Random Walks in Graphs

Thomas Bonald
thomas.bonald@telecom-paristech.fr

May 2017

Consider a weighted, undirected, complete graph G of N nodes. For all $i, j = 1, \ldots, N$, we denote by w_{ij} the weight of edge i, j, by $w_i = \sum_j w_{ij}$ the weight of node i, and by $w = \sum_i w_i$ the total weight of nodes (or edges). We assume that the graph induced by the positive weights is connected (in particular, the node weights are positive) and not bipartite (that is, contains cycles of odd length). Some key metrics useful for ranking and clustering nodes can be expressed in terms of random walks in this graph, that can be seen as electrons moving at random in the corresponding electrical network where each edge i, j has conductance w_{ij}. These notes are mainly based on [1, 2, 3].

1 Markov chain

Consider a random walk in the graph where the probability of moving from node i to node j is $p_{ij} = w_{ij}/w_i$. The successive nodes $X_0, X_1, X_2, \ldots$ visited by the random walk form a Markov chain on $\{1, \ldots, N\}$. We have for all $n \geq 0$:

$$\forall i = 1, \ldots, N, \quad \mathrm{P}(X_{n+1} = i) = \sum_{j=1}^{N} \mathrm{P}(X_n = j) p_{ji}. \tag{1}$$

Since the graph is connected, the Markov chain is irreducible.

A stationary distribution π of the Markov chain satisfies:

$$\forall i = 1, \ldots, N, \quad \pi_i = \sum_{j=1}^{N} \pi_j p_{ji}. \tag{2}$$

It can be easily verified that the distribution defined by $\pi_i = w_i/w$ satisfies these equations. By the Perron-Frobenius theorem, this is the unique solution to these equations. We shall see that it is also the limiting distribution of X_n when $n \to +\infty$, independently of the distribution of X_0.

By the ergodic theorem, we have:

$$\frac{1}{m} \sum_{n=1}^{m} 1_{\{X_n = i\}} \xrightarrow{\text{p.s.}} \pi_i \quad \text{when } m \to +\infty,$$

so that the frequency of visit of each node is proportional to its weight. Similarly,

$$\frac{1}{m} \sum_{n=1}^{m} 1_{\{X_n = i, X_{n+1} = j\}} \xrightarrow{\text{p.s.}} \pi_i p_{ij} \quad \text{when } m \to +\infty,$$

so that the frequency of visit of each edge is also proportional to its weight. Observe that the balance equations (2) can be written

$$\forall i = 1, \ldots, N, \quad \sum_{j=1}^{N} \pi_i p_{ij} = \sum_{j=1}^{N} \pi_j p_{ji}$$

and thus interpreted as the frequency of departures from each node being equal to the frequency of arrivals to this node.

The Markov chain is reversible in the sense that

$$\forall i, j = 1, \dots, N, \quad \pi_i p_{ij} = \pi_j p_{ji}. \tag{3}$$

Thus the frequency of moves from i to j is equal to the frequency of moves from j to i. These equations, called the local balance equations, are stronger than the balance equations (2).

Remark 1 *Any irreducible, reversible Markov chain with N states corresponds to a random walk on a graph of N nodes, with weight $w_{ij} = w\pi_i p_{ij}$ between node i and j, for some positive constant w. The weight of node i is then $w_i = w\pi_i$ and the probability of moving from node i to node j is $w_{ij}/w_i = p_{ij}$.*

If X_0 has the distribution π, then this is the distribution of $X_1, X_2, \dots$ and we have for all $n \geq 0$:

$$\mathrm{P}(X_n = j | X_{n+1} = i) = \frac{\mathrm{P}(X_n = j)}{\mathrm{P}(X_{n+1} = i)} \mathrm{P}(X_{n+1} = i | X_n = j) = \frac{\pi_j}{\pi_i} p_{ji} = p_{ij},$$

so that the random walk in reverse time has the same distribution as the original random walk.

2 Spectral analysis

Let $P = D^{-1}A$ be the transition matrix of the Markov chain $X_0, X_1, X_2, \dots$, where $A = (w_{ij})_{1 \leq i,j \leq N}$ is the adjacency matrix of the graph and $D = \mathrm{diag}(w_1, \dots, w_N)$ the diagonal matrix of node weights. This is a stochastic matrix in the sense that $P \geq 0$ and $P1 = 1$. Writing the distribution of X_n as a vector $\pi(n)$ of dimension N, the equations of evolution (1) and the balance equations (2) can be respectively written in vectorial form,

$$\pi(n+1)^T = \pi(n)^T P \quad \text{and} \quad \pi^T = \pi^T P.$$

In particular,

$$\forall n \geq 0, \quad \pi(n)^T = \pi(0)^T P^n, \tag{4}$$

and π^T is the left eigenvector of P for the eigenvalue 1 such that $\pi^T 1 = 1$.

