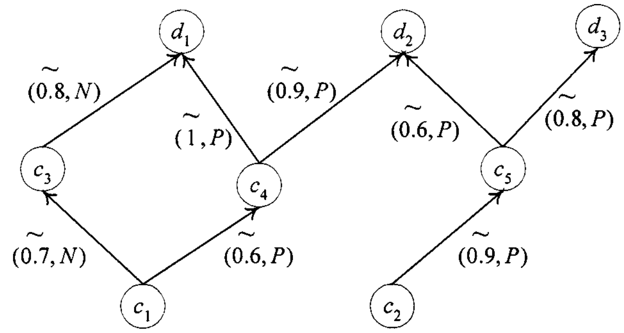

Fig. 7. Fuzzy-valued concept network on host 1.

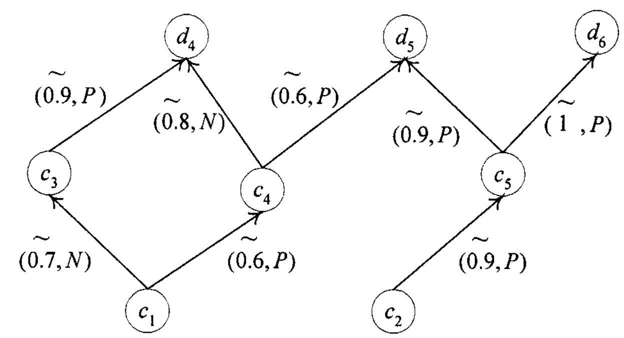

Fig. 8. Fuzzy-valued concept network on host 2.

By the previous discussions, we can see that the relevant matrices and relationship matrices on these two hosts are V and R, respectively, where

$$V = \begin{array}{cc} & \begin{array}{ccccc} c_1 & c_2 & c_3 & c_4 & c_5 \end{array} \\ \begin{array}{c} c_1 \\ c_2 \\ c_3 \\ c_4 \\ c_5 \end{array} & \left[\begin{array}{ccccc} \tilde{1} & \tilde{0} & \widetilde{0.7} & \widetilde{0.6} & \tilde{0} \\ \tilde{0} & \tilde{1} & \tilde{0} & \tilde{0} & \widetilde{0.9} \\ \widetilde{0.7} & \tilde{0} & \tilde{1} & \tilde{0} & \tilde{0} \\ \widetilde{0.6} & \tilde{0} & \tilde{0} & \tilde{1} & \tilde{0} \\ \tilde{0} & \widetilde{0.9} & \tilde{0} & \tilde{0} & \tilde{1} \end{array}\right] \end{array}$$

$$R = \begin{array}{cc} & \begin{array}{ccccc} c_1 & c_2 & c_3 & c_4 & c_5 \end{array} \\ \begin{array}{c} c_1 \\ c_2 \\ c_3 \\ c_4 \\ c_5 \end{array} & \left[\begin{array}{ccccc} P & Z & N & P & Z \\ Z & P & Z & Z & P \\ N & Z & P & Z & Z \\ P & Z & Z & P & Z \\ Z & P & Z & Z & P \end{array}\right] \end{array}.$$

Let the transitive closure of the relevant matrix V be T, and let the transitive closure of the relationship matrix R be L. Then, according to Definition 3.3 and Definition 3.6, we can get T and L shown as follows:

$$T = \begin{array}{cc} & \begin{array}{ccccc} c_1 & c_2 & c_3 & c_4 & c_5 \end{array} \\ \begin{array}{c} c_1 \\ c_2 \\ c_3 \\ c_4 \\ c_5 \end{array} & \left[\begin{array}{ccccc} \tilde{1} & \tilde{0} & \widetilde{0.7} & \widetilde{0.6} & \tilde{0} \\ \tilde{0} & \tilde{1} & \tilde{0} & \tilde{0} & \widetilde{0.9} \\ \widetilde{0.7} & \tilde{0} & \tilde{1} & \widetilde{0.6} & \tilde{0} \\ \widetilde{0.6} & \tilde{0} & \widetilde{0.6} & \tilde{1} & \tilde{0} \\ \tilde{0} & \widetilde{0.9} & \tilde{0} & \tilde{0} & \tilde{1} \end{array}\right] \end{array}$$

$$L = \begin{array}{cc} & \begin{array}{ccccc} c_1 & c_2 & c_3 & c_4 & c_5 \end{array} \\ \begin{array}{c} c_1 \\ c_2 \\ c_3 \\ c_4 \\ c_5 \end{array} & \left[\begin{array}{ccccc} P & Z & N & P & Z \\ Z & P & Z & Z & P \\ N & Z & P & N & Z \\ P & Z & N & P & Z \\ Z & P & Z & Z & P \end{array}\right] \end{array}.$$

