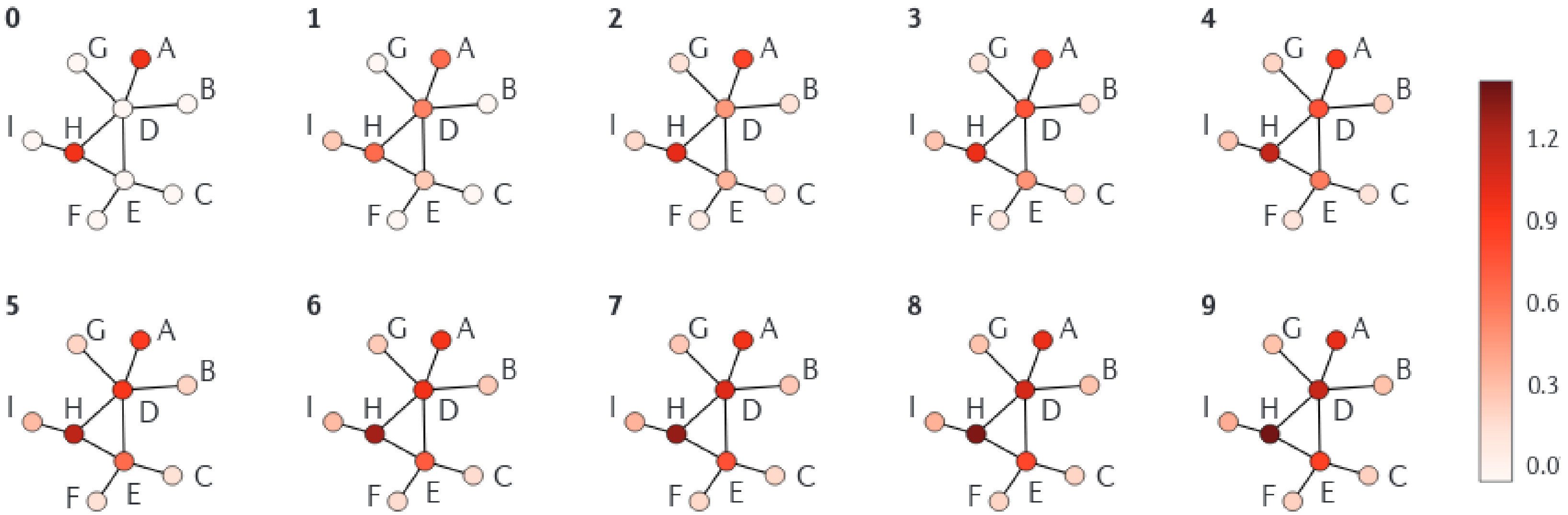
Network Propagation

Network propagation: a universal amplifier of genetic associations

Lenore Cowen[1], Trey Ideker[2], Benjamin J. Raphael[3] and Roded Sharan[4]

