Vol. 4, No. 4 April 2014 ISSN 2225-7217
ARPN Journal of Science and Technology

http://www.ejournalofscience.org

A New Algorithm for Similarity Measures to Pattern Recognition

Henry Chung-Jen Chao
Department of Traffic Science, Central Police University, Taoyuan, Taiwan
una051@mail.cpu.edu.tw

ABSTRACT

Park et al. (2007) published a paper that is related to similarity measures on intuitionistic fuzzy set which was published in Advances in Soft Computing. Park et al. used an example to reveal that Liang and Shi (2003) published in Pattern Recognition Letters sometimes cannot solve pattern recognition problems. We follow their trend to provide an example such that Park et al. (2007) and Liang and Shi (2003) both failed to decide the best pattern for the given sample and then we prepare our approach to create a new recognition algorithm that consists of two new similarity measures. By the same numerical example, we show that our proposed algorithm can solve the pattern recognition problem.

Keywords: *Pattern recognition problem; Similarity measures; Intuitionistic fuzzy set*

1. INTRODUCTION

Similarity measure is a powerful tool to solve pattern recognition problems. We study Liang and Shi (2003) and Park et al. (2007) to point out that sometimes their proposed similarity measures cannot help researchers to decide the best pattern for the proposed sample. Recently, there are many papers that worked on different research areas that related to pattern recognitions. For examples, Ahn et al. (2008), Atanassov et al. (2005), Chen and Li (2010), De et al. (2001), Hung et al. (2010), Li (2010), Li and Cheng (2002), Liu and Wang (2007), Mitchell (2003), Wang et al. (2009), Xu (2009), Xu and Chen (2008), and Yong et al. (2004).

2. REVIEW OF LIANG AND SHI (2003) AND PARK ET AL. (2007)

We recall the similarity measure proposed by Liang and Shi (2003),

$$S_e^p(A,B) = 1 - \frac{1}{\sqrt[p]{n}} \sqrt[p]{\sum_{i=1}^{n} \left(\varphi_{\mu AB}(i) + \varphi_{\nu AB}(i)\right)^p} \tag{1}$$

with

$\varphi_{\mu AB}(i) = |\mu_A(x_i) - \mu_B(x_i)|/2$,

$\varphi_{\nu AB}(i) = |\nu_A(x_i) - \nu_B(x_i)|/2$, for every $x_i \in X$, and the power index, $1 \le p < \infty$.

We construct the following example to illustrate that the similarity measure proposed by Liang and Shi (2003) cannot solve the pattern recognition problem with two patterns A_1 and A_2 was proposed, where

$$A_1 = \{(x_1, 0.4, 0.6), (x_2, 0.1, 0.5), (x_3, 0.4, 0.8)\}, \tag{2}$$

and

$$A_2 = \{(x_1, 0.1, 0.5), (x_2, 0.4, 0.5), (x_3, 0.2, 0.4)\}, \tag{3}$$

and one sample, B, where

$$B = \{(x_1, 0.25, 0.55), (x_2, 0.25, 0.5), (x_3, 0.3, 0.6)\}. \tag{4}$$

We found that for any p with $1 \le p < \infty$,

$$S_e^p(A_1, B) = 1 - \left((1/3)\left(0.1^p + 0.075^p + 0.15^p\right)\right)^{1/p} = S_e^p(A_2, B) \tag{5}$$

to demonstrate that Liang and Shi (2003) cannot decide the pattern of the given sample B.

Next, we recall that Park et al. (2007) provided their new similarity measures for the discrete case as follows,

$$S_g^p(A,B) = 1 - \frac{1}{\sqrt[p]{n}} \sqrt[p]{\sum_{i=1}^{n} \left(\varphi_{\mu AB}(i) + \varphi_{\nu AB}(i) + \varphi_{\pi AB}(i)\right)^p} \tag{6}$$

where $A = \{(x, \mu_A(x_i), \nu_A(x_i)) : i = 1, \ldots, n\}$ and $B = \{(x, \mu_B(x_i), \nu_B(x_i)) : i = 1, \ldots, n\}$,