Unlike the transition matrix $P = D^{-1}A$, the Laplacian matrix $L = D^{-1/2} A D^{-1/2}$ is symmetric, with $D^{-1/2} = \mathrm{diag}(1/\sqrt{w_1}, \dots, 1/\sqrt{w_N})$. There is some matrix Q such that $Q^T Q = I$ and $Q^T L Q = \Lambda$, where $\Lambda = \mathrm{diag}(\lambda_1, \dots, \lambda_N)$ and $|\lambda_1| \geq |\lambda_2| \geq \dots \geq |\lambda_N|$. The columns of Q are the eigenvectors of L for the respective eigenvalues $\lambda_1, \dots, \lambda_N$. Let $U = D^{1/2} Q / \sqrt{w}$ and $V = D^{-1/2} Q \sqrt{w}$. Observe that $U^T V = I$.

Theorem 1 (Spectral decomposition) *We have:*

$$P = V \Lambda U^T. \tag{5}$$

Proof. This follows from:

$$P = D^{-1/2} L D^{1/2} = D^{-1/2} Q \Lambda Q^T D^{1/2} = V \Lambda U^T.$$

□

In view of (5), we have

$$U^T P = \Lambda U^T \quad \text{and} \quad PV = V\Lambda$$

so that the lines of U^T and the columns of V are the left and right eigenvectors of P for the respective eigenvalues $\lambda_1, \dots, \lambda_N$. By Perron-Frobenius Theorem, we have $\lambda_1 = 1 > |\lambda_2| \geq \dots \geq |\lambda_N|$. Let $q_1, \dots, q_N$ be the columns of the matrix Q, $u_1, \dots, u_N$ and $v_1, \dots, v_N$ be the columns of the matrices U and V, respectively. Observing that

$$L D^{1/2} 1 = D^{-1/2} A 1 = D^{-1/2} D 1 = D^{1/2} 1,$$

we get $q_1 = \pm\sqrt{\pi}$. Choosing $q_1 = \sqrt{\pi}$, we obtain $u_1 = \pi$ and $v_1 = 1$. Now for all $n \geq 1$,

$$P^n = V\Lambda^n U^T = v_1 u_1^T + \sum_{k=2}^{N} \lambda_k^n v_k u_k^T. \tag{6}$$

Denoting by Π the matrix $1\pi^T$ (all lines equal to π^T), we get

$$\lim_{n\to+\infty} P^n = \Pi.$$

In view of (4), for any initial distribution $\pi(0)$,

$$\lim_{n\to+\infty} \pi(n)^T = \pi(0)^T \Pi = \pi^T.$$

Moreover, in view of (6), the convergence rate depends mainly on the second largest eigenvalue (in modulus) of P, known as the *spectral gap* of the matrix.

3 Hitting times

Let $\mathrm{P}_i = \mathrm{P}(\cdot|X_0 = i)$ be the probability measure conditioned on the fact that the random walk starts from node i, E_i the corresponding expectation and $T_i = \min\{n \geq 1 : X_n = i\}$ the hitting time of node i. We are interested in the mean hitting time of node j from node i,

$$h_{ij} = \mathrm{E}_i(T_j).$$

When $i = j$, this is the mean return time to node i; since node i is visited at frequency i, we must have

$$h_{ii} = \frac{1}{\pi_i}. \tag{7}$$

This will be proved in section 5.

Proposition 1 *We have*

$$\forall i, j = 1, \ldots, N, \quad h_{ij} = 1 + \sum_{k \neq j} p_{ik} h_{kj}. \tag{8}$$

Proof. The proof follows by first-step analysis:

$$\begin{aligned}
h_{ij} &= \mathrm{E}(T_j|X_0 = i), \\
&= \sum_{k=1}^{N} \mathrm{P}(X_1 = k|X_0 = i)\mathrm{E}(T_j|X_0 = i, X_1 = k), \\
&= \sum_{k=1}^{N} p_{ik}\mathrm{E}(T_j|X_1 = k), \\
&= 1 + \sum_{k \neq j} p_{ik}\mathrm{E}(T_j|X_0 = k), \\
&= 1 + \sum_{k \neq j} p_{ik} h_{kj}.
\end{aligned}$$

□

Now consider the mean hitting time of node i in steady state:

$$h_i = \sum_{j=1}^{N} \pi_j h_{ji}.$$

This defines a measure of centrality of the nodes: node i is more central than node j if $h_i < h_j$. We shall see in section 5 that for all i, j,

$$h_i - h_j = h_{ji} - h_{ij}.$$

Observe that $h_{ij} \neq h_{ji}$ in general.

