International Journal of Innovative
Computing, Information and Control
Volume **9**, Number **5**, May **2013**
 ISSN 1349-4198
pp. 1881–1888

A COMPARISON OF TWO TECHNIQUES FOR PRELIMINARY DIAGNOSIS OF PRIMARY HEADACHES

Jeong Yong Ahn

Department of Statistics (Institute of Applied Statistics)
Chonbuk National University
567 Baekje-daero, Jeonju, Jeonbuk 561-756, Korea
jyahn@jbnu.ac.kr

Received February 2012; revised June 2012

Abstract. *This paper presents an evaluation and a comparison of two diagnosis techniques, the max-min composition method, which especially has the advantage that it is easy to implement, and the distance method, which is used to decrease loss of information. The simulation results show that the two techniques have a significant difference in diagnostic results when we use only the largest value of the max-min composition as a diagnostic measure. However, there is no significant difference when we use the values of more than 0.5 in the method.*

Keywords: Diagnosis techniques, Distance method, Max-min composition, Medical diagnosis

1. **Introduction.** Most of our real life problems in many fields including medical science often involve data which is not necessarily crisp, precise, and deterministic due to various uncertainties associated with these problems [12]. Fuzzy set theory has a number of properties that make it suitable for formalizing uncertain information upon which medical diagnosis and treatment is usually based. It allows us to define inexact medical entities as fuzzy sets [14]. In addition, fuzzy logic enables the formal treatment of making approximate inferences on the basis of fuzzy data and fuzzy relations, and fuzzy relations enable the representation of uncertain medical associations [2, 6].

There are two popular techniques for medical diagnosis using Sanchez's fuzzy relation approach. One is the method that uses the max-min composition rule. The other is the method that uses the distance measure between fuzzy sets for diagnosis of patients. Each method has its own strengths and weaknesses, respectively, and a large number of applications are well researched for the methods [3, 7, 13, 25]. De et al. [13] applied the max-min composition rule to determine the disease of patients as an application of the Sanchez's approach. This max-min method is very visible because it is an intuitive recipe, and easy to work in practice. The method, however, has been known to lead to quite conservative results and to a loss of information because the composition neglects most values except for extreme ones. Szmidt and Kacprzyk [25] indicated the drawbacks of the diagnosis method based on the max-min-max composition rule, and proposed a diagnosis approach based on the distances between diseases and symptoms. The distance method can reduce the loss of information, and make it possible to introduce weights for all symptoms. However, the method also has a drawback in that it provides an unclear diagnosis when the difference in the distances is not large.

Most of these studies introduced each diagnosis method and presented trivial examples using simple fuzzy data sets. However, there is still no valid research that compares with the effectiveness or difference of the diagnosis results of the techniques. The information is an important factor in medical diagnosis because the results can be helpful in the selection of a diagnostic technique. In addition, the results can address how to use the techniques in practice.

In this study, we present simulation results to compare the diagnosis outcome of two diagnosis techniques, max-min composition method and distance method. In Section 2 of this paper, we briefly review the related works on fuzzy sets and the application of intuitionistic fuzzy sets in medical diagnosis. In Section 3, we introduce two diagnosis techniques based on Sanchez's fuzzy relation. In Section 4, we illustrate the diagnosis result of the techniques by a numerical example. The example was applied to differentiate patients according to the three main types of primary headaches (migraine, tension, cluster headache) as an extension of our previous studies [3, 4]. In Section 5, we summarize the simulation results to compare the outcomes of the diagnosis techniques. We finish with a brief discussion in Section 6.

2. **Related Works.**

Fuzzy sets (FS) introduced by Zadeh [29] have been applied in various fields [20, 22, 24]. In the theory of FS, the degrees of membership and non-membership are a single value between 0 and 1. However, in reality, it may not always be certain that the sum of the degrees is just 1 [26]. Intuitionistic fuzzy sets (IFS), as a generalization of FS, have been proposed from understanding the fact [5], and have become a popular topic of investigation in the FS community [8, 23].

Since Zadeh [30] proposed to apply FS to the medical science field, FS theory has been utilized in many approaches to model the diagnostic process. Sanchez [21] represented the physician's medical knowledge as a fuzzy relation between symptoms and diseases. The approach was elaborated by Adlassnig [1] and applied in many studies such as [16, 19].

