ISSN (Online) : 2454-4159
International Journal of Advanced Research in Science Management and Technology
Volume 2, Issue 8, August 2016

An efficient ACO and PSO based method for Web pages Optimization

Abhijit Ahirrao[1], Prof. Kailash Patidar[2], Prof. Manoj Yadav[3]

PG Student, Dept. of CSE, SSSUTMS, Sehore, M.P., India[1]
Professor & Head, Dept. of CSE, SSSUTMS, Sehore, M.P., India[2]
Assistant Professor, Dept. of CSE, SSSUTMS, Sehore, M.P., India[3]

ABSTRACT - The websites popularity is increasing in all areas. Because of ease in use, data gathering, sharing, business everything is possible through World Wide Web (WWW). So the website ranking and optimization determines that which website is open first and shown in the Google search. The traffic determination and their optimization is a challenging issue. So in this paper we have applied association rule mining algorithm has been applied to find the ranking. The ranking has two base one is for individula and another is for the associated mining. Then we apply ACO and Random weight PSO algorithm for the comparison. By this method we will be able to find the optimized values and the associated values both.

KEYWORDS- Domain, WWW, Data Mining Techniques, Optimization.

I. INTRODUCTION

The World Wide Web has grown in the past couple of years from a gathering to the best and most conspicuous system for correspondence and information dissipating. Reliably, the WWW gets to be by around a million electronic pages, adding to the numerous millions starting now on-line [1]. WWW serves as a stage for exchanging diverse sorts of information, reaching out from investigation papers, and enlightening substance, to intuitive media substance and programming [2]. The relentless advancement in the size and the usage of the WWW compels new procedures for changing these huge measures of data [3]. Because of its quick and cluttered improvement, the resulting arrangement of information nonattendances of affiliation and structure [4]. Furthermore, the substance is appropriated in distinctive varying plans. Due to this, customers are feeling here and there puzzled, lost in that information over-weight that returns to augment. Issues that must be overseen of relevant information, including the looking for and indexing of the Web content, the arrangement of some meta learning out of the information which is available on the Web, and the having a tendency to of the individual customers' prerequisites and redirections, by modifying the gave information and administrations[5].

There are a few methods are utilized as a part of the heading of page positioning. K-means grouping, fluffy C-means and enhancement strategies are utilized as a part of a few examination work like [6][7][8][9][10]. This additionally incorporates the utilization of affiliation principle mining as the significant information ought to be separated [11][2][13][14][15].

The same strategy with diverse method for improvement is proposed in [21]. As the mining concentrate the learning from the colossal database and it is additionally helpful in gathering taking into account closeness name's and class which can be a superior apparatus for understanding and portraying the information from the wellspring of activity. This phenomena is likewise proposed in [15][16][17][18][19][20]. Our principle inspiration of this exploration work is to apply enhancement methods with information mining system to attain to aggregate streamlined rank.

II. ACO AND PSO

ISSN (Online) : 2454-4159
International Journal of Advanced Research in Science Management and Technology
Volume 2, Issue 8, August 2016

Ant Colony Optimization (ACO) [34[35] The Ant Colony Optimization (ACO) algorithm is a meta-heuristic which is a grouping of distributed environment, positive feedback system, and systematic greedy approach to find an optimal solution for combinatorial optimization problems. This algorithm tries to mimic the ant's behavior in the real world.

The Ant Colony Optimization algorithm is mainly inspired by the experiments run by Goss et al. [34] which using a grouping of real ants in the real environment. They study and observe the behaviour of those real ants and suggest that the real ants were able to select the shortest path between their nest and food resource, in the existence of alternate paths between the two. The above searching for food resource is possible through an indirect communication known as stigmergy amongst the ants. When ants are travelling for the food resources, ants deposit a chemical substance, called pheromone, on the ground. When they arrive at a destination point, ants make a probability based choice, biased by the intensity of pheromone they smell. This behavior has an autocatalytic effect because of the very fact that an ant choosing a path will increase the probability that the corresponding path will be chosen again by other ants in the next move. After finishing the search ants return back, the probability of choosing the same path is higher because of increasing pheromone quantity. So by the pheromone will be released on the chosen path, it provides the path for the ants. In short we can say that, all ants will select the shortest path.

