Masters Final Oral Defense

Friday, April 13, 2018
1181 Gilman at 3:00 PM

Shruti Sahu

Tensor Decomposition and Filling in Missing Data

The main purpose of this project is to determine the minimum number of simulations that can achieve the same results as the full 1024 simulations produce. There are 2 main steps involved to achieve this: first, to perform tensor decomposition on the initial 5-dimensional tensor. A tensor is nothing but a n-dimensional array. Tensor-decomposition can be described as feature extraction that helps us to interpret the data in a clearer fashion. The type of tensor decomposition explored is 'Tucker Decomposition'. The open-source python package called 'Tensorly' has been used to carry out the experiments for tensor decomposition. This decomposition breaks the tensor into a core tensor and a set of matrices.

After tensor decomposition is done, data of only few simulations will be picked as data samples in order to construct a 'gappy dataset'. Gappy data refers to a dataset where some data points are missing. The second step consists of filling in the missing data in the gappy dataset. There are several many ways to fill in gappy data. Data interpolation and data compression are few methods to fill in missing data. In this project, these methods will be explored in order to determine the best way to fill in the missing data so the data filled in by the algorithm match with the original data that was collected as output of the simulations.

Major Professor:
Baskar Ganapathysubramanian
Samik Basu

Committee Members:
Soumik Sarkar

IOWA STATE UNIVERSITY
Department of Computer Science
www.cs.iastate.edu