Characterization of *E. coli* Gene Regulatory Network and its Topological Enhancement by Edge Rewiring

Satyaki Roy
Department of Computer Science
Missouri University of Science and Technology
sr3k2@mst.edu

Vijay K. Shah
Department of Computer Science
Missouri University of Science and Technology
vksr38@mst.edu

Sajal K. Das
Department of Computer Science
Missouri University of Science and Technology
sdas@mst.edu

ABSTRACT

The innate resilience of biological organisms have long inspired the design of robust systems. Gene regulatory network (GRN) is one such biological network which possesses a gamut of topological properties that contribute to its robustness. In this work, we study *E. coli* GRN as a three-tier topology to characterize such properties and explain why GRN is particularly vulnerable to failure of hub nodes. We also propose an edge rewiring mechanism on existing *E. coli* GRN topology to strengthen its robustness against hub failure. With extensive experiments on *E. coli* GRN, we show that its topological properties improve significantly after applying edge rewiring. Finally, we design wireless sensor networks using original and rewired *E. coli* GRN topologies. Simulation results indicate that rewired GRN has higher packet delivery and lower latency than original GRN.

Categories and Subject Descriptors

C.2.1 [**Network Architecture and Design**]: Network topology; C.2.1 [**Network Architecture and Design**]: Wireless communication

General Terms

Algorithm, Design, Reliability, Performance

Keywords

Robustness, Three-tier topology, Edge rewiring, Motif-based centrality

1. INTRODUCTION

Biological networks exhibit immense resilience against adverse environmental conditions as they perform vital functions necessary for sustenance. In case of gene regulatory network (GRN), robustness is a function of graph topology [1]. While biologists are trying to elucidate the properties that impart robustness to GRN, computer scientists are aspiring to exploit GRN robustness in design of fault-tolerant communication networks.

Before delving further into details of GRN topology, a brief discussion on GRN is imperative. We know that genetic material of any living organism is contained in thread-like structures called *chromosomes*, within the cell nucleus. The chromosomes are made of molecules of *deoxyribonucleic acid* (DNA), and segments of DNA are called *genes* (Figure 1).

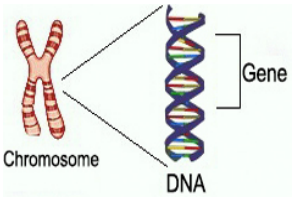

Figure 1: Chromosome, DNA and gene

The gene regulatory network is a biological network formed as a result of interaction between genes. GRN regulates protein synthesis within living cells [2].

Recent studies on GRN explain that its topology exhibits certain interesting graph-theoretic properties. However there are limited instances of well-studied GRN, among which we have two examples of bacteria, namely *E. coli* and yeast [3]. We base our research on the topology of *E. coli* GRN. There is ample evidence showing the existence of properties like scale-freeness, high clustering tendency, preferential attachment and low graph density in *E. coli* GRN, but there is no work explaining the dynamics between these properties.

Our first contribution is to study the *E. coli* GRN as a **three-tier topology** in order to derive better insights into how these properties are related. The three-tier topology not only characterizes the *E. coli* GRN, but also explains its fatal flaw: *GRN breaks into smaller isolated clusters if a few well-connected hub nodes are removed* [4]. **In the context of our paper, robustness is defined as the ability of a network to stay connected despite hub node failures.**

There have been attempts to improve robustness of networks by *edge rewiring*, which refers to the process of removal of edges between certain node pairs and addition of equal number of edges between other node pairs, keeping total number of edges the same. Louzada et al. proposed a method

BICT 2015, December 03-05, New York City, United States

DOI 10.4108/eai.3-12-2015.2262418

to maintain robustness, defined as size of largest connected component, while retaining network efficiency [5]. Yang et al. proposed a 3-step strategy to enhance robustness of complex network while preserving community structure and degree distribution [6]. Xiao et. al showed that robustness of scale-free network can be increased at slightly decreased assortativity coefficient, without changing nodal degree [7].

Inspired by these works, our second contribution is a novel edge rewiring technique on *E. coli* GRN topology to enhance robustness, while preserving its graph properties. **We refer to GRN topology formed as a result of edge rewiring as modified *E. coli* GRN**. While the topology of GRN has been utilized in the past to design wireless sensor network (WSN) [8, 9], our simulations show that the WSN designed using modified GRN topology exhibits better packet delivery rate and lower latency than that of original GRN.

This work has been organized as follows. Section 2 briefly discusses significant topological properties of GRN. Section 3 introduces the three-tier topology. Section 4 covers the mechanism of edge rewiring. Section 5 shows the result of graph theoretic experiments and WSN simulation.

2. GRAPH-THEORETIC PROPERTIES OF E. COLI GRN

This section touches upon the significant graph-theoretic properties of *E. coli* GRN.

2.1 Scale-free property

Scale-free network is a network where very few nodes, called **hubs**, have a very high degree of connectivity and majority of the nodes have a lower degree of connectivity. From the out-degree distribution, it is evident that *E. coli* GRN comes under the ambit of scale-free networks [10].

