Applied Mathematics & Information Sciences
An International Journal

http://dx.doi.org/10.18576/amis/130619

A Fuzzy Responsibility-Based Access Organizer for Leukemia Record Protection using KWatts Algorithm

B. Shuriya[1,*] *and A. Rajendran*[2]

[1]Department of Computer science and Engineering, Adithya Institute Technology, Coimbatore, Tamilnadu, India-641107
[2]Department of Electronics and Communication Engineering, Karpagam college of Engineering, Coimbatore, India-641032

Received: 2 ,ay 2019, Revised: 12 Oct. 2019, Accepted: 28 Oct. 2019
Published online: 1 Nov. 2019

Abstract: The importance of classifying cancer patients into high to low-risk groups has led many research teams to study the application of machine learning (ML) methods. Here, the leukemia dataset is used for classification to diagnosis the disease. The proposed K-Watts classifying algorithm, compared to other classification algorithms, yields the lowest error rate. The purpose of the reproduction for record protection is to represent a clear idea about the new Fuzzy Responsibility- Based Access Organizer (FRBAO). The projected reproduction provides the wider society safe plans, or with access to the organized mode in MultiPlan hypothesis, using fuzzy constraints. The view part of RBAO can be modified with inactive and active authorization duty. The records in the database can be allowed to access depends on the constraints. The existing and new records are making to appeal with the updated record. The proposed reproduction deal with the fuzzy approach along with the datasets. The method is based on fuzzy linguistic variables and Communicating Sequential Processes (CSP) with multi-objective fuzzy supervisory production to identify and to aid in the early detection of cancer cells.

Keywords: Communicating Sequential Processes, Responsibility Based Access Organize, K-medoid, SoD.

1 Introduction

Cancer has been characterized as a heterogeneous disease consisting of different subtypes. The early diagnosis and prognosis of cancer have become a necessity in cancer research, as it can facilitate the subsequent clinical management of the patients records. The exact finding of four heterogeneous childhood cancers, namely, neuroblastoma, non-Hodgkin lymphoma and Ewing sarcoma is significant in light of the fact that they show a comparative analysis of minuscule round blue cell tumors (SRBCTs) and often times leads to misdiagnosis. Be that as it may, because of a little number of tests contrasted with the substantial number of genes in microarray gene expression information, it is difficult to recognize a little subset of a pertinent chromosome that can distinguish between these four subgroups of childhood cancers with mathematical exactness. Responsibility-based Access Organizer (RBAO) made a giant global leap towards access to organizers, with diverse applications pertinent to the safety of submission and protection of records. However, the rising dimension of the problem the domain is represented by information systems that cross organizational borders, safe administration and involving its guarantee masquerade discouraging challenge. RBAO method gives safe administration by implementing the safe plans.

1.1 Literature survey

In paper [1], the author resolves the impression of syndrome clusters in breast cancer scraps using improved K-medoid clustering and he also developed the improved K-medoid clustering which helps to upgrade the clustering performance by reassigning some of the negative average silhouette width (ASW) syndrome to other clusters after initial K-medoid clustering. Consequently, in [2], a reckoning proficient but precise gene ID strategy has been nominated. The t-test technique is antiquated to diminish the measurement of the dataset and after that, the recommended particle swarm optimization - based approach has been utilized to discover helpful genetic code. In [3], a novel hybrid

* Corresponding author e-mail: shuriyasmile@gmail.com

intelligent system on the basis of Association Rule Mining (ARM) and Neural Networks