On the other hand, many approaches are proposed for medical diagnosis. Hung [15] proposed an improved nonprobabilistic entropy approach to support doctors examining the work of the preliminary diagnosing. Lin et al. [18] proposed a wavelet-based fuzzy neural network for classification and medical diagnosis. In addition, some approaches have been continuously proposed to model the diagnostic process [9, 28].

However, there have been no studies comparing the efficiency of the diagnosis techniques of diseases. In this study, we compare two popular diagnosis techniques using Sanchez's fuzzy relation. The features and main contributions of this study relative to previous works are as follows: First, two popular diagnosis techniques of primary headaches are evaluated and compared. The comparison results show that the two techniques have a significant difference in results. Second, we explored the difference in the diagnosis results of the two cases: using the largest value for diagnosis in the max-min composition method and using the values of more than 0.5. Third, we found that there was no significant difference in the diagnosis result of the two techniques when we used the values of more than 0.5 in the max-min method. Lastly, as a result, we recommend using the distance method or the max-min composition method with the values of more than 0.5 for diagnosis of primary headaches.

3. **Two Diagnosis Techniques.**

3.1. **The method based on max-min composition.**

In this section, we present an application of IFS theory in Sanchez's approach using fuzzy relations between symptoms and diseases for medical diagnosis [21].

Let $S = \{S_1, \dots, S_m\}$, $D = \{D_1, \dots, D_n\}$, and $P = \{P_1, \dots, P_q\}$ denote the sets of symptoms, diseases, and patients, respectively. Two fuzzy relations Q and R are defined as Equation (1) and Equation (2).

$$Q = \{< (p, s), \mu_Q(p, s), \nu_Q(p, s) >| \ (p, s) \in P \times S\} \tag{1}$$

$$R = \{< (s, d), \mu_R(s, d), \nu_R(s, d) >| \ (s, d) \in S \times D\} \tag{2}$$

where $\mu_Q(p, s)$ and $\nu_Q(p, s)$ indicate the degrees of the patient's symptoms, i.e., the degrees are the relationship between patient and symptoms (***patient's degrees***). In other words, $\mu_Q(p, s)$ indicates the degree to which the symptom s appears in patient p, and $\nu_Q(p, s)$ indicates the degree to which the symptom s does not appear in patient p. The amount $\pi_Q(p, s) = 1 - [\mu_Q(p, s) + \nu_Q(p, s)]$ is called the hesitation part, which may cater to membership value, non-membership value or both. Similarly, $\mu_R(s, d)$ and $\nu_R(s, d)$ are the relationship between symptoms and diseases (***confirmability degrees***), i.e., $\mu_R(s, d)$ is the degree to which symptom s confirms the presence of disease d, and $\nu_R(s, d)$ the degree to which the symptom s does not confirm the presence of disease d, respectively. Note that Q is defined on the set $P \times S$ and R on the set $S \times D$.

The composition T of R and Q ($T = R \circ Q$) for diagnosis of diseases describes the state of patients in terms of disease as a FR from P to D given by the membership function Equation (3) and non-membership function Equation (4).

$$\mu_T(p, d) = \max_s\{\min[\mu_Q(p, s), \mu_R(s, d)]\} \tag{3}$$

$$\nu_T(p, d) = \min_s\{\max[\nu_Q(p, s), \nu_R(s, d)]\} \tag{4}$$

for all $p \in P$ and $d \in D$.

However, in diagnosis with these max-min compositions, as the compositions were used when looking for T, the dominating symptoms were in fact only taken into account, Therefore, in the next step an improved version is calculated for which the following holds [13]:

$$\mu'_T = \mu_R - \nu_R \pi_R, \tag{5}$$

and Equations (3) and (4) are retained.

As a result, we generally determine the diagnostic labels of patient p for any disease d such that μ'_T is the largest. However, in some cases, the condition $\mu'_T >= 0.5$ is also used for the diagnosis of disease.

3.2. **The method based on distance.** In this section, we summarize an approach for the medical diagnosis of primary headaches originally proposed in Ahn et al. [4]. This approach is divided into four stages:

- Stage 1: Collect the patient's degrees and confirmability degrees for patient's symptoms. Confirmability degrees, the relationship between symptoms and diseases, are presented in the interview chart. The interview chart developed in our previous works [3, 17], and consisted of $23(M1 \sim M23), 17(T1 \sim T17)$ and $15(C1 \sim C15)$ items for the three types of headache (migraine, tension, and cluster), respectively. Table 1 is a part of the interview chart for migraine type. Each item has confirmability degrees with the relation among symptoms and the three types of headache, and has an interval-value in $[0, 1]$ as the degrees.

