

Core-biased random walks in networks

Raúl J. Mondragón

School of Electronic Engineering and Computer Science
Queen Mary University of London
Mile End Road E1 4NS UK
r.j.mondragon@qmul.ac.uk

January 25, 2018

Abstract

A simple strategy to explore a network is to use a random-walk where the walker jumps from one node to an adjacent node at random. It is known that biasing the random jump, the walker can explore every walk of the same length with equal probability, this is known as a Maximal Entropy Random Walk (MERW). To construct a MERW requires the knowledge of the largest eigenvalue λ_1 and corresponding eigenvector $\bar{v}^{(1)}$ of the adjacency matrix $\mathbf{A} = \{a_{ij}\}$, that is global knowledge of the network. When this global information is not available, it is possible to construct a biased random walk which approximates the MERW using only the degree of the nodes, a local property. Here we show that it is also possible to construct a good approximation to a MERW by biasing the random walk via the properties of the network's core. We present some examples based on real and artificial networks showing that the core-biased random walk outperforms the degree-biased random walks. networks, entropy rate random walks, biased random walk, spectral bound

1 Introduction

A simple method to explore a network is via a random walk. Properties of the random walk, like the probability that the walker is in a particular node as time tends to infinity or, the time that it would take the walker to visit at least once all the nodes of the network, are dependent on the network's connectivity. This relationship between the structural properties of the network and the random walk have been used to explore properties of complex networks like community detection [16] or the discovery of the network structure [21]. Here we are interested in biased random walks, and in particular a random walk where the walkers explore every walk with equal probability, known as Maximal Entropy-rate Random Walks (MERW) [4]. The MERW has been used to study well known problems in complex networks like link prediction [10], community

detection [14] and in practical applications, for example in the discovery of salient objects in an image [22] or in the development of routing algorithms in data-centric networks [9]. To construct a MERW requires the knowledge of the largest eigenvalue and corresponding eigenvector of the adjacency matrix [4], these are global properties of the network which require that, at each step, the random walker knows the whole connectivity of the network, often this knowledge is not available.

A good approximation to the MERW using only local properties is the degree biased random walk [8, 7, 17]. In this case the walker jumps from one node to a neighbouring node with a preference based on the degree of the destination node. These degree-biased random walks are also used to understand properties of networks, for example, how to efficiently disperse the information contained in a node [8], in the link prediction problem [11], in the accurate discovery of the network structure in assortative networks [20], in epidemic spreading [15] and more recently the degree biased random walk has been extended to multiplex networks [1].

In here, as an approximation to the MERW, we present a biased random walk where the walker prefers to jump to nodes that are connected to the core of the network. We would justify the proposed biased random walk via a spectral bound of the adjacency matrix and show that this core-biased random walk can outperform the degree-biased random walk.

Next section reviews some results about MERW and its approximations based on local walks of a given length. We use these approximations for comparison purposes with the method introduced here. Section 3 introduces the core biased random walk and how is related to the MERW. Section 4 we present some examples comparing the core biased method with the MERW using some real and synthetic networks. Last section is our conclusions.

2 Biased random walks and the Maximal Rate Entropy Random walk

The connectivity of a finite, undirected and connected network can be described by the symmetric adjacency matrix \mathbf{A} , where $a_{ij} = 1$ if nodes i and j share a link and zero otherwise. In these networks, a random walker would jump from node i to a neighbouring node j with a probability $P_{i \rightarrow j}$. The probability that the walker is in node j at time $t + 1$ is $p_j(t + 1) = \sum_i a_{ij} P_{i \rightarrow j} p_i(t)$ or in matrix notation $\bar{p}(t + 1) = \boldsymbol{\pi} \bar{p}(t)$. If the matrix $\boldsymbol{\pi}$ is primitive then the probability of finding the walker in node i as time tends to infinity is given by the stationary distribution $\bar{p}^* = \{p_i^*\}$. In a network, the jump probability $P_{i \rightarrow j}$ can be expressed as

$$P_{i \rightarrow j} = \frac{a_{ij} f_j}{\sum_j a_{ij} f_j} \quad (1)$$

where f_j is a function of one or several topological properties of the network, in this case the stationary distribution is [8]

$$p_i^* = \frac{f_i \sum_j a_{ij} f_j}{\sum_n f_n \sum_j a_{nj} f_j}. \quad (2)$$

