CHAPTER VI
CONCLUSIONS

This study has been carried out to develop methodologies in centralized as well as in distributed environment to find privacy preserving association rule mining without revealing any private data or information.

Three methodologies are proposed in this thesis work to hide the sensitive item sets in centralized database environment. The first methodology is related to heuristic based approach which utilizes two suggested criteria to efficiently find the victim item and its supporting transactions. The proposed methodology efficiently performs sanitization process especially when overlapping patterns exist in the sensitive item sets.

The proposed modified Inline algorithm efficiently hides the sensitive item sets with minimum side effects by dividing the constraint satisfaction problem into sub constraint satisfaction problems based on the dependency of variables in the constraints with the help of divide and conquer strategy. By solving each sub problem in parallel and individually, the solution for original constraint satisfaction problem is obtained. Experimental results proved that the proposed modified Inline algorithm outperforms heuristic based proposed algorithm in terms of side effects. The modified Inline algorithm not only hides the sensitive item sets but also takes care of not hiding the non sensitive frequent item sets and also avoids generation of new frequent item sets as border revision concepts in the formation of constraint satisfaction problem is considered.

The proposed partition based hybrid hiding methodology efficiently finds the solution for large size database by applying partitioning strategy and parallel concept. Experimental results proved that this methodology outperforms modified Inline and heuristic based approaches in terms of side effects.

The proposed methodology in horizontally partitioned databases with Trusted Party efficiently finds privacy preserving association rule mining by using the adopted sign based secure sum cryptography technique. Another methodology is proposed to handle a situation when no party can be treated as
Trusted Party in the process of finding privacy preserving association rule mining with the help of hash based secure sum cryptography technique. The performance analysis of these two methodologies proved that these are efficient in terms of privacy and communication. The experimental results also proved that they are efficient than the existing considered algorithm.

In case of vertically partitioned databases, the proposed methodology in this work which adopted a scalar product technique, efficiently finds global association rules without revealing individual private data or information. The proposed methodologies in case of two mixed partitioned models efficiently finds global association rules by incorporating tree structure for partitioning.

Each proposed methodology in centralized as well as in distributed environment is analyzed individually but experiments are also conducted on synthetic dataset for comparison analysis purpose.

The proposed methodologies works well for any number of transactions, any number of attributes and for any number of sensitive item sets in case of centralized database environment. The proposed methodologies in distributed database environment efficiently finds global association rules for any number of sites, any size of database with any number of attributes.

**Suggestions For Future Work**

The work carried out in this work may be extended in different ways as follows:

- The methodologies related to heuristic based approach such as randomization, blocking and border based can also be considered as a future work to find privacy preserving association rule mining in centralized database environment.

- The other two cryptography techniques such as secure set union, secure size of set intersection can also be considered to find global association rules in distributed environment by satisfying privacy constraints.

- Privacy preservation in other data mining techniques such as classification and clustering can also be considered as future work in both environments.