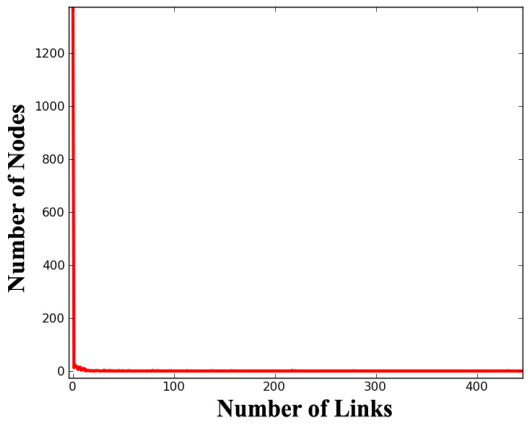

Figure 2: Ecoli GRN: Scale-free out-degree distribution

The disproportionately high degree of the hubs makes *E. coli* GRN *scale-free*, meaning without scale. In Figure 2, we have the degree distribution of *E. coli* GRN showing a very few hub nodes with a high out-degree and majority of nodes with low out-degree of connectivity. The hub nodes form the backbone of scale-free networks.

2.2 Preferential Attachment

When a new node is inducted into a network, it prefers to get attached to a node which has higher degree of connectivity. This property is called preferential attachment.

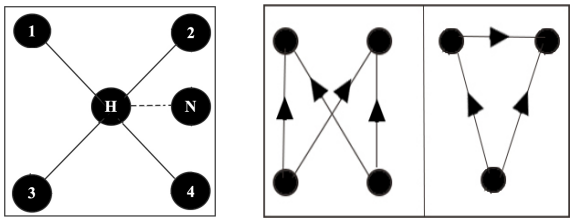

Figure 3: (a)Hub node (labelled H) is preferred candidate of attachment for a new node (labelled N) introduced to network;(b) Two frequently occurring motifs in GRN: bi fan (left) feed forward loop (right)

As a consequence to preferential attachment, the hub nodes tend to acquire more and more links as the network grows. Preferential attachment is an important property in social and biological networks like GRN. The implication of the existence of preferential attachment is depicted in Figure 3(a), where the hub node (labelled H) is the most preferred candidate of attachment for a new node (labelled N). Consequently the hub nodes become increasingly dominant nodes in the GRN by acquiring new edges.

2.3 Motifs and high clustering tendency

GRN of most biological systems are characterized by the existence of recurring patterns of subgraphs consisting of 3 or 4 nodes, called **motifs** (Figure 3(b)). The motifs are called the functional units as they regulate many vital functions within the biological systems via signal exchange [2].

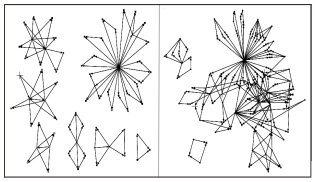

Figure 4: Clustering tendency of motifs: Cluster of feed forward loops (left) and bi fan (right) in the E. coli GRN

The motifs have a overwhelming tendency to form dense clusters resulting in tightly-knit groups in GRN (Figure 4).

2.4 Graph Sparseness

GRN is also characterized by low graph density [2]. If v is number of nodes and e is number of edges, we can define graph density D, for directed graph, on a scale of 0 to 1.

$$D = \frac{e}{v * (v-1)} \tag{1}$$

E. coli GRN is a directed graph with $v = 1564$ and $e = 3758$, so by Equation 1, $D = 0.0015$. This shows the extreme graph sparseness of *E. coli* GRN.

2.5 Vulnerability to hub node failure

GRN, like all scale-free networks, is almost unaffected by the random failure of nodes. However the failure of hub nodes causes the network to disintegrate into smaller components [11].

3. THREE-TIER TOPOLOGY

Biological studies have revealed that GRN is a hierarchical network and it exhibits scale-free degree distribution [12]. Therefore we propose to study *E. coli* GRN as a three-tier topology based on degree distribution, to understand the dynamics between above-mentioned graph properties.

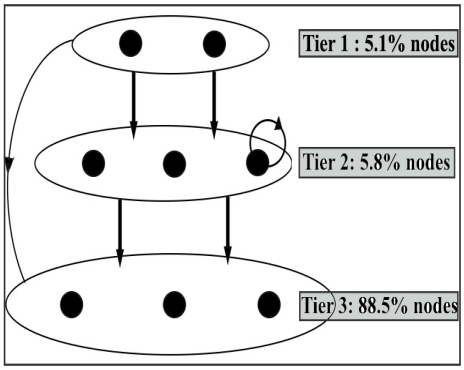

Figure 5: Three-tier network topology of E. coli GRN, where the arrows indicate possible edges between tiers.

Tier 1 consists of all the nodes that have *only out-degree*, **tier 2** is a collection of nodes which *have non-zero in and out-degree*, and finally, there is **tier 3** that consist of nodes with *only in-degree* (Figure 5).

Node distribution

We have observed that out of *1564* nodes in the *E. coli* GRN, only around *5.1 %* lie in tier 1. There are approximately *5.8 %* and *88.5%* nodes in tiers 2 and 3 respectively. Tier 3 nodes therefore are most frequent in *E. coli* GRN.

