Hindawi Publishing Corporation
Mathematical Problems in Engineering
Volume 2016, Article ID 9329812, 14 pages
http://dx.doi.org/10.1155/2016/9329812

Research Article

Binary Classification of Multigranulation Searching Algorithm Based on Probabilistic Decision

Qinghua Zhang[1,2] and Tao Zhang[1]

[1] *The Chongqing Key Laboratory of Computational Intelligence, Chongqing University of Posts and Telecommunications, Chongqing 400065, China*
[2] *School of Science, Chongqing University of Posts and Telecommunications, Chongqing 400065, China*

Correspondence should be addressed to Qinghua Zhang; zhangqh@cqupt.edu.cn

Received 6 June 2016; Revised 5 September 2016; Accepted 26 September 2016

Academic Editor: Kishin Sadarangani

Multigranulation computing, which adequately embodies the model of human intelligence in process of solving complex problems, is aimed at decomposing the complex problem into many subproblems in different granularity spaces, and then the subproblems will be solved and synthesized for obtaining the solution of original problem. In this paper, an efficient binary classification of multigranulation searching algorithm which has optimal-mathematical expectation of classification times for classifying the objects of the whole domain is established. And it can solve the binary classification problems based on both multigranulation computing mechanism and probability statistic principle, such as the blood analysis case. Given the binary classifier, the negative sample ratio, and the total number of objects in domain, this model can search the minimum mathematical expectation of classification times and the optimal classification granularity spaces for mining all the negative samples. And the experimental results demonstrate that, with the granules divided into many subgranules, the efficiency of the proposed method gradually increases and tends to be stable. In addition, the complexity for solving problem is extremely reduced.

1. Introduction

With the rapid development of modern science and technology, the daily information which people are facing is dramatically increasing, and it is urgent to find a simple and effective way to process the complex information. So the rise and development of multigranulation computing [1–8] have been promoted by this demand. Information granulation had attracted researchers' great attention since a paper that focused on discussing information on granulation published by professor Zadeh in 1997 [9]. In 1985, a paper named "Granularity" was published by Professor Hobbs in the International Joint Conference on Artificial Intelligence held in Los Angeles, United States. It focuses on the granulation of decomposition and synthesis and how to obtain and generate different granularity [10]. These studies play a leading role not only in granular computing methodology [11–17], but also in dealing with complex information [18–22]. Subsequently, the number of researches focused on granular computing has rapidly increased. Many scholars successfully use the basic theoretical model of multigranulation computing to deal with practical problems [23–29]. Currently, multigranulation computing becomes a basic theoretical model to solve the complex problems and discover knowledge from mass information [30, 31].

Multigranulation computing method is mainly aimed to establish a multilevels or multidimensional computing model, and then we need to find the solving way and synthesis strategies in different granularity spaces for solving complex problem. Complex information will be subdivided into lots of simple information in the different granularity spaces [32–35]; then, effective solutions will be obtained by data mining and knowledge discovery techniques to deal with simple information. So it can solve the complex problems in different granularity or dimensions. Aiming at large-scale binary classification problem, it is an important issue to determine the class of all objects in domain by using as little times as possible. And this issue attracts a large number of

researchers' attention [36–39]. Supposing we have a binary classification algorithm and N is the number of all objects in domain and p is the probability of negative samples in domain. Then we need to give the class of all objects in domain. The traditional method will search each object one by one by the binary classification algorithm and it is simple but has many heavy workloads when facing a large number of objects set. Therefore, some researchers have proposed group classification that each object is composed of a lot of samples, and this method improves the searching efficiency [40–42]. For example, how can we effectively complete the binary classification tasks with minimum classification times when facing the massive blood analysis? The traditional method is to test per person once. However, it is a heavy workload when facing many thousands of objects. But, to a certain extent, the single-level granulation method (namely, single-level group testing method) can reduce the workload.

In 1986, Professor Mingmin and Junli proposed that using the single-level group testing method can reduce the workloads of massive blood analysis when the prevalence rate p of a sickness is less than about 0.3 [43]. In this method, all objects will be subdivided into many small subgroups, and then every subgroup will be tested. If the testing result of a subgroup is sickness, we will diagnose the result of each object by testing all the objects in this subgroup one by one. Else if its result is health, we can diagnose that all objects in this subgroup are heathy. But for millions and even billions of objects, can this kind of single-level granulation method still effectively solve the complex problem? At present, there are lots of methods about the single-level granulation searching model [44–54], but the studies on the multilevels granulation searching model are few [55–59]. A binary classification of multilevels granulation searching algorithm, namely, establishing an efficient multigranulation binary classification searching model based on hierarchical quotient space structure, is proposed in this paper. This algorithm combines the falsity and truth preserving principles in quotient space theory and mathematical expectations theory. Obviously, on the assuming that p is the probability of negative samples in domain, the smaller p, the higher efficiency of this algorithm. A large number of experimental results indicate that the proposed algorithm has high efficiency and universality.

The rest of this paper is organized as follows. First, some preliminary concepts and conclusions are reviewed in Section 2. Then, the binary classification of multigranulation searching algorithm is proposed in Section 3. Next, the experimental analysis is discussed in Section 4. Finally, the paper is concluded in Section 5.

2. Preliminary

For convenience, some preliminary concepts are reviewed or defined at first.

Definition 1 (quotient space model [60]). Suppose that triplet (X, F, T) describes a problem space or simply space (X, F), where X denotes the universe and F is the structure of universe X. T indicates the attributes (or features) of universe X. Suppose that X represents the universe with the finest granularity. When we view the same universe X from a coarser granularity, we have a coarse-granularity universe denoted by $[X]$. Then we can have a new problem space $([X], [F], [T])$. The coarser universe $[X]$ can be defined by an equivalence relation R on X. That is, an element in $[X]$ is equivalent to a set of elements in X, namely, an equivalence class X. So $[X]$ consists of all equivalence classes induced by R. From F and T, we can define the corresponding $[F]$ and $[T]$. Then we have a new space $([X], [F], [T])$ called a quotient space of original space (X, F, T).

Theorem 2 (falsity preserving principle [61]). *If a problem $[A] \to [B]$ on quotient space $([X], [F], [T])$ has no solution, then problem $A \to B$ on its original space (X, F, T) has no solution either. In other words, if $A \to B$ on (X, F, T) has a solution, then $[A] \to [B]$ on $([X], [F], [T])$ has a solution as well.*

Theorem 3 (truth preserving principle [61]). *A problem $[A] \to [B]$ on quotient space $([X], [F], [T])$ has a solution, if for $[x]$, $h^{-1}([x])$ is a connected set on X, and problem $A \to B$ on (X, F, T) has a solution. $h : X \to [X]$ is a natural projection that is defined as follows:*

$$h(x) = [X],$$
$$h^{-1}(u) = \{x \mid h(x) \in u\}. \tag{1}$$

Definition 4 (expectation [62]). Let X be a discrete random variable. The expectation or mean of X is defined as $\mu = E(X) = \sum_x xp(X = x)$, where $p(X = x)$ is the probability of $X = x$.

In the case that X takes values from an infinite number set, μ becomes an infinite series. If the series converges absolutely, we say the expectation $E(X)$ exists; otherwise, we say that the expectation of X does not exist.

Lemma 5 (see [43]). *Let $f_q(x) = 1/x + 1 - q^x$ $(x = 2, 3, 4, \ldots, q \in (0, 1))$ be a function with integer variable. If $f_q(x) < 1$ always holds for all x, then $q \in (e^{-e^{-1}}, 1)$.*

Lemma 5 is the basis of the following discussion.

Lemma 6 (see [42, 63]). *Let $f(x) = [x]$ denote a top integral function. And let $f_q(x) = 1/x + 1 - q^x$ $(x = 2, 3, 4, \ldots, q \in (0, 1))$ be a function with integer variable. $f_q(x)$ will reach its minimum value when $x = [1/(p + (p^2/2))]$ or $x = [1/(p + (p^2/2))] + 1$, $x \in (1, 1 - 1/\ln q]$.*

Lemma 7. *Let $g(x) = xq^x$ $(e^{-e^{-1}} < q < 1)$ be a function. If $2 \le c \le 1 - 1/\ln q$ and $1 \le i < c/2$, then the inequality $(g(i) + g(c - i))/2 \le (g((c - 1)/2) + g((c + 1)/2))/2 < g(c/2)$ holds.*

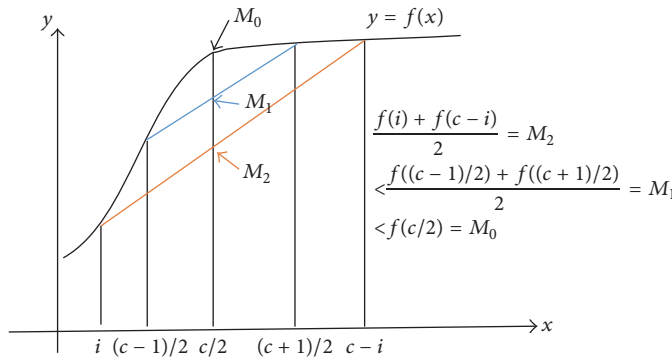

Figure 1: Property of convex function.