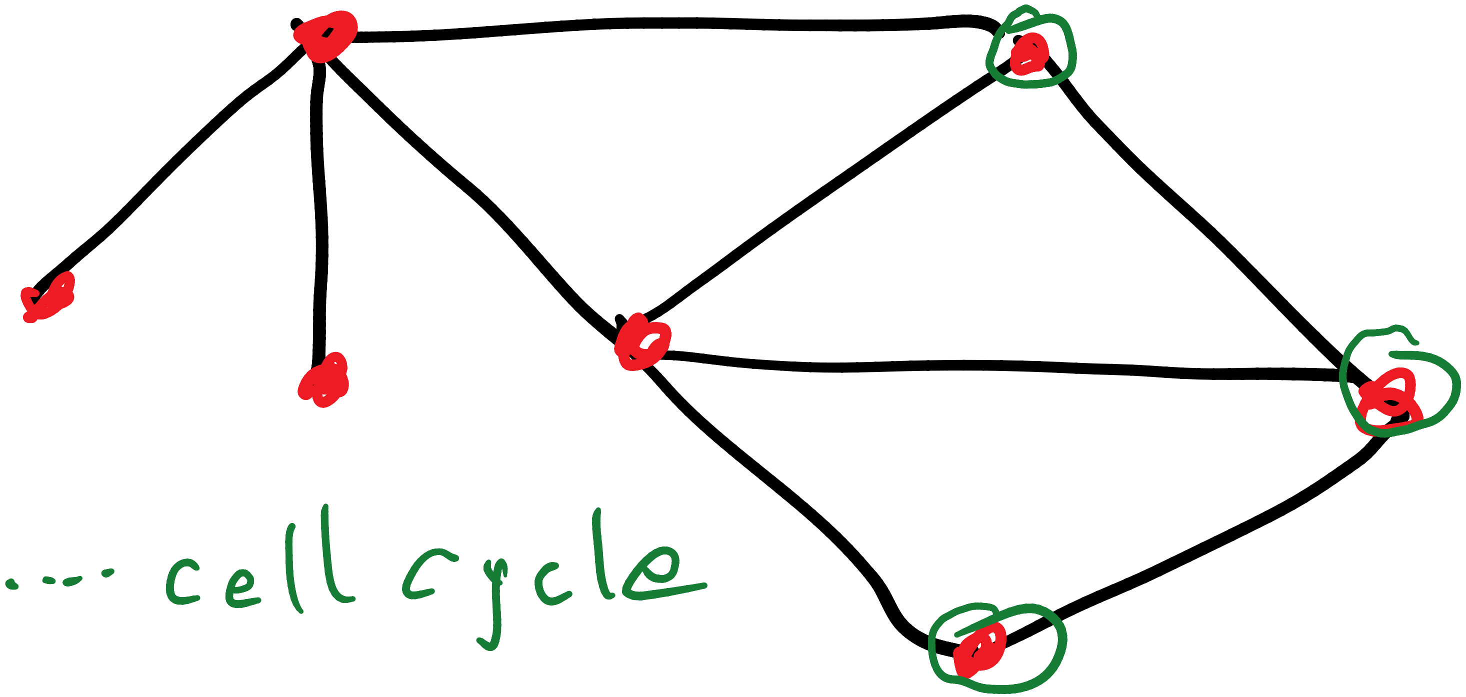


Adapted from: *Cowen et al. (2017) . Nature Reviews Genetics.*

The problem of function prediction

Given:

- protein-protein-interaction network
- Some proteins (=nodes) are annotated with a biological function, e.g. „cell cycle“

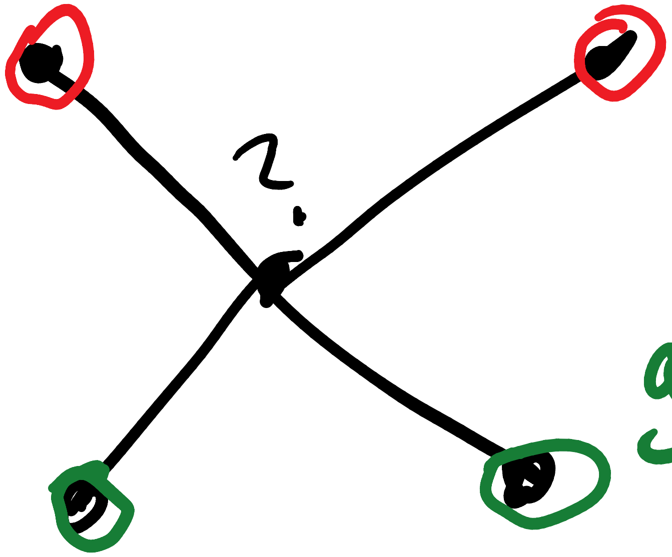


green .... cell cycle

Assumption:

Proteins of the same function (pathway) tend to interact with each other.
I.e., they form a connected component in the PPI graph

red .... cell cycle



green ... apoptosis

Markov chains

(see, e.g., https://www.stat.auckland.ac.nz/~fewster/325/notes/ch8.pdf)

Random variable $X_t, t = 1,2,3 \ldots$ describes where the process is among states 1, …, N at time t.
Transition matrix: $A = (a_{ij})$
Transition matrix describes the probability to jump from state i to state j: $a_{ij} = P(X_{t+1} = j \mid X_t = i)$.