  Patient's degrees, the relationship between patient and symptoms, are assigned by a physician. In other words, confirmability degrees represent the general relationship between symptoms and diseases and patient's degrees represent a particular relationship between a patient and symptoms.
- Stage 2: Calculate the intuitionistic fuzzy weighted arithmetic average (IFWAA) of the patient's degrees and confirmability degrees, respectively, using the aggregate operator of Definition 3.1. A disease in general is presented through many symptoms

Table 1. Interview chart for migraine items

No	Items (Symptoms)	Confirmability degrees					
		migraine		tension		cluster	
		μ_{R_C}	ν_{R_C}	μ_{R_C}	ν_{R_C}	μ_{R_C}	ν_{R_C}
M1	Positive family history...	0.6	0.2	0.3	0.5	0.3	0.6
M2	At least five attacks...	0.9	0.1	0.2	0.7	0.2	0.7
M3	Headache lasting...	0.6	0.3	0.4	0.5	0.3	0.5
:	:	:	:	:	:	:	:
M23	Concurrent with...	0.7	0.2	0.3	0.6	0.3	0.5

Table 2. Patient P_1's degrees: $< \mu_Q(P_1, s), \nu_Q(P_1, s) >$

symptom	M10	M22	M23	T7	C2	C9	C11
μ_Q	0.6	0.7	0.5	0.5	0.5	0.7	0.5
ν_Q	0.2	0.1	0.1	0.3	0.3	0.1	0.2

and the symptoms significantly associated with the disease. Therefore, it is necessary to aggregate the symptoms. Aggregation of intuitionistic fuzzy information has attracted considerable interest from researchers in recent years [27].

- Stage 3: Calculate the distance between intuitionistic fuzzy sets using the measure of Definition 3.2 and the IFWAA calculated in Stage 2.
- Stage 4: Determine the disease of the patient, based on the distance. The lowest distance indicates the most appropriate diagnosis.

Definition 3.1. ***(IFWAA Operator)*** *Let $A = \{< x_i, \mu_A(x_i), \nu_A(x_i) >| \ i = 1, 2, \ldots, n\}$ be a collection of intuitionistic fuzzy values. Then, an IFWAA operator is defined as follows:*

$$IFWAA(A) = (1 - \Pi_{i=1}^{n}(1 - \mu_A(x_i))^{\omega_i}, \Pi_{i=1}^{n}(\nu_A(x_i))^{\omega_i}) \quad (6)$$

where $\omega = (\omega_1, \omega_2, \ldots, \omega_n)^T$ are the weight vectors of set A. In addition, $\omega_i > 0$ and $\sum_{i=1}^{n} \omega_i = 1$.

Definition 3.2. ***(Distance Measure)*** *For any two interval-valued intuitionistic fuzzy sets $A = \{< x_i, \mu_A(x_i), \ \nu_A(x_i) >| \ i = 1, 2, \ldots, n\}$ and $B = \{< x_i, \ \mu_B(x_i), \nu_B(x_i) >| \ i = 1, 2, \ldots, n\}$, the normalized Hamming distance considering the hesitate part is defined as follows:*

$$l_h(A, B) = (1/2n) \sum [\ |\mu_A(x_i) - \mu_B(x_i)| + |\nu_A(x_i) - \nu_B(x_i)| + |\pi_A(x_i) - \pi_B(x_i)|\ ] \quad (7)$$

where π is the degree of hesitation part, i.e., $\pi_A(x_i) = 1 - [\mu_A(x_i) + \nu_A(x_i)]$ and $\pi_B(x_i) = 1 - [\mu_B(x_i) + \nu_B(x_i)]$.

4. **Example.**

In this section, we present an example for diagnosing using the degrees of the patient's symptoms $< \mu_Q(p, s), \nu_Q(p, s) >$ assigned by a physician, and the confirmability degrees $< \mu_R(s, d), \nu_R(s, d) >$ indicated in the interview chart. In the example, we will show different results depending on which techniques were applied.

Let us consider patient P_1. P_1's symptoms are (M10, M22, M23) of migraine, (T7) of tension headache, and (C2, C9, C11) of cluster headache. Table 2 is the degrees for P_1's symptoms assigned by a physician, and Table 3 is the confirmability degrees indicated in the interview chart. Based on Table 2 and Table 3, Table 4 and Table 5 are calculated by applying IFWAA operator Equation (6).