Edge distribution

The direction of arrows in Figure 5 is an indication of the possible edges between each tier. Only possible edges in the *E. coli* GRN are between **tier 1 → tier 2, tier 1 → tier 3, tier 2 → tier 2 and tier 2 → tier 3**. Self-loops, which are a very small fraction of the edges, are found in tier 2.

Table 1: Percentage of edges between tiers

Tiers	Percentage of edges
Tiers 2 → 3	72.6%
Tiers 1 → 3	21.6%
Tiers 1 → 2	4.6%
Tiers 2 → 2	1.2 %

We have summarized the percentage of edges between each of the tiers in Table 1, which shows that over 70 % of total edges are between tiers 2 and 3. At this point, we characterize the graph-theoretic properties using above topology.

Scale-free property: As evident from Figure 5, tier 1 and tier 2 nodes collectively account for less than *12 %* of total nodes but possess all the out-degree edges. Conversely, tier 3 nodes constitute almost *90 %* nodes in the network but have zero out-degree. Since a few nodes have a disproportionately high out-degree, *E. coli* GRN is out-degree scale-free in nature.

Graph Sparseness: The only possible directed edges in the *E. coli* GRN, exist between tiers $1 \rightarrow 2$, $1 \rightarrow 3$, $2 \rightarrow 3$ and $2 \rightarrow 2$. Tier 3 nodes which account for almost 90 % nodes in the network have no edges among them, explaining why the *E. coli* GRN has a low graph density of 0.0015.

Clustering tendency: Clustering property of a network refers to the tendency of the nodes in a graph to form tightly-knit groups. If t(i) is the number of triangles node i participates in and d(i) is its degree, clustering coefficient of node i, is measured as:

$$C(i) = \begin{cases} 0, & \text{if } d(i) < 2 \\ \frac{2 * t(i)}{d(i) * (d(i)-1))}, & \text{otherwise} \end{cases}$$

Overall clustering coefficient can be estimated from this equation by calculating an average of clustering coefficient of all the nodes. Clearly, the clustering coefficient of a network is a function of the number of triangles in the network.

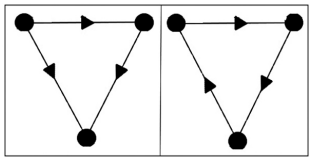

Figure 6: (a) Feed forward loop (FFL): Acyclic triangle (b) Feedback Loop (FBL): Cyclic triangle

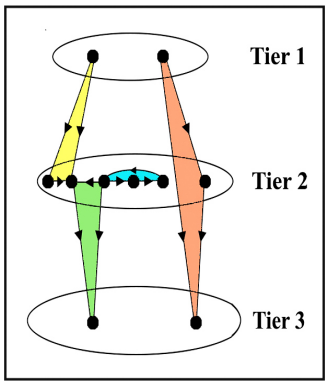

Figure 7: Different colors highlighting some of the possible triangles (based on edges between tiers) in E. coli GRN. For simplicity, possible bidirectional edges in tier 2 have not been shown.

As we have discussed before, the *E. coli* GRN has subnetworks of three or four nodes, called **motifs**. We focus on two types of triangular motifs in GRN: **acyclic triangles (feed forward loop (FFL))** and **cyclic triangles (feedback loop (FBL))**, as shown in Figure 6. These triangular

motifs contribute to the clustering coefficient of *E. coli* GRN.

Now revisiting the possible edge-directions in three-tier topology, it is possible to calculate the number of triangles between each of the tiers. **We can have triangles with the following tier-nodes as endpoints: (tier1, tier2, tier2), (tier1, tier2, tier3), (tier2, tier2, tier2), (tier2, tier2, tier3)**, as demonstrated in Figure 7.

Using this idea, we have estimated the number of cyclic and acyclic triangles in *E. coli GRN* and compared it to the number of triangles in a Erdös-Rényi random graph of 1560 nodes, 3600 edges and probability of edge existence $p = 0.0016$. The triangle-count of the two networks show that the number of triangles in GRN is over 70 times that of the random graph (Table 2). **Consequently, the clustering coefficients of random network is 0.002 and that of *E. coli* GRN is 0.209 (which is more than hundred times of random network with almost the same number of nodes and edges)**.

Table 2: Comparison of the number of triangles in random network and *E. coli* GRN

Network	No. of triangles	Cyclic	Acyclic
E. coli GRN	1404	1401	3
Random Network	19	13	6

Vulnerability to the failure of hub nodes: The *E. coli* GRN topology is very resistant to random failure, but it disintegrates into smaller clusters if well connected hub nodes are knocked off. Again, the three-tier topology shows us that as many as 512 tier 3 nodes are connected to only one hub node in tiers 1 and/or tier 2, which are likely to be isolated from the network when hub nodes are knocked off the *E. coli* GRN (Figure 8(Left)).

