Quantifying robustness of biological networks using NS-2

Bhanu K. Kamapantula*, Ahmed Abdelzaher*, Sajal K. Das† and Preetam Ghosh*

*Department of Computer Science, Virginia Commonwealth University, Richmond, VA 23284.

Email: kamapantulbk@mymail.vcu.edu, abdelzaheraf@mymail.vcu.edu, pghosh@vcu.edu

†Computer Science Department, Missouri University of Science and Technology. Email: sdas@mst.edu

A large section of the work in computational models of biological systems is based on classical chemical kinetic (CCK) formalism based on a set of ordinary differential equations (ODE), also known as reaction rate equations or mass action kinetics [13]. Representing a homogeneous biological system as a set of biochemical reactions, the temporal dynamics of the molecular species is studied in the continuous-deterministic domain. While continuous-deterministic reaction models are capable of capturing behavorial dynamics for spatially homogeneous systems with large number of molecular species, the inherent stochasticity observed in many biological processes (gene expression and protein synthesis) have proven the limitation of CCK in accurately representing biological processes. Arkin et.al in [13] have shown the limitations of CCK in several common biological scenarios, where stochastic reaction dynamics present a more accurate picture of the systems behavior.

[3] describes the advantages of using discrete event simulators for modeling biological systems. A fundamental challenge in computational systems biology [7] is the simplification of the biological system complexity without losing the ensemble dynamic behavior. In the system engineering view of complex processes [14], the key notion is to abstract the complexity of the system as a set of discrete time and space variables (random variables), which capture the behavior of the system in time. The entire system is a collection of functional blocks or modules, which are driven by a set of events, where an event defines a large number of micro level state transitions between a set of state variables accomplished within the event execution time. The underlying assumption driving this abstraction is the segregation of the complete state space into such disjoint sets of independent events which can be executed simultaneously without any interaction. The application of this technique in large complex communication networks has demonstrated the accuracy of the approach for the first and higher order dynamics of the system within the limits of input data and state partitioning algorithms [15]. For example, discrete event based system modeling has been effectively applied for designing routers, the key components responsible for routing traffic through the Internet. Discrete event based simulation techniques have also been used in a wide variety of manufacturing processes and studying the system dynamics of complex industrial processes.

Network simulator, NS-2 [4], is a discrete event simulator widely used for studying wireless networks. NS-2 has been used by researchers to model communication in wireless networks and embedded devices. This simulator continues to evolve with the active support of research community. Taking a step forward, we have used NS-2 as an in-silico platform for quantifying

the robustness of biological networks. Specifically, since the primary objective of a wireless sensor network is information transport to specific sink nodes and they operate under similar noisy and error prone conditions as biological networks, we the define robustness of biological networks by the ability for each node in the network to deliver packets to the sink node with minimal packet loss. Before envisioning a model for any time-varying functional biological system, it is important to illustrate the preliminary model for the biological system in NS-2. While exclusive simulators to model a molecular network are not present currently, existing simulators can be adjusted to model the desired network. It should be noted that this might not be the perfect approach but the opportunity to explore the qualitative and quantitative dynamics of molecular networks is not lost. Scenarios are presented below to demonstrate the use of NS-2 to quantify biological robustness.

Consider a biological network topology derived from a well studied organism, Escherichia coli (E. coli). Sub-networks that are extracted from E. coli comprise of interactions among genes. Let us call this extracted network a Gene Regulatory Network (GRN). Such GRNs comprise two classes of nodes: transcription factors and genes. A transcription factor either up-regulates or down-regulates one (or more) gene. The packet transmission rates are assumed to be identical, in NS-2, for all the non-sink nodes, however, in a real biological setting, such rates are directly proportional to the rate constants associated with every edge in the network along with the concentration of the molecules associated with a node. This however creates a roadblock for existing biological network simulators as each of these rate constants need to be experimentally validated which is not currently feasible for the different sample networks generated in this work. The simulation also assumes all packets transmitted to be identical in type and size which correspond to similar signaling molecules affecting the different nodes in the GRN in the context of biological robustness. Queue limit in NS-2 is useful to limit the number of packets that can be queued at a node. Queue limit in the corresponding mapped GRN represents the half-life of each signal sent from one node to another node. Although this is another approximation in the simulation set-up, it is impossible to characterize all such signaling molecules accurately in the different extracted GRNs. In summary, our proposed NS-2 set-up makes broad assumptions for the pertinent details of biological network signaling but we feel that this is indeed necessary for studying the qualitative dynamics of many sample GRNs wherein such details are not known at length. All these assumptions to use NS-2 for a biological network have been outlined in [1].