TABLE 3. Confirmability degrees: $< \mu_R(s,d), \nu_R(s,d) >$

	migraine		tension		cluster	
symptom	μ_R	ν_R	μ_R	ν_R	μ_R	ν_R
M10	0.6	0.2	0.4	0.5	0.5	0.2
M22	0.8	0.1	0.2	0.7	0.2	0.7
M23	0.7	0.2	0.3	0.6	0.3	0.5
T7	0.3	0.6	0.6	0.3	0.4	0.5
C2	0.3	0.4	0.3	0.5	0.6	0.2
C9	0.2	0.3	0.3	0.3	0.7	0.2
C11	0.3	0.5	0.4	0.3	0.7	0.3

TABLE 4. Patient's degrees: (IFWAA μ_Q, IFWAA ν_Q)

Q	symptom M	symptom T	symptom C
P_1	(0.61, 0.13)	(0.50, 0.30)	(0.58, 0.18)

TABLE 5. Confirmability degrees: (IFWAA μ_R, IFWAA ν_R)

R	Migraine	Tension	Cluster
symptom M	(0.71, 0.16)	(0.30, 0.59)	(0.35, 0.41)
symptom T	(0.30, 0.60)	(0.60, 0.30)	(0.40, 0.50)
symptom C	(0.27, 0.39)	(0.34, 0.36)	(0.67, 0.23)

TABLE 6. Max-min composition for P_1's symptoms

	Migraine	Tension	Cluster
μ'_T	0.57	0.44	0.53

- **The result of max-min composition method**

Table 6 is calculated by applying Equations (3), (4) and then Equation (5) in Table 4 and Table 5. As a result, we can preliminarily diagnose that patient P_1 suffers most likely from migraine headache because μ'_T for migraine is the largest.

- **The result of distance method**

Table 7, the distance between intuitionistic fuzzy sets, is calculated by applying the distance measure Equation (7) to the data from Tables 4 and 5. In this method, as a result, we can preliminarily diagnose that patient P_1 suffers most likely from cluster headache because the distance of cluster is the lowest.

In this example, Tables 6 and 7 lead to a different diagnosis. However, the distance l_h for migraine is only slightly bigger than for cluster headache in Table 7 and the membership degree μ'_T for cluster has a value of more than 0.5. Therefore, additional diagnostic investigations may be considered for a more accurate diagnosis and we can preliminary diagnose that migraine and cluster headache might both be reasonable diagnoses for patient P_1. As a result, it is reasonable to assume that diagnostic result of the two techniques is a coincidence.

5. **Comparison.** In this section we present simulation results to compare the two techniques. The simulation is performed by using the parameters summarized in Table 8. The patient data such as Table 2 was simulated from conditions that occurred in medical

Table 7. Distance for P_1's symptoms

	Migraine	Tension	Cluster
l_h	0.22	0.30	0.21

Table 8. Simulation parameters

Parameters	Meaning	Value
n_1	the number of patients	$1,000$
n_2	the number of simulation runs	$1,000$
s	the number of symptoms	$3 \sim 20$
μ_Q	patient's membership degrees	$0.3 \sim 0.9$

Table 9. Coincidence rate of two methods: using the largest value in max-min composition as a diagnostic measure

N of Symptoms	3	4	5	7	10	13	16	20
Coincidence rate	0.540	0.505	0.474	0.435	0.421	0.426	0.436	0.448

Table 10. Coincidence rate of two methods: using the values of more than 0.5 in max-min composition

N of Symptoms	3	4	5	7	10	13	16	20
Coincidence rate	0.841	0.906	0.939	0.969	0.988	0.996	0.999	1.000

practice during our research. In general, physicians should assign values above 0.3 as a patient's membership degree μ_Q for symptoms that are present in this patient. Based on this additional information, we randomly generate the patient degrees.

Table 9 shows a part of the simulation results when we use the largest value of the max-min composition as a diagnostic measure. The numbers in the cells represent the coincidence of the diagnosis results of the two techniques. The first value of Table 9, 0.540, is the coincidence of the diagnosis results when a patient has three symptoms. Likewise, the value 0.448 is the coincidence of the diagnosis results when a patient has 20 symptoms. The average of coincidences is about 0.445. This result shows that the two techniques have a significant difference in the diagnostic results.

On the other hand, Table 10 shows a part of the simulation results when we use the values of more than 0.5 in the max-min composition. The average of coincidences is about 0.974. As a result, there is no significant difference in the diagnosis results of the two techniques when we use the values of more than 0.5 in the max-min method.