Proof. We can obtain the first and second derivatives of $g(x)$ as follows:

$$g'(x) = q^x (1 + x \ln q),$$

$$g''(x) = q^x \ln q (2 + x \ln q),$$

$$g''(x) = \begin{cases} q^x \ln q (2 + x \ln q) < 0, & 1 \le x < x_0, \\ q^x \ln q (2 + x \ln q) = 0, & x = x_0, \\ q^x \ln q (2 + x \ln q) > 0, & x > x_0, \end{cases} \quad (2)$$

$$\text{where } x_0 = -\frac{2}{\ln q},\ e^{-e^{-1}} < q < 1.$$

So, we can draw a conclusion that $g(x)$ is a convex function (property of convex function is shown in Figure 1) when $1 \le x < 1 - 1/\ln q < x_0 = -2/\ln q$. According to the definition of convex function, the inequality $(g(i) + g(c - i))/2 \le (g((c-1)/2) + g((c+1)/2))/2 < g(c/2)$ holds permanently. So Lemma 7 is proved completely. □

3. Binary Classification of Multigranulation Searching Model Based on Probability and Statistics

Granulation is seen as a way of constructing simple theories out of more complex ones [1]. At the same time, the transformation between two different granularity layers is mainly based on the falsity and truth preserving principles in quotient space theory. And it can solve many classical problems such as scale ball game and branch decisions. All the above instances just clearly embody the ideas to solve complex problems with multigranulation methods.

In the second section, the relevant concepts about multigranulation computing theory and probability theory are reviewed. Of course, a multigranulation binary classification searching model not only solves practical complex problems with less cost, but also can be easily constructed. And this model may also play a certain role in the inspiration for the applications of multigranulation computing theory.

Generally, supposing a nonempty finite set $U = \{x_1, x_2, \ldots, x_n\}$, where x_i $(i = 1, 2, 3, \ldots, n)$ is a binary classification object, the triples $(U, 2^U, p)$ is called a probability quotient space with probability p, where p is the negative sample rate in U. 2^U is the structure of universe U. And let $(U_1, U_2, \ldots, U_t)$ $(\bigcup_{i=1}^{t} U_i = U,\ U_m \cap U_n = \Phi,\ m, n \in \{1, 2, \ldots, t\},\ m \ne n)$ be called a random partition space on probability quotient space. There is an example of binary classification of multigranulation searching model.

Example 8. On the assumption that many people need to do blood analysis of physical examination for diagnosing a disease (there are two classes, normal stands for health and abnormal stands for sickness), the domain $U = \{x_1, x_2, \ldots, x_n\}$ stands for all the people. Let N denote the number of all people, and p stands for the prevalence rate. So the quotient space of blood analysis of physical examination is $(U, 2^U, p)$. Besides, we also know a binary classifier (or a reagent) that diagnoses a disease by testing blood sample. How can we complete all the blood analysis with the minimal classification times? Namely, how can we determine the class of all objects in domain. There are usually three ways as follows.

Method 9 (traditional method). In order to accurately diagnose all objects, every blood sample will be tested one by one, so this method needs to search N times. This method is just like the classification process of machine learning.

Method 10 (single-level granulation method). This method is to mix k blood samples to a group where k may be $1, 2, 3, \ldots, n$; namely, the original quotient space will be random partition to $(U_1, U_2, \ldots, U_t)$ $(\bigcup_{i=1}^{t} U_i = U,\ U_m \cap U_n = \Phi,\ m, n \in \{1, 2, \ldots, t\},\ m \ne n)$. And then each mixed blood group will be tested once. If the testing result of a group is abnormal, and according to Theorem 2, we know that this abnormal group has abnormal object(s). In order to make a diagnosis, all of the objects in this group should be tested once again one by one. Similarly, if the testing result of a group is normal, and according to Theorem 3, we know that all of the objects are normal in this group. Therefore, all k objects in this group only need one time to make a diagnosis. The binary classifier can also classify the new blood sample that consists of k blood samples, in this process.

If every group has been mixed by large-scale blood samples (namely, k is a large number) and when some

groups are tested to be abnormal, that means lots of objects must be tested once again one by one. In order to reduce the classification times, this paper proposes a multilevels granulation model.

Method 11 (multilevels granulation method). Firstly, each mixed blood group which contains k_1 samples (objects) will be tested once, where k_1 may be $1, 2, 3, \ldots, n$; namely, the original quotient space will be random partition to $(U_1, U_2, \ldots, U_t)$ $(\bigcup_{i=1}^{t} U_i = U,\ U_m \cap U_n = \Phi,\ m, n \in \{1, 2, \ldots, t\},\ m \neq n)$. Next, if some groups are tested to be normal, that means all objects in those groups are normal (health). Therefore, all k_1 objects only need one time to make a diagnosis in this group. Else if some groups are tested to be abnormal, those groups will be subdivided into many smaller subsets (subgroups) which contain k_2 $(k_2 < k_1)$ objects: namely, the quotient space of an abnormal group U_i will be random partition to $(U_{i1}, U_{i2}, \ldots, U_{il})$ $(\bigcup_{j=1}^{l} U_{ij} = U_i,\ U_{im} \cap U_{in} = \Phi,\ m, n \in \{1, 2, \ldots, l\},\ m \neq n,\ l < k_1)$. Finally each subgroup will be tested once again. Similarly, if a subgroup is tested to be normal, it is confirmed that all objects are health in corresponding subgroup, and if a subgroup is tested to be abnormal, it will be subdivided into smaller subgroups which contain k_3 $(k_3 < k_2)$ objects once again. Therefore, the testing results of all objects can be ensured by repeating the above process in a group until the number of objects is equal to 1 or its testing result is normal in a subgroup. Then, the searching efficiency of the above three methods is analyzed as follows.

In Method 9, every object has to be tested once for diagnosing a disease, so it must take up N times in total.

In Method 10, the original problem space is subdivided into many disjoint subspaces (subsets). If some subsets are tested to be normal that means all objects need only to be tested once. Therefore, the classification times can be reduced if the probability p is small enough in some degree [9].

In Method 11, the key is trying to find the optimal multigranulation space for searching all objects, so a multilevels granulation model needs to be established. There are two questions. One is grouping strategy: namely, how many objects are contained in a group? The other one is optimal granulation levels: namely, how many levels should be granulated from the original problem space?

In this paper, we mainly solve the two questions in Method 11. According to the truth and falsity preserving principle in quotient space theory, all normal parts of blood samples could be ignored. Hence, the original problem will be simplified to a smaller subspace. This idea not only reduces the complexity of the problem, but also improves the efficiency of searching abnormal objects.

Algorithm Strategy. Example 8 can be regarded as a tree structure which each node (which stands for a group) is an x-tuple. Obviously, the searching problem of the tree has been transformed into a hierarchical reasoning process in a monotonous relation sequence. The original space has been transformed into a hierarchical structure where all subproblems will be solved in different granularity levels.

Table 1: The probability distribution of Y_1.

Y_1	$1/k_1$	$1/k_1 + 1$
$p\{Y_1 = y_1\}$	q^{k_1}	$1 - q^{k_1}$

Firstly, the general rule can be concluded by analyzing the simplest hierarchy and grouping case which is the single-level granulation. Secondly, we can calculate the mathematical expectation of the classification times of blood analysis. Finally, an optimal hierarchical granulation model will be established by comparing with the expectation of classification times.

Analysis. Supposing that there is an object set $Y = \{y_1, y_2, \ldots, y_n\}$, the prevalence rate is p. So the probability of an object that appears normal in blood analysis is $q = 1 - p$. The probability of a group which is tested to be normal is q^{k_1}, and to be abnormal is $1 - q^{k_1}$, where k_1 is the objects number of a group.

3.1. The Single-Level Granulation. Firstly, the domain (which contains N objects) is randomly subdivided into many subgroups, where each subset contains k_1 objects. In other words, a new quotient space $[Y_1]$ is obtained based on an equivalence relation R. Then supposing that the average classification time of each object is a random variable Y_1, so the probability distribution of Y_1 is shown in Table 1.

Thus, the mathematical expectation of Y_1 can be obtained as follows:

$$\begin{aligned} E_1(Y_1) &= \frac{1}{k_1} \times q^{k_1} + \left(1 + \frac{1}{k_1}\right) \times \left(1 - q^{k_1}\right) \\ &= \frac{1}{k_1} + \left(1 - q^{k_1}\right). \end{aligned} \tag{3}$$

Then, the total mathematical expectation of the domain can be obtained as follows:

$$\begin{aligned} N \times E_1(Y_1) = N \\ \times \left\{\frac{1}{k_1} \times q^{k_1} + \left(1 + \frac{1}{k_1}\right) \times \left(1 - q^{k_1}\right)\right\}. \end{aligned} \tag{4}$$

When the probability of p keeps unchanged and k_1 satisfies inequality $E_1(Y_1) < 1$, this single-level granulation method can reduce classification times. For example, if $p = 0.5$ and $k_1 > 1$, and according to Lemma 5, $E_1(k_1) > 1$ no matter what the value of k_1. Then this single-level granulation method is worse than traditional method (namely, testing every object in turn). Conversely, if $p = 0.001$ and $k_1 = 32$, $E_1(Y_1)$ will reach its minimum value, and the classification time of single-level granulation method is less than the

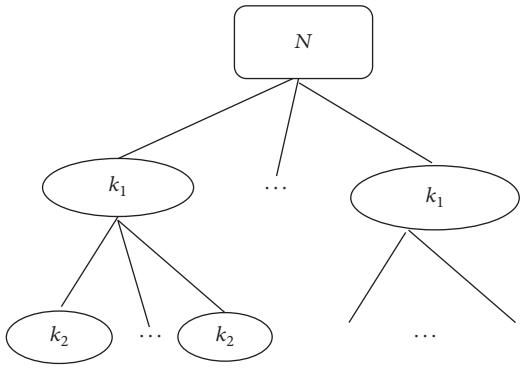

Figure 2: Double-levels granulation.

traditional method. Let $N = 10000$; the total of classification times is approximately equal to 628 as follows:

$$N \times E_1(Y_1) = 10000 \times \left\{\frac{1}{32} \times q^{32} + \left(1 + \frac{1}{32}\right) \times \left(1 - q^{32}\right)\right\} \approx 628. \quad (5)$$

This shows that this method can greatly improve the efficiency of diagnosing and reduce 93.72% classification times in the single-level granulation method. If there is an extremely low prevalence rate, for example, $p = 0.000001$, the total of classification times reaches its minimum value when each group contains 1001 objects (namely, $k_1 = 1001$). If every group is subdivided into many smaller subgroups again and repeating the above method, can the total of classification times be further reduced?