Traditionally, robustness of biological networks has been measured by its static graph theoretic characteristics such as network diameter, average shortest path [12], network efficiency [9] amongst others. A network with negligible change in its diameter is considered to be robust when it loses node(s) after an attack. Similarly, negligible change in average shortest path and network efficiency under network perturbations related to temporal fluctuations in the node and/or link availability is attributed to robust networks. Additionally, we have introduced another metric to measure robustness, packet receival rate at the sink node in [5] to determine their signal transmission dynamics. Packet receival rate is the ratio of the number of packets received in the network to the number of packets sent in the network. Higher the packet receival rate of a GRN, higher its robustness. Sink node selection strategy is critical for optimal GRN performance. In [5], we listed three sink selection strategies: (a) Highest Degree (HD), (b) Highest Coverage (HC) and (c) Motif-based (MB) and identified HD strategy as the best approach to provide higher robustness for NS-2 based simulation of GRNs. Nodes with highest degree are selected as a sink node in the HD strategy. Node involved in any three-node motif is selected as a sink node in the HC strategy. Such three-node

motifs have been earlier identified as the building blocks of robust GRNs [11] from a purely topological perspective and the feed-forward loops were reported to have the most significant impact on GRN robustness. Hence, we also considered nodes involved most in a feed-forward loop (FFL) motif as a sink node in the MB strategy. We considered FFL motifs as they have been identified to play an important role in establishing robustness [8] apart from ensuring important biological functions such as generating signal pulses, and speeding up or delaying response times in target genes [10]. In [6], we compared several GRN-derived networks with randomly generated networks (network sizes 100, 150, 200, 250 and 300) and showed that GRN-derived networks improve the transmission reliability in our NS-2 based simulation setting. In order to test the performance of such GRN-derived networks, comparisons are also made for large-scale predicted GRNs (network size 1500, 1750, 2000, 2250 and 2500) and a similar trend is observed. This might be possible due to the presence of higher number of FFLs in GRN-derived networks as compared to randomly generated networks.

While the above network evaluations establish the significance of GRN-derived networks, only one sink operates in those networks which is not the case in functional GRNs. In [1], we used multiple sink nodes to model GRN communication. An SVM model, built using LibSVM [2], is then used to investigate the relative efficiency of packet receival rates based on topological metrics such as network density, genes coverage, transcription factor network density, motif abundance and genes percentage.

Our simulation setup using NS-2 is generic and can be applied to any GRN (e.g: E. coli, Yeast), and thus provides a common platform to assess dynamic robustness of biological networks. This also allows to sample several extracted and predicted GRN topologies and measure their signal transmission dynamics thereby identifying specific topological and control properties in these networks that impact their robustness. Such a platform will hence allow one to compare the robustness of the GRN topologies of different organisms, design, validate, test and explore different GRN prediction algorithms besides also serving the greater complex networks community by applying such design rules of robust biological networks to create fault-tolerant and efficient engineered systems.

REFERENCES

[1] Ahmed Abdelzaher, Bhanu Kamapantula, Preetam Ghosh, and Sajal Das. A theoretical framework to quantify robustness in biological network topologies. January 2013.

[2] Chih-Chung Chang and Chih-Jen Lin. Libsvm: a library for support vector machines. *ACM Transactions on Intelligent Systems and Technology (TIST)*, 2(3):27, 2011.

[3] S Ghosh, P Ghosh, K Basu, SK Das, and S Daefler. A discrete event based stochastic simulation platform for in silicostudy of molecular-level cellular dynamics. *J Biotechnol Biomaterial S*, 6:2, 2011.

[4] Information-Sciences Institute. Ns-2. http://isi.edu/nsnam/ns.

[5] Bhanu K Kamapantula, Ahmed Abdelzaher, Preetam Ghosh, M Mayo, E Perkins, and Sajal K Das. Performance of wireless sensor topologies inspired by e. coli genetic networks. In *Pervasive Computing and Communications Workshops (PERCOM Workshops), 2012 IEEE International Conference on*, pages 302–307. IEEE, 2012.

[6] Bhanu K Kamapantula, Ahmed Abdelzaher, Preetam Ghosh, Michael Mayo, Edward J Perkins, and Sajal K Das. Leveraging the robustness of genetic networks: a case study on bio-inspired wireless sensor network topologies. *Journal of Ambient Intelligence and Humanized Computing*, pages 1–17.

[7] Hiroaki Kitano. Computational systems biology. *Nature*, 420(6912):206–210, 2002.

[8] Hiroaki Kitano. Towards a theory of biological robustness. *Molecular systems biology*, 3(1), 2007.

[9] Vito Latora and Massimo Marchiori. *The architecture of complex systems*. Oxford UP, 2004.

[10] Shmoolik Mangan and Uri Alon. Structure and function of the feed-forward loop network motif. *Proceedings of the National Academy of Sciences*, 100(21):11980–11985, 2003.

[11] Ron Milo, Shai Shen-Orr, Shalev Itzkovitz, Nadav Kashtan, Dmitri Chklovskii, and Uri Alon. Network motifs: simple building blocks of complex networks. *Science*, 298(5594):824–827, 2002.

[12] Alex KS Ng and Janet Efstathiou. Structural robustness of complex networks. *Physical Review*, 3:175–188, 2006.

[13] Michael S Samoilov and Adam P Arkin. Deviant effects in molecular reaction pathways. *Nature biotechnology*, 24(10):1235–1240, 2006.

[14] Bernard P Zeigler, Herbert Praehofer, Tag Gon Kim, et al. *Theory of modeling and simulation*, volume 19. John Wiley New York, 1976.

[15] Xiang Zeng, Rajive Bagrodia, and Mario Gerla. Glomosim: a library for parallel simulation of large-scale wireless networks. In *Parallel and Distributed Simulation, 1998. PADS 98. Proceedings. Twelfth Workshop on*, pages 154–161. IEEE, 1998.