We define the *mean commute time* between nodes i and j by:

$$\sigma_{ij} = h_{ij} + h_{ji}.$$

The mean commute time is related to the mean escape probability from node i to node j,

$$e_{ij} = \mathrm{P}_i(T_j < T_i).$$

Observe that $e_{ij} > 0$ for any distinct nodes i, j, so that $\sigma_{ij} < \infty$ and $h_{ij} < \infty$ for all nodes i, j.

Proposition 2 *For any distinct nodes i, j,*

$$\sigma_{ij} = \frac{1}{\pi_i e_{ij}}. \tag{9}$$

Proof. Let $S_i = \min\{n \geq 1 : X_n = i, T_j < n\}$. This is the hitting time of i after having visited j. We have:

$$\begin{aligned}
\sigma_{ij} &= \mathrm{E}_i(S_i), \\
&= \mathrm{E}_i(T_i) + \mathrm{E}_i(S_i - T_i), \\
&= \mathrm{E}_i(T_i) + \mathrm{E}_i((S_i - T_i) 1_{S_i > T_i}), \\
&= \mathrm{E}_i(T_i) + \mathrm{P}_i(S_i > T_i)\mathrm{E}_i(S_i - T_i | S_i > T_i), \\
&= \mathrm{E}_i(T_i) + \mathrm{P}_i(T_j > T_i)\mathrm{E}_i(S_i).
\end{aligned}$$

The result then follows from (7) and the fact that $e_{ij} = \mathrm{P}_i(T_j < T_i) = 1 - \mathrm{P}_i(T_j > T_i)$. $\square$

Remark 2 *It follows from (9) that $\pi_i e_{ij} = \pi_j e_{ji}$ for all nodes i, j: the frequency of paths starting from i and hitting j before i is equal to the frequency of paths starting from j and hitting i before j.*

4 Electrical network

We shall see that the mean commute time σ_{ij} between i and j can be interpreted as the *effective resistance* between i and j in the electric network induced by the graph, where each edge is a resistor with conductance equal to the weight of this edge. Consider this network with node i set at electric potential 1 (in V, say) and node j set at electric potential 0. Let V_k be the electric potential of any node k. We have $V_i = 1$ and $V_j = 0$.

By Ohm's law, the current that flows from k to ℓ is

$$I_{k\ell} = (V_k - V_\ell) w_{k\ell}.$$

By Kirchoff's law, we have for any $k \neq i, j$,

$$\sum_{\ell} I_{k\ell} = 0,$$

that is

$$V_k w_k = \sum_\ell V_\ell w_{k\ell}.$$

Finally,

$$\forall k \neq i, j, \quad V_k = \sum_\ell V_\ell P_{k\ell}. \tag{10}$$

Thus the potential in each node $k \neq i, j$ is the weighted average of the potential of its neighbors; this is a *harmonic function* of the graph. We shall see that there is a unique solution to these equations, given the boundary condition $V_i = 1$ and $V_j = 0$. The following result shows that V_k can be interpreted as the probability that the random walk starting from k reaches i before j.

Proposition 3 *For any $k \neq i, j$,*

$$V_k = \mathrm{P}_k(T_i < T_j). \tag{11}$$

Proof. Let $V'_k = \mathrm{P}_k(T_i < T_j)$, for each $k \neq i, j$, $V'_i = 1$ and $V'_j = 0$. Then

$$V'_k = P_{ki} + \sum_{\ell \neq i,j} P_{k\ell} V'_\ell = \sum_\ell V'_\ell P_{k\ell}.$$

Thus V' is a solution to the equations (10). Since the solution is unique, we have $V_k = V'_k$ for all k. □

The *effective conductance* between node i and node j is the current that flows out of node i (equivalently, in node j), that is

$$C^{\mathrm{eff}}_{ij} = \sum_k I_{ik} = \sum_k (1 - V_k) w_{ik} = w_i - \sum_k V_k w_{ik} = w_i (1 - \sum_k V_k P_{ik}).$$

Now in view of (11), we have

$$\sum_k V_k P_{ik} = \sum_k P_{ik} \mathrm{P}_k(T_i < T_j) = \mathrm{P}_i(T_i < T_j),$$

so that, in view of (9),

$$C^{\mathrm{eff}}_{ij} = w_i e_{ij} = \frac{w}{\sigma_{ij}}.$$

This shows that σ_{ij}/w is the effective resistance between i and j.