which utilizes an Evolutionary Algorithm (EA) is introduced to cover the amplitude issue for the identifying of breast cancer. In paper [4], authors have designed a new hybrid technique of Grey Wolf Optimizer (GWO) consolidating with decision tree as a classifier for choosing a minimum number of useful genes from the lots of genes to recognize cancer is designed. In paper [5], the authors proposed a cancer site classification framework by investigating somatic mutations through machine learning approaches. In [6], a mixture of Kernelized fuzzy rough set (KFRS) and semi-supervised support vector machine (S3VM) is affirmed for anticipating growth biomarkers from one miRNA and three gene expression information sets. Biomarkers are utilizing three-feature selection techniques, including KFRS. The adequacy of the recommended KFRS and SVM mixture on the microarray information sets is illustrated, and the cancer biomarkers recognized from miRNA information are accounted for. Besides, biological significance tests are directed miRNA cancer biomarkers. Hence, in [7], authors have used a novel hybrid system on the basis of association rule mining. They also discussed the aspects of data features and aspects of predictive modeling for regression and classification used in data mining. In paper [8], the author implemented the neural network design using backpropagation and radial basis function in simulation studies. In paper [9], authors implemented a neural network by single neuron. The authors have also discussed the complexity of the statistical classifiers. The technique considers the sample size ‘ n’neural network and generates the minimum error rate. In paper [10], the author has discussed a hybrid multi-objective binary biogeography-based optimization with support vector machine introduced for chromosome selection on ten nucleic acid expression datasets. The recommended algorithm can obtain apical accuracy in nine out of ten microarray dataset issues since the multi-objective approach in it can find a distinct result in Pareto optimal set. Paper [11] investigated the dissociation of the model in which the valve of each chromosome is identified separately in free search instances. Authors used an instinctive computation-based search algorithm, called differential evolution, for interpreting the controller of each gene. After analyzing with two artificial GRNs and analyzing a real gene expression profile, they checked the contingency of the developed approach. In paper [12], the author contributed low cost autonomous perceptron neural network. The quantum is computed in neural network. In papers [13,14,15], the evolutionary algorithms are implemented such as SVM, Particle Swarm, Genetic algorithm and other optimization algorithms. In paper [17], the heuristic optimization algorithm is implemented over continuous spaces.

2 Leukemia classifications

Cancers are classified by the type of tissue in which the cancer originates (histological type) and by primary site, or the location in the body where the cancer first developed. In this paper, the classification of cancer cell based on the histological type is implemented.The international standard for the cancer classification and nomenclature of histologies is the International Classification of Diseases for Oncology, Third Edition (ICD-O-3). From a histological standpoint, there are hundreds of different types of cancers, which are grouped into six major categories: Carcinoma, Sarcoma, Myeloma, Leukemia, Lymphoma. Leukemias are cancers of the bone marrow (the site of blood cell production). The word leukemia means ”white blood” in Greek. The disease is often associated with the overproduction of immature white blood cells. These immature white blood cells do not perform as well as they should, therefore the patient is often prone to infection. Leukemia also affects red blood cells and can cause poor blood clotting and fatigue due to anemia for the patients.

A. Dataset description

To compare the data mining classification techniques and comparison analysis, we need the datasets [18]. This research chooses Leukemia data sets. Directly, we can apply these data in the data mining tools (Weka) and predict the results. The chosen dataset ”Testing data” on year 2010 contains 72 leukemia samples (47 ALL and 25 AML).

B. KWatts algorithm

Training procedure for the given labeled training samples with class labels, compute the per-class centroids where the set of indices of samples belongs to class.

Prediction function: the class assigned to an observation is made by its prototype.

The number of data points is D, the mean degree K is assumed as even integer. This model constructs the undirected graph with DK/2 edges as ring lattice and forms the clustering by greedy search by finding optimum solution.

1. Construct the regular ring lattice with D nodes each connected to K neighbours,K/2 on each side i.e., if the nodes are labeled $i_0, i_1 ... i_{n-1}$, there is an edge(i_n, i_m) if and only if $0 < mod((n-m)D-1-(K/2))$.

2. Foe every node $i_0, i_1 ... i_{n-1}$ take every edge i_i to its K/2 on right most side i.e., every edge(i_n, i_m)mod D with $i_n < i_m \leq K/2$

3. Initialize: select k of the i data points as the medoids in the ring lattice

4. Associate each data point to the closest medoid in lattice

5. While the cost of the configuration decreases:

(i). For each medoid m, for each non-medoid data point S:

(a). Swap m and S, recompute the cost (sum of distances of points to their medoid)

(b). If the total cost of the configuration increased in the previous step, undo the swap

6. Rewiring is done with replacing (i_n, i_m) mod D with (i_n, i_l) where l is chosen uniformly at random from nodes in the D by avoiding self loops and duplicates.

3 Modeling safe plans using a fuzzy approach

A. Unclear and Inaccurate Safe Plans

In the organization, the administration does not have proper safe plans with fuzzy security systems [14]. By the natural language, more interpretation happened. Each user and organisation need the terms of privacy, truth and accessibility with their own data. By the increasing of the responsive rank of security with the partitioned rank enter the existing data (EED) and enter the new data (END). The fuzzy linguistic variables provide low, high and very high variables which assign with the value range between 0 and 1. Responsive rank results varied with the terms of mode of inactive and active rules. The inactive mode places with the column side and the active mode values can be unpredictable.