If we analyses the case then we observed that because of the same pheromone laying the shortest path will be chosen. It will be starts with first ants which arrive at the food source are those that took the two shortest branches. After approaching the food destination these ants start their return trip, more pheromone is present on the short branch is the possibility for choosing the shortest one than the one on the Long Branch. This ant behavior was first formulated and arranged as Ant System (AS) by Dorigo et al. [34]. Based on the AS algorithm, the Ant Colony Optimization (ACO) algorithm was proposed [35]. In ACO algorithm, the optimization problem is formulated as a graph G = (C; L), where C is the set of components of the problem, and L is the set of possible connections or transitions among the elements of C. The proposed solution is represented in terms of feasible paths on the graph G, with respect to a set of given constraints and predicate. The population of ants that is also called agents collectively solves the problem under consideration using the graph representation. We assume that the ants are probably very poor of finding a solution, good quality solutions can emerge as a result of collective interaction amongst ants. Pheromone trails encode a long-term memory about the whole ant search process from the starting to the food resource destination. The value depends on the problem formulation, representation and the optimization objective which is different in case to case.

The algorithm presented by Dorigo et al. [35] was given below:

Algorithm ACO meta heuristic ();
while (termination criterion not satisfied)
ant generation and activity();
pheromone evaporation();
daemon actions();
"optional"
end while
end Algorithm

Particle Swarm Optimization (PSO) [36][37]

PSO gained from the situation and utilized it to take care of the improvement issues. In PSO, each one single arrangement is a "fledgling" in the hunt space. We call it "particle". All of particles have wellness values which are assessed by the wellness capacity to be upgraded, and have speeds which administer the flying of the particles. The particles fly through the issue space by

ISSN (Online) : 2454-4159
International Journal of Advanced Research in Science Management and Technology
Volume 2, Issue 8, August 2016

emulating the current ideal particles. PSO is introduced with a gathering of arbitrary particles (arrangements) and afterward hunt down optima by redesigning eras. In every cycle, every particle is overhauled by taking after two "best" values. The first is the best arrangement (wellness) it has accomplished as such. (The wellness quality is likewise put away.) This worth is called pbest. An alternate "best" esteem that is followed by the particle swarm analyzer is the best esteem, acquired so far by any particle in the populace. This best esteem is a worldwide best and called gbest. At the point when a particle partakes of the populace as its topological neighbors, the best esteem is a neighborhood best and is called lbest.

In the wake of discovering the two best values, the particle upgrades its speed and positions with taking after comparison (an) and (b).

v[] = v[] + c1 * rand() * (pbest[] - present[]) + c2 * rand() * (gbest[] - present[]) (a)

present[] = present[] + v[] (b)

v[] is the particle velocity, present[] is the current particle (solution). pbest[] and gbest[] are defined as stated before. rand () is a random number between (0,1). c1, c2 are learning factors. usually c1 = c2 = 2.

The pseudo code of the procedure is as follows

```
For each particle
Initialize particle
END
Do
   For each particle
      Calculate fitness value
      If the fitness value is better than the best fitness value (pBest) in history
         set current value as the new pBest
   End
 Choose the particle with the best fitness value of all the particles as the gBest
   For each particle
     Calculate particle velocity according equation (a)
      Update particle position according equation (b)
   End
```

While maximum iterations or minimum error criteria is not attained Particles' velocities on each dimension are clamped to a maximum velocity Vmax. If the sum of accelerations would cause the velocity on that dimension to exceed Vmax, which is a parameter specified by the user. Then the velocity on that dimension is limited to Vmax.

III. IMPLEMENTATION

The data has been gathered from google trends. This data is first extracted from google trends. This data is then converted into rows and columns. Then association rule mining algorithm has been applied to find the ranking. The ranking has two base one is for individula and another is for the associated mining. The principle purpose of web utilization mining is to watch the behavioral examples of clients interrelating with a site. The found examples are by and large described as a gathering of pages, articles or

ISSN (Online) : 2454-4159
International Journal of Advanced Research in Science Management and Technology
Volume 2, Issue 8, August 2016

assets which are consistently gotten to by gatherings of clients with normal diversions. So the investment areas have been found by Apriori calculation.

