Chapter 2

STATISTICAL PROPERTIES OF SOCIAL NETWORKS

Mary McGlohon
School of Computer Science
Carnegie Mellon University
mmc gloho@cs.cmu.edu

Leman Akoglu
School of Computer Science
Carnegie Mellon University
lakoglu@cs.cmu.edu

Christos Faloutsos
School of Computer Science
Carnegie Mellon University
christos@cs.cmu.edu

Abstract In this chapter we describe patterns that occur in the structure of social networks, represented as graphs. We describe two main classes of properties, static properties, or properties describing the structure of snapshots of graphs; and dynamic properties, properties describing how the structure evolves over time. These properties may be for unweighted or weighted graphs, where weights may represent multi-edges (e.g. multiple phone calls from one person to another), or edge weights (e.g. monetary amounts between a donor and a recipient in a political donation network).

Keywords: Power laws, network structure, weighted graphs

What do social networks look like on a global scale? How do they evolve over time? How do the different components of an entire network form? What
happens when we take into account multiple edges and weighted edges? Can we identify certain patterns regarding these weights?

There has been extensive work focusing on static snapshots of graphs, where fascinating properties have been discovered, the most striking ones being the ‘small-world’ phenomenon [38] (also known as ‘six degrees of separation’ [24]) and the power-law degree distributions [3, 12]. Time-evolving graphs have attracted attention only recently, where even more fascinating properties have been discovered, like shrinking diameters, and the so-called densification power law [18]. Moreover, we find interesting properties in terms of multiple edges between nodes, or edge weights.

In this chapter we will describe some of the most important properties apparent in social networks, with a particular emphasis on dynamic properties, and some of the newer findings with respect to edge weights.

The questions of interest are:

- **What do social networks look like, on a large scale?** Do most nodes have few connections, with several “hubs” or is the distribution more stable? What sort of clustering behavior occurs?
- **How do networks behave over time?** Does the structure vary as the network grows? In what fashion do new entities enter a network? Does the network retain certain graph properties as it grows and evolves? Does the graph undergo a “phase transition”, in which its behavior suddenly changes?
- **How do the non-giant weakly connected components behave over time?** One might argue that they grow, as new nodes are being added; and their size would probably remain a fixed fraction of the size of the GCC. Someone else might counter-argue that they shrink, and they eventually get absorbed into the GCC. What is happening, in real graphs?
- **What distributions and patterns do weighted graphs maintain?** How does the distribution of weights change over time—do we also observe a densification of weights as well as single-edges? How does the distribution of weights relate to the degree distribution? Is the addition of weight bursty over time, or is it uniform?

Answering these questions is important to understand how natural graphs evolve, and to (a) spot anomalous graphs and sub-graphs; (b) answer questions about entities in a network and what-if scenarios; and (c) discard unrealistic graph generators.

Let’s elaborate on each of the above applications: Spotting anomalies is vital for determining abuse of social and computer networks, such as link-spamming in a web graph, fraudulent reputation building in e-auction systems [29], detection of dwindling/abnormal social sub-groups in a social-networking site like Yahoo-360 (360.yahoo.com), Facebook (www.facebook.com) and LinkedIn