

Yilin Feng

The Pennsylvania State University
ypf5071@psu.edu

EDUCATION

The Pennsylvania State University, U.S.

Aug. 2019 - Present

Ph.D. in Computer Science, advisor: Dr. Kandemir Mahmut and Dr. Chita Das, GPA: 3.52/4.0

Nanjing University, Nanjing, China

Sept. 2015 - Mar. 2019

Master's degree in Computer Science and Technology, advisor: Dr. Junyuan Xie, GPA: 86.10/100

China University of Mining and Technology, Xuzhou, China

Sept. 2011- Jul. 2015

Bachelor's degree in Computer Science and Technology, GPA: 91.36/100, Ranking: 2/148

ACADEMIC RESEARCH

FPGA-optimized Adaptive Banded Event Alignment

Oct. 2022 – Present

Advisor: Dr. Kandemir Mahmut, Distinguished Professor at Department of Computer Science and Engineering, Penn State

Adaptive Banded Event Alignment (ABEA) is a crucial algorithmic component of sequence polishing, aligning events from raw signals to k-mers of base-called reads via a dynamic programming strategy.

- Propose a fine-grained two-level parallel ABEA algorithm based on the analysis of data independence in consecutive iterations in ABEA, achieving cell-level computation parallelism within a band and band-level computation parallelism among multiple reads within a FPGA kernel function.
- Adapt data structures including reads, events, k-mer models to streams for communication, forming dataflow pattern and achieving task-level parallelism.
- Divide reads into multiple partitions according to the number of bands and transfer each partition to the corresponding top function on FPGA to perform alignment, achieving load balance on FPGA and heterogeneous processing between CPU and FPGA.
- Made an estimation our approach can achieve comparable performance with GPU-implementation with less energy consumption.

GPU-accelerated *Nanopolish* Call-methylation Module

Apr. 2020 - Jun. 2022

Advisor: Dr. Kandemir Mahmut

Nanopolish is a software package for signal-level analysis of Oxford Nanopore sequencing data. Call-methylation module of *Nanopolish* can detect methylation based on Hidden Markov Model(HMM). This module consists of four steps: event detection, alignment, final calibration, and the methylation score calculation.

- Parallelized and optimized the methylation score calculation step on GPU. The score matrix is divided into multiple blocks and blocks are processed by threads in parallel, solving data dependence problems and achieving load balance among threads.
- Pipelined the four steps of this module to increase the concurrency of CPU and GPU as well as hardware resource utilization.
- Experimental results show our work achieves good speedup with previous work. Our GPU implementation of methylation score calculation outperforms the CPU implementation of that by -2x. On average, our proposed approach can achieve 3x-5x speedup compared to *Nanopolish*, and reduce the total execution time by 35% compared to the state-of-the-art work *f5c*, which proposed parallel alignment execution.

Hybrid CUDA Solution for AllPairs Similarity Search on High-dimensional Sparse Datasets

Nov. 2017 - Jul. 2018

Advisor: Dr. Jie Tang, Associate Professor at Department of Computer Science and Technology, Nanjing University

- Developed a Parallel Cosine Similarity Search Algorithm on Word2Vec in CUDA environment, combining a feature-parallel scan and a pair-parallel scan on different parts of vectors, trading off between memory accessing and dot product computing.
- Proposed a parameter tuning method for a user-defined p parameter for best performance. The method avoids a large number of atomic operations, considering statistical characteristics of high-dimensional sparse datasets.
- Experimental results show our algorithm could solve the problem much faster than existing methods on several benchmark

datasets, and achieve the speedup of 14x-85x against cuSPARSE and 1.5x-23x against the state-of-the-art parallel algorithm, while maintaining a relatively stable running time with different values of the threshold.

LANGUAGES

- C++, CUDA, JAVA

PUBLICATION

- **Yilin Feng, Zheyu Li**, Gulsum Gudukbay Akbulut, Vijaykrishnan Narayanan, Mahmut Kandemir, Chita R. Das. FPGA Implementation of Adaptive Band Event Alignment using High-Level Synthesis. In *Workshop on ACCELERATOR ARCHITECTURE IN COMPUTATIONAL BIOLOGY AND BIOINFORMATICS, in conjunction with 50th IEEE International Symposium on Computer Architecture*, 2023.
- **Yilin Feng**, Gulsum Gudukbay Akbulut, Xulong Tang, Jashwant Raj Gunasekaran, Amatur Rahman, Paul Medvedev, Mahmut Kandemir. GPU-Accelerated and Pipelined Methylation Calling. *Bioinformatics Advances*. 2022;2(1):vbac088.
Link: <https://doi.org/10.1093/bioadv/vbac088>
Github repo: <https://github.com/fengyilin118/Galaxy-methyl>
- **Yilin Feng**, Jie Tang, Chongjun Wang and Junyuan Xie. CuAPSS: A Hybrid CUDA Solution for AllPairs Similarity Search. In *International Conference on Algorithms and Architectures for Parallel Processing*, pp. 421-436. Springer, Cham, 2018.
Link: https://link.springer.com/chapter/10.1007/978-3-030-05051-1_29
- **Yilin Feng**, Jie Tang, Meilin Liu, Chongjun Wang and Junyuan Xie. Fast Document Cosine Similarity Self-Join on GPUs. In *2018 IEEE 30th International Conference on Tools with Artificial Intelligence (ICTAI)*, pp. 205-212. IEEE, 2018.
Link: <https://ieeexplore.ieee.org/document/8576038/>

AWARDS

- National Scholarship, Ministry of Education of the People's Republic of China (Top 2%) Nov. 2014 and Nov. 2012
- Outstanding Graduate Award, China University of Mining and Technology Jul. 2015
- First-class Academic Scholarship, Nanjing University Nov. 2015