The measure which tell us the minimum amount of information needed to describe the stochastic walk is the entropy rate $h = \lim_{t \rightarrow \infty} S_t/t$, where S_t is the Shannon entropy of all walks of length t . There is another definition of entropy rate [5]. If $\{x_i\}$ denotes the stochastic process of the random walker, then if the stationary probability of this process exist, the rate entropy $h' = \lim_{t \rightarrow \infty} S(x_t|x_{t-1}, \dots, x_1)$, is equal to $h = \lim_{t \rightarrow \infty} S_t/t$ [5]. This alternative definition is used when describing random walks. If the random walk is described by a stationary Markov chain then $\lim_{t \rightarrow \infty} S(x_t|x_{t-1}, \dots, x_1) = S(x_2|x_1)$ and [5]

$$h = - \sum_{i=1}^N p_i^* \sum_{j=1}^N P_{i \rightarrow j} \ln(P_{i \rightarrow j}). \quad (3)$$

The maximal rate entropy h_{\max} corresponds to random walks where all the walks of the same length have equal probability. The value of h_{\max} can be expressed in terms of the spectral properties of the network as follows. If $|\lambda_1| \geq |\lambda_2| \geq \dots \geq |\lambda_N|$ are the eigenvalues of the adjacency matrix \mathbf{A} , then the number of walks of length t is $W_t = \sum_{ij} (\mathbf{A}^t)_{ij}$ where $(\mathbf{A}^t)_{ij}$ denotes the i, j entry of the adjacency matrix \mathbf{A} raised to the power t . The number of walks can be expressed in terms of the spectra as $W_t = \sum_i a_i^2 \lambda_i^t = a_1^2 \lambda_1^t (1 + (a_2/a_1)^2 (\lambda_2/\lambda_1)^t + \dots)$, where the a_i s are constants. From this last expression the maximal rate entropy is [4]

$$h_{\max} = \lim_{t \rightarrow \infty} \frac{\ln \sum_{ij} (\mathbf{A}^t)_{ij}}{t} = \lim_{t \rightarrow \infty} \left(\frac{1}{t} \right) \ln (a_1^2 \lambda_1^t (1 + (a_2/a_1)^2 (\lambda_2/\lambda_1)^t + \dots)) = \ln(\lambda_1). \quad (4)$$

For the MERW the probability $P_{i \rightarrow j}$ is such that all the walks of the same length have equal probability. This probability is determined by considering the proportion of all walks of length t that start in node i and pass through node j relative to all the walks of length t that start in node i , and then take the limit as the walk length tends to infinity [4], that is

$$P_{i \rightarrow j} = \lim_{t \rightarrow \infty} \left(\frac{a_{ij} \sum_{k=1}^N (\mathbf{A}^{t-1})_{jk}}{\sum_{k=1}^N a_{ik} \sum_{m=1}^N (\mathbf{A}^{t-1})_{km}} \right) = \lim_{t \rightarrow \infty} \left(\frac{a_{ij} (\mathbf{A}^{t-1} \cdot \mathbf{1})_j}{\sum_{k=1}^N a_{ik} (\mathbf{A}^{t-1} \cdot \mathbf{1})_k} \right) \quad (5)$$

where the inner sums are expressed as the multiplication of the matrix \mathbf{A}^{t-1} with the vector $\mathbf{1}$ which has all its entries equal to 1. The vector $\mathbf{1}$ can be expressed as a linear combination of the eigenvectors of \mathbf{A} as $\mathbf{1} = \sum_{j=1}^N b_j \bar{v}^{(j)}$ where $\bar{v}^{(j)}$ is the j -th eigenvector and the b_j s are constants. The expression $\mathbf{A}^{t-1} \cdot \mathbf{1} = \sum_{j=1}^N b_j \lambda_j^{t-1} \bar{v}^{(j)}$ in the limit as t tends to infinity is $\lim_{t \rightarrow \infty} (\mathbf{A}^{t-1} \cdot \mathbf{1}) = \sum_{j=1}^N b_j \lambda_j^{t-1} \bar{v}^{(j)}$

