they can be experimentally adjusted by applying targeted drugs. Under the constraint that the control parameters can be perturbed in given ranges, attractor networks are constructed, providing complete information about the feasibility of driving the system from one attractor to another and consequently a quantitative understanding of the controllability of the underlying network. There is also preliminary evidence that noise can enhance the controllability significantly.

CONCLUSION

The ability to control complex networks is of uttermost importance to many critical problems in science and engineering, and has the potential to generate great technological breakthroughs. We argue that it is possible to develop a controllability framework for complex, non-linear dynamical networks based on the idea of attractor networks. For the field of complex dynamical systems, the framework will lead to landscape changes, revolutionizing our ability to control the systems. Another field that will benefit tremendously from this is systems and synthetic biology, where a basic problem is to control GRNs. To be able to control complex, non-linear networks will also have significant impacts on fields such as computer networks, wireless networks, cybersecurity, biological networks, and even social and economical networks.

Our framework of attractor networks has the appealing features that (a) it is applicable to non-linear dynamical networks in general, (b) the attractor network can possess quite simple structure even for large, complex networks, and (c) noise can enhance the controllability. The benefit of noise, while counterintuitive, has its origin in well-known phenomena in non-linear dynamical systems such as stochastic resonance and coherence resonance. There are, however, difficulties with the attractor network framework. For example, for large networks the construction of an attractor network may be quite challenging—the scalability issue. In addition, the structure of the attractor network in general depends on the system parameters. While we emphasize the need to focus on physically meaningful and experimentally accessible parameter perturbations, there can still be a large number of attractor networks depending on the parameters, making it difficult to formulate a rigorous mathematical framework. We believe that these issues can and will be satisfactorily addressed in the near future, finally realizing the grand goal of controlling non-linear dynamical networks.

ACKNOWLEDGEMENTS

I thank Mr. Riqi Su and Ms. Lezhi Wang for discussions. The idea of ‘attractor network’ was originated from their current Ph.D. research at Arizona State University.

FUNDING

This work was supported by Air Force Office of Scientific Research (FA9550-10-1-0083) and by National Sanitation Foundation (CD11-1026710).

Yang-Cheng Lai
School of Electrical, Computer and Energy Engineering, Arizona State University, USA
E-mail: Ying-Cheng.Lai@asu.edu

REFERENCES

1. Ott, E, Grebogi, C and Yorke, JA. Phys Rev Lett 1990; 64: 1196–9.
2. Boccaletti, S, Grebogi, C, Lai, YC et al. Phys Rep 2000; 329: 103–87.
3. Lombardi, A and Hörnquist, M. Phys Rev E 2007; 75: 056110.
4. Liu, YY, Slotine, JJ and Barabási, AL. Nature 2011; 473: 167–73.
5. Yuan, ZZ, Zhao, C, Di, ZR et al. Nat Commun 2013; 4: 2447.
6. Kalman, RE. SIAM J Appl Math 1963; 1: 152–92.
7. Lin, CT. IEEE Trans Autom Control 1974; 19: 201–8.
8. Slotine, JJ and Li, W. Applied Nonlinear Control. New Jersey: Prentice-Hall, 1991.
9. Wang, XF and Chen, G. Physica A 2002; 310: 521–31.
10. Li, X, Wang, XF and Chen, G. IEEE Trans. Circuits Syst. I 2004; 51: 2074–87.
11. Sorrentino, F, di Bernardo, M, Garofalo, F et al. Phys Rev E 2007; 75: 046103.
12. Lai, YC and Tél, T. Transient Chaos—Complex Dynamics on Finite-Time Scales. New York: Springer, 2011.

do:10.1093/nsr/nwu023
Advance access publication 14 July 2014

Fair recently, considerable efforts were made to address the controllability issue of complex networks [1]. As a key notion in control theory, controllability concerns our ability to drive a dynamic system from any initial state to any final state in finite time [2], which agrees well with our intuitive notion of control. To model complex networks as dynamic systems, we normally adopt the so-called nodal dynamics, i.e. we associate each node in a network with a state variable, whose time evolution crucially depends on the state variables of the node itself and/or its neighbors.
Nodal dynamics is natural for modeling dynamic processes on many real systems, where the state variable of a node often has a clear physical meaning, e.g. the expression level of genes in transcriptional regulatory networks or the population of species in food webs. Thanks to a combination of tools from graph theory, control theory and statistical physics, a series of results about network controllability with nodal dynamics were obtained (see Fig. 1). For example, the minimum number of driver nodes, whose time-dependent control can guide the system’s dynamics, is mainly determined by the degree distribution of the network. And surprisingly, driver nodes tend to avoid hubs, i.e. the nodes with high connectivity. Sparse and heterogeneous networks are harder to be controllable than dense and homogeneous networks [1]. These results open new avenues to deepen our understanding of complex networked systems and have triggered a burst of research activities [3–10].

A complementary approach in studying the controllability of complex networks is to use edge dynamics, where each edge is associated with a state variable [9]. This approach differs from the conventional large-scale systems theory where intra-connections are typically not variables. The edge dynamics is more suitable for modeling networks where nodes are not passive elements but active components with information-processing capabilities. For example, in social communication networks, a node (i.e. an individual) constantly processes the information received from its upstream neighbors and makes decisions that are communicated to its downstream neighbors. The information received and passed by a node can then be represented by the state variables on its incoming and outgoing edges, respectively. Naturally the node itself acts as a switchboard-like device, which maps the signals of the incoming edges onto that of the outgoing edges. Neural network on the macroscopic level is also an ideal candidate for the switchboard-like modeling approach, where the nodes represent anatomically distinct areas of the brain, each consisting of hundreds of thousands of neurons. It has been shown that the controllability properties of this process significantly differ from simple nodal dynamics. For example, driver nodes prefer hubs with large out-degree and heterogeneous networks have better controllability properties than homogeneous networks [9].

