An Improvement of The Response Speed by the Reference Prediction on WWW

Yasumitsu Makino

School of Information Science,
Japan Advanced Institute of Science and Technology

February 15, 1999

Keywords: Internet, WWW, Caching Proxy, Acceess Trend, Reference Prediction.

In the Internet, The World Wide Web (here after abbreviated as WWW) has become an essential tool at all. I want to acquire information at high speed on WWW. However, it consists of a complicated network between servers. Therefore, an object cannot necessarily be aquired at high speed. Especially, It is very late if both bandwidth is narrow.

For example, when a big animation file tends to get, information acquisition at the speed exceeding bandwidth is impossible. And when using the reverse-side side of the earth, and the link using the satellite, delay is not avoided however there may be bandwidth. It considers as the method of easing delay of the response speed by the network composition and WWW server load like the above-mentioned example, andthe method of using cache objects, such as caching proxy, is used widely.

By using a cache object, an object can receive at high speed. And it is not influenced network composition or WWW server load. At this time, the response speed on the appearance seen from the user is quick.

However, in a passive cache system, a portion reducible with use of a cache object is the amount of traffic of the object used by overlapping. Although an effect is high if the same object is referred to repeatedly, an effect is low if there is little number of times of reference. An effect is high if the same object is referred to repeatedly. But, an effect is low if there is little number of times of reference.

In this paper, it proposes about cache object directions effective in the improvement in response speed from the tendency of WWW access to the beginning.

The following things were checked as a result of examination.

- The difference of such transmission time exists in the reference frequency for every URL, and object elapsed time at a reusable object.

- In the present cacheing proxy(exsample for Squid), object management cost is applied uniformly,without taking the difference of such transmission time into consideration to the reference frequency for every URL, and object elapsed time.

Consequently, the high speed technique with consideration to such cost by object Elapsed time. And the effective improvement in response speed was obtained. In this paper, We defines object acquision cost. And reference prediction for high speed selection and updating prediction are performed.

- The selection defines the index of acquisition cost every WWW server. The large object of the cache effect is selected as an index using mean transmission time.

- By the reference prediction mechanism, change of the number of times of reference is supervised for every URL, and it compares with qualitative WWW access tendency. From a comparison result, it determines whether to be referred to to a certain URL from now on.

- By the renewal prediction mechanism, the suitable renewal interval of an object is predicted from a past updating history, and the renewal of an original object is reflected in a cache object.

The effect was verified about each three mechanisms above.

And three mechanisms werw combined. The contents of cache can be maintained at the newest by doing so. And a cache object can be used as it is.

Refreshment operation is performed by updating prediction at a suitable interval only to what is actually referred to by reference prediction to that to which the method for maintaining the contents of cache at the newest was limited by the selection for improvement in the speed.

Moreover, it is detectable to hove stopped using a cache object from reference prediction. And as compared with LRU, the object not used can be deleted effectively.

To this proposal system, the simulation of operation was performed the access log of caching proxy to origin, and the opportunity to obtain an object from cache as compared with the conventional caching proxy increased. Consequently, the variation in response speed has been decreased as compared with the conventional caching proxy.

It became clear that it is effective in the improvement in response speed by performing preponderant improvement in the speed as mentioned above by introduction of the selection for improvement in the speed, reference prediction, and an updating prediction mechanism.

When it considered becoming slowly to WWW server with the slowest speed in the case of the information which consists of an object which exists in two or more WWW servers, such as banner, especially, it became clear that the big improvement in response speed is realizable.