

An Agent System for Improving Reliability of WWW Multimedia Repository by Monitoring Its Capacity

Kyung-Hee Lee

Dept. of Computer Science

Chungbuk National University

San 48 Kaeshing-dong Heungduk-ku

Chongju city, Republic of Korea

khlee@trut.chungbuk.ac.kr

Jeong-Hye Han

Statistical Training Center

National Institute of Professional Administration

124 Kajung-dong Yoosung-ku

Taejeon city, Republic of Korea

hanjh@nipa.go.kr

ABSTRACT

The scale of web usage is stressing the capacity of the internet infrastructure and leads to poor performance and low reliability of service. We consider an environment that a WWW multimedia repository is dealing with a certain different resource type such as audio and video data called CM(Continuous Media) better than text and image data. Recently, several techniques have been adopted to reduce the HTTP response time, to improve their reliability and local-balance among multimedia repositories. One way to improve their performances and reliabilities is to replicate popular multimedia resources among different repositories. The replication allows the balance of the clients' requests among different repositories and enables "cost-conscious capability" of multimedia services whereby a surge in a repository load can be handled by dynamically replicating hot multimedia resources on additional repositories.

Most replication selection algorithms aim at selection of the nearby replicas to either reduce response time or the load on network links with the HTTP latency time. However, the HTTP latency time can not sufficiently consider the environment where sizes of the requested resources are large such as CM data, because the replication repository is selected by the shortest average HTTP latency time without considering its excess capacity even when serving a enormous sizes of the requested resources.

This paper focuses on the agent system embedding a new dynamic selection algorithm of the WWW multimedia replication repositories inspired by the need of improving QoS(Quality of Services) of delivering multimedia resources without incurring high access delays. It is called the

capacity algorithm with an additional metric, the dispersion of the sizes of multimedia resources. It must be a metric that considers the capacity of the repository depended on its many parts such as CPU, memory size, OS, network equipment and so on.

We design the agent architecture that consists of three components, Analyzer, Knowledge Base, and Automaton with the capacity algorithm which is able to automatically select to multimedia repositories with good performance and high reliability. The Analyzer collects web-logs of replicated repositories, calculates some figures by monitoring their capacities, and then periodically refreshes them on the Knowledge Base. The Knowledge Base also contains information on the specification of the repositories such as the upper limit of the number of the allowable concurrent users. The Automaton decides which is the best one for the client's request by the capacity algorithm that adapted two metrics, the average HTTP latency time and dispersion of the sizes of the requested multimedia resources.

As the major achievement of our work, we demonstrate that it is possible to consider the HTTP latency time as well as the additional metric, PCI(Process Capability Index) C_p , containing dispersion of the sizes of multimedia resources. It is obtained by very simple numerical calculation under some assumptions. The capacity algorithm with PCI gives small selection probability for a replicated repository with large variation of size of the requested multimedia resources, and it gives large one for otherwise.

We will give the quantitative comparisons with the other algorithms for our further study. Moreover, we believe that it is extremely important to estimate the HTTP latency average time and the parameter of the characteristic. So this issue is a subject of our further work with the percentile method, some bootstrap methods, and EWMA(Exponentially Weighted Moving Average) model.

RESUME

- Kyung-Hee Lee received the B.S. degree and the M.S. degree in Computer Science from Chungbuk National Univ. in 1995, 1999, respectively. Since 2000 she is preparing her doctoral thesis in Computer Science from Chungbuk National Univ. and serves as a extraordinary professor of Computer Science Dept. of Keukdong College. Her research interests include analysis of algorithms, computer network, and XML.
- Jeong-Hye Han received the B.S. degree in Statistics and the Ph.D. degree in Computer Science from Chungbuk National Univ. in 1992, 1998, respectively. From 1998 to 1999, she served as a post-doctorial researcher at Industrial System Engineering Dept. of Yonsei Univ. in Seoul. From July 1999 to 2001, she was a professor in the Statistical Training Center of National Institute of Professional Administration. In March 2001, she joined Dept. of Computer Education of Chongju National University of Education where currently she is a professor. Her research interests include network management, Human-Computer Interaction, Electronic Commerce, and Data Mining.