

Figure 4: in this graph the users in each group is shown after fuzzy clustering this shows that most user does not shows similarity with others.

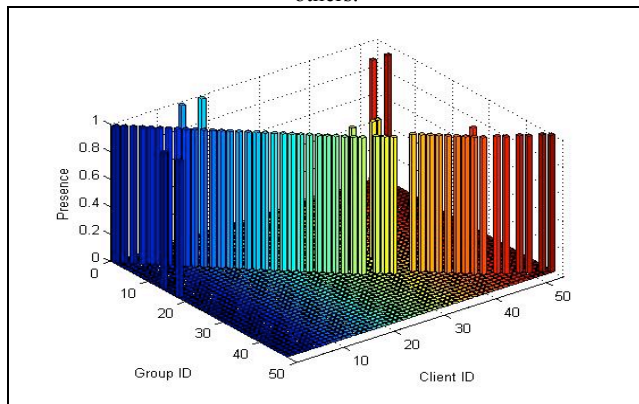


Figure 5: shows how the clients are arranged in groups the presence axis only have binary value which represents presence or absents of client in particular group.

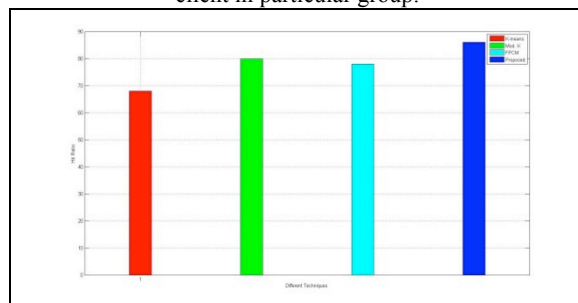


Figure 6: a comparative analysis of hit ratio is shown and the blur bar which is for proposed algorithm reaches up to 95% which is 5% higher than the previous best (in yellow 90%)

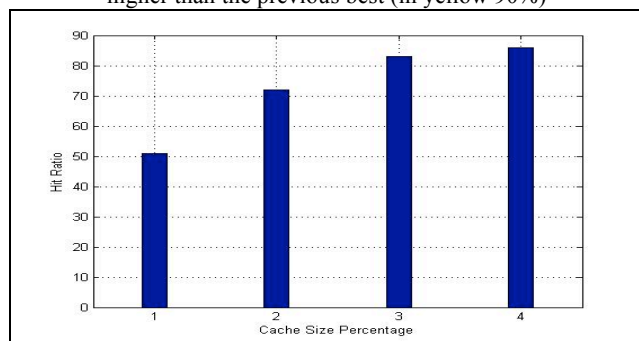


Figure 7: the performance of the proposed algorithm with different cache configuration is shown it shows that the hit ration greatly increases with cache percentage for lower values of cache and gets saturate quickly after 4%.

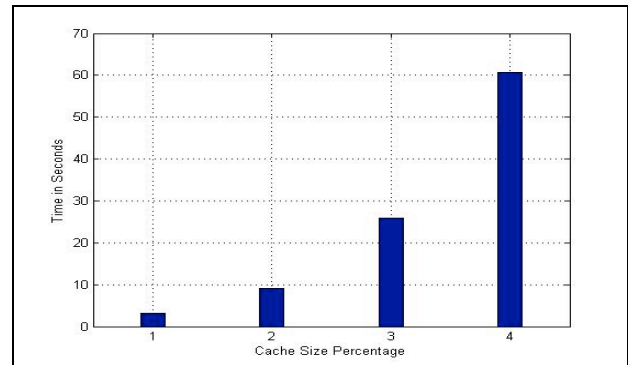


Figure 8: as shown in figure 6 that the performance increases with but it also increases the training and predicting time exponentially.

V. CONCLUSION AND SCOPE

Conclusion: The simulation result shows that the proposed algorithm can provide the hit ratio of 50% by just using 1% of cache and increases very quickly to about 90% in just increasing the cache to 5%, the simulation results also shows that it takes just a few seconds in training although the time increases exponentially but it is still manageable it also show that when it is operated irrespective of cache percentage it could give the hit ratio up to 95%, secondly the proposed emission matrix provides a good similarity measuring ground which could be further used for other methods. Finally it can be said the proposed algorithm works well in terms of hit ratio, latency reduction while requiring minimum resources.

Future Work: the current work has some possibilities of enhancement in future which are

The HMM model could be optimizing for reduction of rare emissions and states.

The Fuzzy clustering can also extend for non linear grouping relations.

Some other Machine learning techniques can also be test.

REFERENCES

- [1] <http://maya.cs.depaul.edu/~mobasher/webminer/survey/node23.html>, "Web Usage Mining Architecture".
- [2] B. Mobasher, N. Jain, E. Han, and J. Srivastava, "Web mining: Pattern discovery from world wide web transactions", Technical Report TR 96-050, University of Minnesota, Dept. of Computer Science, Minneapolis, 1996.
- [3] R. Cooley, B. Mobasher, and J. Srivastava, "Grouping web page references into transactions for mining world wide web browsing patterns", Technical Report TR 97-021, University of Minnesota, Dept. of Computer Science, Minneapolis, 1997.
- [4] Jiawei Han, Micheline Kamber, "Data Mining Concepts and Techniques", Second Edition, China Machine Press, Beijing, 2007.
- [5] Jing-Gong Li, Xiang-Gong Wang, "Fuzzy Set Theory And Application", Chain Science Press, Beijing, 2005.
- [6] J Vellingiri, S.Chenthur Pandian, " A Survey on Web Usage Mining", Global Journal of Computer Science and Technology Volume 11 Issue 4 Version 1.0 March 2011.
- [7] Kobra Etmnani, Mohammad-R. Akbarzadeh, Noorali Raejji Yanehsari, "Web Usage Mining: users' navigational patterns extraction from web logs using Ant-based Clustering Method" IFSA-EUSFLAT 2009.
- [8] Chun-Jung Lin, Fan Wu, I-Han Chiu, "Using Hidden Markov Model to Predict the Surfing User's Intention of Cyber Purchase on the Web".
- [9] Freitag, D., and McCallum, A., "Information Extraction with HMM Structures Learned by Stochastic Optimization", Just Research 2009.