Big Data Research Using NJIT Cloud Computing Systems

Professor/Supervisor: Dr. Jason T. L. Wang
Student: Yasser Abduallah
Department: Computer Science
Date: Wednesday, January 11, 2017

Abstract:

Gene regulation is a series of processes that control gene expression and its extent. The connections among genes and their regulatory molecules, usually transcription factors, and a descriptive model of such connections, are known as gene regulatory networks (GRNs). Elucidating GRNs is crucial to understand the inner workings of the cell and the complexity of gene interactions. To date, numerous algorithms have been developed to infer gene regulatory networks. However, as the number of identified genes increases and the complexity of their interactions is uncovered, networks and their regulatory mechanisms become cumbersome to test. Furthermore, prodding through experimental results requires an enormous amount of computation, resulting in slow data processing. Therefore, new approaches are needed to expeditiously analyze copious amounts of experimental data resulting from cellular GRNs. To meet this need, cloud computing is promising as reported in the literature. Here we discuss our experience of using the cloud computing platforms available at NJIT to perform GRN inference in cells. We also present some recent results of comparing Hadoop and Spark in the cloud computing platforms at NJIT.