http://ansinet.com/itj

ITJ

ISSN 1812-5638

INFORMATION TECHNOLOGY JOURNAL

ANSInet
Asian Network for Scientific Information
308 Lasani Town, Sargodha Road, Faisalabad - Pakistan

Information Technology Journal 12 (19): 5377-5385, 2013
ISSN 1812-5638 / DOI: 10.3923/itj.2013.5377.5385

Application of Flexible Logic Average Operation Model on Selection of Approximate Support Vectors on [0,8) Interval

[1]Yang Zhi-xiao, [1]Fan Yan-feng and [2]He Hua-can
[1]College of Information Science and Engineering, Henan University of Technology, Henan, 450001, Zhengzhou, China
[2]College of Computer Science, Northwest Polytechnical University, Shanxi, 710072, Xi'an, China

Abstract: In flexible logics, operation models are continuously variable operator clusters along with both generalized self-correlation coefficient k and generalized correlation coefficient h in their existential domain. Based on the principal of modifying central base models by generator integrity clusters, this study uses exponential N/T-generator integrity clusters to define the flexible average operation models of propositional connectives on [0,8) interval, designs an algorithm for Approximate Support Vectors (ASVs) selection and applies 0-level flexible average operation model on [0,8) to the selection of ASVs in Support Vector Machine (SVM) training.

Key words: Flexible logics, generator integrity clusters, average operation models, [0,8) space, approximate support vectors

INTRODUCTION

Preliminaries of flexible logics: Fuzzy logics, the true value of proposition is a real number $x\in[0,1]$. Fuzzy logic can solve the fuzziness of true value and has been widely used in many fields. But in our opinion, not only the continuous changeability of propositional true value but also the continuous changeability of relations among propositions affects the operation model of propositional connectives. We call the former truth value flexibility while the later relational flexibility. Flexible logics (also called universal logics) mainly research the relational flexibility. Relational flexibility is caused by two independent factors as follows (He *et al*., 2005):

Measurement error of true value: The measurement error affects the truth value calculation of NOT propositions, so it affects all the logical operations. The relativity between propositions and their NOT propositions is called generalized self-correlativity. The continuous generalized self-correlativity coefficient $k\in[0,1]$ is used to describe the size of generalized self-correlativity.

Relationship between propositions: Relationship between propositions affects the truth value calculation of binary composite propositions. The relativity is called generalized correlation which can change continuously from the maximal correlation to the minimal correlation. The continuous generalized correlativity coefficient $h\in[0, 1]$ is used to describe the size of generalized correlativity.

Flexible logics operation models are generated based on the generator integrity clusters which modify each logic operator central base model: N-generator integrity cluster modifies the effect to propositional truth value caused by generalized self-correlativity (measure errors); T(S)-generator integrity cluster modifies the effect to relations of the propositions caused by generalized correlativity. Take *AND* operation in flexible logics for example:

$$T(x, y, h, k) = F^{-1}(\max(F(0, h, k), F(x, h, k) + F(y, h, k)-1), h, k)$$

in which $F(x, h, k) = F_0(\Phi(x, k), h)$, $F_0(x, h)$ is called T-generator integrity cluster, $\Phi(x, k)$ is called N-generator integrity cluster; k and h are generalized self-correlativity coefficient and generalized correlativity coefficient respectively. So, flexible logic operation models are continuously variable operator clusters along with both generalized self-correlation coefficient k and generalized correlation coefficient h in their existential domain. Because of space constraints, for detailed research about flexible logics (He *et al*., 2005).

Corresponding Author: Yang Zhi-Xiao, College of Information Science and Engineering, Henan University of Technology, Henan, 450001, Zhengzhou, China

Research background: At present, researches on flexible logics are mainly based on [0,1] base space. For over 10 years since flexible logics has been established, many researchers have made deep study on several different theoretical aspects (Ma and He, 2006; Luo and He, 2005; Xue *et al*., 2008; Zhang *et al*., 2006; Chen *et al*., 2011). Flexible logics have also made certain achievements applied to ternary optical computer, data mining, etc. (Jin *et al*., 2004; Jin *et al*., 2005; Jia, 2009). But all these researches are mainly on [0,1] base space.

However, in some important disciplines, such as life science, social science, thinking science, intelligent science, ecological system, meteorology system, as well as complex and large systems, it is difficult to transfer problem domain into [0,1] space. The difficulties are mainly due to:

- In complex nonlinear systems, it is not easy to determine the membership function because of the interaction among many variables
- In essence, the process to decide the membership function should be objective, but everyone's understanding to the same fuzzy concept is different, it is subjective to determine the membership function
- Complex system has higher sensitivity to initial values, the transformation between different domains will lead to information loss. The small deviation of the initial value will cause the extremely deviation of the results

Chen (2004), Mao *et al.* (2006), Chen *et al.* (2006a) and Chen *et al.* (2006b) studied the models of binary propositional connectives in [a, b] space with certain lower/upper limit. In comparison with [a, b] space with certain lower/upper limit, the properties of no upper limit problem changes greatly, especially the construction of N/T/S generators and the generation method of corresponding norms are different.

