

Design of a Near-Minimal Dynamic Perfect Hash Function on Embedded Device

Derek Pao, Xing Wang and Ziyang Lu

*Department of Electronic Engineering,
City University of Hong Kong, HONG KONG*

E-mail: d.pao@cityu.edu.hk, {xingwang4, ziyanglu2}@student.cityu.edu.hk

Abstract— There has been a general opinion that it is difficult to construct perfect hash tables with high load factor for large datasets having a million records. The problem is even more challenging if new records can be added to the hash table incrementally. In this article, we shall demonstrate the design of a dynamic perfect hash function on embedded device based on simple bit-shuffle and bit-extraction operations. The achievable load factor can be up to 100%, and the amortized memory cost of the hash function is about 7 to 15 bits per key for 32-bit keys. Incremental updates to the hash table are allowed. The perfect hash function for a dataset with 1 million keys can be constructed in a few seconds of CPU time.

Keywords—Dynamic Perfect Hash Table, Embedded System, Pipelined Architecture, Searching.



Derek Pao received the B.Sc.(Eng.) degree in Electrical Engineering from the University of Hong Kong, and the M. Comp. Sc. and Ph.D. degree in Computer Science from Concordia University, Canada. He is an Associate Professor with the Electronic Engineering Department, City University of Hong Kong. His research interests include architectures for network processing, and high speed pattern matching for network and system security.



Xing Wang received his M.Sc. degree from Peking University in 2007. He is currently an Engineer in Peking University Shenzhen Graduate School. His research interests include network security, architecture for network processing and computer network.



Ziyang Lu received the Bachelor of Engineering in Electronic Engineering and Information Science from the University of Science and Technology of China in 2009. She is working towards the PhD in Electronic Engineering at the City University of Hong Kong. Her current research interest is IP address lookup.