Let the document descriptor relevant matrix and the document descriptor relationship matrix used to model the local fuzzy-valued concept network on host 1 be E_1 and F_1, respectively, and let the document descriptor relevant matrix and the document descriptor relationship matrix used to model the local fuzzy-valued concept network on host 2 be E_2 and F_2, respectively. Then, from Figs. 7 and 8, we can get E_1, F_1, E_2, and F_2, shown as follows:

$$E_1 = \begin{array}{cc} & \begin{array}{ccccc} c_1 & c_2 & c_3 & c_4 & c_5 \end{array} \\ \begin{array}{c} d_1 \\ d_2 \\ d_3 \end{array} & \left[\begin{array}{ccccc} \tilde{0} & \tilde{0} & \widetilde{0.8} & \tilde{1} & \tilde{0} \\ \tilde{0} & \tilde{0} & \tilde{0} & \widetilde{0.9} & \widetilde{0.6} \\ \tilde{0} & \tilde{0} & \tilde{0} & \tilde{0} & \widetilde{0.8} \end{array}\right] \end{array}$$

$$F_1 = \begin{array}{cc} & \begin{array}{ccccc} c_1 & c_2 & c_3 & c_4 & c_5 \end{array} \\ \begin{array}{c} d_1 \\ d_2 \\ d_3 \end{array} & \left[\begin{array}{ccccc} Z & Z & N & P & Z \\ Z & Z & Z & P & P \\ Z & Z & Z & Z & P \end{array}\right] \end{array}$$

$$E_2 = \begin{array}{cc} & \begin{array}{ccccc} c_1 & c_2 & c_3 & c_4 & c_5 \end{array} \\ \begin{array}{c} d_4 \\ d_5 \\ d_6 \end{array} & \left[\begin{array}{ccccc} \tilde{0} & \tilde{0} & \widetilde{0.9} & \widetilde{0.8} & \tilde{0} \\ \tilde{0} & \tilde{0} & \tilde{0} & \widetilde{0.6} & \widetilde{0.9} \\ \tilde{0} & \tilde{0} & \tilde{0} & \tilde{0} & \tilde{1} \end{array}\right] \end{array}$$

$$F_2 = \begin{array}{cc} & \begin{array}{ccccc} c_1 & c_2 & c_3 & c_4 & c_5 \end{array} \\ \begin{array}{c} d_4 \\ d_5 \\ d_6 \end{array} & \left[\begin{array}{ccccc} Z & Z & P & N & Z \\ Z & Z & Z & P & P \\ Z & Z & Z & Z & P \end{array}\right] \end{array}.$$

Let $E_1^* = E_1 \odot T$, then E_1^* contains the implicit relevant degrees between documents and concepts of the local fuzzy-valued concept networks in host 1. Let $F_1^* = F_1 * L$, then F_1^* contains the implicit relationships between documents and concepts of the local fuzzy-valued concept networks in host 1. Let $E_2^* = E_2 \odot T$, then E_2^* contains the implicit relevant degrees between documents and concepts of the local fuzzy-valued concept networks in host 2. Let $F_2^* = F_2 * L$, then F_2^* contains the implicit relationships between documents and concepts of the local fuzzy-valued concept networks in host 2, where E_1^*, F_1^*, E_2^*, and F_2^* are shown as follows:

$$E_1^* = \begin{array}{cc} & \begin{array}{ccccc} c_1 & c_2 & c_3 & c_4 & c_5 \end{array} \\ \begin{array}{c} d_1 \\ d_2 \\ d_3 \end{array} & \left[\begin{array}{ccccc} \widetilde{0.7} & \tilde{0} & \widetilde{0.8} & \tilde{1} & \tilde{0} \\ \widetilde{0.6} & \widetilde{0.6} & \widetilde{0.6} & \widetilde{0.9} & \widetilde{0.6} \\ \tilde{0} & \widetilde{0.8} & \tilde{0} & \tilde{0} & \widetilde{0.8} \end{array}\right] \end{array}$$

$$F_1^* = \begin{array}{cc} & \begin{array}{ccccc} c_1 & c_2 & c_3 & c_4 & c_5 \end{array} \\ \begin{array}{c} d_1 \\ d_2 \\ d_3 \end{array} & \left[\begin{array}{ccccc} P & Z & N & P & Z \\ P & P & N & P & P \\ Z & P & Z & Z & P \end{array}\right] \end{array}$$