Despite those exciting results and the fact that controllability issue is fundamental for network control, designing a practical control law that guides the system to a desired final state in an optimal way is pressed for. We are just at the beginning of a fantastic journey. Yet, there are lots of challenges down the road.

First of all, the dynamics processes on many real-world networks, especially biological networks, are very difficult to quantify. To quantitatively formalize the dynamics of a complex networked system, one has to assume a dynamical model, which usually gets involve a large number of system parameters. The parameter identifiability issue is the very first obstacle one has to surmount [11]. If some parameters are not identifiable even with perfect experimental measurements, the whole dynamical model makes no sense for further control analysis.

Secondly, to design a control law for a dynamic system, one apparently has to know the values of those system parameters, particularly those parameters in edge dynamics which classical large-scale systems theory typically did not consider. But in reality, especially biological networks such as gene regulatory networks, the model itself and the associated parameters are largely unknown, which fundamentally limits our ability to design a suitable control law.

Fortunately, modern molecular biology technologies have been used to generate increasingly large amounts of high-throughput systems-level measurement data for many organisms under a variety
of experimental conditions. One can apply statistical inference and computational learning theory to infer the model structure and eventually formalize the dynamics rules governing biological processes.

Finally, the non-trivial topology of complex networks brings an intrinsic layer of complexity to the control problem. From the advances towards understanding complex networks accumulated in the last decade, we know that network topology fundamentally affects many dynamical processes on it, from epidemic spreading to synchronization phenomenon. We also know that even in the case of linear control, the topological characteristics of the networks have a big impact on their controllability [1]. It is fair to expect that the network topology would definitely affect its controllability in the non-linear case. For example, it has been shown that finding a control strategy leading to the desired global state for Boolean dynamics is computationally intractable (NP-hard) in general, but it can be solved in polynomial time if the network has a tree structure [12].

In sum, our ultimate goal is to develop the mathematical underpinning behind the control of complex networks, unifying under a single theoretical foundational framework. This is a problem that given its complexity and depth of applications will probably engage network science and control community for the next decade.

Yang-Yu Liu1,2
1Channing Division of Network Medicine, Department of Medicine, Brigham and Women’s Hospital, Harvard Medical School, USA
2Center for Cancer Systems Biology, Dana-Farber Cancer Institute, USA
E-mail: yyl@channing.harvard.edu

REFERENCES
1. Liu, YY, Slotine, JJ and Barabási, AL. Nature 2011; 473: 167–73.
2. Kalman, RE. SIAM J Appl Math 1963; 1: 152–92.
3. Chapman, A and Mesbahi, M. On strong structural controllability of networked systems: a constrained matching approach. In: American Control Conference (ACC), Washington, DC, 2013. p. 6126.
4. Yuan, Z, Zhao, C, Di, Z et al. Nat Commun 2013; 4. 2447.
5. Jia, T, Liu, YY, Csoóka, E et al. Nat Commun 2013; 4. 3002.
6. Liu, YY, Slotine, JJ and Barabási, AL. PLoS One 2012; 7. e44459.
7. Liu, YY, Slotine, JJ and Barabási, AL. Proc Natl Acad Sci USA 2013; 110: 2460–5.
8. Yan, G, Ren, J, Lai, YC et al. Phys Rev Lett 2012; 108: 218703.
9. Nepusz, T and Vicsek, T. Nat Phys 2012; 8: 568–73.
10. Wang, WX, Ni, X, Lai, YC et al. Phys Rev E 2012; 85: 1–5.
11. Ljung, L and Glad, T. Automatica 1994; 30: 265–76.
12. Akutsu, T, Hayashida, M, Ching, WK et al. J Theor Biol 2007; 244: 670–9.

doi: 10.1093/nsr/nwu025
Advance access publication 12 July 2014

Special Topic: Network Science

Scientometrics: untangling the topics
Ádám Szántó-Várnagy1, Péter Pollner2,3, Tamás Vicsek1,2,3 and Illés J. Farkas2,3,*

Measuring science is based on comparing articles to similar others. However, keyword-based groups of thematically similar articles are dominantly small. These small sizes keep the statistical errors of comparisons high. With the growing availability of bibliographic data, such statistical errors can be reduced by merging methods of thematic grouping, citation networks and keyword co-usage.

Pieces of our collective human scientific knowledge are constantly defined and modified through our global scientific communication. The most common units of this process are publications, also called articles or papers. These units (i) provide ‘road signs’ for newcomers to a field and (ii) allow the scientific community to steer its work toward consensus-based goals given the available resources. Due to the size of science automated measurements are necessary to achieve these two goals. In particular, the steering aspect involves decisions about manuscript acceptance and science funding, which includes even jobs of scientists. Thus, it seems reasonable to move to the public domain not only scientometric algorithms but also bibliographic data [1]. With more data in the public domain, our current assumptions about the data itself may be challenged.

To measure science, one needs to measure the scientific communication process, which is a network of articles (nodes) connected by citations (directed links) and tagged with article keywords. Most current scientific metrics are built on article-level metrics (ALMs) and the most common ALM is the (total) citation number. The citation number—similarly to other mention-counting ALMs—has the following major properties. First, there are more publications every year (Fig. 1a) and the number of references per publication is growing too (Fig. 1b). Second, papers with an earlier publication date have had until now more time to receive citations. Third, the citation count by itself blanks out citation context [2], which includes citing paper quality. In summary, the citation number tends to favor papers that appeared close (in time and topic) to the origins of large and still active research areas. Improvements...