(what I think I have understood so far about) Random walks on networks

Vincenzo (Enzo) Nicosia[1]

[1]School of Mathematical Sciences, Queen Mary University of London, UK

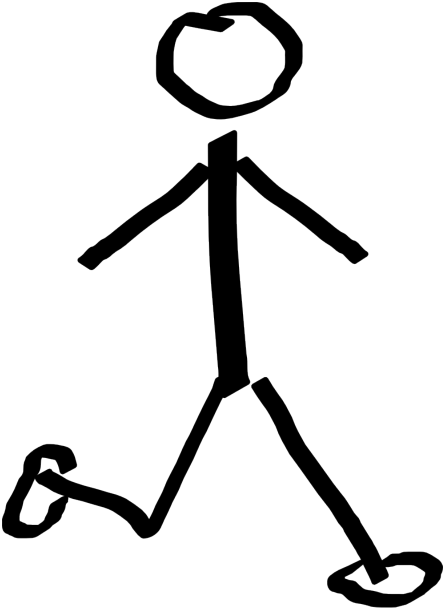
Sept. 30^{th} 2015

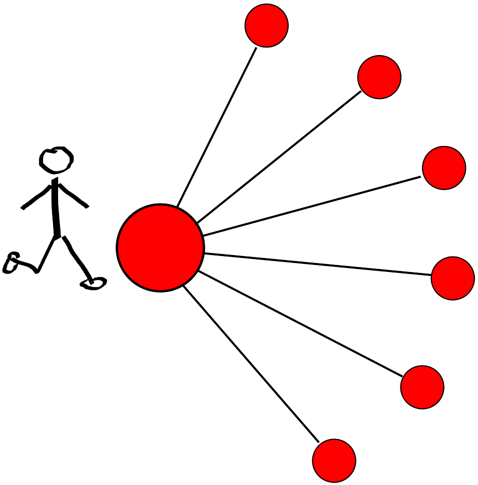
Outline

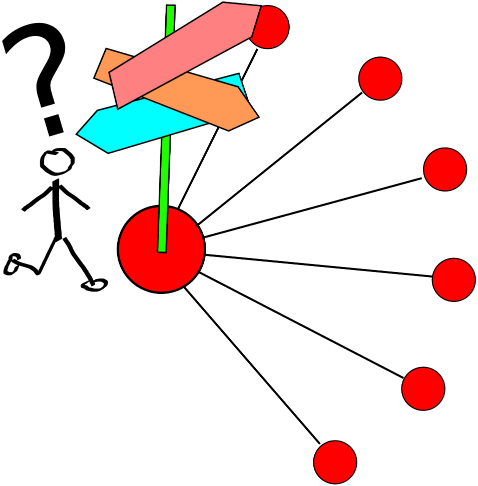
1. Introduction and definitions
2. Plain random walks
3. Variations on the theme
4. Using random walks

Outline

1. Introduction and definitions
2. Plain random walks
3. Variations on the theme
4. Using random walks

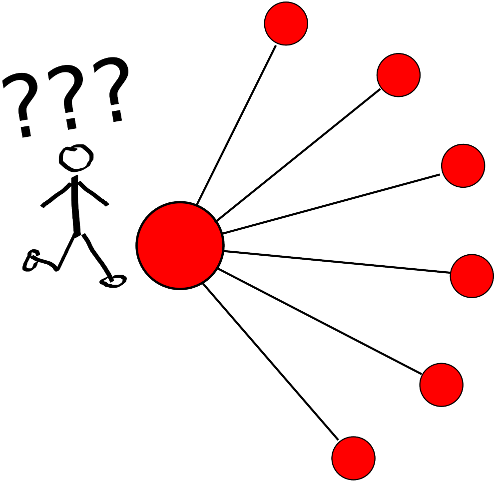


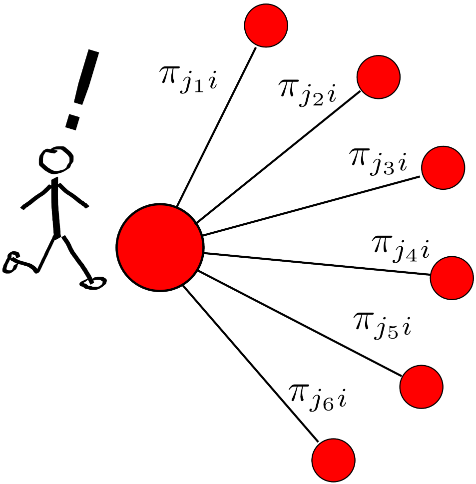




Queen Mary University of London

- π_{ji}: probability for a walker on node i to “jump” on node j in one time step
- $$\sum_j \pi_{ji} = 1$$




!
$\pi_{j_1 i}$
$\pi_{j_2 i}$
$\pi_{j_3 i}$
$\pi_{j_4 i}$
$\pi_{j_5 i}$
$\pi_{j_6 i}$

So, given...

- $G(V, E)$: a simple graph, with $N = |V|$ nodes and $K = |E|$ edges (no self-loops, no multiple edges)

So, given...

- $G(V, E)$: a simple graph, with $N = |V|$ nodes and $K = |E|$ edges (no self-loops, no multiple edges)
- Adjacency matrix: $A = \{a_{ij}\}$
  $a_{ij} = 1$ if node i and node j are connected by an edge ($a_{ij} = 0$ otherwise)

So, given...

- $G(V, E)$: a simple graph, with $N = |V|$ nodes and $K = |E|$ edges (no self-loops, no multiple edges)
- Adjacency matrix: $A = \{a_{ij}\}$
  $a_{ij} = 1$ if node i and node j are connected by an edge ($a_{ij} = 0$ otherwise)
- Transition matrix: $\Pi = \{\pi_{ij}\}$
  π_{ji}: probability for a walker on node i to "jump" on node j in one time step

So, given...

- $G(V, E)$: a simple graph, with $N = |V|$ nodes and $K = |E|$ edges (no self-loops, no multiple edges)
- Adjacency matrix: $A = \{a_{ij}\}$
  $a_{ij} = 1$ if node i and node j are connected by an edge ($a_{ij} = 0$ otherwise)
- Transition matrix: $\Pi = \{\pi_{ij}\}$
  π_{ji}: probability for a walker on node i to "jump" on node j in one time step
- Occupation probability: $p_i(t)$ is the probability of finding a walker on node i at time t

...a random walk is a...

