

Inconsistent Data Processing using Vague Set and Neutrosophic Set for Justifying Better Outcome

Soumitra De and Jaydev Mishra
Computer Science & Engineering Department
College of Engineering and Management, Kolaghat
Purba Medinipur, 721171, West Bengal, India
soumitra@cemk.ac.in, jsm03@cemk.ac.in

Abstract- Here, we have focused to process inconsistent data through imprecise queries from a database which consists of neutrosophic and vague set. Neutrosophic set is based on truth, indeterminacy and false membership values where as vague set is based on truth and false membership value. Firstly, we have applied two different similarity measure formulas to measure closeness of two neutrosophic data and vague data which in turn are used to get the similarity value for each row with the query for vague and neutrosophic set and the query is totally based on imprecise data. Next, we provide a similarity measure based α -cut value to generate a SQL command for the imprecise query which is act upon both set. The query return back different result sets for different α -cut values for checking the closeness between two different set. Finally, we have given suitable examples of vague and neutrosophic set where we established that neutrosophic set gives much more accurate result for any kind of uncertain query rather than vague set.

Keywords- *vague set; neutrosophic set; vague data; neutrosophic data; similarity measure; α -cut .*

I INTRODUCTION

Smarandache [1] introduced the neutrosophic set in 2001, which is used for handling problems involving imprecise data. A neutrosophic set, which is considered as next advancement of vague set [2], is based on truthness, indeterminacy and falseness instead of truth and false memberships used in vague sets. A vague set, which came from the fuzzy set [3], uses the thought of interval-based membership instead of point-based membership in fuzzy set. The three interval-based memberships in neutrosophic sets is more expressive for making decision than two interval based memberships in vague set.

Research works on neutrosophic set based applications in different fields are progressing fast

and we have used the neutrosophic set boundary as per our belief. A few applications using neutrosophic set has been published [4, 5, 6, 7, 8]. In the present work, we have attempted to use two different similarity measure formulas which are worked upon vague and neutrosophic data to perform uncertain queries. Several authors have used fuzzy or vague set [9, 10, 11, 12, 13] to execute imprecise queries but no such work has been reported literature using neutrosophic set. In this paper, our objectives are to execute a query using vague and neutrosophic data and both data are uncertain in nature. So, first we have used a similarity measure (S.M.) formula for finding similarity value of two neutrosophic data also we used other similarity measure formula for getting similarity value of two vague data. Then for these two different set of similarity measure values, we considered a tolerance (α -cut) value for which the given imprecise query is executed. For different α values, different output table will be generated from the main database. In this purpose, we have chosen a **Student** database to perform imprecise query using vague and neutrosophic data.

We have planned to organize the paper as follows. Definition of vague and neutrosophic set both are mentioned in Section II. Different Similarity measure formulas between two vague values and two neutrosophic values are defined in section III. Section IV executes an imprecise queries using similarity measure values of two different set which are came from vague and neutrosophic data oriented student database. Section V is the conclusion part of the paper.

II VAGUE AND NEUTROSOPHIC SET

Let U be the universe of discourse where an element of U is denoted by u .

Definition 1

A vague set is characterized by two membership functions given by:

- (i) truth membership function $t_x : U \rightarrow [0,1]$,
- (ii) false membership function $f_x : U \rightarrow [0,1]$ and

The rules are $t_x(u)+f_x(u) \leq 1$ and vague set is bounded by a subinterval $[t_x(u), 1-f_x(u)]$ of $[0,1]$ i.e, $t_x(u) \leq \mu_x(u) \leq 1-f_x(u)$

Definition 2

A neutrosophic set is characterized by three membership functions given by:

- (i) truth membership function $t_x : U \rightarrow [0,1]$,
- (ii) false membership function $f_x : U \rightarrow [0,1]$ and
- (iii) indeterminacy membership function $I_x : U \rightarrow [0,1]$ such that $t_x(u)+f_x(u) \leq 1$ and $t_x(u) + i_x(u) + f_x(u) \leq 2$.

III SIMILARITY MEASURE OF TWO NEUTROSOPHIC VALUES

In this paper, we have mentioned a similarity measure formula using neutrosophic data to measure closeness between two neutrosophic values which is given below.

Definition 3

Similarity Measure (S.M.) between two neutrosophic values:

Let x and y be any two neutrosophic values such that $x = [t_x, i_x, f_x]$ and $y = [t_y, i_y, f_y]$ where $0 \leq t_x \leq 1, 0 \leq i_x \leq 1, 0 \leq f_x \leq 1$ and $0 \leq t_y \leq 1, 0 \leq i_y \leq 1, 0 \leq f_y \leq 1$ with $0 \leq t_x + f_x \leq 1, 0 \leq t_y + f_y \leq 1, t_x + i_x + f_x \leq 2, t_y + i_y + f_y \leq 2$.

Let $SE(x, y)$ denote the similarity measure between x and y .