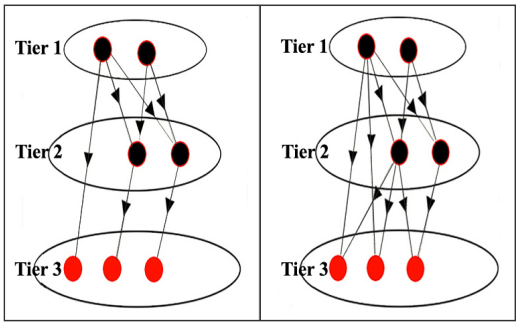

Figure 8: (Left) Before edge rewiring: Some tier 3 nodes connected to one node from tiers 1 or 2 (Right) After edge rewiring: All tier 3 nodes connected to atleast two nodes from tiers 1 or 2

4. EDGE REWIRING

The most important takeaway from the observations on the *E. coli* GRN using three-tier topology is that despite the presence of significant graph properties, the *E. coli* GRN is vulnerable to the failure of hub nodes . In an attempt to overcome this limitations without compromising its strengths, we perform edge rewiring on existing *E. coli* GRN.

edge rewiring is the process of addition and removal of edges in a network, such that the total number of edges in modified network is same as that in original network. In modified *E. coli* GRN, each tier 3 node has in-degree of atleast 2. We must also ensure that modified *E. coli* GRN preserves the graph-theoretic properties of original GRN (Figure 8(Right)).

4.1 Edge addition

We have established that 512 tier 3 nodes in *E. coli* GRN are connected to only one hub node. The objective of edge-addition is to make sure that every tier 3 node is connected to atleast two nodes from tier 1 and/or tier 2, and are therefore less likely to be isolated from the network when hub nodes are knocked off. The edge-addition mechanism is based on the preferential attachment property of GRN.

Preferential attachment growth model: Preferential attachment of any network can be expressed using a growth function f [13]. If d is the average degree of a node, in case of a random network, $f(d) = 1$, implying that the preferential attachment of a node is independent of its degree. On the contrary, for a scale-free network like *E. coli* GRN, $f(d) = d$, so the preferential attachment of a node is proportional to its degree. This is the very principle we use to add edges to all tier-3 nodes with in-degree less than 2.

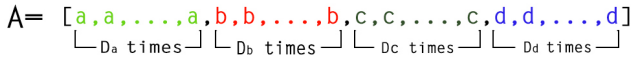

Figure 9: Edge addition: Structure of list A

We propose a linear time edge-addition algorithm. In this approach we define a list A which constitutes all the tier 1 and 2 nodes. Each node i appears D_i times in list A, where D_i is the out-degree of node i (Figure 9).

Algorithm 1 Edge-Addition algorithm

Input: List A which contains node labels, where the frequency of a label is equal to the node out-degree
Result: Every tier 3 node is connected to atleast two nodes from tiers 1 and/or 2.

```
1: procedure EDGE–ADDITION
2:     for each node j ∈ tier 3 do
3:         Choose random element i from list A
4:         Introduce edge (i, j) to E. coli GRN topology
5:     end for
6: end procedure
```

Analysis of edge addition: In course of edge-addition, for every tier 3 node with in-degree less than 2, we introduce a new link between any tier 1 or 2 node randomly chosen from the list A. Edge addition algorithm chooses nodes from list A with uniform randomness. If D_i is the out-degree of node i in list A and D_{sum} is the sum of degrees of all nodes in list A, the probability p_i of node i being the chosen node is $p_i = \frac{D_i}{D_{sum}}$. Probability p_i will be higher for a higher value of D_i. Therefore addition of new edges to *E. coli* GRN topology is based on the preferential attachment property.

Consequences of edge addition: Before edge-addition,

the *E. coli* GRN has 1564 nodes and 3758 edges. Before edge addition, there are 512 tier-3 nodes which have in-degree less than 2. **Therefore, after edge addition there are** 512 **more edges in the *E. coli* GRN and there are no tier-3 nodes with in-degree less than 2.**

4.2 Edge deletion

The edge deletion procedure will ensure that the number of edges in modified *E. coli* GRN is the same as the original network. Therefore the graph density of modified *E. coli* GRN should be same as that of original GRN.

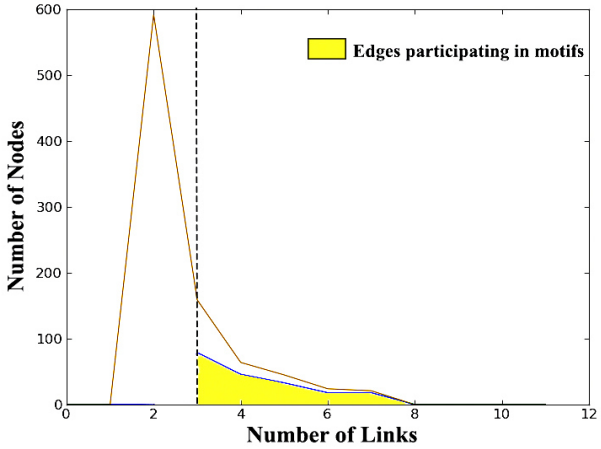

Figure 10: E. coli GRN in-degree distribution: The dashed line indicates the nodes with in-degree $\geq$ 3. The region highlighted (in yellow) shows the edges participating in motifs.