- discrete-time (jumps occur at equally-spaced time steps)

...a random walk is a...

- discrete-time (jumps occur at equally-spaced time steps)
- time-invariant ($\pi_{ji}(t_1) = \pi_{ji}(t_2) \forall t_1, t_2$)

...a random walk is a...

- discrete-time (jumps occur at equally-spaced time steps)
- time-invariant ($\pi_{ji}(t_1) = \pi_{ji}(t_2) \forall t_1, t_2$)
- stochastic process (at each step the identity of the destination node is a random variable)

...a random walk is a...

- discrete-time (jumps occur at equally-spaced time steps)
- time-invariant ($\pi_{ji}(t_1) = \pi_{ji}(t_2) \forall t_1, t_2$)
- stochastic process (at each step the identity of the destination node is a random variable)

A random walk on a graph $G(V, E)$ is a Markov chain defined by the transition matrix Π on the state space V.

Motion rule

- $P(t) = \{p_i(t)\}$ is the occupation probability distribution at time t

Motion rule

- $P(t) = \{p_i(t)\}$ is the occupation probability distribution at time t
- $\sum_{i \in V} p_i(t) = 1 \, \forall t$

Motion rule

- $P(t) = \{p_i(t)\}$ is the occupation probability distribution at time t
- $\sum_{i \in V} p_i(t) = 1 \; \forall t$
- One-step evolution:

$$p_i(t+1) = \sum_j \pi_{ij} \times p_j(t)$$

- Or equivalently

$$P(t+1) = \Pi P(t)$$

Motion rule

- $P(t) = \{p_i(t)\}$ is the occupation probability distribution at time t
- $\sum_{i \in V} p_i(t) = 1 \; \forall t$
- One-step evolution:

$$p_i(t+1) = \sum_j \pi_{ij} \times p_j(t)$$

- Or equivalently

$$P(t+1) = \Pi P(t)$$

- τ-step evolution:

$$P(t+\tau) = \Pi P(t+\tau-1) = \Pi^2 P(t+\tau-2) = \ldots = \Pi^\tau P(t)$$

Question: If we know that the walker was at node i at time 0, where can we find it after t time steps??

Stationary occupation probability

- $$P(t+1) = \Pi P(t)$$

Stationary occupation probability

- $$P(t+1) = \Pi P(t) \quad (1)$$
- Does the limit:

$$\lim_{t\to\infty} P(t)$$

exist?

Stationary occupation probability

- $$P(t+1) = \Pi P(t) \tag{1}$$
- Does the limit:
  $$\lim_{t\to\infty} P(t)$$
  exist?
- Can we find a P^* such that $P^* = \Pi P^*$, i.e., a fixed point for the dynamics of Eq. (1)?

Perron-Frobenius theorem...

Given a non-negative irreducible (aperiodic) matrix $M = \{m_{ij}\}$

- The largest eigenvalue $\lambda_{\max}$ of M (in modulus) is real and positive
- $\lambda_{\max}$ is simple
- The eigenvector associated to $\lambda_{\max}$ is the only positive eigenvector of M

...plus the power method...

- If M is a non-negative irreducible aperiodic matrix, then the sequence of vectors

$$x(t+1) = Mx(t)$$

converges to a vector $\tilde{x}$ which is parallel to the eigenvector associated to the largest eigenvalue of M

...give the answer

- $$P(t+1) = \Pi P(t) \tag{2}$$

...give the answer

- $$P(t+1) = \Pi P(t) \quad (2)$$

- If Π is irreducible and aperiodic (G is connected and contains one odd cycle)

...give the answer

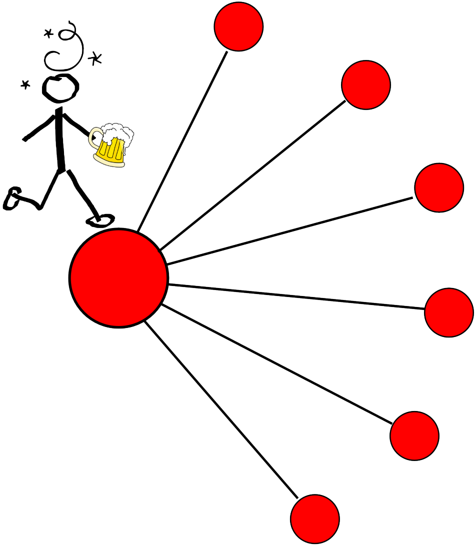
- $$P(t+1) = \Pi P(t) \qquad (2)$$
- If Π is irreducible and aperiodic (G is connected and contains one odd cycle)
- $\Longrightarrow$
  $$\lim_{t\to\infty} P(t)$$
  exists and is equal to the first eigenvector of Π

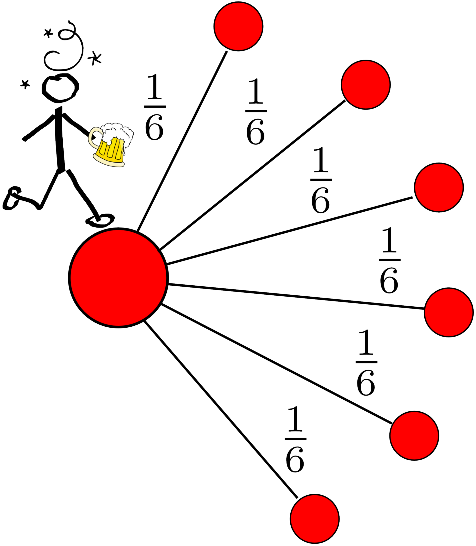
...give the answer

- $$P(t+1) = \Pi P(t) \tag{2}$$
- If Π is irreducible and aperiodic (G is connected and contains one odd cycle)
- $\Longrightarrow$
  $$\lim_{t\to\infty} P(t)$$
  exists and is equal to the first eigenvector of Π
- which is positive and is called the stationary occupation probability distribution associated to Π

Outline

Simplest example: "The Drunken"




$\frac{1}{6}$
$\frac{1}{6}$
$\frac{1}{6}$
$\frac{1}{6}$
$\frac{1}{6}$
$\frac{1}{6}$

The drunken equation (Plain random walk)

- Transition probability:

$$\pi_{ji} = \frac{a_{ij}}{\sum_j a_{ij}} = \frac{a_{ij}}{k_i}$$

The drunken equation (Plain random walk)

