Nano Communication Networks 6 (2015) 133–144

Contents lists available at ScienceDirect

Nano Communication Networks

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

Contribution of canonical feed-forward loop motifs on the fault-tolerance and information transport efficiency of transcriptional regulatory networks

Ahmed F. Abdelzaher [a], Michael L. Mayo [b], Edward J. Perkins [b], Preetam Ghosh [a,*]

[a] Department of Computer Science, Virginia Commonwealth University, Richmond, VA, USA
[b] Environmental Laboratory, US Army Engineer Research and Development Center, Vicksburg, MS, USA

ARTICLE INFO

Article history:
Received 16 March 2015
Accepted 1 April 2015
Available online 21 April 2015

Keywords:
Shortest path
Centrality
Graph randomization
Clustering coefficient
Feed-forward loop

ABSTRACT

Motifs and degree distribution in transcriptional regulatory networks play an important role towards their fault-tolerance and efficient information transport. In this paper, we designed an innovative in silico canonical feed-forward loop motif knockout experiment in the transcriptional regulatory network of *E. coli* to assess their impact on the following five topological features: average shortest path, diameter, closeness centrality, global and local clustering coefficients. Additional experiments were conducted to assess the effects of such motif abundance on *E. coli*'s resilience to nodal failures and the end-to-end transmission delay. The purpose of this study is two-fold: (i) motivate the design of more accurate transcriptional network growing algorithms that can produce similar degree and motif distributions as observed in real biological networks and (ii) design more efficient bio-inspired wireless sensor network topologies that can inherit the robust information transport properties of biological networks.

Specifically, we observed that canonical feed forward loops demonstrate a strong negative correlation with the average shortest path, diameter and closeness centralities while they show a strong positive correlation with the average local clustering coefficient. Moreover, we observed that such motifs seem to be evenly distributed in the transcriptional regulatory network; however, the direct edges of multiple such motifs seem to be stitched together to facilitate shortest path based routing in such networks.

1. Introduction

The ability for networks to rewire its links was introduced by biologists when they realized that biological networks can resist external perturbations, yet proceed with their natural activities. This property was referred to as 'Biological Robustness' [1], and was mainly attributed to these network's topological features [2]. For example in the proposed bio-inspired self organizing wireless sensor and actuator network [3], edge rewiring must guarantee optimal topological preservation in case of nodal collapse.

Among the features that aid in network dynamics, is the fact that biological networks are sparse [4], or loosely connected. In such networks, degree distributions can be expressed using a power law, $p(k) \sim k^{-\gamma}$ [5], meaning, a steep negative slope will result from the bi-log plot of the different nodal degrees vs. the nodes that have such degrees. Consequently, few nodes have degrees much higher

* Corresponding author.
E-mail addresses: abdelzaheraf@vcu.edu (A.F. Abdelzaher), Michael.L.Mayo@usace.army.mil (M.L. Mayo), Edward.J.Perkins@usace.army.mil (E.J. Perkins), pghosh@vcu.edu (P. Ghosh).

http://dx.doi.org/10.1016/j.nancom.2015.04.002

than the average degree (i.e. hubs) [6], while the bulk of the remaining nodes have degrees much lower than the average, resulting in loosely connected components. For example, most gene nodes in the transcriptional network of Escherichia coli (herein E. coli), have single incoming links and no outgoing ones.

Most biological networks have $2 < \gamma < 3$ [5]—a property which classifies a network as 'scale-free' [6]. The significance of networks having hub to low degree orientations, can be expressed through the reduced probability of having a detrimental random attack. Most random attacks will result in loss of minimal links [7]. Similarly with edge rewiring, the relative overall damage to the entire network will not be too high. However intentional hub attacks can be very costly and can result in disconnecting the network [8].

Other network classifications fall under two other major categories, namely random (ER) [9] and small-world (SW) [10] networks. The preceding considers links connecting two nodes at equal probabilities during growth, while the latter is inclined towards minimizing the number of hops between pairs of nodes. In contrast with scale-free networks, systems grow using 'the rich get richer and the poor get poorer' phenomena. ER and SW networks typically have tightly connected components, which are different from that of scale-free networks. Nodal degrees are almost equal to the average degree, therefore making random attacks as equally costly as intentional attacks [6].

