

Pattern Recognition Algorithm Based on Closeness Degree of Triangle Fuzzy Number

Yu-e Bao* and Erdun Bai

Inner Mongolia University of Nationalities, Inner Mongolia Tongliao, 028043, China

*Corresponding author

Abstract—In this paper, for determining recognition target class problem in inaccurate or imprecise environments, we designed triangular fuzzy number pattern recognition algorithm based on closeness degree. Firstly, the attribute values of each standard test set and the sample set are converted into triangular fuzzy numbers. Secondly, the *EW*-type closeness formula is used to calculate closeness degree between the attributes of each standard set and sample set, and establish the relative closeness degree matrix. Thirdly, the target identification of class model is determined based on the relative closeness degree matrix. By identifying 20 kinds of perfume test data sets with the tag to verify the validity and rationality of the proposed algorithm.

Keywords—*EW*-type closeness degree; triangular fuzzy numbers; relative approach degree; pattern recognition

I. INTRODUCTION

Pattern recognition can be understood as a process of identification of the input data. The purpose of fuzzy pattern recognition system is assigning each input data to a possible model class [1-4]. The different input values with similar characteristics can be grouped into the same category, enter a value different characteristics into different categories. Fuzzy pattern recognition can be divided into two types: fuzzy pattern recognition direct method (maximum membership principle) and fuzzy pattern recognition indirect method (degree of closeness close principle). Indirect fuzzy pattern recognition is to investigate a fuzzy set and fuzzy sets which are known's close degree are the best.

In Ref. [5], it uses closeness theory of fuzzy sets and interval number to study the closeness degree of fuzzy numbers, and gives type integral similarity measures can reasonably effectively reflect the degree of two fuzzy numbers close. Triangular fuzzy number is a special kind of fuzzy numbers, has good usability when used to describe the lack of data, data accuracy is low, etc [6-8]. We use type integral closeness degree of fuzzy numbers to discussion on the triangular fuzzy number pattern recognition algorithm. Establish a pattern recognition algorithm based on type triangle fuzzy numbers close degree. And as an application of algorithm identify 20 kinds of perfume test data set with labels.

II. PREREQUISITE

Let R be the real numbers field. A fuzzy set

$$u : R \rightarrow [0,1]$$

is called a fuzzy number if it is normal, convex, and upper semi-continuous and support set is a compact set. Let F_0 be the set of all fuzzy numbers, and is called fuzzy number space. In particular, for $u \in F_0$, if

$$u(x) = \begin{cases} \frac{x-a}{b-a}, & a \leq x \leq b \\ \frac{c-x}{c-b}, & b < x \leq c \\ 0, & \text{else} \end{cases}$$

then u is called the triangular fuzzy numbers, and denoted as

$$u = \langle a, b, c \rangle$$

For $u \in F_0$, the r -level set of u :

$$[u]^r = \begin{cases} \{x \in R : u(x) \geq r\}, & 0 < r \leq 1, \\ \text{supp } u, & r = 0. \end{cases}$$

is a nonempty bounded closed interval on the R , which is denoted as

$$[u]^r = [\underline{u}(r), \bar{u}(r)]$$

And

$$E([u]^r) = \frac{\underline{u}(r) + \bar{u}(r)}{2}$$

$$w([u]^r) = \frac{\bar{u}(r) - \underline{u}(r)}{2}$$

are respectively the expectation and width of bounded closed interval $[u]^r$.

Theorem 1.1^[5] Let mapping $N_{EW}^p : F_0 \times F_0 \rightarrow [0,1]$ defined as

$$N_{EW}^p(u, v) = \left(\int_0^1 \left(n_{EW}([u]^r, [v]^r) \right)^p dr \right)^{\frac{1}{p}} \quad (p \geq 1)$$

Where $u, v \in F_0$ and

$$n_{EW}([u]^r, [v]^r) = \frac{1}{1 + \left(\left| E([u]^r) - E([v]^r) \right|^2 + \frac{1}{3} |W([u]^r) - W([v]^r)|^2 \right)^{\frac{1}{2}}}$$

then

$$N_{EW}^p(u, v)$$

becomes closeness degree between fuzzy numbers u and v , and the closeness degree is called as EW – Type integral closeness degree.

III. TRIANGULAR FUZZY NUMBER PATTERN RECOGNITION ALGORITHM

A. Problem Description

In some consider inaccurate or imprecise environments, how to identify the target belongs to which one of l class.