This study focuses on the construction of flexible Average propositional connective on [0,8) interval by using exponential integrity clusters, designs a algorithm to apply the model to the selection of Approximate Support Vectors (ASVs). The detailed organization of each part is as following:

Part 2: Basic set of exponential integrity clusters of NT operation model on [0,8) interval
Part 3: Construction of flexible Average propositional connective and proves the four special Average operators on [0,8) interval
Part 4: Designs an algorithm for applying 0-level flexible Average operation model to the selection of ASVs before SVM training process
Part 5: Conclusion of the study

For the purpose of abbreviation, the following discussion is on [0,8) space unless otherwise specified.

BASICSET OF EXPONENTIAL INTEGRITY CLUSTERS OF OPERATION MODEL

As mentioned above, flexible logic operation models are generated based on the generator integrity clusters. Directly substituting N-generator integrity cluster $\Phi(x, k)$ into the base model of NOT propositional connective, we can obtain the operation model of NOT propositional connective to achieve the definition of flexible NOT propositional connective. For [0,8) interval, refer to (Fan *et al*., 2012). Similarly, if directly substituting T-generator integrity cluster $F_0(x,h)$ (or S-generator integrity cluster $G_0(x,h)$) into the NT base model (or NS base model) of binary propositional connective, we can obtain the 0-level operation model of binary propositional connective; if directly substituting N-generator integrity cluster $\Phi(x,k)$ and T-generator integrity cluster $F_0(x,h)$ (or S-generator integrity cluster $G_0(x,h)$) into the base model of binary propositional connective at the same time, we can obtain the 1-level operation model of binary propositional connective, thus we can achieve the definition of flexible binary propositional connectives.

Here, the basic elements for establishment of binary propositional connective operation model on [0,8) interval will be discussed, in which exponential N-generator integrity cluster, exponential T-generator integrity cluster and NT operation model are used.

Exponential N-generator integrity cluster:

Definition 1: Suppose $\Phi(x,k)$ is a N-cluster in [0,8), in which $k\in[0, 1]$. For a certain $k_1\in[0,1]$, $\phi(x) = \Phi(x, k_1)$ is a N-generator. If $\Phi(x, k)$ satisfies:

- $\Phi(x, k)$ is continuous and strictly monotone decreasing with k
- $k = \Phi^{-1}(x, k)/(1+\Phi^{-1}(x, k))$ and if $k = 0.5$, then $\Phi(x, k) = \Phi_1 = x$
- When $k = 1$, $\Phi(x, k)\rightarrow\Phi_3'$; when $k = 0$, $\Phi(x, k)\rightarrow\Phi_0'$
- For $k_1,k_2\in[0, 1]$, there is $k_{21}\in[0, 1]$ which makes $\Phi(x, k_{21}) = \Phi(\Phi(x, k_1), k_2)$
- For $k_1\in[0, 1]$, there is $k_1'\in[0, 1]$ which makes $\Phi^{-1}(x, k_1) = \Phi(x, k_1')$

Then $\Phi(x, k)$ is called N-generator integrity cluster, whose abbreviation is N-generator cluster. In which $\Phi_3' = \Phi_0'^{-1} = \text{ite}\{8|x\rightarrow 8, 0\}$, $\Phi_0' = \Phi_3'^{-1} = \text{ite}\{0|x = 0, 8\}$.

Here $\Phi^{-1}(x, k)$ indicates the inverse of $\Phi(x, k)$ to the variable x.

Theorem 1: When $x\rightarrow[0, 8)$, the exponential function cluster $\Phi_2(x, k) = x^n/((1+x)^n-x^n)$, $n>0$, $k = 2^{-1/n}$ is N-generator integrity cluster.

Proof: We can easily get that $\Phi_2(x, k)$ is a continuous and strictly monotone function on [0, 8) increasing along with x and a continuous and strictly monotone function decreasing along with k. And we have $k = \Phi_2^{-1}(1, k)/(1+\Phi_2^{-1}(1, k))$ (Fan *et al*., 2012). Moreover, when $k = 1$, $\Phi_2(x, k)\rightarrow\Phi_3'$; when $k = 0$, $\Phi_2(x, k)\rightarrow\Phi_0'$; when $k = 0.5$, $\Phi_2(x, k) = \Phi_1 = x$.

For any two generators in the clusters:

$$\Phi_2(x, k_1) = x^{n1}/((1+x)^{n1}-x^{n1})$$

and:

$$\Phi_2(x, k_2) = x^{n2}/((1+x)^{n2}-x^{n2})$$

their composite operation:

$$\Phi_2(\Phi_2(x,k_1),k_2) = \frac{(x^{n1}/((1+x)^{n1}-x^{n1}))^{n2}}{(1+x^{n1}/((1+x)^{n1}-x^{n1}))^{n2}-(x^{n1}/((1+x)^{n1}-x^{n1}))^{n2}}$$
$$= x^{n1n2}/((1+x)^{n1n2}-x^{n1n2}) = \Phi_2(x,k_3)$$

and inverse operation:

$$\Phi_2^{-1}(x, k_1) = x^{1/n1}/((1+x)^{1/n1}-x^{1/n1}) = \Phi_2(x, k_1')$$

are also in $\Phi_2(x, k)$ cluster, so $\Phi_2(x, k)$ cluster is self-closed in composition and inverse operation.

Therefore $\Phi_2(x, k)$ is N-generator integrity cluster.

Exponential T-generator integrity cluster:

Definition 2: If T-generator f(x) is continuous along with generalized correlation coefficient h on [0, 8), f(x) is called T-generator integrity cluster, denoted as f(x, h).