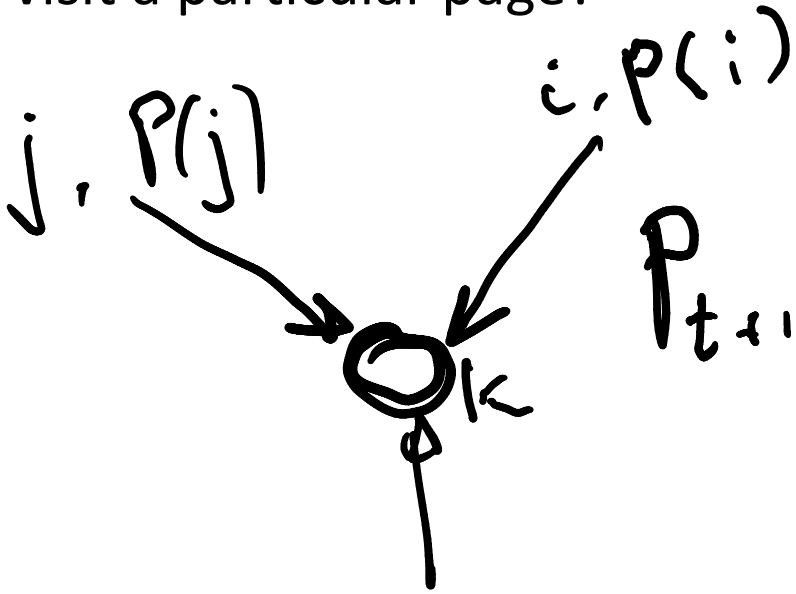
A is stochastic, i.e. the rows sum up to 1.

Google PageRank algorithm (wikipedia „PageRank“)

Which web pages are „important“? Formally: Which pages will be visited
by a randomly clicking through the web more often?

You start at one page and you can follow any of its links with equal probability.

What is the probability to visit a particular page?



$$P_{t+1}(k) = a_{jk} P_t(j) + a_{ik} P_t(i) + \ldots\ldots$$

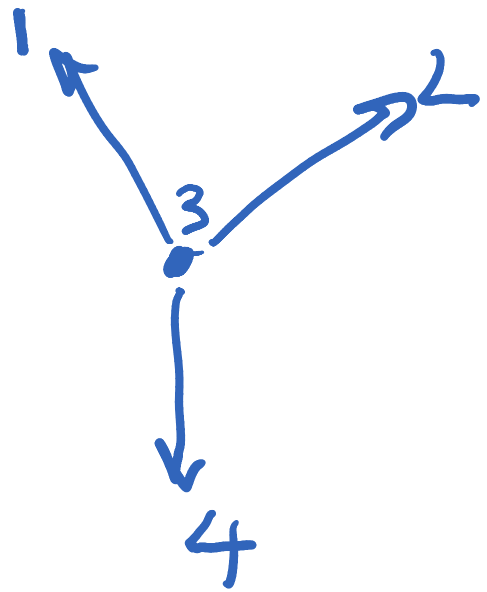
$\vec{p}_t$ ...... N-dim. vector of probabilities of states

$\vec{p}_{t+1} = A p_t$ [correct: $\vec{p}_{t+1} = \vec{p}_t A$]

$p_0 \rightarrow p_1 = A p_0 \rightarrow p_2 = A p_1 = A^2 p_0 \rightarrow$

$\rightarrow p_3 = A^3 p_0 \rightarrow \dots\dots \rightarrow \pi$

Definition of A (preliminary)



$$a_{31} = \frac{1}{3} = \frac{\#\text{edges } 3 \to 1}{\text{outdegree}(3)}$$

$a_{ij} = 0$ if $\nexists$ edge $i \to j$

With this definition: $\sum_{\ell} a_{k\ell} = 1 \quad \forall k$

After the convergence, the process is in equilibrium:

$$\pi = A\pi$$

i.e., π is an eigenvector

(A is stochastic – eigenvalue with π is 1)

New nomenclature:

A …. Adjacency matrix of the graph

For a directed graph: (i,j) entry is one when node i has an edge to node j.
For an undirected graph: Interpret as having edges in both directionsm,
i.e. both (i,j) and (j,i) are 1 (symmetric matrix).

W … Adjacency matrix made stochastic: entries are inverse of node degree.
In matrix notation:
Let D be the diagonal matrix with the node (out)degrees on the diagonal.

$$W = AD^{-1}$$

Internet browsing as a random walk:

$$p_{t+1} = W\, p_t \rightarrow \cdots \rightarrow \ \pi = W\pi$$

After infinitely long web surfing

π is the vector of probabilities of each web page
π corresponds to the importance of a page

Why does this converge to the eigenvector (to the largest eigenvalue)?

- Power method:

  For many kinds of matrices (not necessarily stochastic):

  $W^k\, v \rightarrow a\ multiple\ of\ the\ eigenvector\ (to\ largest\ eigenvalue)$

  Here v can be any random starting vector!

  See wikipedia „Power iteration“.

- Perron-Frobenius theorem: guarantees that this works for a Markov transition matrix and that it converges to probability vector (non-negative, sum 1; see wikipedia „Perron–Frobenius theorem“)

„The transition matrix W defines a random walk on the graph.“

What about the nodes of outdegree 0 („sink“)?