- Transition probability:

$$\pi_{ji} = \frac{a_{ij}}{\sum_j a_{ij}} = \frac{a_{ij}}{k_i}$$

- Stationary probability distribution:

$$p_i^* = \frac{k_i}{2K}$$

P^* for plain random walks

- Probability of going from i to j in t steps:

$$W_{i\to j}(t) = \sum_{j_1,j_2,\ldots,j_{t-1}} \pi_{j_1,j} \times \pi_{j_2,j_1} \times \ldots \times \pi_{j,j_{t-1}}$$

P^* for plain random walks

- Probability of going from i to j in t steps:

$$W_{i\to j}(t) = \sum_{j_1,j_2,\ldots,j_{t-1}} \pi_{j_1,j} \times \pi_{j_2,j_1} \times \ldots \times \pi_{j,j_{t-1}}$$

$$W_{i\to j}(t) = \sum_{j_1,j_2,\ldots,j_{t-1}} \frac{a_{ij_1}}{k_i} \times \frac{a_{j_1j_2}}{k_{j_1}} \times \ldots \times \frac{a_{j_{t-1}j}}{k_{j_{t-1}}}$$

P^* for plain random walks

- Probability of going from i to j in t steps:

$$W_{i \to j}(t) = \sum_{j_1, j_2, \ldots, j_{t-1}} \pi_{j_1, j} \times \pi_{j_2, j_1} \times \ldots \times \pi_{j, j_{t-1}}$$

$$W_{i \to j}(t) = \sum_{j_1, j_2, \ldots, j_{t-1}} \frac{a_{ij_1}}{k_i} \times \frac{a_{j_1 j_2}}{k_{j_1}} \times \ldots \times \frac{a_{j_{t-1} j}}{k_{j_{t-1}}}$$

- Probability of going from j to i in t steps:

$$W_{j \to i}(t) = \sum_{j_1, j_2, \ldots, j_{t-1}} \frac{a_{jj_1}}{k_j} \times \frac{a_{j_1 j_2}}{k_{j_1}} \times \ldots \times \frac{a_{j_{t-1} i}}{k_{i_{t-1}}}$$

...

By comparing $W_{i\to j}$ and $W_{j\to i}$ (undirected graphs):

$$\sum_{j_1,j_2,\ldots,j_{t-1}} \frac{a_{ij_1}}{k_i} \times \ldots \times \frac{a_{j_{t-1}j}}{k_{j_{t-1}}} \qquad \sum_{j_1,j_2,\ldots,j_{t-1}} \frac{a_{jj_1}}{k_j} \times \ldots \times \frac{a_{j_{t-1}i}}{k_{j_{t-1}}}$$

...

By comparing $W_{i\to j}$ and $W_{j\to i}$ (undirected graphs):

$$\sum_{j_1,j_2,\ldots,j_{t-1}} \frac{a_{ij_1}}{k_i} \times \ldots \times \frac{a_{j_{t-1}j}}{k_{j_{t-1}}} \qquad \sum_{j_1,j_2,\ldots,j_{t-1}} \frac{a_{jj_1}}{k_j} \times \ldots \times \frac{a_{j_{t-1}i}}{k_{j_{t-1}}}$$

we get:

$$W_{i\to j}k_i = W_{j\to i}k_j$$

...

By comparing $W_{i\to j}$ and $W_{j\to i}$ (undirected graphs):

$$\sum_{j_1,j_2,\ldots,j_{t-1}} \frac{a_{ij_1}}{k_i} \times \ldots \times \frac{a_{j_{t-1}j}}{k_{j_{t-1}}} \qquad \sum_{j_1,j_2,\ldots,j_{t-1}} \frac{a_{jj_1}}{k_j} \times \ldots \times \frac{a_{j_{t-1}i}}{k_{j_{t-1}}}$$

we get:

$$W_{i\to j} k_i = W_{j\to i} k_j$$

but if a stationary probability distribution exists, then

$$p_i^* = \lim_{t->\infty} W_{i\to j} \quad (\forall i \in V)$$

...

By comparing $W_{i\to j}$ and $W_{j\to i}$ (undirected graphs):

$$\sum_{j_1,j_2,\ldots,j_{t-1}} \frac{a_{ij_1}}{k_i} \times \ldots \times \frac{a_{j_{t-1}j}}{k_{j_{t-1}}} \qquad \sum_{j_1,j_2,\ldots,j_{t-1}} \frac{a_{jj_1}}{k_j} \times \ldots \times \frac{a_{j_{t-1}i}}{k_{j_{t-1}}}$$

we get:

$$W_{i\to j} k_i = W_{j\to i} k_j$$

but if a stationary probability distribution exists, then

$$p_i^* = \lim_{t->\infty} W_{i\to j} \quad (\forall i \in V)$$

$$p_j^* k_i = p_i^* k_j$$

...

By comparing $W_{i\to j}$ and $W_{j\to i}$ (undirected graphs):

$$\sum_{j_1,j_2,\ldots,j_{t-1}} \frac{a_{ij_1}}{k_i} \times \ldots \times \frac{a_{j_{t-1}j}}{k_{j_{t-1}}} \qquad \sum_{j_1,j_2,\ldots,j_{t-1}} \frac{a_{jj_1}}{k_j} \times \ldots \times \frac{a_{j_{t-1}i}}{k_{j_{t-1}}}$$

we get:

$$W_{i\to j} k_i = W_{j\to i} k_j$$

but if a stationary probability distribution exists, then

$$p_i^* = \lim_{t->\infty} W_{i\to j} \quad (\forall i \in V)$$

$$p_j^* k_i = p_i^* k_j$$

and by imposing the normalisation condition $\sum_j p_j^* = 1$ we get:

$$\sum_j p_j^* k_i = \sum_j p_i^* k_j \quad \Rightarrow \quad p_i^* = \frac{k_i}{2K}$$

Let's time our walker!

- First passage (hitting) time τ_{ij}: the average number of steps required to a walker to travel from node i to node j (notice: $\tau_{ij} \neq \tau_{ji}$!!!)

Let's time our walker!

- First passage (hitting) time τ_{ij}: the average number of steps required to a walker to travel from node i to node j (notice: $\tau_{ij} \neq \tau_{ji}$!!!)
- Return time r_i: the average number of steps required to a walker started at node i to come back to i ($r_i = \tau_{ii}$)

Let's time our walker!

