CHAPTER 7

CONCLUSION AND SCOPE FOR FUTURE WORK

7.1 CONCLUSION

This chapter summarises the final conclusion and discussion of the results obtained from the thesis. The thesis proposed different procedures for web client clustering; finding Maximal frequent itemsets, using hash table with linear probing; mining closed frequent itemsets in a parallel manner and utilize the advantage of generating grades for different items in the dataset along with support to generate closed sequence patterns. The results are provided in the corresponding chapters and it can be pointed out that the proposed algorithms perform better when compared to the existing systems.

The third chapter discussed algorithm which forms, validates, verifies and updates the web client clusters using pre-fetching and socket clone methods. The performance of the algorithm is evaluated and is shown in the Figures 3.4 and 3.5. It can be observed that the performance of the system with clustering is more successful when there is similarity between web client requests when compared to the system without clustering.

In the fourth chapter, the thesis proposed the algorithm HBMFI-LP to efficiently find maximal frequent itemsets. The performance of the hash based mining algorithm for maximal frequent itemsets, using linear probing (HBMFI-LP) for connect-4 dataset is given in Figure 4.9. A similar comparison is made using mushroom and chess datasets, also given in Figures 4.10 and 4.11. The results show that the proposed HBMFI-LP
algorithm outperforms 2 to 3 times better when compared to the existing algorithm-MAFIA. The datasets used for experimentation are provided in Appendix-A. The main advantage of this system is linear probing method. Linear probing is easy and efficient for implementation and also for searching, insertion etc.

In chapter five, the thesis proposed the algorithm for association rule mining with weighted support and closed itemset lattices in parallel. Here, to improve the performance, parallel algorithm – count distribution is used to generate closed frequent itemsets and rules are also generated in parallel. The characteristics like scaleup, sizeup and speedup are examined and compared with the existing count distribution algorithm. The results from Figure 5.1 to 5.6 clearly indicate that the proposed algorithm performs better when compared to the count distribution algorithm.

In chapter six, the proposed algorithm GSSM- Graded support sequential pattern mining method is compared with Clospan and WIS algorithms, in terms of time taken to generate the frequent subsequences. The results clearly show that the proposed algorithm – GSSM performs better than the Clospan and WIS algorithm and the same is given in the Figure 6.3. The advantage of this algorithm is that it considers the grade of each item in generating frequent sets without losing any interesting patterns and does not generating uninterested patterns.

7.2 SCOPE FOR FUTURE WORK

The modern era being the age of virtual computing, it is more appropriate if future enhancements have at least in part have something to do with virtual computing. Virtual servers to service separate clusters is a distinct possibility. Improved caching is another paradigm in consideration. Heuristic approaches towards clustering of clients are another approach. The complete
elimination of dispatchers is a definite line of thought and it has intelligent agents roaming the network to absorb patterns and trends. These systems are complex and hard to implement. However, it must be stated that these kinds of systems will generally have a higher rate of accuracy than the existing systems.