CHAPTER 5

Investigation of Control Configuration in Biological Networks

5.1 Introduction and motivation

Biological processes are governed by physical and chemical interactions between proteins. Cells respond to environmental changes that influence the signalling network. For instance, in order to drive the human cancer signalling network from a cancerous state to a healthy state, a natural conversion of a cancer cell into a normal cell is not possible. This can be achieved by perturbing the system using drugs or other external hues.

The maximum matching model developed to identify driver nodes, nodes with which we can achieve full control, predict the existence of multiple control configurations, in turn prompting us to classify each node in the network based on its role in control. In Chapter 2, we classified a node as critical, ordinary, redundant/non-driver according as its presence in all, some, none of the sets of driver nodes. This classification of driver nodes has lead to identification of two distinct control modes/configurations for a network. The control mode of a network can be altered through small structural perturbations [36]. In this chapter we perturb the network configuration, classify the driver nodes in the perturbed network to gain insights into the dynamics of the system represented by the initial network. We use this technique to identify the control modes of some biological networks.
5.2 Control modes in networks

The role of each node in controlling a network is identified by classifying each node into one of the three categories based on its presence in minimum driver node set (MDS). Critical, meaning the node is always a driver node, i.e. it is present in all MDS; redundant if it never acts as a driver node and therefore not a part of any MDS; ordinary if it acts as driver node in some but not all MDS. This classification leads to bifurcation phenomenon, predicting that a bimodal behaviour determines the controllability of many real world networks [36]. This bimodality uncovers two control modes, centralized and distributed. The control modes are determined by fraction of redundant nodes.

Studies on network models have shown that for networks with symmetric in- and out-degree distributions, the fraction of redundant nodes \( n_r = N_R/N \), where \( N_R \) is number of redundant nodes and \( N \) is total nodes in the network) undergoes a bifurcation at a critical mean degree \( \langle k \rangle_c \). In particular it has been shown that for low mean degree \( \langle k \rangle \) \( n_r \) is uniquely determined by \( \langle k \rangle \), but when \( \langle k \rangle \) exceeds \( \langle k \rangle_c \) there exist two different values for \( n_r \), one with very high and the other with low values leading to bimodal behaviour (Figure 5.1). Hence for large \( \langle k \rangle \) two control modes coexist [36]. These modes are:

1. Centralized control: For networks that follow the upper branch of the bifurcation diagram (Figure 5.1), most nodes are redundant nodes (Figure 5.1(c)) and \( n_r \) is very high. In these networks, only a small fraction of nodes are required for control, i.e., \( n_c + n_i \) is small, where \( n_c \) is the fraction of critical nodes and \( n_i \) is the fraction of ordinary nodes in the network. This mode of control is like that of an organisational setting where leadership is concentrated in the hands of a few supervisors and the employees are only executors.

2. Distributed control: In networks that follow the lower branch of the bifurcation diagram \( n_c + n_i \) can exceed 90% of the nodes, meaning most nodes participate as driver nodes in some control configuration. This mode of con-
Figure 5.1: Emergence of bimodality in controlling complex networks. (a) Bimodality in network for symmetric in- and out-degree distribution ($\gamma_{in} = \gamma_{out} = 3$) for high $\langle k \rangle$. The plot is of $n_r$ and $n_c$ vs $\langle k \rangle$ in scale-free networks. (b) $n_r$ in networks with asymmetric degree distribution ($\gamma_{in} = 2.67, \gamma_{out} = 3$, upper branch and $\gamma_{in} = 3, \gamma_{out} = 2.67$ lower branch). The control mode is predetermined by their degree asymmetry. (c) and (d) are networks representing centralized or distributed control. Both networks have $N_D = 4$ and $N_c = 1$ (red node) but different number of redundant nodes (uncoloured node), $N_r = 23$ in (c) and $N_r = 3$ in (d). From Y.Liu et al., 2016.

Control is like that of an organisation where different employees take leadership roles at different times as in case of change in shift.

