International Journal of Hybrid Information Technology
Vol. 9, No.7 (2016), pp. 79-88
http://dx.doi.org/10.14257/ijhit.2016.9.7.09

Variable Threshold Valued Tolerance Relation in Incomplete Information System

Liu ping[1], Xiong xiaoxiao[1] and Yang Junping[2*]

[1]*The school of Information Engineering, Nanchang University, Nanchang 330031, China*
[2,*]*Jiangxi University of Traditional Chinese Medicine*
[1]*liuping@ncu.edu.cn*
[2]*yyfyjp0507@126.com*

Abstract

Aiming at the incomplete information systems on the condition of no prior domain knowledge, several known model extension based on the rough set theory are introduced at first, such as the tolerance relation, non-symmetric similarity relation, limited tolerance relation and valued tolerance relation. Then the merits and drawbacks of several existing valued tolerance relations are compared in this article. Next, the experiments on some UCI data sets have been done ,based on the experimental result, the author discuss the relationship between the threshold selection and classification accuracy of statistical valued tolerance relation(SVT) . Directing at the difficulty of selecting a suitable threshold, the author presents a new improved valued tolerance relation (NVT) which can choose proper threshold automatically on the basis of each data set's feature. Experiment results indicate that the new relation can get better classification accuracy than the other extension models in dealing with the incomplete system which has small incomplete degree

***Keywords**: Incomplete information system, rough set, New Valued Tolerance relation (NVT), statistical valued tolerance relation (SVT), threshold selection*

1. Introduction

The classical rough set theory raised by Poland mathematician in 1982 was based on the complete information system (CIS). However, in our daily life there is various reasons lead to incomplete information system (IIS), such as the errors of data measuring, the limitations of acquiring data, some human factors, *etc.* Therefore, in recent years, many scholars and experts focus on how to manage incomplete information. To find out the solution, they have presented several extended rough set model [1-15]. However, they all have merits and faults.

The tolerance relation [1] provides us an incomplete information system (IIS),

$IIS = <U, C \cup D, V, f>$ $B \subseteq C$, The unknown attribute value is denoted by "*". And definition of tolerance relation is as follow:

$$\forall x, y \in U \times U \ T(x, y) \Leftrightarrow \forall b \in B \ b(x) = b(y) \vee b(x) = * \vee b(y) = * \quad (1)$$

The classical tolerance relation has its own limitations [2]. By an example, the object $X = (3,*,1,*,2,*,2,*)$ and the object $Y = (*,10,*,5,*,8,*,3)$ are divided into the tolerance class, but actually there are no equal certain attribute value between the two objects. The definition of classical tolerance relation is imprecision that "*" can be regarded as any attribute value.

The definition of non-symmetric similarity relation is defined as [3]

$$S_B(x, y) = \{(x, y) \in U \times U \mid \forall_{b \in B}(b(x) = * \vee b(x) = b(y))\} \quad (2)$$

According to the definition of non-symmetric similarity relation, the object X= (0, *, 1, 6, 9, 8, 12, 3) and the object Y= (*, 0, 1, 6, 9, 8, 12, 3) can be distinguished, so they are not

ISSN: 1738-9968 IJHIT

satisfied with the similar relation. In fact, object X has six attributes equaled with object Y among its all eight attributes. The similarity between them can not be ignored.

To overcome the shortcomings of classical tolerance relation, Wang Guoying put forward limited tolerance relation [4].

$$L_B(x,y)=\{(x,y)\in U\times U \mid \forall_{b\in B}(b(x)=b(y)=*)\vee(P(x)\cap P(y)\neq\phi)\wedge\forall_{b\in B}((b(x)\neq *)\wedge(b(y)\neq *)\rightarrow(b(x)=b(y)))\} \quad (3)$$

However, the limited tolerance relation still has limitations as classical tolerance relation in some way. In the light of limited tolerance relation's definition, though object $X=(*,*,*,*,*,*,*,*,*,2)$ and object $Y=(*,*,*,*,*,*,*,*,*,2)$ only has one equal attribute value, they still in the same tolerance class. Actually, the similarity of these two objects is quite small.