- First passage (hitting) time τ_{ij}: the average number of steps required to a walker to travel from node i to node j (notice: $\tau_{ij} \neq \tau_{ji}$!!!)
- Return time r_i: the average number of steps required to a walker started at node i to come back to i ($r_i = \tau_{ii}$)
- Average first passage time T:

$$T = \frac{1}{N(N-1)} \sum_i \sum_j \tau_{ij}$$

Let's time our walker!

- First passage (hitting) time τ_{ij}: the average number of steps required to a walker to travel from node i to node j (notice: $\tau_{ij} \neq \tau_{ji}$!!!)
- Return time r_i: the average number of steps required to a walker started at node i to come back to i ($r_i = \tau_{ii}$)
- Average first passage time T:

$$T = \frac{1}{N(N-1)} \sum_i \sum_j \tau_{ij}$$

- Average return time R:

$$R = \frac{1}{N} \sum_i r_i$$

Let's time our walker!

- First passage (hitting) time τ_{ij}: the average number of steps required to a walker to travel from node i to node j (notice: $\tau_{ij} \neq \tau_{ji}$!!!)
- Return time r_i: the average number of steps required to a walker started at node i to come back to i ($r_i = \tau_{ii}$)
- Average first passage time T:

$$T = \frac{1}{N(N-1)} \sum_i \sum_j \tau_{ij}$$

- Average return time R:

$$R = \frac{1}{N} \sum_i r_i$$

- Coverage time: average time needed for a walker to visit all the nodes of the graph

Interesting facts

- Return time

$$r_i = \frac{1}{p_i^*}$$

Interesting facts

- Return time

$$r_i = \frac{1}{p_i^*}$$

- .... so for plain random walks we have:

$$r_i = \frac{2K}{k_i}$$

Interesting facts

- Return time

$$r_i = \frac{1}{p_i^*}$$

- .... so for plain random walks we have:

$$r_i = \frac{2K}{k_i}$$

- Fundamental matrix $Z = \{z_{ij}\}$

$$Z = (I - \Pi^\intercal + W)^{-1}$$

$$\tau_{ij} = \frac{z_{jj} - z_{ij}}{p_j^*}$$

Interesting facts

- Return time

$$r_i = \frac{1}{p_i^*}$$

- .... so for plain random walks we have:

$$r_i = \frac{2K}{k_i}$$

- Fundamental matrix $Z = \{z_{ij}\}$

$$Z = (I - \Pi^{\intercal} + W)^{-1}$$

$$\tau_{ij} = \frac{z_{jj} - z_{ij}}{p_j^*}$$

- Coverage time: I am sure you don't want to know!

Entropy rate

- $(i_0, i_1, i_2, \ldots, i_t)$

Entropy rate

- $(i_0, i_1, i_2, \ldots, i_t)$
- $P(i_0, i_1, i_2, \ldots, i_t)$

Entropy rate

- $(i_0, i_1, i_2, \ldots, i_t)$
- $P(i_0, i_1, i_2, \ldots, i_t)$
- $$\sum_{i_0, i_1, \ldots, i_t} P(i_0, i_1, i_2, \ldots, i_t) = 1$$

Entropy rate

- $(i_0, i_1, i_2, \ldots, i_t)$
- $P(i_0, i_1, i_2, \ldots, i_t)$
- $$\sum_{i_0, i_1, \ldots, i_t} P(i_0, i_1, i_2, \ldots, i_t) = 1$$
- Entropy rate:

$$h = \lim_{t\to\infty} -\frac{1}{t} \sum_{i_0, i_1, \ldots, i_t} P(i_0, i_1, i_2, \ldots, i_t) \log P(i_0, i_1, i_2, \ldots, i_t)$$

Entropy rate

- $(i_0, i_1, i_2, \ldots, i_t)$
- $P(i_0, i_1, i_2, \ldots, i_t)$
- $$\sum_{i_0, i_1, \ldots, i_t} P(i_0, i_1, i_2, \ldots, i_t) = 1$$
- Entropy rate:

$$h = \lim_{t\to\infty} -\frac{1}{t} \sum_{i_0, i_1, \ldots, i_t} P(i_0, i_1, i_2, \ldots, i_t) \log P(i_0, i_1, i_2, \ldots, i_t)$$

- h measures the dispersiveness of the walk

Entropy rate

- $(i_0, i_1, i_2, \ldots, i_t)$
- $P(i_0, i_1, i_2, \ldots, i_t)$
- $$\sum_{i_0, i_1, \ldots, i_t} P(i_0, i_1, i_2, \ldots, i_t) = 1$$
- Entropy rate:

$$h = \lim_{t\to\infty} -\frac{1}{t} \sum_{i_0, i_1, \ldots, i_t} P(i_0, i_1, i_2, \ldots, i_t) \log P(i_0, i_1, i_2, \ldots, i_t)$$

- h measures the dispersiveness of the walk
- h is maximal when all the trajectories of infinite length have equal probability

Entropy rate

- $(i_0, i_1, i_2, \ldots, i_t)$
- $P(i_0, i_1, i_2, \ldots, i_t)$
- $$\sum_{i_0, i_1, \ldots, i_t} P(i_0, i_1, i_2, \ldots, i_t) = 1$$
- Entropy rate:

$$h = \lim_{t\to\infty} -\frac{1}{t} \sum_{i_0, i_1, \ldots, i_t} P(i_0, i_1, i_2, \ldots, i_t) \log P(i_0, i_1, i_2, \ldots, i_t)$$

- h measures the dispersiveness of the walk
- h is maximal when all the trajectories of infinite length have equal probability

