Department of Computer Science
香港城市大學
City University of Hong Kong

COMPUTER SCIENCE COLLOQUIUM

DyTIS: A Dynamic Dataset Targeted Index Structure Simultaneously Efficient for Search, Insert, and Scan

SPEAKER **Prof. Sam Noh**

Professor
Department of Computer Science,
Virginia Tech, USA

DATE 24 May, 2023 (Wed)
TIME 3:30 PM - 4:30 PM
VENUE Room 5-206, 5/F., Lau Ming Wai Academic Building, City University of Hong Kong, 83 Tat Chee Avenue, Kowloon Tong, Hong Kong

ABSTRACT

Many datasets in real life are complex and dynamic, that is, their key densities are varied over the whole key space and their key distributions change over time. It is challenging for an index structure to efficiently support all key operations for data management, in particular, search, insert, and scan, for such dynamic datasets. In this paper, we present DyTIS (Dynamic dataset Targeted Index Structure), an index that targets dynamic datasets. DyTIS, though based on the structure of Extendible hashing, leverages the CDF of the key distribution of a dataset, and learns and adjusts its structure as the dataset grows. The key novelty behind DyTIS is to group keys by the natural key order and maintain keys in sorted order in each bucket to support scan operations within a hash index. We also define what we refer to as a dynamic dataset and propose a means to quantify its dynamic characteristics. Our experimental results show that DyTIS provides higher performance than the state-of-the-art learned index for the dynamic datasets considered.

BIOGRAPHY

Sam H.(Hyuk) Noh received the BE degree in computer engineering from the Seoul National University, Seoul, Korea, in 1986, and the PhD degree from the Department of Computer Science, University of Maryland, College Park, MD, in 1993. He held a visiting faculty position at the George Washington University, Washington, DC, from 1993 to 1994 before joining Hongik University, in Seoul, Korea, where he was a professor in the School of Computer and Information Engineering until the Spring of 2015. During this period, he served as the Chair of the Department of Computer Engineering as well as the Head of the School from September of 2013 through February of 2015. From August 2001 to August 2002, he was also a visiting associate professor with the University of Maryland Institute of Advanced Computer Studies (UMIACS), College Park, MD. Starting from the Fall of 2015 he joined UNIST (Ulsan National Institute of Science and Technology), a young science and tech focused national university, where he was a Professor at the Department of Computer Science and Engineering and served as the inaugural Dean of the Graduate School of Artificial Intelligence in the College of Information and Biotechnology from August 2020 through March 2023. He also served as the Dean of the School of Electrical and Computer Engineering from January of 2016 through June of 2018. As of January 2023, he is a Professor at the Computer Science Department at Virginia Tech. He has served/serves as General Chair, Program Chair, and Program Committee Member on a number of technical conferences and workshops including ACM ASPLOS (2024, 2018 (ERC), 2016), ACM SOSP (2023, 2019), ACM Eurosys (2023, 2020, 2018), USENIX FAST (2024, 2023, 2020 (co-chair), 2019, 2018, 2017, 2016, 2105, 2014, 2013), ACM Systor (2023, 2022, 2018, 2017), USENIX OSDI (2022, 2016), IEEE RTAS (2022, 2019), USENIX ATC (2021, 2019), ACM EMSOFT (2021, 2020, 2019, 2018), ACM/USENIX HotStorage (2021, 2020, 2018, 2016 (co-chair), 2015), ACM LCTES (2015 General Chair), IEEE ICPADS (2013 subtrack co-chair), and WWW (2003), among others. He also served on the Steering Committee of LCTES from 2016 through 2020. He is currently the Chair of the Steering Committee for ACM HotStorage and a Steering Committee member of USENIX FAST and IEEE NVMSA. He also served as Editor-in-Chief of the ACM Transactions on Storage from 2016 through 2022. His research interests include system software issues pertaining to computer systems in general and storage systems in particular, with a focus on the use of new memory technologies such as flash memory and persistent memory. He is a Fellow of the ACM and IEEE and a member of USENIX and KIISE (Korean Institute of Information Scientists and Engineers).

All are welcome!

In case of questions, please contact Prof. Jason Xue at jasonxue@cityu.edu.hk, or visit the CS Departmental Seminar Web at https://www.cs.cityu.edu.hk/events/cs-seminars/recent-cs-colloquiums.