The completeness tolerance relation is defined as [5].

$$M_B(x,y)=\{(x,y)\in U\times U \mid \forall_{b\in B}(p(x)\neq\phi)\wedge(p(y)\neq\phi)$$
$$\wedge\frac{|P(x)\cap P(y)|}{\min(|p(x)|,|p(y)|)}\geq\gamma$$
$$\wedge\forall_{b\in B}((b(x)\neq *)\wedge b(y)\neq *)\rightarrow(b(x)=b(y)))\}$$
(4)

$p(x)=\{b\in B \mid b(x)\neq *\}$, $\gamma(x_i)=\frac{|p(x_i)|}{|B|}$, $\gamma=\sum_{i=1}^{|U|}\frac{\gamma(x_i)}{|U|}$

Two objects whose certain attribute value are the same but have low completeness are avoidable divided into one tolerance class.

MA Xi-ao also analyses some limitations of completeness tolerance relation [2]. For example, two objects $x=(*,*,*,*,6,*,*,*,*,*)$ and $y=(*,*,*,*,6,*,*,9,*,*)$ which are satisfied with the completeness tolerance relation's definition with the condition: $\frac{|p(x)\cap p(y)|}{\min(|p(x)|,|p(y)|)}=1$. Meanwhile, the similarity of these two objects is quite small. Analyzing the characteristics of the two objects, there is inclusion relation in the known attribute values.

Stefanowski presented the valued tolerance relation to measure the tolerance degree between two objects with a numerical value [6]. In next section, some related concepts of valued tolerance relation will be introduced.

2. The Comparison of Some Known Valued Tolerance Relation

2.1. The Valued Tolerance Relation Raised By Stefanowski

Stefanowski represents the definition of valued tolerance relation and tolerance class [6]. Let $S=<U,C\cup D,V,f>$ be an IIS, where $B\subseteq C$ is a subset of the attribute. For any $x\in U$, $Q(x)$ is defined as $Q_B(x)=\{(y,P_B(x,y)) \mid y\in U, 0<P_B(x,y)\leq 1\}$, the tolerance class of x, is a fuzzy set by using the tolerance degree of reference element as membership function.

And the $P_B(x,y)$ represents the membership function that y belongs to the tolerance class of x and it is based on the subset of attribute set B, which means the tolerance degree between y and x within B. And according to distribution characteristics of data set in IIS, the measurement method of probability among the objects has various forms, which causes the multiple forms of valued tolerance relations.

The valued tolerance relation [6] is based on the assumption that there exists a uniform probability distribution among the possible values on each attribute.

For IIS $S=<U,C\cup D,V,f>$, $B\subseteq C$

$$T_B^\lambda=\{<x,y>\in U^2 \mid P_B(x,y)\geq\lambda\}\cup I_U, \quad I_U=\{<x,x>\mid x\in U\} \quad (5)$$

$$P_B(x,y)=\prod_{b\in B}p_{\{b\}}(x,y) \quad (6)$$

$$P_{\{b\}}(x,y)=\begin{cases}1 & b(x)=b(y)\wedge b(x)\neq *\wedge b(y)\neq * \\ 0 & b(x)\neq b(y)\wedge b(x)\neq *\wedge b(y)\neq * \\ 1/|V_b| & b(x)=*\vee b(y)=*\end{cases} \tag{7}$$

λ in the formula is a given threshold ,while $P_{\{b\}}(x,y)$ represents the probability that x is similar to y on b. And $P_B(x,y)$ is the probability that x is similar to y on B, $P_B(x,y)$ is called tolerance degree, T_B^λ is the valued toleration relation. Similarly, V_b denotes the domain of the attribute b.

The range of tolerance degree described by classical tolerance relation is $\{0,1\}$, that is to say, the tolerance degree of two objects is only 0 or 1 such two discrete value[7].Nevertheless, the range is $[0..1]$, a continuous value range, in the valued tolerance relation raised by Stefanowski and Tsoukis. In this way, the tolerance degree has a quantitative standard. However, this method still has its own limitations that the attribute value needs to be independent and evenly distributed.