In this example, responsive rank schedule patient ID and the name are inactive mode. The disease and the patient ID are highly responsive data of a patient and accessing those should be controlled. Therefore, the disease and patient ID columns are assigned with Very High responsive values. However, the highly responsive disease column depends on the type of the disease. For example, sensitivity level of a record of an Aids patient is more responsive than that of a Cancer patient. Enter the existing data (EED) and Enter the new data (END) values to make the resolutions as shown in Figure 1.

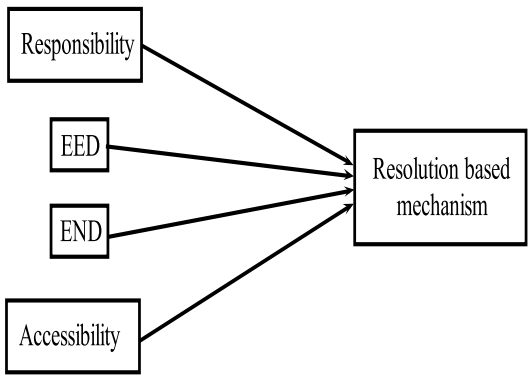

Fig. 1: Safe plans

In most of the information systems, commands of records include different entries of existing or new records into database resources. For example in a hospital application, a doctor may want to exist or new to the disease field of a patient record than a nurse. The proposed model assigns an EED and an END value for each data field in the database. Since these EED and END values are determined based on the human decision, those are defined using fuzzy linguistic variables.

B. FRBAO Mechanism

Authorization duty mechanism is indicated with the incorporate a fuzzy ending. FRBAO representation divided as Inactive duty assignment (IDA) and Active duty assignment (ADA) both contain safe plan using a fuzzy move towards the record in the data base. In the FRBAO model, both IDA and ADA are mediated by a component called Fuzzy Plan Surveyor (FPS). The FPS includes two mechanisms; the plan mechanism and the resolution mechanism. RBAO director requests for an authorization duty, by triggering IDA. The plan module of the FPS takes the role that is responsible for performing the operation and the operation to be assigned as the inputs and calculates the maximum, the average and the minimum values of sensitivity, EED and END values of the record resources affected by that authorization duty. These values are calculated using the relationship between the requested operation and the affected register. During the calculation linguistic variables are converted to their corresponding decimal values, e.g., Low = 0.4 and Very High = 0.9. Then those calculated values are passed to the resolution factor. Then, the resolution factor calculates the Access value and a No Access value which decides whether to grant the permission or not to the user. It uses some straight forward fuzzy multi-objective conclusion production (FMOC) to make the decision.

C. Evaluation Resolution

The resolution factor uses the average responsibility, EED and END as objectives O= responsibility, EED, END and Access and No Access values change in the active mode with the values of FMOC [4].

A= Access, No Access For example, assume that after the RBAO director requests for a permission assignment, the following values are calculated by the plan factor:

Average Responsibility = 0.42

Average EED = 0.65

Average END =0.53

Then the resolution factor assigns the Access and No Access values for the objectives as follows:

Responsibility = 0.56, 0.42

EED = 0.65, 0.42

END= 0.53, 0.45

Here the Access value for responsibility taken as 1-Responsibility, because resources are 0.42 responsive which means that they cannot be accessed in 0.42 times and can be accessed in 0.56 times.

By using FMOD, the decision is calculated as follows:

Conclusion = max [min(Access), min (No Access)]

Here, according to the FMOC, the decision is Access. Then by considering the maximum, the minimum and the average values of the responsibility, the EED and the END

and the conclusion, RBAO director can choose to give the authorization or not.

Min (EED)=0.3, Max(EED)=0.8, Avg(EED)=0.66

Min (END)=0.4, Max(END)=0.8, Avg(END)=0.6

Min (responsibility) =0.09, Max (responsibility) =0.8, Avg (responsibility) =0.37 Then, it evaluates the access and Noaccess values using fuzzy conclusion.

Responsibility= 0.69, 0.39

EED = 0.68, 0.34

END= 0.71, 0.31

The Access value= 0.62, No Access value=0.3

The minimum responsibility= 0.1, maximum responsibility = 0.7, responsibility = 0.38 and the Access value = 0.62 and the No Access value = 0.3, the RBAO director will conclude to give the authorization to use the operation new Medical Reocords () to the doctor or not. In the AAD, when the user selects an operation, the request is passed to the FPS. Assume that user 1 who is a doctor, wants to execute the operation view Medical Records () and he wants to access the record of a patient with the patient ID 007. The record 007 has a dynamic mode tracked. The FPS calculates access and No Access values by a process with AAD using the active responsibility of the disease field given in the active regulation. Those, records are as follows: Maximum responsibility = 0.9, Access Value = 0.58, No Access Value = 0.30. Since the maximum sensitivity value is 0.8, the access is denied for the user 1.