3.2. The Double-Levels Granulation. After the objects of domain are granulated by the method of Section 3.1, the original objects space becomes a new quotient space in which each group has k_1 objects. According to the falsity and truth preserving principles in quotient space theory, if the group is tested to be abnormal, it can be granulated into many smaller subgroups. The double-levels granulation can be shown in Figure 2.

Then, the probability distribution of the double-levels granulation is discussed as follows.

If each group contains k_1 objects and tested once in the 1st layer, the average of classification times is $1/k_1$ for each object. Similarly, the average of classification times of each object is $1/k_2$ in the 2nd layer. When a subgroup contains k_2 objects and is tested to be abnormal, every object has to be retested one by one once again in this subgroup, so the total of classification times of each object is equal to $1/k_2 + 1$.

For simplicity, suppose that every group in the 1st layer will be subdivided into two subgroups which, respectively, contain k_{21} and k_{21} objects in the 2nd layer.

The classification time is shown in Table 2 (✗ represents the testing result which is abnormal and ✓ represents normal).

Table 2: The average classification times of each object with different results.

Times	Objects		
	k_1	k_{21}	k_{22}
	Result		
1	✓	✓	✓
$3 + k_{21}$	✗	✗	✓
$3 + k_{22}$	✗	✓	✗
$3 + k_1$	✗	✗	✗

Table 3: The probability distribution of Y_2.

Y_2	$1/k_1$	$1/k_1 + 1/k_2$	$1/k_1 + 1/k_2 + 1$
$p\{Y_2 = y_2\}$	q^{k_1}	$\left(1 - q^{k_1}\right) \times q^{k_2}$	$\left(1 - q^{k_1}\right) \times \left(1 - q^{k_2}\right)$

For instance, let $k_1 = 8$, $k_{21} = 4$, and $k_{22} = 4$; there are four kinds of cases to happen.

Case 1. If a group is tested to be normal in the 1st layer, so the total of classification times of this group is $k_1 \times 1/k_1 = 1$.

Case 2. If a group is tested to be abnormal in the 1st layer and its one subgroup is tested to be abnormal, and the other subgroup is also tested to be abnormal in the 2nd layer, so the total of classification times of this group is $k_{21} \times (1/k_1 + 1/k_{21} + 1) + k_{22} \times (1/k_1 + 1/k_{22}) = 3 + k_{21} = 7$.

Case 3. If a group is tested to be abnormal in the 1st layer, its one subgroup is tested to be normal, and the other subgroup is tested to be abnormal in the 2nd layer, so the total of classification times of this group is $k_{21} \times (1/k_1 + 1/k_{21}) + k_{22} \times (1/k_1 + 1/k_{22} + 1) = 3 + k_{22} = 7$.

Case 4. If a group is tested to be abnormal in the 1st layer, and its two subgroups are tested to be abnormal in the 2nd layer, so the total of classification times of this group is $k_{21} \times (1/k_1 + 1/k_{21} + 1) + k_{22} \times (1/k_1 + 1/k_{22} + 1) = 3 + k_1 = 11$.

Suppose each group contains k_1 objects in the 1st layer, and their every subgroup has k_2 objects in the 2nd layer. Supposing that the average classification times of each object is a random variable Y_2, then the probability distribution of Y_2 is shown in Table 3.

Thus, in the 2nd layer, the mathematical expectation of Y_2 which is the average classification times of each object is obtained as follows:

$$E_2(Y_2) = \frac{1}{k_1} \times q^{k_1} + \left(\frac{1}{k_1} + \frac{1}{k_2}\right) \times \left(1 - q^{k_1}\right) \times q^{k_2} + \left(1 + \frac{1}{k_1} + \frac{1}{k_2}\right) \times \left(1 - q^{k_1}\right) \times \left(1 - q^{k_2}\right) = \frac{1}{k_1} + \left(1 - q^{k_1}\right) \times \left(\frac{1}{k_2} + 1 - q^{k_2}\right). \quad (6)$$

As long as the number of granulation levels increases to 2, the average classification times of each object will be further reduced: for instance, when $p = 0.001$ and $N = 10000$.

Table 4: The probability distribution of Y_i.

Y_i	$p\{Y_i = y_i\}$
$\frac{1}{k_1}$	q^{k_1}
$\frac{1}{k_1} + \frac{1}{k_2}$	$\left(1 - q^{k_1}\right) \times q^{k_2}$
$\vdots$	$\vdots$
$\sum_{j=1}^{i} \frac{1}{k_j}$	$\left(1 - q^{k_1}\right) \times \left(1 - q^{k_2}\right) \times \cdots \times \left(1 - q^{k_{i-1}}\right) \times q^{k_i}$
$\sum_{j=1}^{i} \frac{1}{k_j} + 1$	$\left(1 - q^{k_1}\right) \times \left(1 - q^{k_2}\right) \times \cdots \times \left(1 - q^{k_{i-1}}\right) \times \left(1 - q^{k_i}\right)$

As we know, the minimum expectation of the total of classification times is about 628 with $k_1 = 32$ in the single-level granulation. And according to (6) and Lemma 6, $E_2(Y_2)$ will reach minimum value when $k_2 = 16$. The minimum mathematical expectation of each object's average classification times is shown as follows:

$$\begin{aligned} N \times E_2(X) = N \times \Bigg\{ & \frac{1}{k_1} \times q^{k_1} + \frac{1}{k_1} + \frac{1}{k_2} \times \left(1 - q^{k_1}\right) \\ & \times q^{k_2} + \left(1 + \frac{1}{k_1} + \frac{1}{k_2}\right) \times \left(1 - q^{k_1}\right) \times \left(1 - q^{k_2}\right) \Bigg\} \\ & \approx 338. \end{aligned} \tag{7}$$

The mathematical expectation of classification times can save 96.62% compared with traditional method and save 46.18% compared with single-level granulation method. Next we will discuss i-levels granulation ($i = 3, 4, 5, \ldots, n$).

3.3. The i-Levels Granulation. For blood analysis case, the granulation strategy in ith layer is concluded by the known objects number of each group in previous layers (namely, $k_1, k_2, \ldots, k_{i-1}$ are known and just only k_i is unknown). According to the double-levels granulation method, and supposing that the classification time of each object is a random variable Y_i in the i-levels granulation, so the probability distribution of Y_i is shown in Table 4.

Obviously, the sum of probability distribution is equal to 1 in each layer.

Proof.

Case 1 (the single-level granulation). One has

$$q^{k_1} + 1 - q^{k_1} = 1. \tag{8}$$

Case 2 (the double-levels granulation). One has

$$\begin{aligned} & q^{k_1} + \left(1 - q^{k_1}\right) \times q^{k_2} + \left(1 - q^{k_1}\right) \times \left(1 - q^{k_2}\right) \\ & = q^{k_1} + \left(1 - q^{k_1}\right) \times \left(q^{k_2} + 1 - q^{k_2}\right) \\ & = q^{k_1} + \left(1 - q^{k_1}\right) \times 1 = 1. \end{aligned} \tag{9}$$

Case 3 (the i-levels granulation). One has

$$\begin{aligned} & q^{k_1} + \left(1 - q^{k_1}\right) \times q^{k_2} + \cdots + \left(1 - q^{k_1}\right) \times \left(1 - q^{k_2}\right) \\ & \quad \times \cdots \times \left(1 - q^{k_{i-1}}\right) \times q^{k_i} + \left(1 - q^{k_1}\right) \times \left(1 - q^{k_2}\right) \\ & \quad \times \cdots \times \left(1 - q^{k_i}\right) = q^{k_1} + \left(1 - q^{k_1}\right) \times q^{k_2} + \cdots \\ & \quad + \left(1 - q^{k_1}\right) \times \left(1 - q^{k_2}\right) \times \cdots \times \left(1 - q^{k_{i-1}}\right) \\ & \quad \times \left(q^{k_i} + \left(1 - q^{k_i}\right)\right) = q^{k_1} + \left(1 - q^{k_1}\right) \times q^{k_2} + \cdots \\ & \quad + \left(1 - q^{k_1}\right) \times \left(1 - q^{k_2}\right) \times \cdots \times \left(1 - q^{k_{i-1}}\right) = q^{k_1} \\ & \quad + \left(1 - q^{k_1}\right) \times q^{k_2} + \cdots + \left(1 - q^{k_1}\right) \times \left(1 - q^{k_2}\right) \\ & \quad \times \cdots \times \left(q^{k_{i-1}} + \left(1 - q^{k_{i-1}}\right)\right) = \cdots = q^{k_1} + \left(1 - q^{k_1}\right) \\ & \quad \times 1 = 1. \end{aligned} \tag{10}$$

The proof is completed. □

Definition 12 (classification times expectation of granulation). In a probability quotient space, a multilevels granulation model will be established from domain $U = \{x_1, x_2, \ldots, x_n\}$ which is a nonempty finite set, the health rate is q, the max granular levels is L, and the number of objects in ith layer is k_i, $i = 1, 2, \ldots, L$. So the average classification time of each objects is $E_i(Y_i)$ in ith layer.

$$\begin{aligned} E_i(Y_i) = & \frac{1}{k_1} + \sum_{i=2}^{L} \left[\frac{1}{k_i} \times \prod_{j=1}^{i-1} \left(1 - q^{k_j}\right) \right] \\ & + \prod_{i=1}^{L} \left(1 - q^{k_i}\right). \end{aligned} \tag{11}$$

In this paper, we mainly focus on establishing a minimum granulation expectation model of classification times by multigranulation computing method. For simplicity, the mathematical expectation of classification times will be regarded as the measure of contrasting with the searching efficiency. According to Lemma 5, the multilevels granulation model can simplify the complex problem only if the prevalence rate $p \in (0, 1 - e^{-e^{-1}})$ in the blood analysis case.

