Journal of Research in Science and Engineering (JRSE)
ISSN: 1656-1996 Volume-7, Issue-5, May 2025
DOI: 10.53469/jrse.2025.07(05).15

Fuzzy Rough Set Theory for Decision Support Systems

Mohammad Hasan Algarhy

Department of Computer Science & Engineering, Adikavi Nannaya University, Rajamahendravaram, Andhra Pradesh, India

Abstract: ***Decision Making is an art of obtaining optimal (sometimes satisfactory) solution of a problem. This procedure may not be always conventional or logical and sometimes it may involve irrational approaches, tacit knowledge, beliefs and faith. In this paper, we deal with the decision-making process when the datum is missing in the information system which possesses fuzzy decision attribute using the concept of rough sets.***

Keywords: Decision Making, rough sets, incomplete information systems, fuzziness

1. Introduction

For about a few decades, enormous work has been carried out in implementing fuzzy concepts in decision making because, due to complexity and constraints, we may not be able to arrive any crisp or clear decision through the available decision rules. In this regard, it is noteworthy to mention that Ganesan et.al. [1,2] proposed a mathematical frame work to deal with implementing the concept of rough approximations in the information systems with fuzzy decision attributes.

Whereas, in case of incomplete information systems [3,4,6], retrieval of the unknown value(s) is of complex in nature, as the missing value may be anything other than our predictions. Here, some of the common methods [5] which are in practise to fix the unknown values of the decision tables such as (a) Most Common Attribute Value Method (b) Maximum relative frequency method, or maximum conditional probability method (c) C4.5 Method (d) Event-Covering Method etc. In case of Incomplete Systems, there are several mathematical and statistical approaches as mentioned above are in practise to fix the unknown values of the records. However, all those methods do not deal with when, the decisions may be of fuzzy in nature.

2. Proposed Model

Considering this in view, Two methods are proposed here namely Fuzzy Similarity and Fuzzy Dissimilarity Approaches.

Here, we confine only with the information systems which possess only one unknown or missing value. For convenience, we name the records in which all the values are known as Complete Records and the records in which only one value is unknown or missing as incomplete records.

In All the methods, we propose the following Mathematical approach. For instance, the table is as follows:

	d_1	d_2	...	d_{j-1}	d_j	d_{j+1}	....	d_n
w_1	$x_{1,1}$	$x_{1,2}$	...	$x_{1,j-1}$	$x_{1,j}$	$x_{1,j+1}$	...	$x_{1,n}$
w_2	$x_{2,1}$	$x_{2,2}$	...	$x_{2,j-1}$	$x_{2,j}$	$x_{2,j+1}$	...	$x_{2,n}$
.	...	...	...	...	...	...	...	...
.	...	...	...	...	...	...	...	...
w_{i-1}	$x_{i-1,1}$	$x_{i-1,2}$	...	$x_{i-1,j-1}$	$x_{i-1,j}$	$x_{i-1,j+1}$	...	$x_{i-1,n}$
w_i	$x_{i,1}$	$x_{i,2}$	...	$x_{i,j-1}$	*	$x_{i,j+1}$	...	$x_{i,n}$
w_{i+1}	$x_{i+1,1}$	$x_{i+1,2}$	...	$x_{i+1,j-1}$	$x_{i+1,j}$	$x_{i+1,j+1}$	...	$x_{i+1,n}$
.	...	...	...	...	...	...	...	...
.	...	...	...	...	...	...	...	...
w_m	$x_{m,1}$	$x_{m,2}$	...	$m_{1,j-1}$	$x_{m,j}$	$x_{m,j+1}$	...	$x_{m,n}$

In the above table, $x_{i,j}$ is unknown and for which the value needs to be fixed or approximated in order to proceed the indexing algorithms which were discussed in the earlier chapters.

