Hardware-Accelerated Time Series Data Mining Algorithms

SPEAKER
Dr Philip BRISK
Associate Professor
Department of Computer Science and Engineering
Bourns College of Engineering
University of California, Riverside
USA

DATE
13 July 2018 (Friday), 12 July 2018 (Thursday)

TIME
10:30 am - 11:30 am, 3:15 pm – 4:15 pm

VENUE
B5-114, 5th Floor, Blue Zone
CSE Conference Room B6605
6th Floor, Blue Zone
Yeung Kin Man Academic Building
City University of Hong Kong
83 Tat Chee Avenue
Kowloon Tong

ABSTRACT
An immense amount of today’s digital data takes the form of time series. In almost every scientific field, measurements are performed over time. With the emergence of the Internet of Things (IoT), billions of sensors worldwide produce information in the form of time series which is stored (and often ignored) in data centers worldwide. Wearable health devices, including but not limited to FitBits, smart watches, and medical monitors, produce time series data relevant to human health; often, this data must be acted upon immediately by a networking and cloud infrastructure that was not designed for critical real-time applications. It is quite likely that computing as a public utility will be needed to manage emerging fleets of self-driving automobiles, and much of the sensory data that is processed will once again take the form of time series.

This talk will describe recent and ongoing efforts to accelerate the way that time-series data is processed:

The first part of the talk will consider the Dynamic Time Warping (DTW) similarity measure for time series in the context of real-time health monitoring. At its core, DTW is a quadratic-time dynamic programming algorithm with origins in speech recognition, and is amenable to acceleration as dedicated systolic array in hardware; in the real-time monitoring context, DTW is used as a thresholding mechanism to detect the presence or absence of an anomalous event, which may require immediate medical attention. Here, DTW is combined with normalization (a necessity) and software optimizations which can rapidly discern the absence of an anomalous event without completing the dynamic programming task. The resulting algorithm has irregular control flow which is highly data-dependent, making it difficult to parallelize effectively. To improve performance and energy efficiency, we developed an application-specific processor that features custom instructions to accelerate both the control and data-dominated portions of the algorithm; to reduce cost, our solution eschews systolic acceleration. Further improvements in performance and energy efficiency were obtained by replacing floating-point operators with logarithmic arithmetic.

The second part of the talk will introduce the Matrix Profile, a data structure that annotates a time series, and an algorithm to rapidly compute the Matrix Profile via GPU acceleration. Using the Matrix Profile many traditionally hard problems in time series data mining become trivial or easy, such as motif discovery, clustering, etc. The Matrix Profile is exact, parameter free, space-efficient, is incrementally maintainable, allows anytime algorithms, does not require user-specified thresholds, can handle missing data, and can be constructed in deterministic time. This talk will describe two algorithms to compute the matrix profile: STAMP, which is an anytime algorithm, and STOMP, which is not anytime, but has a lower time complexity and is also amenable to GPU acceleration.

Collectively, the two parts of the talk both justify the need for hardware acceleration of time series data mining algorithms and show how the need for hardware acceleration ranges from low-cost microcontroller-class processors to the data center scale.

BIOGRAPHY
Philip Brisk received, the BS, MS, and PhD Degrees, all in Computer Science, from UCLA in 2002, 2003, and 2006 respectively. From 2006-2009 he was a Postdoctoral Scholar at EPFL in Switzerland. He has been with UC Riverside since 2009. His research interests include programmable microfluidics and lab-on-a-chip technology, FPGAs and reconfigurable computing, and other forward-looking applications of computer engineering principles.

All are welcome!