Theorem 13. *Let the prevalence rate $p \in (0, 0.3)$; if a group is tested to be abnormal in the 1st layer (namely, this group contains abnormal objects), the average classification times of each object will be further reduced by subdividing this group once again.*

Proof. The expectation difference between the single-level granulation $E_1(Y_1)$ and the double-levels granulation $E_2(Y_2)$ can adequately embody their efficiency. Under the conditions

of $e^{-e^{-1}} < q < 1$ and $1 \le k_2 < k_1$, and according to (3) and (6), the expectation difference $E_1(Y_1) - E_2(Y_2)$ is shown as follows:

$$\begin{aligned} &E_1(Y_1) - E_2(Y_2) \\ &= \frac{1}{k_1} + \left(1 - q^{k_1}\right) \\ &\quad - \left\{\frac{1}{k_1} + \left(1 - q^{k_1}\right) \times \left(\frac{1}{k_2} + 1 - q^{k_2}\right)\right\} \\ &= \left(1 - q^{k_1}\right) \times \left\{1 - \left(\frac{1}{k_2} + 1 - q^{k_2}\right)\right\} > 0. \end{aligned} \tag{12}$$

According to Lemma 5, $(1 - q^{k_1}) > 0$ and $f_q(k_2) = 1/k_2 + 1 - q^{k_2} < 1$ always hold; then we can get $1 - (1/k_2 + 1 - q^{k_2}) > 0$. So the inequality $E_1(X) - E_2(X) > 0$ is proved successfully. □

Theorem 13 illustrates that it can reduce classification times by continuing to granulate the abnormal groups into the 2nd layer when $k_1 > 1$. There is attempt to prove that the total of classification times will be further reduced by continuously granulating the abnormal groups into ith layers until a group's number is no less than 1.

Theorem 14. *Supposing the prevalence rate $p \in (0, 0.3)$, if a group is tested to be abnormal (namely, this group contains abnormal objects), the average classification times of each object will be reduced by continuously subdividing the abnormal group until the objects number of its subgroup is no less than 1.*

Proof. The expectation difference between $(i - 1)$-levels granulation $E_{i-1}(Y_{i-1})$ and i-levels granulation $E_i(Y_i)$ reflects their efficiency. On the condition of $e^{-e^{-1}} < q < 1$ and $1 \le k_i < k_{i-1}$, and according to (11), the expectation difference $E_{i-1}(Y_{i-1}) - E_i(Y_i)$ is shown as follows:

$$\begin{aligned} &E_{i-1}(Y_{i-1}) - E_i(Y_i) = \frac{1}{k_1} + \sum_{l=2}^{i-1}\left[\frac{1}{k_l} \times \prod_{j=1}^{l-1}\left(1 - q^{k_j}\right)\right] \\ &+ \prod_{l=1}^{i-1}\left(1 - q^{k_l}\right) - \left\{\frac{1}{k_1} + \sum_{l=2}^{i}\left[\frac{1}{k_i} \times \prod_{j=1}^{l-1}\left(1 - q^{k_j}\right)\right]\right. \\ &\left. + \prod_{l=1}^{L}\left(1 - q^{k_l}\right)\right\} = \left(1 - q^{k_1}\right) \times \cdots \times \left(1 - q^{k_{i-1}}\right) \\ &\times \left\{1 - \left(\frac{1}{k_i} + 1 - q^{k_i}\right)\right\} > 0. \end{aligned} \tag{13}$$

Because $(1 - q^{k_1}) \times \cdots \times (1 - q^{k_{i-1}}) > 0$ is known, according to Lemma 5 and $k_i \ge 1$, we can get $(1/k_i + 1 - q^{k_i}) < 1$: namely $1 - (1/k_i + 1 - q^{k_i}) > 0$. So $E_{i-1} - E_i > 0$ is proved successfully. □

Theorem 14 shows that this method will continuously improve the searching efficiency in the process of granulating abnormal groups from 1st layer to ith layer because $E_{i-1}(Y_{i-1}) - E_i(Y_i) > 0$ always holds. However, it is found that the classification times cannot be reduced when the objects number of an abnormal group is less than or equal to 4, so the objects of this abnormal group should be tested one by one. In order to achieve the best efficiency, then we will explore how to determine the optimum granulation, namely, how to determine the optimum objects number of each group and how to obtain the optimum granulation levels.

3.4. The Optimum Granulation. It is a difficult and key point to explore an appropriate granularity space for dealing with a complex problem. And it not only requires us to keep the integrity of the original information but also simplify the complex problem. So we take the blood analysis case as an example to explain how to obtain the optimum granularity space in this paper. Suppose the condition $e^{-e^{-1}} < q < 1$ always holds.

Case 1 (granulating abnormal groups from the 1st layer to the 2nd layer). (a) If k_1 is an even number, every group which contains k_1 objects in 1st layer will be subdivided into two subgroups into 2nd layer.

Scheme 15. Supposing the one subgroup of the 2nd layer has i $(1 \le i < k_1/2)$ objects, according to formula (6), the expectation of classification times for each object is $E_2(i)$. And the other subgroup has $(k_1 - i)$ objects, so the expectation of classification times for each object is $E_2(k_1 - i)$. The average expectation of classification times for each object in the 2nd layer is shown as follows:

$$\frac{i \times E_2(i) + (k_1 - i) \times E_2(k_1 - i)}{k_1}. \tag{14}$$

Scheme 16. Suppose every abnormal group in 1st layer is average subdivided into two subgroups: namely, each subgroup has $k_1/2$ objects in the 2nd layer. According to formula (6), the average expectation of classification times for each object in the 2nd layer is shown as follows:

$$\frac{2 \times k_1/2 \times E_2(k_1/2)}{k_1} = \frac{k_1 \times E_2(k_1/2)}{k_1}. \tag{15}$$

The expectation difference between the above two schemes embodies their efficiency. In order to prove that Scheme 16 is more efficient than Scheme 15, we only need to prove that the following inequality is correct: namely,

$$\begin{aligned} &\frac{(i \times E_2(i) + (k_1 - i) \times E_2(k1_1 - i))}{k_1} - \frac{k_1 \times E_2(k_1/2)}{k_1} \\ &> 0, \quad \left(e^{-e^{-1}} < q < 1,\ k_1 > 1\right). \end{aligned} \tag{16}$$

Table 5: The changes of average expectation with different objects number in two groups.

(k_{21}, k_{22})	(1, 15)	(2, 14)	(3, 13)	(4, 12)	(5, 11)	(6, 10)	(7, 9)	(8, 8)
E_2	0.07367	0.07329	0.07297	0.07270	0.07249	0.07234	0.07225	0.07222

Proof. Let $g(x) = xq^x$ $(e^{-e^{-1}} < q < 1)$, and according to Lemma 7, then, we have

$$\begin{aligned}
&\frac{g(i)+g(c-i)}{2} < g\left(\frac{c}{2}\right) \Longrightarrow \\
&\frac{\left(i\times q^i + (k_1-i)\times q^{(k_1-i)}\right)}{2} < \frac{k_1}{2}\times q^{k_1/2} \Longrightarrow \\
&\left(i\times q^i + (k_1-i)\times q^{(k_1-i)}\right) < k_1\times q^{k_1/2} \Longrightarrow \\
&i\times\left(\frac{1}{k_1} + \left(1-q^{k_1}\right)\times\left(1+\frac{1}{i}-q^i\right)\right) + (k_1-i) \\
&\quad\times\left(\frac{1}{k_1} + \left(1-q^{k_1}\right)\times\left(1+\frac{1}{(k_1-i)}-q^{(k_1-i)}\right)\right) \\
&\quad> k_1\left(\frac{1}{k_1} + \left(1-q^{k_1}\right)\times\left(1+\frac{2}{k_2}-q^{k_1/2}\right)\right) \Longrightarrow \\
&\frac{(i\times E_2(i) + (k_1-i)\times E_2(k1_1-i))}{k_1} - \frac{k_1\times E_2(k_1/2)}{k_1} \\
&\quad> 0.
\end{aligned} \tag{17}$$

The proof is completed. □

Therefore, if every group has k_1 (k_1 is an even number and $k_1 > 1$) objects in the 1st layer that need to be subdivided into two subgroup, Scheme 16 is more efficient than Scheme 15.

The experiment results have verified the above conclusion in Table 5. Let $p = 0.004$ and $k_1 = 16$. When every subgroup contains 8 objects in the 2nd layer, the expectation of classification times obtains minimum value for each object, where k_{21} is the number of the one subgroup in the 2nd layer, k_{22} is the number of the other subgroup, and E_2 is the corresponding expectation of classification times for each object.

(b) If k_1 is an even number, every group which contains k_1 objects in 1st layer will be subdivided into three subgroups into 2nd layer.

