Mathematical Statistics
Stockholm University
Research Report **2018:20**,
http://www.math.su.se

Directed preferential attachment models

Tom Britton*

October 2018

Abstract

The directed preferential attachment model is revisited. A new exact characterization of the limiting in- and out-degree distribution is given by two *independent* pure birth processes that are observed at a common exponentially distributed time T (thus creating dependence between in- and out-degree). The characterization gives an explicit form for the joint degree distribution, and this confirms previously derived tail probabilities for the two marginal degree distributions. The new characterization is also used to obtain an explicit expression for tail probabilities in which both degrees are large. A new generalised directed preferential attachment model is then defined and analysed using similar methods. The two extensions, motivated by empirical evidence, are to allow double-directed (i.e. undirected) edges in the network, and to allow the probability to connect an ingoing (outgoing) edge to a specified node to also depend on the out-degree (in-degree) of that node.

Keywords: Preferential attachment, directed network, birth processes, tail distribution.

1 Introduction and models

The (undirected) preferential attachment model (PA) is a random network model defined by Barabási and Albert [1]. To start off, the network consists of one single node without any edge. At each time step $k = 1, 2, \dots$, a new node with m (a fixed integer-parameter in the model) new edges connected to it, is added. Each of the new edges of the node is connected, independently, to existing nodes, and the probability to connect to a specific node with current degree i is proportional to i. Two novel features, as compared to most other network models at the time, were that it was defined sequentially, thus with nodes having different ages, and that the degree distribution of nodes in a large network were shown to have power law tails rather than exponentially decaying tail probabilities.

In 2003, Bollobás et al [3] defined a related model, but now for a network in which edges are directed rather than undirected. As in the undirected model, edges/nodes are entered at each discrete time step. However, now one of three different possibilities may happen: 1) either a new node with an *outgoing* edge is added (with probability α), or 2) a new

*Stockholm University, Sweden; tom.britton@math.su.se