Abstract

This thesis is devoted to the study of biological networks at the sub-cellular level. Network inference is one of the branches of systems biology that aims to contribute in this direction. It focuses on the inference of interactions between the molecular species in biological networks. Computational approaches are employed to retrieve these interactions from the molecular interaction data. The objective of this thesis is to develop mathematical learning approaches for network inference and contribute to the understanding of biological subsystems.

We can classify the work done in this thesis along three dimensions. The first is biological networks studied. We particularly looked at signalling networks and gene regulatory networks. The second is the organisms studied. We looked at signalling pathways and expression data sets of organisms such as *E. coli*, *S. cerevisiae* and *A. thaliana*. The third is computational models used for the network inference task. All work done in this thesis revolves around transition-based models (specifically Petri nets).

Though a large amount of biological data is available, mostly, this data is sparse, noisy, and limited, which creates a challenge in the network inference procedure. Probabilistic extensions of the transition-based approaches are developed to tackle these challenges. The probabilistic formulation of the approaches tries to model the inherent stochastic behaviour of the biological subsystem and includes a provision for the incorporation of domain knowledge that helps in dealing with the limited data problem. To model the nature of interactions in the two subsystems, i.e., signalling networks and gene regulatory networks, we proposed two network reconstruction approaches, Probabilistic Logic-Guarded Transition System (PLGTS) and Probabilistic Extended Petri Net for Gene Regulatory Networks (PEPN-GRN), respectively for each subsystem. We also tried to fit the probabilistic formulation of Petri nets in the family of graphical models and proved their equivalence with the Dynamic Bayesian Networks for a particular class of transition systems that exhibit the regulation mechanism.

The ultimate goal of systems biology is to provide a holistic view of a biological system. The unification of proposed modelling approaches PLGTS and PEPN-GRN in a single framework is a step towards achieving this goal, where along with the inference of individual subsystems, the

interplay between the different subsystems is also illustrated.