Given the fact that the self-loops in *E.coli* GRN are not consequential, they can be removed. Therefore we have a simple formula to predetermine surplus edges to be removed in edge deletion procedure.

$$|Surplus\ Edges| = |Edges\ added\ | - \ |Self\ loops| \quad (2)$$

Even though we added 512 edges, we only remove 465 edges, because there are 47 self-loops. Now, we have to take stock of a few GRN properties as we carry out edge deletion.

In-degree distribution: Figure 10 shows the in-degree distribution of *E. coli GRN*, where the dashed line demarcates edges associated with nodes having in-degree greater than 2. **The edges available for deletion are on the right of dashed line and are incident to nodes with in-degree more than 2. Deletion of any edge associated with node of in-degree less or equal to 2, will cause that node to be connected to only one hub node**.

Furthermore, some of the edges available for deletion participate in triangular motifs as shaded (in yellow) in Figure 10. If edges are removed from shaded zone, we are likely to lose motifs, that contribute to the clustering tendency of GRN.

Motif-based centrality: Motif based centrality of an edge is the number of motifs each edge participates in. The triangular motifs contribute to the overall clustering coefficient of *E. coli* GRN. Therefore to reduce the loss of motifs during edge-deletion, it is imperative to estimate its motif-based centrality value. In course of edge-removal, our intention is to remove edges with lowest possible motif-based centrality.

If N is the set of nodes, $\forall a, b, c \in N$ and $\Delta(a, b, c)$ represents a triangle with end-nodes a,b and c, we formalize motif-based centrality C of each edge $e(a, b)$ with end-nodes a and b as:

$$C(e(a,b)) = \sum_{\substack{c \in N \\ c \neq a,b}} \Delta(a,b,c) \quad (3)$$

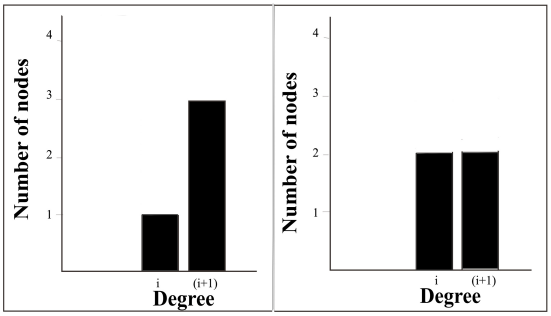

Figure 11: (Left) Before edge removal: There are one and three nodes of degrees i and $(i+1)$;(Right) After removal of one edge of degree $(i+1)$: The number of nodes with degree i increases by 1

Effect of edge-removal on degree-distribution: If a node of in-degree i loses an incoming edge, then its in-degree becomes $(i-1)$. In the in-degree distribution curve, if node count of in-degree i is reduced by 1, the node count for in-degree $(i-1)$ is incremented by 1 as shown in Figure 11.

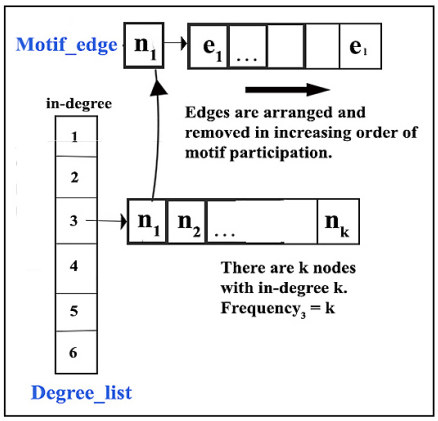

Figure 12: Edge Deletion algorithm

Keeping all these factors in mind, we propose edge deletion algorithm. In this approach, we define a list **Degree list** which consists of all in-degree values (ranging from 0 to maximum in-degree). Each in-degree entry in **Degree list** contains a list of nodes with that in-degree. For every node, we also define a list **Motif edge**, containing all edges incident to that node, arranged in the increasing order of motif participation as shown in Figure 12.

The edge deletion algorithm predetermines the number of edges to be removed from nodes with i^{th} degree, $\forall i > 2$, depending on two factors:

Number of nodes with degree i (defined as f_i): Higher the value of f_i, more edges are available to be removed from i^{th} degree nodes.

Number of nodes with degree $i+1$ (defined as f_{i+1}): As Figure 13 shows, for higher value of f_{i+1}, deletion of edges from $(i+1)^{th}$ degree nodes will in turn increase the frequency of i^{th} degree nodes. Therefore we can formulate the number of edges removed from i^{th} degree nodes T_i as follows:

$$T_i = \frac{a * f_i + (1-a) * f_{i+1}}{Total\ in-degree\ sum} * |Surplus\ edges| \quad (4)$$

In equation 4, scale constant a, which lies between 0 and 1, determines the weightage of f_i and f_{i+1} in calculating T_i.