Scheme 17. In the 2nd layer, if the first subgroup has i $(1 \le i < k_1/2)$ objects, the average expectation of classification times for each object is $E_2(i)$. If the second subgroup has j $(1 \le j < k_1/2)$ objects, the expectation of classification times for each object is $E_2(j)$. Then the third subgroup has $(k_1 - i - j)$ objects, and the average expectation of classification times for each object is $E_2(k_1 - i - j)$. So the average expectation of classification times for each object in the 2nd is shown as follows:

$$\frac{i\times E_2(i) + j\times E_2(j) + (k_1-i-j)\times E_2(k_1-i-j)}{k_1}. \tag{18}$$

Similarly, it is easy to be prove that Scheme 16 is also more efficient than Scheme 17. In other words, we only need to prove the following inequality: namely,

$$\begin{aligned}
&\frac{(i\times E_2(i) + j\times E_2(j) + (k_1-i-j)\times E_2(k_1-i-j))}{k_1} \\
&-\frac{k_1\times E_2(k_1/2)}{k_1} > 0, \quad \left(e^{-e^{-1}} < q < 1,\ k_1 > 1\right).
\end{aligned} \tag{19}$$

Proof. Let $g(x) = xq^x$ $(e^{-e^{-1}} < q < 1)$, and according to Lemma 7, then we have

$$\begin{aligned}
&\frac{g(i)+g(c-i)}{2} < g\left(\frac{c}{2}\right) \Longrightarrow \\
&\frac{\left(t\times q^i + (k_1-t)\times q^{(k_1-t)}\right)}{2} < \frac{k_1}{2}\times q^{k_1/2} \Longrightarrow \\
&\frac{\left(i\times q^i + j\times q^j + (k_1-i-j)\times q^{(k_1-i-j)}\right)}{2} < \frac{k_1}{2}\times q^{k_1/2} \Longrightarrow \\
&\left(i\times q^i + j\times q^j + (k_1-i-j)\times q^{(k_1-i-j)}\right) < k_1\times q^{k_1/2} \Longrightarrow \\
&i\times\left(\frac{1}{k_1} + \left(1-q^{k_1}\right)\times\left(1+\frac{1}{i}-q^i\right)\right) + j \\
&\quad\times\left(\frac{1}{k_1} + \left(1-q^{k_1}\right)\times\left(1+\frac{1}{j}-q^j\right)\right) + (k_1-i-j) \\
&\quad\times\left(\frac{1}{k_1} + \left(1-q^{k_1}\right)\times\left(1+\frac{1}{(k_1-i-j)}-q^{(k_1-i-j)}\right)\right) \\
&\quad> k_1\times\left(\frac{1}{k_1} + \left(1-q^{k_1}\right)\times\left(1+\frac{2}{k_2}-q^{k_1/2}\right)\right) \Longrightarrow \\
&\frac{(i\times E_2(i) + j\times E_2(j) + (k_1-i-j)\times E_2(k_1-i-j))}{k_1} \\
&\quad-\frac{k_1\times E_2(k_1/2)}{k_1} > 0.
\end{aligned} \tag{20}$$

The proof is completed. □

Therefore, if every group which contains k_1 (is an even number and $k_1 > 1$) objects needs to be subdivided into three subgroups in the first layer, Scheme 16 is more efficient than Scheme 17.

The experimental results have verified the above conclusion in Table 6. Let $p = 0.004$ and $k_1 = 16$. When every subgroup contains 8 objects in the 2nd layer, the average expectation of classification times reaches minimum value for each object. In Table 6, the first line stands for the objects number of first group in the 2nd layer and the first row stands for the objects number of second group, and data stands for the corresponding average expectation of classification times. For example, (1, 1, 7.7143) expresses that the objects number

Table 6: The changes of average expectation with different objects number of three groups.

Objects	Objects				
	1	2	3	4	5
	Expectation ($\times 10^{-2}$)				
1	7.7143	7.6786	7.6488	7.6250	7.6072
2	7.6786	7.6458	7.6189	7.5980	7.5831
3	7.6488	7.6189	7.5950	7.5770	7.5651
4	7.6250	7.5980	7.5770	7.5620	7.5530
5	7.6072	7.5831	7.5651	7.5530	7.5470
6	7.5953	7.5742	7.5591	7.5500	7.5470
7	7.5894	7.5712	7.5591	7.5530	7.5530
8	7.5894	7.5742	7.5651	7.5620	7.5651
9	7.5953	7.5831	7.5770	7.5770	7.5980

of three groups, respectively, is 1, 1, and 14, and the average classification time for each object is $E_2 = 0.077143$ in the 2nd layer.

(c) When an abnormal group contains k_1 (even) objects and it needs to be further granulated into the 2nd layer, Scheme 16 still has the best efficient.

There are two granulation schemes. Scheme 18 is that the abnormal groups are randomly subdivided into s ($s < k_1$) subgroups in the 1st layer, and Scheme 16 is that the abnormal groups are averagely subdivided into two subgroups in the 1st layer.

Scheme 18. Supposing an abnormal group will be subdivided into s ($s < k_1$) subgroups. The first group has x_1 ($1 \le x_1 < k_1/2$) objects and the average expectation of classification times for each object is $E_2(x_1)$; the 2nd subgroup has x_2 ($1 \le x_2 < k_1/2$) objects and the average expectation of classification times for each object is $E_2(x_2), \ldots$; ith subgroup has x_i ($1 \le x_i < k_1/2$) objects and the average expectation of classification times for each object is $E_2(x_i), \ldots$; sth subgroup has x_s ($1 \le x_s < k_1/2$) objects and the average expectation of classification times for each object is $E_2(x_s)$. Hence, the average expectation of classification times for each object in the 2nd layer is shown as follows:

$$\frac{1}{k_1} \times \sum_{j=1}^{s} x_j \times E_2\left(x_j\right). \tag{21}$$

Similarly, in order to prove that Scheme 16 is more efficient than Scheme 18, we only need to prove the following inequality: namely,

$$\frac{1}{k_1} \times \sum_{j=1}^{s} x_j \times E_2\left(x_j\right) - \frac{k_1 \times E_2\left(k_1/2\right)}{k_1} > 0, \quad \left(e^{-e^{-1}} < q < 1,\ k_1 > 1\right). \tag{22}$$

Proof. Let $g(x) = xq^x$ ($e^{-e^{-1}} < q < 1$), and according to Lemma 7. Then, we have

$$\begin{aligned} &\frac{\sum_{i=1}^{s} g\left(x_i\right)}{2} < g\left(\frac{c}{2}\right) \Longrightarrow \\ &\frac{\sum_{i=1}^{s} x_i \times q^{x_i}}{2} < \frac{k_1}{2} \times q^{k_1/2} \Longrightarrow \\ &\sum_{j=1}^{s} x_j \times \left(\frac{1}{k_1} + \left(1 - q^{k_1}\right) \times \left(1 + \frac{1}{x_j} - q^{x_j}\right)\right) \\ &\quad > k_1 \times \left(\frac{1}{k_1} + \left(1 - q^{k_1}\right) \times \left(1 + \frac{2}{k_2} - q^{k_1/2}\right)\right) \Longrightarrow \\ &\frac{1}{k_1} \times \sum_{j=1}^{s} x_j \times E_2\left(x_j\right) - \frac{k_1 \times E_2\left(k_1/2\right)}{k_1} > 0. \end{aligned} \tag{23}$$

The proof is completed. □

Therefore, when every abnormal group which contains k_1 (which is an even number and $k_1 > 1$) in the 1st layer objects needs to be granulated into many subgroups, Scheme 16 is more efficient than other schemes.

(d) In a similar way, when every abnormal group which contains k_1 (k_1 is an odd number and $k_1 > 1$) objects in the 1st layer will be granulated into many subgroups, the best scheme is that every abnormal group is uniformly subdivided into two subgroups: namely, each subgroup contains $(k_1 - 1)/2$ or $(k_1 + 1)/2$ objects in the 2nd layer.

Case 2 (granulating abnormal groups from the 1st layer to ith layer)

Theorem 19. *In ith layer, if the objects number of each abnormal group is more than 4, then the total of classification times can be reduced by keeping on subdividing the abnormal groups into two subgroups which contain equal objects as far as possible. Namely, if each group contains k_i objects in ith layer, then each subgroup may contain $k_i/2$ or $(k_1 - 1)/2$ or $(k_1 + 1)/2$ objects in $(i + 1)$th layer.*

Proof. In the multigranulation method, the objects number of each subgroup in the next layer is determined by the objects number of group in the current layer. In other words, the objects number of each subgroup in the $(i + 1)$th layer is determined by the known objects number of each group in ith layer.

According to recursive idea, the process of granulating abnormal group from ith layer into $(i + 1)$th layer is similar to that 1st layer into 2nd layer. It is known that the best efficient way is as far as possible uniformly subdividing an abnormal group in current layer into two subgroups in next layer when granulating abnormal group from 1st layer into 2nd layer. Therefore, the best efficient way is also as far as possible uniformly subdividing each abnormal group in ith layer into two subgroups in $(i + 1)$th layer. The proof is completed. □

Based on k_1 which is the optimum objects number of each group in 1th layer, then the optimum granulation levels

Table 7: The best testing strategy in different layers with the different prevalence rates.

p	E_i	$(k_1, k_2, \ldots, k_i)$
0.01	0.157649743271	(11, 5, 2)
0.001	0.034610328332	(32, 16, 8, 4)
0.0001	0.010508158027	(101, 50, 25, 12, 6, 3)
0.00001	0.003301655870	(317, 158, 79, 39, 19, 9, 4)
0.000001	0.001041044160	(1001, 500, 250, 125, 62, 31, 15, 7, 3)

and their corresponding objects number of each group could be obtained by Theorem 19. That is to say, $k_{i+1} = k_i/2$ (or $k_{i+1} = (k_i - 1)/2$ or $k_{i+1} = (k_i + 1)/2$), where k_i $(k_i > 4)$ is the objects number of each abnormal group in ith $(1 \le i \le s - 1)$ layer, and s is the optimum granulation levels. Namely, in this multilevels granulation method, the final structure of granulating abnormal group from the 1st layer to last layer is similar to a binary tree, and the origin space can be granulated to the structure which contains many binary trees.