Then we apply ACO and Random weight PSO algorithm. Each ants and particle knows its best value so far (pbest) and its position in case of Iteartive PSO. Here the ants and the particles are assumed to as the item set values. This data is similarity of individual encounters of every molecule. Also, every molecule knows the best esteem so far in the gathering (gbest) among pbests. This data is relationship of information of how alternate particles around them have performed. Every molecule tries to adjust its position in light of current positions, current speeds, remove between the current position and pbest and separate between the current position and gbest.

The primary target of this research paper is to discover the upgraded patterns in a year which can help to anticipate alternate pages on the same website in this admiration through the month. For this first apriori calculation is connected. By apriori calculation continuous examples are discover so that in light of an obliged least bolster esteem a set can be qualified. Then we apply aco and pso for checking the optimized value up to the 11 month specified. The algorithm is shown below. The termenology used are PS for particleset, PS1, PS2 is for particleset1, particleset2 etc. GS for groupset, GS1,GS2 for the group set values. RV is used for result values. S_{AC} is for the self accuracy.

ACO [33]

Input:

- PS(ps1,ps2….psn)
- GS(Gs1,Gs2….Gsn)

Output:

- RV1…….RVn

Step 1: Input Set

Step 2: Initialize pheromone as the trends obtained month wise

Step 3: Check the PS pheromone set for the relevancy

For 1 to 5

$T_V=(PS_1 + PS_2 + PS_3 +…. + PS_n)/n$

$P_t= T_v - E_p$

$E_p = \{0.2, 0.4, 0.6, 0.8\}$

If $(P_{t1} > P_{tn-1})$

$P_{t1} = P_{tn-1}$

Step 4: Final set

For 1 to 5

$T_V=(T_1 + T_2 + T_3 +…. + T_n)/n$

$P_t= T_v - E_p$

$E_p = \{0.2, 0.4, 0.6, 0.8\}$

If $(P_{t1} > P_{tn-1})$

$P_{t1} = P_{tn-1}$

Step 5: Overall Accuracy

$S_{AC}=\sum P_{t1} / n$

Step 6: Finish

Then we apply particle swarm optimization for the comparison. The algorithm is shown below:

Random PSO

PSO

Input:

- PS(ps1,ps2….psn)
- OS(Os1,Os2….Osn)

ISSN (Online) : 2454-4159
International Journal of Advanced Research in Science Management and Technology
Volume 2, Issue 8, August 2016

Output:

- ET1…….ETn

Ps→ Particles
OS→Optimal Set
ET→Efficient Trails
V→ Velocity
RV→Random Velocity
RV_p → Previous Random Velocity
Step 1: Google Trends Values
Step 2: Initialize particle
Step 3: Random Velocity Calculation
for i=0 ;i<=5;i++
RV_i=Math.random();
Step 4: Distribute PS for the below Iteration

do
$E_V=(PS_1*RV_1 + PS_2* RV_2 + PS_3 * RV_3 +…. + PS_n * RV_n)/n$
If ($V_{t1} > V_{tn-1}$)
$V_{t1} = V_{tn-1}$
$RV_p = RV_i$
while;

