Wu L, Li M, Wang JX *et al.* Controllability and its applications to biological networks. JOURNAL OF COMPUTER SCIENCE AND TECHNOLOGY 34(1): 16–34 Jan. 2019. DOI 10.1007/s11390-019-1896-x

Controllability and Its Applications to Biological Networks

Lin Wu[1], Min Li[2], Jian-Xin Wang[2], and Fang-Xiang Wu[1,2,3,*], *Senior Member*, *IEEE*

[1] *Division of Biomedical Engineering, University of Saskatchewan, Saskatoon, SK S7N 5A9, Canada*

[2] *School of Information Science and Engineering, Central South University, Changsha 410083, China*

[3] *Department of Mechanical Engineering, University of Saskatchewan, Saskatoon, SK S7N 5A9, Canada*

E-mail: liw557@mail.usask.ca; {limin, jxwang}@csu.edu.cn; faw341@mail.usask.ca

Received July 1, 2018; revised November 8, 2018.

Abstract Biological elements usually exert their functions through interactions with others to form various types of biological networks. The ability of controlling the dynamics of biological networks is of enormous benefits to pharmaceutical and medical industry as well as scientific research. Though there are many mathematical methods for steering dynamic systems towards desired states, the methods are usually not feasible for applying to complex biological networks. The difficulties come from the lack of accurate model that can capture the dynamics of interactions between biological elements and the fact that many mathematical methods are computationally intractable for large-scale networks. Recently, a concept in control theory — controllability, has been applied to investigate the dynamics of complex networks. In this article, recent advances on the controllability of complex networks and applications to biological networks are reviewed. Developing dynamic models is the prior concern for analyzing dynamics of biological networks. First, we introduce a widely used dynamic model for investigating controllability of complex networks. Then recent studies of theorems and algorithms for having complex biological networks controllable in general or specific application scenarios are reviewed. Finally, applications to real biological networks manifest that investigating the controllability of biological networks can shed lights on many critical physiological or medical problems, such as revealing biological mechanisms and identifying drug targets, from a systematic perspective.

Keywords biological network, network controllability, steering node

1 Introduction

Biological systems are composed of biological elements that can interact with one another. The structure of biological systems can be described by biological networks in which nodes are biological elements and edges connect biological elements that have interactions. Biological processes, which are vital for living organisms to live, are usually carried out by complicated interactions among a variety of biological elements. Therefore, studying biological elements and their interactions is critical for understanding the roles of biomolecules within cells and uncovering the mechanisms of biological processes. With the development of biomedical techniques, such as high throughput technologies and MRI (magnetic resonance imaging), various types of biological data have been acquired in a large amount, which benefits the reconstructions of different types of biological networks[1−4]. Many efforts have been devoted to excavating underlying relationships among biological elements based on the topology of biological networks and substantial progresses have been made, such as drug target identification[5,6], human disease gene prediction[7], and protein complex identification[8].

One of the final goals for investigating a network is to control its behaviour or state. For biological networks, acquiring the ability of controlling their behaviours implies the capability of changing phenotypes of biological systems as desired, which is vital for im-

Survey

Special Section of NSFC Joint Research Fund for Overseas Chinese Scholars and Scholars in Hong Kong and Macao 2014–2017

The work was supported by the Natural Sciences and Engineering Research Council of Canada (NSERC), the National Natural Science Foundation of China under Grant Nos. 61772552 and 61622213, and Chinese Scholarship Council (CSC).

*Corresponding Author

proving human lives. Control theory is a relatively well established subject in engineering dealing with the control of dynamic systems. In 1960s, Kalman pioneered the state-space approach to systems and introduced the notions of controllability and observability[9], which have become the bases of modern control theory. In the state-space representation, individual nodes have their own state variables which have specific physical meanings (e.g., gene expression levels in gene regulatory networks). Because of the interactions among nodes in a network, actuating the states of some nodes can affect other nodes, which may change the state of the network. Controllability, measuring the ability to steer a network around in its state space by actuating certain steering nodes, considers whether it is possible to achieve certain control objectives by actuating the determined steering nodes. Therefore, understanding controllability is critical in the implementation of controlling networks. It should be mentioned that the problems related to the actual control processes, such as the determination of amplitude and frequency of control signals, control time or control trajectories, are not related to controllability, but in the scope of control theory.

In this review we focus on the controllability of complex networks. Several fundamental questions are raised naturally for the controllability of biological networks. 1) How to model the dynamics of complex biological networks? 2) To what extent the topology of biological networks is related to the dynamics and controllability? 3) How to select a set of steering nodes in order to steer a network to a desired state? In this review, steering nodes refer to the nodes which should be directly actuated by input control signals in different control scenarios (e.g., complete controllability, output controllability, and transittability). 4) How to select steering nodes under realistic constraints? 5) What information can we obtain from biological networks based on network controllability? To answer these questions, we review recent studies on the controllability of complex networks. In addition, applications of network controllability to biological networks are discussed specifically, though the developed network control methods can be applied to other complex networks.

Fig.1 shows the topics and contents of the following sections, which are based on the proposed questions and flow for analyzing the controllability of biological

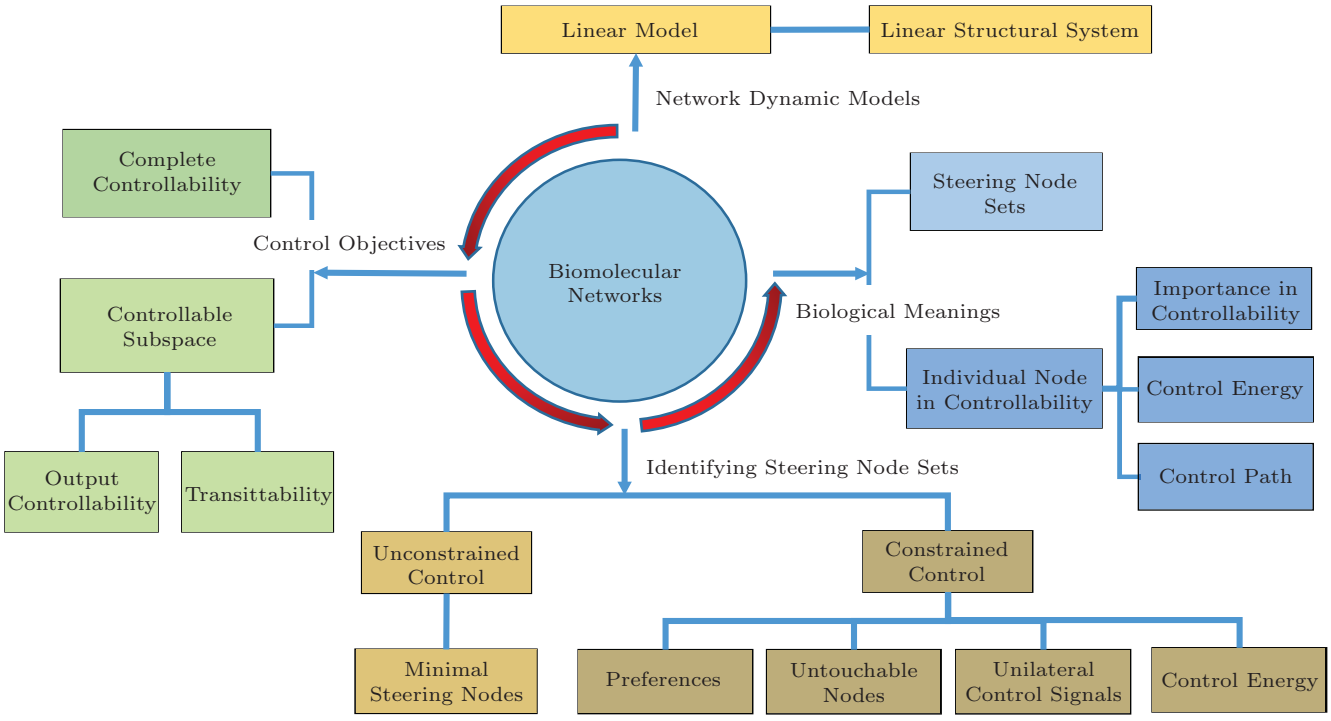


Fig.1. Contents in the article and flow of analyzing controllability of biological networks. For the dynamic models, we first introduce dynamic models for representation of biological networks. We illustrate the reasons of focusing on the linear dynamic model in this article. Then we introduce control theorems related to the complete controllability and controllable subspace of complex networks. Based on the control theorems, methods for identifying steering node sets for different control objectives are discussed. Finally, we review studies of biological networks from the perspective of network controllability. Biological meanings of specific steering node sets and individual nodes that play different roles in network controllability are discussed.

networks. To address question 1, in Section 2, we introduce the linear dynamic models and their corresponding structural systems. Section 3 introduces the network controllability theorems based on the structural controllability of structural linear dynamic models[10]. Network topology and dynamic rules between network elements are two fundamental factors for fully understanding the controllability of networks[11]. However, dynamic rules between biological elements are usually unavailable or unreliable due to the limits of knowledge and the difficulties of estimating model parameters. Therefore, focusing on the structurally controllability, which investigates the controllability based on the network topology, could give an answer to question 2. Section 4 introduces the proposed methods to identify steering node sets for different control objectives, which provide solutions to question 3 and question 4. In Section 5, recent studies on the controllability of biological networks are reviewed, which correspond to question 5. Applications demonstrate that analyzing the controllability of biological networks provides a novel perspective for many biological questions, such as inferring disease associated pathways[12], identifying drug targets[13], and predicting functions of biological elements[14].

2 Dynamic Model of Biological Networks

To understand the controllability of biological networks, it is important to make clear the dynamic models. In this section, we present the linear dynamic models and the corresponding structural systems of biological networks which are widely used in analyzing controllability of complex networks.

2.1 Linear Model

In a dynamic network, each node has their own state value. Assuming that the state change rate of one node is a linear combination of the states of nodes pointing to it, we can represent the dynamics of the network by a set of linear equations. Although the dynamics of biological systems are nonlinear, linear models have been applied to describe the dynamics of many biological networks such as gene regulatory networks[15]. To study the controllability of biological networks, it is reasonable to represent the dynamics of biological networks by linear dynamic models. First, there are a large number of tools available from control theory to study systems with linear dynamics. For example, a sufficient and necessary condition for the controllability of general linear systems has been developed by Kalman[9]. Second, the controllability of nonlinear systems is structurally similar to that of linear systems in many aspects. If a network is structurally controllable, then it is controllable for almost all possible parameter realizations[10]. Therefore, the structural controllability of linear system can provide a sufficient condition for the controllability of most nonlinear systems[11,16]. Actually, to develop strategies for controlling nonlinear networks, the first step is to investigate the controllability of the locally linearized system[17]. Last but not the least, there is an intuitive connection between the network topology and the state transition matrix of a linear dynamic model, which makes it possible to create dynamic models for large-scale biological networks based on their topology. Owing to the strong correlations between the linear model and the network topology, studying on the linear dynamic model can provide a vision of a previously proposed question, which is how much the controllability of biological networks is related to their topological features.