5.3 Identification of control modes in biological networks

Identification of control mode and altering the state of a network can be achieved by its transpose. The transposed network of a given network is obtained by re-
versing the direction of each link of the given network. The control mode is captured by comparing \( n_r \) of the network with that of its transposed network \( n_r^T \). If \( \Delta n_r = n_r - n_r^T > 0 \) then the network is said to be in centralized mode. If \( \Delta n_r < 0 \) then the network is said to be in distributed and if \( \Delta n_r = 0 \) then the mode of control of the network cannot be determined. For some biological networks we have determined the control mode as tabulated below (Table 5.1).

<table>
<thead>
<tr>
<th>Network</th>
<th>Nodes</th>
<th>Edges</th>
<th>( n_r )</th>
<th>( n_r^T )</th>
<th>( \Delta n_r )</th>
<th>Control mode</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cancer signaling</td>
<td>1232</td>
<td>3060</td>
<td>0.396104</td>
<td>0.350649</td>
<td>0.045455</td>
<td>centralized</td>
</tr>
<tr>
<td>Directed human PPI</td>
<td>6339</td>
<td>34813</td>
<td>0.41505</td>
<td>0.425304</td>
<td>-0.010254</td>
<td>undetermined</td>
</tr>
<tr>
<td>HIV-human molecular</td>
<td>6361</td>
<td>40625</td>
<td>0.416352</td>
<td>0.427449</td>
<td>-0.011097</td>
<td>undetermined</td>
</tr>
<tr>
<td>T-cell activation</td>
<td>121</td>
<td>255</td>
<td>0.520661</td>
<td>0.528926</td>
<td>-0.008265</td>
<td>undetermined</td>
</tr>
<tr>
<td>HIV-T-cell activation</td>
<td>137</td>
<td>367</td>
<td>0.49635</td>
<td>0.613139</td>
<td>-0.116789</td>
<td>distributed</td>
</tr>
<tr>
<td>E. coli transcription</td>
<td>423</td>
<td>578</td>
<td>0.096926</td>
<td>0.269504</td>
<td>-0.172578</td>
<td>distributed</td>
</tr>
</tbody>
</table>

Table 5.1: Control mode of some biological networks

The human signalling network follows a centralized mode of control. One possible reason could be the nature of the network. In this network, cells receive information from neighbouring cells through receptors which in turn activates or inhibits proteins down stream of it. In this way many proteins coordinate with one another to maintain cellular information processing. Since there is a kind of hierarchy among the proteins, this could be a reason for why such systems follow a centralized mode of control.

The HIV-1 t-cell activation and the E. coli transcription networks have distributed control. In the case of HIV-1 t-cell activation network, the HIV-1 proteins interact with human proteins for its replication. Further due to infection, the human proteins independently activate the immune system to combat the virus, thus allowing for a distributed mode of control. So is the case with the E. coli transcription network which comprises motifs that have a specific function in determining different gene expressions. In the case of the directed human PPI network, HIV-1 human molecular network and T-cell activation network, the mode of control cannot be determined.
5.4 Conclusion

Identifying the control mode of a network throws light on its dynamics. This tells us the ease with which one can control such systems. Biological networks are in general difficult to control due to their underlying complexity. By identifying the mode, we can perhaps find ways to effectively control the system. Further, identifying the control mode of a biological network will add to better understanding of mechanisms underlying its nature. For most biological networks, many molecular details like function, interactions, kinetic parameters among genes and proteins are unknown. Despite the lack of a thorough knowledge of the entities in the network, the control mode can shed some light on the mechanisms of the network.

The fact that control modes are suited for different task raises an important question as to what is the least modification required to change the mode of a network from centralized to distributed and vice versa? It is to be noted that if the original network is centralised then its transpose is distributed [36]. It is of interest to identify as few edges as possible that when reversed can change the mode of control. Biological processes can now be reversed at a cellular level. For instance tumorigenesis can be reversed through targeted inactivation of oncogenes [100]. Study by Cho et.al. have shown that the network can be rewired by inducing reshaping of the attractor landscape (a molecular phase portrait describing the dynamics of a molecular regulatory network) and phenotype landscape (that is determined by the steady states of particular output molecules) in the attractor landscape [101]. One should bear in mind that, the purpose of control determines the control mode best suited for the system.