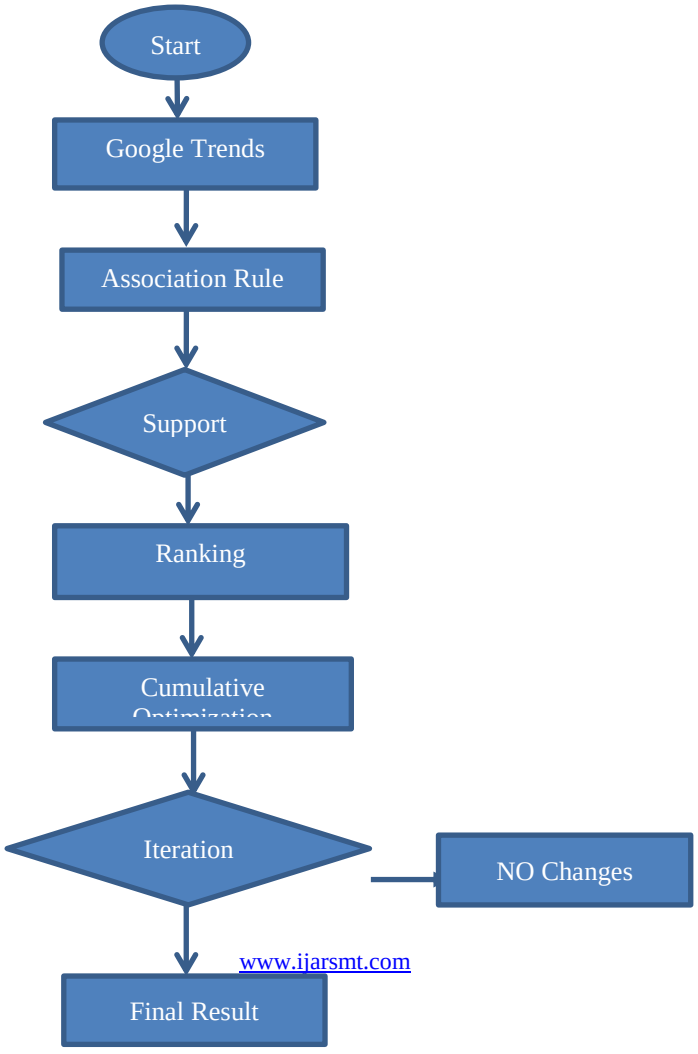
For 2 to 5

$T_V=(PS_1*RV_1 + PS_2* RV_2 + PS_3 * RV_3 +…. + PS_n * RV_n)/n$ - value(RV_p)
$V_{t1} = V_{tn-1}$
If ($V_{t1} > V_{tn-1}$)
$V_{t1} = V_{tn-1}$

Step 5: Overall Accuracy

$O_{AC}=\sum PS_i / n$

Step 6: Finish



ISSN (Online) : 2454-4159
International Journal of Advanced Research in Science Management and Technology
Volume 2, Issue 8, August 2016

Fig. 1 Flowchart

1. Result

The optimized iteration values after ACO and Random weight PSO are shown below.