$\mathbf{1}) = b_1 \lambda_1^{t-1} \bar{v}^{(1)}$ as $|\lambda_1| \geq |\lambda_i|$ for all i , then from this observation, Eq. (5) is [4]

$$P_{i \rightarrow j} = \frac{a_{ij} v_j}{\sum_j a_{ij} v_j} = \frac{a_{ij} v_j}{\lambda_1 v_i}, \quad (6)$$

where v_i is the i -th entry of the eigenvector $\bar{v}^{(1)}$. For clarity, in the above equation we have drop the superscript labelling the eigenvectors, and from now on, the entries v_i would always refer to the eigenvector $\bar{v}^{(1)}$. The stationary probability for the MERW is $p_i^* = v_i^2$, which can be verified by evaluating that the condition of stationary $p_j^* = \sum_i a_{ij} (v_j / \lambda_1 v_i) p_i^*$ is satisfied. From this stationary probability, Eq. (6) and Eq. (3) it can be confirmed that the MERW maximises the rate entropy. Also it is not difficult to verify that the probability of the walk of length t starting from i and finishing in j is $p(W_{ij}^{(t)}) = (v_i / (\lambda_1 v_a)) (v_a / \lambda_1 v_b) \dots (v_y / (\lambda_1 v_z)) (v_z / (\lambda_1 v_j)) = v_i / (\lambda_1^t v_j)$, showing that the probability of the walk $W_{ij}^{(t)}$ only depends on the initial and final node and the length of the walk.

The implementation of the MERW requires global knowledge of the network connectivity as we need to evaluate the largest eigenvalue-eigenvector pair. If this global knowledge is unavailable then it is possible to approximate the maximal entropy random walk via the node's degree k_i [8, 7] which is a local network property. In this case walks of length $t = 2$ have equal probability, as $(\mathbf{A}^{t-1} \cdot \mathbf{1})_i = (\mathbf{A} \cdot \mathbf{1})_i = k_i$ then the biased jump is

$$P_{i \rightarrow j} = \frac{a_{ij} (\mathbf{A} \cdot \mathbf{1})_j}{\sum_{k=1}^N a_{ik} (\mathbf{A} \cdot \mathbf{1})_k} = \frac{a_{ij} k_j}{\sum_j a_{ij} k_j}. \quad (7)$$

This degree biased random walk has been extended to

$$P_{i \rightarrow j}(\alpha) = \frac{a_{ij} k_j^\alpha}{\sum_j a_{ij} k_j^\alpha} \quad (8)$$

where α is a free parameter [3, 17, 1]. The parameter α can be adjusted to find the best approximation to the maximal entropy rate. We would use both, Eqs. (7) and (8) for comparison purposes Section 4.

It is possible to improve even more the approximation to the maximal rate entropy by considering walks of length three. In this case $t = 3$, $(\mathbf{A}^2 \cdot \mathbf{1})_i$ is the number of walks of length three from node i . The number of walks of length three from node i is equal to the sum of all walks of length two starting from the neighbouring nodes of i , if we denote the neighbours of i with i_q then $(\mathbf{A}^2 \cdot \mathbf{1})_i = \sum_{q=1}^{k_i} k_{i_q} = k_i \sum_{q=1}^{k_i} k_{i_q} / k_i = k_i k_i^{(1)}$ where $k_i^{(1)}$ is the average neighbours degree of node i , in this case the random jump probability is [17]

$$P_{i \rightarrow j} = \frac{a_{ij} k_j k_j^{(1)}}{\sum_j a_{ij} k_j k_j^{(1)}}. \quad (9)$$

We would also consider this average neighbours degree random jump for comparison purposes in Section 4.

