DUO LU

691 North Road, B.C., Canada | 778-223-4264 | luduo1999@126.com

https://github.com/Duo-Lu

EDUCATION

Simon Fraser University, Burnaby, B.C., Canada   *Jan 2019 - May 2022*
- Bachelor of Science - **Computer Science** (CGPA: 3.98 / 4.33)
- Courses: Database Systems (grad-level); Distributed and Cloud Systems (grad-level); Data Mining; Information Retrieval and Web Search; Operating Systems; Multimedia Systems; Intelligent Systems; Computer systems

AWARDS

SFU Undergraduate Open Scholarship
SFU Dean's & President's Honour Roll
SFU Faculty of Applied Science Alumni Scholarship
Best Poster Award at Inaugural CS Undergraduate Research Symposium
SFU VP Research - Undergraduate Student Research Award - Science ($9000+, Supervisor: Dr. Jiguo Cao)

PUBLICATION

Evaluating Persistent Memory Range Indexes: Part Two.
Yuliang He, <u>Duo Lu</u>, Kaisong Huang, Tianzheng Wang
48th International Conference on Very Large Data Bases (VLDB 2022)

RESEARCH EXPERIENCE

Undergraduate Research Assistant (Supervisor: Dr. Tianzheng Wang)   *Jan 2021 - May 2022*
Implementation and Evaluation of FPTree (Fingerprinting Persistent Tree) [code]   *SFU*
- Utilized **C++** to implement and evaluate **multi-threaded** B^+-tree based index called FPTree according to a paper from SIGMOD that builds on top of the next-generation hardware, Intel Optane DCPMM
- Optimized operations throughput by issuing cache line alignment and prefetch to **reduce cache miss**, SIMD instructions (Intel AVX512) on hashed keys to facilitating search
- Employed Hardware Transactional Memory (Intel TSX proposed in paper) as the **concurrency control** mechanism to simplify the synchronization without sacrificing scalability, and debug intensively to solve deadlock issues
- Implemented **crash recovery** that the whole index could be rebuilt quickly from system restart and power failure
- Contacted the original author to verify the performance and implementation details, open-sourced and used by several top conference papers as the baseline index

A closer look at synchronization protocols on FPTree [code]   *SFU*
- Implemented two alternative **synchronization** methods (Lock Coupling and Optimistic Lock Coupling) to achieve higher, more stable throughput, especially under high contention workloads (skewed insert, scan)

Evaluating Persistent Memory Range Indexes: Part Two [paper] [code]   *VLDB 2022*
- Evaluated 5 state-of-the-art persistent indexes using a unified benchmark framework (PiBench) under uniform, skewed, mixed, variable length key, latency, NUMA workloads with varying numbers of threads
- Incorporated persistent allocator (PMDK) support into each index to ensure fairness and reliability
- Profiled benchmark results using **Linux Perf** through CPUs utility, memory statistics, system calls and bandwidth usages to identify bottlenecks and inefficiency inside index implementations and utilized this information to verify the performance of the proposed design in each index
- Modified the benchmark framework and collected real-world datasets (URLs, names, article titles) to compare persistent indexes performance on **variable length keys** according to the paper revision advice

Undergraduate Research Assistant (Supervisor: Dr. Martin Ester)   *May 2020 - May 2021*
Enhance feature representations for Cross-domain Few-shot Learning   *SFU*
- Improved the model benchmark performance from one of the popular papers Selecting from Universal Representations (SUR) in the Cross-domain Few-shot Classification area by a more flexible feature selection strategy
- Applied feature selection on the entire set of feature vectors to increase the flexibility of feature representation
- Contacted the original author to discuss issues in the paper, reproduced results, and improved benchmark results

WORK EXPERIENCE

Embedded Software Developer   *May 2022 - Present*
Fortinet, Full-time Permanent (FortiOS, router team)   *Vancouver*

- Maintained and resolved bugs for the open-sourced **operating system** (FortiOS) written by **C** and primarily focused on enterprise dynamic routing protocols such as BGP, OSPF, RIP, ISIS
- Implemented features such as Virtual Routing and Forwarding (VRF) and Bidirectional Forwarding Detection (BFD) support on static routes via IPSec VPN tunnel interfaces
- Modified the Linux kernel to handle the networking traffic of new features from the user space

Data Scientist Intern (Supervisor: Dr. Markus Roggon)   *Aug 2020 - Dec 2020*
DELIC Labs, Full-time Co-op [website]   *Vancouver*

- Built statistical regression models to optimize cannabis extraction processes by identifying the best set of conditions (temperature, pressure, and flow-rate) to maximize throughput
- Performed data cleaning and imputing by eliminating data with extreme pressure or temperature and inserting missing extraction data points with KNN, to get a more smooth and useful dataset for modelling
- Derived significant features such as density and residence time of $CO2$ (extraction medium) through self-learnt chemistry and cannabis extraction domain knowledge
- Discovered the optimal cannabis extraction conditions by building polynomial/logistic regressions models to fit raw extraction conditions, derived features and cannabinoid concentrations throughput
- Utilized Matplotlib and Plotly generating 2D/3D figures, and box plots to visualize the relationship between the extractor machine conditions and output concentrations especially when they reached maximum output

TEACHING ASSISTANT

Operating Systems I
SFU CMPT 300 | Spring 2021 | with Dr. Tianzheng Wang

Introduction to Computer Systems
SFU CMPT 295 | Summer 2021 | with Dr. Arrvindh Shriraman

Special Topics in Computing Systems (Software validation/testing in Rust)
SFU CMPT 479/982 | Fall 2021 | with Dr. Steven Ko

TALK & POSTER

Evaluating Persistent Memory Range Indexes: Part Two   *Apr. 2022*
Yuliang He*, Duo Lu*, Kaisong Huang, Tianzheng Wang (*Presenters)
Presented at:
Inaugural CS Undergraduate Research Symposium (Poster)
SFU Faculty of Applied Sciences, School of Computing Science

Selecting from Universal Representations for Cross-domain Few-shot Learning   *Mar. 2021*
Duo Lu, Haoyu Peng
Presented at:
Special Research Project, ESTER LAB
SFU Faculty of Applied Sciences, School of Computing Science

SERVICE

Peer Mentor   *Sept. 2021 - Dec. 2021*
SFU Centre for Accessible Learning   *Vancouver*

- Instructed a complete vision loss (blindness) female student to study Computer Systems and Data Structures
- Awarded $2600+ funding by StudentAidBC, Ministry of Advanced Education

SKILLS

Languages: English (proficient), Mandarin (native)
Programming Languages: C/C++, Python, Java, JavaScript, SQL, Rust, LaTeX
Libraries: Numpy, Pandas, Matplotlib, Scikit-learn, Pthreads, Intel oneTBB, PMDK, React.js, Node.js, Mongoose
Tools: Android Studio, Git, Jira, MongoDB, Studio 3T, ParaView, AWS, Docker, Kubernetes