Swati Ahlawat al. International Journal of Recent Research Aspects ISSN: 2349-7688, Vol. 4, Issue 4, Dec 2017, pp. 88-92

Fuzzy Scheduling in cloud computing for Improving QoS level

Swati Ahlawat

M. Tech Scholar, Dept of Computer Science, NorthCap University Gurugram

(Formerly ITM University)

Abstract: **The refined Bayesian scheduling algorithm is the combination of hybrid clustering and the parallel job scheduling algorithm that divides the jobs into many queue forms and address solving resources matching that have better concurrent granularity by calculating percentage of the jobs using the hybrid clustering optimization algorithm. When the futile virtual machine nodes are found, job servers' duty is to recognize its vacant state with respect to the data calculated and make one choice from the queue of largest job for diversifying by making use of Bayesian scheduling algorithm. If the job is considered to be good with respect to its probability density, the execution start scheduling the jobs; if the job is called the waiting job, the probability of the coming job is calculated until the correct job is recognized. During some execution one may find some queue for which no job is scheduled. Then our task is to switch to another largest queue to search for appropriate schedulable task.**

***Keywords*— cloud computing, Fuzzy Scheduling, QoS level.**

I. INTRODUCTION

Scheduling is the allocation of resources to separate tasks to ensure completion for all tasks. Currently, there are many established algorithms for process scheduling. However, there is always a demand for greater efficiency for scheduling to maximize processing resources. Some criteria for determining the best algorithm include turnaround time and waiting time, amongst others. Not only is there a demand for meeting such criteria, it is also helpful to offload decision-making to the scheduler, rather than creating hard rules that don't always give the results that are desirable.

Types of Scheduling: -

- **First come first serve: -** FCFS algorithm for concurrent parallel processing aims at the task scheduling with the minimum waiting time of the queue. The CloudSim toolkit is used for the internal scheduling of the jobs by making use of First Come First Serve (FCFS) scheduling strategy. The responsibility of the virtual machine provisioned component is to allocate the application based virtual machines to Hosts in cloud based data center.. First-Come-First-Serve (FCFS) is the default policy implemented by the virtual machine provisioned to allocate the hardware to the specific task. The non preemptive property is the main disadvantage of FCFS. The shortest tasks have to wait for a long time if the any large task is in the front of it in the queue which decreases the turnaround and the response time.
- **Round Robin Scheduling: -** Round Robin (RR) algorithm is better than FCFS. Using RR algorithm in place of FCFS increase the response time and throughput. To store the jobs RR use the ring fashion with the token assign to them. Each and every task in the queue has equal time of execution and they all will execute in sequence. If in any case the job is not able to complete in its turn, and for its next turn, the job will be stored in the waiting queue. The main merit of Round Robin is every job will run in its turn and if it doesn't finish in time then other job does not have to wait for the predecessor job to get completed. But if the average weight seems to be extra heavy, Round Robin algorithm will have large amount time to finish all queued jobs. Like FCFS algorithm the CloudSim toolset also supports Round Robin scheduling used for internal scheduling of jobs. The main demerit of RR is that, for large number of jobs the time taken for completion is pity high.
- **Generalized Priority Algorithm: -** Sometimes according to the customer's demand priority have to be defined on the basis of parameters of cloudlet like size, memory, bandwidth etc. In the traditional way, the jobs are given priority on the basis of their size. One task that has highest size is assigned the top rank. The Virtual Machines are given priority on the basis of the MIPS value. The virtual machine with largest MIPS value is assigned the top rank and vice versa. Thus, the main idea for giving the priority to the tasks and to VMs is size and MIPS respectively. This policy is far better than FCFS and Round Robin scheduling. Take 5 Virtual Machines and represent them by their Ids and MIPS like V = {{1, 350}, {2, 2000}, {3, 355}, {4, 505}, {5, 350}}. Here Vm2 will get first preference because it has highest MIPS, second preference is given to Vm4 and then to Vm3, rest Vm1 and Vm5 get further preferences.
- **Max-Min Algorithm**: This algorithm selects the task with highest size and then this executes it first.
- **Min-Min Algorithm:** This algorithm selects the smaller task first and then executes it first. After that in increasing order of the tasks the are executed.