Algorithm 2 Edge-Deletion algorithm

Input: Graph G, Motif_edge, Degree_list, max_indegree
Output:

```
1: procedure EDGE–DELETION
2:     for i = 3 to max_indegree do
3:         Calculate T(i) using equation 4
4:         for j = 1 to T(i) do
5:             N(j) = Degree_list(i, j)
6:             //N(j) is the j^th node of degree i
7:             Remove first edge e from Motif_edge(N(j))
8:             Remove edge e from G
9:         end for
10:    end for
11: end procedure
```

Analysis of edge deletion algorithm: For every in-degree value i ($0 \leq i \leq max\_indegree$), the algorithm removes the edge with lowest motif participation from T_i nodes of in-degree i. For each of the T_i nodes, edge deletion algorithm removes one edge with lowest motif participation.

Consequence of edge deletion algorithm: As a result of edge deletion, the modified *E. coli* GRN has the same number of edges as original GRN.

5. RESULTS

In subsection 5.1 we discuss and compare the graph-theoretic properties of original and modified *E. coli* GRN, to understand whether edge rewiring preserves the graph properties of *E. coli* GRN. In section 5.2, we discuss the performance of WSN topology designed using both GRN topologies.

5.1 Graph theoretic results

Scale-free property: The modified *E. coli* GRN retains the out-degree scale-freeness of original GRN, as shown in Figure 13.

Network efficiency and clustering coefficient: We now compare the two graph parameters- network efficiency and clustering coefficient of the original and modified GRN.

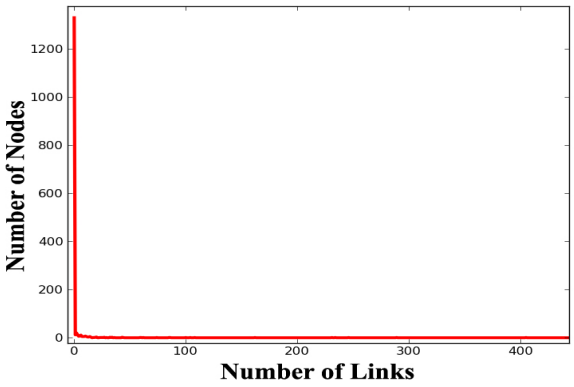

Figure 13: Out-degree distribution of modified E. coli gene regulatory network

Average network efficiency is a measure of how effectively a network exchanges information. For a graph of n nodes, if $d(i,j)$ is the shortest path between a pair of nodes i and j, average efficiency E can be calculated as:

$$E = \frac{2}{n(n-1)} \sum_{i<j}^{n} \frac{1}{d(i,j)} \quad \forall i,j \in n \quad (5)$$

As Table 3 shows, the efficiency of *E. coli* GRN is 0.0064, and the modified GRN has efficiency of 0.0086.

Table 3: Comparison of graph theoretic properties

Parameters	Original GRN	Modified GRN
Sparseness	0.0015	0.0015
Clustering Coefficient	0.209	0.320
Network Efficiency	0.0064	0.0086

Clustering coefficient of node i, $C(i)$ can also be defined as:

$$C(i) = \frac{|\text{Triangles connected to node i}|}{|\text{Triples centered around node i}|} \quad (6)$$

In equation 6, a triple is set of two edges connected to node i. In the modified *E. coli* GRN we have 1217 triangles whereas the original network has 1404 triangles, however the clustering coefficient of the modified network is 0.32 while that of original *E. coli* GRN is 0.209. We can attribute this increase to the drop in the number of triples in modified *E. coli* GRN.

Random node and hub node failure: We study the robustness of the original and modified *E. coli* GRN. For original and modified *E. coli* GRN, we knock off (i) 10 nodes randomly and (ii) 10 hub nodes with out-degree more than 20. For both networks, we compare the number of isolated nodes as a result of random and hub node failure.

The failure of 10 *E. coli* GRN nodes causes the isolation of 37 nodes in the original network, whereas in the modified network only 15 nodes are lost (Figure 14). When 10 hub nodes (i.e. nodes with out-degree greater than 20) are knocked off, 137 nodes are knocked off the network in the original GRN, while in the modified GRN we lose 68 nodes