According to reasons discussed above, in this article, we will focus on the controllability of linear systems and biological networks represented by linear dynamic models. For a linear time-invariant network with n nodes, the dynamics can be described by the equation:

$$\dot{\boldsymbol{x}}(t) = \boldsymbol{A}\boldsymbol{x}(t) + \boldsymbol{B}\boldsymbol{u}(t), \tag{1}$$

where $\boldsymbol{x}(t) = (x_1(t), ..., x_n(t))^{\mathrm{T}} \in \mathbb{R}^n$ is an n-dimensional vector that describes the states of all n nodes in the network. $\boldsymbol{A}$ is an $n \times n$ state transition matrix, whose structure is determined by the adjacent matrix of the network, indicating the regulatory relationships between nodes in the network. Entry a_{ij} in matrix $\boldsymbol{A}$ indicates the intensity of influence from node j to node i. $\boldsymbol{u}(t)$ is an m-dimensional vector of m independent input control signals. The $n \times m$ matrix $\boldsymbol{B}$ is an input matrix indicating nodes which are directly actuated by input control signals. A network system described by (1) is denoted as system $(\boldsymbol{A}, \boldsymbol{B})$.

2.2 Structural System and Graph Representation of Linear Systems

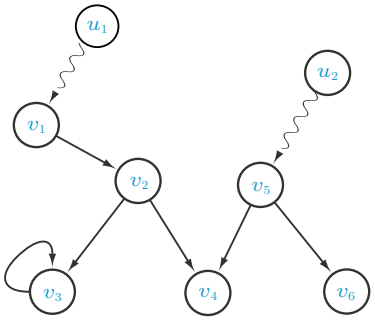
When modeling the dynamics of complex networks, the nonzero entries in matrix $\boldsymbol{A}$ indicate the strengths of relationships between nodes in the networks. However, in many scenarios, it is not feasible to obtain the values of nonzero entries in matrix $\boldsymbol{A}$ precisely. For example, although it is feasible to qualify whether there is

a regulatory relationship between two nodes in biological networks, it is difficult to quantify the intensity of the regulation. In addition, though Kalman's controllability rank condition has been proposed to test the controllability of linear systems[9], calculating the rank of controllability matrix of a large-scale network is computationally intractable. Therefore, it is difficult to test the controllability of a network directly by Kalman's controllability theorem. To address these issues, recent studies on the controllability of complex networks are mainly based on the framework of structural systems, which was proposed by Lin in 1974[10]. In Lin's study, the controllability of structural systems was studied and the sufficient and necessary condition for completely structural controllability of structural systems was given. Lin's result has been proved in different ways[18−21] and generalized to controllable subspaces[22−24].

When entries in matrices $\boldsymbol{A}$ and $\boldsymbol{B}$ are either fixed zero or independent free parameters, matrices $\boldsymbol{A}$ and $\boldsymbol{B}$ are called structural matrices and the corresponding system $(\boldsymbol{A}, \boldsymbol{B})$ is called a structural system. A structural system $(\boldsymbol{A}, \boldsymbol{B})$ is called completely structurally controllable if the Kalman's controllability condition can be satisfied by freely choosing the values of the independent free parameters in matrices $\boldsymbol{A}$ and $\boldsymbol{B}$[10]. Besides completely structural controllability, structural output controllability[24] and structural transittability[25] have been studied, respectively. In this review, the controllability of structural systems (such as completely structural controllability, structural output controllability, and structural transittability) is referred to as structural controllability.

The rationale of investigating the controllability of networks based on the structural system comes from two aspects. First, the structural linear dynamic model of a network can be created only based on its topology and each nonzero entry in $\boldsymbol{A}$ corresponds to an edge in the network. Therefore, for modeling biological networks, there is no need to consider the types of biological networks, the kinetic models regulating the dynamics as well as plenty of unknown parameters. Second, if a structural system $(\boldsymbol{A}, \boldsymbol{B})$ is structurally controllable, most of its parameter realizations which are denoted as admissible systems $(\widetilde{\boldsymbol{A}}, \widetilde{\boldsymbol{B}})$ are controllable, where $(\widetilde{\boldsymbol{A}}, \widetilde{\boldsymbol{B}})$ can be obtained by assigning some specific values to the free parameters of $(\boldsymbol{A}, \boldsymbol{B})$[10]. Therefore, if a network is structurally controllable, no matter how to choose the values of unknown regulatory strengths, the probability that the network is controllable is almost 100%, except some cases that the unknown regulatory strengths satisfy some constraints (equations). Therefore, structural controllability analyses can provide reliable results for real networks even though their parameters are unknown. When a structural network is controllable with all combinations of values assigned to the free parameters, the network is called strong structurally controllable[26]. However, studies based on strong structural controllability are usually less computational efficient than methods based on structural controllability, which are not suitable for investigating large-scale complex networks[27]. Therefore, we mainly focus on applications to biological networks in the sense of structural controllability in this review.

Each structural system $(\boldsymbol{A}, \boldsymbol{B})$ can be represented by a digraph $G(\boldsymbol{A}, \boldsymbol{B}) = \{V, E\}$, where $V = V_A \bigcup V_U$ is a node set and E is an edge set. Nodes in $V_A = \{v_1, ..., v_n\}$ correspond to nodes in the network under investigation and nodes in $V_U = \{u_1, ..., u_m\}$ correspond to the input control signals represented by $\boldsymbol{u}(t)$. $E = \{v_j \to v_i, u_k \to v_l | a_{ij} \neq 0,\ b_{lk} \neq 0\}$ consists of edges among nodes and edges from control signals to nodes. The subgraph of $G(\boldsymbol{A}, \boldsymbol{B})$ induced by the node set V_A is denoted as $G(\boldsymbol{A})$, which is the original network without input control signals. Fig.2 is an example of the system $(\boldsymbol{A}, \boldsymbol{B})$ and its graph representation.



(a)

$$\boldsymbol{A} = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 \\ a_{21} & 0 & 0 & 0 & 0 & 0 \\ 0 & a_{32} & a_{33} & 0 & 0 & 0 \\ 0 & a_{42} & 0 & 0 & a_{45} & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & a_{65} & 0 \end{pmatrix}, \quad \boldsymbol{B} = \begin{pmatrix} b_{11} & 0 \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \\ 0 & b_{52} \\ 0 & 0 \end{pmatrix}$$

(b)

Fig.2. Graph representation of a network system. (a) $G(\boldsymbol{A}, \boldsymbol{B})$ corresponds to system $(\boldsymbol{A}, \boldsymbol{B})$. (b) State transition matrix $\boldsymbol{A}$ and input matrix $\boldsymbol{B}$ of the system $(\boldsymbol{A}, \boldsymbol{B})$.

With graph representation of structural systems, algebraic structural controllability conditions can be

converted to graph-theoretic forms. Therefore, various graph-theoretic algorithms can be applied to investigate the structural controllability of complex networks, which is more computationally feasible for large-scale networks compared with algebraic methods. Taking the advantages of structural controllability, Liu *et al.*[11] recently have applied the concept of completely structural controllability to complex networks. Liu *et al.*'s research inspired several research progresses on network controllability, such as the controllability of networks with edge dynamics (each edge has a state variable)[28], nodal dynamics (each node has a self-loop) controllability[29], robustness of network controllability[30] and enhancing network controllability (reducing input control signals) via minimal structural perturbations[31] have also been studied. In addition, a comprehensive platform has been developed for analyzing the controllability of complex networks[32].

3 Network Controllability Theorems

Since linear systems have been deeply studied in control theory, a variety of controllability theorems have been proposed, which paves the way to understand the controllability of complex networks. In addition, taking the advantages of structural controllability, connections between the network topology and controllability can be established. In this section, we introduce some important theorems of network controllability.

3.1 Complete Controllability of Complex Networks

A network is completely controllable if it can be steered from any initial state $\boldsymbol{x}(0)$ to any desired final state $\boldsymbol{x}(t_f)$ in finite time t_f with appropriate control signals. The condition for complete controllability is given by the following theorem.

Theorem 1 (Kalman's Controllability Theorem[9]). *System* $(\boldsymbol{A}, \boldsymbol{B})$ *is completely controllable if and only if the* $n \times nm$ *controllability matrix*

$$\mathfrak{C} = (\boldsymbol{B} \quad \boldsymbol{AB} \quad \boldsymbol{A}^2\boldsymbol{B} \quad \ldots \quad \boldsymbol{A}^{n-1}\boldsymbol{B}),$$

has full row rank of n.

To interpret this criterion, (1) can be solved in the following form:

$$\boldsymbol{x}(t) = e^{\boldsymbol{A}t}\boldsymbol{x}(0) + \int_0^t e^{\boldsymbol{A}(t-\tau)}\boldsymbol{B}\boldsymbol{u}(\tau)\mathrm{d}\tau. \tag{2}$$

On the right-hand side of (2), the first term corresponds to the state that the network will be without any control signals and the second term represents the effect of control signals on the network. $e^{\boldsymbol{A}(t-\tau)}\boldsymbol{B}$ can be expanded in series, which is a linear combination of the columns in controllability matrix $\mathfrak{C}$. When a network is completely controllable, the final state $\boldsymbol{x}(t_f)$ could be any state in the n-dimensional state space. On the one hand, if $rank(\mathfrak{C}) < n$, columns in $\mathfrak{C}$ will not contain a full basis to span the entire n-dimensional state space (see Fig.3). Then there exist some final states $\boldsymbol{x}_{t_f}$, such that by letting $\boldsymbol{x}(t_f) = \boldsymbol{x}_{t_f}$, (2) has no solution for $\boldsymbol{u}$. On the other hand, if $rank(\mathfrak{C}) = n$, columns in $\mathfrak{C}$ contain a full basis. Given any desired final state $\boldsymbol{x}_{t_f}$ and let $\boldsymbol{x}(t_f) = \boldsymbol{x}_{t_f}$, an appropriate input vector $\boldsymbol{u}$ can always be solved based on (2). Therefore, the system is completely controllable.

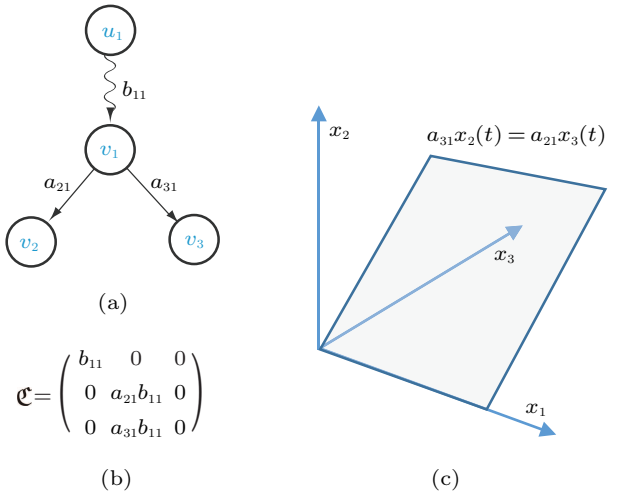


Fig.3. Uncontrollable network system $(\boldsymbol{A}, \boldsymbol{B})$. (a) $G(\boldsymbol{A}, \boldsymbol{B})$ corresponds to system $(\boldsymbol{A}, \boldsymbol{B})$. (b) Controllability matrix of system $(\boldsymbol{A}, \boldsymbol{B})$. (c) Controllable subspace. Suppose $\boldsymbol{x}(0) = \boldsymbol{0}$, the state of the network will be kept in a subspace, which is the plane $a_{31}x_2(t) = a_{21}x_3(t)$, no matter how to choose the input control signal $u_1(t)$.