2.2. Statistical Valued Tolerance Relation (SVT)

Given an IIS $S=<U,C\cup D,V,f>$, $b\in C$, the domain of the attribute b is defined as $V_b=\{k_b^1b_1,k_b^2b_2,\ldots k_b^m b_m\}$, where $b_1,b_2,\ldots,b_m$ is all possible known values of b, and $k_b^i(i=1,2,\ldots,m)$ denotes the cardinality of the set $\{x\in U\mid b(x)=b_i\}$.

For $\forall x\in U$, the probability that $b(x)=b_i$ is $k_b^i/(k_b^1+k_b^2+\ldots+k_b^m)$, $P_B(x,y)$ is the probability that x is similar to y on B is defined as

$$P_B(x,y)=\prod_{b\in B}p_{\{b\}}(x,y) \tag{8}$$

$$ST_B^\lambda=\{<x,y>\in U^2\mid P_B(x,y)\geq\lambda\}\cup I_U \tag{9}$$

$$P_{\{b\}}(x,y)=\begin{cases}1 & b(x)=b(y)\wedge b(x)\neq *\wedge b(y)\neq * \\ 0 & b(x)\neq b(y)\wedge b(x)\neq *\wedge b(y)\neq * \\ k_b^i/\sum_{j=1}^m k_b^j & (b(x)=b_i\wedge b(y)=*)\vee(b(x)=*\wedge b(y)=b_i) \\ \sum_{i=1}^m(k_b^i/\sum_{j=1}^m k_b^j)^2 & b(x)=*\wedge b(y)=*\end{cases} \tag{10}$$

$$I_U=\{<x,x>\mid x\in U\}$$

$$ST_B^\lambda(x)=\{y\mid(x,y)\in ST_B^\lambda, y\in U\} \tag{11}$$

ST_B^λ represents the statistical valued tolerance relation(SVT), now we analyze the advantages of SVT[7]:

On the basis of the definition of ST_B^λ, it can be seen that statistical valued tolerance relation ST_B^λ is a special case of Stefanowski's valued tolerance relation T_B^λ ; ST_B^λ can degenerate into T_B^λ in the condition of $k_b^1=k_b^2=\ldots=k_b^m$ which means that all known values of each attribute is evenly distributed.

On account of the analysis of the two valued tolerance relation, we can know that the selection of threshold is essential while judging whether two objects is compatible or not. In the meantime, the selection of threshold is a key problem for valued tolerance relation.

2.3. Relation between the Selection of Threshold and the Classification Accuracy in SVT

To test the relation between the selection of threshold and the classification accuracy in SVT, we did the test below. We have used the following formula to compute the classification accuracy [8].

$$\mu_R=\frac{\sum_{x\in U}\frac{|[x]_E\cap R(x)|}{|[x]_E\cup R(x)|}}{|U|} \tag{11}$$

$$\mu_R = \frac{\sum_{x \in U} \frac{|[x]_E \cap (K_R(x) \cup K_R^{-1}(x))|}{|[x]_E \cup (K_R(x) \cup K_R^{-1}(x))|}}{|U|} \quad (12)$$

$[x]_E$ represents the equivalent class of x in U, while R is denoted as a generalized indiscernibility relation on U. If R is symmetric, use formula (11) to compute μ_R, else use formula (12) to compute μ_R.

In our experiments, four complete data sets (Balance, Tic-Tac-Toe, Chess) in UCI have been used. For each data set, randomly selected from 5%, 10%, 30%, 50%, 70% data for missing values, so 20 incomplete data sets (Balance-5%, Balance-10%, Balance-30%, Balance-50%, Balance-70%, Tic-Tac-Toe-5%,Tic-Tac-Toe-10%,Tic-Tac-Toe-30%,Tic-Tac-Toe-50%,Tic-Tac-Toe-70%, Chess-5%, Chess-10%, Chess-30%, Chess-50%, Chess-70%, Car-5%, Car-10%, Car-30%, Car-50%, Car-70%) have been got.

