Growing Graphs from Hyperedge Replacement Grammars

Salvador Aguinaga Corey Pennycuff Rodrigo Palacios[†] Tim Weninger
University of Notre Dame
[†]California State University – Fresno
{saguinag, cpennycu, tweninge}@nd.edu

ABSTRACT

Lorem ipsum dolor sit amet, consectetuer adipiscing elit. Ut purus elit, vestibulum ut, placerat ac, adipiscing vitae, felis. Curabitur dictum gravida mauris. Nam arcu libero, nonummy eget, consectetuer id, vulputate a, magna. Donec vehicula augue eu neque. Pellentesque habitant morbi tristique senectus et netus et malesuada fames ac turpis egestas. Mauris ut leo. Cras viverra metus rhoncus sem. Nulla et lectus vestibulum urna fringilla ultrices. Phasellus eu tellus sit amet tortor gravida placerat. Integer sapien est, iaculis in, pretium quis, viverra ac, nunc. Praesent eget sem vel leo ultrices bibendum. Aenean faucibus. Morbi dolor nulla, malesuada eu, pulvinar at, mollis ac, nulla. Curabitur auctor semper nulla. Donec varius orci eget risus. Duis nibh mi, congue eu, accumsan eleifend, sagittis quis, diam. Duis eget orci sit amet orci dignissim rutrum.

Categories and Subject Descriptors

H.4 [**Information Systems Applications**]: Miscellaneous; H.2.8 [**Database Management**]: Database Applications-Data mining

General Terms

Algorithm, Theory

1. EXPERIMENTS

Hyperedge replacement grammars enables us to learn how networks grow and how to predict future instances of the network in question. In this work, we examine well established generative models to compare and distinguish the benefits our approach provides. We consider the properties that underlie a number of real-world networks and compare the distribution of graphs generated using generators for random graphs, such as Erdos-Renyi-Gilber, Newman-watts-Strogatz, Barabasi-Albert, Kronecker, ERGM, and Forestfire graphs.

Experiments were carried out by implementing hyperedge replacement grammars graph induction using the following high-level programming laguages: Python, R, and C++.

1.1 Real Networks Examined

We cosider real-world networks that exibit properties that are both common to many networks across different fields and distinct properties inherent only to some networks. These network examples help frame the context of this work.

Table 1: S

Type	**Dataset Name**	**G(n,m)**
toprule Social Networks	Karate Club	$n = 34, m = 78$
	Some other	
Knowledge Network	Wikipedia	
Scientific Collaboration	Microsoft Academic Search Bibliographic	
	DBLP Bibliographic	

1.2 Erdos-Renyi-Gilbert

A bionomial gaph that choses eah of the possible edges with a probablity p.

1.3 Newman-Watts-Strogatz

Models a small-world graph by specifying the number of nodes, where each node is joined by its k nearest neighbors in a ring topology with shortcuts created by the addition of new edges of each edge having a probability p

1.4 Preferntial Attachement

The Barabasi-Albert graph model is configured to have n nodes and m edges that preferentially attach from a new node to existing nodes of high degegree [**?**].

1.5 Kronecker Graphs

1.6 Exponential Random Graph Models

2. RELATED WORK

Well-accepted and useful applications of graph theory includes those used to model networks in the real world. Understanding the organizing principles of natural and engineered networks: symmetry, invariance, density, and regularity– is what mathematicians, engineers, and computer scientists feast on. How networks grow is an area of interest in physics, biology, sociology, and engineering, but it not limited to this set by far. In the social sciences social structure is explored in a study of a problem with runaways, which dates back to the 1930s [**?**]. Interest in generative models has been around for a long time. Significant advancement in the study of network structure using models for generating random graphs was pioneered in by Edgar Gilber, Paul Erdos and Alfred Renyi around 1960 [**?**, **?**]. Henceforth, a number of generative models have been developed on all fronts, which are designed to exhibit the network properties that best explains the available data [**?**].

The challenge lies in balancing competing forces of simplicity of design, a concept tightly coupled with the need to capture salient features of the network, and the challenges of analyzing the model– if want it to be good or at least useful.

Here we describe relevant prior art in two research areas our contribution is perched on: generative modes and hyper-edge replacement grammars.

2.1 Generative Models

Generative models that underlie the dynamically growing Web graph (as in the World Wide Web) have received a great deal attention for some time now. This graph's nodes and edges have been shown to exhibit power law and lognormal distributions in empirical studies [**?**, **?**]. The Web graph is a real-world graph, with pages and its hyperlinks correspond this graph's nodes and edges, respectively. We begin by looking at *preferenital attachment*, which is an example of a graph that grows by having new added objects attach to popular objects. *Add more detail?*

Kronecker graphs is another class of a generative network models obeying many of the static and temporally evolved network patterns empirically observed in real-world networks including the Web, internet topology, peer-to-peer networks, etc.) [**?**, **?**, **?**].

Exponential random graphs models (ERGMs) belong to a class of statistical models, also known as p^* models, that has been used extensively to model social behavior in humans and animals. More recently, ERGMs are being used to model the structure as well as other complex neurological interactions of the brain bullmoreSporns2009complex,10.1371/Simpson2011.pone.0020039,GOODREAU2009. Goldenberg *et al.* survey statistical models and discuss how ERGMs are an extension of of the Erdos-Renyi-Gilbert model to account for popularity, expansiveness and network effects due to reciprocation [**?**, **?**].

2.2 Hyper-edge Replacement Grammars