$\varphi_{\pi AB}(i) = \dfrac{|\pi_A(x_i) - \pi_B(x_i)|}{2}$, with $\varphi_{\mu AB}(i)$ and $\varphi_{\nu AB}(i)$ in equation (1). For the continuous case, they assumed that

$$S_k^p(A,B) = 1 - \frac{1}{\sqrt[p]{n}} \sqrt[p]{\int_a^b \left(\varphi_{\mu AB}(x) + \varphi_{\nu AB}(x) + \varphi_{\pi AB}(x)\right)^p} \tag{7}$$

where $A = \{(x, \mu_A(x), \nu_A(x)) : x \in [a,b]\}$ and

$B = \{(x, \mu_B(x), \nu_B(x)) : x \in [a,b]\}$

Vol. 4, No. 4 April 2014 ISSN 2225-7217
ARPN Journal of Science and Technology

http://www.ejournalofscience.org

with

$$\varphi_{\mu AB}(x)=\frac{|\mu_A(x)-\mu_B(x)|}{2},$$
$$\varphi_{\nu AB}(x)=\frac{|\nu_A(x)-\nu_B(x)|}{2},$$
$$\varphi_{\pi AB}(x)=\frac{|\pi_A(x)-\pi_B(x)|}{2} \quad \text{for} \quad x\in[a,b] \quad \text{and}$$

$1\le p<\infty$, for the continuous case.

For the same numerical example, we evaluate S_g^p to point out that

$$S_g^p(A_1,B)=1-\left((1/3)\left(0.15^p+0.15^p+0.2^p\right)\right)^{1/p}=S_g^p(A_2,B) \tag{8}$$

that is independent of the value of p. The above example demonstrates that Park et al. (2007) cannot help researcher select the best suitable pattern for the sample B.

Consequently, in the next section, we will construct our approach to solve the dilemma appeared in equations (5) and (8).

3. OUR NEW ALGORITHM FOR SIMILARITY MEASURES

We will develop a new algorithm that consists of three similarity measure such that we can construct a series of examination to help decision makers solve pattern recognition problems. First, based on the traditional geometric formula of distance, we apply the following similarity measure,

$$S_1(A,B)=1-\left[\frac{1}{2}\sum_{j=1}^{n}w_j\left(\left|\mu_A(x_j)-\mu_B(x_j)\right|^{\alpha}+\left|\nu_A(x_j)-\nu_B(x_j)\right|^{\alpha}+\left|\pi_A(x_j)-\pi_B(x_j)\right|^{\alpha}\right)\right]^{1/\alpha} \tag{9}$$

where $w_j\ge 0$ is the weight for element x_j in the universe of discourse with $\sum_{j=1}^{n}w_j=1$. We should point out that our result from equation (9) corresponds to equation (6) of Park et al. (2007). The difference is that we used a first exponential, then sum by our approach as opposed to the first sum, then exponential approach of Park et al. (2007). If we consider the special case when $\alpha=1$, $p=1$ and $w_j=1/n$, for $j=1,...,n$, then the two similarity measures of equations (6) and (9) are identical.

Next, we will apply the cosine similarity measure to define the second similarity measure,

$$S_2(A,B)=\frac{\sum_{j=1}^{n}w_j\left(\mu_A(x_j)\cdot\mu_B(x_j)+\nu_A(x_j)\cdot\nu_B(x_j)+\pi_A(x_j)\cdot\pi_B(x_j)\right)}{\sqrt{\sum_{j=1}^{n}w_j\left(\mu_A^2(x_j)+\nu_A^2(x_j)+\pi_A^2(x_j)\right)}\sqrt{\sum_{j=1}^{n}w_j\left(\mu_B^2(x_j)+\nu_B^2(x_j)+\pi_B^2(x_j)\right)}} \tag{10}$$

with the same condition as equation (9).

Due to the fact that the arithmetic mean will create a differentiated problem, we modified the first similarity measure to construct our third similarity measure as follows,

$$S_3(A,B)=1-\left[\frac{\sum_{j=1}^{n}w_j\left(\left|\mu_A(x_j)-\mu_B(x_j)\right|^{\alpha}+\left|\nu_A(x_j)-\nu_B(x_j)\right|^{\alpha}+\left|\pi_A(x_j)-\pi_B(x_j)\right|^{\alpha}\right)}{\sum_{j=1}^{n}w_j\left(\left|\mu_A(x_j)+\mu_B(x_j)\right|^{\alpha}+\left|\nu_A(x_j)+\nu_B(x_j)\right|^{\alpha}+\left|\pi_A(x_j)+\pi_B(x_j)\right|^{\alpha}\right)}\right]^{1/\alpha} \tag{11}$$

with the same condition of equation (9) where the adjusted coefficient, $1/2$, is to restrict the total sum within one that will be appeared both in the numerator and denominator so they are cancelled out. We will now apply our new algorithm to the previous unsolvable pattern recognition problems.