$$E_2^* = \begin{array}{cc} & \begin{array}{ccccc} c_1 & c_2 & c_3 & c_4 & c_5 \end{array} \\ \begin{array}{c} d_4 \\ d_5 \\ d_6 \end{array} & \left[\begin{array}{ccccc} \widetilde{0.7} & \tilde{0} & \widetilde{0.9} & \widetilde{0.8} & \tilde{0} \\ \widetilde{0.6} & \widetilde{0.9} & \widetilde{0.6} & \widetilde{0.6} & \widetilde{0.9} \\ \tilde{0} & \widetilde{0.9} & \tilde{0} & \tilde{0} & \tilde{1} \end{array}\right] \end{array}$$

$$F_2^* = \begin{array}{cc} & \begin{array}{ccccc} c_1 & c_2 & c_3 & c_4 & c_5 \end{array} \\ \begin{array}{c} d_4 \\ d_5 \\ d_6 \end{array} & \left[\begin{array}{ccccc} N & Z & P & N & Z \\ P & P & N & P & P \\ Z & P & Z & Z & P \end{array}\right] \end{array}.$$

Then, E_1^*, F_1^*, E_2^*, and F_2^* form the basis for computing the similarities between documents and users' queries.

Assume that a user formulates his/her query expression in the fuzzy information retrieval system based on the network-type fuzzy-valued concept network on host 1 shown in Fig. 6. First, the user's query expression is handled by the method presented in Section IV. If the desired documents are not found in host 1, the system can choose other hosts from a list of hosts. Then, the user's query is sent to the other hosts chosen by the user automatically by the system. Assume host 2 is chosen, then the user's query is handled on host 2 to see if the desired documents are located on host 2. If the desired documents do not exist on host 2, then the other hosts are chosen to process the user's query. The above processes are done repetitively until the desired documents are found or they do not exist.

Example 5.2: As in Example 5.1, the fuzzy-valued concept network in host 1 and the fuzzy-valued concept network in host 2 are shown in Figs. 7 and 8, respectively. Assume that the user sets his/her query first in host 1 and hopes the retrieved document should contain concept 2 (the degree of strength is about 0.8) and concept 4 (the degree of strength is about 0.9), but should not contain concept 5 (the degree of strength is about 1). Then, the user's query Q can be represented by a query descriptor relevant vector $\overline{qv_1}$ and query descriptor relationship vector $\overline{qr_1}$ shown as follows:

$$\overline{qv_1} = \langle \overset{c_1}{-} \quad \overset{c_2}{\widetilde{0.8}}, \quad \overset{c_3}{-}, \quad \overset{c_4}{\widetilde{0.9}}, \quad \overset{c_5}{1} \rangle$$
$$\overline{qr_1} = \langle \overset{c_1}{-}, \quad \overset{c_2}{P}, \quad \overset{c_3}{-}, \quad \overset{c_4}{P}, \quad \overset{c_5}{N} \rangle.$$

Then, based on the results of Example 5.1 and (10), we can get

$$RS(d_1) = \frac{0 + 0.9 + 0}{3} = \frac{0.9}{3} = 0.3$$
$$RS(d_2) = \frac{0.8 + 1 + 0}{3} = \frac{1.8}{3} = 0.6$$
$$RS(d_3) = \frac{1 + 0 + 0}{3} = \frac{1}{3} = 0.333.$$

Assume that the information retrieval threshold value $\alpha = 0.3$, then the sequential order from the highest retrieval status value to that with the lowest retrieval status value is $d_2 > d_3 > d_1$. In this case, document d_2 is the best choice for the user's query.

Assume that the desired documents are not found on host 1 (although they exist on host 1 in this example), and assume that the user's query is sent to host 2. Then, based on the results of Example 5.1 and (10), we can get

$$RS(d_4) = \frac{0 + 0 + 0}{3} = \frac{0}{3} = 0$$
$$RS(d_5) = \frac{0.9 + 0.7 + 0}{3} = \frac{1.6}{3} = 0.533$$
$$RS(d_6) = \frac{0.9 + 0 + 0}{3} = \frac{0.9}{3} = 0.3.$$

Assume that the information retrieval threshold value $\alpha = 0.3$, then the sequential order from the highest retrieval status value to that with the lowest retrieval status value is $d_5 > d_6 > d_4$. In this case, document d_5 is the best choice for the user's query, and document d_4 will not be retrieved in this example due to the fact that its degree of satisfaction is smaller than 0.3.

VI. Conclusion

In this paper, we have presented a new method for document retrieval using fuzzy-valued concept networks. The fuzzy-valued concept networks allow the values of the relevant degree between concepts to be arbitrary shapes of fuzzy numbers, and the relevant relationships between concepts not only to be fuzzy positive association relationship but also fuzzy negative association relationship. The fuzzy information retrieval systems based on the fuzzy-valued concept networks can be designed in a more flexible and more intelligent manner. Moreover, we also allow the users' queries to be represented by arbitrary shapes of fuzzy numbers and to use fuzzy positive association relationship and fuzzy negative association relationship for formulating their queries for increasing the flexibility of fuzzy information retrieval systems. Furthermore, we also extend the proposed fuzzy-valued concept network architecture to the network-type fuzzy-valued concept networks and present a fuzzy information retrieval method based on the network-type fuzzy-valued concept networks in the Internet environment.