Separation of duty with communicating sequential processes (CSP)

Separation of Duty (SoD) [17] means to record without any error by allocating responsibilities connected with record users. Record users must prevent external threats. Threats often reside within allowed coordination users from the error. SoD which ensures record truthful.

A. User identification

Let the name of the process as N and n E N.The set of process records P is indicated.

$P ::= \sigma \to p|skip|stop|n|p.p|p \cap p|p||p; p$, where $\sigma^{\varepsilon} \sum$

Let the U be user members and T a set of tasks. For a task t E T and a user u E U, we call an event of the form t:u a (task) completing event and is denoted by $X = \{t.u|tET; uEU\}$ the set of all completing events. A completion event t.u which required completing the task with record user.

B. Composing workflows and access control

Authorized duty evaluates with process A. Workflow process s W which access rules determined with the process A by evaluating A with that properly $W||XA$. Role-based Access organizes (RBAO) have dynamic workflows and interior quality, the disintegration of the user authorization duty with user-role and a role-authorization duty assignment. RBAO construction tuple with the record of organisation which have user access and process access. User access belongs to the multiples of user and roles for the user duty. Process access belongs to the multiples of roles and tasks for the authorization duty. Let R be a set of roles and the tuples (UA, PA). For the safe Workflows (SW), $SW(UA, PA) = W||XRBAO(UA, PA)$

For example, maintaining the medical records with the format prescription and the specialist.

An RBAO process is parameterized by an RBAO configuration (UA, PA) and hold in every completing result t.u if u is authoritative to complete t with respect to (UA, PA). Also, RBAO method conversion in user- duty relative is processed by engaging in administrative events and may terminate at any time.

C. SoDA Syntax (explain)

For the workflow SoDA grammar G with the user group $U = \{u_1, u_2...u_m\}$ and roles group $R = \{r_1, r_2....r_n\}$ and constraints are (A,B,C,D)

$A = \{D, AT, PT, GT, US, UR, U, R\}$ with the nonterminal code .

$B = \{`, ', (,), \{, \}, \square, \}UUUR$ with the terminal code.

$D ::= AT|PT$

$AT ::= (ATA)|(ATA)|(AA)|(A.A)|(GT)+$

$PT ::= GT|(AT.AT)|(AT \cap AT)|AT$

$UR ::= U|U, UR$

$U ::= u1|...|um$

$R ::= r1|...|rn$

The SoD-safe workflow procedure indicates the symbol SSWf. The SSWf method engages in three kinds of events: execution events, executive measures, and event X. Therefore implementation events and explain why every task instance in with sequential execution system accepted by SSWf.

Consider the SoD-safe workflow procedure $SSW\phi(UA, PA) = (W||RBAC(UA, PA))||SOD\phi(UA)$, for a SoDA term an RBAC configuration (UA,PA), and a workflow process W that models a workflow w. Assume that (SSWf (UA,PA)) corresponds to an unfinished workflow instance of w.

SoD enforcement method acknowledged with workflow. Sometimes the workflow this contented with the sketch group of SoD enforcement methods without administration procedures with the record codes.

Result and Discussion

Leukemia data from high-density Affymetrix oligonucleotide arrays were previously analyzed in Golub et al. (1). There were 7,129 genes and 34 samples: 20 in class ALL (acute lymphocytic leukemia) and 14 in class AML (acute mylogenous leukemia). The results for the methods of Golub et al., nearest shrunken centroids and k-medoid are shown in Table 1.

The author [18] reports a test error rate of 4/34 for their procedure using 50 genes to obtain their results. We found that if the number of genes was reduced to less than 47, the test error increased. The results for the nearest shrunken centroid method and K- Watts shows that minimum cross-validation error occurs near$\triangle = 1.4$ but leaves about 1,000 genes. In practice, this minimum error solution might be of interest. Here, to obtain a more manageable set of genes, we chose instead $\triangle = 4.06$, the

Table 1: Comparison of leukemia classification methods

Method	10-Fold CV error	Test error	No. of genes
Golub et al. (18)	3/38	4/34	50
Nearest shrunken centroid	1/38	2/34	21
Kwatts	1/38	1.5/34	19

point at which the cross-validation error starts to rise quickly, yielding only 21 genes.