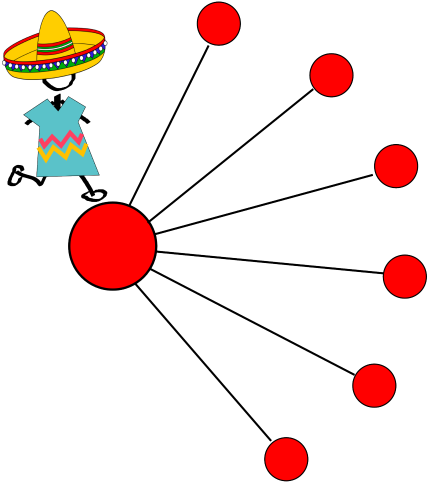
Outline

First Problem

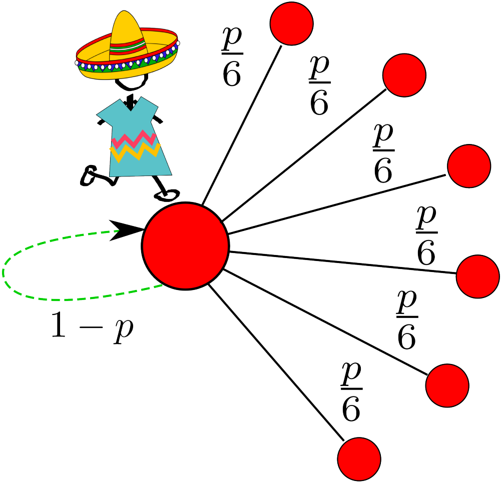
G is not primitive $\Longrightarrow$ no stationary occupation probability (limit cycles)

The Lazy

Lazy random walk



Lazy random walk



Moves with probability p

Remains still with probability

$1-p$

The “Lazy” equation (lazy RW)

- $$P(t+1) = ((1-p)I + p\Pi)P(t)$$

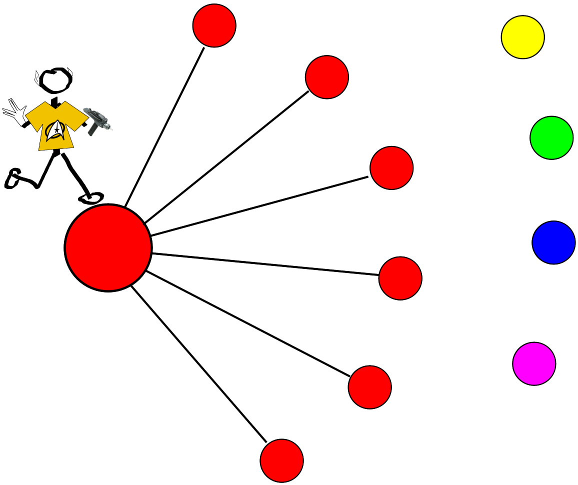
The "Lazy" equation (lazy RW)

- $$P(t+1) = ((1-p)I + p\Pi)P(t)$$
- If G is connected then $M = (1-p)I + p\Pi$ is primitive (degenerate odd-length cycles)

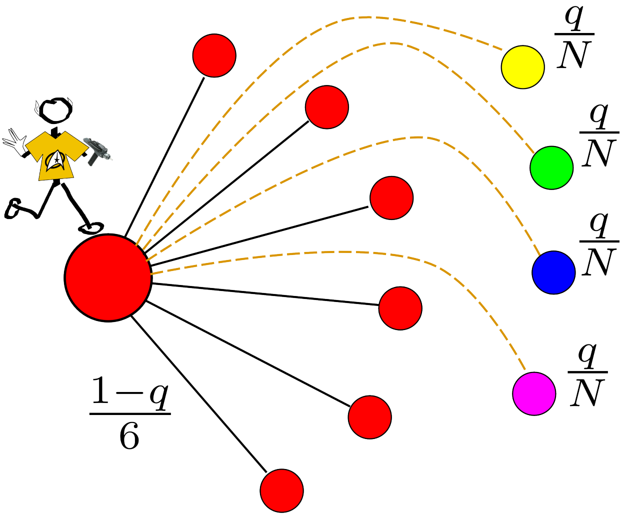
Second problem:

G is directed $\Longrightarrow$ walkers will remain trapped on nodes having $k_{out} = 0$ and nodes with $k_{in} = 0$ will never be visited

The "Smart"



Queen Mary
University of London



The “Smart” equation (RW with teleportation)

- The walker moves to one node uniformly chosen at random in the graph (teleportation) with probability q...

The "Smart" equation (RW with teleportation)

- The walker moves to one node uniformly chosen at random in the graph (teleportation) with probability q...
- ... or it uses the standard transition matrix of the plain random walk with probability $(1 - q)$

The "Smart" equation (RW with teleportation)

- The walker moves to one node uniformly chosen at random in the graph (teleportation) with probability q...
- ... or it uses the standard transition matrix of the plain random walk with probability $(1-q)$
- In formula:

$$\phi_{ji}(q) = \frac{q}{N} + (1-q)\pi_{ji}$$

The “Smart” equation (RW with teleportation)

- The walker moves to one node uniformly chosen at random in the graph (teleportation) with probability q...
- ... or it uses the standard transition matrix of the plain random walk with probability $(1-q)$
- In formula:

$$\phi_{ji}(q) = \frac{q}{N} + (1-q)\pi_{ji}$$

- which in vectorial form reads:

$$\Phi(q) = \frac{q}{N}\mathbf{1}\,\mathbf{1} + (1-q)\Pi$$

The "Smart" equation (RW with teleportation)

- The walker moves to one node uniformly chosen at random in the graph (teleportation) with probability q...
- ... or it uses the standard transition matrix of the plain random walk with probability $(1-q)$
- In formula:

$$\phi_{ji}(q) = \frac{q}{N} + (1-q)\pi_{ji}$$

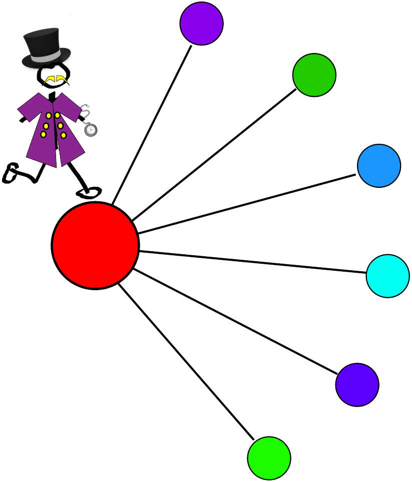
- which in vectorial form reads:

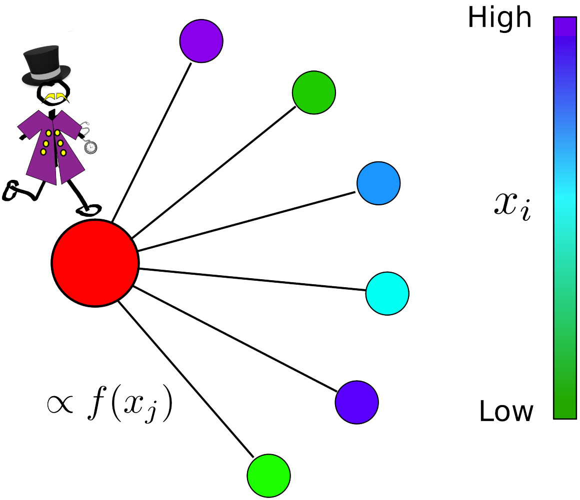
$$\Phi(q) = \frac{q}{N}\mathbf{1}\,\mathbf{1} + (1-q)\Pi$$

Third problem:

In plain walks $p_i^* \propto k_i$ (WWW)

The "Snob"




$\propto f(x_j)$
High
x_i
Low

The "Snob" equation (Biased RW)

- Transition probability:

$$\pi_{ji} = \frac{a_{ij} f_j}{\sum_\ell a_{i\ell} f_\ell}$$

where $f_i = f(x_i)$

The "Snob" equation (Biased RW)

- Transition probability:

$$\pi_{ji} = \frac{a_{ij} f_j}{\sum_\ell a_{i\ell} f_\ell}$$

where $f_i = f(x_i)$

- Stationary probability distribution:

$$p_i^* = \frac{c_i f_i}{\sum_j c_j f_j}$$

where:

$$c_i = \sum_\ell a_{i\ell} f_\ell$$

The "Snob" equation (Biased RW)

- Transition probability:

$$\pi_{ji} = \frac{a_{ij} f_j}{\sum_\ell a_{i\ell} f_\ell}$$

where $f_i = f(x_i)$

- Stationary probability distribution:

$$p_i^* = \frac{c_i f_i}{\sum_j c_j f_j}$$

where:

$$c_i = \sum_\ell a_{i\ell} f_\ell$$

- Entropy rate:

$$h = \frac{\sum_i f_i \sum_j a_{ij} f_j \log(f_j) - \sum_i f_i c_i \log(c_i)}{\sum_i c_i f_i}$$

Degree-biased random walks

- If $f_i = k_i^{\alpha}$ we get degree-biased random walks

Degree-biased random walks

- If $f_i = k_i^{\alpha}$ we get degree-biased random walks
- Transition probability

$$\pi_{ji} = \frac{a_{ij} k_j^{\alpha}}{\sum_{\ell} a_{i\ell} k_{\ell}^{\alpha}}$$

Degree-biased random walks

- If $f_i = k_i^\alpha$ we get degree-biased random walks
- Transition probability

$$\pi_{ji} = \frac{a_{ij} k_j^\alpha}{\sum_\ell a_{i\ell} k_\ell^\alpha}$$

- $\alpha > 0$: the walker moves preferentially towards hubs

Degree-biased random walks

- If $f_i = k_i^\alpha$ we get degree-biased random walks
- Transition probability

$$\pi_{ji} = \frac{a_{ij} k_j^\alpha}{\sum_\ell a_{i\ell} k_\ell^\alpha}$$

- $\alpha > 0$: the walker moves preferentially towards hubs
- $\alpha < 0$: the walker moves preferentially towards poorly-connected nodes

Degree-biased random walks

- If $f_i = k_i^\alpha$ we get degree-biased random walks
- Transition probability

$$\pi_{ji} = \frac{a_{ij} k_j^\alpha}{\sum_\ell a_{i\ell} k_\ell^\alpha}$$

- $\alpha > 0$: the walker moves preferentially towards hubs
- $\alpha < 0$: the walker moves preferentially towards poorly-connected nodes
- $\alpha = 0$: plain random walk

Stationary distribution of degree-biased walks

- $$p_i^* = \frac{c_i f_i}{\sum_\ell c_\ell f_\ell}$$

Stationary distribution of degree-biased walks

- $$p_i^* = \frac{c_i f_i}{\sum_\ell c_\ell f_\ell}$$

- $$p_i^* = \frac{k_i^\alpha \sum_j a_{ij} k_j^\alpha}{\sum_j k_j^\alpha \sum_\ell a_{j\ell} k_\ell^\alpha}$$

Stationary distribution of degree-biased walks

- $$p_i^* = \frac{c_i f_i}{\sum_\ell c_\ell f_\ell}$$

- $$p_i^* = \frac{k_i^\alpha \sum_j a_{ij} k_j^\alpha}{\sum_j k_j^\alpha \sum_\ell a_{j\ell} k_\ell^\alpha}$$

- The stationary probability p_i^* depends both on the degree of i and on the degree of the neighbours of i

Stationary distribution of degree-biased walks

- $$p_i^* = \frac{c_i f_i}{\sum_\ell c_\ell f_\ell}$$
- $$p_i^* = \frac{k_i^\alpha \sum_j a_{ij} k_j^\alpha}{\sum_j k_j^\alpha \sum_\ell a_{j\ell} k_\ell^\alpha}$$
- The stationary probability p_i^* depends both on the degree of i and on the degree of the neighbours of i
- $\Longrightarrow$ p^* is determined by degree-degree correlations!

Interesting facts

- If G has no degree correlations:

$$p_i^* = \frac{k_i^{\alpha+1}}{N\langle k^{\alpha+1}\rangle}$$

Interesting facts

- If G has no degree correlations:

$$p_i^* = \frac{k_i^{\alpha+1}}{N\langle k^{\alpha+1}\rangle}$$

- If $\alpha = -1$:

$$p_i^* = \frac{1}{N}$$

Interesting facts

- If G has no degree correlations:

$$p_i^* = \frac{k_i^{\alpha+1}}{N\langle k^{\alpha+1}\rangle}$$

- If $\alpha = -1$:

$$p_i^* = \frac{1}{N}$$

  Return time:

$$r_i = N$$

  The mean first-passage time T is minimal

Outline

1. Introduction and definitions
2. Plain random walks
3. Variations on the theme
4. Using random walks

The Explorer

Random walks and modularity

- A partition $\mathcal{P}$ of the graph in communities

Random walks and modularity

- A partition $\mathcal{P}$ of the graph in communities
- Plain random walk $\pi_{ji} = \frac{a_{ij}}{k_i}$

Random walks and modularity

- A partition $\mathcal{P}$ of the graph in communities
- Plain random walk $\pi_{ji} = \frac{a_{ij}}{k_i}$
- $P(\mathcal{C}, t)$ is the probability that a walker which started inside the community $\mathcal{C}$ is found inside community $\mathcal{C}$ after t timesteps

Random walks and modularity

- A partition $\mathcal{P}$ of the graph in communities
- Plain random walk $\pi_{ji} = \frac{a_{ij}}{k_i}$
- $P(\mathcal{C}, t)$ is the probability that a walker which started inside the community $\mathcal{C}$ is found inside community $\mathcal{C}$ after t timesteps
- $P(\mathcal{C}, \infty)$ is the probability that a walker started within $\mathcal{C}$ is found in C after an infinite number of steps...