3 Core-biased random walk

If the largest eigenvalue-eigenvector pair is not known, the MERW results suggests that a good approximation to the largest eigenvector $\bar{v}^{(1)}$ could be used to construct a biased random walk. In here, instead of approximating the MERW by considering walks of increasing length, we use a lower bound of λ_1 based on the network's core, to construct an approximation to the MERW.

A lower bound for the largest eigenvalue is $\lambda_1 \geq (W_t/W_0)^{1/t}$, $t = 1, \dots$ [19] where $W_t = \bar{u}^T \mathbf{A}^t \bar{u}$ is the total number of walks of length t , \mathbf{A} is the adjacency matrix and \bar{u} is a vector with all its entries equal to one. A lower bound for the number of walks is $W_t \leq \sum_{i=1}^N k_i^t$ [6] where the equality is true only if $t \leq 2$.

In a network where the nodes are ranked in decreasing order of their degree, the connectivity of the network can be described with the degree sequence $\{k_r\}$ and the sequence $\{k_r^+\}$ where k_r^+ is the number of links that node r shares with nodes of higher rank. A bound of the largest eigenvalue in terms of the $\{k_r^+\}$ sequence is $\lambda_1 \geq 2\langle k^+ \rangle_r$, where $\langle k^+ \rangle_r = (1/r) \sum_{i=1}^r k_i^+$ is the average number of links shared by the top ranked r nodes [12]. If $r = N$ then we recover the well known bound $\lambda_1 \geq W_1/W_0 = 2(\sum_i^N k_i^+)/N = L/N$ where $W_0 = N$ is the total number of nodes and $W_1 = L$ is the total number of links.

Consider the adjacency matrix \mathbf{A} of the network where the nodes are ranked in decreasing order of their degree and $\bar{u}(r)$ a vector where its first r entries are set to one and the rest to zero. The vector $\bar{z}(r) = \mathbf{A}\bar{u}(r)$ has entries $z_i(r) = K_i^+(r)$, where $K_i^+(r)$ is the number of links that node i shares with the top r ranked nodes, see Fig. 1. Also notice that $2\langle k^+ \rangle_r = (\bar{u}^T(r) \mathbf{A} \bar{u}(r))/r$. As $2\langle k^+ \rangle_r$ is a lower approximation to λ_1 , then $\bar{z}(r)$ gives an approximation to the eigenvector $\bar{v}^{(1)}$.

This bound based on $\{k_i^+\}$ suggests a core biased random walk. If the top r ranked nodes are the core of the network, then a core-biased random jump is $P_{i \rightarrow j}(r) = a_{ij}K_j^+(r)/\sum_j(a_{ij}K_j^+(r))$. But it is possible that $K_j^+(r) = 0$ if node j has no links with the network's core and then the random-walk will be ill-defined. To avoid this situation we propose

$$P_{i \rightarrow j}(r) = \frac{a_{ij}(K_j^+(r) + 1)}{\sum_j a_{ij}(K_j^+(r) + 1)}. \quad (10)$$

As we want to have the best possible approximation to the maximal rate entropy h_{\max} we define the core as the value of r which maximises the value of $h(r)$, that is

$$r_c = \underset{r}{\operatorname{argmax}} (h(r)). \quad (11)$$

The core are the nodes ranked from 1 to r_c . The value of $h(r)$ is evaluated numerically from the core biased random jump $P_{i \rightarrow j}(r)$ given by Eq. (1) with $f_j = K_j^+(r) + 1$, then evaluating the stationary distribution $\{p_i^*\}$ (Eq. 2) and from this distribution the rate entropy $h(r)$ (Eq. (3)).