4. OUR ALGORITHM FOR THE UNSOLVED DILEMMA

For the unsolved problem of Liang and Shi (2003) and Park et al. (2007), we apply our algorithm to evaluate that

$$S_1(A_1,B)=1-\left[0.5\left(w_1\left(3^{\alpha}+1+2^{\alpha}\right)+2w_2 3^{\alpha}+2^{\alpha}w_3\left(2+2^{\alpha}\right)\right)0.05^{\alpha}\right]^{1/\alpha} \tag{12}$$

ARPN Journal of Science and Technology

http://www.ejournalofscience.org

$= S_1(A_2, B)$,

for any selection of w_j for $j = 1,2,3$ with $w_j \geq 0$, for $j = 1,2,3$ and $\sum_{j=1}^{3} w_j = 1$, and any choice of α with $1 \leq \alpha < \infty$.

The above computation reveals that the first similarity measure cannot decide the best pattern for the given sample.

Next, with uniformly weight with $w_1 = w_2 = w_3 = 1/3$, we estimate $S_2(A_1, B)$ and $S_2(A_2, B)$ to imply that

$$S_2(A_1, B) = 0.937 < S_2(A_2, B) = 0.944, \tag{13}$$

to imply that the sample B should be assigned to pattern A_2.

5. CONCLUSIONS

In this paper, we have illustrated that the similarity measures of Liang and Shi (2003) and Park et al. (2007) are useless for our proposed example. Hence, we constructed our proposed algorithm that consists of three similarity measures that can improve the deficiency in the similarity measures of Liang and Shi (2003) and Park et al. (2007). Our proposed algorithm provides a new approach to solve pattern recognition problems that will help researchers to construct proper similarity measures for their investigated topic.

REFERENCES

[1] Ahn, J.Y., Mun, K.S., Kim, Y.H., Oh, S.Y., Han, B.S., 2008. A fuzzy method for medical diagnosis of headache. IEICE Transactions on Information and Systems E91-D (4), 1215-1217.

[2] Atanassov, K.T., 1984. Intuitionistic fuzzy sets. in: V. Sgurev (Ed), VII ITKR's Session. Sofia (June 1983 Central Sci and Techn Library Bulg Academy of Sciences).

[3] Atanassov, K.T., 1986. Intuitionistic fuzzy sets. Fuzzy Sets and Systems 20, 87–96.

[4] Atanassov, K.T., Pasi, G., Yager, R., 2005. Intuitionistic fuzzy interpretations of multi-criteria multi-person and multi-measurement tool decision making. International Journal of Systems Science 36 (14), 859-868.

[5] Chen, T.Y., Li, C.H., 2010. Determining objective weights with intuitionistic fuzzy entropy measures: a comparative analysis. Information Sciences 180 (21), 4207–4222.

[6] De, S.K., Biswas, R., Roy, A.R., 2001. An application of intuitionistic fuzzy sets in medical diagnosis. Fuzzy Sets and Systems 117 (2), 209–213.

[7] Hung, K.C., Wou, Y.W., Julian, P., 2010. Discussion on "A fuzzy method for medical diagnosis of headache". IEICE Transactions on Information and Systems E93-D (5), 1307-1308.

[8] Li, D.F., 2010. A ratio ranking method of triangular intuitionistic fuzzy numbers and its application to MADM problems. Computers and Mathematics with Applications 60 (6), 1557-1570.

[9] Li, D.F., Cheng, C.T., 2002. New similarity measures of intuitionistic fuzzy sets and application to pattern recognition. Pattern Recognition Letters 23, 221-225.

[10] Liang, Z., Shi, P., 2003. Similarity measures on intuitionistic fuzzy sets, Pattern Recognition Letters 24, 2687-2693.

[11] Liu, H.W., Wang, G.J., 2007. Multi-criteria decision-making methods based on intuitionistic fuzzy sets. European Journal of Operational Research 179 (1), 220-233.

[12] Mitchell, H.B., 2003. On the Dengfeng–Chuntian similarity measure and its application to pattern recognition. Pattern Recognition Letters 24 (16), 3101–3104.

[13] Park, J.H., Park, J.S., Kwun, Y.C., Lim, K.M., 2007. New similarity measures on intuitionistic fuzzy sets, Advances in Soft Computing 40, 22-30.

[14] Wang, Z., Li, K.W., Wang, W., 2009. An approach to multiattribute decision making with interval-valued intuitionistic fuzzy assessments and incomplete weights. Information Sciences 179 (17), 3026–3040.

[15] Xu, Z.S., 2009. Intuitionistic fuzzy hierarchical clustering algorithms. Journal of Systems Engineering and Electronics 20 (1), 90□97.

[16] Xu, Z.S., Chen, J., 2008. An overview of distance and similarity measures of intuitionistic fuzzy sets. International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems 16 (4), 529□555.

[17] Yong, D., Wenkang, S., Feng, D., Qi, L., 2004. A new similarity measure of generalized fuzzy numbers and its application to pattern recognition. Pattern recognition Letters 25, 875-883.