Random walks and modularity

- A partition $\mathcal{P}$ of the graph in communities
- Plain random walk $\pi_{ji} = \frac{a_{ij}}{k_i}$
- $P(\mathcal{C}, t)$ is the probability that a walker which started inside the community $\mathcal{C}$ is found inside community $\mathcal{C}$ after t timesteps
- $P(\mathcal{C}, \infty)$ is the probability that a walker started within $\mathcal{C}$ is found in C after an infinite number of steps...
- Stability of a partition $\mathcal{P}$:

$$R(t) = \sum_{\mathcal{C}\in\mathcal{P}} P(\mathcal{C}, t) - P(\mathcal{C}, \infty)$$

- Stability of a partition $\mathcal{P}$ at time t:

$$R(t) = \sum_{\mathcal{C}\in\mathcal{P}} P(\mathcal{C}, t) - P(\mathcal{C}, \infty)$$

Queen Mary University of London

- Stability of a partition $\mathcal{P}$ at time t:

$$R(t) = \sum_{\mathcal{C}\in\mathcal{P}} P(\mathcal{C}, t) - P(\mathcal{C}, \infty)$$

- At the stationary state $P(\mathcal{C}, \infty)$ is the probability of finding two walkers in the same community $\mathcal{C}$
- If we consider $t = \Delta\tau = 1$ (discrete-time random walk), then we have:

$$R(1) = \sum_{i,j} \frac{a_{ij}}{k_i}\frac{k_i}{2K}\delta_{c_i,c_j} - p_i^* p_j^* \delta_{c_i,c_j} = \frac{1}{2K}\sum_{i,j}\left(a_{ij} - \frac{k_i k_j}{2K}\right)\delta_{c_i,c_j}$$

Queen Mary University of London

- Stability of a partition $\mathcal{P}$ at time t:

$$R(t) = \sum_{\mathcal{C}\in\mathcal{P}} P(\mathcal{C}, t) - P(\mathcal{C}, \infty)$$

- At the stationary state $P(\mathcal{C}, \infty)$ is the probability of finding two walkers in the same community $\mathcal{C}$
- If we consider $t = \Delta\tau = 1$ (discrete-time random walk), then we have:

$$R(1) = \sum_{i,j} \frac{a_{ij}}{k_i}\frac{k_i}{2K}\delta_{c_i,c_j} - p_i^* p_j^* \delta_{c_i,c_j} = \frac{1}{2K}\sum_{i,j}\left(a_{ij} - \frac{k_i k_j}{2K}\right)\delta_{c_i,c_j}$$

- So the modularity of a partition is just the one-step stability of a plain random walk on the graph with that partition

The Engineer

Node centrality

Problem: how we measure the relative importance of nodes?

PageRank

- The "Smart" transition matrix (RW with teleportation):

$$\Phi(q) = \frac{q}{N}\mathbf{1}\,\mathbf{1} + (1-q)\Pi$$

- $\Phi(q)$ satisfies all the hypotheses of Perron-Frobenius
- $\Longrightarrow$ the eigenvector of $\Phi(q)$ associated to the first eigenvalue is the centrality induced by $\Phi(q)$

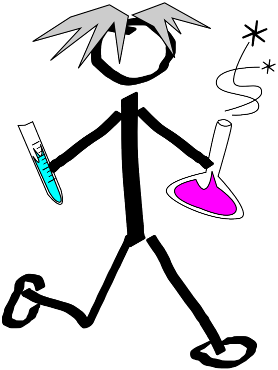
PageRank

- The "Smart" transition matrix (RW with teleportation):

$$\Phi(q) = \frac{q}{N}\mathbf{1}\,\mathbf{1} + (1-q)\Pi$$

- $\Phi(q)$ satisfies all the hypotheses of Perron-Frobenius
- $\Longrightarrow$ the eigenvector of $\Phi(q)$ associated to the first eigenvalue is the centrality induced by $\Phi(q)$
- This eigenvector is used to compute the PageRank score of a node

The Alchemist



- A plain random walk W defined by Π

- A plain random walk W defined by Π
- $(i_0, i_1, i_2, \ldots)$: nodes visited by W