Figure 1 shows a plot of the simulation results. Using the largest value in the max-min composition, the diagnosis result shows a slight decrease of coincidence when the number of symptoms is increased. Using the values of more than 0.5, however, the coincidences are getting close to 1.000 when the number of symptoms is increased.

6. **Conclusion.** In recent years, fuzzy set theory and fuzzy logic have been applied successfully in medical diagnosis systems, and many techniques have been proposed. However, evaluation and comparison on the techniques is not sufficient.

In this paper, we compared two popular techniques, the max-min composition method and the distance method, for medical diagnosis of primary headaches. The main part

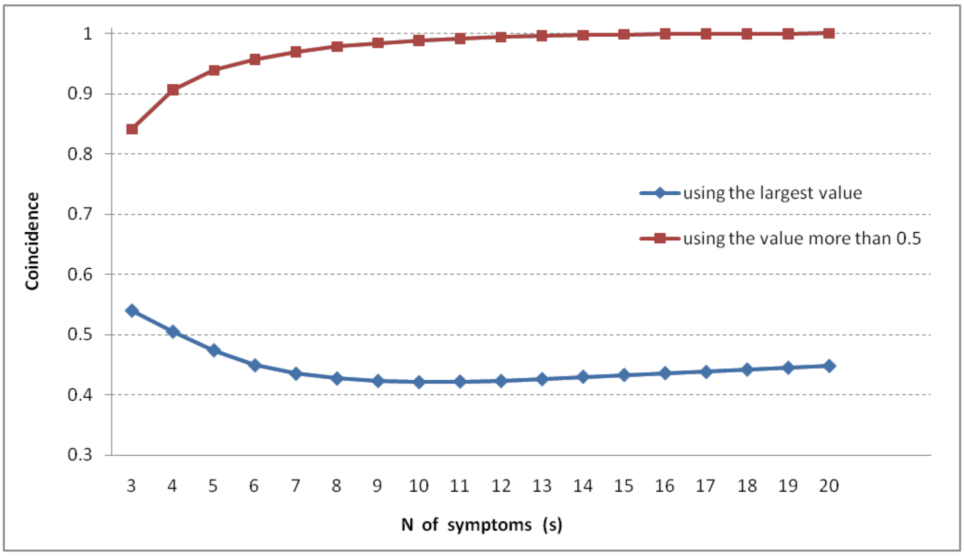

FIGURE 1. Coincidences of two methods

of this paper contains a simulation study to compare the diagnosis outcome. The results of the simulation study show that the two techniques have a significant difference in diagnostic results when we use only the largest value in the result of the max-min composition method. However, there is no significant difference if we use the values of more than 0.5 in the max-min method. As a result of this study, we recommend using the distance method or the max-min composition method with the values of more than 0.5 for diagnosis of primary headaches. We expect that the results of this study can be helpful in the selection of a diagnostic technique.

There are several remaining problems for future explorations. First, practical comparison based on real data for a detailed evaluation of the techniques is needed. We are preparing real data with a physician. Also, more diagnosis techniques for generalization should be considered, even if the techniques considered in this study are the most common methods.

REFERENCES

[1] K. P. Adlassnig, Fuzzy set theory in medical diagnosis, *IEEE Trans. on Systems, Man and Cybernetics*, vol.16, pp.260-265, 1986.
[2] K. P. Adlassnig, G. Kolarz and W. Scheithauer, Present state of the medical expert system CADIAG-2, *Methods of Information in Medicine*, vol.24, pp.13-20, 1985.
[3] J. Y. Ahn, K. S. Mun, Y. H. Kim, S. Y. Oh and B. S. Han, A fuzzy method for medical diagnosis of headache, *IEICE Trans. on Information and Systems*, vol.E91-D, no.4, pp.1215-1217, 2008.
[4] J. Y. Ahn, K. S. Han, S. Y. Oh and C. D. Lee, An application of interval-valued intuitionistic fuzzy sets for medical diagnosis of headache, *International Journal of Innovative Computing, Information and Control*, vol.7, no.5(B), pp.2755-2762, 2011.
[5] K. Atanassov, Intuitionistic fuzzy sets, *Fuzzy Sets and Systems*, vol.20, pp.87-96, 1986.
[6] F. Aversa, E. Gronda, S. Pizzuti and C. Aragno, A fuzzy logic approach to decision support in medicine, *Proc. of the Conf. on Systemics, Cybernetics and Informatics*, 2002.
[7] L. Bobrowski, Induction of similarity measures and medical diagnosis support rules through separable, linear data transformations, *Methods of Information in Medicine*, vol.45, pp.200-203, 2006.
[8] J. J. Buckley, *Fuzzy Probability and Statistics*, Springer, 2006.
[9] T. Chaira and T. Chaira, Intuitionistic fuzzy sets: Application to medical image segmentation, *Studies in Computational Intelligence*, vol.85, pp.51-68, 2008.
[10] S. M. Chen, A weighted fuzzy reasoning algorithm for medical diagnosis, *Decision Support Systems*, vol.11, pp.37-43, 1994.
[11] S. M. Chen, Measures of similarity between vague sets, *Fuzzy Sets Systems*, vol.74, pp.217-223, 1995.