B. Dataset Description

For each standard test set

$$C_k \quad (k = 1, 2, \dots, l)$$

we randomly choose $n(k)$ samples from each standard test set. By observation and evaluation of a property, we obtain c_i^k which is the observation of the i th sample in k th standard test set. Data set can be obtained as follows:

$$\begin{aligned} C_1 &: [c_1^1, c_2^1, c_3^1, \dots, c_{n(k)}^1] \\ C_2 &: [c_1^2, c_2^2, c_3^2, \dots, c_{n(k)}^2] \\ &\dots \dots \\ C_l &: [c_1^l, c_2^l, c_3^l, \dots, c_{n(k)}^l] \end{aligned}$$

For each test set

$$T_j \quad (j = 1, 2, \dots, m)$$

we randomly choose $\bar{n}(j)$ samples from each test set. By observation and evaluation of a property, we obtain t_i^j which is the observation of the i th of samples in j th standard test set of standards. Data set can be obtained as follows:

$$\begin{aligned} T_1 &: [t_1^1, t_2^1, t_3^1, \dots, t_{\bar{n}(j)}^1] \\ T_2 &: [t_1^2, t_2^2, t_3^2, \dots, t_{\bar{n}(j)}^2] \\ &\dots \dots \\ T_m &: [t_1^m, t_2^m, t_3^m, \dots, t_{\bar{n}(j)}^m] \end{aligned}$$

Next, a triangular fuzzy number pattern recognition algorithm based on the EW – type integral close degree is given.

Step 1.

For each

$$C_k \quad (k = 1, 2, \dots, l)$$

calculate the mean and standard deviation by the following formulas:

$$\mu^k = \frac{1}{n(k)} \sum_{i=1}^{n(k)} c_i^k \quad (k = 1, 2, \dots, l) \quad (1)$$

$$\sigma^k = \sqrt{\frac{1}{n(k)-1} \sum_{i=1}^{n(k)} (c_i^k - \mu^k)^2} \quad (k = 1, 2, \dots, l) \quad (2)$$

and construct triangular fuzzy number

$$u^k \quad (k = 1, 2, \dots, l)$$

where

$$u^k(x) = \begin{cases} \frac{x - (\mu^k - 2\sigma^k)}{2\sigma^k}, & \mu^k - 2\sigma^k \leq x \leq \mu^k \\ \frac{x - (\mu^k + 2\sigma^k)}{2\sigma^k}, & \mu^k < x \leq \mu^k + 2\sigma^k \\ 0, & \text{else} \end{cases} \quad (3)$$

Step 2.

For each T_j ($j = 1, 2, \dots, m$), calculate the mean and standard deviation by the following formulas:

$$\bar{\mu}^j = \frac{1}{n(j)} \sum_{i=1}^{n(j)} t_i^j \quad (j = 1, 2, \dots, m) \quad (4)$$

$$\bar{\sigma}^j = \sqrt{\frac{1}{n(j)-1} \sum_{i=1}^{n(j)} (t_i^j - \bar{\mu}^j)^2} \quad (j = 1, 2, \dots, m) \quad (5)$$

and construct triangular fuzzy number \bar{u}^j ($j = 1, 2, \dots, m$):

$$\bar{u}^j(x) = \begin{cases} \frac{x - (\bar{\mu}^j - 10\bar{\sigma}^j)}{10\bar{\sigma}^j}, & \bar{\mu}^j - 10\bar{\sigma}^j \leq x \leq \bar{\mu}^j \\ \frac{x - (\bar{\mu}^j + 10\bar{\sigma}^j)}{10\bar{\sigma}^j}, & \bar{\mu}^j < x \leq \bar{\mu}^j + 10\bar{\sigma}^j \\ 0, & \text{else} \end{cases} \quad (6)$$

Step 3.

Calculate

$$N_{EW}^p(\bar{u}^j, u^k) \quad (k = 1, 2, \dots, l) \quad (j = 1, 2, \dots, m)$$

which is the closeness degree between \bar{u}^j and u^k , and establish the closeness degree matrix $N_{m \times l}$. The element of j th row and k th column of $N_{m \times l}$ is

$$N_{EW}^p(\bar{u}^j, u^k) = \left(\int_0^1 \left(n_{EW} \left([\bar{u}^j]^r, [u^k]^r \right) \right)^p dr \right)^{\frac{1}{p}} \quad (p \geq 1, k = 1, 2, \dots, l, j = 1, 2, \dots, m)$$

Step 4.