Definition 3: If S-generator g(x) is continuous along with generalized correlation coefficient h on [0, 8), g(x) is called S-generator integrity cluster, denoted as g(x, h).

Definition 4: For T-norm or S-norm on [0, 8), if there is no measure error, i.e., no effect of truth value error k, T(x, y, h) and S(x, y, h) are called 0-level T-norm integrity cluster and 0-level S-norm integrity cluster respectively. If there is effect of truth value error k, they are called 1-level T-norm integrity cluster and 1-level S-norm integrity cluster respectively, denoted as T(x, y, h, k) and S(x, y, h, k), respectively.

Theorem 2: Exponential function cluster:

$$f(x) = x^m/((1+x)^m-x^m)$$

is 0-level T-generator integrity cluster on [0, 8), in which $m = (3-4h)/4h(1-h)$, $m\rightarrow R$, $h\rightarrow[0, 1]$.

Proof:

- Exponential function cluster f(x) is a continuous and strictly monotone function on [0, 8):
  - When $m>0$, f(x) is increasing and $f(0) = 0$, $f(8)\rightarrow 8$, so f(x) is an automorphism increasing T-generator
  - When $m<0$, f(x) is decreasing and $f(0) = -1$, $f(8)\rightarrow 8$, so f(x) is an extensional increasing T-generator

- By equation:

$$m = (3-4h)/4h(1-h)$$

we know that if $h = 0$, then $m\rightarrow 8$; if $h = 0.5$, then $m = 1$; if $h = 0.75$, then $m = 0$; if $h = 1$, then $m\rightarrow -8$. That is h changes from 0 to 1, then m changes from 8 to -8. m is continuously and strictly monotonely changeable along with h, so f(x) is a continuous and strictly monotone function along with generalized correlation h and denoted as:

$$F_0(x, h) = x^m/((1+x)^m-x^m)$$

Since, the effect of generalized self-correlation k is not considered, i.e., the effect of measure error is not considered, so $F_0(x, h)$ is called 0-level T-generator integrity cluster.

How to define and compute the generalized correlation coefficient h in T/S-norm integrity clusters is both theoretically and practically important. There are three ways in research to define the relation between h and m, see reference [3] for detailed discussion.

If the effect of generalized self-correlation coefficient k is considered on 0-level T-generator integrity cluster, we can obtain 1-level T-generator integrity cluster. Sequentially, 1-level T-norm integrity cluster can be achieved.

Basic set of exponential integrity clusters of NT operation model: The basic set of exponential NT model integrity clusters includes 0-level N/T-generator integrity clusters, 1-level N/T-generator integrity clusters, 0-level N/T-norm integrity clusters and 1-level N/T-norm integrity clusters. Directly substituting these basic integrity clusters into NT-base models of NOT and, OR, Implication, Equivalence, Average, Combination and other self-defined propositional connectives on [0, 8), we can obtain 0-level an 1-level flexible operation models on [0, 8).

Basic set of exponential integrity clusters of NT operation model on [0, 8) interval is as follows:

- 0-level N-generator integrity cluster:

$$\phi(x) = x \text{ (i.e., with no measure error, } k = 0.5, n = 1)$$

- 0-level N-norm integrity cluster:

$$N(x) = 1/x \text{ (i.e., central N norm)}$$

- 0-level T-generator integrity cluster:

$$F_0(x, h) = x^m/((1+x)^m-x^m) = 1/((1+1/x)^m-1)$$

- 0-level T-norm integrity cluster:

$$T(x, y, h) = F_0^{-1}(\max(F_0(0, h), (F_0(x, h)F_0(y, h)-1)/(2+F_0(x, h)+F_0(y, h))), h)$$

- 1-level N-generator integrity cluster:

$$\phi_2(x, k) = x^n/((1+x)^n-x^n)$$

- 1-level N-norm integrity cluster:

$$N(x, k) = ((1+x)^n-x^n)^{1/n}/((1+x)-((1+x)^n-x^n)^{1/n})$$

- 1-level T-generator integrity cluster:

$$F(x, h, k) = F_0(\phi_2(x, k), h) = x^{mn}/((1+x)^{mn}-x^{mn}) = 1/((1+1/x)^{mn}-1)$$

- 1-level T-norm integrity cluster:

$$T(x, y, h, k) = F^{-1}(\max(F(0, h, k), (F(x, h, k)F(y, h, k)-1)/(2+F(x, h, k)+F(y, h, k))), h, k)$$

Explanation of several important parameters and their relations:

n = Position mark parameter of N-generator integrity cluster
k = Generalized self-correlation coefficient
m = Position mark parameter of T-generator integrity cluster
h = Generalized correlation coefficient
n = $-1/\log_2 k$, $k \rightarrow [0, 1]$
k = $2^{-1/n}$, $n \rightarrow R_+$
m = $(3-4h)/4h(1-h)$, $h \rightarrow [0, 1]$
h = $((1+m)-((1+m)^2-3m)^{1/2})/(2m)$, $m \rightarrow R$

For the purpose of the integrity of the study, here only gives a simple introduction. Detailed discussion about N-generator integrity cluster and N-norm integrity cluster on [0, 8) can be referred in reference (Fan *et al.*, 2012). More researches about T/S-generator integrity clusters and T/S-norm integrity clusters have been discussed in other studys which are in publication process.