Procedure of Scheduling

There are three stages to schedule the tasks in the cloud environment and they are:

1. Discover a get resource and filter them in sequence.
2. The decision stage is to select a target resource.

3. Last stage is to submit a particular task to a target resource.

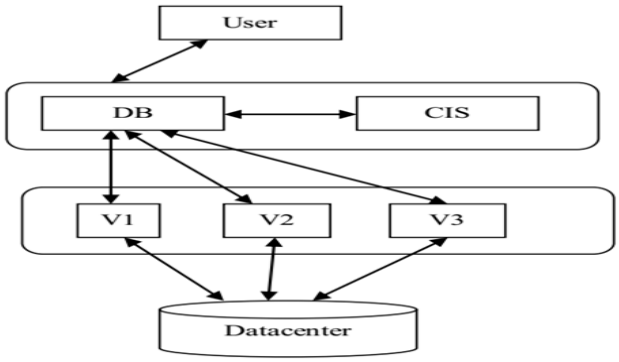


Figure 2.3 Stages of scheduling

DB: - Datacenter Broker
CIS: - Cloud Information services
V1, V2, V3 are the virtual machines

II. FUZZY SCHEDULING

The proposed fuzzy Scheduling uses fuzzy logic. Fuzzy logic does not have disciplined assigning of the elements to sets such as binary logic in place of that degree of membership is assigned to the elements. This degree of membership is denoted in between 0 and 1. A fuzzy system needs to be constructed to apply fuzzy set theory to a particular problem in cloud like scheduling between cloudlets or virtual machines. The construction of fuzzy system mainly has two following steps:

1) **Fuzzyfication**: In fuzzification, the degree of membership to the element as input is assigned to fuzzy sets which is given by Y→ [0, 1], where Y is the set of values that are used as input. So therefore each and every input value that is to be set is mapped to a value between zero and one.

2) **De-Fuzzyfication**: In defuzzification, the numerical output value is always gained from the output set of values after going through the fuzzy control system. Defuzzification is the complementary of fuzzifation.

Where variable {x1,x2,….xm} and {r1,r2,……rm} are job vector and resource vector respectively and K1 and V1 are the fuzzy cluster where k1 is the key and v1 is the value

2.3.1 Fuzzy control system

The difference between set theory and fuzzy logic is that the fuzzy set theory makes use of fuzzy logic to represent uncertainties whereas the set theory ignores this aspect. Fuzzy set theory accepts circumstances that are either partially true or false. Fuzzy logic works good for calculating the logics as it considered the random uncertainty i.e. some prediction of events that the not going to happen or some that get ignored due to some circumstances. A fuzzy control system is said to be the rule-based system which focus on the rules specified, which shows a control mechanism that takes decision to adjust the effects. The main goal of the fuzzy control system is to succeed a skillful human operator with fuzzy logic system based on fuzzy rules. On the basis of network's current state interference engine, an algorithm is used to design fuzzy rules that determine the internet based decision that adjust the manner of the system that is acceptable. Mainly there are two types of fuzzy logic controller: one is feedback controller that accepts the input from the user and according to that produce the output which act as an output to the fuzzy controller which further used to adjust the dimensions of the system. There are series of steps used to design the fuzzy control system. The first pace in fuzzy control comprises of defining the control variables and input variables. Each control and input variable needs to be quantified. After that each quantification of the variable is specifically assigned a membership function. After assigning member function a fuzzy rule should be designed to determine the action that needs to take place under given inputs. The fuzzy rules used are described in an "if-then" format. To execute the individual 'if and then rules' in the rule, an implication formula is used. To obtain sum of the result of rule to have a fuzzy output set a composition rule is used. In specified fuzzy system, the Mamdani minimum inference method is considered as the fuzzy inference method.

