Network statistics and processes on networks

Edoardo M. Airoldi Department of Statistics Harvard University airoldi@fas.harvard.edu

Abstract

Networks are ubiquitous in science and have become a focal point for discussion in everyday life. Formal statistical models for the analysis of network data have emerged as a major topic of interest in diverse areas of study, and most of these involve a collections of measurements on pairs of objects. Probability models on graphs date back to 1959. Along with empirical studies in social psychology and sociology from the 1960s, these early works generated an active network community and a substantial literature in the 1970s. This effort moved into the statistical literature in the late 1970s and 1980s, and the past decade has seen a burgeoning network literature in statistical physics and computer science. The growth of the World Wide Web and the emergence of online networking communities such as Facebook and LinkedIn, and a host of more specialized professional network communities has intensified interest in the study of networks and network data. In this talk, I will review a few ideas that are central to this burgeoning literature, and then focus on a few problems including network representation, sampling and inference about processes on networks.