For a structural system $(\boldsymbol{A}, \boldsymbol{B})$, the rank of $\mathfrak{C}$ is a function of independent free parameters in $\boldsymbol{A}$ and $\boldsymbol{B}$. The maximum value of the rank of $\mathfrak{C}$ is defined as the generic dimension of the controllable subspace of structural system $(\boldsymbol{A}, \boldsymbol{B})$ and denoted by $GDCS(\boldsymbol{A}, \boldsymbol{B})$. A structural system $(\boldsymbol{A}, \boldsymbol{B})$ is completely structurally controllable if and only if $GDCS(\boldsymbol{A}, \boldsymbol{B}) = n$, which means it is possible to choose the values of the free entries in matrices $\boldsymbol{A}$ and $\boldsymbol{B}$ such that the Kalman's controllability rank condition is satisfied.

A graph-theoretic condition for completely structural controllability (Theorem 2) has been developed in previous studies[10,22,23]. Before presenting Theorem 2, we introduce following two definitions (see Fig.4 for example).

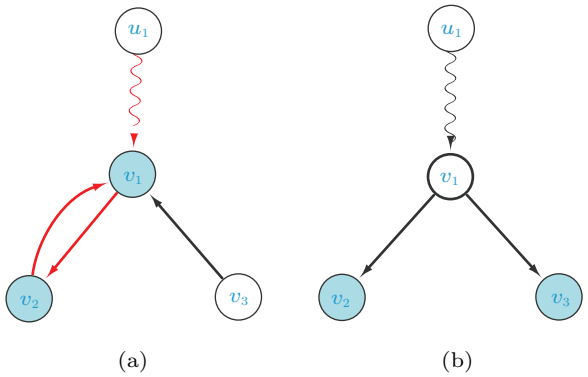


Fig.4. Inaccessible nodes and dilation. (a) There is no path from u_1 to v_3. Therefore, node v_3 is inaccessible. Nodes v_1 and v_2 are accessible. (b) Considering a set $S = \{v_2, v_3\}$, we have $T(S) = \{v_1\}$. Because $|T(S)| < |S|$, there exists a dilation. Based on Theorem 2, systems in (a) and (b) are both structurally uncontrollable.

Definition 1 (Accessibility[10,33]). *In digraph $G(\boldsymbol{A}, \boldsymbol{B})$, a node v_i in V_A is called accessible if and only if there exists a directed path from the input vertices V_U to v_i; otherwise v_i is inaccessible.*

Definition 2 (Dilation[10,33]). *The digraph $G(\boldsymbol{A}, \boldsymbol{B})$ contains a dilation if and only if there is a subset S of V_A such that $|T(S)| < |S|$, where $T(S) = \{v_j \ | (v_j \to v_i) \in E \text{ and } v_i \in S\}$ and E is the edge set of $G(\boldsymbol{A}, \boldsymbol{B})$. The input nodes are not allowed to belong to S but belong to $T(S)$. $|S|$ or $|T(S)|$ is the cardinality of set S or $T(S)$, respectively.*

Theorem 2 (Completely Structural Controllability Theorem[10,33]). *A structural system $(\boldsymbol{A}, \boldsymbol{B})$ is completely structurally controllable if and only if:*

1) *there is no dilation in the digraph $G(\boldsymbol{A}, \boldsymbol{B})$;*

2) *all nodes in V_A are accessible.*

There is an equivalent expression of condition 1: all the nodes in V_A can be covered by node disjoint simple cycles or simple paths starting from nodes in V_U. In a graph, a simple path is a sequence of edges $\{(v_1 \to v_2), (v_2 \to v_3), ..., (v_{k-1} \to v_k)\}$ where all the nodes $\{v_1, v_2, ..., v_k\}$ are distinct. If $v_1 = v_k$ and other nodes are distinct, the sequence of edges is called a simple cycle.

3.2 Control in Subspaces

In many practical problems, it is neither feasible nor necessary to completely control a network, which prompts researchers to develop methods for controlling parts of a network. Though a system may not be completely controllable sometimes, it remains controllable within a subspace (see example in Fig.3(c)). Having a system controllable within a subspace is enough for many real applications. In addition, it is natural that ensuring the controllability within a restricted subspace will require fewer steering nodes being actuated by input control signals than ensuring controllability within the whole state space. Therefore, several approaches have been proposed to investigate the controllability of networks within subspaces.

3.2.1 *Controllable Subnetwork*

For structural system $(\boldsymbol{A}, \boldsymbol{B})$, the dimension of its controllable subspace is measured by $GDCS(\boldsymbol{A}, \boldsymbol{B})$, which is the maximum rank of the controllability matrix $\mathfrak{C}$ by arbitrarily choosing the values of independent free parameters. Hosoe[22] proved that if all nodes in a network system $(\boldsymbol{A}, \boldsymbol{B})$ are accessible, then

$$GDCS(\boldsymbol{A}, \boldsymbol{B}) = \max_{G \in G^*} \{|E(G)|\},$$

where G^* denotes the set of subnetworks of $G(\boldsymbol{A}, \boldsymbol{B})$ which can be spanned by a collection of vertex-disjoint cycles and at most m simple paths (corresponding to m control signals). $|E(G)|$ is the number of edges in G. Actually, each subnetwork in G^* is completely controllable. Therefore, the dimension of its controllable subspace $GDCS(\boldsymbol{A}, \boldsymbol{B})$ equals the number of edges in the largest controllable subnetwork in set G^*. Considering the structural system $(\boldsymbol{A}, \boldsymbol{B})$ in Fig.3, the corresponding G^* consists of two subnetworks of $G(\boldsymbol{A}, \boldsymbol{B})$ induced by node sets $\{u_1, v_1, v_2\}$ and $\{u_1, v_1, v_3\}$, respectively. Based on Hosoe's controllable subspace theorem, the $GDCS(\boldsymbol{A}, \boldsymbol{B})$ of system in Fig.3 is 2, which suggests the whole network can be steered in a 2-dimensional state space. Suppose the network is at the origin at time $t = 0$, it can be observed that the states of nodes v_2 and v_3 must satisfy the equation $a_{31}x_2(t) = a_{21}x_3(t)$. But if we only need to control the subnetwork induced by nodes $\{v_1, v_2\}$ or $\{v_1, v_3\}$, it is enough by actuating node v_1 alone.

Based on Hosoe's controllable subspace theorem, recent studies investigated the controllable subspaces or completely controllable subnetworks from different perspectives, which supplement the theoretical foundation of structural controllability of complex networks. Liu *et al.*[34] defined the control centrality to measure the ability of individual nodes to control a network. Control centrality of node i is defined as $GDCS(\boldsymbol{A}, \boldsymbol{b}(i))$, where $\boldsymbol{b}(i)$ is a vector with a single nonzero i-th entry. The higher control centrality of node means by actuating only the node with input control signal, the whole network can be steered in a larger dimension of

its state space or a larger subnetwork can be completely controlled. Control centrality can be extended to cases in which more than one node is actuated by control signals. Based on this observation, Iudice *et al.*[35] introduced the network permeability, which measures the propensity of a network to be controllable. To calculate the permeability, Indice *et al.*[35] solved a problem related to control centrality at first: identifying m steering nodes from a network with n nodes, such that the corresponding $GDCS(\boldsymbol{A}, \boldsymbol{B}_m)$ is maximized, where $\boldsymbol{B}_m$ is an $n \times m$ controllability matrix corresponding to m steering nodes. Then the permeability is defined as

$$\mu = \frac{\int_0^n (GDCS(\boldsymbol{A}, \boldsymbol{B}_m) - m)\mathrm{d}m}{\int_0^n (n-m)\mathrm{d}m} = \frac{2}{n^2}\int_0^n (GDCS(\boldsymbol{A}, \boldsymbol{B}_m) - m)\mathrm{d}m.$$

Based on the definition, for a network with a high permeability, a large controllable subspace can be obtained or a large subnetwork can be completely controlled by actuating a relatively small set of steering nodes. In order to find a subnetwork which is easy to be controlled with less steering nodes, Liu and Pan[36] proposed a method to choose subnetworks that are important and easy to be controlled in network systems. Then the authors applied this method to multiple real networks and discovered that nodes in the subnetworks chosen by this method tend to be essential. In another study, Commault *et al.*[37] claimed that though the dimension of the controllable subspace is constant for almost any parameter realization of a structural system $(\boldsymbol{A}, \boldsymbol{B})$, the subspace itself is a function of these parameters. Therefore, the authors defined a concept called fixed controllable subspace, which is the intersection of the controllable subspaces of all parameter realizations whose dimension of controllable subspace equals $GDCS(\boldsymbol{A}, \boldsymbol{B})$.

3.2.2 Output (Target) Controllability

In real applications, we are interested in controlling a specific subset of nodes or a subnetwork of interest. Since the subset of nodes can be considered as the output of the network, Wu *et al.*[38] formulated the problem of controlling a predefined subset of nodes in a network as a network output controllability problem. Gao *et al.*[39] proposed the same idea in an independent study, in which they referred to as target control.

The outputs of a linear dynamic system $(\boldsymbol{A}, \boldsymbol{B})$ can be described by the following equation:

$$\boldsymbol{y}(t) = \boldsymbol{C}\boldsymbol{x}(t), \tag{3}$$

where $\boldsymbol{y}(t) = (y_1(t), ..., y_p(t))^{\mathrm{T}}$ is an output vector in which each entry represents an output. $\boldsymbol{C}$ is a $p \times n$ matrix that indicates the outputs of the network. A system described by (1) and (3) is denoted by matrix triplet $(\boldsymbol{A}, \boldsymbol{B}, \boldsymbol{C})$. For target controllability, the outputs are defined as the states of a set of nodes in the network. Then it is assumed that there is one and only one nonzero entry in each row of $\boldsymbol{C}$ such that $\boldsymbol{y}(t)$ is a p-dimensional vector that each entry corresponds to the state of one node. Therefore, target controllability is a special case of output controllability.

A network is output controllable if its outputs can be steered from any initial state $\boldsymbol{y}(0)$ to any desired final state $\boldsymbol{y}(t_f)$ in finite time t_f with appropriate control signals. To test the output controllability of a system $(\boldsymbol{A}, \boldsymbol{B}, \boldsymbol{C})$, a $p \times mn$ output controllability matrix is defined as

$$\mathfrak{o}\mathfrak{C} = (\boldsymbol{C}\boldsymbol{B} \quad \boldsymbol{C}\boldsymbol{A}\boldsymbol{B} \quad \boldsymbol{C}\boldsymbol{A}^2\boldsymbol{B} \quad ... \quad \boldsymbol{C}\boldsymbol{A}^{n-1}\boldsymbol{B}).$$

The condition of output controllability is given by the following theorem in control theory.

Theorem 3 (Output Controllability Theorem[40]). *System $(\boldsymbol{A}, \boldsymbol{B}, \boldsymbol{C})$ is output controllable if and only if* $\mathrm{rank}\,(\mathfrak{o}\mathfrak{C}) = p$.