More generally, consider some set of nodes $B \subset \{1, \ldots, n\}$ with $|B| \neq 1, n$ and some potential function $V : B \to \mathbb{R}$. By the same argument as above, the potential function in any other node satisfies:

$$\forall k \notin B, \quad V_k = \sum_\ell V_\ell P_{k\ell}. \tag{12}$$

Proposition 4 *There is a unique solution to the equations* (12).

Proof. We first prove that $\max_k V_k = \max_{k \in B} V_k$ and $\min_k V_k = \min_{k \in B} V_k$. Let k be any node such that V_k is maximum. If $k \notin B$, it follows from (12) that V_ℓ is maximum for all neighbors ℓ of k. If no neighbor of k belongs to B, we apply again this argument until we reach a node in B; such a node exists because the graph is connected. The proof is similar for the minimum.

Now consider two solutions V' and V'' of (12). Then $V = V' - V''$ is a solution of (12) such that $V_k = 0$ for all $k \in B$. We deduce that $V_k = 0$ for all k (because $\max_k V_k = \min_k V_k = 0$) so that $V' = V''$. □

Proposition 3 can then be extended as follows:

Proposition 5 *Let $T = \min\{n \geq 1 : X_n \in B\}$ be the hitting time of the set B. For any $k \notin B$,*

$$V_k = \sum_{i \in B} \mathrm{P}_k(T_i = T) V_i. \tag{13}$$

Proof. Define for all $k \notin B$,

$$V'_k = \sum_{i \in B} \mathrm{P}_k(T_i = T) V_i,$$

and $V'_k = V_k$ for all $k \in B$. Then

$$V'_k = \sum_{i \in B} \left(P_{ki} V_i + \sum_{\ell \notin B} P_{k\ell} \mathrm{P}_\ell(T_i = T) V_i \right) = \sum_{i \in B} P_{ki} V_i + \sum_{\ell \notin B} P_{k\ell} V'_\ell = \sum_\ell V'_\ell P_{k\ell}.$$

Thus V' is a solution to the equations (10). Since the solution is unique, we have $V_k = V'_k$ for all k. □

5 Spectral embedding

Define the *fundamental matrix* of the Markov chain as

$$Z = I + \sum_{n \geq 1} (P^n - \Pi).$$

The entry i, j of the matrix $Z - I$ is the limit when $n \to +\infty$ of the difference between the mean number of visits to node j starting from node i in n steps and the mean number of visits to node j starting from the steady state in n steps. This limit exists and is finite in view of the following result.

Theorem 2 *The matrix Z exists and is the inverse of $I - P + \Pi$.*

Proof. The proof relies on the fact that $\Pi P = P \Pi = \Pi$, so that $\Pi P^n = P^n \Pi = \Pi$ for all $n \geq 1$. We first prove that the matrix $I - P + \Pi$ is invertible. Let u be some vector such that $(I - P + \Pi)u = 0$. Then for all $n \geq 1$,

$$P^n u - P^{n+1} u + \Pi u = 0.$$

Taking the limit, we get $\Pi u = 0$, so that $u = Pu$. By Perron-Frobenius Theorem, this implies that $u = \alpha 1$ for some $\alpha \in \mathbb{R}$. Finally, $\alpha = 0$ because $\Pi u = \alpha 1 \pi^T 1 = 0$.

To conclude the proof, we observe that for all $n \geq 1$,

$$(P - \Pi)^n = \sum_{k=0}^n \binom{n}{k} P^k (-\Pi)^{n-k} = P^n + \Pi \sum_{k=0}^n (-1)^k = P^n - \Pi.$$