III. PROPOSED ALGORITHM

The proposed work explores about the task scheduling mechanism in the cloud network that makes use of linear programming. The vital idea of this system is to schedule the task expressed by the user in the cloud computing environment to the resources that are available in a way that the user needs to pay the minimal amount to the service provided by the cloud. In return the provider accomplishes the task of executing the user's request. There is another way to fulfill this task by making use of the batch image. This system consists of providing the batch images as tasks to the cloud service provider. The task of scheduler is to schedule the jobs by making use of computing resources. For this thing Fuzzy C-Means (FCM) algorithm is used as clustering technique to make the clusters that is far better than the linear programming.

3.1.1 Fuzzy C-Means Algorithm

Fuzzy C-Means (FCM) Algorithm is a clustering technique that gives the optimized solution and it is the reciprocal of the linear programming. FCM algorithm has the capability to use more than two set of data points in more than one cluster. In this algorithm the set of data points that are selected from some data sets may belong to more than two clusters at the same time. In Fuzzy C Mean clustering technique, each and every data point is assigned to a set of values that is called membership values. Membership value is given between 0 and 1 at any given point with respect to specified. Thus on the basis of FCM algorithm these membership values are assigned and on the basis of these values, each and every data point is assigned to the cluster.

Implementation of FCM algorithm:

1. Let there be n number of elements and they are given name as S = {s1, s2,…..s_n} and they are to be grouped by using FCM in t number of clusters which are given name as C = {C1, C2,….. C_t}

2. Predecide the number of clusters that should be made from clustering.

3. Choose any membership value 'w_{ij}' that shows the belongingness of point s_i to cluster C_j and the membership value should be random.

Where, i = {1,2,3,……,n} and j = {1,2,3,4,……..,t}

3. Repeat this step unless the difference between memberships values in next coming loops is not mare than the threshold, ϵ.

a) Calculate the centroid by making use of the following formula.

$$c_j = \frac{\sum_{i=1}^{n} w_{ij}^{p} s_i}{\sum_{i=1}^{m} w_{ij}^{p}}$$

Here,
cj = centroid of the j[th] cluster and
P = level of fuzziness that decides the membership value.

b) In the next step calculate membership value by making use of the following formula.

$$w_{ij} = \frac{\left(1/dist(s_i,c_j)^2\right)^{\frac{1}{p-1}}}{\sum_{q=1}^{t}\left(1/dist(s_i,c_q)^2\right)^{\frac{1}{p-1}}}$$

For each point the membership value is calculated that help us to compute the centroids and corresponding clusters

IV. RESULT ANALYSIS

In the first module a cloud environment is made which have the set of virtual machines and the tasks that are entered with their respective size. To access the cloud environment username and password are to be set which are handled by the admin.

Figure 4.1 Cloud login page

2. After login with the specified username and password we can access the cloud.
In next step add task name with the execution time specified as size.

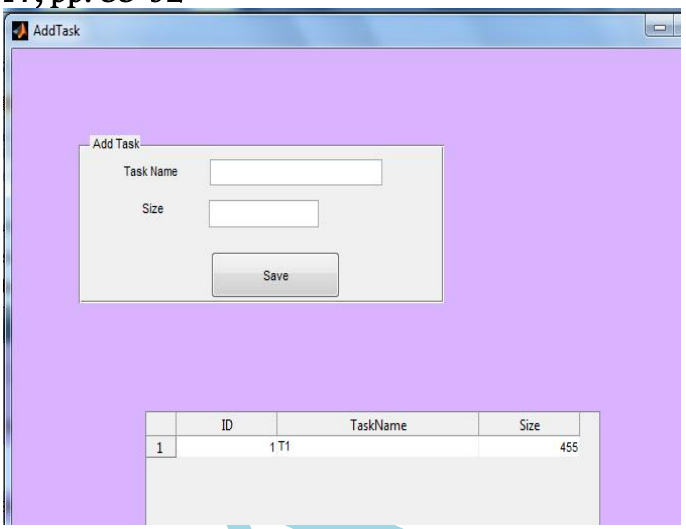


Figure 4.2 Adding task in cloud environment

3. In next step add the virtual machine with the size that can handle maximum number of tasks. These are embedded in the SQL at the backhand. The list of all the virtual machine with their respective size is assigned and can be checked in through SQL. The data can be viewed in SQL.