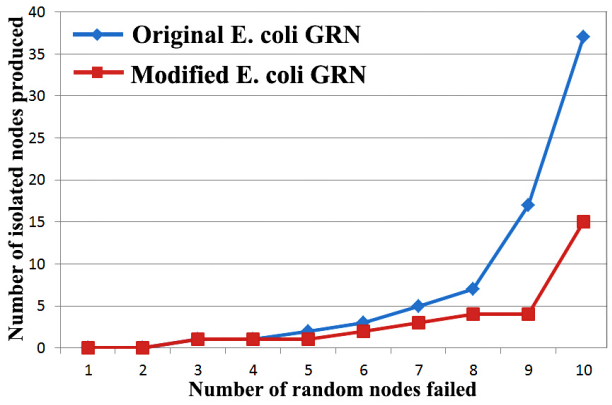

Figure 14: Random node failure: Original vs. E. coli modified GRN

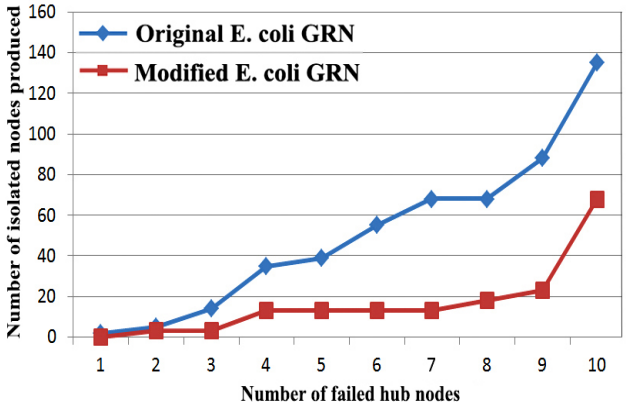

Figure 15: Hub node failure: original vs. modified GRN

(Figure 15). In both cases, number of isolated nodes in case of modified *E. coli* GRN is less than 50% of original GRN.

5.2 Deployment of wireless sensor network

As we have pointed out, the *E. coli* GRN has been utilized in the past as a model for the design of WSN [9]. We compare the performance of WSN designed using original and modified GRN, with and without node failure.

Topological properties and simulation tools: As evidenced by the three-tier topology, *E. coli* GRN is a hierarchical network. Therefore it lends itself to the design of WSN where data is transferred from source to sink nodes. We have implemented WSN topology for original and modified GRN on OMNET++ simulator using the Collection Tree Protocol (CTP), which is a distance vector routing protocol specifically designed for sensor network communication [14].

We know that the tier-3 nodes constitute almost 90% of the entire network, therefore the edge directions have been reversed to make tier-3 nodes the source nodes and the tier-1 nodes the sinks (with out-degree now equal to 0 due to edge reversal). Only 5 % of total nodes in each network, are sinks.

Conditions of simulation experiment: We have extracted 5 *E. coli* GRN subnetworks of 400 nodes each, using the software called GeneNetWeaver [15]. For both our experiment, we have also generated 5 Erdös Rényi random graphs of 400 nodes each, with probability of existence of edge $p = 0.004$. Nodes have been deployed uniformly over an area of 50×50 square metres and simulation duration has been varied from 450 to 1800 seconds. Each experiment has been repeated 5 times and the average score for each experiment has been plotted.

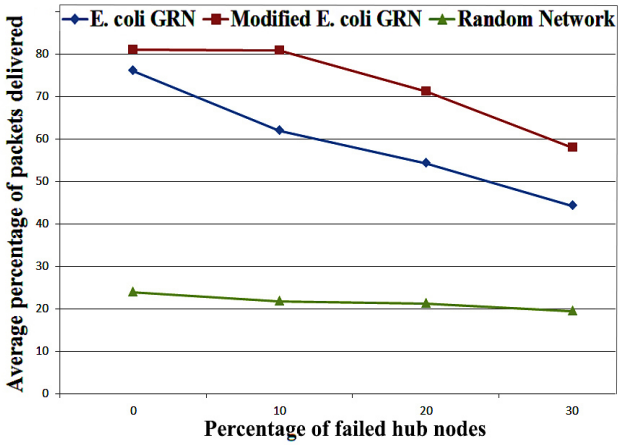

Figure 16: Experiment I: Comparison of average number of packets delivered per node for varying failure of hub nodes

Experiment-I: We know that average packet delivery and latency are effective performance metrics for communication in networks. The first experiment calculates the percentage packet delivery of random network, *E. coli* GRN and modified *E. coli* GRN. We have simulated failure of 0%, 10%, 20% and 30% hub nodes. Duration of simulation is 1800 seconds.

The results show that modified *E. coli* GRN topology demonstrates highest packet delivery percentage of the three networks. Also, the decline in packet delivery rate in modified GRN is significantly less than the original GRN, as shown in Figure 16.

Experiment-II: In the second experiment, we estimate the average number of packets dropped and latency per node for simulation duration ranging from 450 to 1800 seconds, for random network, original and modified *E. coli* GRN. with a fixed failure of 10 % of hub nodes.

The results show that the number of packets dropped per node is the lowest for modified *E. coli* GRN. Also, the gap between the number of packets dropped for the three networks widened gradually as time progresses as we see in Figure 17. We also observe that even for a fixed failure of 10 % hub nodes, the rate of packet loss increases with time.

For the same set-up, we have calculated the average latency in packet delivery for the random network, original and modified *E. coli* GRN. The results show that the average latency

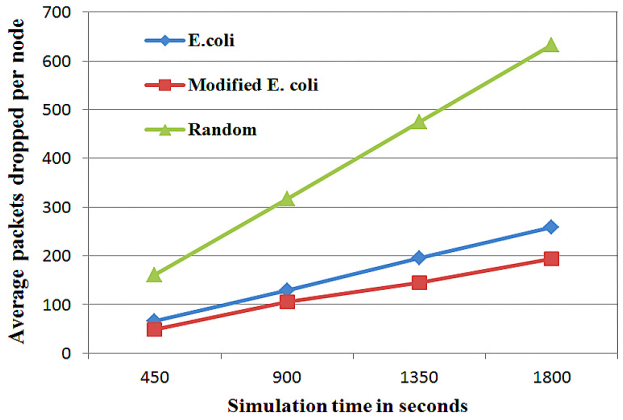

Figure 17: Experiment II: Comparison of average number of packets dropped for random network, original and modified E. coli GRN