References

[1] G. Bordogna, P. Carrara, and G. Pasi, "Query term weights as constraints in fuzzy information retrieval," *Inf. Process. Manage.*, vol. 27, no. 1, pp. 15–26, 1991.

[2] S. M. Chen, "A fuzzy reasoning technique based on the α-cuts operations of fuzzy numbers," in *Proc. 2nd Int. Conf. Automation Technology*, vol. 3, Taipei, Taiwan, R.O.C., July 1992, pp. 147–154.

[3] S. M. Chen and J. Y. Wang, "Document retrieval using knowledge-based fuzzy information retrieval techniques," *IEEE Trans. Syst., Man, Cybern.*, vol. 25, pp. 793–803, May 1995.

[4] S. M. Chen, W. H. Hsiao, and Y. J. Horng, "A knowledge-based method for fuzzy query processing for document retrieval," *Cybern. Syst.*, vol. 28, pp. 99–119, Jan./Feb. 1997.

[5] S. M. Chen and Y. J. Horng, "Fuzzy query processing for document retrieval based on extended fuzzy concept networks," *IEEE Trans. Syst., Man, Cybern. B*, vol. 29, pp. 126–135, Feb. 1999.

[6] D. E. Comer, *Computer Networks and Internets*. Englewood Cliffs, NJ: Prentice-Hall, 1997.

[7] J. Giarratano and G. Riley, *Expert Systems: Principles and Programming*. Boston, MA: PWS, 1994.

[8] G. T. Her and J. S. Ke, "A fuzzy information retrieval system model," in *Proc. 1983 Nat. Computer Symp.*, Taiwan, R.O.C., Dec. 1983, pp. 147–1551.

[9] Y. J. Horng and S. M. Chen, "Document retrieval based on extended fuzzy concept networks," in *Proc. 4th Nat. Conf. Defense Management*, vol. 2, Taipei, Taiwan, R.O.C., Mar. 1996, pp. 1039–1050.

[10] ——, "Finding inheritance hierarchies in fuzzy-valued concept networks," *IEEE Trans. Syst., Man, Cybern. B*, vol. 29, pp. 126–135, Feb. 1999.

[11] Y. J. Horng, S. M. Chen, and C. H. Lee, "A fuzzy information retrieval method using fuzzy-valued concept networks," in *Proc. 10th Int. Conf. Tolls Artificial Intelligence*, Taipei, Taiwan, R.O.C., Dec. 1998, pp. 104–111.

[12] A. Kandel, *Fuzzy Mathematical Techniques with Applications*. Reading, MA: Addison-Wesley, 1986.

[13] M. Kracker, "A fuzzy concept network model and its applications," in *Proc. 1st IEEE Int. Conf. Fuzzy Systems*, Mar. 1992, pp. 761–768.

[14] M. Kamel, B. Hadfield, and M. Ismail, "Fuzzy query processing using clustering techniques," *Inf. Process. Manage.*, vol. 26, no. 2, pp. 279–293, 1990.

[15] D. Lucarella and R. Morara, "FIRST: Fuzzy information retrieval system," *J. Inf. Sci.*, vol. 17, no. 1, pp. 81–91, 1991.

[16] S. Miyamoto, "Information retrieval based on fuzzy associations," *Fuzzy Sets Syst.*, vol. 38, pp. 191–205, 1990.

[17] T. Murai, M. Miyakoshi, and M. Shimbo, "A fuzzy document retrieval method based on two-valued indexing," *Fuzzy Sets Syst.*, vol. 30, pp. 103–120, 1989.

[18] T. Radechi, "Generalized Boolean methods of information retrieval," *Int. J. Man–Mach. Stud.*, vol. 18, no. 5, pp. 409–439, 1983.

[19] G. Salton and M. J. Mcgill, *Introduction to Modern Information Retrieval*. New York: McGraw-Hill, 1983.

[20] V. Tahani, "A fuzzy model of document retrieval system," *Inf. Process. Manage.*, vol. 12, pp. 177–187, 1976.

[21] A. S. Tanenbaum, *Computer Networks*. Englewood Cliffs, NJ: Prentice-Hall, 1989.

[22] L. A. Zadeh, "Fuzzy sets," *Inf. Contr.*, vol. 8, pp. 338–353, 1965.