According to Theorem 19, multigranulation strategy can be used to solve the blood analysis case. When facing the different prevalence rates, such as $p_1 = 0.01$, $p_2 = 0.001$, $p_3 = 0.0001$, $p_4 = 0.00001$, and $p_5 = 0.000001$, the best searching strategy is that the objects number of each group in the different layers is shown in Table 7 (k_i stands for the objects number of each groups in ith layer and E_i stands for the average expectation of classification times for each object).

Theorem 20. *In the above multilevels granulation method, if p which is the prevalence rate of a sickness (or the negative sample ratio in domain) tends to 0, the average classification times for each object tend to be* $1/k_1$*; in other words, the following equation always holds:*

$$\lim_{p \to 0} E_i = \frac{1}{k_1}. \tag{24}$$

Proof. According to Definition 12, let $q = 1 - p$, $q \to 1$; we have

$$\begin{aligned} E_i = {} & \frac{1}{k_1} + \left(1 - q^{k_1}\right) \times \left(\frac{1}{k_2} + \left(1 - q^{k_2}\right) \times \left(\frac{1}{k_3} + \left(1 - q^{k_3}\right)\right.\right. \\ & \times \left(\frac{1}{k_4} + \left(1 - q^{k_4}\right)\right. \\ & \left.\left.\left.\times \left(\cdots \times \left(\frac{1}{k_{i-1}} + \left(1 - q^{k_{i-1}}\right) \times \left(\frac{1}{k_i} + \left(1 - q^{k_i}\right)\right)\right) \cdots\right)\right)\right)\right). \end{aligned} \tag{25}$$

According to Lemma 6, $k_1 = [1/(p + (p^2/2))]$ or $k_1 = [1/(p + (p^2/2))] + 1$. And then let

$$\begin{aligned} T = {} & \frac{1}{k_2} + \left(1 - q^{k_2}\right) \times \left(\frac{1}{k_3} + \left(1 - q^{k_3}\right) \times \left(\frac{1}{k_4} + \left(1\right.\right.\right. \\ & \left.\left.\left. - q^{k_4}\right) \times \left(\cdots\right.\right.\right. \\ & \left.\left.\left. \times \left(\frac{1}{k_{i-1}} + \left(1 - q^{k_{i-1}}\right) \times \left(\frac{1}{k_i} + \left(1 - q^{k_i}\right)\right)\right)\right.\right. \\ & \left.\left.\left. \cdots\right)\right)\right), \end{aligned} \tag{26}$$

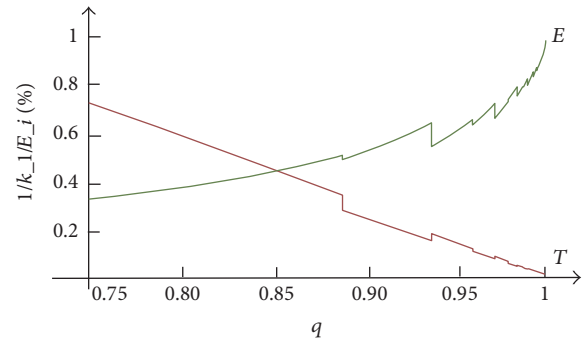

Figure 3: The changing trend about T and E with q.

so $\lim_{q \to 1} T_i = 0$ and $\lim_{q \to 1} E_i = 1/k_1$. The proof is completed. □

Let $E = (1/k_1)/E_i$. The changing trend between T and E with the variable q is shown in Figure 3.

3.5. Binary Classification of Multigranulation Searching Algorithm. In this paper, a kind of efficient binary classification of multigranulation searching algorithm is proposed through discussing the best testing strategy of the blood analysis case. The algorithm is illuminated as follows.

Algorithm 21. Binary classification of multigranulation searching algorithm (BCMSA).

Input. A probability quotient space $Q = (U, 2^U, p)$.

Output. The average classification times expectation of each object E.

Step 1. k_1 will be obtained based on Lemma 6. $i = 1$, $j = 0$ and searching_numbers = 0 will be initialized.

Step 2. Randomly dividing U_{ij} into s_i subgroups $U_{i1}, U_{i2}, \ldots, U_{is_i}$. ($s_i = N_{ij}/k_i$, where N_{ij} stands for the number of objects in U_{ij}, $U_{10} = U_1$).

Step 3. For i to $\lfloor \log_2^{N_{ij}} \rfloor$. ($\lfloor \log_2^{N_{ij}} \rfloor$ stands for $\log_2^{N_{ij}}$ and will round down to the nearest integer which is less than it).

For j to s_i.

If Test(U_{ij}) > 0 and $N_{ij} > 4$. Then searching_numbers + 1, $U_{i+1} = U_{ij}$, $i + 1$, go to Step 2 (Test function is a searching method).

If Test(U_{ij}) = 0. Then searching_numbers + 1, $i + 1$.

If $N_{ij} \le 4$. Go to Step 4.

Step 4. searching_numbers+$\sum U_{ij}$, E = (searching_numbers+ UN)/N.

Step 5. Return E.

The algorithm flowchart is shown in Figure 4.

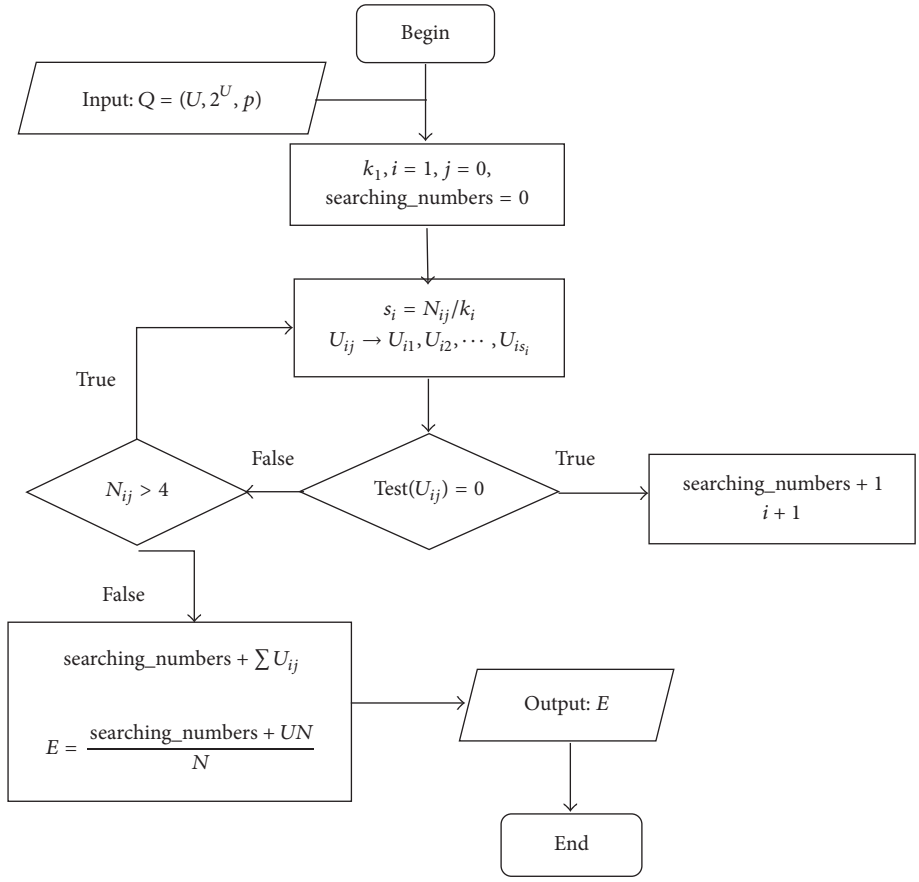

Figure 4: Flowchart of BCMSA.

Complexity Analysis of Algorithm 21. In this algorithm, the best case is p where the prevalence rate tends to be 0, and the classification time is $N * E_i \approx N/k_1 \approx N * (p + (p^2/2))$, which tends to 1, so the time complexity of computing is $O(1)$. But the worst case is p which tends to be 0.3, and the classification times tend to N, so the time complexity of computing is $O(N)$.

4. Comparative Analysis on Experimental Results

In order to verify the efficiency of the proposed BCMSA, in this paper, suppose there are two large domains $N_1 = 1 \times 10^4$, $N_2 = 100 \times 10^4$ and five kinds of different prevalence rates which are $p_1 = 0.01$, $p_2 = 0.001$, $p_3 = 0.0001$, $p_4 = 0.00001$, and $p_5 = 0.000001$. In the experiment of blood analysis case, the number "0" stands for a sick sample (negative sample) and "1" stands for a healthy sample (positive sample), then randomly generating N numbers, in which the probability of generating "0" denoted as p and the probability of generating "1" denoted as $1 - p$ stand for all the domain objects. The binary classifier is that, summing all numbers in a group (subgroup), if the sum is more than 1, it means this group is tested to be abnormal, and if the sum equals 0, it means this group has been tested to be normal.

Experimental environment is 4 G RAM, 2.5 GHz CPU, and WIN 8 system, the program language is Python, and the experimental results are shown in Table 8.

