CHAPTER 6

CONCLUSION AND SCOPE FOR FUTURE WORK

6.1 CONCLUSION

This research work has redefined a new rule set, namely the informative rule set that presents prediction sequences equal to those presented by the way of association rule set using the confidence value. The informative rule set is significantly smaller than the association rule set, especially when the minimum support is small. The proposed work has characterized the relationships between the informative rule set and the non-redundant association rule set, and revealed that the informative rule set is a subset of the non-redundant association rule set.

The work takes into consideration, the upward closure properties of informative rule set for omission of uninformative association rules, and presented a direct algorithm to efficiently generate the informative rule set without generating all frequent itemsets. The informative rule set generated in proposed work is significantly smaller than both the association rule set and the non-redundant association rule set for a given database that can be generated more efficiently than the association rule set.

The enhancement of Apriori is done by increasing the efficiency of candidate pruning phase by reducing the number of candidates that are generated to further verification. The basic elements considered in the development of the algorithm are number of transactions, average size of
transaction, average size of the maximal large itemsets, number of items, and distribution of occurrences of large itemsets.

The efficiency improvement results from that the generation of the informative rule set needs fewer candidates and database accesses, than that of the association rule set and large memory usage like some other algorithms does. Application of parallel pruning results in reduced candidate itemset generation compared to that of traditional Apriori. The number of database accesses of the proposed algorithm is significantly fewer than other direct methods for generating association rules on all items. Next stage is the fuzzy model deployment for association rule mining in improving the scalability of the transactional database.

The FOSSP approach for mining ARM presented in this work produces faster execution of the ARM and accurate rules being generated. Thread mechanism presented in the FOSSP approach reduces the memory faintness which in turn increases the execution speed. The feedback schemes introduced in FOSSP approach for the partitioned fuzzy data set reduces redundancy and minimize the sensitivity of rules being generated for fuzzy member attribute in one partitioned domain to the other. The experimental result shows that FOSSP approach has 25% better execution speed than the fuzzy adoption variants of Apriori. In addition accuracy of rule generation is also better than the fuzzy Apriori.

The proposed FOSSP approach presented for fuzzy transactions, in which the items are considered as being normal data values, shows that it is easy to formulate and use fuzzy based association rule mining. The proposed model can be employed in mining distinct types of patterns, ordinary fuzzy association rules, fuzzy and approximate functional dependencies, and gradual rules.
The performance measure and experimental results conducted with the real and repository data sets showed that FOSSP is better in terms of pruning time reduction for forming the maximal informative association rules compared to that of fuzzy Apriori model. In addition FOSSP also results in reduced search space for mining large size datasets with more number of unique items generated compared to its fuzzy Apriori counterpart.

6.2 SCOPE FOR FUTURE WORK

In Future, Genetic Algorithm (GA) can be applied on FOSSP to improve the scalable association rule mining with larger itemsets. The element of GA property is utilized to store more number of items in its Gene. The reduced candidate itemset from FOSSP for pruning is maintained in the genes. Genetic operations such as selection, mutation and cross over can be applied to evaluate more appropriated candidate item for respective levels of pruning frequent items in parallel.