For any complete record w_t and the incomplete record w_i, the ***relative deviation*** (w_t, w_i) is given by

$$(w_t, w_i) = \frac{\sqrt{\left(x_{t,1} - x_{i,,1}\right) + \left(x_{t,2} - x_{i,,2}\right) + ... + \left(x_{t,j-1} - x_{i,,j-1}\right) + \left(x_{t,j+1} - x_{i,,j+1}\right) + \left(x_{t,j+2} - x_{i,,j+2}\right) + ... + \left(x_{t,n} - x_{i,,n}\right)}}{n-1}$$

The ***relative minimum quotient*** $\left\lfloor \frac{w_t}{w_i} \right\rfloor$ is given by

$$\left\lfloor \frac{w_t}{w_i} \right\rfloor = \min\left(\frac{x_{t,1}}{x_{i,1}}, \frac{x_{t,2}}{x_{i,2}}, ..., \frac{x_{t,j-1}}{x_{i,j-1}}, .\frac{x_{t,j+1}}{x_{i,j+1}}, ..., \frac{x_{t,n}}{x_{i,n}} \right)$$

The ***relative maximum quotient*** $\left\lceil \frac{w_t}{w_i} \right\rceil$ is given by

$$\left\lceil \frac{w_t}{w_i} \right\rceil = \max\left(\frac{x_{t,1}}{x_{i,1}}, \frac{x_{t,2}}{x_{i,2}}, ..., \frac{x_{t,j-1}}{x_{i,j-1}}, .\frac{x_{t,j+1}}{x_{i,j+1}}, ..., \frac{x_{t,n}}{x_{i,n}} \right)$$

For Example, consider a 4 tuple complete record A(4,6,5,3) and a 4 tuple incomplete record B(7,*,3,8).

The relative deviation (A,B) is given by

$$(A, B) = \frac{\sqrt{(4-7)^2 + (5-3)^2 + (3-8)^2}}{3} = 2.0548$$

The relative minimum quotient $\left\lfloor \frac{A}{B} \right\rfloor$ is given by

$$\left\lfloor \frac{A}{B} \right\rfloor = \min\left(\frac{4}{7}, \frac{5}{3}, \frac{3}{8}\right) = \min(0.57, 1.67, 0.38) = 0.38$$

The relative maximum quotient $\left\lceil \frac{A}{B} \right\rceil$ is given by

$$\left\lceil \frac{A}{B} \right\rceil = \max\left(\frac{4}{7}, \frac{5}{3}, \frac{3}{8}\right) = \max(0.57, 1.67, 0.38) = 1.67$$

Now, using these mathematical tools, we shall look into the methods of fixing/ approximating the unknown/missing value in the incomplete decision tables which possess fuzzy or intuitionistic fuzzy decision attributes.

2.1 Incomplete Information Systems with Fuzzy Decision Attributes

Consider the following decision Table with E as the decision variable.

	d_1	**d_2**	**...**	**d_{j-1}**	**d_j**	**d_{j+1}**	**....**	**d_n**	**μ_E**
w_1	$x_{1,1}$	$x_{1,2}$	...	$x_{1,j-1}$	$x_{1,j}$	$x_{1,j+1}$	...	$x_{1,n}$	$\mu_E(w_1)$
w_2	$x_{2,1}$	$x_{2,2}$	...	$x_{2,j-1}$	$x_{2,j}$	$x_{2,j+1}$	...	$x_{2,n}$	$\mu_E(w_2)$
.	...	...	...	...	...	...	...	...	.
.	...	...	...	...	...	...	...	...	.
w_{i-1}	$x_{i-1,1}$	$x_{i-1,2}$	...	$x_{i-1,j-1}$	$x_{i-1,j}$	$x_{i-1,j+1}$	...	$x_{i-1,n}$	$\mu_E(w_{i-1})$
w_i	$x_{i,1}$	$x_{i,2}$	...	$x_{i,j-1}$	*	$x_{i,j+1}$	...	$x_{i,n}$	$\mu_E(w_i)$
w_{i+1}	$x_{i+1,1}$	$x_{i+1,2}$	...	$x_{i+1,j-1}$	$x_{i+1,j}$	$x_{i+1,j+1}$	...	$x_{i+1,n}$	$\mu_E(w_{i+1})$
.	...	...	...	...	...	...	...	...	.
.	...	...	...	...	...	...	...	...	.
w_m	$x_{m,1}$	$x_{m,2}$	...	$m_{1,j-1}$	$x_{m,j}$	$x_{m,j+1}$	...	$x_{m,n}$	$\mu_E(w_m)$