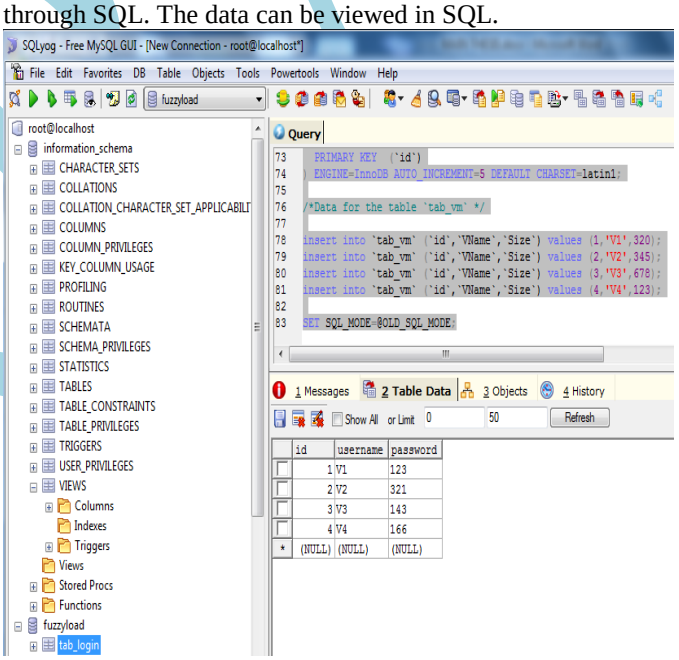


Figure 4.3 Virtual machine data in SQL

After adding the virtual machine name and the required parameter, it gets stored in the SQL in the form of table which can go under certain number of changes like adding VMs, deleting VMs and so on.

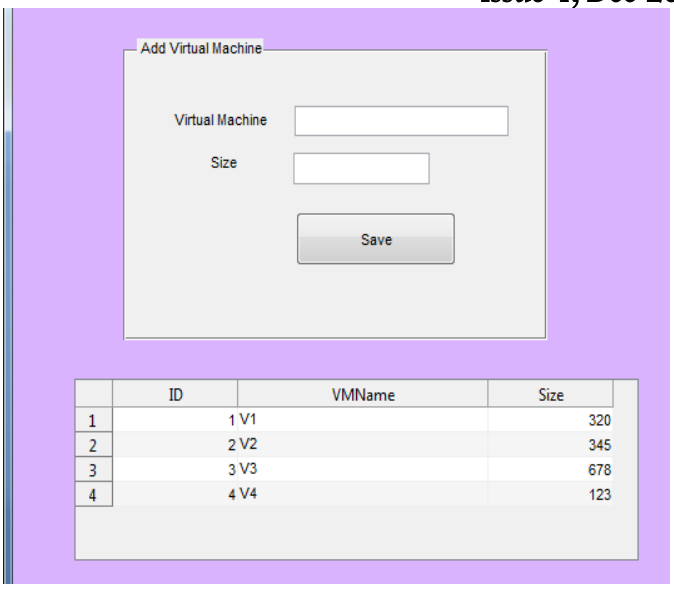


Figure 4.4 Add virtual machines

4. After adding the task name and the virtual machine name, fuzzy clustering is applied on the tasks.
To make fuzzy clustering possible the number of cluster and the threshold (between 0 and 1) are specified. This forms the graph by making use of the member function.

