International Journal of Innovative Computing, Information and Control
Volume **5**, Number **3**, March **2009**

 ISSN 1349-4198
pp. 821–828

ANALYSIS OF AN EFFICIENT PRE-FETCHING ALGORITHM FOR SCALABLE CONTENTS OVER CONTENT DELIVERY NETWORKS

Zhou Su, Jiro Katto and Yasuhiko Yasuda

School of Science and Engineering
Waseda University
3-4-1 Okubo Shinjuku-ku Tokyo 169-8555, Japan
zhousu@asagi.waseda.jp

Received January 2008; revised June 2008

Abstract. *Recently it has been shown that scalable (layered) image is appropriate for the Internet because of its better flexibility and functionality. At the same time, the pre-fetching technology also proves to be an efficient method to improve the performance of image-delivery with the low delay and adjustable quality. However, how to deploy the pre-fetching to deliver scalable contents has not been mentioned. Therefore, this paper presents an analysis of the pre-fetching algorithm to deliver layered image contents over the Internet. Based on the analysis of web request probability and the relations between different layers in the scalable imaging delivering system, a new pre-fetching algorithm is proposed to drastically improve cache performance. This proposal is a request frequency-based approach, in which a progressive image format can be used.*
Keywords: Contents delivery, Hierarchical image coding, Pre-fetching, Web caching

1. **Introduction.** To deliver the multimedia contents over the current network [18-25], caching systems are commonly introduced to reduce data traffic and improve response time in the Internet, because the data is cached temporarily in the intermediate server and utilized many times. However, current system employs a "hard" caching strategy: an image is stored in a cache or not even if its data is quite big.

Scalable (layered) coding is playing an important role to overcome the above problem. Scalable image format has been widely used since the first hierarchical coding method was proposed [1]. Many other kinds of hierarchical formats were also designed such as pyramids [2], wavelets [3] and sub-band coding [4].

Researchers and engineers have argued that scalable format is appropriate for the Internet because of its better flexibility and functionality [5-10]. Basically, scalable format assumes scalable coding, in which the original signal is coded into several layers, from the lowest (base) to highest (enhancement) layers. This scalable coding is supported by some image compression standards, such as *JPEG2000*. Then, the scalable image can transmit each layer over different channels (different protocols, different error correction codes, or different paths) [12,13,15-17].

In this paper, therefore, to decide which layer of which image should be pre-fetched in an image delivery network, we firstly quantitatively clarifies on the probability analysis of web request. Then, we assign different lowest ranking of images to be pre-fetched (T_k) adaptively when the user is asking for a different layer (kth) of the image. Next, we derive a relationship, $T_{K+1} = T_K \cdot \sqrt[\alpha]{P}$, where α is a parameter of Zipf-like distribution and p, called Greedy Degree, is a probability that the user still wants to get the $(j+1)$th layer after having got the jth layer of this image. Finally, according to the above analyses, we propose a request frequency-based pre-fetching algorithm, which chooses the pre-fetch-ranking