Table 1.Four Complete Data Sets in UCI

Data sets	No. of objects	No. of Condition attributes	No. of Decision attributes	Whether contains the missing values
Balance	625	4	1	NO
Tic-Tac-Toe	958	9	1	NO
Chess	3196	36	1	NO
Car	1728	6	1	NO

Table 2.Experiment Results

Threshold / Data sets	λ =0	λ =0.01	λ =0.05	λ =0.1	λ =0.2	λ =0.3	λ =0.4
Balance-5%	0.521	0.532	0.639	0.639	0.755	1.000	1.000
Balance-10%	0.242	0.300	0.505	0.505	0.678	1.000	1.000
Balance-30%	0.021	0.154	0.667	0.671	0.779	1.000	1.000
Balance-50%	0.005	0.289	0.905	0.921	0.942	1.000	1.000
Balance-70%	0.002	0.579	0.984	0.997	0.997	1.000	1.000
Car-5%	0.499	0.502	0.533	0.647	0.676	0.850	1.000
Car-10%	0.215	0.239	0.316	0.553	0.627	0.829	1.000
Car-30%	0.014	0.121	0.476	0.803	0.906	0.959	1.000
Car-50%	0.002	0.396	0.873	0.974	0.994	0.998	1.000
Car-70%	0.001	0.830	0.995	0.999	1.000	1.000	1.000
Chess-5%	0.793	0.804	0.837	0.866	0.913	0.944	0.966
Chess-10%	0.587	0.662	0.778	0.857	0.935	0.971	0.987
Chess-30%	0.081	0.776	0.985	0.997	1.000	1.000	1.000
Chess-50%	0.006	0.997	1.000	1.000	1.000	1.000	1.000
Chess-70%	0.002	1.000	1.000	1.000	1.000	1.000	1.000
Tic-Tac-Toe-5%	0.913	0.917	0.950	0.976	1.000	1.000	1.000
Tic-Tac-Toe-10%	0.678	0.732	0.887	0.957	1.000	1.000	1.000
Tic-Tac-Toe-30%	0.062	0.541	0.968	0.988	1.000	1.000	1.000
Tic-Tac-Toe-50%	0.007	0.895	0.996	0.999	1.000	1.000	1.000
Tic-Tac-Toe-70%	0.002	0.998	1.000	1.000	1.000	1.000	1.000

The experiment results are shown in Table 2, where the first row and column of the table denotes the selection of threshold and the incomplete data sets. By analyzing the data in the table 2, we can draw the conclusions below.

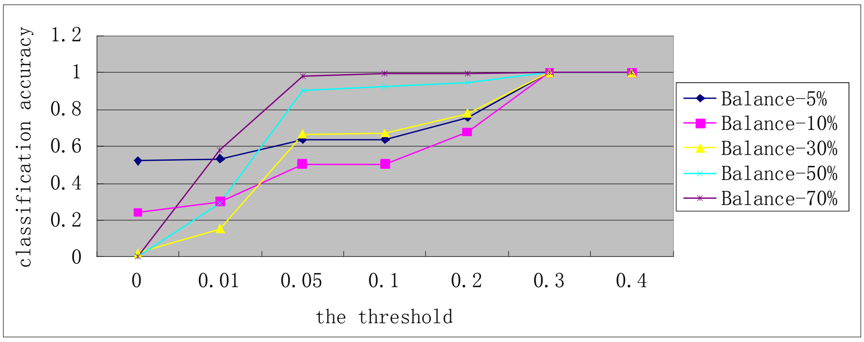

Figure 1. The Relationship between Threshold and Classification Accuracy of Balance Data Set

The horizontal ordinate in Figure1 represents the value of threshold while the vertical coordinate denotes the classification accuracy. And each polyline in Figure 1 is denoted as the variety of the classification accuracy with the change of the threshold towards one incomplete data set. For instance, we observe the yellow polyline, the data set of Balance-30%, which contains 30% missing attribute values. While the threshold changes from 0 to 0.4, the corresponding classification accuracy grows from 0.021 to 1. To find what causes this problem, we need to go back to definition of the classification accuracy.