Introduce a pseudo-transition-probability (like the pseudocounts in PWMs and HMMs):

Every transition is possible, although with a small probability

$$\frac{1-\alpha}{N}$$

Other interpretation: In addition to clicking on a link, I can enter a new URL to jump to. With equal probability. **„Random surfer model“**

Taken together:

Let $\mathbf{1}$ be a vector of all 1s.

$$p_t = \frac{1-\alpha}{N}\mathbf{1} + \alpha \, \mathrm{W} \, p_{t-1}$$

$\mathbb{R}^n$

$\vec{v}_2$

$\lambda \vec{v}_1 + (1-\lambda) v_2$

(observation)

$\vec{v}_1$

(prejudice)

$\lambda = 1 \longrightarrow \vec{v}_1$

$\lambda = 0 \longrightarrow \vec{v}_2$

"convex combination"

Is this a Markov chain? What is the transition matrix?

$$p_t = \frac{1-\alpha}{N}\mathbf{1} + \alpha\, \mathrm{W}\, p_{t-1}$$

α is typically 0.85

Fun fact: Let p be a probability vector (sum=1) and let E be a matrix of all 1s. Then $\mathbf{1} = Ep$

Therefore we can write

$$p_t = \frac{1-\alpha}{N}\mathbf{1} + \alpha\, \mathrm{W}\, p_{t-1} = \frac{1-\alpha}{N}\mathrm{E}\, p_{t-1} + \alpha\, \mathrm{W}\, p_{t-1} = \left(\frac{1-\alpha}{N}\mathrm{E} + \alpha\, \mathrm{W}\right) p_{t-1}$$

$\frac{1-\alpha}{N} E + \alpha W$ is a matrix.

But is it stochastic?

$$\frac{1}{N} E = \begin{pmatrix} \frac{1}{N} & \cdots & \frac{1}{N} \\ \vdots & & \\ \frac{1}{N} & \cdots & \frac{1}{N} \end{pmatrix} \cdots\cdots \text{stochastic}$$

W contains $\frac{1}{\text{degree}}$ for edges, 0 elsewhere

– also stochastic

Where do we stand?

Google PageRank defines a Markov chain on the graph of web links.

This Markov chain has an equilibirium distribution,
i.e. eventually there exists a certain probability to visit a particular page.
Or, in other words:
Of my time in the internet, I will spend this proportion on that page.

This equilibrium probabilty vector represents the importance ranking of the web pages.

Let us return to PPI networks and functional annotation

What does the network look like from the viewpoint of a few nodes with given annotation? What is the probability to reach another node from this set of nodes?

Define a starting probability p_0 which has weight exactly on the nodes of the given category. Define RANDOM WALK WITH RESTART (rwr):

$$p_k = (1 - \alpha)p_0 + \alpha W p_{k-1}$$

Yields a ranking for all other nodes indicating how connected an unannotated node is to the given category.

We can also choose p_0 to represent weights (e.g., number of mutations the gene in cancer samples).

Definition of W --- PPI network is undirected

D …. Diagonal matrix with node degrees on the diagonal

A …. Adjacency matrix of the graph

$$W = A\,D^{-1}$$

Like before, we have two ways of solving this:

1) Direct solution:

$$p = \alpha(I - (1-\alpha)W)^{-1}p_0$$

The matrix $\mathrm{F} = \alpha(I - (1-\alpha)W)^{-1}$ is sometimes called diffusion matrix.

2) Iterative

Network Propagation Summary

Random Walk (with restart)