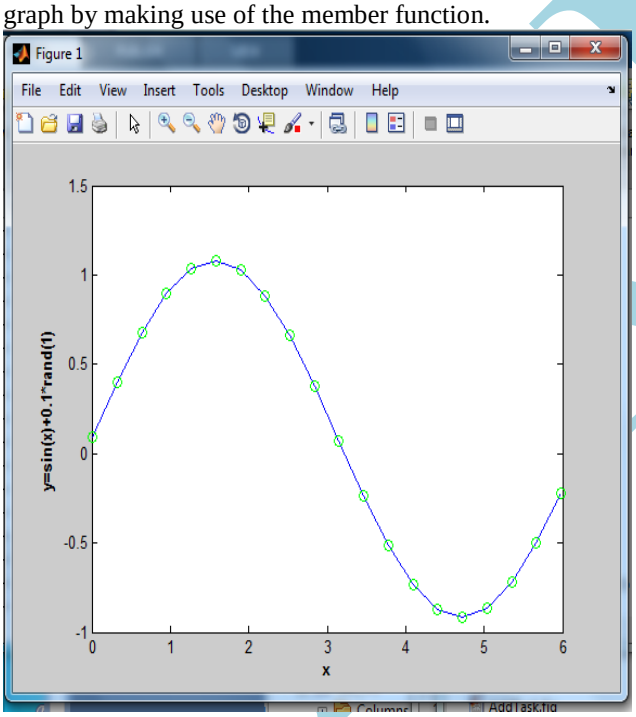
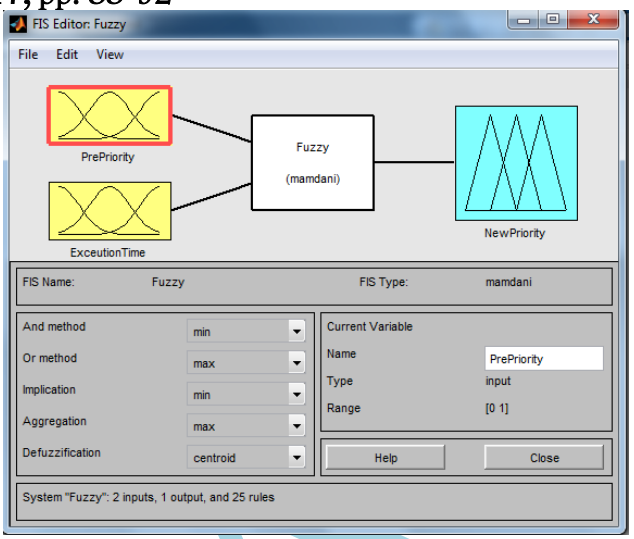


Figure 4.5 Function used for clustering

5. Specify the fuzzy rules by making use of MATLAB tools. These fuzzy rules when collaborated with the member function proceeds towards the scheduling process. Fuzzy rules are specified for making the fuzzy clusters.
6. The prepriority and execution time are the parameters used as the input. Fuzzy rules are applied to them to form fuzzy clusters. The new priority is obtained as the output.
7. Fuzzy rules are implemented by making use of Fuzzy logic design tools in MATLAB.



.Figure 4.6 Fuzzy design

8. In this fuzzy system, there are 2 inputs and one output is obtained by making use of 25 rules. The rules are described in the below given figure

