Chapter 17
Multi-Agent-Based Simulation of University Email Communities

David M.S. Rodrigues
ISCTE – Lisbon University Institute, Portugal

ABSTRACT

In this chapter, a study on informal communication network formation in a university environment is presented. The teacher communication network is analyzed through community detection techniques. It is evident that informal communication is an important process that traverses the vertical hierarchical structure of departments and courses in a university environment. A multi-agent model of the case study is presented here, showing the implications of using real data as training sets for multi-agent-based simulations. The influence of the “social neighborhood,” as a mechanism to create assortative networks of contacts without full knowledge of the network, is discussed. It is shown that the radius of this social neighborhood has an effect on the outcome of the network structure and that in a university’s case this distance is relatively small.

INTRODUCTION

Networks are pervasive in our lives. They are everywhere, from the Internet, to biology, to social relations or economics (M. Newman, Barabasi, & Watts, 2006). The notion of a connected world is one that we assume is prevalent as a foundation. The connectedness of daily things prevails, even when one segments it into sub-networks.

DOI: 10.4018/978-1-60960-171-3.ch017

Everything seems related, in some sense, to everything else.

In communication, the notion of networks is always present. Formal and informal relations arise from the interplay of actors during communication processes. Traditionally, networks have been categorized into four types, or classes: lattice networks, that are very regular and rigid, where a certain pattern is repeated ad infinitum; random networks, in which every connection is established according to some probability \( p \); small-world networks that are somewhat in
between random networks and lattice networks, and have high transitivity and short average path lengths; and finally scale-free networks that have the same type of structure at different levels, with a characteristic hub and spoke structure, where every connection is made according to the degree of existing connections. In this scenario, informal communication networks seem to be formed according to other types of rules, as they can’t truly be mapped into one of those four types of networks. These non-trivial networks, arise from the “social” aspect of these kinds of networks and several authors have discussed the problems of those four types of networks in failing to explain social networks. The problems that informal social communication networks present make them well-suited to be tested under new models of network formation and actor interplay. Multi-agent-based simulation is a popular field where these ideas can be tested and where ideas can be benchmarked.

In the next sections, we discuss the mechanisms presently available for community detection, mainly those developed with networks in mind, and we discuss some application of these algorithms to informal email communication systems through a case study. Also, we present a multi-agent model developed for exploring the influence of using real data in simulation and to test the idea of a “social neighborhood” in the formation of informal assortative communication networks.

COMMUNITY DETECTION ALGORITHMS

Community detection can be very useful for performing an exploratory analysis of data, and its usage transverses several domains, from statistics to computer science, biology or psychology. In every science, it is necessary to deal with empirical data, and one of the first classifications that one tries is to group the data according to some property that might manifest itself similarly inside the groups. Several algorithms and techniques have been devised to accomplish this partitioning (Fortunato, 2010), but in practice all are faced with situations where a good partitioning isn’t accomplished, and new methods have to be devised. Some methods are robust, and can be used effectively to classify groups with sets of data that are very heterogeneous. On the other hand, some are very specific to certain problems and need initial conditions that are particular to make its results appropriate (Shortreed, 2006).

The span of techniques and algorithms that tackle the problem of classification and identification of communities in graph representations of data has seen a great amount of interest and developments in the past few years. The field isn’t confined only to traditional methods like graph partitioning, hierarchical clustering, partitional clustering and spectral clustering. There is a new set of divisive methods based on modularity, dynamical algorithms, spectral algorithms, and based on statistical inference that populated this field with several possible approaches on how to obtain information about structure (mainly social, but not limited to) that can be organized in the form of a graph (Fortunato, 2010).

Looking into more detail at some clustering techniques for networks, these can be divided into two main classes, according to the approach they take to the partitioning problem: global or local. In global strategies, the network is taken as a whole and usually a general property is used to divide the network, separating all of its members into clusters. One example of these techniques is hierarchical clustering algorithms like Girvan-Newman (Girvan & Newman, 2001) that use edge betweenness as the property of interest. In local algorithms, the strategy uses some local patterns when considering which points belong to each cluster. Clique percolation (Palla, Derényi, Farkas, & Vicsek, 2005) uses the notion of cliques to identify groups or modules.

Hierarchical Clustering is a type of partitioning strategy that produces a dendrogram from the