For a structural system, the rank of $\mathfrak{o}\mathfrak{C}$ can reach a maximum value by arbitrarily choosing the values of independent free parameters in $\boldsymbol{A}$, $\boldsymbol{B}$ and $\boldsymbol{C}$. The maximum value is defined as the generic dimension of the controllable output subspace of structural system $(\boldsymbol{A}, \boldsymbol{B}, \boldsymbol{C})$ and denoted by $GDCOS(\boldsymbol{A}, \boldsymbol{B}, \boldsymbol{C})$. The structural system $(\boldsymbol{A}, \boldsymbol{B}, \boldsymbol{C})$ is called structurally output controllable if $GDCOS(\boldsymbol{A}, \boldsymbol{B}, \boldsymbol{C}) = p$[13,24]. Though Theorem 3 presents conditions for output controllability, there is no method to calculate $GDCOS(\boldsymbol{A}, \boldsymbol{B}, \boldsymbol{C})$ of structural system $(\boldsymbol{A}, \boldsymbol{B}, \boldsymbol{C})$. Murota and Poljak[24] have developed a method to determine the upper and the lower bounds of $GDCOS(\boldsymbol{A}, \boldsymbol{B}, \boldsymbol{C})$.

3.2.3 Transittability

Output controllability measures the ability of a predefined subset of nodes that can be steered by input control signals. However, the states of nodes out of the predefined subset are not considered during control processes. On the other hand, Wu *et al.*[25] introduced a new concept called the transittability of networks, which measures the ability of transition between two specific states of complex networks. Transittability takes the states of all nodes into consideration and reduces the required steering nodes compared with complete controllability.

For system $(\boldsymbol{A}, \boldsymbol{B})$, it is called transittable between these two specific states $\boldsymbol{x}_0$ and $\boldsymbol{x}_1$ if there exist input control signals $\boldsymbol{u}(t),\ t \in [0,\ t_f]$, by which the system $(\boldsymbol{A}, \boldsymbol{B})$ can be transited between two specific states $\boldsymbol{x}(0) = \boldsymbol{x}_0$ and $\boldsymbol{x}(t_f) = \boldsymbol{x}_1$. A sufficient and necessary condition for transittability controllability is given by the following theorem.

Theorem 4 (Transittability Theorem[25]). *With either specific state $\boldsymbol{x}_0$ or $\boldsymbol{x}_1 \in span\{\mathfrak{C}\}$, system $(\boldsymbol{A}, \boldsymbol{B})$ is transittability between $\boldsymbol{x}_0$ and $\boldsymbol{x}_1$ if and only if*

$$\mathrm{rank}(\mathfrak{C}) = \mathrm{rank}(\bar{\mathfrak{C}}),$$

where $\bar{\mathfrak{C}} = (\bar{\boldsymbol{B}} \quad \boldsymbol{A}\bar{\boldsymbol{B}} \quad \boldsymbol{A}^2\bar{\boldsymbol{B}} \quad ... \quad \boldsymbol{A}^{n-1}\bar{\boldsymbol{B}})$ *and* $\bar{\boldsymbol{B}} = (\boldsymbol{x}_0 - \boldsymbol{x}_1, \boldsymbol{B})$.

Similar to completely structural controllability, a structural system $(\boldsymbol{A}, \boldsymbol{B})$ is called structurally transittable between two specific structural states $\boldsymbol{x}_0$ and $\boldsymbol{x}_1$ if there exists an admissible system $(\widetilde{\boldsymbol{A}}, \widetilde{\boldsymbol{B}})$ (with respect to $(\boldsymbol{A}, \boldsymbol{B})$) and admissible states $\widetilde{\boldsymbol{x}_0}$ and $\widetilde{\boldsymbol{x}_1}$ (with respect to $\boldsymbol{x}_0$ and $\boldsymbol{x}_1$, respectively) such that the system $(\widetilde{\boldsymbol{A}}, \widetilde{\boldsymbol{B}})$ is transittable between states $\widetilde{\boldsymbol{x}_0}$ and $\widetilde{\boldsymbol{x}_1}$. In fact, for structure systems, the transittability between two structure states actually measures the ability to control a subset of nodes in the network without disturbing other nodes.

4 Identification of Steering Node Sets

In order to control a network, the first step is to identify a set of steering nodes which should be actuated by input control signals. A network system is completely controllable if each node is directly actuated by a distinct input control signal. However, it is costly and impractical for large networks. Therefore, methods are required to identify minimal steering node sets such that the control objective can be satisfied. The identification of steering nodes for controlling networks can be viewed as problems of determining appropriate control matrix $\boldsymbol{B}$ when a network, which is represented by $\boldsymbol{A}$, is given. Theorems 2–4 provide conditions to judge if a structural system $(\boldsymbol{A}, \boldsymbol{B})/(\boldsymbol{A}, \boldsymbol{B}, \boldsymbol{C})$ is completely structurally controllable, structurally output controllable or transittable between two specific states respectively. However, for a given network (matrix $\boldsymbol{A}$), controllability theorems do not indicate a set of steering nodes (matrix $\boldsymbol{B}$) such that the network system is controllable. A brute-force search for a minimal steering node set would require checking the controllability conditions for almost 2^n distinct controllability matrices $\boldsymbol{B}$, which is computationally prohibited. In this section, methods for identifying steering node sets for different control objectives are reviewed.

4.1 Steering Nodes for Complete Controllability

The minimum driver node set (MDS)[11] and the minimum steering node set (MSS)[33] are two mostly investigated steering node sets for completely controlling networks. Recently, graph-theoretic methods have been proposed to identify MDSs and MSSs of networks based on Theorem 2.

The MDS is a minimum set of nodes in which each node should be actuated by an independent control signal such that the condition i ("no dilation" condition) of Theorem 2 can be satisfied. However, applying independent control signals to an MDS does not guarantee complete controllability of the network and it is a necessary condition for completely controlling a network. In [11], the identification of MDS has been formulated as a maximum matching problem in an undirected bipartite graph corresponding to the original network. A matching on an undirected graph is a set of edges without common nodes and a maximum matching is a matching with the largest size. To identify an MDS, a bipartite graph which contains node sets $R = \{r_1, \cdots, r_n\}$ and $C = \{c_1, \cdots, c_n\}$ is constructed. The nodes r_i and c_i correspond to the node i of $G(\boldsymbol{A})$. If there is a directed edge from node i to node j in $G(\boldsymbol{A})$, there is an edge in the bipartite graph connecting r_j and c_i. A maximum matching in the bipartite graph can be solved by the Hoproft-Karp algorithm[41]. Then the MDS is corresponding to the nodes in R that are not connected to any matching edges (see Fig.5(b)). It can be verified that if each node in an MDS is actuated by an input control signal, which means adding a control node u_i for each node i in the MDS, the resulting graph $G(\boldsymbol{A}, \boldsymbol{B})$ will have no dilation. Since MDSs of a network are not unique, Zhang *et al.*[42] proposed a preferential matching algorithm to identify MDSs that have a specific degree property. Zhang *et al.*'s work provides an inspiration that the MDS can be selected with preference, by which realistic information can be taken into consideration.

The MSS is a minimum set of nodes in a network which should be actuated by control signals to completely structurally control the network. Compared with the MDS, applying independent control signals to an MSS guarantees the complete controllability of the network and it is a sufficient and necessary condi-

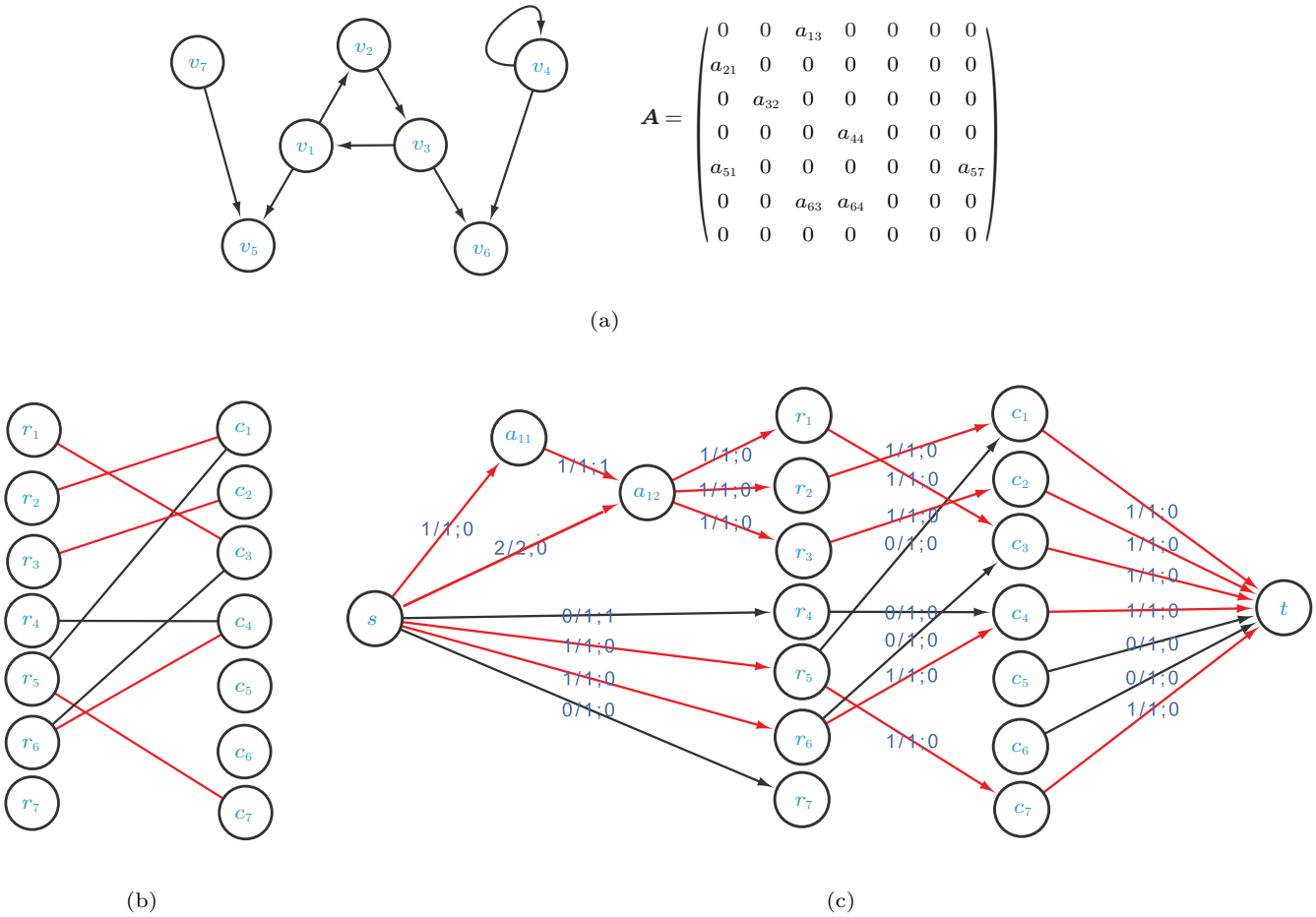


Fig. 5. Identification of an MDS and an MSS by maximum matching and minimum cost maximum flow method, respectively. (a) A network $G(\boldsymbol{A})$ and its corresponding system state transition matrix. (b) The corresponding undirected bipartite graph and the maximum matching. Nodes r_4 and r_7 in node sets R are not matched in the maximum matching, which suggests v_4 and v_7 make up an MDS. (c) A directed graph constructed based on the structure of $G(\boldsymbol{A})$ and the minimum cost maximum flow. The labels on edges represent flow, capacity and cost per unit flow, respectively. There is no flow passing through r_4 and r_7, which suggests v_4 and v_7 belong to an MSS. The flow passes through one edge with cost 1, which means there is an additional steering node which should be chosen from the corresponding source strong connective component (SCC). Then nodes v_4, v_7 and v_i $(i = 1, 2, 3)$ make up an MSS.

tion for completely controlling a network, which satisfies both conditions of Theorem 2. Therefore, each MSS contains a set of nodes which make up an MDS. In [33], the bipartite graph for the identification of MDS has been extended to a directed graph. The authors[33] proved that a minimum cost maximum flow (MCMF) in the constructed digraph corresponds to an MSS of the network (see Fig.5(c)). The algorithm for solving the MCMF problem can be found in [43]. Similar to the MDS, MSSs of a network are not unique as well. Therefore, Wu *et al.*[44] developed an approach to identify MSSs with preference, such that the average preference value of nodes in the identified MSS is the maximum among all possible MSSs of the network. When properly assigning preference values to the nodes in a network, the algorithm is able to find a most suitable MSS for controlling the states of the network in practical applications.