Firstly, we discuss the extreme case that $\lambda \geq 0.3$ which means the tolerance degree between two objects need to come to at least 0.3, And for each object's tolerance class $R(x) = \{x\}$ in the data set "Balance-30%" is only the object itself, $[x]_E \cap R(x) = [x]_E \cup R(x) = x$, therefore, we can derive $\mu_R = 1$; and Secondly, when $\lambda < 0.05$, we can see that there are numerous quantitative tolerance class $R(x)$ which means denominator $|[x]_E \cup R(x)|$ increased and numerator is still the same while the threshold is small so that classification accuracy decreases from the definition of μ_R. Finally, classification accuracy has a qualitative leap, growing from 0.154 to 0.667 when $0.01 \leq \lambda \leq 0.05$ meanwhile classification accuracy changes from 0.667 to 0.779 while $0.05 \leq \lambda \leq 0.2$. Therefore, for the data set "Balance-30%", the threshold should select from the domain 0.05 to 0.2. Consequently, we can acquire the threshold's suitable domain of the data set above:

Table 3.The Domain of Suitable Threshold towards Different Data Set

Data set	Domain of suitable threshold	Data set	Domain of suitable threshold
Balance-5%	[0.05,0.2]	Chess-5%	[0.2,0.4]
Balance-10%	[0.05,0.2]	Chess-10%	[0.2,0.4]
Balance-30%	[0.05,0.2]	Chess-30%	[0.05,0.1]
Balance-50%	[0.05,0.2]	Chess-50%	>0.01
Balance-70%	[0.05,0.2]	Chess-70%	>0.01
Car-5%	[0.2,0.3]	Tic-Tac-Toe-10%	[0.01,0.1]
Car-10%	[0.2,0.3]	Tic-Tac-Toe-30%	[0.01,0.1]
Car-30%	[0.2,0.3]	Tic-Tac-Toe-50%	[0.01,0.1]
Car-50%	[0.1,0.3]	Tic-Tac-Toe-70%	[0.01,0.1]
Car-70%	[0.05,0.1]		

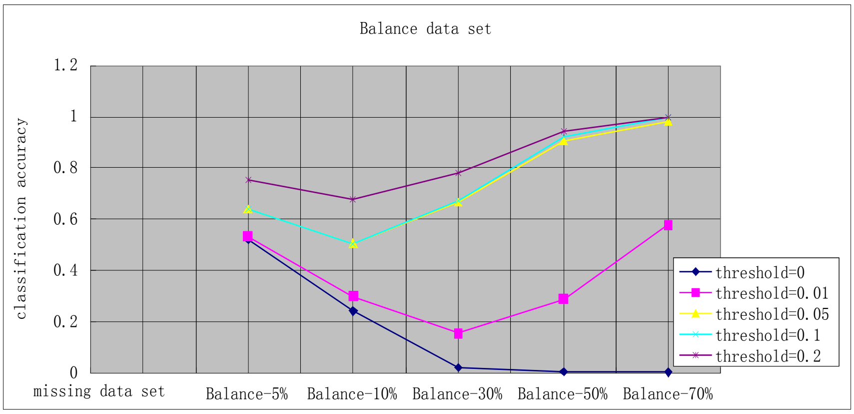

Figure 2. Given a Threshold, the Classification Accuracy of Different Balance Missing Data Set

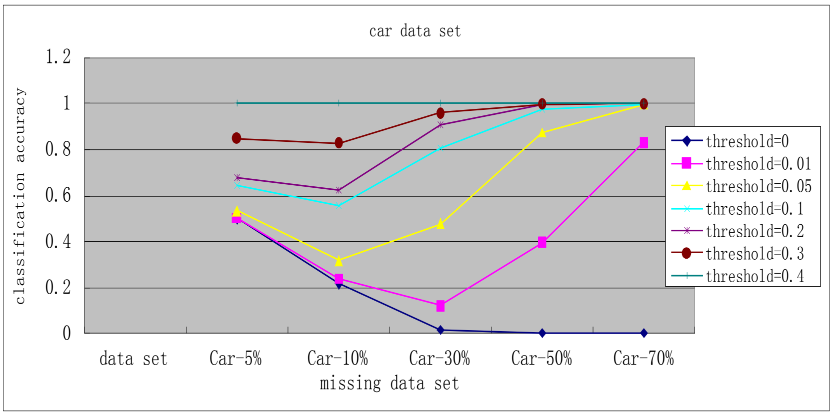

Figure 3. Given a Threshold, The Classification Accuracy Of Different Car Missing Data Set.

