	Network Science	
Lecture 6: Aug 4, 2009		Lecturer: Henning Mortveit
	Scribe: Guanhong Pei	

Network Dynamical Process: SDS II

6.1 Functional Linkage Networks

6.1.1 Preliminary

Two genes may have the same function if their protein products interact or if they have very similar patterns of gene expression.

Definition 1 *FLN*
An *FLN* [1, 2, 3] is a powerful medium for inferring gene function by integrating the evidence captured by proteinprotein interactions and gene-expression data. An FLN is a graph in which each node represents a protein or a gene. A node is labelled by the set of functions that annotate the gene; an edge in an FLN connects two genes if some experimental or computational procedure suggests that these genes might share the same function. Each edge in the FLN has a real-valued weight; the sign of the weight indicates whether the connected genes share or do not share the function, while the magnitude of the weight reflects our confidence in the edge.

6.1.2 Gene Annotation Using Functional Linkage Networks

Goal: Determine if a collection of genes have some set of biological functions.

Idea: Use existing knowledge of what genes express what functions to predict functions of new genes.

Modeling approach:

1. Represent genes as vertices and connect genes if there is sufficient experimental/computational evidence indicating that they share biological functions as in Figure 6.1.
2. Each edge has a weight $w \in [-1, 1]$ encoding the degree of of co-expression.
3. A new biological function is given, and it is known that some subset of the genes expresses this function.

4. Assign a state +1 to genes that express the new function and −1 to genes that do not express function f. Assign state 0 to states of remaining vertices.

$$k = \begin{cases} -1 & - \quad \text{not expressed;} \\ 0 & - \quad \text{no idea;} \\ 1 & - \quad \text{expressed.} \end{cases}$$

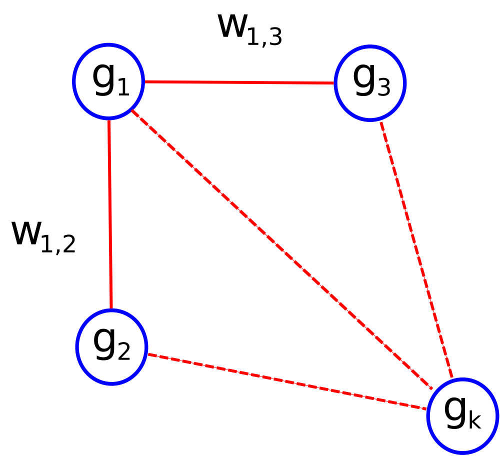

Figure 6.1: Gene function prediction.

Given the functional linkage network (FLN) and the initial configuration $x = \{x_i\}$, assign states ± 1 to all the genes with state 0 so as to minimize the energy function

$$E(x) = -\sum_{i,j} w_{i,j} x_i x_j$$

Observation: Locally, neighbor genes with identical states should be connected by a positive edge and neighbor genes with opposite states should be connected by a negative edge.

The algorithm in [3] computes an approximation to such a minimal state through an iterative sequential scheme.

1. Select (at random) a permutation π of the set V of vertices initially in state 0.
2. Repeat: Update the state of each of vertex $v \in V$ asynchronously using the update sequence π and the vertex function

$$x_i \xrightarrow{f_i} sign \sum_{i,j} w_{j,i} x_i.$$

The algorithm stops when successive configurations are identical, i.e., "convergence".

Observation: This algorithm corresponds to a fixed point computation of a threshold SDS.

Fact: The SDS has only fixed points as attractors (i.e. no periodic orbits of period no smaller than 2).

6.2 NOR Permutation SDS

Definition 2 *NOR Function*

$$nor(x_1, x_2, \dots, x_k) = \begin{cases} 1 & \text{if all x's are 0;} \\ 0 & \text{otherwise.} \end{cases}$$

Can we have a fixed point?
For sequence π, function nor_π, x is a fixed point if $nor(x) = x$.
There is no fixed point.

Note: *transient length* is at most 1 for any point. Periodic points are in a one to one correspondence with the independent sets of the graph Y.

What characterizes a periodic point for *nor* permutation SDS?
Initial state: $x = (x_1, x_2, \dots, x_k)$.
To be periodic, a state can only contain isolated 1's.

6.3 Invertibility

When is $F_\pi = F_{\pi(n)} \circ F_{\pi(n-1)} \circ \dots \circ F_{\pi(2)} \circ F_{\pi(1)}$ invertible? Assume $k = \{0, 1\}$.

Definition 3 *Invertible*
take y, there is a unique x, such that $F(x) = y$. Each F_i should be invertible.

$$F_i(x) = (x_1, x_2, \dots x_{i-1}, f_i(x[i]), x_k)$$

That happens if $g_i = f(-; \cap x_i)$, $K \to K$ is a bijection function for each fixed $\cap x_i$. There are 4 functions.

Implication: If $k \in \{0, 1\}$, and F_π is invertible, $F_\pi^{-1} = F_{inv(\pi)}$. In general, $(F_\pi|_p)^{-1} = F_{inv(\pi)}|_p$.

Fixed point ada independent.

Proposition 1 .
Let F_π be a permutation SDS map and assume $x \in Fix(F_\pi)$. Then $x \in Fix(F_{\pi'})$ for all $\pi \in S_Y$.

Proposition 2 .
A threshold SDS only has fixed points as attractors/limit sets.

Threshold function: $f_k\{0, 1\}^m \to \{0, 1\}$.

$$F_{inv(\pi)} \circ F_\pi(x) = x$$

6.4 Schedule Equivalence

6.4.1 Functional Equivalence

Let $\pi, \sigma \in S_Y$, where $F_\pi = F_\sigma$.

$$F_\pi = F_{\pi(n)} \circ \ldots \circ F_{\pi(1)}$$

If π and σ differ by a transportation of consequtive, non-adjacent orientations, they induce the same SDS. Alternatively, π and σ induce the same SDS map if O_Y^π and O_Y^σ are the same acyclic orientations.

Example: in a $Circ_4$, $F_\pi = F_{\pi(1)} \circ F_{\pi(2)} \circ F_{\pi(4)} \circ F_{\pi(3)}$ and $F_\sigma = F_{\sigma(1)} \circ F_{\sigma(4)} \circ F_{\sigma(2)} \circ F_{\sigma(3)}$ are functional equivalent.

The largest number of distinct permutation SDS maps $\alpha(Y) = |Acyc(Y)|$. Since

$$\alpha(Circ_n) = 2^n - 2,$$

we have $\leq 2^n - 2$ SDS maps.

α is turtle canonical. $\alpha(Y) = \alpha(Y'_e) + \alpha(Y''_e)$, where $\alpha(Y'_e) = Y \setminus \{e\}$ and $\alpha(Y''_e) = Y/\{e\}$.

6.4.2 Other Equivalence

Dynamical Equivalence

Orbit Equivalence

6.5 List of Supplemental Notes

- circ4_nor_equivalences.pdf
- SDS_Application.pdf

References

[1] N. Massjouni, C. Rivera, and T. M. Murali, "Virgo: Computational prediction of gene functions," *Nucleic Acids Research*, vol. 34, pp. W340–W344, 2006.

[2] M. Henning, "Graph dynamical systems - a mathematical framework for interactionbased systems, their analysis and simulations," in *Discrete Models in Systems Biology Workshop*, December 2008.

[3] U. Karaoz, T. M. Murali, S. Letovsky, Y. Zheng, C. Ding, C. R. Cantor, and S. Kasif, "Whole-genome annotation by using evidence integration in functional-linkage networks," *Proceedings of the National Academy of Sciences*, pp. 2888–2893, 2004.