Several studies investigated the identification of steering nodes under constraints, which are common in real applications. Pequito *et al.*[45] proved that the minimum constrained input selection (minCIS) problem, which selects the minimum number of inputs from a given set of possible inputs, is NP-hard. When there are n possible inputs and each input can actuate a distinct node of the network, the minCIS problem reduces to the problem of the identification of MSS, which could be solved in polynomial time.

Stepping out of structural controllability, some studies considered constraints from the aspects of input control signals and control energy. Lindmark and Altafini[46] studied the controllability of complex net-

works with unilateral inputs, which assumes that an input control signal is either negative or positive, but not both. Conditions for unilateral controllability have been formulated algebraically in terms of eigenspaces of the system matrix $\boldsymbol{A}$. Compared with unconstrained control, more steering nodes are required to achieve complete controllability for unilateral control. By studying several instances with randomly weights assigned to the edges, they discovered that the number of additional steering nodes for unilateral control is strongly related to the number of roots and dilations in a network. Then a lower bound of the minimum number of steering nodes required for unilateral controllability can be determined by network structure alone. This study provides methodology for having networks controllable with unilateral control signals, which are common in real scenarios. For instance, input control signals of biological networks are usually drugs or chemical molecules, which can only either activate or inhibit the steering nodes.

In many cases, though actuating an MSS can completely control a network theoretically, the associated control cost can be unbearably large, which prevents actual control from being realized physically. The control cost can be measured by control energy, which is defined as

$$E(t_f) = \int_0^{t_f} (\boldsymbol{u}_t^{\mathrm{T}} \cdot \boldsymbol{u}_t)\mathrm{d}t,$$

where $\boldsymbol{u}_t$ represents input control signals[47]. Wang *et al.*[48] proposed physical controllability which considers the probability of achieving control practically. By investigating control energy for controlling chain structures in networks, they provided strategies to make physically uncontrollable networks physically controllable by properly adding additional steering nodes. Li *et al.*[49] studied the problem of identifying a fixed number of steering nodes, such that a network can be completely controllable with the minimum energy. They formulated the original problem as an optimization problem and developed two methods to solve it. These studies take control energy into consideration when identifying steering nodes, which can save the cost of the control processes.

4.2 Steering Nodes for Output Controllability

It has been proved that identifying the minimum number of steering nodes for structural output controllability is an NP-hard problem[50], where the outputs are defined as the states of a set of nodes. In Wu *et al.*'s study[13], the lower bound of $GDCOS(\boldsymbol{A}, \boldsymbol{B}, \boldsymbol{C})$[24] has been applied to design an algorithm to identify steering nodes for output controllability, which guarantees that the network is output controllable by actuating identified steering nodes. Therefore, actuating the identified steering node set is a sufficient but may not be a necessary condition for structural output controllability. The identification of steering nodes for output controllability has been formulated to maximum weight complete matching problem in a bipartite graph constructed based on network topology and a predefined set of nodes to be controlled. The maximum weight complete matching problem can be solved by the Kuhn-Munkres (KM) algorithm[51]. Fig.6 is an illustrative example for identifying steering nodes for controlling a subset of nodes in a network.

In Gao *et al.*'s study[39], a greedy algorithm has been developed to identify steering nodes which are sufficient for target control. Several further studies developed algorithms to identify steering nodes for target controllability by reducing the number of steering nodes or considering realistic constraints. Instead of using the greedy algorithm, Zhang *et al.*[52] developed an algorithm which elaborately rearranges the matching order of the nodes such that the required number of steering nodes for target control can be significantly reduced. The comparison results on model generated networks and real networks indicate that the proposed algorithm outperforms Gao *et al.*'s algorithm[39]. Because the functions of network systems intensively depend on the connections between nodes, Liu *et al.*[53] investigated the target controllability of giant connected components of directed networks by selecting target nodes from giant connected components, which are the connected components of networks that have constant fractions of nodes in networks. In the study, the relationships between the number of steering nodes for controlling giant connected components and the parameters of model generated networks are explored. Piao *et al.*[54] considered controlling a subnetwork called target community of a complex network when the whole topological structure of the network is not available. They argued that though a target community is controllable with steering nodes identified by structural controllability analysis, determining input control signals that are able to achieve a given control goal can be very difficult. It is because the process of controlling target communities would be influenced by signals from the remainder network, but the topology and state of the remainder

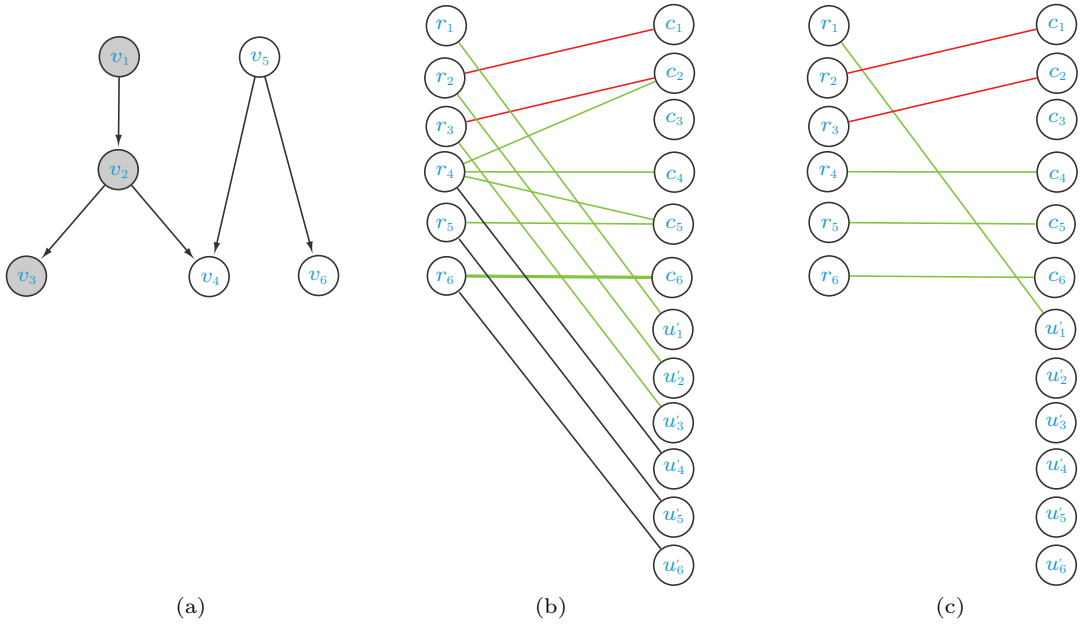


Fig.6. Identifying steering nodes for output controllability. (a) A network. The outputs of the system are the states of grey nodes v_1, v_2 and v_3. (b) A corresponding weighted bipartite graph. The weights of red lines, green lines and black lines are 1, 0 and -1, respectively. (c) A maximum weight complete matching in (b). Since r_1 matches a node in set $U' = \{u'_i | i = 1, \ldots, 6\}$, node v_1 makes up a steering node set for output controllability, which suggests that the states of nodes v_1, v_2 and v_3 can be controlled by actuating node v_1 only.

network is not available. To deal with this issue, they defined a type of steering nodes, which refer to as immune nodes, for blocking signals transmitting from the remainder network. Then they proposed a method to reduce the total number of steering nodes and immune nodes such that the subnetwork is completely controllable and the signals from the rest parts of the network can be blocked.

By considering some practical constraints, Guo *et al.*[55] proposed the concept called constrained target controllability of complex networks. Because not all nodes are possible to be steering nodes in reality, the constrained target controllability concerns the target controllability by selecting steering nodes from a predefined constrained node set. Then they developed an algorithm to identify steering nodes from a constrained node set for controlling a set of target nodes. Iudice *et al.*[35] also investigated the target controllability of networks by not only considering the constraints on selection of steering nodes, but also introducing a set of untouchable nodes, whose states should not be perturbed during the control process. Selecting steering nodes from constrained node set improves the applicability of control strategies while keeping states of untouchable nodes unchanged reduces side effects or undesired effects during control processes.

4.3 Steering Nodes for Transittability

In addition, Wu *et al.*[25] developed an algorithm to identify steering nodes with a given network $G(\boldsymbol{A})$ and a set of target nodes whose states are supposed to be changed during state transition. Identification of steering nodes for transittability has been formulated to maximum weight complete matching problem in a bipartite graph constructed based on network $G(\boldsymbol{A})$ and structural states. Fig.7 is an illustrative example of identifying steering nodes for state transittability. The result indicates that by actuating nodes v_1 and v_3 with input control signals, the states of nodes v_1, v_2 and v_3 can be controlled without affecting nodes v_4, v_5 and v_6, which is different from the example of output controllability in Fig.6: the states of nodes v_1, v_2 and v_3 can be controlled by actuating v_1; however, the state of v_4 might be perturbed as well. Transittability usually requires less steering nodes than completely controlling the whole network, which is more efficient and practical. Compared with output controllability, transittability needs more steering nodes for controlling the states of target nodes but causes less undesired effects to the non-target nodes, which makes a balance between control efficiency and side effects.

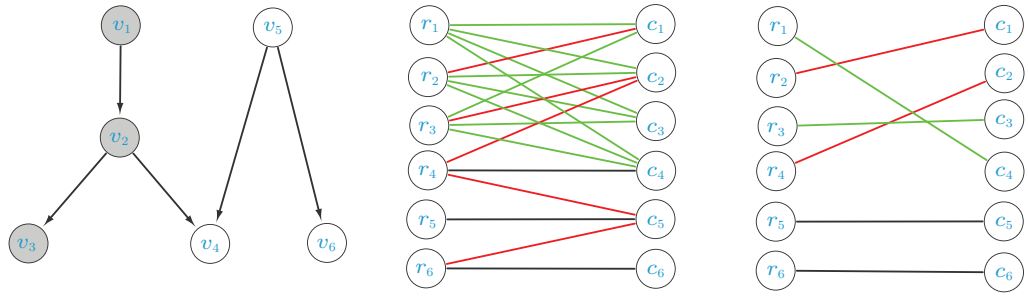


Fig.7. Identifying steering nodes for state transittability. (a) A network. Grey nodes v_1, v_2 and v_3 could be changed to any states in finite time while states of white nodes would not be affected by control signals at the end of the control process. (b) A corresponding weighted bipartite graph. The weights of red lines, green lines and black lines are 1, λ and 0, respectively, where λ is a small enough positive number. (c) A maximum weight complete matching in (b). Since r_1 and r_3 match green lines, nodes v_1 and v_3 make up a steering node set for state transittability.