Practically to implement this method from the perspective of the random-walker, the walker needs to know the number of neighbours that node j has

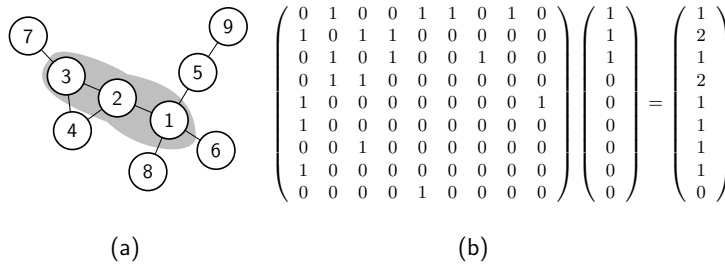


Figure 1: (a) Example of simple network where the nodes are ranked in decreasing order of their degree. The core is the top three ranked nodes, shown in the shaded area. (b) Multiplying the adjacency matrix of this network by the vector with the top three entries equal to 1 and the rest equal to 0 gives the number of links K_i^+ that the nodes share with the core. Notice that node 1, even that is the best connected node, has few connections with the core and node 4, which is not in the core, is well connected with the core. Node 9 has $K_9^+ = 0$ as it does not shares a link with the core.

with the core. The number of neighbours of node j is a local property but, to measure how many of these neighbours are in the core requires the rank of the neighbouring nodes and the rank r_c which defines the core, which are global properties. If there is no core $r_c = 0$ then $K_i^+ = 0$ and we recover a random walk with $P_{i \rightarrow j} = 1/k_i$. The other extreme is if $r = N$, that is all nodes are the core, in this case there is no point to check if the second neighbours of j are in the core, as they will be, so $K_j^+ = k_j$ and $P_{i \rightarrow j} = a_{ij}(1 + k_j)/(k_i + \sum_j k_j)$, which takes into consideration the walks of length one and two.

The evaluation of the core is done by searching the value of r that maximises $h(r)$ (Eq. (11)). This operation requires the full knowledge of the network connectivity hence the computational complexity is $O(L)$ where L is the number of links in the network, which is the same computational complexity as for the MERW. This compares unfavourably with the degree-biased random walk which has computational complexity of $O(\langle k \rangle)$ [17] as it requires, on average, the degree of the neighbouring nodes. However, it is possible to create a core biased random walk which has $O(\langle k \rangle)$ computational complexity. If the overall ranking of the nodes it is not known but we assume that the core are the nodes with degree greater than k^* , then we can redefine $K_j^+(k^*)$ as the number of links that node j has with nodes of degree higher or equal to k^* and

$$P_{i \rightarrow j} = \frac{a_{ij}(K_j^+(k^*) + 1)}{\sum_j a_{ij}(K_j^+(k^*) + 1)} \quad (12)$$

is the jump probability of a biased random walk. Clearly this will not give the best approximation to h_{\max} as we have not put any restriction on the de-

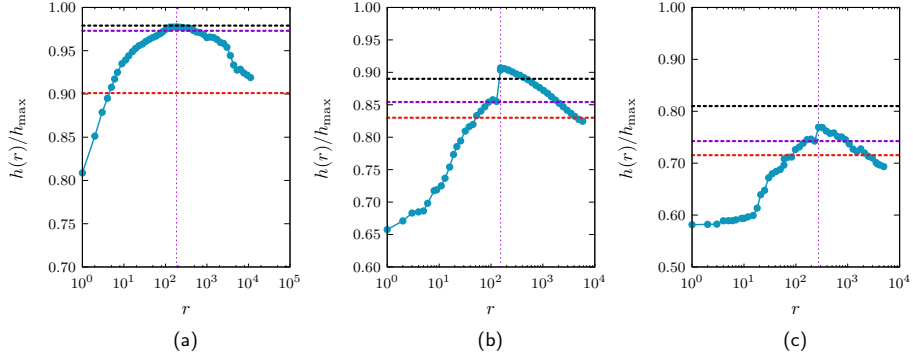


Figure 2: Ratio h/h_{\max} for (a) the AS-Internet, (b) HepTh and (c) Power networks. The vertical lines show the value of r when h is maximal. The horizontal lines are different maximal values of h obtained from degree-biased method (Eq. (8)) (bottom and middle) and average-neighbours degree biased method (Eq. (9)) (top line). The middle horizontal line was obtained when $\alpha = 0.64$, $\alpha = 1.39$ and $\alpha = 1.60$ which maximises the value of h for the AS-Internet, HepTh and Power networks respectively.

gree k^* but as shown in next section this biased random walk can give a good approximation to h_{\max} .