In Table 8, item "p" stands for the prevalence rate, item "levels" stands for the granulation levels of different methods, and item "$E(X)$" stands for the average expectation of classification times for each object. Item "k_1" stands for the objects number of each group in the 1st layer. Item "ℓ" stands for the degree of $E(X)$ close to $1/k_1$. $N_1 = 1 \times 10^4$ and $N_2 = 1 \times 10^6$, respectively, stand for the objects number of two original domains. Items "Method 9" and "Method 10," respectively, stand for the improved efficiency where Method 11 compares with Method 9 and Method 10.

Form Table 8, diagnosing all objects needs to expend 10000 classification times by Method 9 (traditional method), 201 classification times in Method 10 (single-level grouping method), and only 113 classification times in Method 11 (multilevels grouping method) for confirming the testing results of all objects when $N_1 = 1 \times 10^4$ and $p = 0.0001$. Obviously, the proposed algorithm is more efficient than Method 9 and Method 10, and the classification times can, respectively, be reduced to 98.89% and 47.33%. At the same time, when the probability p is gradually reduced, BCMSA has gradually become more efficient than Method 10, and ℓ tends to 100%; that is to say, the average classification time for each object tends to $1/k_1$ in the BCMSA. In addition,

Table 8: Comparative result of efficiency among 3 kinds of methods.

p	Levels	$E(X)$	k_1	ℓ	N_1	N_2	Method 9	Method 10
0.01	Single-level	0.19557083665	11	46.48%	1944	195817	81.44%	—
	2 levels	0.15764974327		57.67%	1627	164235	83.58%	16.13%
0.001	Single-level	0.06275892424	32	49.79%	633	62674	94.72%	—
	4 levels	0.03461032833		90.29%	413	41184	96.82%	34.62%
0.0001	Single-level	0.01995065634	101	49.62%	201	19799	98.00%	—
	6 levels	0.01050815802		94.22%	113	11212	98.89%	47.33%
0.00001	Single-level	0.00631957079	318	49.76%	633	6325	99.37%	—
	7 levels	0.00330165587		95.26%	333	3324	99.67%	47.75%
0.000001	Single-level	0.00199950067	1001	49.96%	15	2001	99.80%	—
	9 levels	0.00104104416		95.96%	15	1022	99.89%	47.94%

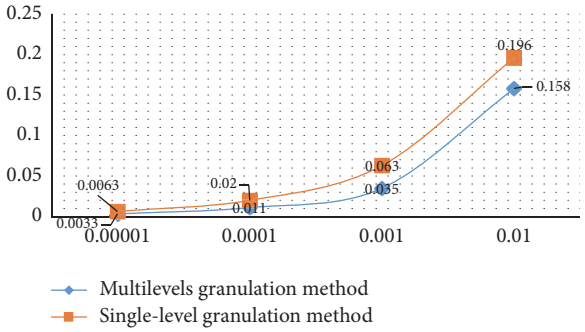

Figure 5: Comparative analysis between 2 kinds of methods.

the BCMSA can save 0%~50% of the classification times compared with Method 10. The efficiency of Method 10 (single-level granulation method) and Method 11 (multilevels granulation method) is shown in Figure 5; the x-axis stands for prevalence rate (or the negative sample rate) and the y-axis stands for the average expectation of classification times for each object.

In this paper, BCMSA is proposed, and it can greatly improve searching efficiency when dealing with complex searching problems. If there is a binary classifier, which is not only valid to an object, but also valid to a group with many objects, the efficiency of searching all objects will be enhanced by BCMSA, such as blood analysis case. As the same time, it may play an important role for promoting the development of granular computing. Of course, this algorithm also has some limitations. For example, if the prevalence rate of a sickness (or the occurrence rate of event A) $p > 0.3$, it will have no advantage compared with traditional method. In other words, the original problem need not be subdivided into many subproblems when $p > 0.3$. And when the prevalence rate of a sickness (or the negative sample rate in domain) is unknown, this algorithm needs to be further improved so that it can adapt to the new environment.

5. Conclusions

With the development of intelligence computation, multigranulation computing has gradually become an important tool to process the complex problems. Specially, in the process of knowledge cognition, granulating a huge problem into lots of small subproblems means to simplify the original complex problem and deal with these subproblems in different granularity spaces [64]. This hierarchical computing model is very effective for getting a complete solution or approximate solution of the original problem due to its idea of divide and conquer. Recently, many scholars pay their attention to efficient searching algorithms based on granular computing theory. For example, a kind of algorithm for dealing with complex network on the basis of quotient space model is proposed by L. Zhang and B. Zhang [65]. In this paper, combining hierarchical multigranulation computing model and principle of probability statistics, a new efficient binary classifier of multigranulation searching algorithm is established on the basis of mathematical expectation of probability statistics, and this searching algorithm is proposed according to recursive method in multigranulation spaces. Many experimental results have shown that the proposed method is effective and can save lots of classification times. These results may promote the development of intelligent computation and speed up the application of multigranulation computing. However, this method also causes some shortcomings. For example, on the one hand, this method has strict limitation for the probability value of p: namely, $p < 0.3$. On the contrary, if $p > 0.3$, the proposed searching algorithm probably is not the most effective method, and the improved methods need to be found. On the other hand, it needs a binary classifier, which is not only valid to an object, but also valid to a group with many objects. In the end, with the decrease of probability value of p (even it infinitely closes to zero), for every object, its mathematical expectation of searching time will gradually close to $1/k_1$. In our future research, we will focus on the issue on how to granulate the huge granule space without any probability value of each object and try our best to establish a kind of effective searching algorithm under which we do not know the probability of negative samples in domain. We hope these researches can promote the development of artificial intelligence.

Competing Interests

The authors declared that they have no conflict of interests related to this work.

Acknowledgments

This work is supported by the National Natural Science Foundation of China (no. 61472056) and the Natural Science Foundation of Chongqing of China (no. CSTC2013jjb 40003).

References

[1] A. Gacek, "Signal processing and time series description: a perspective of computational intelligence and granular computing," *Applied Soft Computing Journal*, vol. 27, pp. 590–601, 2015.

[2] O. Hryniewicz and K. Kaczmarek, "Bayesian analysis of time series using granular computing approach," *Applied Soft Computing*, vol. 47, pp. 644–652, 2016.

[3] C. Liu, "Covering multi-granulation rough sets based on maximal descriptors," *Information Technology Journal*, vol. 13, no. 7, pp. 1396–1400, 2014.

[4] Z. Y. Li, "Covering-based multi-granulation decision-theoretic rough sets model," *Journal of Lanzhou University*, no. 2, pp. 245–250, 2014.

[5] Y. Y. Yao and Y. She, "Rough set models in multigranulation spaces," *Information Sciences*, vol. 327, pp. 40–56, 2016.

[6] J. Xu, Y. Zhang, D. Zhou et al., "Uncertain multi-granulation time series modeling based on granular computing and the clustering practice," *Journal of Nanjing University*, vol. 50, no. 1, pp. 86–94, 2014.

[7] Y. T. Guo, "Variable precision β multi-granulation rough sets based on limited tolerance relation," *Journal of Minnan Normal University*, no. 1, pp. 1–11, 2015.

[8] X. U. Yi, J. H. Yang, and J. I. Xia, "Neighborhood multigranulation rough set model based on double granulate criterion," *Control and Decision*, vol. 30, no. 8, pp. 1469–1478, 2015.

[9] L. A. Zadeh, "Towards a theory of fuzzy information granulation and its centrality in human reasoning and fuzzy logic," *Fuzzy Sets and Systems*, vol. 19, pp. 111–127, 1997.

[10] J. R. Hobbs, "Granularity," in *Proceedings of the 9th International Joint Conference on Artificial Intelligence*, Los Angeles, Calif, USA, 1985.

[11] L. Zhang and B. Zhang, "Theory of fuzzy quotient space (methods of fuzzy granular computing)," *Journal of Software*, vol. 14, no. 4, pp. 770–776, 2003.

[12] J. Li, C. Mei, W. Xu, and Y. Qian, "Concept learning via granular computing: a cognitive viewpoint," *Information Sciences*, vol. 298, no. 1, pp. 447–467, 2015.

[13] X. Hu, W. Pedrycz, and X. Wang, "Comparative analysis of logic operators: a perspective of statistical testing and granular computing," *International Journal of Approximate Reasoning*, vol. 66, pp. 73–90, 2015.

[14] M. G. C. A. Cimino, B. Lazzerini, F. Marcelloni, and W. Pedrycz, "Genetic interval neural networks for granular data regression," *Information Sciences*, vol. 257, pp. 313–330, 2014.

[15] P. Hońko, "Upgrading a granular computing based data mining framework to a relational case," *International Journal of Intelligent Systems*, vol. 29, no. 5, pp. 407–438, 2014.

[16] M.-Y. Chen and B.-T. Chen, "A hybrid fuzzy time series model based on granular computing for stock price forecasting," *Information Sciences*, vol. 294, pp. 227–241, 2015.

[17] R. Al-Hmouz, W. Pedrycz, and A. Balamash, "Description and prediction of time series: a general framework of Granular Computing," *Expert Systems with Applications*, vol. 42, no. 10, pp. 4830–4839, 2015.

[18] M. Hilbert, "Big data for development: a review of promises and challenges," *Social Science Electronic Publishing*, vol. 34, no. 1, pp. 135–174, 2016.

[19] T. J. Sejnowski, S. P. Churchland, and J. A. Movshon, "Putting big data to good use in neuroscience," *Nature Neuroscience*, vol. 17, no. 11, pp. 1440–1441, 2014.