[12] B. Chetia and P. K. Das, An application of interval-valued fuzzy soft sets in medical diagnosis, *International Journal of Contemporary Mathematical Sciences*, vol.5, pp.1887-1894, 2010.
[13] S. K. De, R. Biswas and A. R. Roy, An application of intuitionistic fuzzy sets in medical diagnosis, *Fuzzy Sets and Systems*, vol.117, pp.209-213, 2001.
[14] M. Fathi-Torbaghan and D. Meyer, MEDUSA: A fuzzy expert system for medical diagnosis of acute abdominal pain, *Methods of Information in Medicine*, vol.33, pp.522-529, 1994.
[15] K. C. Hung, Medical pattern recognition: Applying an improved intuitionistic fuzzy cross-entropy approach, *Advances in Fuzzy Systems*, 2011.
[16] P. R. Innocent and R. I. John, Computer aided fuzzy medical diagnosis, *Information Sciences*, vol.162, pp.81-104, 2004.
[17] Y. H. Kim, S. K. Kim, S. Y. Oh and J. Y. Ahn, A fuzzy differential diagnosis of headache, *Journal of the Korean Data & Information Science Society*, vol.18, pp.429-438, 2007.
[18] C. J. Lin, C. Chen and C. Lee, Classification and medical diagnosis using wavelet-based fuzzy neural networks, *International Journal of Innovative Computing, Information and Control*, vol.4, no.3, pp.735-748, 2008.
[19] A. P. Rotshtein and H. B. Rakytyanska, Diagnosis problem solving using fuzzy relations, *IEEE Trans. on Fuzzy Systems*, vol.16, no.3, pp.664-675, 2008.
[20] T. Samatsu, K. Tachikawa and Y. Shi, GUI form for car retrieval systems using fuzzy theory, *ICIC Express Letters*, vol.2, no.3, pp.245-249, 2008.
[21] E. Sanchez, Medical diagnosis and composite fuzzy relations, in *Advances in Fuzzy Set Theory and Applications*, M. M. Gupta, R. K. Ragade and R. R. Yager (eds.), 1979.
[22] C. Schuh, Fuzzy sets and their application in medicine, *Proc. of the North American Fuzzy Information Society*, pp.86-91, 2005.
[23] R. Seising, A history of medical diagnosis using fuzzy relations, *Proc. in the Conf. on Fuzziness*, 2004.
[24] T.-S. Shih, J.-S. Su and J.-S. Yao, Fuzzy linear programming based on interval-valued fuzzy sets, *International Journal of Innovative Computing, Information and Control*, vol.5, no.8, pp.2081-2090, 2009.
[25] E. Szmidt and J. Kacprzyk, Intuitionistic fuzzy sets in intelligent data analysis for medical diagnosis, *Lecture Notes in Computer Science*, vol.2074, pp.263-271, 2001.
[26] W. Wang and X. Xin, Distance measure between intuiitionistic fuzzy sets, *Pattern Recognition Letters*, vol.26, pp.2063-2069, 2005.
[27] Z. Xu and X. Cai, Recent advances in intuitionistic fuzzy information aggregation, *Fuzzy Optimization and Decision Making*, vol.9, pp.359-381, 2010.
[28] K. Yamada, Diagnosis under compound effects and multiple causes by means of the conditional causal possibility approach, *Fuzzy Sets and Systems*, vol.145, pp.183-212, 2004.
[29] L. A. Zadeh, Fuzzy sets, *Information and Control*, vol.8, pp.338-353, 1965.
[30] L. A. Zadeh, Biological applications of the theory of fuzzy sets and systems, *Proc. of an International Symposium on Biocybernetics of the Central Nervous System*, pp.199-206, 1969.