$$p_k = \alpha p_0 + (1-\alpha) W p_{k-1}$$

$$\vdots$$

$$\boxed{p} = \boxed{\alpha(I-(1-\alpha)W)^{-1}}\boxed{p_0}$$

$$\boxed{\begin{pmatrix} S'_1 \\ \vdots \\ S'_n \end{pmatrix}}$$

$$\boxed{W = AD^{-1}}$$

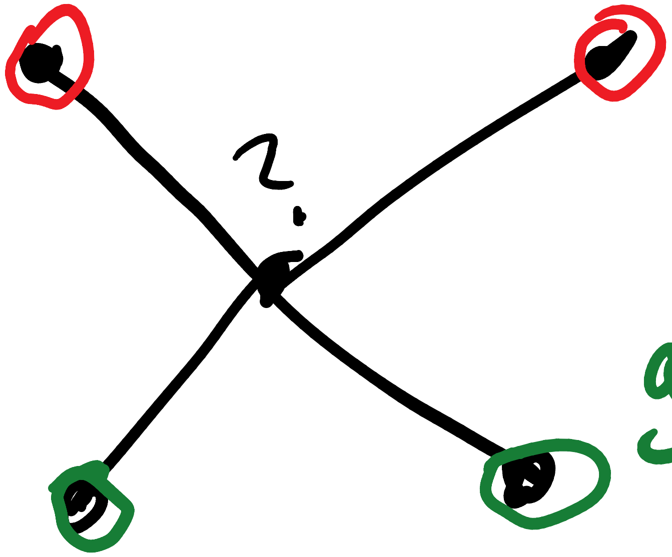
$$\boxed{\begin{pmatrix} S_1 \\ \vdots \\ S_n \end{pmatrix}}$$

$$\boxed{\begin{pmatrix} A_{11} & \dots & A_{1n} \\ \vdots & \ddots & \vdots \\ A_{n1} & \dots & A_{nn} \end{pmatrix}}$$

$$\boxed{\begin{pmatrix} D_1 & 0 & 0 \\ 0 & \ddots & 0 \\ 0 & 0 & D_n \end{pmatrix}}$$

Annotating several categories

Competition between propagated annotation?



red .... cell cycle

green ... apoptosis

Network propagation as clustering in graphs

Try this: All the nodes that are „close“ to an annotated category form a „module“.

Introduce a cut-off and declare everything above the cutoff to belong to one module?

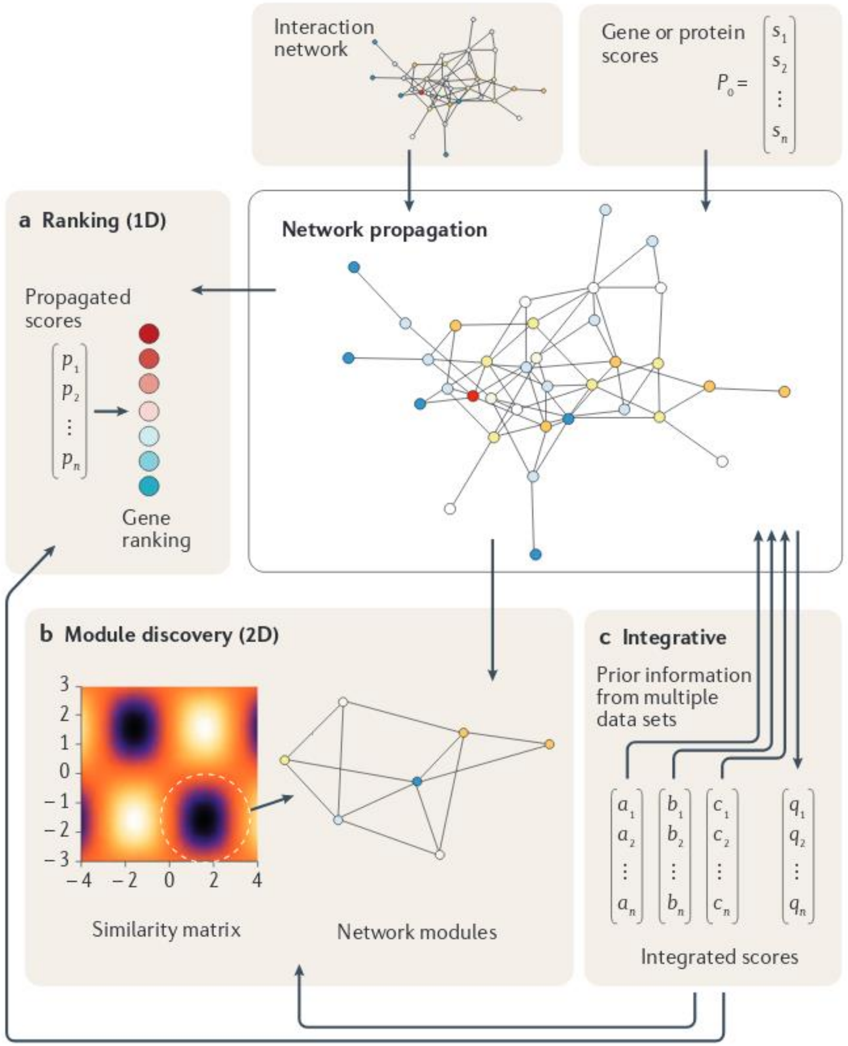
What does hotnet2 do?

Odds and ends:

Heat equation, Graph Laplacian

Diffusion kernel

Network Propagation



Cowen et al. Nature Reviews Genetics 18, 551–562 (2017)


CAVA