5 Applications to Biological Networks

Studies introduced in Section 4 offer methods to identify steering node sets for controlling biological networks in different control scenarios. The steering nodes are identified based on the control theory, which guarantees the controllability of networks theoretically. However, theoretical analyses on structural controllability of complex networks do not explain how to manipulate the steering nodes to steer networks from one state to another, which depends on the details of connections and interactions in the networks. Though structural controllability methods are not able to provide explicit strategies to control complex networks, they still offer useful approaches to investigate the controllability of complex networks when only the network topology is available. For biological networks, investigating the meanings of biological elements that play different roles in controlling biological networks is a way to interpret the controllability of biological network. Conversely, reasonable functions of biological elements could support the validity of controllability analyses. In this section, we review recent studies that explored various types of biological networks, including biomolecular networks, neuronal networks and brain networks, based on network controllability. Particularly, we mainly focus on understanding the functions of nodes in the various biological networks from viewpoints of controllability.

5.1 Steering Node Sets in Biological Networks

By applying input control signals to steering node sets, networks can be steered to the desired states. Therefore, recent studies explored the biological meanings of steering node sets in biological networks for different control objectives, which are complete controllability[11], output controllability[13], and transittability[25].

5.1.1 *Biological Properties of Steering Nodes*

MDS and MSS of biological networks are two of the most investigated steering node sets for completely controlling networks. There are many studies identifying MDSs or MSSs of biological networks and the biological functions of biological elements in the steering node sets are examined.

Khazanchi *et al.*[56] compared driver nodes in MDSs and hub nodes of four different protein-protein interaction (PPI) networks. They found that hub nodes are more likely to be lethal proteins while driver nodes tend to be transcription factors. In addition, they found that driver nodes are enriched in first-degree neighbors of hubs, which suggests that one should control the nodes interacting with the hubs, instead of controlling the hubs directly, to control networks. Khazanchi *et al.*'s research indicates that when controlling PPI networks, it is better not to perturb the lethal hub proteins, but the proteins near the hubs.

Badhwar and Bagler[57] identified the MDS of *C. elegans* neuronal network, which consists of 302 neurons that are connected through chemical synapses and gap junctions. Based on the functions, neurons can be classified as sensory neurons, motor neurons and inter neurons. By investigating the functional and the spacial correlations of driver neurons in the MDS, they found that most driver neurons are motor neurons that are located in the ventral nerve cord. In addition, the driver neurons mostly participate in the biological reproduc-

tion. This study demonstrates the importance of driver neurons and their ability of controlling the behaviours of the organism.

Noori *et al.*[58] constructed a comprehensive neurochemical network of the rat brain and identified an MDS of the rat brain network. Interestingly, one of the four steering nodes in the identified MDS, subthalamic nucleus (STh), has already been proved to be crucial in global circuit dynamics[59] and treatment of Parkinson disease as well as other disorders[60] by numerous deep brain stimulation studies. This observation manifests an agreement between the structural controllability and the function of neuronal networks.

For the MSSs of biological networks, Wu *et al.*[33] applied their method of identifying MSS to the *S.cerevisiae* cell cycle networks[61,62], Epithelial to Mesenchymal Transition (EMT) network[63] and myeloid differentiation regulatory network[64]. It has been discovered that steering nodes in MSSs of these networks play critical roles in triggering cell division process, maintaining homeostasis of epithelial or regulating early myeloid development as well as hematopoietic stem cells, respectively. Since the identified MSSs in these networks are closely related to dynamic behaviours of the networks, it is fair to suggest their importance in controlling the networks.

For transittability, Wu *et al.*[25] employed different biological systems with different phenotypes to validate the applicability of identified steering node sets for transittability. For example, T helper cells (Th cells), which play an important role in the immune system, are a sub-group of lymphocytes. A network has been constructed by Mendoza[65] to model the differentiation of Th cells. Matured Th cells can be classified as Th0 (precursor), Th1 and Th2 (effector) cells, which correspond to three different states of Th differentiation network. Steering node sets for state transitions among these three phenotypes have been identified by the proposed algorithm. According to the transittability analyses, actuating steering nodes SOCS1 and T-bet can steer the network between Th0 and Th1 and actuating nodes IL-4 and GATA3 can steer the network between Th0 and Th2, which is in agreement with existing knowledge[66,67]. Actuating steering nodes T-bet and GATA3 can cause the transition between Th1 and Th2, which is completely in agreement with the experimental data[68]. The proof of applicability suggests further applications for drug target identification or medical treatments based on transittability: by actuating steering nodes with input control signals (e.g., drugs), the biological systems can be steered from disease states to healthy states.

5.1.2 Identification of Drug Targets

There is an intuitive application of identifying steering nodes for the controllability of biological networks, which is drug target identification. By perturbing the steering nodes, biological networks can be steered to desired states. Therefore, steering nodes in a disease related biological network could be potential drug targets.

Wu *et al.*[13] formulated the problem of drug target identification as a problem of identifying steering node set for output controllability of biological networks. In the study, disease biomolecules and biomolecules whose state changes would lead to side effects are defined as the outputs of the network, which takes both efficiency and safety into consideration. The steering nodes for controlling these two types of biomolecules are considered as potential drug targets. The method has been applied to several real biological networks. The identified potential drug targets are targets of approved drugs or in agreement with existing research results, which indicates the feasibility of the method.

By considering the constrained target controllability, Guo *et al.*[55] applied the developed algorithm for identifying steering nodes to a gene regulatory network related to type 1 diabetes. By defining the five genes related to type 1 diabetes as the target nodes and all FDA-approved drug targets as the constrained node sets, they found that FASLG and CD80 are steering nodes for controlling the target nodes related to type 1 diabetes, which is supported by previous wet experiments. In another study, Kanhaiya *et al.*[69] built three PPI networks for breast, pancreatic and ovarian cancer, respectively. They considered survivability-essential proteins specific to each cancer type as control targets in each network. A method was also proposed to identify steering nodes among FDA-approved drug target nodes. Different from the method in Guo *et al.*'s study[55], the selection of steering nodes from FDA-approved drug targets is not obligatory, but is preferred in Kanhaiya *et al.*'s study[69]. The results indicate that many steering nodes are known drug targets for cancer therapies, but some of them are not the drug targets corresponding to cancer types.

By considering a more realistic constraint, Wu *et al.*[70] improved the algorithm of identifying MSSs by considering the preference of individual nodes, such that nodes in the identified MSS have higher preference

values compared with nodes in other eligible MSSs of a network. The algorithm has been applied to study MSSs with drug binding preference of some biological networks. The biomolecules in the MSSs with binding preference are enriched with known drug targets and are likely to have more chemical-binding opportunities with existing drugs compared with randomly chosen MSSs, suggesting novel applications for drug target identification and drug repositioning: steering nodes for having a disease related network controllable could be potential drug targets whereas drugs that can bind to the steering nodes might be used for the treatment of other diseases.

5.1.3 Prediction of Node Functions

Instead of analyzing known functions of steering nodes, Yan *et al.*[14] predicted the involvement of neurons in the *C. elegans* neuronal network by formalizing the responsive mechanism of *C. elegans* to external stimuli as a target control problem. The predictions based on the target control of network have been validated by their experiments. For example, it has been predicted that three neurons (DD04, DD05, or DD06) in the DD motor neuronal class should affect locomotion when ablated individually. Their experimental validation shows that ablations of DD04 or DD05 have impacts specifically on posterior body movements, whereas ablations of other neurons in the DD motor neuronal class (DD02 or DD03) do not affect the locomotion. Yan *et al.*'s study[14] not only provides a novel method to unveil how the structure of neuron network affects its functions based on controllability perspective, but also offers the first experimental proof of the validity of network structural controllability analyses.

5.2 Roles of Individual Nodes in Controllability

Instead of focusing on specific steering node sets for different control objectives, several studies proposed methods to quantify the importance or analyze the roles of individual nodes in controlling networks and then investigated biological meanings of nodes based on the proposed methods. The analyses were mainly based on the importance of nodes in network controllability, control energy or control paths. Applications show that nodes in biological networks with different biological functions can be distinguished by certain network controllability measures. Therefore, investigating the controllability of biological networks provides new methods for understanding the biological properties of nodes in biological networks.

5.2.1 Qualifying and Quantifying Importance of Nodes in Network Controllability

Since the MDSs or MSSs of a network are not unique, the algorithms for identifying MSSs or MDSs do not result in a unique set of MDS or MSS. Therefore, some studies attempted to figure out the importance of nodes in network controllability by classifying nodes into different categories or assigning centrality values to nodes according to certain rules.

Jia *et al.*[71] classified a node in the network as critical, intermittent or redundant if it acts as a driver node in all, some or none of all possible MDSs, respectively. By classifying nodes in a human signaling network, Liu and Pan[72] discovered that critical nodes are enriched in the group of ligands, intermittent nodes are enriched in cell surface receptors, and redundant nodes are enriched in intracellular signaling proteins. They also found that cancer-associated genes are enriched in redundant nodes, which suggests that controlling the regulators of the cancer-associated genes could be more feasible than controlling the cancer-associated genes directly.

In a related work of the classification, Jia and Barabási[73] proposed a concept called control capacity, which is defined as the likelihood that a node is a driver node in an arbitrary MDS. Liu and Pan[74] calculated the control capacity of nodes in a human liver metabolic network and classified nodes into critical, high-frequency and low-frequency nodes based on their control capacity values. They found that in the metabolic network, critical metabolites are likely to be essential metabolites while the high-frequency metabolites tend to participate in different metabolic pathways.

Though the MDSs of a network may not be unique, the cardinality of all the MDSs is the same. Vinayagam *et al.*[75] classified a node in a network as indispensable, neutral or dispensable, which is correlated to increasing, no effect, or decreasing the cardinality of the MDSs of the network by removing that node and edges which are connected to the node. Then the authors[75] applied their classification strategy to a directed human PPI network and found that indispensable proteins or corresponding genes are enriched in essential genes, human virus targets, drug targets or disease-causing mutations. Their study provides a novel classification strategy based on network controllability. Nodes in diffe-

rent categories show distinct biological properties in the context of essentiality, evolutionary conservation, and regulation of translational or post-translational modifications. In fact, before the work of Vinayagam *et al.*[75], Matsuoka *et al.*[76] identified the indispensable nodes, which they called "critical node" in their study, of an influenza A virus life cycle network. They found that the indispensable nodes are important factors of the viral life cycle, which are known drug targets or could be potential therapeutic targets. In another study, Uhart *et al.*[77] studied a directed phosphorylation-based PPI network by analyzing the biological characteristics of indispensable nodes. Because post-translational modification and inhibition of transduction by miRNAs are two important mechanisms of regulation in eukaryotic cells, it is meaningful to evaluate the relationship between proteins that are important in controlling the network and these two mechanisms. It has been discovered that indispensable nodes are more enriched in post-translational modifications and miRNA targets, which indicates that indispensable nodes are targets of intense biological regulation. Uhart *et al.*'s study[77] provides a deeper understanding of the controllability of biological networks and bridges the controllability theorems and cell regulation processes, such as post-translational modification, in a phosphorylation-based PPI network.