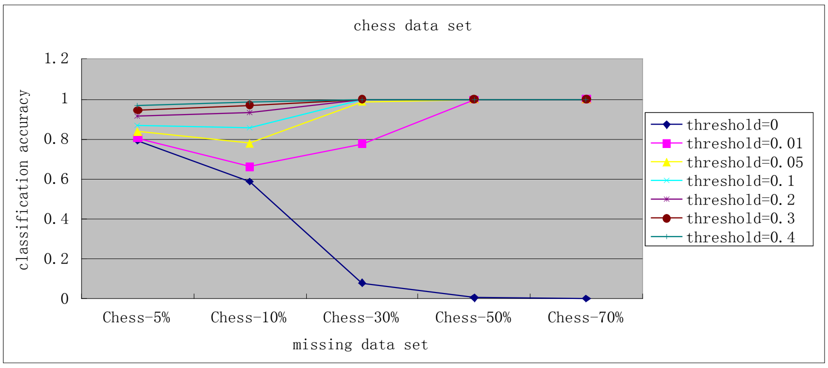

Figure 4. Given a Threshold, the Classification Accuracy of Different Chess Missing Data Set

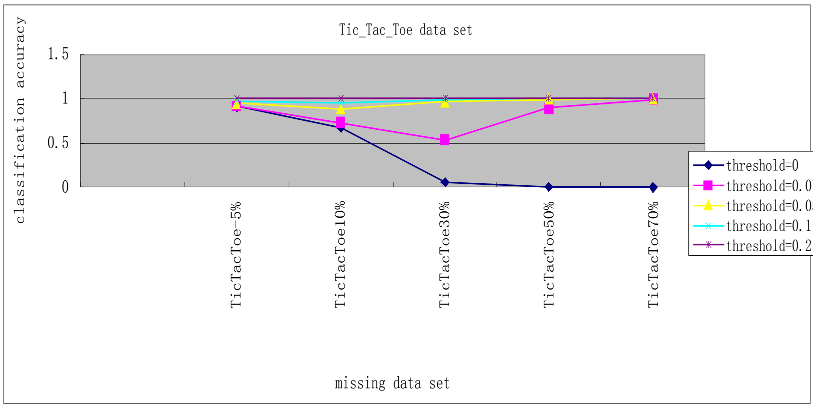

Figure 5. Given a Threshold, the Classification Accuracy of Different Tic_Tac_Toe Missing Data Set

The following analysis can be drawn from observing Figure 2 to 5:

First: For the data set which contains missing values, classification accuracy has a sharp decline while the incomplete degree increases if threshold equals to zero. What causes this phenomenon?

If $\lambda = 0$, $\forall x \in U$, $R(x) = \{y \mid P_B(x, y) > 0, y \in U\} = U$, each object can be regarded as any other objects' tolerance class which means each object's tolerance class $R(x)$ has a sharp increase, $[x]_E \cap R(x) \approx [x]_E$; $[x]_E \cup R(x) = U$; the numerator is still the same while denominator has a increase so that the classification accuracy $\mu_R(x)$ gets a decline.

Second: With the incomplete degree increase, classification accuracy declines at first and then rises up gradually. Now we take the yellow polyline in Figure 3 as an example, which represents the threshold is 0.05. And while the increase of incomplete degree of the data set Car, especially reaching a lowest value when the incomplete degree at 10%, the classification accuracy reaches a decline. Then when the incomplete degree grows from 10% to 50% gradually, the classification accuracy gradually rising.

3. A New Improved Valued Tolerance Relation

On the basis of the experiment and analysis, we represent a new improved valued tolerance relation.

