CHAPTER 7

CONCLUSION AND FUTURE WORKS

7.1 SUMMARY

The work carried out in this thesis is aimed at the improvement of efficiency of information retrieval system using an integrated approach using various techniques. The effectiveness of the proposed techniques on information sharing through co-operative clients and transparent proxy with cluster based prediction and pre-fetching is demonstrated through a real time implementation. The results obtained during this research have shown that the integrated system is a feasible way to improve the efficiency.

Thus, the highlights of the present investigation include the following

1. The proxy cache manager works in two situations. The offline training phase and online interactive phase.
   a. Offline phase is responsible for training the proxy server through SVM machine learning algorithm.
   b. The intelligent approaches are executed in the online phase. It creates an interactive communication between co-operative clients and also with the proxy server.

2. Clustering based prediction is used in the proxy content. A cluster does not require any prior knowledge of the system. It
simply maintains domain specific clusters in a linked data structure.

3. The developed system supports both the temporal locality and spatial locality. The temporal locality of web objects very much reduces the user perceived latency, decreases server overloading and network congestion. The spatial locality decreases the server under utilization.

4. The performance of the system is compared before and after pre-fetching.

5. Real time implementation of an integrated information retrieval system.

7.2 MAJOR FINDINGS

- In the simulated environment, the user perceived access latency is compared with traditional approach.

- It is observed that the access latency reduces remarkably for subsequent access.

- The same is compared for the number of clients. It shows that, when the count of co-operative clients increases, the access latency gets decreased.

- Web objects are shared among authenticated co-operative clients. Each client is unaware of the objects in the other clients.

- Some of the client level browser objects are moved into SOS, which enhances object sharing among authenticated clients. The object movement is based on user willingness alone. So the developed system has the facility to maintain data privacy.
• SOS also maintains an inter cluster of objects which enhances spatial locality of object access among co operative clients.

• SOS cache can be managed by hybrid based cache management.

• Proxy server maintains clusters, based on the access log of entire group.

• Various web caching; web pre fetching metrics are tested with our system, which shows a performance improvement.

7.4 CONCLUSION

• Object sharing among authenticated group causes the system to improve the hit ratio and decrease the bandwidth usage.

• Data privacy is also maintained. Each client has the facility of secrecy to maintain some of their objects in the browser cache without having to share with others.

• Each client maintains a cluster based on its access behavior, and proxy maintains a cluster based on global behavior of network of clients, which improves the performance.

• Proxy cache is allowed to store only the seed URLs, which avoid duplicate storage of the object.

• Proxy server will maintain all active client lists to avoid accessing a dead node’s objects.
7.5 SCOPE FOR FUTURE WORK

- The developed system has a good performance for static contents objects. Extend the concept for managing dynamic contents.

- Improve this system which includes hiding the identities of the client senders and receivers.

- Instead of downloading the sharable objects from the neighbor client, with the help of virtual memory concept, access the object from the master copy itself.

- A need for content-based web caching algorithm extended to improve for content based network sharing.

- May include suggested query option, and spell check option in query integrator.