INDRAPRASTHA INSTITUTE of INFORMATION TECHNOLOGY DELHI

Invitation to MTech Thesis Defense of Divyanshu Srivastava July 14, 2018 (Saturday): 03.00-04.30 IST

In Partial Fulfillment of the Requirements for the Degree of
M.Tech CB

Divyanshu Srivastava (MT16125)

Will defend his thesis

Title: "Graph Signal Processing based analysis of biological networks"

IIIT-D Faculty and Students are invited

Date: July 14th 2018 (Saturday)
Time: 03.00-04.30 IST
Place: Meeting Room, 3rd Floor(NAB)

Examiner:	**Internal:**	**Debarka Sengupta**
	External/Internal:	**K K Chtaurvedi (CABin)**
	Advisor:	**Vibhor Kumar**

****

Abstract
Complex biological network analysis is a widely studied topic in computational biology. This thesis tries to investigate two most common and widely studied complex biological networks, the Genetic Interactions Networks and Protein Residue Contact Networks from a fairly recent approach of Graph Signal Processing (GSP). With signal processing techniques applied over complex biological networks, not only the traditional network analysis techniques are seen to improve, much deeper insights about the network are also witnessed. These insights about the biological mechanisms, structures, modules and interactions within the network tend to give an explanation of the emergent behaviour of such complex systems. Here, graph signal processing has been applied on genetic interaction networks for prokaryotic and eukaryotic model organisms. Microarray-based gene expression information from the DREAM5 Network Inference Challenge datasets are modelled into interaction networks using existing methods. GSP is then used to filter out genuine interactions from meaningful and noisy experimental data. Novel expression filtering techniques, along with different ways to extract most informative segments from the spectrum of interaction graphs has enabled us to improve the predictions made by standard off-the-shelf statistical methods, as well as more sophisticated approaches like ARACNE. Filtered networks are seen to predict genuine interactions matching with the golden standards with up to 2-4 times higher than raw networks. In a separate experiment in this thesis, GSP based methods have also been shown to discern relevant segment of signals on Protein residue Contact Networks (PCN). This relevant segment is further verified to show correlation with biophysical properties of the protein, namely its folding rate. A set of 52 single domain two-state folding proteins with their folding rate information is taken, and different PCN models are generated using their three-dimensional structure information. A multivariable linear regression model has been used here to estimate the the collective role played by different signals on protein folding. It has been shown here in absence of known structure that GSP based method can also be used on residue co-evolution network of proteins to model their folding rate. The achieved accuracy for modelling folding rate was between 0.55 - 0.8. With these two experiments, we have shown that modern signal processing technique can be used on biological data sets to yield meaningful results. Keywords: Graph Signal Processing, Gene Interaction Network, Protein Contact Networks.