Here, using strong threshold approach, the decision attribute can be converted into 1 or 0. Obviously, the incomplete record too holds the decision value either 1 or 0. The complete records which hold the decision as same as that of incomplete record, they are said to be similar records with respect to that incomplete record w_i and that collection is denoted as ***SIM(w_i)*** and the other complete records are called the dissimilar records with respect to that incomplete record w_i and that collection is denoted as ***DISSIM(w_i)***.

For example, consider the following decision table which possesses only the Boolean decisions:

	d_1	**d_2**	**d_3**	**d_4**	**Decision(E)**
w_1	40	30	15	56	1
w_2	45	70	32	56	0
w_3	12	10	*	42	1
w_4	45	67	68	12	0
w_5	37	32	23	56	0

In this table, the logical decision of the incomplete record w_3 matches with w_1 and contradicts with w_2,w_4 and w_5. Hence, ***SIM(w_3)***= (w_1} and ***DISSIM(w_3)***= { w_2,w_4,w_5 }

It is to be noted that if ***SIM(w_i)=Φ***, then the Similarity procedure cannot be followed and if ***DISSIM(w_i)=Φ***, then the Dissimilarity procedure cannot be followed.

As there procedures are exclusive of each other, both of them cannot be empty at the same time.

2.1.1. Fuzzy Similarity using one Threshold

In this section, we propose a method of approximating or fixing the unknown value in the incomplete record with fuzzy decision attribute using one threshold approach using similarity.

Consider the following decision Table with E as the decision variable.

	d_1	**d_2**	**...**	**d_{j-1}**	**d_j**	**d_{j+1}**	**....**	**d_n**	**μ_E**
w_1	$x_{1,1}$	$x_{1,2}$	...	$x_{1,j-1}$	$x_{1,j}$	$x_{1,j+1}$	...	$x_{1,n}$	$\mu_E(w_1)$
w_2	$x_{2,1}$	$x_{2,2}$	...	$x_{2,j-1}$	$x_{2,j}$	$x_{2,j+1}$	...	$x_{2,n}$	$\mu_E(w_2)$
.	...	...	...	...	...	...	...	...	.
.	...	...	...	...	...	...	...	...	.
w_{i-1}	$x_{i-1,1}$	$x_{i-1,2}$	...	$x_{i-1,j-1}$	$x_{i-1,j}$	$x_{i-1,j+1}$	...	$x_{i-1,n}$	$\mu_E(w_{i-1})$
w_i	$x_{i,1}$	$x_{i,2}$	...	$x_{i,j-1}$	*	$x_{i,j+1}$	...	$x_{i,n}$	$\mu_E(w_i)$
w_{i+1}	$x_{i+1,1}$	$x_{i+1,2}$	...	$x_{i+1,j-1}$	$x_{i+1,j}$	$x_{i+1,j+1}$	...	$x_{i+1,n}$	$\mu_E(w_{i+1})$
.	...	...	...	...	...	...	...	...	.
.	...	...	...	...	...	...	...	...	.
w_m	$x_{m,1}$	$x_{m,2}$	...	$m_{1,j-1}$	$x_{m,j}$	$x_{m,j+1}$	...	$x_{m,n}$	$\mu_E(w_m)$

The Procedure being followed is

a) Let α be a threshold. Using this threshold, we shall discretarize the decision values to be either 1 or 0.
b) Compute ***SIM(w_i)***.
c) If ***SIM(w_i)=Φ***, conclude that the Specified Method cannot be followed.
d) Let, ***SIM(w_i)***={S_1,S_2,..., S_t}
e) Compute ***relative deviation*** (S_j, w_i) for each j=1,2,...t
f) Let S_k be the complete record which has the least ***relative deviation***

g) Compute ***relative minimum quotient*** $\left\lfloor \frac{w_i}{S_k} \right\rfloor$

h) Fix the missing value $x_{i,j}$ of w_i as the product of ***relative minimum quotient*** $\left\lfloor \frac{w_i}{S_k} \right\rfloor$ and j[th] coefficient of the record S_k.