FLEXIBLE AVERAGE LOGIC OPERATION MODEL

Generally Average operation only exits in numerical analysis and decision analysis. There is no Average propositional connective in the traditional logics. Due to the affection of two-value logic, it seems that there in no need to consider Average problems in logic. It is senseless to compute the average value of two propositions with different values (one is true, the other is false). But in continuous-value logics, the Average operation can not be neglected. As their truth value is likely to be any value in their definition domain, however the result of AND operation is not more than the minimum and the result of OR operation is not less than the maximum. Their must be an operation to describe the logic compromise between the minimum and the maximum that is Average operation.

The physical meaning of flexible Average operation is: the result of twice observation and testing on the same object is generally different and the value should be obtained in the logic compromise of two observation results. There are various Average calculation methods, such as arithmetical Average, geometrical Average, harmonic mean and exponential Average, in which there is the difference of isobar and non-isobar. In flexible logics, logic operators are continuous operator clusters between maximum operator and minimum operator, so flexible Average operation can generate all the Average operation above.

Definition of flexible Average logic operation model: All central operation models can be expressed not only as NT (NOT-AND) base model, but also as NS (NOT-OR) base model. Generator integrity clusters are different and expressions of the base model are also different, but the operation models generated by them are the same. The following is Average operation models generated through putting N-generator integrity cluster and T-generator integrity cluster into NT-base models on [0, 1] base space:

$$M(x, y, h, k) = N(F^{-1}(F(N(x, k), h, k)/2+ F(N(y, k), h, k)/2, h, k), k)$$

As discussed above, we substitute 0-level T-generator integrity cluster into NT base model of central Average operation (generated through unilateral infinite expanding, [0, 1]→[0, 8): x' = x/(1-x), the middle element e' = 1), the following definition is proposed.

Definition 5: Substituting 0-level T-generator integrity cluster:

$$F_0(x, h) = x^m/((1+x)^m-x^m)$$

and 0-level N-norm integrity cluster:

$$N(x) = 1/x$$

into 0-level NT base model of Average operation:

$$M(x,y,h) = N\left(F^{-1}\left(\frac{F(N(x),h)+F(N(y),h)+2F(N(x),h)F(N(y),h)}{2+F(N(x),h)+F(N(y),h)},h\right)\right)$$

the flexible operation is called 0-level Average operation, denoted by ®$_h$, where x, y→[0, 8), m = (3-4h)/4h(1-h), h→[0, 1], m→R.

Definition 6: Substituting 1-level T-generator integrity cluster:

$$F(x, h, k) = x^{mn}/((1+x)^{mn}-x^{mn})$$

and 1-level N-norm integrity cluster:

$$N(x, k) = ((1+x)^n-x^n)^{1/n}/((1+x)-((1+x)^n-x^n)^{1/n})$$

into 1-level NT base model of Average operation:

$$\begin{aligned}&M(x,y,h,k)\\&= N(F^{-1}((F(N(x,k),h,k)+F(N(y,k),h,k)\\&\quad +2F(N(x,k),h,k)F(N(y,k),h,k))\\&\quad /(2+F(N(x,k),h,k)+F(N(y,k),h,k)),h,k),k)\end{aligned}$$

the flexible operation is called 1-level Average operation, denoted by è$_{h,\ k}$, where x, y→[0, 8), m = (3-4h)/4h(1-h), h→[0, 1], m→R; n = -1/log$_2$k, k→[0, 1], n→R$_+$.

For other 0-level and 1-level logic operation models, the same method could be employed.

Four special flexible Average operators: In order to ensure the 0-level integrity of Average flexible propositional connective operation model, it is continuous from maximal Average operator, passing through the probability Average operator and central Average operator to the minimal Average operator, there are several special operators in the operator clusters. For example, when k = 0.5(no measure error) and h = 1, h = 0.75, h = 0.5, h = 0 respectively, the operator clusters are corresponding to the four special operators: Zadeh operator (maximal operator), probability operator, bounded operator (central operator) and drastic operator (minimal operator).

Theorem 3: If h = 1, then m→-8, binary 0-level flexible Average operator is maximal Average operator, also called Zadeh Average operator:

$$M(x, y, 1) = \max(x, y)$$

Proof: As $M(x, y, 1) = (2(1+x)^m(1+y)^m/((1+x)^m+(1+y)^m))^{1/m}-1$:

$$\lim_{m\to-\infty}(2(1+x)^m(1+y)^m/((1+x)^m+(1+y)^m))^{1/m}$$

$$= \lim_{m\to-\infty} e^{\ln(2(1+x)^m(1+y)^m/((1+x)^m+(1+y)^m))^{1/m}} = \lim_{m\to-\infty} e^{\frac{\ln(2(1+x)^m(1+y)^m/((1+x)^m+(1+y)^m))}{m}}$$

$$= e^{\lim_{m\to-\infty}\frac{\ln(2(1+x)^m(1+y)^m/((1+x)^m+(1+y)^m))}{m}}$$

By L'Hospital's rule:

$$\lim_{m\to-\infty}\frac{\ln(2(1+x)^m(1+y)^m/((1+x)^m+(1+y)^m))'}{m'}$$

$$= \lim_{m\to-\infty}\frac{(1+y)^m\ln(1+x)+(1+x)^m\ln(1+y)}{(1+x)^m+(1+y)^m}$$

$$= \lim_{m\to-\infty}\left(\left(\frac{1}{1+(\frac{1+x}{1+y})^m}\right)\ln(1+x)+\left(\frac{1}{1+(\frac{1+y}{1+x})^m}\right)\ln(1+y)\right) \quad (1)$$

- if x = y, Eq. 1 is ln(1+x) = ln(1+y), so M(x, y, 1) = x = y
- if x>y, there is 1+x>1+y, Eq. 3 is ln(1+x), so M(x, y, 1) = x
- if x<y, there is 1+x<1+y, Eq. 3 is ln(1+y), so M(x, y, 1) = y

  So, M(x, y, 1) = max(x, y).