In recent work, Ravindran *et al.*[78] combined two types of classification strategies and investigated a cancer signaling network. Nodes are classified as critical, intermittent or redundant based on Jia *et al.*'s classification strategy[71], and indispensable, neutral or dispensable based on Vinayagam *et al.*'s classification strategy[75]. Then the authors[75] analyzed the distribution of cancer genes and targets of anti-cancer drugs in each node class. Enrichment analyses show that redundant nodes, especially indispensable redundant nodes are enriched in both cancer genes and anti-cancer drug targets, which implies a strong correlation between indispensable redundant nodes and cancer development or cancer treatment. This study[78] indicates that the two classification strategies can capture the roles that individual nodes play in controlling a network from different aspects. Therefore, it is likely to obtain more comprehensive results by combining these two classification strategies.

By investigating topological features of steering nodes in MDSs, Ruths and Ruths[79] found that each driver node in MDSs corresponds to one of three topological features: source nodes, external dilations and internal dilations. Sources nodes are nodes that have no incoming edges and the number of source nodes is denoted as N_{source}. External dilations appear when sink nodes, which are nodes without outgoing edges, outnumber source nodes. The number of sink nodes is denoted as N_{sink} and then the number of external dilations equals $N_{\text{external}} = \max(0, N_{\text{sink}} - N_{\text{source}})$. Internal dilations are dilations other than external dilations and the number of internal dilation is denoted as N_{internal}. Then the cardinality of MDSs $N_{\text{MDS}} = N_{\text{source}} + N_{\text{external}} + N_{\text{internal}}$, which is the sum of the three topological features. Then the driver nodes can be classified into three categories based on their corresponding topological features. The authors[79] discovered that the MDSs of a network are usually dominated by a specific topological feature. Based on the proportions of each type of driver nodes, a network can be classified as source dominated, external-dilation dominated or internal-dilation dominated. The classification of networks has been tested on various of real networks. The results offer insights into the relationship between the topology and the functions of complex networks. For example, neuronal networks are source-dominated, which tend to allow relatively uncorrelated behaviors and are suitable for distributed processing.

5.2.2 *Control Energy*

Control energy is another aspect needed to be considered in the controllability of biological networks. Gu *et al.*[4] studied the controllability of a human brain network, in which each node represents a region of interest (ROI) of the human brain. Three types of measures are developed to quantify the importance of nodes in controlling the brain network: average controllability measures the ability of brain regions to steer the system state with less energy input, modal controllability identifies brain regions that steer the system to states which require substantial input energy, and boundary controllability identifies brain regions that locate at boundaries between network communities and control the segregation and integration of cognitive systems. Though Tu *et al.*[80] have different opinions on the results of this study, this study provides a novel perspective to understand the cognitive processes from the control energy and network controllability. The proposed methods and measures based on control energy could provide insights into studies on the controllability of other types of biological networks.

5.2.3 Control Paths

To understand the disease etiology from the perspective of network control, Wang *et al.*[12] defined a concept called perturbation influence, which is a subset of nodes based on the control paths (vertex-disjoint cycles and simple paths starting from steering nodes), to identify and quantify the ways by which disease genes perturb human regulatory networks. Intuitively, for a certain disease, the perturbation influences of different disease genes can be considered as the significant pathways related to the disease, which are etiologically essential. In addition, perturbation influence can be applied to prioritize disease genes based on the similarities of perturbation influences between nodes and known disease genes. The validation of the prioritizing method on 112 diseases shows that this method outperforms the state-of-the-art method PRINCE[81]. Similar to perturbation influence, concepts such as control range[82] or vertex domination centrality[83], which are defined based on control paths as well, have been proposed to study the controllable subspaces of nodes or measure the importance of nodes in controlling networks. The proposed controllability concepts based on control paths enrich analytical tools for understanding roles of nodes in controlling network subspaces.

6 Conclusions

In this article, we reviewed recent advances on the controllability of complex networks and the applications to biological networks. First, dynamic models of complex networks were briefly reviewed. Though the linear dynamic models are simple models that cannot describe many behaviours of biological networks, the structural control theorems based on the linear dynamic models provide a frame to study the controllability of biological networks solely from their underlying network structures. Since nonlinear dynamics would enhance the controllability of biological networks, studying the controllability based on the linear dynamics provides a sufficient condition for having biological networks controllable.

Started from the study of the completely controllability of networks based on structural linear dynamic models, we then reviewed different control objectives such as output controllability and transittability. Algorithms to identify steering node sets for different control objectives were reviewed. In order to make the investigation more useful and practical for real biological networks, we introduced the studies that consider controllability of biological networks under realistic constraints, such as selecting steering nodes (biomolecules) based on drug binding preference, minimizing control energy or controlling biological networks by using unilateral control signals. When more constraints are considered, the analyzing results are more reliable and the developed methods are closer to real applications.

In network controllability studies, methods have been proposed to identify steering nodes for having biological networks controllable. Because the controllability studies mainly based on the simple structural linear dynamic models, approaches to verifying whether perturbing the steering nodes can steer biological networks to desired states are lack. Therefore, studying the biological meanings of nodes which play different roles in controlling biological networks is a practical way to verify the validity of the proposed methods and to interpret the controllability of biological networks. Studies on the controllability of biological network showed that controllability is an innovative perspective to analyze biological networks and suggested a variety of promising applications.

In the future, other practical constraints such as control trajectories can be considered in order to avoid some forbidden or fatal states of biological networks during control processes. Besides controllability, other concepts in control theory could also shed lights on our ability to understand or manipulate biological networks, which is worthy for future investigation. For example, observability, which is a mathematical dual problem of controllability, can be applied to measure the states of biological networks by monitoring a specific set of biological elements. It is believed that controlling biological networks will be increasingly feasible and effective when our knowledge of control theory is enhanced and our understanding of dynamics of biology systems is deepened.

References

[1] Ito T, Chiba T, Ozawa R *et al.* A comprehensive two-hybrid analysis to explore the yeast protein interactome. *Proceedings of the National Academy of Sciences*, 2001, 98(8): 4569-4574.

[2] Sprinzak E, Margalit H. Correlated sequence-signatures as markers of protein-protein interaction. *Journal of Molecular Biology*, 2001, 311(4): 681-692.

[3] Liu L Z, Wu F X, Zhang W J. Reverse engineering of gene regulatory networks from biological data. *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*, 2012, 2(5): 365-385.

[4] Gu S, Pasqualetti F, Cieslak M *et al*. Controllability of structural brain networks. *Nature Communications*, 2015, 6: Article No. 8414.

[5] Csermely P, Agoston V, Pongor S. The efficiency of multi-target drugs: The network approach might help drug design. *Trends in Pharmacological Sciences*, 2005, 26(4): 178-182.

[6] Dai Y F, Zhao X M. A survey on the computational approaches to identify drug targets in the postgenomic era. *BioMed Research International*, 2015, 2015: Article No. 239654.

[7] Wang X, Gulbahce N, Yu H. Network-based methods for human disease gene prediction. *Briefings in Functional Genomics*, 2011, 10(5): 280-293.

[8] Chen B, Fan W, Liu J *et al*. Identifying protein complexes and functional modules — From static PPI networks to dynamic PPI networks. *Briefings in Bioinformatics*, 2013, 15(2): 177-194.

[9] Kalman R E. Mathematical description of linear dynamical systems. *Journal of the Society for Industrial and Applied Mathematics Control, Series A*, 1963, 1(2): 152-192.

[10] Lin C T. Structural controllability. *IEEE Transactions on Automatic Control*, 1974, 19(3): 201-208.

[11] Liu Y Y, Slotine J J, Barabási A L. Controllability of complex networks. *Nature*, 2011, 473(7346): 167-173.

[12] Wang B, Gao L, Zhang Q *et al*. Diversified control paths: A significant way disease genes perturb the human regulatory network. *PLoS One*, 2015, 10(8): Article No. e0135491.

[13] Wu L, Shen Y, Li M *et al*. Network output controllability-based method for drug target identification. *IEEE Transactions on Nano Bioscience*, 2015, 14(2): 184-191.

[14] Yan G, Vértes P E, Towlson E K *et al*. Network control principles predict neuron function in the caenorhabditis elegans connectome. *Nature*, 2017, 550(7677): 519-523.

[15] D'haeseleer P, Wen X, Fuhrman S *et al*. Linear modeling of mRNA expression levels during CNS development and injury. *Pacific Symposium on Biocomputing*, 1999, 4: 41-52.

[16] Slotine J J, Li W. Applied Nonlinear Control. Pearson, 1991.

[17] Liu Y Y, Barabási A L. Control principles of complex systems. *Reviews of Modern Physics*, 2016, 88(3): Article 035006.

[18] Shields R, Pearson J. Structural controllability of multiinput linear systems. *IEEE Transactions on Automatic Control*, 1976, 21(2): 203-212.

[19] Glover K, Silverman L. Characterization of structural controllability. *IEEE Transactions on Automatic Control*, 1976, 21(4): 534-537.

[20] Hosoe S, Matsumoto K. On the irreducibility condition in the structural controllability theorem. *IEEE Transactions on Automatic Control*, 1979, 24(6): 963-966.

[21] Linnemann A. A further simplification in the proof of the structural controllability theorem. *IEEE Transactions on Automatic Control*, 1986, 31(7): 638-639.

[22] Hosoe S. Determination of generic dimensions of controllable subspaces and its application. *IEEE Transactions on Automatic Control*, 1980, 25(6): 1192-1196.

[23] Poljak S. On the generic dimension of controllable subspaces. *IEEE Transactions on Automatic Control*, 1990, 35(3): 367-369.

[24] Murota K, Poljak S. Note on a graph-theoretic criterion for structural output controllability. *IEEE Transactions on Automatic Control*, 1990, 35(8): 939-942.

[25] Wu F X, Wu L, Wang J *et al*. Transittability of complex networks and its applications to regulatory biomolecular networks. *Scientific Reports*, 2014, 4: Article No. 4819.

[26] Mayeda H, Yamada T. Strong structural controllability. *SIAM Journal on Control and Optimization*, 1979, 17(1): 123-138.

[27] Tu C. Strong structural control centrality of a complex network. *Physica Scripta*, 2015, 90(3): Article No. 035202.

[28] Nepusz T, Vicsek T. Controlling edge dynamics in complex networks. *Nature Physics*, 2012, 8(7): 568-573.

[29] Cowan N J, Chastain E J, Vilhena D A *et al*. Nodal dynamics, not degree distributions, determine the structural controllability of complex networks. *PLoS One*, 2012, 7(6): Article No. e38398.

[30] Nie S, Wang X, Zhang H *et al*. Robustness of controllability for networks based on edge-attack. *PLoS One*, 2014, 9(2): Article No. e89066.

[31] Wang W X, Ni X, Lai Y C *et al*. Optimizing controllability of complex networks by minimum structural perturbations. *Physical Review E*, 2012, 85(2): Article No. 026115.

[32] Wu L, Li M, Wang J *et al*. CytoCtrlAnalyser: A cytoscape app for biomolecular network controllability analysis. *Bioinformatics*, 2018, 34(8): 1428-1430.

[33] Wu L, Li M, Wang J *et al*. Minimum steering node set of complex networks and its applications to biomolecular networks. *IET Systems Biology*, 2016, 10(3): 116-123.

[34] Liu Y Y, Slotine J J, Barabási A L. Control centrality and hierarchical structure in complex networks. *PLoS One*, 2012, 7(9): Article No. e44459.