2.1.2 Fuzzy Dissimilarity using one Threshold

Now, we propose a method of approximating or fixing the unknown value in the incomplete record with fuzzy decision attribute using one threshold approach using dissimilarity.

Consider the following decision Table with E as the decision variable.

	d_1	d_2	...	d_{j-1}	d_j	d_{j+1}	....	d_n	μ_E
w_1	$x_{1,1}$	$x_{1,2}$	...	$x_{1,j-1}$	$x_{1,j}$	$x_{1,j+1}$	...	$x_{1,n}$	$\mu_E(w_1)$
w_2	$x_{2,1}$	$x_{2,2}$	...	$x_{2,j-1}$	$x_{2,j}$	$x_{2,j+1}$	...	$x_{2,n}$	$\mu_E(w_2)$
.	...	...	...	...	...	...	...	...	.
.	...	...	...	...	...	...	...	...	.
w_{i-1}	$x_{i-1,1}$	$x_{i-1,2}$	...	$x_{i-1,j-1}$	$x_{i-1,j}$	$x_{i-1,j+1}$	...	$x_{i-1,n}$	$\mu_E(w_{i-1})$
w_i	$x_{i,1}$	$x_{i,2}$	...	$x_{i,j-1}$	*	$x_{i,j+1}$	...	$x_{i,n}$	$\mu_E(w_i)$
w_{i+1}	$x_{i+1,1}$	$x_{i+1,2}$	...	$x_{i+1,j-1}$	$x_{i+1,j}$	$x_{i+1,j+1}$	...	$x_{i+1,n}$	$\mu_E(w_{i+1})$
.	...	...	...	...	...	...	...	...	.
.	...	...	...	...	...	...	...	...	.
w_m	$x_{m,1}$	$x_{m,2}$	...	$m_{1,j-1}$	$x_{m,j}$	$x_{m,j+1}$	...	$x_{m,n}$	$\mu_E(w_m)$

The Procedure being followed is

a) Let α be a threshold. Using this threshold, we shall discretarize the decision values to be either 1 or 0.

b) Compute ***DISSIM(w_i)***.

c) If ***DISSIM(w_i)=Φ***, conclude that the Specified Method cannot be followed.

d) Let, ***DISSIM(w_i)***=$\{S_1,S_2,..., S_t\}$

e) Compute ***relative deviation*** (S_j, w_i) for each j=1,2,...t

f) Let S_k be the complete record which has the least ***relative deviation***

g) Compute ***relative maximum quotient*** $\left\lceil \frac{w_i}{S_k} \right\rceil$

h) Fix the missing value $x_{i,j}$ of w_i as the product of ***relative maximum quotient*** $\left\lceil \frac{w_i}{S_k} \right\rceil$ and j[th] coefficient of the record S_k.

Note: _Since the comparison with more records will provide the accuracy, it is to be noted that

- If |***SIM(w_i)***| > |***DISSIM(w_i)***|, then Similarity approach is to be used
- If |***DISSIM(w_i)***| **>** |***SIM(w_i)***|, then Dissimilarity approach is to be used
- If |***SIM(w_i)***| **=** |***DISSIM(w_i)***|, then any approach may be followed

Example: Consider the following decision table

	d_1	d_2	d_3	d_4	d_5	d_6	μ_E
w_1	10	20	15	18	35	40	0.7
w_2	30	28	35	42	43	32	0.6
w_3	34	34	28	56	76	43	0.2
w_4	21	45	31	32	78	51	0.3
w_5	45	44	39	51	67	44	0.7
w_6	67	56	*	34	31	87	0.9
w_7	54	64	21	65	62	64	0.7
w_8	67	35	23	45	65	53	0.5
w_9	44	34	21	78	67	72	0.6
w_{10}	21	44	33	22	76	41	0.4

Here, w_6 has the missing value.