Establish a relative closeness matrix $P_{m \times l}$. The element of j th row and k th column of $P_{m \times l}$ is

$$p(\bar{u}^j, u^k) = \frac{N_{EW}^p(\bar{u}^j, u^k)}{\sum_{k=1}^l N_{EW}^p(\bar{u}^j, u^k)} \quad (7)$$

which is the relative closeness degree for tags of j -test data equivalent to tags of k -standard data.

Step 5.

Analyzing tag test data, if

$$p(\bar{u}^j, u^{k_0}) = \max_{k=1,2,\dots,l} \{p(\bar{u}^j, u^k)\} \quad (8)$$

then the tags of j -test data equivalent to tags of k_0 -standard data.

IV. NUMERICAL EXPERIMENTS

In this section, we will introduce the experimental process of triangular fuzzy number pattern recognition algorithm based on the type integral Close degree, and compare it with other pattern recognition algorithms, and show the rationality and effectiveness of the algorithms.

Datasets are Perfume Data Set data sets which is the publicly available data sets [9] published in the UCI website. Data set is provided to the site by Professors Yousif Al-Bastaki in Kingdom of Bahrain and Professor Bekir KARLIK in Konya Turkey. The data was obtained from 20 different perfumes by using a handheld odor meter (OMX-GR sensor). Second record data, a total of 28 seconds. It includes standard data sets and test data sets and each data set have its labels.

Next, we use the triangular fuzzy number pattern recognition algorithm based on the EW -type integral proximity degree to recognize 20 kinds of perfume test data sets, and compare them with the original algorithm. The identification process is as follows:

Step 1.

Using the formulas (1), (2), (4), (5) to calculate the expectation and variance of standard data sets and test data sets with labels. See Table 1.

TABLE I. THE EXPECTATION AND VARIANCE OF STANDARD DATA SETS AND TEST DATA SETS WITH LABELS

	μ	σ	$\bar{\mu}$	$\bar{\sigma}$
1	64.4674	0.0703	64.2389	0.2157
2	60.7788	0.2111	61.2327	0.1808
3	57.8271	0.1748	56.7201	0.5227
4	71.8107	0.3596	71.8976	0.3042
5	68.1160	0.1001	68.1934	0.0361
6	70.2926	0.4163	70.0835	0.0356
7	60.4489	0.2395	70.0835	0.0356
8	68.8180	0.2275	68.3761	0.1945
9	72.2070	0.2374	71.8021	0.2181
10	73.9805	0.2556	72.4985	0.0730
11	62.8561	0.1270	63.0347	0.7765
12	46.0148	0.0000	47.0155	1.0381
13	82.4257	0.0354	82.3579	0.0003
14	79.4276	0.2540	79.9276	0.0688
15	85.0560	0.0000	85.0556	0.0000
16	71.6419	0.2381	72.2133	0.2487
17	66.5734	0.2790	68.6455	3.7527
18	64.1674	0.9844	62.0563	0.5925
19	66.8483	0.2388	66.2218	0.0000
20	65.6467	0.2596	66.2877	0.1279

Step 2.

Using the formulas (3) and (6) respectively to convert the data in Table 1 into triangular fuzzy numbers, and obtain Table 2.

TABLE II. TRIANGULAR FUZZY NUMBER OF STANDARD DATA SETS AND TEST DATA SETS WITH LABELS

	Standard data sets	test data sets
1	<63.7644, 64.4674, 65.1703>	<62.0817, 64.2389, 66.3960>
2	<58.6681, 60.7788, 62.8895>	<59.4244, 61.2327, 63.0410>
3	<56.0794, 57.8271, 59.5747>	<51.4931, 56.7201, 61.9471>
4	<68.2145, 71.8107, 75.4070>	<68.8552, 71.8976, 74.9401>
5	<67.1152, 68.1160, 69.1168>	<67.8324, 68.1934, 68.5545>
6	<66.1336, 70.2969, 74.4603>	<69.7277, 70.0835, 70.4393>
7	<58.0539, 60.4489, 62.8439>	<69.7277, 70.0835, 70.4393>
8	<66.5433, 68.8180, 71.0928>	<66.4312, 68.3761, 70.3210>
9	<69.8335, 72.2070, 74.5806>	<69.6210, 71.8021, 73.9832>
10	<71.4248, 73.9805, 76.5363>	<71.7685, 72.4985, 73.2285>
11	<61.5863, 62.8561, 64.1260>	<55.2702, 63.0347, 70.7993>
12	<46.0148, 46.0148, 46.0148>	<36.6348, 47.0155, 57.3962>

TABLE III. PATTERN RECOGNITION RESULTS

1	2	3	4	5	6	6	8	9	10	11	12	13	14	15	16	17	18	19	20
1	2	3	4	5	6	6	8	9	9	11	12	13	14	15	16	8	11	19	20

Step 5.