Theorem 4: If h = 0.75, then m→0, binary 0-level flexible Average operator is probability Average operator:

$$M(x, y, 0.75) = ((1+x)(1+y))^{1/2}-1$$

Proof: Similar to Theorem 1, we have:

$$\lim_{m\to 0}\frac{(1+y)^m\ln(1+x)+(1+x)^m\ln(1+y)}{(1+x)^m+(1+y)^m}$$

$$=\lim_{m\to 0}\frac{\ln(1+x)+\ln(1+y)}{2}=\lim_{m\to 0}\ln(1+x)^{1/2}(1+y)^{1/2}$$

So, $M(x, y, 0.75) = ((1+x)(1+y))^{1/2}-1$.

Theorem 5: If h = 0.5, then m = 1, binary 0-level flexible Average operator is bounded Average operator, i.e., central Average operator:

$$M(x, y, 0.5) = (x+y+2xy)/(2+x+y)$$

Proof: As:

$$M(x, y, 0.5) = (2(1+x)^m(1+y)^m/((1+x)^m+(1+y)^m))^{1/m}-1 = (x+y+2xy)/(2+x+y)$$

So, $M(x, y, 0.5) = (x+y+2xy)/(2+x+y)$.

Theorem 6: If h = 0, then $m\to 8$, binary 0-level flexible Average operator is minimal Average operator, also called drastic operator Average operator:

$$M(x, y, 0) = \min(x, y)$$

Proof: Similar to Theorem 1, the detailed proof is omitted.

APPROXIMATE SUPPORT VECTORS(ASVs) SELECTION

It is known that only support vectors in SVM make contributions to the classification decision. Support vectors are the samples which are close to or on the classification margin. If the support vectors or ASVs (samples which have possibility to be support vectors) could be selected before training, training samples will be reduced greatly and then the training efficiency will be raised. Based on the characteristics of forces between negative and positive electric charges-like charges repel each other, but opposite charges attract, we design the following algorithm.

Definition of attraction between samples: Suppose there are two classes of samples called positive samples and negative samples respectively. If x_i belongs to positive class, then $y_i = 1$; otherwise $y_i = -1$. If x_i and x_j are linear separable, then the distance r_{ij} between x_i and x_j is Euclidean distance, $r_{ij} = \|x_i-x_j\|_2$; If xi and xj are linear un-separable, then the distance r_{ij} between x_i and x_j is Hilbert space distance, $r_{ij}=\sqrt{K(x_i,x_i)+K(x_j,x_j)-2K(x_i,x_j)}$, $K(x_i, x_j)$ is kernel function. For any x_i in the training samples, the attraction it received from the same class samples is denoted as AF_i, the attraction it received from the different class samples is denoted as UAF_i. Figure 1a shows the force distribution in which the attraction is from samples in the different class; Fig. 1b shows the force distribution in which the repulsion is from samples in the same class; In Fig. 1a, the darker the color is, the bigger the attraction is from other class; In Fig. 1b, the

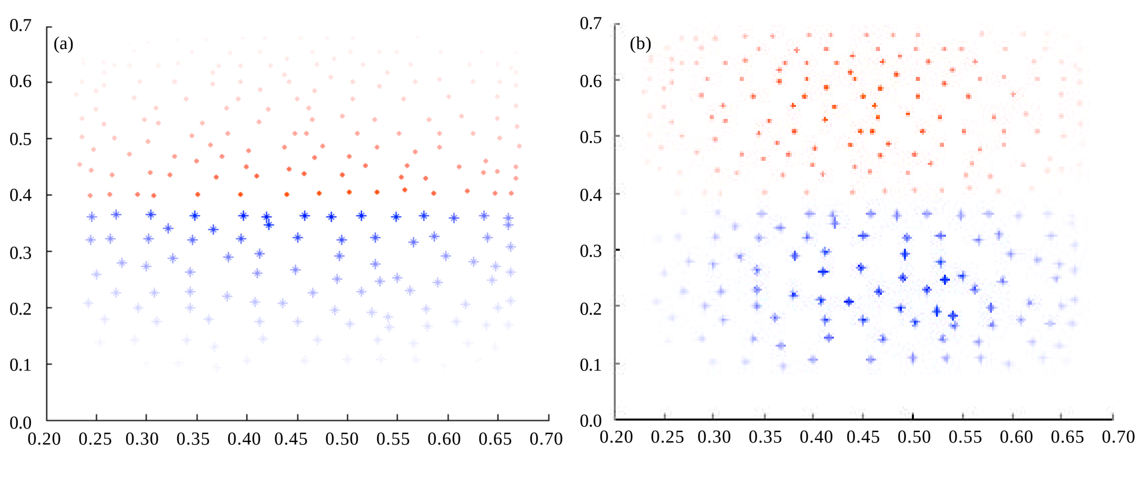

Fig. 1(a-b): Force distribution chart (a) Forces distribution among samples in different classes and (b) Forces distribution among samples in same classes darker the color is, the bigger the repulsion is from the same class. The selection of ASVs is to maintain the samples close to the boundary of different classes as far as possible. Generally, the selection of ASVs is a compromise of the attraction and repulsion. Because we will compute the average value of the two factors as the criterion of the ASVs selection, here we revise AF and consider the effect of UAF to the selection is bigger than that of AF. AF_i and UAF_i are defined as follows:

$$AF_i = \max\left(\sum_{x_k,x_j \in C}\left(\frac{|y_k y_j|}{r_{kj}^2}\right)\right) - \sum_{x_i,x_j \in C}\left(\frac{|y_i y_j|}{r_{ij}^2}\right) + \min\left(\sum_{x_k,x_j \in C}\left(\frac{|y_k y_j|}{r_{kj}^2}\right)\right)$$

in which C is the sample set in which samples and x_i belong to the same class; $i \rightarrow j$, $k \rightarrow j$; i, j, k = 1, 2, ..., m_+, m_+ is the number of samples in set C:

$$UAF_i = \sum_{x_i \in C, x_{j1} \in \bar{C}} \frac{|y_i y_{j1}|}{r_{ij1}^2} + \max\left(\sum_{x_k,x_{j2} \in C}\left(\frac{|y_k y_{j2}|}{r_{kj2}^2}\right)\right)$$

in which C is the sample set in which samples and x_i belong to the same class; $\bar{C}$ is the sample set in which samples and x_i belong to the different class; $i \rightarrow j1$, $k \rightarrow j2$; j1 = 1, 2, ..., m_-, m_- is the number of samples in set $\bar{C}$; j2 = 1, 2, ..., m_+, m_+ is the number of samples in set C.

Since, $y_i\, y_j = 1$, $y_i\, y_{j1} = -1$, $y_i\, y_{j2} = 1$, the above equationscould be expressed as:

$$AF_i = \max\left(\sum_{x_k,x_j \in C}\left(\frac{1}{r_{kj}^2}\right)\right) - \sum_{x_i,x_j \in C}\left(\frac{1}{r_{ij}^2}\right) + \min\left(\sum_{x_k,x_j \in C}\left(\frac{1}{r_{kj}^2}\right)\right)$$

$$UAF_i = \sum_{x_i \in C, x_{j1} \in \bar{C}} \frac{1}{r_{ij1}^2} + \max\left(\sum_{x_k,x_{j2} \in C}\left(\frac{1}{r_{kj2}^2}\right)\right)$$

Algorithm of ASVs selection based on flexible Average operator

Algorithm description: Because the definitions of attraction AF and UAF between samples are based on the square of distance, the definition domain of all the data is in [0, 8) interval. With the comprehensive consideration of the two kinds of attraction, the algorithm of ASVs selection based on flexible Average operator (ASVs_FAO) is described as follows:

Step 1: Determine the threshold value δ of data reduction percentage, let Sup_Vec = Φ, Sup_Vec_p = Φ, Sup_Vec_n = Φ, MAF_p = Φ, MAF_n = Φ

Step 2: Compute UAF_{ip}, AF_{ip} of positive sample x_i and UAF_{in}, AF_{in} of negative sample x_i respectively

Step 3: Sort UAF_{ip}, UAF_{in}, AF_{ip} and AF_{in} in descending order and assign the results to sets $SUAF_p$, $SUAF_n$, SAF_p and SAF_n respectively

Step 4: Compute generalized correlation coefficient h between UAF_{ip} and AF_{ip} of positive sample x_i; compute generalized correlation coefficient h between UAF_{in} and AF_{in} of negative sample x_i

Step 5: For any $x_i \rightarrow C_p$

if h = 1, $MF_{ip} = \max(UAF_{ip}, AF_{ip})$

if h = 0.75, $MF_{ip} = ((1+UAF_{ip})(1+AF_{ip}))^{1/2}-1$

if h = 0.5, $MF_{ip} = (UAF_{ip}+AF_{ip} +2.UAF_{ip}.AF_{ip})/(2+UAF_{ip}+AF_{ip})$

if h = 0, $MF_{ip} = \min(UAF_{ip}, AF_{ip})$

Otherwise:

$$MF_{ip} = \left(\frac{2(1+UAF_{ip})^m (1+AF_{ip})^m}{(1+UAF_{ip})^m + (1+AF_{ip})^m}\right)^{1/m} - 1$$

Similarity, for any $x_i \rightarrow C_n$, compute Mf_{in}. Here, m = (3-4h)/4h(1-h), $h \rightarrow [0, 1]$.

Step 6: For any $x_i \rightarrow C_p$, $MAF_{ip} = MAF_{ip} \rightarrow \{MF_{ip}\}$; for any $x_i \rightarrow C_n$, $MAF_{in} = MAF_{in} \rightarrow \{MF_{in}\}$

Step 7: Sort MAF_{ip} and MAF_{in} in descending order

Step 8: For any $x_i \rightarrow C_p$, suppose q is the serial number in MAF_{ip}. If q = δ.$|MAF_{ip}|$, $Sup\_Vec_p = Sup\_Vec_p \rightarrow \{x_i\}$; Similarity, for any $x_i \rightarrow C_n$, If q = δ.$|MAF_{in}|$, $Sup\_Vec_n = Sup\_Vec_n \rightarrow \{x_i\}$

Step 9: $Sup\_Vec = Sup\_Vec_p \rightarrow Sup\_Vec_n$, finish execution

Here, $|MAF_{ip}|$, is the number of positive samples, $|MAF_{in}|$ is the number of negative samples. According to the different number and distribution of positive samples and negative samples, δ could be defined as different value.