Another important aspect that aids in biological robustness, is the existence of 3–6 nodal substructures known as 'motifs' [11]. For example, ecologists believe that synthetic communities forming motifs can be inserted with weak interactions to increase the community's stability [12]. These substructures are labeled 'significant' because their abundance in the real networks is much higher as compared with their numbers in multiple randomized versions of these networks [11,13]. One of the major goals of this paper is to demonstrate that motif abundance aids in the overall information transport in such transcriptional regulatory networks (herein TRNs).

Understanding the underlying architecture of TRNs can provide insights in disease dynamics and drug development [14,15]. In a TRN, nodes portray the genes in a cell, and a set of directed links that correspond to interacting pairs or genes [16]. Interactions could either represent translation or transcription [17]. Unlike engineered networks, TRNs exhibit biological robustness [1,2] due to their tolerance of noise during gene expression [18]. This phenomena arises from feed-back control nodal arrangements and repetitive substructures [18], or motifs.

Among the significant 3-node motifs in TRNs, the feed-forward loop (or FFL) has received the most attention. An FFL consists of a transcription factor A that regulates another transcription factor B and a gene C, while B co-regulates C. This topology allows it to deliver essential tasks like generating pulses, irreversible speed ups and signal delays [19].

At the nodal level, research has been directed towards correlating robustness with disturbances in the network. For example, the effect of nodal collapse has been modeled to show a decrease in network efficiency [20,21]—which is defined as the inverse of the average shortest path in the network [22]. The effects of successive random node deletions on the network diameter were studied for scale-free and ER networks [6,7]. Others model disturbances in the form of connection failures [23], or partial inactivation [24], where the length of the links in the TRN of E. coli were increased to resemble interaction delays.

At the motif level, much attention has been focused towards understanding motif functions. Different motif configurations have been investigated using mathematical models of transcription and translation to understand the relationship between coupling and function of embedded motifs [25,26]. Experiments with E. coli have been conducted where transcription factors were rewired and the tolerance of the bacteria was analyzed [27].

However, little is known regarding the *motif distribution* in the network and its contributions towards robustness, particularly in terms of information transport. In this paper, we address this issue by designing *in-silico* FFL knock-outs in E. coli, while preserving the individual in-, out- and cumulative degrees of the nodes. After every successive FFL deletion, different topological features of the resulting TRNs were recorded. Then, Pearson's correlation coefficient is used to determine the correlation between each pair of metrics studied here.

2. Motivation

2.1. Bio-inspired wireless sensor networks

Wireless sensor networks form a special class of engineered systems wherein sensor nodes forward data packets that are routed through adjacent sensors to a sink capable of processing the sensed information. Resemblance between gene regulation systems and wireless sensor networks (herein WSNs) can be described through transcription, where genes process signals from adjacent neighbors in the form of transcription factors that excite/repress other genes by generating mRNA molecules. Nodes in a TRN interface by conveying signals (transcription factors), that are then processed into output signals (mRNAs). WSNs operate in a similar manner, where sensor nodes send signals to others in the form of data packets. Packets at destination nodes convey forwarding instructions, which in return relays such packets to other sensors.

Recently, we have shown that wireless sensor networks adopting the transcriptional regulatory topologies (of E. coli), designated as bio-inspired WSNs, are more efficient than those adopting ER topologies of the same size in terms of conveying packets to sink nodes [28–30]. A support vector machine model was constructed with ~90% accuracy to predict packet receipt rates using the topological features of the networks as input [31,32] that includes the average degree, network density, as well as the abundance of FFLs. Each of these three topological features was ranked higher than the other ones. It is hence important to study how FFL abundance positively or negatively correlates with the other topological features in the network. Such a study will motivate the design of smart WSN topologies that exhibit similar FFL abundance and possibly FFL distribution as observed in the TRNs of E. coli and hence will have better efficiency in terms of their average packet receipt rates under node/link failures and channel noise.