Thesis: Modeling And Analysis Of Affiliation Networks With Subsumption

Abstract: An affiliation (or two-mode) network is an abstraction commonly used for representing systems with group interactions. It consists of a set of nodes and a set of their groupings called affiliations. We introduce the notion of affiliation network with subsumption, in which no affiliation can be a subset of another. A network with this property can be modeled by an abstract simplicial complex whose facets are the affiliations of the network. We propose and new model for generating affiliation networks with and without subsumption (represented as simplicial complexes and hypergraphs). In this model, at each iteration, a constant number of affiliations is sampled uniformly at random and then nodes are selected from these affiliations with fixed probability. This implicit preferential attachment growth results in a power-law in the degree distribution.

We develop a theoretical model of this network generation procedure, prove that the degree distribution in the hypergraph case is governed by the Yule-Simon distribution, then find the exponent of its power-law tail. Similarly, we show that in the simplicial complex case, the degree distribution also has a power-law tail, we develop a numerical technique for computing its exponent. We show that the affiliation size distributions can be concisely described via their generating functions. We develop two numerical techniques for solving the resulting functional equations, find the generating functions and compute their PMFs. Furthermore, it turns out that a shifted Poisson distribution is a limiting case of the affiliation size distribution and can serve as its approximation in practice.

After that, we study the process of a giant component formation, develop a theoretical estimate of the critical threshold when a giant component appears and compare it with experiments. For a more complete description of the proposed network generation procedure, we study the average pairwise distance in the network, its assortativity, clustering coefficient, and use Q-analysis methods to compare our generated networks with other synthetic networks and real-world networks.

Committee:

- Professor Saad Mneimneh, Mentor, Hunter College
- Professor Felisa Vázqyez-Abad, Huntern College
- Professor Amotz Barnoy – Brooklyn College

Outside Member:

- Ram Ramanathn, goTenna Inc.