Parameter estimation and experiment results: The problem is different, the definition of generalized correlation coefficient h is different. How to determine h is a key problem in the practical applications. In ASVs selection algorithm, parameter h is defined as follows.

For any $x_i \rightarrow C_p$:

- If $j1 \le m_+.\mu$, h = 1
- If $j2 \ge m_+.\mu$, h = 0
- Otherwise, s1 = 1-(j1-1)/(m_+-1), s2 = 1-(j2-1)/(m_+-1), h = (s1+s2)/2.

Here, j1 is the serial number in $SUAF_p$, j2 is the serial number in SAF_p; 0 = μ = 1, usually 10% = μ = 20%. Similarity, the same for any $x_i \rightarrow C_n$.

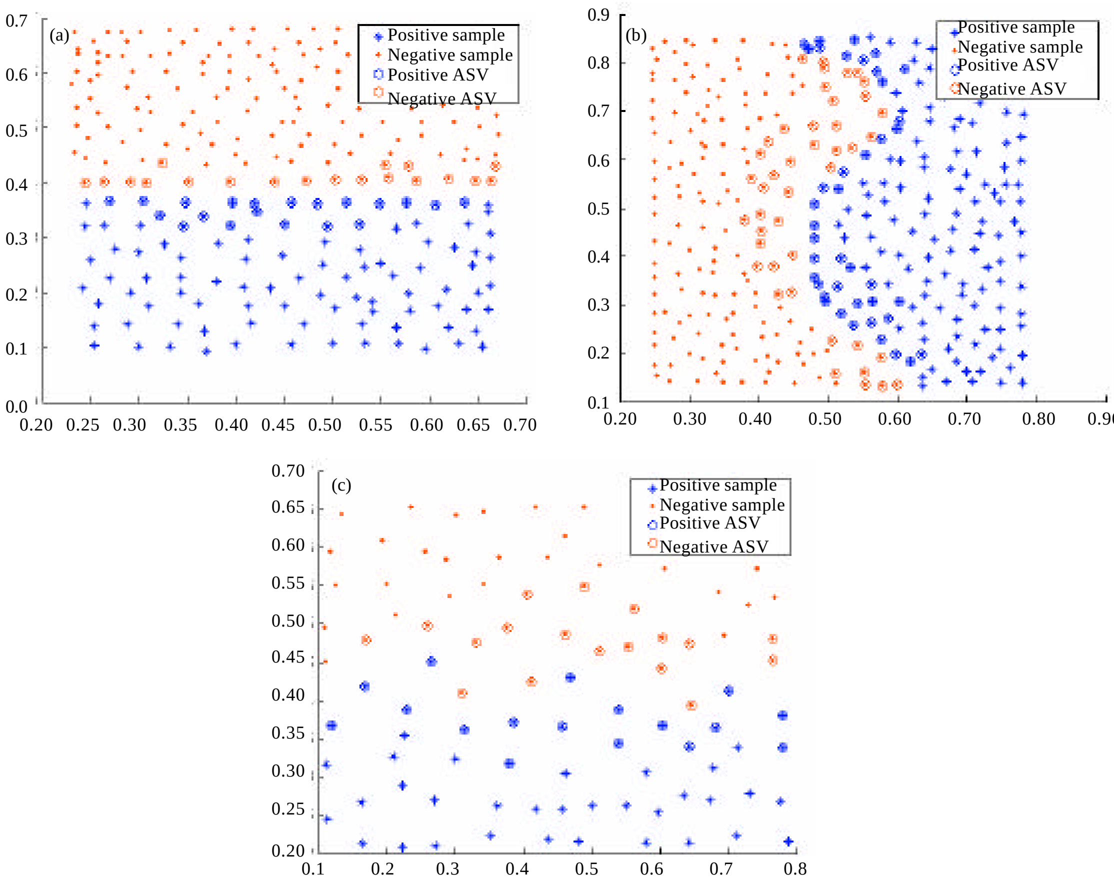

Fig. 2(a-c): Sample distribution chart (a) Linear separable and boundary clear, 102 positive samples and 159 negative samples, δ = 0.2, (b) linear un-separable and boundary clear, 165 positive samples and 174 negative samples, δ = 0.25 and (c) linear un-separable and boundary unclear, 49 positive samples and 52 negative samples, δ = 0.35

We use the parameter estimation method for h described above to apply the flexible Average operation model on [0, 8) to the ASVs selection algorithm. Meanwhile, for simplicity and without losing the universality, we use three different random sample distributions in two dimension space to illustrate, as shown in Fig. 2. Here, each simple has two attributes (i.e. two - dimensional coordinate) and belongs to one class ("." or "*"). There are three cases, classification surfaces in (a) and (b) are clear, but samples in (a) are linear separable and samples in (b) are linear un-separable; classification surfaces in (c) is unclear and the samples are linear un-separable; From Fig. 2, we can see that ASVs close to or on the boundary could be selected for both cases. If these ASVs are used to train the SVM, the majority of the samples will be reduced, the training efficiency will be raised greatly. And we have proven this in other experiments.