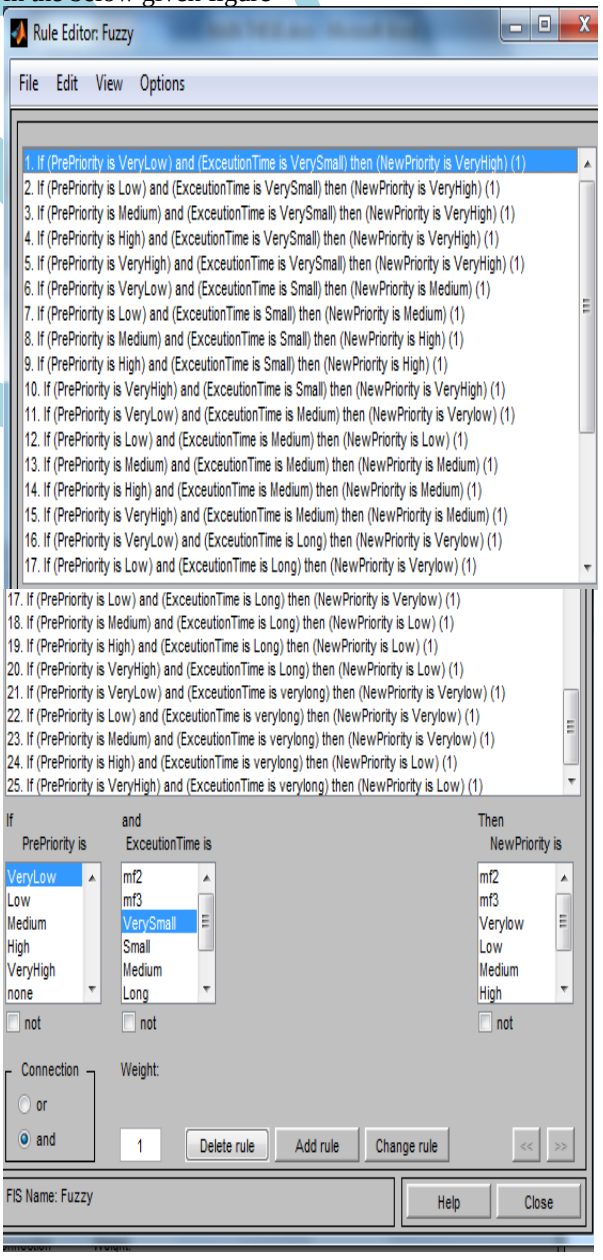


Figure 4.7 Fuzzy rules

8. On the basis of rules made they can be viewed. The graph and plotting change according to the input that we enter.

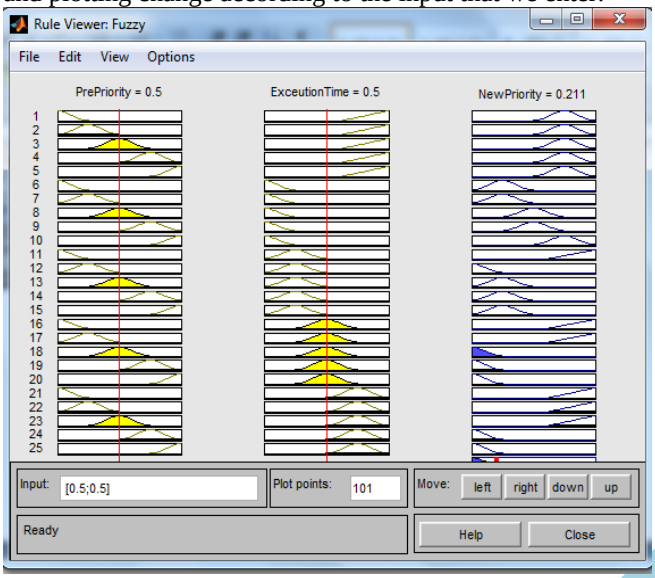


Figure 4.8 Rules view window

9. After the clustering of tasks the SJF scheduling algorithm is used.

The algorithm used is as follow

i. Take number of fuzzy clusters to be inserted.
ii. Select the cluster which has shortest burst time among all process will execute first.
iii. If two cluster have same burst time length then FCFS (First come First Serve) scheduling algorithm used.
iv. Start with first process, selection as above and other processes are to be in queue.
v. Calculates Burst total number of time.
vi. Display the Related values.
vii. Now Close process and the assignment is completed.

10. On the basis of this scheduling the output is obtained which is better than the existing scheduling methods. The graph shows the improved algorithm by making use of SJF scheduling. It shows the increased efficiency on increasing number of tasks. The graph shows a high performance ratio in comparison to other scheduling algorithms like FCFS, RR, Bayesian algorithm and many more that are considered in the existing work.

