Invited Talk

Department of Computer Science and Engineering

Indian Institute of Technology Kanpur

Venue: RM 101

Time: 11:30 am -12:30 pm, March 18, 2019

Low rank approximation of Wasserstein distance kernel for malware detection*

Amir Averbuch

Department of Applied Mathematics, School of Mathematical Sciences, Tel Aviv University
School of Computer Science, Tel Aviv University

The goal of this work is to detect malware (anomaly) in an unsupervised way. The algorithm exhibits a high detection rate and a low false positive rate where the input data, which is called the training data, is arranged as a kernel matrix. The distances among the kernel entries are measured as Wasserstein distances also called Earth Mover's Distance. The malware detection is based on embedding the kernel matrix into a lower dimensional space represented by a manifold. Then, a newly arrived multidimensional data point, which did not participate in the training data and is embedded into the lower dimensional space, is classified as normal or anomaly if it lies in the manifold or deviate from it, respectively. The embedded space is determined by the eigenvalues and eigenvectors of the kernel matrix. Kernel matrices are huge since their sizes are determined by the size of the training data. The performance of the embedding is determined by the decay of the eigenvalues of the kernel matrix which is usually huge. The focus in this paper is to use numerical linear algebra with randomized algorithms for kernel computation by SVD decomposition. Therefore, this paper describes several methods, which are based on SVD-based low rank matrix decomposition, to approximate efficiently the kernel. In addition, inherent parallelism in the algorithms are uncovered and they enable to port the computation into several GPUs to achieve substantial speedup. The outcomes of the algorithms are presented as ROC graphs that show their dependencies between true positive and false positive rates. The performances of the algorithms in this paper are evaluated from the applications of anomalies (malwares) detections algorithms to monitored processes in a computer.

*** Work done with Tatiana Osokin, Gil Shabat**

Bio: Amir Averbuch is well known for his research in applied and computational harmonic analysis, Big Data processing and analysis, wavelets, signal/image processing and scientific computing. At the Tel Aviv University, Prof. Averbuch has supervised fourteen post-doctoral researchers, as well as more than 30 PhD students and 100 MSc students. He is a founder of Thetaray, and before joining Tel Aviv University, he was a research staff member with IBM's TJ Watson Research Center in New York.