4 Examples

We consider four examples. The first example is about the behaviour of the rate entropy $h(r)$ as a function of the rank r for different real networks. We show that the core biased random walk gives a better approximation to the MERW than the degree biased method. More interestingly, in some networks the core biased gives a better approximation to the average neighbours degree method. The second example is similar to the first example as we compare the approximations to the MERW for different real and synthetic networks. However in this example, we also show that the approximation based on the guess of which nodes form the core can outperform the degree biased method. The third and fourth examples are related to properties of two characteristic times associated with the random walk. The third example considers the *mixing time* for the network of airports. The mixing time is the number of steps for the initial distribution $\{p_i(0)\}$ to reach the stationary distribution $\{p_i^*\}$. This time has practical applications, for example, the mixing time is one of the methods to decide how long it will take for the random walker to collect information about the network and get a sample result which is good enough. The example illustrates that the MERW has the fastest convergence compared to the approximations and the normal random walk. The fourth and last example

presents the return time for the Airports network. This time is the average time that takes to return to the starting node of the walk. In some networks the MERW has the property that the random walker can spend a very long time visiting a selection of nodes compared with a non-biased walk [4]. This behaviour would be reflected in the return time as there would be starting nodes that take a relatively very long time for the walker to return to them. We compare the return time of the MERW against its approximations.

In the first example we consider three real networks, the AS-Internet, the network of co-authors for Physicist researching in High Energy Physics (HepTh) and an electrical distribution network (Power network). Figure 2 shows $h(r)/h_{\max}$ against the size of the core r for these networks. The maximal entropy $h_{\max} = \ln \lambda_1$ was evaluated from the diagonalization of the adjacency matrix \mathbf{A} . In the figure, for comparison purposes, the bottom horizontal line shows the ratio h/h_{\max} for the degree-biased random walk (Eq. (7)), the middle horizontal line for the degree-biased random walk where the value of α maximises h (Eq. (8)) and the top horizontal line the case when the average neighbours degree is included (Eq. (9)).

In the figure the circles show the ratio $h(r)/h_{\max}$ against the size of the core for the core-biased random jump given by Eq. (10). For the three networks there is a large interval in the rank r where the core-biased random walk performs better than the degree-biased random walk (bottom and middle horizontal lines), and in the HepTh network there are values of r where the core-biased random walk is better than the method considering the average neighbours degree (top line). For the Power network the approximation to h_{\max} is not as good as for the AS-Internet or the HepTh. The reason is that the vector $\bar{z}(r)$ with entries $z_i(r) = K_i^+(r)$ gives a poor approximation to the eigenvector $\bar{v}^{(1)}$ [12]. The vertical lines show the rank r where $h(r)$ is maximal. If we define the core of the network at this value then the core size is relatively small. The ratio of number of nodes in the core to the total number of nodes, r_c/N , is 1.6% for the AS-Internet, 2.6% for the HepTh and 5.5% for the Power network.

The results for the second example are in table 1 which compares h/h_{\max} for some real and artificial networks when using different biased walks. In the table, the first five network entries are data from real networks and the last five are synthetic networks. The entry Random is a Erdős-Rényi random network, “Power law” refers to power law networks with different assortativity coefficient and Regular is a random regular network. In general the biased random walks based on the average neighbours degree (Eq. (9)) outperforms the core biased method except for the Les-mis network. The core biased method outperforms the degree biased and the weighted degree biased methods. The last column shows the ratio h/h_{\max} where h is obtained from Eq. (12) for $k^* = 4$, that is, the nodes with degree higher than 4 are considered the core of the network. This crude definition of the core can give a better approximation of h_{\max} than the degree-biased method. We also noticed that the assortativity of the network or the size of the core are not correlated with the values of the rate entropy h .

