“Inferring the large-scale structure of networks”

Networks form the backbone of a wide variety of complex systems, ranging from food webs, gene regulation, social networks, transportation and the internet. However, due to the sheer size and complexity of many of these systems, it remains an open challenge to formulate general descriptions of their structures, and to extract such information from data. Since networks are high-dimensional relational objects, they cannot be directly inspected using basic tools, and instead require new methodology. In this talk, I will describe a principled approach to this task, based on the elaboration of probabilistic generative models, and their statistical inference from data. In particular, I will present a general class of generative models that describe the multilevel modular structure of network systems, as well as efficient algorithms to infer their parameters. I will highlight the common pitfalls present in more heuristic methods of capturing this type of structure, and demonstrate the efficacy of more principled methods based on Bayesian statistics and statistical physics.

Throughout the talk I will show applications using many empirical networks such as the internet at the autonomous systems level, the global airport network, the network of actors and films, social networks, citations among websites, co-occurrence of disease-causing genes and others.