Information Sciences 288 (2014) 220–245

Contents lists available at ScienceDirect

Information Sciences

journal homepage: www.elsevier.com/locate/ins

# Probabilistic inequalities for evaluating structural network measures

Veronika Kraus [a], Matthias Dehmer [a,b,*], Frank Emmert-Streib [c]

[a] Division for Bioinformatics and Translational Research, UMIT, Eduard Wallnoefer Zentrum 1, 6060 Hall in Tyrol, Austria
[b] Department of Computer Science, Universität der Bundeswehr München, Werner-Heisenberg-Weg 39, 85577 Neubiberg-München, Germany
[c] Computational Biology and Machine Learning Laboratory, Center for Cancer Research and Cell Biology, School for Medicine, Dentistry and Biomedical Sciences, Queen's University Belfast, Belfast, United Kingdom

ARTICLE INFO

*Article history:*
Received 1 March 2013
Received in revised form 18 January 2014
Accepted 17 July 2014
Available online 12 August 2014

*2010 MSC:*
62D99
05C75
68R10
90B10

*Keywords:*
Quantitative graph theory
Information inequality
Statistics
Random network models

ABSTRACT

Proving interrelations between structural graph measures analytically has been intricate. Generally, relations between structural graph measures describe the interplay between measures which turned out to be useful for better understanding the properties of such quantities. The results which have been achieved so far are restricted to simple measures or special graph classes such as trees. In this paper, we introduce a probabilistic approach for establishing inequalities between quantitative network measures on random networks. Those inequalities between different graph measures lead to a deeper understanding of the mathematical apparatus and, in particular, to a discussion of quality aspects of structural graph measures, which is a major contribution of this paper.


## 1. Introduction

Structural network measures have been studied extensively for several decades [32,59,35,46]. However, when studying the existing body of literature, we see that most of those contributions deal with applications rather than exploring mathematical properties of structural graph measures. Typical problems in this context are, for example, to investigate the ability of a measure to distinguish graphs structurally, which is often referred to as uniqueness or degeneracy (see Section 2.6), sensitivity properties, the structural interpretation of a measure and its behaviour on random graphs [5,17,22,20,68]. Yet, an important problem that has received little attention so far is to prove interrelations between network measures that lead to a deeper understanding of the mathematical apparatus [24,23,74]. As already outlined by Dehmer and Mowshowitz [24], knowing the relations between quantitative based measures is of great practical importance for analyzing complex networks as they describe interplay between graph measures which is usually not known. Such inequalities can be inferred for

* Corresponding author. Address: Fakultät für Informatik, Universität der Bundeswehr München, Werner-Heisenberg-Weg 39, 85577 Neubiberg, Germany. Tel.: +43 (0)50 8648 3851.
*E-mail addresses:* veronika.kraus@gmx.at (V. Kraus), matthias.dehmer@umit.at (M. Dehmer), v@bio-complexity.com (F. Emmert-Streib).

http://dx.doi.org/10.1016/j.ins.2014.07.018
0020-0255/

general graphs as well as for special graph classes, see [24,23]. To explain the problem more precisely let us assume two structural graph measures with a given structural interpretation (e.g., branching, cyclicity, connectedness and so forth, see [22]). Formally, a graph measure is a mapping $M : \mathcal{G} \longrightarrow \mathbb{R}$ where $\mathcal{G}$ is a class of graphs. Then, inequalities of the form, e.g.,

$$M_1(G) \underset{(\geqslant)}{<} M_2(G), \tag{1}$$

describe interrelations between two graph measures $M_1$ and $M_2$ possessing the interpretation $i_1$ and $i_2$, respectively. For instance, that means a branching measure (e.g., $i_1$) can be bounded above by a cyclicity measure (e.g., $i_2$) for a certain class of graphs. To infer such interrelations is valuable as many of those (possible) relations have not been explored so far in view of the vast amount of existing graph measures. As related work in this area demonstrates [18,25,29,23], inferring inequalities for graphs analytically has been challenging. As a consequence, we therefore propose a probabilistic framework for deriving such inequalities by using random graphs which allow to draw conclusions on various graph measures more meaningfully.

Note that some inequalities for simple graph measures have already been achieved for measures which are based on distances or vertex degrees [52,74]. Proving analytical results (inequalities) for various network measures and general graphs turned out to be a daunting problem though. For example, recent work has shown that it is quite challenging to obtain general results for exploring interrelations between graph entropies, see [23,51].

In this paper, we propose a probabilistic approach to establish inequalities between structural graph measures. In order to do so, we use several categories of graph measures which have been investigated in the literature. Also, this choice can be motivated by emphasizing that the measures may capture structural information very differently if their definitions (categories) are based on distinct principles (e.g., degree-based vs. distance-based, entropy-based vs. eigenvalue-based, etc.) More precisely, we statistically infer such inequalities for Barabási–Albert and Watts–Strogatz graphs and claim that the corresponding inequality is true with a high probability for two such random graphs. This extends related work in this area significantly for three reasons. First, we establish probabilistic inequalities by using structural graph measures for which analytical results are too difficult to obtain. Second, other graph classes can be employed easily leading to results for several graph classes and, hence, to deeper insights [18,25,29,23,53]. Third, the probabilistic approach allows to draw conclusions on the measures by taking the structure of the underlying networks into account [18,25,29,23]. In particular, this apparatus also extends the recent results on *information inequalities for graphs* [18,25,29,23] (see Section 2.2) as information-theoretic graph measures can be employed easily. Apart from the novel inequalities for several graph measures, we use those results to extend the discussion about the *quality* of structural graph measures (see Section 2.6).

This paper is structured as follows:

- Section 1 explains the state of the art as well as the underlying problem we deal with.
- Section 2 introduces the random models for generating the graphs under consideration. Also, this section explains the definition and meaning of probabilistic inequalities for graphs in general.
- Sections 3–7 discuss the inference and interpretation of probabilistic inequalities for distance-based measures (Section 3), degree-based measures (Section 4), geometric–arithmetic indices (Section 5), entropy-based measures (Section 6), and eigenvalue-based measures (Section 7).
- The paper finishes with a summary and discussion in Section 8.

## 2. Fundamentals and problem

The problem investigated in this paper is inspired by the theory of information inequalities in information theory [33,72,73], which stimulated the need of establishing information inequalities by using information-theoretic graph measures [18,25,29,23,51]. The aim of the present paper is to study a variety of graph measures on random networks to infer probabilistic and statistical inequalities. We expect that this leads to new insights when studying the measures and to a deeper understanding regarding the structural interpretation and relation between those measures. In order to start, we first introduce basic random graph models we are going to use in this paper. Then, we explain the type of probabilistic inequalities for graphs we want to infer.

### 2.1. Random network models

Several models for random graphs have been developed [3,8,34,37,58]. In this paper, we study two such models, both of which are based on a random generation process depending on some free parameters. We emphasize that the structural properties of Barabási–Albert and Watts–Strogatz networks have been explored extensively. Therefore there exist various theoretical results to describe the structure of these graphs depending on their free parameters [3,8,34,37,58]. This fact supports that our results for inferring probabilistic inequalities for networks can be interpreted straightforwardly.

We start by explaining the Barabási–Albert model [3] and note that those networks model scale-free networks, whose degree-distribution follows a power law, that is, the fraction of vertices of degree $k$ decays by

$$P(k) = k^{-\gamma},$$