With $M = P - \Pi$, we have

$$(I - M)(I + M + \ldots + M^{n-1}) = I - M^n = I - P^n + \Pi.$$

Since $I - M$ is invertible and $P^n \to \Pi$ when $n \to +\infty$, the matrix

$$\sum_{n \geq 0} M^n = I + \sum_{n \geq 1} (P^n - \Pi)$$

is well defined and is the inverse of $I - M$. □

Proposition 6 *We have:*

$$Z = v_1 u_1^T + \sum_{k=2}^{n} \frac{1}{1-\lambda_k} v_k u_k^T. \tag{14}$$

Proof. For all $n \geq 1$,

$$P^n = V\Lambda^n U^T = v_1 u_1^T + \sum_{k=2}^{N} \lambda_k^n v_k u_k^T$$

and since $U^T V = V U^T = I$,

$$I = v_1 u_1^T + \sum_{k=2}^{N} v_k u_k^T$$

The proof follows from the fact that $v_1 u_1^T = 1\pi^T = \Pi$ $\square$

Let $H = (h_{ij})_{1 \leq i,j \leq N}$ be the matrix of mean hitting times. In view of (8), we have

$$H = 11^T + P(H - d(H)), \tag{15}$$

where for any square matrix M, $d(M)$ is the diagonal matrix which has the same diagonal as M.

Theorem 3 *There is a unique solution to* (15)*, given by*

$$H = (I - Z + 11^T d(Z))d(\Pi)^{-1}. \tag{16}$$

Proof. We first prove that the matrix H as defined by (16) satisfies (15). Observing that $d(H) = d(\Pi)^{-1}$, we get

$$H - d(H) = (-Z + 11^T d(Z))d(\Pi)^{-1}$$

and

$$H - P(H - d(H)) = (I - Z + PZ)d(\Pi)^{-1}.$$

Using the fact that $(I - P + \Pi)Z = I$ and $\Pi Z = \Pi$ (because $\pi^T Z = \pi^T$), we obtain

$$H - P(H - d(H)) = \Pi d(\Pi)^{-1} = 1\pi^T d(\Pi)^{-1} = 11^T.$$

Now take any solution H' of (15). Since

$$\pi^T H' = \pi^T (11^T + P(H' - d(H')) = 1^T + \pi^T H' - \pi^T d(H'),$$

we have $\pi^T d(H') = 1^T$ so that $d(H') = d(\Pi)^{-1} = d(H)$. Now let $\Delta = H - H'$. We have $d(\Delta) = 0$ and, in view of (15), $\Delta = P\Delta$. This implies that each column of Δ is proportional to the vector 1. Since $d(\Delta) = 0$, we get $\Delta = 0$ and $H' = H$. $\square$

Let $h = (h_i)_{1 \leq i \leq N}$ be the vector of mean hitting times in steady state, $h^T = \pi^T H$. In view of (16), we have:

$$h^T = 1^T d(Z) d(\Pi)^{-1} \tag{17}$$

and

$$H = d(\Pi)^{-1} - Zd(\Pi)^{-1} + 1h^T. \tag{18}$$

Using (14) and the equality $wD^{-1}U = V$, we get

$$Zd(\Pi)^{-1} = v_1 v_1^T + \sum_{k=2}^{n} \frac{1}{1-\lambda_k} v_k v_k^T.$$

Now let $(\vec{v_1}, \ldots, \vec{v_N})$ be the rows of matrix V. This is the embedding of the nodes in the vectorial space $\mathbb{R}^N$ equipped with the basis $(u_1, \ldots, u_N)$, the left eigenvectors of P. Define the scalar product:

$$< \vec{a}, \vec{b} >= a_1 b_1 + \sum_{k=2}^{N} \frac{1}{1-\lambda_k} a_k b_k.$$

It follows from (17) and (18) that

$$\forall i, \quad h_i =< \vec{v}_i, \vec{v}_i >= ||\vec{v}_i||^2$$

and

$$\forall i \neq j, \quad h_{ij} = - < \vec{v}_i, \vec{v}_j > + h_j, \quad \forall i, \quad h_{ii} = \frac{1}{\pi_i}.$$

Moreover,

$$\sigma_{ij} = ||\vec{v}_i||^2 - 2 < \vec{v}_i, \vec{v}_j > + ||\vec{v}_j||^2 = ||\vec{v}_i - \vec{v}_j||^2.$$

Thus $\sqrt{\sigma_{ij}}$ is the distance between i and j in this vectorial space, while $\sqrt{h_i}$ is the distance between i and the origin (which may be considered as the embedding of a random node, drawn from the distribution π).

References

[1] P. Brémaud. *Markov chains: Gibbs fields, Monte Carlo simulation, and queues*, volume 31. Springer Science & Business Media, 2013.

[2] L. Lovász. Random walks on graphs. *Combinatorics, Paul Erdos is eighty*, 2:1–46, 1993.

[3] P. Snell and P. Doyle. Random walks and electric networks. *Free Software Foundation*, 2000.