Case 1:

Let α=0.55

On applying α cut, we obtain $SIM(w_6)=\{w_1,w_2,w_5,w_7,w_9\}$ and $DISSIM(w_6) =\{w_3,w_4,w_8,w_{10}\}$

Since, $|SIM(w_6)|$ is greater than $|DISSIM(w_6)|$, we use the Similarity Approach.

Now, consider the similar records and compute (S_j, w_6)

	d_1	d_2	d_4	d_5	d_6	(S_j, w_6)
w_1	10	20	18	35	40	16.76425
w_2	30	28	42	43	32	14.67787
w_5	45	44	51	67	44	12.74676
w_7	54	64	65	62	64	10.36147
w_9	44	34	78	67	72	13.37161

Among all, the minimum relative deviation occurs for the record w_7.

	d_1	d_2	d_4	d_5	d_6	$\left\lfloor \frac{w_6}{w_7} \right\rfloor$
w_6	67	56	34	31	87	
w_7	54	64	65	62	64	
	1.240741	0.875	0.523077	0.5	1.359375	0.5

Here we obtain $\left\lfloor \frac{W_6}{W_7} \right\rfloor = 0.5$

Thus, the missing coefficient is fixed as 0.5 x 21=10.5

Case 2:

Let α=0.65

On applying α cut, we obtain $SIM(w_6)=\{w_1,w_5,w_7\}$ and $DISSIM(w_6)=\{w_2,w_3,w_4,w_8,w_9, w_{10}\}$

Since, $|DISSIM(w_6)|$ is greater than $|SIM(w_6)|$, we use the Dissimilarity Approach.

Now, consider the dissimilar records and compute (S_j, w_6)

	d_1	d_2	d_4	d_5	d_6	(S_j, w_6)
w_2	30	28	42	43	32	14.67787
w_3	34	34	56	76	43	15.51515
w_4	21	45	32	78	51	15.16047
w_8	67	35	45	65	53	10.72194
w_9	44	34	78	67	72	13.37161
w_{10}	21	44	22	76	41	16.18023

Among all, the maximum relative deviation occurs for the record w_{10}.

	d_1	d_2	d_4	d_5	d_6	$\left\lceil \frac{w_6}{w_{10}} \right\rceil$
w_6	67	56	34	31	87	
w_{10}	21	44	22	76	41	
	3.19048	1.273	1.545	0.408	2.122	3.19048

Here, $\left\lceil \frac{w_6}{w_{10}} \right\rceil = 3.19048$

Hence, the missing value is fixed as 3.19048 x 33=105.2858

Thus, the missing value is fixed using the proposed method.

3. Conclusion

In this paper, we developed a mathematical model leading into Two approaches namely fuzzy similarity and fuzzy dissimilarity which helps to fix the unknown value in the decision table.

References

[1] Ganesan G, et.al., *Feature Selection using Fuzzy Decision Attributes*, Journal of INFORMATION, Vol 9, No3, pp:381-394, 2006

[2] Ganesan G et.al., *Rough Set: Analysis of Fuzzy Sets using thresholds*, Computational Mathematics, Narosa, India,. pp: 81-87, 2005

[3] Kryszkiewicz M. *Rules in incomplete information systems*, Information Sciences, pp: 271-292, 1999

[4] Latkowski, R. *On decomposition for incomplete data*, Fundamenta Informaticae, pp:1-16, 2003

[5] Little R.J.A. and Rubin D.B. Statistical Analysis with Missing Data, Second Edition, J. Wiley & Sons, Inc., 2002.

[6] Stefanowski J. and Tsoukias A. *Incomplete information tables and rough classification*, Computational Intelligence, pp: 545-566, 2001