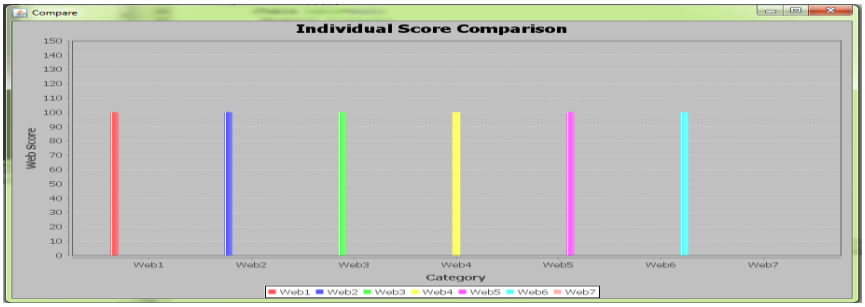


Fig. 2 Associated Rank

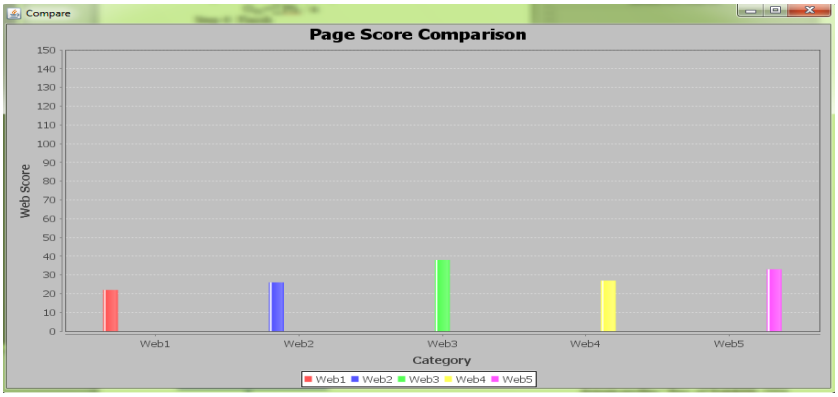


Fig. 3 Without Optimization

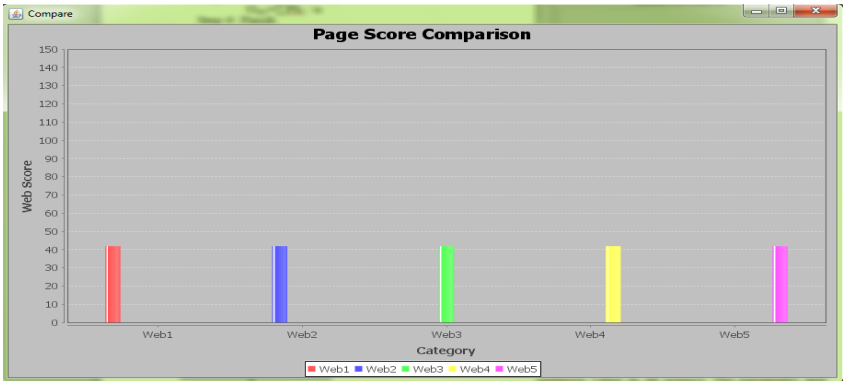


Fig. 4 With ACO Optimization

ISSN (Online) : 2454-4159
International Journal of Advanced Research in Science Management and Technology
Volume 2, Issue 8, August 2016

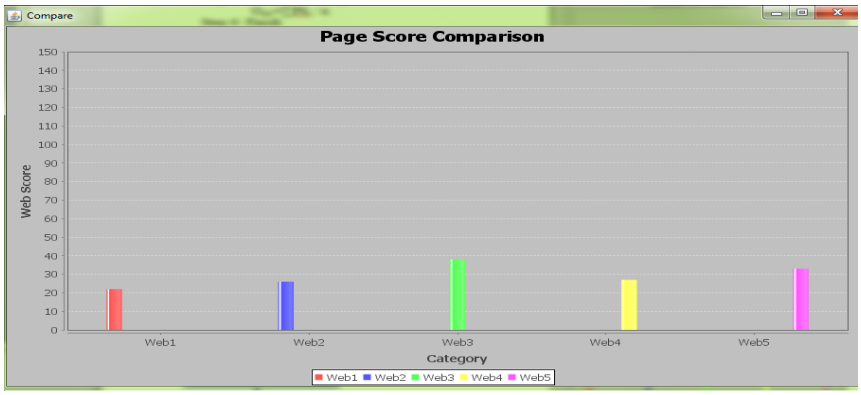


Fig. 5 Without Optimization

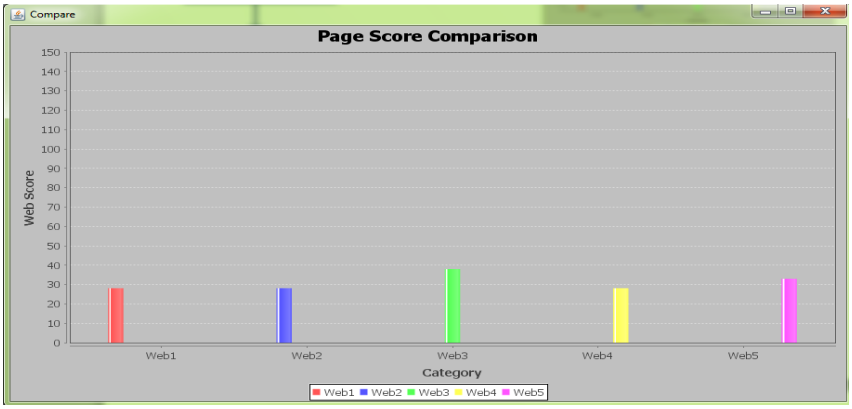


Fig. 6 With Random weight PSO Optimization

IV. CONCLUSION

This research paper provides a direction of web mining and web traffic optimization so that the impact have been discussed and analyzed. This research emphasize the focus on the data that can be seen to enhance the optimize value in all respect. The parametric and probabilistic changes has been notices and discussed. We have applied ACO and Random weight PSO to analyzed the web data and find the overall optimized ranking. The result suggests the chances of optimized ranking are equal with ACO and Random weight PSO.

REFERENCES

[1] Aktas, Mehmet S., Mehmet A. Nacar, and Filippo Menczer. "Personalizing PageRank based on domain profiles." Proc. of WebKDD. 2004.

