Tuning Topology Generators Using Spectral Distributions

Hamed Haddadi[1], Damien Fay[2], Steve Uhlig[3], Andrew Moore[2], Richard Mortier[4], Almerima Jamakovic[3], and Miguel Rio[1]

[1] University College London
[2] University of Cambridge
[3] Delft University of Technology
[4] Vipadia Ltd

Abstract. An increasing number of synthetic topology generators are available, each claiming to produce representative Internet topologies. Every generator has its own parameters, allowing the user to generate topologies with different characteristics. However, there exist no clear guidelines on tuning the value of these parameters in order to obtain a topology with specific characteristics.

In this paper we optimize the parameters of several topology generators to match a given Internet topology. The optimization is performed either with respect to the link density, or to the spectrum of the normalized Laplacian matrix. Contrary to approaches in the literature that rely only on the largest eigenvalues, we take into account the set of all eigenvalues. However, we show that on their own the eigenvalues cannot be used to construct a metric for optimizing parameters. Instead we present a weighted spectral method which simultaneously takes into account all the properties of the graph.

Keywords: Internet Topology, Graph Spectrum.

1 Introduction

Today's Internet is formed from more than 25,000 Autonomous Systems (ASes), each of which can contain tens or hundreds of routers. Constant evolution and change in the Internet, due to failures and bugs in the short term, and growth and death of networks in the long term, has made it difficult for scientists to produce representative Internet topologies at either AS or router level. However, such maps are essential for the simulation and analysis of ideas including new and improved routing protocols, and peer-to-peer, media-streaming applications. Since obtaining accurate, timely maps of the Internet topology is difficult, and development of new protocols and systems requires understanding their performance over a range of scenarios, researchers use synthetic topology generators.

There are many such generators, each of which is parameterized, often with multiple parameters, giving rise to a plethora of potential synthetic graphs. Understanding and generating those graphs, useful because they accurately

S. Kounev, I. Gorton, and K. Sachs (Eds.): SIPEW 2008, LNCS 5119, pp. 154–173, 2008.

represent features of the true underlying Internet graph, is difficult. Existing approaches to tuning the generator parameters range from selection of particular metrics of interest, e.g., link count, and tuning to match that particular metric, to simply using the default parameters encoded in the particular release of the generator package in use!

The core problem is to select an appropriate cost function which reflects those aspects of the graph that are important to the user and weights those aspects accordingly. Such a selection process is inherently subjective: there is no "best" cost function in general. Once a suitable cost function is selected, it is a simple matter to tune the available parameters of the topology generator to produce output that optimally matches said cost function.

In the light of this, our contributions in this paper are as follows:

- We propose a new cost function, the *weighted spectrum*, constructed from the eigenvalues of the normalized Laplacian matrix, or graph spectrum;
- We demonstrate that the graph spectrum alone is unsatisfactory as a cost function;
- We provide an efficient approximation of the weighted spectrum;
- We use this approximation to tune parameters for a set of Internet topology generators, enabling us to use these generators to effectively match a particular measured Internet topology.

The graph spectrum is a useful starting point for such a cost function as it yields a set of invariants about a graph that encode all the properties of that graph [8]. Our proposed cost function improves on the simple graph spectrum because it incorporates the knowledge that not all eigenvalues are equally important, and weights toward those that are considered to encode more significant aspects of the graph's structure. The basis of our algorithm is to provide a way to measure the difference between two graphs with respect to a common reference, a suitable regular graph.[1]

After reviewing related work in Section 2, we outline background theory in Section 3 before introducing the topology generators we use in Section 4. In Section 5 we present the results of our analysis and in Section 6 we compare topologies generated at optimal values of the parameters with an observed dataset. Finally, we conclude the paper in Section 7 and discuss future work.

2 Related Work

Zegura *et al.* [27] analyze topologies of 100 nodes generated using pure random, Waxman [25], exponential and several locality based models of topology such as Transit-Stub [6]. They use metrics such as average node degree, network diameter, and number of paths between nodes, and use the number of edges as the metric of choice for optimization of the tuning parameter. However as we show in this paper, the number of links is not an ideal choice particularly in random

[1] A regular graph is one where all nodes have the same degree.