Definition1: Suppose $S = <U, C \cup D, V, f>$ is an incomplete information system (IIS). In the subset $B \subseteq C$, the completeness of the object $x_i \in U$ is defined as: $\gamma_B(x_i) = |T(x_i)| / |B|$. $T(x_i) = \{b \mid b \in B \wedge b(x_i) \neq *\}$, $|B|$ represents the cardinality of B set, Then the completeness of IIS in the B is defined as follows:

$$\gamma_B = \sum_{i=1}^{|u|} \gamma_B(x_i) / |U|$$

Definition 2: For the IIS $IIS = <U, C \cup D, V, f>$, the average value of toleration degree is defined as $\lambda = \frac{\sum P_B(x, y)}{C_{|U|}^2}$, x, y are two arbitrary objects in U, $x, y \in U$; $|U|$ is cardinality of the set U. $C_{|U|}^2$ represents permutation and combination, it shows that select any two from $|U|$ objects.

Definition 3: New Valued Tolerance relation $IIS = <U, C \cup D, V, f>$, $B \subseteq C$, $\exists b \in B$. New Valued Tolerance relation is defined as

$$NVT_B^{\lambda} = \{(x, y) \in U \times U \mid (P_B(x, y) > \lambda) \wedge (\frac{\gamma(x) + \gamma(y)}{2} \geq \gamma_B) \wedge (f(x, D) = f(y, D))\} \cup I_U \quad (13)$$

$I_U = \{<x, x> \mid x \in U\}$

Definition 4: New valued tolerance class:

$$NVT_B^{\lambda}(x) = \{y \mid (x, y) \in NVT_B^{\lambda}, y \in U\} \quad (14)$$

The improved valued tolerance relation considers the virtue of the complete tolerance relation and the valued tolerance relation. From the concept of improved valued tolerance relation, we can see that we discipline that the completeness degree of two objects need to greater than the average completeness degree of the whole object set and the tolerance degree of two objects need to greater than the average tolerance degree of the whole object set. That is to say, for any incomplete information system, the threshold, measuring two objects whether satisfies the definition of valued tolerance relation is defined as the average value of any two objects in the IIS. And only two objects that have the same decision value and satisfy the conditions above can be the tolerance class.

And next, we'll prove the legitimacy of our definition through experiments.

In our experiments, four complete data sets (Balance, Tic-Tac-Toe, Chess) in UCI have been used. For each data set, randomly selected from 5%, 10%, 30% data for missing values, so 12 incomplete data sets (Nursery-5%, Nursery -10%, Nursery -30%, Molecular -5%, Molecular -10%, Molecular -30%, Chess-5%, Chess-10%, Chess-30%, Car-5%, Car-10%, Car-30%) have been got.

Table 4.Four Complete Data Sets in UCI

Data sets	No. of objects	No. of Condition attributes	No. of Decision attributes
Nursery	12960	8	1
Molecular	3190	60	1
Chess	3196	36	1
Car	1728	6	1

Table 5.Experiment Results

Data sets	TR	LTR	SVT(0.5)	SVT(0.1)	NVT
Nursery-5%	0.483	0.483	0.403	0.577	1.000
Nursery-10%	0.197	0.197	0.159	0.505	0.751
Nursery-30%	0.010	0.010	0.190	0.860	0.765
Molecular-5%	0.998	0.998	0.944	0.960	0.978
Molecular-10%	0.994	0.994	0.921	0.953	0.952
Molecular-30%	0.983	0.983	0.972	0.981	0.942
Car-5%	0.581	0.581	0.533	0.647	1.000
Car-10%	0.308	0.308	0.316	0.553	0.709
Car-30%	0.032	0.033	0.476	0.803	0.250
Chess-5%	0.815	0.815	0.837	0.866	0.943
Chess-10%	0.629	0.629	0.778	0.857	0.825
Chess-30%	0.114	0.114	0.985	0.997	0.408

The experiment results has shown in table 5, where TR, LTR, SVT, NVT denote respectively tolerance relation, limited tolerance relation, statistical valued tolerance relation and new valued tolerance relation. SVT (0.5) denotes statistical valued tolerance relation and the threshold equals 0.5. On the basis of the analysis of the data from table 5, the classification accuracy of NVT is higher than that of TR, LTR; the classification accuracy of NVT is higher than that of SVT (0.5) and SVT (0.1) when incomplete degree of data set is less than 30%, while incomplete degree is more than 30%, the performance of NVT is poor. The reason is that there are more restrictive conditions in NVT.