[2] Berendt, Bettina. "Understanding Web usage at different levels of abstraction: coarsening and visualizing sequences." Proc. of the Mining Log Data across All Customer TouchPoints Workshop (WEBKDD '01), San Francisco, CA. 2001.

[3] Berendt, Bettina. "Using site semantics to analyze, visualize, and support navigation." Data Mining and Knowledge Discovery 6.1 (2002): 37-59.

[4] Baraglia, Ranieri, and Fabrizio Silvestri. "An online recommender system for large web sites." Web Intelligence, 2004. WI 2004. Proceedings. IEEE/WIC/ACM International Conference on. IEEE, 2004.

[5] Deng, Lin, et al. "Spying out real user preferences for metasearch engine personalization." Proc. of ACM WEBKDD. 2004.

[6] Chen, K. and Liu. L. A random rotation perturbation approach to privacy data classification. In Proc of IEEE Intl. Conf. on Data Mining (ICDM), pp. 589-592, 2005.

[7] Kai Li, Lijuan Cui, "A Kernel Fuzzy Clustering Algorithm with Generalized Entropy Based on Weighted Sample", International Journal of Advanced Computer Research (IJACR), Volume-4, Issue-15, June-2014, pp.596-600.

[8] Kai Li, Peng Li, " A Selective Fuzzy Clustering Ensemble Algorithm", International Journal of Advanced Computer Research (IJACR), Volume-3, Issue-13, December-2013, pp.1-6.

[9] Shyi-Ching Liang, Yen-Chun Lee and Pei-Chiang Lee, "The Application of Ant Colony Optimization to the Classification Rule Problem", 2011 IEEE International Conference on Granular Computing.

[10] Dubey, Ashutosh Kumar, Umesh Gupta, and Sonal Jain. "A Survey on Breast Cancer Scenario and Prediction Strategy." In Proceedings of the 3rd International Conference on Frontiers of Intelligent Computing: Theory and Applications (FICTA) 2014, pp. 367-375. Springer International Publishing, 2015.

[11] Anshuman Singh Sadh, Nitin Shukla," Association Rules Optimization: A Survey", International Journal of Advanced Computer Research (IJACR), Volume-3 Number-1 Issue-9 March-2013.

[12] Ashutosh Kumar Dubey, Animesh Kumar Dubey, Vipul Agarwal, Yogeshver Khandagre, "Knowledge Discovery with a Subset-Superset Approach for Mining Heterogeneous Data with Dynamic Support", Conseg-2012.

[13] Animesh Dubey, Rajendra Patel and Khyati Choure, " An Efficient Data Mining and Ant Colony Optimization technique (DMACO) for Heart Disease Prediction " , International Journal of Advanced Technology and Engineering Exploration (IJATEE), Volume-1, Issue-1, December-2014 ,pp.1-6.

[14] Prashant Sharma, " Association Rule Mining with enhancing List Level Storage for Web Logs: A Survey ", International Journal of Advanced Technology and Engineering Exploration (IJATEE), Volume-1, Issue-1, December-2014, pp.15-20.

[15] Anshuman Singh Sadh, Nitin Shukla, "Apriori and Ant Colony Optimization of Association Rules", International Journal of Advanced Computer Research (IJACR), Volume-3, Issue-10, June-2013, pp.35-42.

[16] Shawana Jamil, Azam Khan, Zahid Halim and A. Rauf Baig," Weighted MUSE for Frequent Subgraph Pattern Finding in Uncertain DBLP Data", IEEE 2011.

[17] Ashwin C S, Rishigesh.M and Shyam Shankar T M," SPAAT-A Modern Tree Based Approach for sequential pattern mining with Minimum support", IEEE 2011.

[18] K. Zuhtuogullari and N. Allahverdi ,"An Improved Itemset Generation Approach for Mining Medical Databases", IEEE 2011.

[19] Ruchita Gupta, C.S.Satsangi, " An Efficient Range Partitioning Method for Finding Frequent Patterns from Huge Database ", International Journal of Advanced Computer Research (IJACR), Volume-2, Issue-4, June-2012 ,pp.62- 69.