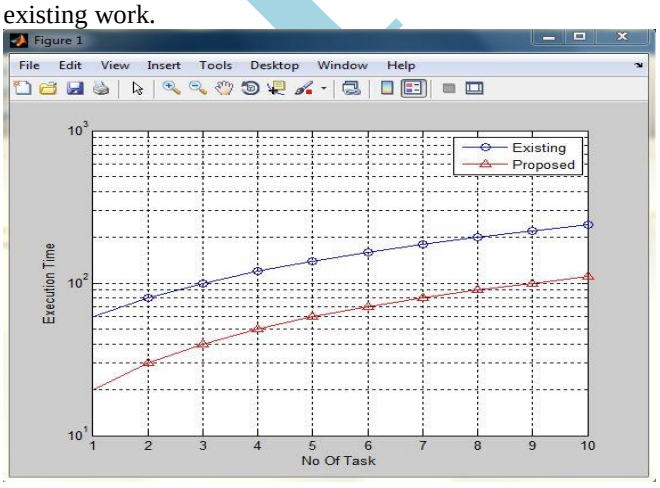


Figure 4.9 output

V. CONCLUSION

There are number of scheduling algorithms that are described till now. Each and every algorithm has its own specification and characteristics, merits and demerits, and application environment. None of the algorithm takes the responsibility of reliability and the efficiency without a fault. These are the optimized algorithms and they work on the input sets and further research work is done in this field. As per this model fuzzy clustering is said to be the best clustering technology and it is better than the algorithms discussed in the literature survey. In future Fuzzy Clustering can be used more effectively with the hybrid algorithm to improve efficiency and performance in cloud network and more optimized algorithm will be obtained.

REFERENCES

[1]. Godoy.D and Amandi. A, "User Profiling in Personal Information Agents: A Survey," Knowledge Eng. Rev., vol. 20, no. 4, pp. 329 361, 2005.

[2]. J.A. Iglesias, A. Ledezma, and Sanchis .A, "Creating User Profiles from a Command Line Interface: A Statistical approach," Proc. Int"l Conf. User Modeling, Adaptation, and Personalization (UMAP), pp. 90 101, 2009.

[3]. Schonlau .M, Dumouchel.W, W.H. Ju, Karr, and Theus, "Computer Intrusion: Detecting Masquerades," Statistical Science, vol. 16, pp. 58 74, 2001.

[4]. R.A. Maxion and Townsend .T.N, "Masquerade Detection Using Truncated Command Lines," Proc. Int"l Conf. Dependable Systems and Networks (DSN), pp. 219 228, 2002.

[5]. Alaniz Macedo .A, K.N. Truong, Camacho Guerrero , and Graca Pimentel, "Automatically Sharing Web Experiences through a Hyperdocument Recommender System," Proc. ACM Conf. Hypertext and Hypermedia (HYPERTEXT "03), pp. 48 56, 2003.

[6]. Pepyne .D.L, J. Hu, and W. Gong, "User Profiling for Computer Security," Proc. Am. Control Conf., pp. 982 987, 2004.

[7]. Godoy .D and A. Amandi, "User Profiling for Web Page Filtering," IEEE Internet Computing, vol. 9, no. 4, pp.56 64, July/Aug. 2005.

[8]. Anderson .J, Learning and Memory: An Integrated Approach. John and Sons, 1995.

[9]. Y. Horman and G.A. Kaminka, "Removing Biases in Unsuper vised Learning of Sequential Patterns," Intelligent Data Analysis, vol. 11, no. 5, pp. 457 480, 2007.

[10]. Riley .P and M.M. Veloso, "On Behavior Classification in Adversarial Environments," Proc. Int"l Symp. Distributed Autonomous Robotic Systems (DARS), pp. 371 380, 2000.

[11]. E. Fredkin, "Trie Memory," Comm. ACM, vol. 3, no. 9, pp. 490 499,1960.

[12]. Iglesias .J.A, A. Ledezma, and Sanchis .A, "Sequence Classification Using Statistical Pattern Recognition," Proc. Int"l Conf.Intelligent Data Analysis (IDA), pp. 207 218, 2007.

[13]. Kaminka G.A, M. Fidanboylu, A. Chang, and M.M. Veloso,"Learning the Sequential Coordinated

Behavior of Teams from Observations," Proc. RoboCup Symp., pp. 111 125, 2002.

[14]. J.A. Iglesias, A. Ledezma, and Sanchis .A, "A Comparing Method of Two Team Behaviours in the Simulation Coach Competition,"Proc. Int"l Conf. Modeling Decisions for Artificial Intelligence (MDAI),pp. 117 128, 2006.

[15]. Agrawal .R and Srikant .R, "Mining Sequential Patterns," Proc.Int"l Conf. Data Eng., pp. 3 14, 1995.