4. Conclusion

Focusing on data sets in UCI, this paper conducted experiments on the relation between the selection of threshold and the classification accuracy in SVT. And on the basis of the experiment results, the relation between the selection of threshold and the classification accuracy is discussed and some suitable thresholds of actual data sets are proposed. Then, a new improved valued tolerance relation that combines the advantages of respectively completeness tolerance relation and valued tolerance relation comes into being. We present a simple and reasonable method of the selection of the threshold in the end. And the experiment results also prove the rationality of this new defined tolerance relation.

Acknowledgements

This research is supported by the National Natural Science Foundation of China under grant No. 81260578 and No. 81460769.

References

[1] M. Kryszkiewicz, “Rough set approach to incomplete information systems”, Information Sciences, vol. 112, **(1998)**, pp. 39-49.
[2] X.-A. Ma, G.-Y. Wang, Q.-H. Zhang and Q.-S. Yang, “Extended rough set model based on improved complete tolerance relation”, Journal of Computer Applications, vol. 30, no. 7, **(2010)**, pp. 1873-1882.
[3] J. Stefamowski and A. Tsoukeas, “On the extension of rough sets under incomplete information”, International Journal of Intelligent System, vol. 16, no. 1, **(2000)**, pp. 29-38.
[4] G.Y. Wang, “Extension of rough set under incomplete information systems”, Journal of Computer Research and Development, vol. 39, no. 10, **(2002)**, pp. 1238-1243.
[5] L. Sheng and H.-Z. Yang, “Extended rough set model based on completed tolerance relation”, Control and Decision, vol. 23, no. 3, **(2008)**, pp. 258-262.
[6] J. Stefanowski and A. Tsouki`as, “On the Extension of Rough Sets under Incomplete Information”, In: Zhong, N., Skowron, A., Ohsuga, S. (eds.) RSFDGrC,LNCS (LNAI), vol. 1711, **(1999)**, pp. 73-82
[7] L.-H. Guan, “processing incomplete information methods based on rough set”, Journal of Chongqing University of Posts and Telecommunications (Natural Science Edition), vol. 21, no. 4, **(2009)**, pp. 461-466.
[8] G. Wang and L. Guan, “Data-Driven Valued Tolerance Relation”, RSKT, (2012), LNAI 7414; pp. 11-19.
[9] C. Ding and L. L, “Improved ROUSTIDA algorithm based on similarity relation vector”, Computer Engineering and Applications, vol. 50, no.13, **(2014)**, pp. 133-136.
[10] J. Zhou and X. Yang, “Study on characteristic consistent blocks based rough set approach to incomplete system”, IEEE International Conference on Granular Computing, **(2012)**.
[11] J. Dai, “Rough set approach to incomplete numerical data. Information Sciences”, vol. 241, **(2013)**, pp. 43-57.
[12] X. Li, S.-G. Wang, D.-Y. Li, X.-P. Kang and Y.-h. Zhai, “Knowledge Acquisition in Incomplete Information System Based on Formal Concept Analysis”, Computer Science, vol. 41, no. 7, **(2014)**, pp. 250-260.
[13] D. V. Nguyen, Y. Koichi and U. Muneyuki, “Rough Set Model Based on Parameterized Probabilistic Similarity Relation in Incomplete Decision Tables”, SCIS-ISIS, November 20-24; Kobe, Japan, **(2012)**.
[14] S. Huang and M. Li, “Limited and Variable Precision Rough Set Model”, Journal of Information & Computational Science, vol. 11, no. 10, (2014), pp. 3493-3501.
[15] W. Xu, C.Wu and X.B. Yang, “Incomplete variable precision multi-granulation rough sets based on tolerance relation”, Application Research of Computer, vol. 30, no.6, **(2013)**, pp. 1712-1715.