[20] G. George, M. R. Haas, and A. Pentland, "BIG DATA and management," *Academy of Management Journal*, vol. 30, no. 2, pp. 39–52, 2014.

[21] M. Chen, S. Mao, and Y. Liu, "Big data: a survey," *Mobile Networks and Applications*, vol. 19, no. 2, pp. 171–209, 2014.

[22] X. Wu, X. Zhu, G. Q. Wu, and W. Ding, "Data mining with big data," *IEEE Transactions on Knowledge & Data Engineering*, vol. 26, no. 1, pp. 97–107, 2014.

[23] Y. Shuo and Y. Lin, "Decomposition of decision systems based on granular computing," in *Proceedings of the IEEE International Conference on Granular Computing (GrC '11)*, pp. 590–595, Garden Villa Kaohsiung, Taiwan, 2011.

[24] H. Hu and Z. Zhong, "Perception learning as granular computing," *Natural Computation*, vol. 3, pp. 272–276, 2008.

[25] Z.-H. Chen, Y. Zhang, and G. Xie, "Mining algorithm for concise decision rules based on granular computing," *Control and Decision*, vol. 30, no. 1, pp. 143–148, 2015.

[26] K. Kambatla, G. Kollias, V. Kumar, and A. Grama, "Trends in big data analytics," *Journal of Parallel & Distributed Computing*, vol. 74, no. 7, pp. 2561–2573, 2014.

[27] A. Katal, M. Wazid, and R. H. Goudar, "Big data: issues, challenges, tools and good practices," in *Proceedings of the 6th International Conference on Contemporary Computing (IC3 '13)*, pp. 404–409, IEEE, New Delhi, India, August 2013.

[28] V. Cevher, S. Becker, and M. Schmidt, "Convex optimization for big data: scalable, randomized, and parallel algorithms for big data analytics," *IEEE Signal Processing Magazine*, vol. 31, no. 5, pp. 32–43, 2014.

[29] J. Fan, F. Han, and H. Liu, "Challenges of big data analysis," *National Science Review*, vol. 1, no. 2, pp. 293–314, 2014.

[30] Q. H. Zhang, K. Xu, and G. Y. Wang, "Fuzzy equivalence relation and its multigranulation spaces," *Information Sciences*, vol. 346-347, pp. 44–57, 2016.

[31] Z. Liu and Y. Hu, "Multi-granularity pattern ant colony optimization algorithm and its application in path planning," *Journal of Central South University (Science and Technology)*, vol. 9, pp. 3713–3722, 2013.

[32] Q. H. Zhang, G. Y. Wang, and X. Q. Liu, "Hierarchical structure analysis of fuzzy quotient space," *Pattern Recognition and Artificial Intelligence*, vol. 21, no. 5, pp. 627–634, 2008.

[33] Z. C. Shi, Y. X. Xia, and J. Z. Zhou, "Discrete algorithm based on granular computing and its application," *Computer Science*, vol. 40, pp. 133–135, 2013.

[34] Y. P. Zhang, B. luo, Y. Y. Yao, D. Q. Miao, L. Zhang, and B. Zhang, *Quotient Space and Granular Computing The Theory and Method of Problem Solving on Structured Problems*, Science Press, Beijing, China, 2010.

[35] G. Y. Wang, Q. H. Zhang, and J. Hu, "A survey on the granular computing," *Transactions on Intelligent Systems*, vol. 6, no. 2, pp. 8–26, 2007.

[36] J. Jonnagaddala, R. T. Jue, and H. J. Dai, "Binary classification of Twitter posts for adverse drug reactions," in *Proceedings of the Social Media Mining Shared Task Workshop at the Pacific Symposium on Biocomputing*, pp. 4–8, Big Island, Hawaii, USA, 2016.

[37] M. Haungs, P. Sallee, and M. Farrens, "Branch transition rate: a new metric for improved branch classification analysis," in *Proceedings of the International Symposium on High-Performance Computer Architecture (HPCA '00)*, pp. 241–250, 2000.

[38] R. W. Proctor and Y. S. Cho, "Polarity correspondence: a general principle for performance of speeded binary classification tasks," *Psychological Bulletin*, vol. 132, no. 3, pp. 416–442, 2006.

[39] T. H. Chow, P. Berkhin, E. Eneva et al., "Evaluating performance of binary classification systems," US, US 8554622 B2, 2013.

[40] D. G. Li, D. Q. Miao, D. X. Zhang, and H. Y. Zhang, "An overview of granular computing," *Computer Science*, vol. 9, pp. 1–12, 2005.

[41] X. Gang and L. Jing, "A review of the present studying state and prospect of granular computing," *Journal of Software*, vol. 3, pp. 5–10, 2011.

[42] L. X. Zhong, "The predication about optimal blood analyze method," *Academic Forum of Nandu*, vol. 6, pp. 70–71, 1996.

[43] X. Mingmin and S. Junli, "The mathematical proof of method of group blood test and a new formula in quest of optimum number in group," *Journal of Sichuan Institute of Building Materials*, vol. 01, pp. 97–104, 1986.

[44] B. Zhang and L. Zhang, "Discussion on future development of granular computing," *Journal of Chongqing University of Posts and Telecommunications: Natural Science Edition*, vol. 22, no. 5, pp. 538–540, 2010.

[45] A. Skowron, J. Stepaniuk, and R. Swiniarski, "Modeling rough granular computing based on approximation spaces," *Information Sciences*, vol. 184, no. 1, pp. 20–43, 2012.

[46] J. T. Yao, A. V. Vasilakos, and W. Pedrycz, "Granular computing: perspectives and challenges," *IEEE Transactions on Cybernetics*, vol. 43, no. 6, pp. 1977–1989, 2013.

[47] Y. Y. Yao, N. Zhang, D. Q. Miao, and F. F. Xu, "Set-theoretic approaches to granular computing," *Fundamenta Informaticae*, vol. 115, no. 2-3, pp. 247–264, 2012.

[48] H. Li and X. P. Ma, "Research on four-element model of granular computing," *Computer Engineering and Applications*, vol. 49, no. 4, pp. 9–13, 2013.

[49] J. Hu and C. Guan, "Granular computing model based on quantum computing theory," in *Proceedings of the 10th International Conference on Computational Intelligence and Security*, pp. 156–160, November 2014.

[50] Y. Shuo and Y. Lin, "Decomposition of decision systems based on granular computing," in *Proceedings of the IEEE International Conference on Granular Computing (GrC '11)*, pp. 590–595, IEEE, Kaohsiung, Taiwan, November 2011.

[51] F. Li, J. Xie, and K. Xie, "Granular computing theory in the applicatiorlpf fault diagnosis," in *Proceedings of the Chinese Control and Decision Conference (CCDC '08)*, pp. 595–597, July 2008.

[52] Q.-H. Zhang, Y.-K. Xing, and Y.-L. Zhou, "The incremental knowledge acquisition algorithm based on granular computing," *Journal of Electronics and Information Technology*, vol. 33, no. 2, pp. 435–441, 2011.

[53] Y. Zeng, Y. Y. Yao, and N. Zhong, "The knowledge search base on the granular structure," *Computer Science*, vol. 35, no. 3, pp. 194–196, 2008.

[54] G.-Y. Wang, Q.-H. Zhang, X.-A. Ma, and Q.-S. Yang, "Granular computing models for knowledge uncertainty," *Journal of Software*, vol. 22, no. 4, pp. 676–694, 2011.

[55] J. Li, Y. Ren, C. Mei, Y. Qian, and X. Yang, "A comparative study of multigranulation rough sets and concept lattices via rule acquisition," *Knowledge-Based Systems*, vol. 91, pp. 152–164, 2016.

[56] H.-L. Yang and Z.-L. Guo, "Multigranulation decision-theoretic rough sets in incomplete information systems," *International Journal of Machine Learning & Cybernetics*, vol. 6, no. 6, pp. 1005–1018, 2015.

[57] M. A. Waller and S. E. Fawcett, "Data science, predictive analytics, and big data: a revolution that will transform supply chain design and management," *Journal of Business Logistics*, vol. 34, no. 2, pp. 77–84, 2013.

[58] R. Kitchin, "The real-time city? Big data and smart urbanism," *GeoJournal*, vol. 79, no. 1, pp. 1–14, 2014.

[59] X. Dong and D. Srivastava, *Big Data Integration*, Morgan & Claypool, 2015.

[60] L. Zhang and B. Zhang, *Theory and Applications of Problem Solving*, Quotient Space Based Granular Computing (The Second Version), Tsinghua University Press, Beijing, China, 2007.

[61] L. Zhang and B. Zhang, "The quotient space theory of problem solving," in *Rough Sets, Fuzzy Sets, Data Mining, and Granular Computing*, G. Wang, Q. Liu, Y. Yao, and A. Skowron, Eds., vol. 2639 of *Lecture Notes in Computer Science*, pp. 11–15, Springer, Berlin, Germany, 2003.

[62] J. Sheng, S. Q. Xie, and C. Y. Pan, *Probability Theory and Mathematical Statistics*, Higher Education Press, Beijing, China, 4th edition, 2008.

[63] L. Z. Zhang, X. Zhao, and Y. Ma, "The simple math demonstration and rrecise calculation method of the blood grouptest," *Mathematics in Practice and Theory*, vol. 22, pp. 143–146, 2010.

[64] J. Chen, S. Zhao, and Y. Zhang, "Hierarchical covering algorithm," *Tsinghua Science & Technology*, vol. 19, no. 1, pp. 76–81, 2014.

[65] L. Zhang and B. Zhang, "Dynamic quotient space model and its basic properties," *Pattern Recognition and Artificial Intelligence*, vol. 25, no. 2, pp. 181–185, 2012.