[35] Iudice F L, Garofalo F, Sorrentino F. Structural permeability of complex networks to control signals. *Nature Communications*, 2015, 6: Article No. 8349.

[36] Liu X, Pan L. Controllability of the better chosen partial networks. *Physica A: Statistical Mechanics and Its Applications*, 2016, 456: 120-127.

[37] Commault C, van der Woude J, Boukhobza T. On the fixed controllable subspace in linear structured systems. *Systems & Control Letters*, 2017, 102: 42-47.

[38] Wu L, Shen Y, Li M *et al*. Drug target identification based on structural output controllability of complex networks. In *Proc. the 10th International Symposium Bioinformatics Research and Applications*, June 2014, pp.188-199.

[39] Gao J, Liu Y Y, D'Souza R M *et al*. Target control of complex networks. *Nature Communications*, 2014, 5: Article No. 5415.

[40] Ogata K. Modern Control Engineering (3rd edition). Prentice Hall, 1996.

[41] Hopcroft J E, Karp R M. An $n^{5/2}$ algorithm for maximum matchings in bipartite graphs. *SIAM Journal on Computing*, 1973, 2(4): 225-231.

[42] Zhang X, Lv T, Yang X *et al*. Structural controllability of complex networks based on preferential matching. *PLoS One*, 2014, 9(11): Article No. e112039.

[43] Goodrich M T, Tamassia R. Algorithm Design: Foundation, Analysis and Internet Examples. John Wiley & Sons, 2006.

[44] Wu L, Tang L, Li M *et al*. The MSS of complex networks with centrality based preference and its application to biomolecular networks. In *Proc. the 2016 IEEE International Conference on Bioinformatics and Biomedicine*, December 2016, pp.229-234.

[45] Pequito S, Kar S, Aguiar A P. On the complexity of the constrained input selection problem for structural linear systems. *Automatica*, 2015, 62: 193-199.

[46] Lindmark G, Altafini C. Controllability of complex networks with unilateral inputs. *Scientific Reports*, 2017, 7: Article No. 1824.

[47] Rugh W J, Kailath T. Linear System Theory (2nd edition). Pearson, 1995.

[48] Wang L Z, Chen Y Z, Wang W X *et al*. Physical controllability of complex networks. *Scientific Reports*, 2017, 7: Article No. 40198.

[49] Li G, Tang P, Wen C *et al*. Boundary constraints for minimum cost control of directed networks. *IEEE Transactions on Cybernetics*, 2017, 47(12): 4196-4207.

[50] Czeizler E, Gratie C, Chiu W K *et al*. Target controllability of linear networks. In *Proc. the 14th International Conference on Computational Methods in Systems Biology*, September 2016, pp.67-81.

[51] Kuhn H W. The Hungarian method for the assignment problem. *Naval Research Logistics Quarterly*, 1955, 2(1/2): 83-97.

[52] Zhang X, Wang H, Lv T. Efficient target control of complex networks based on preferential matching. *PLoS One*, 2017, 12(4): Article No. e0175375.

[53] Liu X, Pan L, Stanley H E *et al*. Controllability of giant connected components in a directed network. *Physical Review E*, 2017, 95(4): Article No. 042318.

[54] Piao X, Lv T, Zhang X *et al*. Strategy for community control of complex networks. *Physica A: Statistical Mechanics and Its Applications*, 2015, 421: 98-108.

[55] Guo W F, Zhang S W, Wei Z G *et al*. Constrained target controllability of complex networks. *Journal of Statistical Mechanics: Theory and Experiment*, 2017, 2017(6): Article No. 063402.

[56] Khazanchi R, Dempsey K, Thapa I *et al*. On identifying and analyzing significant nodes in protein-protein interaction networks. In *Proc. the 23rd IEEE International Conference on Data Mining Workshops*, December 2013, pp.343-348.

[57] Badhwar R, Bagler G. Control of neuronal network in *caenorhabditis elegans*. *PLoS One*, 2015, 10(9): Article No. e0139204.

[58] Noori H R, Schöttler J, Ercsey-Ravasz M *et al*. A multiscale cerebral neurochemical connectome of the rat brain. *PLoS Biology*, 2017, 15(7): Article No. e2002612.

[59] Deisseroth K. Circuit dynamics of adaptive and maladaptive behaviour. *Nature*, 2014, 505(7483): 309-317.

[60] Kringelbach M L, Jenkinson N, Owen S L *et al*. Translational principles of deep brain stimulation. *Nature Reviews Neuroscience*, 2007, 8(8): 623-635.

[61] Li F, Long T, Lu Y *et al*. The yeast cell-cycle network is robustly designed. *Proceedings of the National Academy of Sciences of the United States of America*, 2004, 101(14): 4781-4786.

[62] Davidich M I, Bornholdt S. Boolean network model predicts cell cycle sequence of fission yeast. *PLoS One*, 2008, 3(2): Article No. e1672.

[63] Moes M, Le Béchec A, Crespo I *et al*. A novel network integrating a miRNA-203/SNAI1 feedback loop which regulates epithelial to mesenchymal transition. *PLoS One*, 2012, 7(4): Article No. e35440.

[64] Krumsiek J, Marr C, Schroeder T *et al*. Hierarchical differentiation of myeloid progenitors is encoded in the transcription factor network. *PLoS One*, 2011, 6(8): Article No. e22649.

[65] Mendoza L. A network model for the control of the differentiation process in Th cells. *Biosystems*, 2006, 84(2): 101-114.

[66] Lee H J, Takemoto N, Kurata H *et al*. Gata-3 induces T helper cell type 2 (Th2) cytokine expression and chromatin remodeling in committed Th1 cells. *Journal of Experimental Medicine*, 2000, 192(1): 105-116.

[67] Szabo S J, Kim S T, Costa G L *et al*. A novel transcription factor, T-bet, directs Th1 lineage commitment. *Cell*, 2000, 100(6): 655-669.

[68] Hwang E S, Szabo S J, Schwartzberg P L *et al*. T helper cell fate specified by kinase-mediated interaction of T-bet with GATA-3. *Science*, 2005, 307(5708): 430-433.

[69] Kanhaiya K, Czeizler E, Gratie C *et al*. Controlling directed protein interaction networks in cancer. Technical Report, Turku Centre for Computer Science, 2017. http://tucs.fi/publications/attachment.php?fname=tKaCzGrPe16a.full.pdf, Nov. 2018.

[70] Wu L, Tang L, Li M *et al*. Biomolecular network controllability with drug binding information. *IEEE Transactions on Nano Bioscience*, 2017, 16(5): 326-332.

[71] Jia T, Liu Y Y, Csóka E *et al*. Emergence of bimodality in controlling complex networks. *Nature Communications*, 2013, 4: Article No. 2002.

[72] Liu X, Pan L. Identifying driver nodes in the human signaling network using structural controllability analysis. *IEEE/ACM Transactions on Computational Biology and Bioinformatics*, 2015, 12(2): 467-472.

[73] Jia T, Barabási A L. Control capacity and a random sampling method in exploring controllability of complex networks. *Scientific Reports*, 2013, 3: Article No. 2354.

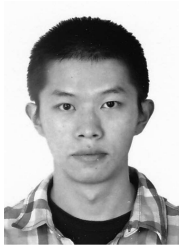
[74] Liu X, Pan L. Detection of driver metabolites in the human liver metabolic network using structural controllability analysis. *BMC Systems Biology*, 2014, 8(1): Article No. 51.

[75] Vinayagam A, Gibson T E, Lee H J *et al*. Controllability analysis of the directed human protein interaction network identifies disease genes and drug targets. *Proceedings of the National Academy of Sciences of the United States of America*, 2016, 113(18): 4976-4981.

[76] Matsuoka Y, Matsumae H, Katoh M *et al*. A comprehensive map of the influenza A virus replication cycle. *BMC Systems Biology*, 2013, 7(1): Article No. 97.

[77] Uhart M, Flores G, Bustos D. M. Controllability of protein-protein interaction phosphorylation-based networks: Participation of the hub 14-3-3 protein family. *Scientific Reports*, 2016, 6: Article No. 26234.

[78] Ravindran V, Sunitha V, Bagler G. Identification of critical regulatory genes in cancer signaling network using controllability analysis. *Physica A: Statistical Mechanics and Its Applications*, 2017, 474: 134-143.
[79] Ruths J, Ruths D. Control profiles of complex networks. *Science*, 2014, 343(6177): 1373-1376.
[80] Tu C, Rocha R P, Corbetta M *et al*. Warnings and caveats in brain controllability. *Neuroimage*, 2017, 176: 83-91.
[81] Vanunu O, Magger O, Ruppin E *et al*. Associating genes and protein complexes with disease via network propagation. *PLoS Computational Biology*, 2010, 6(1): Article No. e1000641.
[82] Wang B, Gao L, Gao Y. Control range: A controllability-based index for node significance in directed networks. *Journal of Statistical Mechanics: Theory and Experiment*, 2012, 2012(04): Article No. P04011.
[83] Wang B, Gao L, Gao Y *et al*. Controllability and observability analysis for vertex domination centrality in directed networks. *Scientific Reports*, 2014, 4: Article No. 5399.



Lin Wu received his B.Sc. degree in computer science and technology from Central South University, Changsha, in 2012. Currently, he is working toward his Ph.D. degree in the Division of Biomedical Engineering at the University of Saskatchewan, Saskatoon, Canada. His current research interests include analysis of large-scale biological data, modeling, analysis, and control of biomolecular networks.

Min Li received her B.S. degree in communication engineering and her M.S. and Ph.D. degrees in computer science, all from Central South University, Changsha, in 2001, 2004, and 2008, respectively. She is currently a professor at the School of Information Science and Engineering, Central South University, Changsha. Her main research interests include bioinformatics and systems biology.

Jian-Xin Wang received his B.Eng. and M.Eng. degrees in computer engineering from Central South University, Changsha, in 1992 and 1996, respectively, and his Ph.D. degree in computer science from Central South University, Changsha, in 2001. He is a professor in School of Information Science and Engineering, Central South University, Changsha. His current research interests include algorithm analysis and optimization, parameterized algorithm, bioinformatics, and computer networks.

Fang-Xiang Wu received his B.Sc. and M.Sc. degrees in applied mathematics from Dalian University of Technology, Dalian, in 1993 and 1990, respectively, and two Ph.D. degrees, the first one in control theory and its applications from Northwestern Polytechnical University, Xi'an, in 1998, and the second one in bioinformatics and computational biology at the University of Saskatchewan (U of S), Saskatoon, in August 2004. During Sept. 2004–Aug. 2005, he worked as a post-doctoral fellow (PDF) with the Laval University Biomedical Research Center. He is currently a full professor in School of Information Science and Engineering, Central South University, Changsha. His research mainly focuses on bioinformatics and computational systems biology, specifically, biomolecular network analytics, machine learning in bioinformatics, big biological/clinical data analytics, complex bionetwork controllability for treatment and diagnosis of complex diseases, brain networks, and nonlinear biodynamic analytics. So far he has published more than 270 refereed journal/conference papers. He has been invited to give research talks in about 60 international conferences or academic institutes worldwide. Dr. Wu is serving as the editorial board member of five international journals, the guest editor of numerous international journals, and the program committee chair or member of over 30 international conferences. He is a senior member of IEEE.