Then,

$$SE(x,y) = \sqrt{1 - \frac{|(t_x - t_y) - (i_x - i_y) - (f_x - f_y)|}{3}} \left(1 - \frac{|(t_x - t_y) + (i_x - i_y) + (f_x - f_y)|}{3} \right)$$

We are using the similarity measure formula [13] for

two vague data to find the similarity value between two vague data.

IV PROCESS IMPRECISE QUERY OF VAGUE AND NEUTROSOPHIC SET

In the current portion, we have given an example through which we have made two different column which are based on vague and neutrosophic value from a specific column value then using the formulas we got two different set of similarity measure value on which a given uncertain query have been worked and retrieved two different set of result as per the tolerance value is concerned. To demonstrate this fact, we have used the following **Student** relational database as given in Table I and uncertain query is mentioned below:

Stu_N	Stu_A (yrs)	Stu_M (100)
Rohit	18	57
Komal	20	79
Bikas	17	78
Titas	22	65
Milon	24	76
Ajit	21	80
Dipak	23	72

Table I. **Student Relation**

Uncertain query

“List the name of the students whose marks is very close to 78”.

Solution

Firstly, we have made the vague and neutrosophic attributes of student marks column (Stu_M). These two newly formed columns are used to finding the similarity measure between two vague values and two neutrosophic values. Here, vague and neutrosophic attribute of Stu_M is shown in the Table II.

Stu_N	Stu_M (100)	Vague form of Stu_M	Neutrosophic form of Stu_M
Rohit	57	<57, [44, 44]>	<57, [.44, .55, .52]>
Komal	79	<79, [.98, .98]>	<79, [.98, .025, .01]>
Bikas	78	<78, [1, 1]>	<78, [1, 0, 0]>
Titas	65	<65, [.80, .80]>	<65, [.8, .2, .18]>
Milon	76	<76, [.96, .96]>	<76, [.96, .10, .03]>
Ajit	80	<80, [.92, .92]>	<80, [.92, .15, .06]>
Dipak	72	<72, [.85, .85]>	<72, [.85, .3, .12]>

Table II. **Vague and Neutrosophic Representation of Stu_M attribute in Student Relation**

Next, we calculate the closeness value for vague and neutrosophic representation of Stu_M attribute separately which are shown in Table III.

For example, let us consider the two vague data $x = \langle 78, [1, 1] \rangle$ and $y = \langle 79, [.98, .98] \rangle$.

Here, $t_x = 1, f_x = 0, t_y = .98, f_y = .02$

Then

$$S.M.(x,y) = \sqrt{\left(1 - \frac{|(1-.98)-(0-.02)|}{2}\right) (1-|(1-.98)+(0-.02)|)}$$

$$= \sqrt{\left(1 - \frac{.04}{2}\right) (1-1.02)}$$

$$= \sqrt{1-.04} = \sqrt{0.98} = 0.989$$

Again, for $x = \langle 78, [1, 1] \rangle$ and $y = \langle 80, [.92, .92] \rangle$, $t_x = 1, f_x = 0, t_y = .92, f_y = .08$. This gives

$$S.M.(x,y) = \sqrt{\left(1 - \frac{|(1-.92)-(0-.08)|}{2}\right) (1-|(1-.92)+(0-.08)|)}$$

$$= \sqrt{\left(1 - \frac{.16}{2}\right) (1-1.08)}$$

$$= \sqrt{1-.16} = \sqrt{0.92} = 0.959$$

and so on.

For example, let us consider the two neutrosophic data $x = \langle 78, [1, 0, 0] \rangle$ and $y = \langle 79, [.98, .025, .01] \rangle$.

Here,

$t_x = 1, i_x = 0, f_x = 0, t_y = .98, i_y = .025, f_y = .01$

Then

$$S.M.(x,y) = \sqrt{\left(1 - \frac{|(1-.98)-(0-.025)-(0-.01)|}{3}\right) (1-|(1-.98)+(0-.025)+(0-.01)|)}$$

$$= \sqrt{\left(1 - \frac{.055}{3}\right) (1-.015)} = \sqrt{0.98166 \times 0.985} = \sqrt{.96693} = 0.983$$

Again, for $x = \langle 78, [1, 0, 0] \rangle$ and $y = \langle 80, [.92, .15, .06] \rangle$,

$t_x = 1, i_x = 0, f_x = 0, t_y = .92, i_y = .15, f_y = .06$.

This gives

$$S.M.(x,y) = \sqrt{\left(1 - \frac{|(1-.92)-(0-.15)-(0-.06)|}{3}\right) (1-|(1-.92)+(0-.15)+(0-.06)|)}$$

$$= \sqrt{\left(1 - \frac{.29}{3}\right) (1-.13)} = \sqrt{0.7859} = 0.886$$

and so on.

The calculated similarity measure for the both cases is shown in Table III.

