BIOGRAPHICAL SKETCH

Md Mosharaf Hossain was born in Rajbari, Bangladesh. He received his Bachelor of Science Degree in Computer Science & Engineering from Bangladesh University of Engineering & Technology in 2009. Later, he got opportunity to work in Samsung Bangladesh R&D Center in November, 2010. In March 2012, he joined at the Business Intelligence Department in Grameenphone Ltd., and worked there nearly four years. Throughout his exciting professional career, he had experiences in application development, databases, data mining, and machine learning. He entered in the graduate program at the Tennessee Tech University in Fall, 2016 and expecting to complete his Master's degree in Computer Science in December 2018. His research interests include High Performance Computing, Big Data, and Machine Learning.

EDUCATION

M.S., Computer Science
Tennessee Technological University

B.S., Computer Science and Engineering
Bangladesh University of Engineering & Technology

College of Engineering

TENNESSEE TECH

The Department of

Computer Science

Announces the Thesis Defense

Of

Md Mosharaf Hossain

In Partial Fulfillment of the Requirements

For the degree of

Master's of Science in Computer Science

1st November, 2018 at 9 am

Held at

Bruner Hall, Room 406

University Drive

Tennessee Technological University

Cookeville, TN, 38505

FIELD OF STUDY

Computer Science

THESIS TOPIC

Flexible-Block Partitioning for Parallel Matrix Multiplication Routines in Shared Memory Environment for the Edge Case Matrices

EXAMINING COMMITTEE

Dr. Sheikh Ghafoor (Chairperson)

Dr. Mike Rogers (Committee member)

Dr. Doug Talbert (Committee member)

ABSTRACT

Linear algebra operations are at the heart of many applied domains such as machine learning, natural language processing, computer vision, and engineering simulations. Basic Linear Algebra Subprograms (BLAS) prescribes some routines to perform these operations. Matrix-matrix operations are categorized as Level-3 operations by the BLAS specification. However, two commonly used matrix-matrix operations are General Matrix-Matrix multiplication (GEMM) and Symmetric Rank-k update (SYRK). The SYRK routine is a specialization of the GEMM routine, where some multiplications are skipped as the resultant matrix is known to be symmetric. In this thesis, we first present a comprehensive analysis of the SYRK routine under popular dense linear algebra libraries such as OpenBLAS, Intel MKL, and BLIS, particularly focusing on edge cases of dense matrices (strictly thin or strictly fat shapes) in shared memory machine. The primary reason we concentrate on the edge cases is that the traditional linear algebra libraries are highly optimized for squarish matrices. Therefore, it presents a big opportunity to improve scalability of the routines in multicore environment when dealing with the edge case matrices. Our work identifies noticeable performance drop of the SYRK routine for the edge case matrices as we increase threads to a considerable number. Then, we propose new blocking approach, we name it Flexible-blocking. In contrast to the contemporary libraries, our approach formulates the blocks of the input matrices based on the shapes of the matrices as well as number of threads used in the implementation. Our proposed technique shows noticeable performance improvement on multicore shared-memory machines for the edge case matrices.