CONCLUSION

Average operation reflects the logical compromise of two different truth value x, y to the same proposition. The result of the compromise is required between x and y and satisfies the idempotent property. Obviously, for the flexible Average operation, the range of the value is on [x, y]; if x, y is partial false, then M(x, y, h, k) is partial false; if x, y is partial true, then M(x, y, h, k) is partial true, but it is not satisfied on the contrary; if M(x, y, h, k) is partial false, then min(x, y) is partial false; if M(x, y, h, k) is partial true, then max(x, y) is partial true, but it is not satisfied on the contrary. These cannot be expressed by logic AND or OR.

In practical applications, many logic reasoning controls must be accomplished in their own definition domains. Physical meaning and the operation rule are clearer if the satisfactory degrees are expressed naturally.

For example, it is more direct to use [0, 8) to represent the degree in no upper limit problems. So, the definitions of flexible logic operation models on [0, 8) interval are necessary.

In flexible logics, operation models are generated based on the principal of modifying central base models by generators. This study generate [0, 8) value flexible Average operation model by using [0, 8) value exponential T-generator cluster to modify the central Average operation model. Meanwhile, the study designs an algorithm to use the model for ASVs selection. Reference (Jia, 2009) discussed the equal/unequal weighted flexible Average operation model on [0, 1] and [a, b] space. In fact, the same problems on [0, 8) interval also deserve further research.

Through the Average operation model generation process, we noticed that flexible logic operation models seem too complex, but that is truly what they are. Usually there are two ways to reduce the complexity in applications-by software or by hardware. We can program standard modules to achieve the models' function; we can also implement the operations models by physical apparatus.

ACKNOLEDGMENTS

This study is supported by the National High Technology Research and Development Program of China under Grant No. 2012AA101608, the Science-Technology Project of Zhengzhou of China under Grant No. 2010GYXM364 and the High Level Talent Foundation of Henan University of Technology under Grant No.2010BS027.

REFERENCES

Chen, J.L., H.C. He, C.X. Liu and M.X. Luo, 2011. Integrity studies on 0-level universal operation models of flexible logic. J. Beijing Univ. Posts Telecommun., 34: 10-13.

Chen, Z.C., 2004. Studies on correlation reasoning of fractal, chaos and logic in complex system. Northwestern Polytechnical University, Xi'an.

Chen, Z.C., M.Y. Mao, H.C. He and W.K. Yang, 2006. Generalized T-norm and Fractional "AND" Operation Model. In: Rough Sets and Knowledge Technology, Wang, G.Y., J.F. Peters, A. Skowron and Y. Yao (Eds.). Springer-Verlag, Berlin, Heidelberg, ISBN: 978-3-540-36297-5, pp: 586-591.

Chen, Z.C., M.Y. Mao, W.K. Yang and H.C. He, 2006. A New Uniform OR Operation Model based on Generalized S-norm. In: Computational Intelligence, Huang, D.S., K. Li and G.W. Irwin (Eds.). Springer-Verlag, Berlin, Heidelberg, ISBN: 978-3-540-37274-5, pp: 96-101.

Fan, Y.F., H.C. He and L.R. Ai, 2012. N-Norm on [0,8) and method for calculating generalized self-correlation coefficient k. J. Northwestern Polytech. Univ., 28: 270-275.

He, H.C., H. Wang and Y.H. Liu, 2005. Principle of Universal Logics. Science Press, Beijing, China.

Jia, P.T., 2009. Studies on time series data mining based on the flexible logic. Northwestern Polytechnical University, Xi'an.

Jin, Y., H.C. He and L.R. Ai, 2004. Lane of parallel through carry in ternary optical adder. Sci. China Ser. F: Inform. Sci., 48: 107-116.

Jin, Yi, H.C. He and Y.T. Lv, 2005. Ternary optical computer architecture. Phys. Scripta, T118: 98-101.

Luo, M.X. and H.C. He, 2005. A Propositional Calculus Formal Deductive System L^U of Universal Logic and Its Completeness. In: Fuzzy Systems and Knowledge Discovery, Wang, L. and Y.C. Jin (Eds.). Springer-Verlag, Berlin, Heidelberg, ISBN: 978-3-540-28312-6, pp: 31-40.

Ma, Y.C. and H.C. He, 2006. The Axiomatization for 0-Level Universal Logic. In: Advances in Machine Learning and Cybernetics, Yeung, D.S., Z.Q. Liu, X.Z. Wang and H. Yan (Eds.). Springer-Verlag, Berlin, Heidelberg, ISBN: 978-3-540-33584-9, pp: 367-376.

Mao, M.Y., Z.C. Chen and H.C. He, 2006. A new uniform neuron model of generalized logic operators based on [a,b]. Int. J. Pattern Recogn. Artif. Intell., 20: 159-171.

Xue, Z.A., Y.C. Ma and H.C. He, 2008. A new kind of the flexible interval r-implication. Int. J. Comput. Sci. Knowledge Eng., 2: 117-136.

Zhang, X.H., H.C. He and Y. Xu, 2006. A fuzzy logic system based on Schweizer-Sklar t-norm. Sci. China Ser. F: Inform. Sci., 49: 175-188.