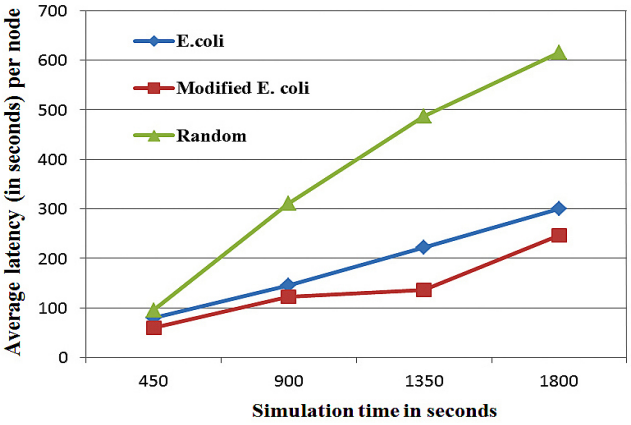

Figure 18: Experiment II: Comparison of average latency in packet delivery for random network, original and modified E. coli GRN

in packet delivery for the modified *E. coli* GRN is consistently lower than the other two networks (Figure 18).

6. CONCLUSION

In this work we characterize the *E. coli* GRN topology and enhance its robustness against hub node failure, without compromising its existing topological attributes. Results show that modified *E. coli* GRN exhibits improvement in topological properties and WSN performance metrics over original GRN. Thus it is our intuition that modified GRN could be an effective model for design of robust communication networks including WSN.

In the future, we would like to study average packet delivery, latency and energy consumption of original and modified E. coli GRN for WSNs of different sizes. Our objective will be to analyse how node failure affects the performance for varying size of networks. Also, we want to perform extensive analysis of the effects of edge addition and/or deletion on the overall topological robustness of GRN, and come up with mathematical formulation for the same.

7. ACKNOWLEDGMENTS

Satyaki Roy and Vijay K. Shah are grateful to Nitish Uplavikar, Dr. Mayank Raj and the members of CReWMaN for their constant support.

8. REFERENCES

[1] Nasimul Noman, Taku Monjo, Pablo Moscato, and Hitoshi Iba. Evolving robust gene regulatory networks. *PloS one*, 10(1), 2015.

[2] Uri Alon. *An introduction to systems biology: design principles of biological circuits*. CRC press, 2006.

[3] Nabil Guelzim, Samuele Bottani, Paul Bourgine, and François Képès. Topological and causal structure of the yeast transcriptional regulatory network. *Nature genetics*, 31(1):60–63, 2002.

[4] Agustino Martínez-Antonio. Escherichia coli transcriptional regulatory network. *Network Biology*, 1(1):21–33, 2011.

[5] Vitor HP Louzada, Fabio Daolio, Hans J Herrmann, and Marco Tomassini. Generating robust and efficient networks under targeted attacks. In *Propagation Phenomena in Real World Networks*, pages 215–224. Springer, 2015.

[6] Yang Yang, Zhoujun Li, Yan Chen, Xiaoming Zhang, and Senzhang Wang. Improving the robustness of complex networks with preserving community structure.

[7] S Xiao, G Xiao, TH Cheng, S Ma, X Fu, and H Soh. Robustness of scale-free networks under rewiring operations. *EPL (Europhysics Letters)*, 89(3):38002, 2010.

[8] Nour El-Mawass, Nada Chendeb, and Nazim Agoulmine. Robust self-organized wireless sensor network: A gene regulatory network bio-inspired approach. In *Genetic and Evolutionary Computing*, pages 105–114. Springer, 2014.

[9] Azade Nazi. Robust design of wireless sensor networks using gene regulatory networks.

[10] Reka Albert. Scale-free networks in cell biology. *Journal of cell science*, 118(21):4947–4957, 2005.

[11] Attila Gursoy, Ozlem Keskin, and Ruth Nussinov. Topological properties of protein interaction networks from a structural perspective. *Biochemical Society Transactions*, 36(6):1398, 2008.

[12] Douglas H Erwin and Eric H Davidson. The evolution of hierarchical gene regulatory networks. *Nature Reviews Genetics*, 10(2):141–148, 2009.

[13] Jérôme Kunegis, Marcel Blattner, and Christine Moser. Preferential attachment in online networks: measurement and explanations. In *Proceedings of the 5th Annual ACM Web Science Conference*, pages 205–214. ACM, 2013.

[14] Omprakash Gnawali, Rodrigo Fonseca, Kyle Jamieson, David Moss, and Philip Levis. Collection tree protocol. In *Proceedings of the 7th ACM Conference on Embedded Networked Sensor Systems*, pages 1–14. ACM, 2009.

[15] Thomas Schaffter, Daniel Marbach, and Dario Floreano. Genenetweaver: in silico benchmark generation and performance profiling of network inference methods. *Bioinformatics*, 27(16):2263–2270, 2011.