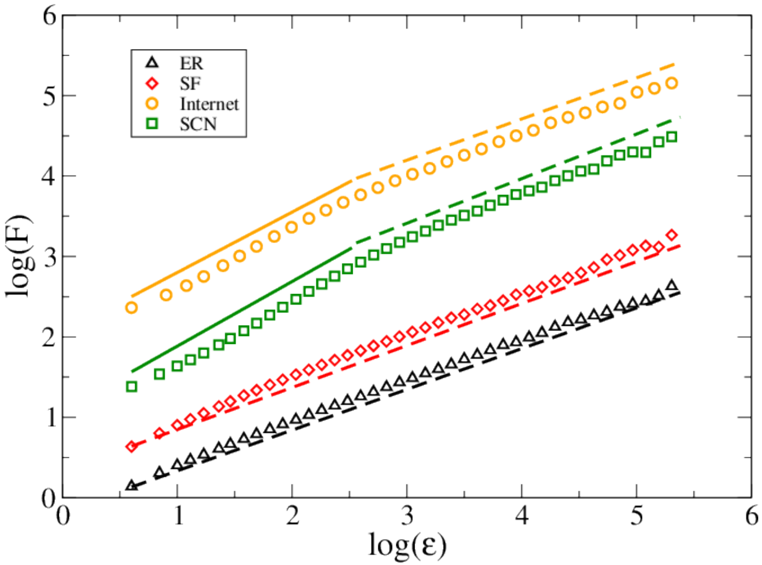
Queen Mary University of London

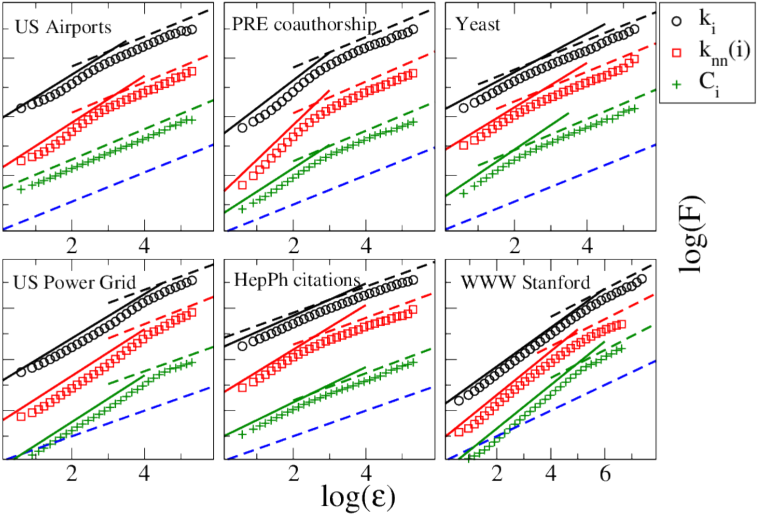
- A plain random walk W defined by Π
- $(i_0, i_1, i_2, \ldots)$: nodes visited by W
- $(H_{i_0}, H_{i_1}, H_{i_2}, \ldots)$, where H_i is k_i, or k_i^{nn}, or C_i

- A plain random walk W defined by Π
- $(i_0, i_1, i_2, \ldots)$: nodes visited by W
- $(H_{i_0}, H_{i_1}, H_{i_2}, \ldots)$, where H_i is k_i, or k_i^{nn}, or C_i
- Look at the typical fluctuations of H_i in an interval of length ε and call them $F(\varepsilon)$

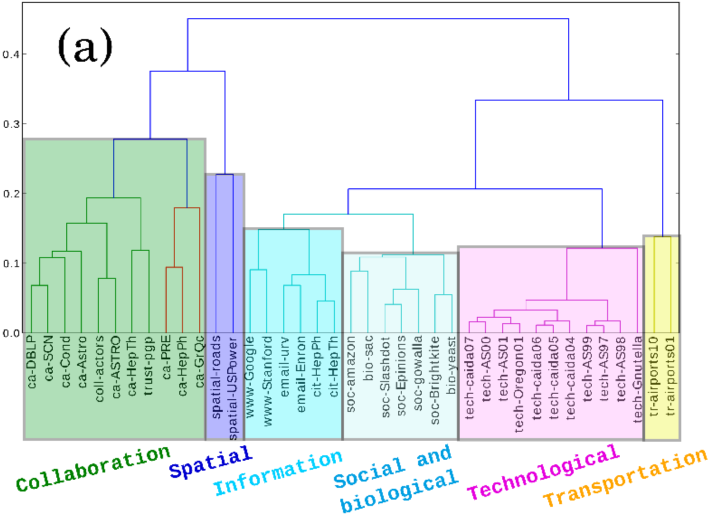
- A plain random walk W defined by Π
- $(i_0, i_1, i_2, \ldots)$: nodes visited by W
- $(H_{i_0}, H_{i_1}, H_{i_2}, \ldots)$, where H_i is k_i, or k_i^{nn}, or C_i
- Look at the typical fluctuations of H_i in an interval of length ε and call them $F(\varepsilon)$
- Look at how $F(\varepsilon)$ scales with the size ε (Detrended fluctuation analysis)

Queen Mary University of London




US Airports
PRE coauthorship
Yeast
k_i
$k_{nn}(i)$
C_i
US Power Grid
HepPh citations
WWW Stanford
2
4
6
log(F)
log(ε)

Taxonomy


(a)
0.4
0.3
0.2
0.1
0.0
ca-DBLP
ca-SCN
ca-Cond
ca-Astro
coll-actors
ca-ASTRO
ca-HepTh
trust-pgp
ca-PRE
ca-HepPh
ca-GrQc
spatial-roads
spatial-USPower
www-Google
www-Stanford
email-urv
email-Enron
cit-HepPh
cit-HepTh
soc-amazon
bio-sac
soc-Slashdot
soc-Epinions
soc-gowalla
soc-Brightkite
bio-yeast
tech-caida07
tech-AS00
tech-AS01
tech-Oregon01
tech-caida06
tech-caida05
tech-caida04
tech-AS99
tech-AS97
tech-AS98
tech-Gnutella
tr-airports10
tr-airports01
Collaboration
Spatial
Information
Social and biological
Technological
Transportation

Take-home message

Random walks at 9:00 am might be boring...

Take-home message

Random walks at 9:00 am might be boring...
...but are a nice process to understand the structure and dynamics of complex networks ;)

THANK YOU!

References

- J. Noh, H. Rieger, "Random Walks on Complex Networks", Phys. Rev. Lett. 92, 118701 (2004).
- J. Gómez-Gardeñes, V. Latora, "Entropy rate of diffusion processes on complex networks", Phys. Rev. E 78, 065102(R) (2008).
- R. Sinatra, J. Gomez-Gardenes, R. Lambiotte, V. Nicosia, V. Latora "Maximal-entropy random walks in complex networks with limited information", Phys. Rev. E , 83, 030103R (2011).
- J. Gomez-Gardenes, V. Nicosia, R. Sinatra, V. Latora "Motion-induced synchronization in metapopulations of mobile agents", Phys. Rev. E 87, 032814 (2013).
- M. Bonaventura, V. Nicosia, V. Latora "Characteristic times of biased random walks on complex networks", Phys. Rev. E 89, 012803 (2014).
- M. Rosvall, C. T. Bergstrom, "Maps of random walks on complex networks reveal community structure", Proceedings of the National Academy of Sciences, 105, 1118-1123 (2008).
- R. Lambiotte, J. C. Delvenne, M. Barahona, "Laplacian Dynamics and Multiscale Modular Structure in Networks", arxiv:0812.1770 (2008).
- S. Brin, L. Page "The Anatomy of a Large-Scale Hypertextual Web Search Engine", `http://infolab.stanford.edu/~backrub/google.html`
- V. Nicosia, M. De Domenico, V. Latora "Characteristic exponents of complex networks", EPL - Europhysics Letters 106, 58005 (2014).