Reconstructing the Stochastic Evolution Diagram of Dynamic Complex Systems

Navid Bazzazzadeh, Benedikt Brors, Roland Eils
German Cancer Research Center (DKFZ)
Im Neuenheimer Feld 580, D-69120 Heidelberg, Germany
{n.bazzazzadeh, b.brors, r.eils}@dkfz.de
Tel: +49 (0)6221 42 3604

Abstract

The behavior and dynamics of complex systems are in focus of many research fields. The complexity of such systems comes not only from the number of their elements, but also from the unavoidable emergence of new properties of the system, which are not just a simple summation of the properties of its elements. The behavior of complex systems can be fitted with a number of well developed models, which, however, do not incorporate the modularity and the evolution of a system simultaneously. In this work, we propose a generalized model that addresses this issue. Our model is developed within the Random Set Theory's framework and allows for reconstructing the stochastic evolution diagrams of complex systems.

Recent advancements of technology in many scientific fields have resulted in production of datasets with massive amounts of variables that characterize the properties of complex systems. One can find the respective examples in such research fields as systems biology, neuroscience, social science, economics, astronomy, etc. (Smith et al. 2006; Vázquez et al. 2004; Newman 2006; Hartwell et al. 1999). There is a variety of methods to mathematically describe the data obtained from observations on complex systems. Representation of the system under study in the form of a network which shows the associations between variables (nodes) has attained particularly much attention.

Modularity and Dynamics in Complex Networks

A complex system is comprised of an unknown number of processes (components) whose dynamics is not derivable from the summation of the dynamics of individual processes. Complex systems are studied in a variety of research areas. In many of these areas, complex systems demonstrate strong similarities, with the large topological change and natural division into a modular structure being the most common features (Vázquez et al. 2004; Newman 2006; Hartwell et al. 1999; Luscombe et al. 2004). The problem of detecting and characterizing the dynamics of modular structure of a complex system has an outstanding importance (Newman 2006), and has motivated the present work.

A modular network is a network which is well divided into modules such that there are dense internal connections between nodes within modules but only sparse connections between different modules. In many real complex systems the topology of the underlying network is not static and can rapidly evolve over time. Assume that a module in a network is a subset of the set of nodes. Then a network (regardless of edges) is a set of subsets of nodes, and, regarding to the dynamic topology of complex networks, we can briefly say that a network at each time step is a random set of random sets (modules). (It should be emphasized that in this paper we examine an underlying network or a module by considering only nodes not the edges.)

Random Finite Sets (RFS)

Let $U = \{u_1, \ldots, u_n\}$ and $n < \infty$. Then a random variable which takes its value from the universal sample space 2^U is called a random finite set (Mahler 2007). A measure ($\mathbf{X} : 2^U \rightarrow [0 \quad 1]$) can be defined by assigning probabilities $m_x(A) \triangleq P(X = A) = P_X(\{A\})$ directly to each $A \in 2^U$, and the belief mass function for a random set A is defined as $\beta_X(A) \triangleq P(X \subset A) = \sum_{B \subset A} m_x(B)$. The belief mass function plays the same role in random finite set statistics as the cumulative distribution plays in random vector statistics (Mahler 2007).

Let us define the phase as a period of time during which the elements of a module (a RFS) does not change, but the state of each element is allowed to evolve over this period. Then the phase transition can be defined as an event when new elements appear or old elements disappear in the module. A random-cluster approach is widely used to model systems which have phase transitions or, more generally, systems with a graph structure. Cluster processes are a concept in the theory of point processes, and are described as a superposition of point processes of a cluster (Swain and Clark 2010). Loosely speaking, analogous to a Markov process, a cluster process is a memoryless time-varying RFS, i.e. $p(mo_{k+1}|mo_k, \ldots, mo_1) = p(mo_{k+1}|mo_k)$, where mo_k is the state of a module of the network at time k .

Also analogous to the hidden Markov model (HMM), a hidden-set Markov model (HSMM) is a model in which the system is assumed to undergo a cluster process with unobserved (hidden) states. The Mahler's finite set statistics (Mahler 2007) generalizes the Bayesian framework for the study of random set-valued variables, and provides a means to estimate the state of a time sequence of random finite sets which are generated from an assumed cluster process.

Network as a Random Finite Set of Modules

We have shown how to characterize the uncertainty of a module in a network by modeling the module state and the module measurements as random finite sets (RFS). We have also formulated the corresponding motion model and observation model, and mentioned that a network itself is a random finite set of modules with its own dynamics. Therefore, understanding the dynamics of a complex system that has a dynamic-topology underlying network, is the problem of characterizing the uncertainty of the underlying network, or in other words, detecting, identifying, classifying and estimating (tracking) the states of modules and their nodes at each time point.