[20] Yadav, M.P.; Feeroz, M.; Yadav, V.K., "Mining the customer behavior using web usage mining in e-commerce," Computing Communication & Networking Technologies (ICCCNT), 2012 Third International Conference on , vol., no., pp.1,5, 26-28 July 2012

[21] Shushma Lata, "An Iterative PSO for Web worth Optimization through random velocity ", International Journal of Advanced Technology and Engineering Exploration (IJATEE), Volume-2, Issue-3, February-2015, pp.31-36.

[22] Sharma, K.; Shrivastava, G.; Kumar, V., "Web mining: Today and tomorrow," Electronics Computer Technology (ICECT), 2011 3rd International Conference on, pp.399, 403, 8-10 April 2011.

[23] Manjusha, R.; Ramachandran, R., "Web mining framework for security in e-commerce," Recent Trends in Information Technology (ICRTIT), 2011 International Conference on , pp.1043,1048, 3-5 June 2011.

[24] Sampath, P.; Ramesh, C.; Kalaiyarasi, T.; Banu, S.S.; Selvan, G.A., "An efficient weighted rule mining for web logs using systolic tree," Advances in Engineering, Science and Management (ICAESM), 2012 International Conference on , pp.432,436, 30-31 March 2012.

[25] Nassar, O.A.; Al Saiyd, N.A., "The integrating between web usage mining and data mining techniques," Computer Science and Information Technology (CSIT), 2013 5th International Conference on, pp.243,247, 27-28 March 2013.

[26] Azad, H.K.; Abhishek, K., "Entropy measurement and algorithm for Semantic-Synaptic web mining," Data Mining and Intelligent Computing (ICDMIC), 2014 International Conference on , pp.1,5, 5-6 Sept. 2014.

[27] Azad, Hiteshwar Kumar, and Kumar Abhishek. "Semantic-Synaptic Web Mining: A Novel Model for Improving the Web Mining." Communication Systems and Network Technologies (CSNT), 2014 Fourth International Conference on. IEEE, 2014.

[28] Yadav, M.P.; Feeroz, M.; Yadav, V.K., "Mining the customer behavior using web usage mining in e-commerce," Computing Communication & Networking Technologies (ICCCNT), 2012 Third International Conference on , vol., no., pp.1,5, 26-28 July 2012.

[29] Apeksha Khabia and M. B. Chandak, " A Cluster Based Approach for Classification of Web Results " , International Journal of Advanced Computer Research (IJACR), Volume-4, Issue-17, December-2014 ,pp.934-938.

[30] Huang Qinglan; Duan Longzhen, "Multi-level Association Rule Mining Based on Clustering Partition," Intelligent System Design and Engineering Applications (ISDEA), 2013 Third International Conference on , vol., no., pp.982,985, 16-18 Jan. 2013.

[31] Abul Kalam Mohammad Azad, Md Kabirul Islam, Mohammed Shamsul Hoque, " An analysis of popularity of Consumer to Consumer websites " , International Journal of Advanced Computer Research (IJACR), Volume-4, Issue-16, September-2014 ,pp.828-833.

[32] Tempaiboolkul, J., "Mining rare association rules in a distributed environment using multiple minimum supports," Computer and Information Science (ICIS), 2013 IEEE/ACIS 12th International Conference on , vol., no., pp.295,299, 16-20 June 2013.

[33] Suman Mishra, Prateek Gupta, "An efficient Optimization method for Data Classification" , International Journal of Advanced Computer Research (IJACR), Volume-4, Issue-14, March-2014 ,pp.383-388.

[34] S. Goss, S. Aron, J. L. Deneubourg, and J. M. Pasteels. Self-organized Shorcuts in the Argentine Ant. Naturwissenschaften, 76:579–581, 1989.

[35] M. Dorigo, G. Di Caro, and L. M. Gambardella, "Ant algorithms for discrete optimization," Artif. Life, vol. 5, no. 2, pp. 137–172, 1999.

[36] Keon-Myung Lee,Mining Generalized Fuzzy Quantitative Associaton Rules wih Fuzzy Generalization Hierarchies, IEEE 2011.

[37] Kennedy, James. "Particle swarm optimization." In Encyclopedia of Machine Learning, pp. 760-766. Springer US, 2010.