EXTREMES OF STOCHASTIC SEQUENCES WITH APPLICATION TO WEB GRAPHS

NATALIA MARKOVICH

V. A. Trapeznikov Institute of Control Sciences of Russian Academy of Sciences, Moscow, 117997 Russia
e-mail: nat.markovich@gmail.com

The fast finding of most influential nodes in a network graph and its declustering constitute important research problems of the analysis of Web communication, social and complex networks. PageRank [5] and a Max-linear model [2] are considered as two indices of the influence of network nodes. An extremal index of PageRank determines the first hitting time, i.e. a minimal time to reach the first influential node by means of a PageRank random walk. Following [1], [5], we consider the PageRank of a random Web page as an autoregressive process with a random number of random coefficients that depend on ranks of incoming nodes and their out-degrees, and a user preference term. The coefficients are assumed to be regularly varying distributed with different tail indices. It is proved that the tail and extremal indices are the same for both PageRank and the Max-linear model and the values of the extremal index depending on the tail indices are found [4]. The results are based on the study of stochastic sequences of random lengths [3] and the comparison of the distributions of their maxima and linear combinations. The exposition is accompanied by some examples based on simulation and the analysis of graph data stemming from real Web graphs.

References