RFS Bayesian Estimators

The optimal Bayesian random finite set estimator is capable of recursive propagation of the RFS posterior density in time (Vo, Singh, and Doucet 2005). However, it is not practical to obtain a sequence of states of RFS due to computational issues. Several computationally feasible approaches have been proposed as an alternative to approximate RFS Bayesian recursive estimator (Mahler 2007). Analogous to the Kalman filter, which is the most successful approximation method for matching the two first order moments (mean and covariance) of the Bayesian estimator, the first moment of the recursive RFS Bayesian estimator is the *Probability Hypothesis Density* (PHD) (Mahler 2003).

The Stochastic Evolution Diagram

A group of elements (e.g., a module composed of nodes in a network) unavoidably develops properties which are not a simple summation of the properties of its elements (a phenomenon widely known as emergence). Hypothetically, the nodes of a module are coordinated. The way they are coordinated can change over time, and can be described with the help of a "virtual leader". A wide variety of parameters can be used as virtual leaders, for instance geometric centroid (Clark and Godsill 2007) or parameters of the probability distribution of nodes of a module. We have shown how to estimate the life-time parameters of a module and the corresponding virtual leader. Life-time parameters of a module, such as the module's birth-time, death-time, and spawning times, are considered as a phase transition.

Similar to the multivariate Markov model of time series that constructs parallel Markov chains, the multiple hidden-set Markov model reconstructs the *stochastic evolution diagram*. We introduce this term to denote a collection of Markov chains, in which some of the chains are tied together at certain time points. Each edge in the graph is a Markov chain. A schematic example of a stochastic evolution diagram is illustrated in Figure 1.

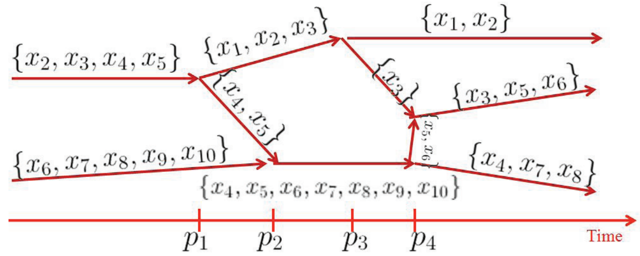

Figure 1: Schematic illustration of a stochastic evolution diagram, where an edge is a trajectory of a module's virtual leader, and the labels over the edges are modules' elements, and p's are phase transitions.

References

Clark, D., and Godsill, S. 2007. Group Target Tracking with the Gaussian Mixture Probability Hypothesis Density Filter. *2007 3rd International Conference on Intelligent Sensors Sensor Networks and Information* 149–154.

Hartwell, L. H.; Hopfield, J. J.; Leibler, S.; and Murray, A. W. 1999. From molecular to modular cell biology. *Nature* 402(6761):C47–52.

Luscombe, N. M.; Babu, M. M.; Yu, H.; Snyder, M.; Teichmann, S. A.; and Gerstein, M. 2004. Genomic analysis of regulatory network dynamics reveals large topological changes. *Nature* 431(7006):308–312.

Mahler, R. P. S. 2003. Multitarget bayes filtering via first-order multitarget moments. *Ieee Transactions On Aerospace And Electronic Systems* 39(4):1152–1178.

Mahler, R. P. S. 2007. *Statistical multisource-multitarget information fusion*. Artech House Publishers.

Newman, M. E. J. 2006. Modularity and community structure in networks. *Proceedings of the National Academy of Sciences of the United States of America* 103(23):8577–8582.

Smith, V. A.; Yu, J.; Smulders, T. V.; Hartemink, A. J.; and Jarvis, E. D. 2006. Computational Inference of Neural Information Flow Networks. *PLoS Computational Biology* 2(11):14.

Swain, A., and Clark, D. E. 2010. First-moment filters for spatial independent cluster processes. *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series*.

Vázquez, A.; Dobrin, R.; Sergi, D.; Eckmann, J. P.; Oltvai, Z. N.; and Barabasi, A. L. 2004. The topological relationship between the large-scale attributes and local interaction patterns of complex networks. *Proceedings of the National Academy of Sciences of the United States of America* 101(52):17940–17945.

Vo, B.-N.; Singh, S.; and Doucet, A. 2005. Sequential monte carlo methods for multi-target filtering with random finite sets. *Ieee Transactions On Aerospace And Electronic Systems* 41(4):1224–1245.