4 Conclusion

K-Watts classification method is proposed to classify the Leukemia disease for diagnosis in which the level of the disease is identified. The results are compared with the other two methods. The safety plans which belongs to the Boolean values is based on the resolution results. The fuzzy logic with the basic procedure for multiple running processes. The fuzzy linguistic models executes with the values of EED and END by the Active and Inactive mode. Communication Sequential process with the security mode by the method of Separation of Duty for the truth and access control based on the mechanism involved in it. By the novel scheme combination, the user can manage the record with the procedures. So that the highest sensitivity level of the diseases is measured and given priority to the highest level sensitivity record to be treated.

References

[1] Q. Ping, C. C. Yang, S. A. Marshall, N. E. Avis and E. H. Ip, Breast cancer symptom clusters derived from social media and research study data using improved *k*-medoid clustering. IEEE transactions on computational social systems, **3**, 63, (2016).

[2] S. Kar, K. D. Sharma and M. Maitra, in Proc. 2nd International Conference on Control, Instrumentation, Energy & Communication (CIEC), 130-134, (2016).

[3] A. Boutorh and A. Guessoum, in 2015 12th International Symposium on Programming and Systems (ISPS), 1-9, (2015).

[4] M. Vosooghifard and H. Ebrahimpour, in Proc. 5th International Conference on Computer and Knowledge Engineering (ICCKE). IEEE, 147-151, (2015).

[5] Y. Chen, J. Sun, L.-C. Huang, H. Xu and Z. Zhao, Classification of cancer primary sites using machine learning and somatic mutations. BioMed research international 2015, (2015).

[6] D. Chakraborty and U. Maulik, Identifying cancer biomarkers from microarray data using feature selection and semisupervised learning. IEEE journal of translational engineering in health and medicine, **2**, 1, (2014).

[7] J. David, Hand. Principles of data mining. Drug safety, **30**, 621, (2007).

[8] H. B. Demuth, M. H. Beale, O. De Jess and M. T. Hagan, Neural network design. Martin Hagan, (2014).

[9] Š. Raudys, Evolution and generalization of a single neurone:: II. Complexity of statistical classifiers and sample size considerations. Neural Networks, **11**, 297, (1998).

[10] X. Li and M. Yin, Multiobjective binary biogeography based optimization for feature selection using gene expression data. IEEE Transactions on NanoBioscience, **12**, 343, (2013).

[11] N. Noman, L. Palafox and H. Iba, in Natural computing and beyond. Springer, 93-103, (2013).

[12] M. Zidan, A.-H. Abdel-Aty, A. El-Sadek, E. Zanaty and M. Abdel-Aty, in AIP in Proc. Conference Proceedings. AIP Publishing, **1905**, 020005, (2017).

[13] M. A. Lones, in Proc. Companion Publication of the 2014 Annual Conference on Genetic and Evolutionary Computation. ACM, 1419-1422, (2014).

[14] M. Hauschild and M. Pelikan, An introduction and survey of estimation of distribution algorithms. Swarm and evolutionary computation, **1**, 111, (2011).

[15] S. Das and P. N. Suganthan, Differential evolution: A survey of the state-of-the-art. IEEE transactions on evolutionary computation, **15**, 4, (2010).

[16] D. F. Ferraiolo, R. Sandhu, S. Gavrila, D. R. Kuhn and R. Chandramouli, Proposed NIST standard for role-based access control. ACM Transactions on Information and System Security (TISSEC), **4**, 224, (2001).

[17] R. Storn and K. Price, Differential evolution–a simple and efficient heuristic for global optimization over continuous spaces. Journal of global optimization, **11**, 341, (1997).

[18] T. R. Golub, D. K. Slonim, P. Tamayo, C. Huard, M. Gaasenbeek, J. P. Mesirov, H. Coller, M. L. Loh, J. R. Downing and M. A. Caligiuri, Molecular classification of cancer: class discovery and class prediction by gene expression monitoring. Science, **286**, 531, (1999).

B. Shuriya has seven years of professional experience in teaching, student management, and soft computing. She holds an ME in Computer Science and Engineering. She is currently working with Engineering College as an Assistant Professor. She has five publications including in national, international journals and conferences.

A. Rajendran is a Professor with over 23 years experience in both academic and administrative positions in engineering institutions. He holds an MS by research and Ph.D. in Information and Communication Engineering. He has presented, published 26 papers in national and international journals and conferences.