Stu_N	Vague form of Stu_M	S.M with $\langle 78, [1,1] \rangle$	Neutrosophic form of Stu_M	S.M with $\langle 78, [1,0,0] \rangle$
Rohit	$\langle 57, [44,44] \rangle$.66	$\langle 57, [44,55,52] \rangle$.47
Komal	$\langle 79, [.98, .98] \rangle$.99	$\langle 79, [.98, .025, .01] \rangle$.98
Bikas	$\langle 78, [1,1] \rangle$	1	$\langle 78, [1,0,0] \rangle$	1
Titus	$\langle 65, [.80, .80] \rangle$.89	$\langle 65, [.8, .2, .18] \rangle$.81
Milon	$\langle 76, [.96, .96] \rangle$.98	$\langle 76, [.96, .10, .03] \rangle$.93
Ajit	$\langle 80, [.92, .92] \rangle$.96	$\langle 80, [.92, .15, .06] \rangle$.89
Dipak	$\langle 72, [.85, .85] \rangle$.92	$\langle 72, [.85, .3, .12] \rangle$.77

Table III. S. M. calculations of vague and neutrosophic attribute of Stu_M

Next, we process the same imprecise query for neutrosophic sets:

SELECT * FROM STUDENT WHERE S.M. \geq 0.9

With the tolerance value 0.90, the above SQL command of imprecise query for vague and neutrosophic set will be generated separately.

The resultant tuples thus retrieved are shown in Table IV and V for vague and neutrosophic data separately.

Stu_N	Stu_A	Stu_M
Komal	20	79
Bikas	17	78
Milon	24	76
Ajit	21	80
Dipak	23	72

Table IV. Resultant tuples of a given imprecise Query for vague set at T.V $\alpha = 0.90$

Stu_N	Stu_A	Stu_M
Komal	20	79
Bikas	17	78
Milon	24	76

Table V. Resultant tuples of a given imprecise Query for neutrosophic set at T.V $\alpha = 0.90$

So, from the outcome table IV and V we observed that the imprecise query has given less number of tuples for neutrosophic data rather than vague data and the output domain also very much closure for neutrosophic data rather than vague data for a certain tolerance value.

V CONCLUSION

Similarity Measure formula for calculating similarity value between two vague and neutrosophic data, is used but the formulas are differ in respect to number of parameters and their membership values. Here we have clearly seen that neutrosophic representation of data gives us better result than vague data as per imprecise query and the closeness of similarity values and data domain. SQL command is generated for the imprecise query which has returned back appropriate result from the database for neutrosophic data rather than vague data as per number of tuples came for outcome and the closeness of similarity measure values.

REFERENCES

- [1] F. Smarandache, "First International Conference on Neutrosophy, Neutrosophic Probability, Set, and Logic", University of New Mexico, Vol. 1, No. 3, 2001.
- [2] W. L. Gau, D. J. Buehrer "IEEE Trans. Syst. Man, Cybernetics", Vague Sets, Vol. 23, No. 2, pp. 610-614, 1993.
- [3] A. L. Zadeh, "Fuzzy Sets", Information and Control, Vol. 8, No. 3, pp. 338-353, 1965.
- [4] M. Arora, R. Biswas, "Deployment of Neutrosophic technology to retrieve answers for queries posed in natural language", in 3rd International Conference on Computer Science and Information Technology ICSSIT 2010, Vol. 3, pp. 435-439, 2010.
- [5] M. Arora, U.S.Pandey, "Supporting Queries with Imprecise Constraints in Total Neutrosophic Databases", Global Journal of Science Frontier Research Vol. 11 No. 8, 2011.
- [6] S. Broumi, "Generalized Neutrosophic Soft Set", International Journal of Computer Science, Engineering and Information Technology, Vol. 3, No. 2, pp. 17-30, 2013.
- [7] S. De, J. Mishra "Query Processing of Inconsistent Data using Neutrosophic Set", IEEE International Conference on Computing Communication and Automation, April 29 -30, Pno:120-124, 2016.
- [8] S. De, J. Mishra "Compare Different Similarity Measure Formula Based Imprecise Query on Neutrosophic Data", International Conference on Advanced Computing and Intelligent Engineering, December 21 -23, Pno:330-334, 2016.
- [9] F. Zhao, Z. M. Ma and L. Yan, "A Vague Relational Model and Algebra", Fourth International Conference on Fuzzy Systems and Knowledge Discovery (FSKD 2007), Vol. 1, pp. 81-85, 2007.
- [10] L. Zadeh, "Fuzzy Sets", Information and Control, Vol. 8, No. 3, pp. 338-353, 1965.
- [11] M. I. Sözat and A. Yazici "A complete axiomatization for fuzzy functional and multivalued dependencies in fuzzy database relations", Fuzzy Sets and Systems, Vol. 117, No. 2, pp.161-181, 2001.
- [12] F. Zhao and Z. M. Ma, "Vague Query Based on Vague Relational Model", AISC 61, pp. 229-238, Springer- Verlag Berlin Heidelberg, 2009.
- [13] J. Mishra, S. Ghosh, "Uncertain Query Processing using Vague Sets or Fuzzy Set: Which one is Better?", International Journal of Computers, Communications and Control (IJCCC), Vol. 9, No. 6, pp. 730-740, 2014.