Seminars in Cell & Developmental Biology 51 (2016) 44–52

Contents lists available at ScienceDirect

Seminars in Cell & Developmental Biology

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

Review

Evaluating network inference methods in terms of their ability to preserve the topology and complexity of genetic networks☆

Narsis A. Kiani*, Hector Zenil*, Jakub Olczak, Jesper Tegnér

Unit of Computational Medicine, Science For Life Laboratory (SciLifeLab), Center for Molecular Medicine, Department of Medicine Solna, Karolinska Institute, Stockholm, Sweden

ARTICLE INFO

Article history:
Received 7 January 2016
Accepted 7 January 2016
Available online 3 February 2016

Keywords:
Network reverse engineering
Network reconstruction
Evaluation of networks
Information content
Shannon entropy
Algorithmic complexity

ABSTRACT

Network inference is a rapidly advancing field, with new methods being proposed on a regular basis. Understanding the advantages and limitations of different network inference methods is key to their effective application in different circumstances. The common structural properties shared by diverse networks naturally pose a challenge when it comes to devising accurate inference methods, but surprisingly, there is a paucity of comparison and evaluation methods. Historically, every new methodology has only been tested against *gold standard* (true values) purpose-designed synthetic and real-world (validated) biological networks. In this paper we aim to assess the impact of taking into consideration aspects of topological and information content in the evaluation of the final accuracy of an inference procedure. Specifically, we will compare the best inference methods, in both graph-theoretic and information-theoretic terms, for preserving topological properties and the original information content of synthetic and biological networks. New methods for performance comparison are introduced by borrowing ideas from gene set enrichment analysis and by applying concepts from algorithmic complexity. Experimental results show that no individual algorithm outperforms all others in all cases, and that the challenging and non-trivial nature of network inference is evident in the struggle of some of the algorithms to turn in a performance that is superior to random guesswork. Therefore special care should be taken to suit the method to the purpose at hand. Finally, we show that evaluations from data generated using different underlying topologies have different signatures that can be used to better choose a network reconstruction method.

Contents

☆ Dr. Ali Masoudi-Nejad, sole Guest Editor for this paper.

* Corresponding authors.
E-mail addresses: Narsis.Kiani@ki.se (N.A. Kiani), Hector.Zenil@ki.se (H. Zenil).

http://dx.doi.org/10.1016/j.semcdb.2016.01.012
1084-9521/

1. Introduction

Many real-world networks, such as complex technological and social networks, belong in the category of so-called 'complex networks', and display a number of the properties that govern the formation and evolution of complex networks [23,1,2]. Studying biological networks encompasses network analysis, comparisons, modeling, and alignments aimed at discovering a relationship between network topology on the one hand and biological function and disease on the other. The accurate inference of networks from biological data is an open challenge, and over the past few years it has developed into a broad field of study, driven by the application of ever more sophisticated techniques.

The *Dialogue for Reverse Engineering Assessment and Methods* (DREAM) challenge [34,35] has fostered significant progress. The DREAM challenge aims to fairly compare the strengths and weaknesses of inference methods. Network inference methods have complementary pros and cons under different conditions. Ideally, the validation and interpretation of GRN models must keep pace with new knowledge and experimental data available for modeling, and thus it is important to illustrate all aspects and capacities of a network inference method.

Comparing large cellular networks, which is analogous to genetic sequence comparison, will revolutionize biological understanding. However, comparing large networks is computationally infeasible due to the NP-completeness of the underlying subgraph isomorphism problem. Thus, large network analyses and comparisons rely on heuristics, commonly called network parameters or properties. Furthermore, our understanding of gene regulatory networks is still only partial [4,22].

Generally, researchers produce artificial networks and simulated data for method assessment. Synthetic data do not usually reflect the complexity of a real biological system if no prior biological information is introduced. Although the exact details may differ, most methods of evaluation of network reconstruction consider the sensitivity, specificity, precision, and in some cases, a receiver operating characteristic (ROC) curve in illustrating the performance of a method. The analysis of biological networks has led to the realization that the architecture of these networks shares many features with other complex networks. They show non-trivial topological properties such as modular structure and long-tail degree distribution [29].

The common structural properties shared by diverse networks naturally pose a challenge when it comes to devising more accurate inference methods capable of preserving them. Surprisingly, there has been no evaluation or comparison of different models from this point of view.

Understanding the advantages and limitations of different network inference methods is necessary for their effective application in specific circumstances. In this paper we address this question, evaluating the similarity between the structural features of a true network and those of an inferred network. We have chosen six different inference algorithms from among the best-performing algorithms in past DREAM challenges. These methods have been studied using statistical performance measures such as the *F*-score [22] or area under the receiver operator curve (AUROC) [18,15]. Attempts have been made to consider aspects of the overall properties of an inferred network other than the specific number of false and true positive/negative edge inference cases [32].

In this article we analyze network inference methods, employing topological measures and indices in combination with ensemble data in order to assess their performance. An effective similarity metric is needed for scoring network inference methods, one which, given two complex networks, evaluates the degree of similarity between their structural features, beyond just looking at individual numbers. We borrow ideas from gene set enrichment analysis (GSEA) [37,19,5] to formulate an intelligent method which we offer as a new way to measure the topological similarity of two complex networks. We benchmark them using synthetic transcriptional networks proposed by Mendes et al. [28] using a Hill equation method. These networks consist of 100 "genes" with 200 connections among them, exemplifying three different topologies: Erdös-Rényi [8], small world (SW) [38] and scale-free [29] (SF). Mendes et al. have used these networks with well-defined topologies to run in-silico experiments simulating real laboratory gene expression values. We compared ARACNe [25], CLR (Context Likelihood of Relatedness [9], GENIE3 [16], INFERELATOR [11], TIGRESS [12] and Correlation on the basis of diameter, average shortest path length, clustering and centrality scores.

2. Methods

Six different network inference algorithms are considered in this study and will be discussed in the following section. Table 1 summarizes the differences between the models analyzed.

2.1. Network inference methods

Several methods have been proposed for inferring gene regulatory interactions from measured gene expression levels. Approaches employed include Bayesian networks, Boolean models, auto-regressive models, correlation-based models, clustering techniques and differential equation models, among others [10,26,6,18,24]. Some of them are static, while others take into account the dynamic aspects of the dependencies. Mutual information network inference methods are a class of network inference methods which infer regulatory interactions between genes based on pairwise mutual information. The low computational complexity and the low number of required samples are the main advantages of mutual information based inference methods. We have examined two commonly used state-of-the-art network inference methods based on pairwise mutual information: Algorithm for the Reconstruction of Accurate Cellular Networks, ARACNe [25] and Context Likelihood of Relatedness, CLR [9]. ARACNe is based on an information theoretical approach that uses the concept of mutual information, MI, a measure of entropy, to determine the pairwise interaction between nodes by assessing the MI between them. It then applies a data processing inequality (DPI) to eliminate indirect interactions. The CLR algorithm is an extension of the network relevance approach. It is another information theoretic approach and computes the MI between two nodes, comparing it to the empirical background distributions of MI. Regression based network inference methods comprise one of the largest network inference sub-categories, and we have studied three of the best regression based methods: GENIE3, TIGRESS and INFERELATOR. GENIE3 or GENIE [16] decomposes the prediction of a regulatory network between p genes into p different regression problems such that in each, the expression pattern of one of the genes may be