The first line is the original tags of test data sets, and the second line is the tags obtained by our proposed algorithm. The comparison results of Recognition accuracy rate between the proposed algorithm and other four pattern recognition algorithms can be seen in Table 4.

TABLE IV. COMPARISON OF RECOGNITION ACCURACY RATE

Name	Accuracy
Pattern recognition algorithm based on EW-type triangle fuzzy numbers close degree	85%
Bayes Classifier	70%
Logistic regression	70%
K-NN	65%
C4.5	55%

V. CONCLUSIONS

The design of triangular fuzzy number pattern recognition algorithm based on EW – type integral Close degree is based on data set of multiple measurements on each dimension of each sample, so the algorithm is better than other algorithms in data description.

ACKNOWLEDGMENTS

This work is supported by the National Science Fund of China (11461052)

REFERENCES

[1] Chen Wen, Zhang Qiang. The Partner Selection in Virtual Enterprises Based on Fuzzy Pattern Recognition. Fuzzy Systems and Mathematics. 2006, vo 2. pp.146-151.

13	<82.0714, 82.4257, 82.7800>	<82.3553, 82.3579, 82.3605>
14	<76.8879, 79.4276, 81.9673>	<79.2397, 79.9276, 80.6154>
15	<85.0560, 85.0560, 85.0560>	<85.0556, 85.0556, 85.0556>
16	<69.2609, 71.6419, 74.0228>	<69.7265, 72.2133, 74.7001>
17	<63.7839, 66.5734, 69.3630>	<31.1181, 68.6455, 106.1730>
18	<54.3235, 64.1674, 74.0112>	<56.1310, 62.0563, 67.9816>
19	<64.4603, 66.8483, 69.2362>	<66.2217, 66.2218, 66.2219>
20	<63.0504, 65.6467, 68.2431>	<65.0091, 66.2877, 67.5662>

Step 3.

Using the formula (7) to calculate closeness degree and establish closeness degree matrix $N_{20 \times 20}$.

Step 4.

Using the formula (8) to calculate relative closeness degree and establish relative closeness degree matrix $P_{20 \times 20}$.

By the above formulas and the relative closeness degree matrix (9), the Recognition results is in Table 3.

[2] Zhang Cai-qing, Lin Ming, Qi Ruo-nan. Atmospheric Quality Assessment in Coal-fired Power Plants Based on Fuzzy Pattern Recognition. Proceedings of the CSEE. 2009, vo 29. pp.30-34

[3] Zhou Li-juan, Xiao Man-sheng, Wen Zhi-qiang. Fuzzy identification method of incomplete text image based on grid closeness degree. Computer Engineering and Design. 2013, vo 12. pp. 4326-4329

[4] Shyi-Ming Chen, Shou-Hsiung Cheng, Tzu-Chun Lan. A novel similarity measure between intuitionistic fuzzy sets based on the centroid point so frans formed fuzzy numbers with applications to pattern cognition. Information Sciences. 2016, vo 343. pp.15-40

[5] Lei Yiming , Dai Bing , Bao Yu-e. The research of degree of closeness about fuzzy numbers. Pure and Applied Mathematics. 2016, VO 31. pp.611-619(in Chinese)

[6] Liu Sen, Zhao Xinan. Feature Recognition Based on individual advantage triangular fuzzy numbers. Statistics and Decision. 2014, VO 21. pp.22-25. (in Chinese)

[7] Liu Xiu-mei, Zhao Ke-qin, Wang Chuan-bin. New multiple attribute decision-making model with triangular fuzzy numbers based on connection numbers. Systems Engineering and Electronics. 2009, VO10. pp.2399-2403. (in Chinese)

[8] Yu Chun-hai, Fan Zhi-ping. A Clustering Method for Multiple Attribute Information with Triangular Fuzzy Numbers. Systems. Engineering Theory Methodology Applications. 2004, vo5. pp. 467-470

[9] <http://archive.ics.uci.edu/ml/datasets/Perfume+Data>