In this third example we consider the mixing time of the Airports networks. Here we are only considering networks which are undirected, connected and not

Table 1: Comparison of the ratio h/h_{\max} obtained using the core-biased method (h_c/h_{\max}), degree-biased method (h_k/h_{\max} , $\alpha = 1$), weighted degree-biased (h_α/h_{\max}) and average neighbours degree biased method ($h/h_{k_k(1)}$). The column labelled by α shows the value of this parameter that maximises the entropy for h_α . Next column shows the relative size of the core r/N for the core-biased method. The parameter ρ is the assortativity coefficient [13]. The last column is $h_{k^*=4}/h_{\max}$ obtained from the core $k^* = 4$ biased method (Eq. (12)).

Network	h_c/h_{\max}	h_k/h_{\max}	h_α/h_{\max}	$h/h_{k_k(1)}$	r/N	α	ρ	$h_{k^*=4}/h_{\max}$
Airports	0.999	0.991	0.995	0.999	0.136	0.755	-0.267	0.994
CondMat	0.945	0.885	0.905	0.953	0.039	1.291	0.157	0.891
NetSci	0.914	0.890	0.891	0.934	0.137	1.047	-0.081	0.897
Football	0.998	0.998	0.998	0.999	0.913	1.472	0.162	0.996
LesMis	0.997	0.981	0.982	0.993	0.350	0.941	-0.165	0.986
Random	0.983	0.981	0.982	0.992	0.449	1.086	-0.045	0.980
Power law	0.972	0.960	0.967	0.989	0.227	1.145	-0.004	0.980
Power law	0.970	0.965	0.965	0.985	0.489	1.000	-0.245	0.970
Power law	0.980	0.951	0.968	0.991	0.126	1.291	0.222	0.954
Regular	1.000	1.000	1.000	1.000	1.000	-	-	1.000

bipartite, that means that in the transition $\bar{p}(t+1) = \boldsymbol{\pi}\bar{p}(t)$ the matrix $\boldsymbol{\pi}$ is primitive and as a consequence as time tends to infinity the probability that the walker is in node i is the stationary probability p_i^* . The transition probability is a Markov chain that can be written as $\bar{p}(t) = \boldsymbol{\pi}^t\bar{p}(0)$ where $\bar{p}(0)$ is the initial distribution at time $t = 0$. The mixing time is measured by considering how ‘close’ is the initial distribution to the stationary distribution after time t . Let i be a starting node at time $t = 0$ and $P_{i \rightarrow j}^{(t)}$ the distribution of the walk at time equal t . The variational distance $d(t) = \max_i (1/2 \sum_j |P_{i \rightarrow j}^{(t)} - p_j^*|)$, with $i = 1, \dots, N$, tell us how close is the initial distribution to the stationary distribution at time t .

The *mixing rate* is how fast the random walk converges to the stationary distribution and it is given by the maximal eigenvalue of $\boldsymbol{\pi}$ which is strictly smaller than 1. The eigenvalues ω_i of $\boldsymbol{\pi}$ can be ordered such that $\omega_1 = 1 > \omega_2 > \omega_3 \dots$, then the convergence is $d(t) \sim c\omega_2^t$ as t tends to infinity, where c is a constant. The term ω_2^t is small if t is larger than $-1/\log(\omega_2)$. However, is common that if the state space of the Markov chain is large then ω_2 is close to 1, so to make $d(t)$ small the time t should be bigger than $t_{\text{rel}} = 1/(1 - \omega_2)$. If t_{mix} is the time such that $d(t_{\text{mix}}) \leq \epsilon$ with $0 < \epsilon < 1$ then [18, 2]

$$(t_{\text{rel}} - 1) \log\left(\frac{1}{2\epsilon}\right) \leq t_{\text{mix}}(\epsilon) \leq \log\left(\frac{1}{\epsilon p_{\min}^*}\right) t_{\text{rel}} \quad (13)$$

where p_{\min}^* is the minimal value of p_i^* for $i = 1, \dots, N$. We evaluate this mixing

Table 2: Mixing time for the airport network for different biased random walks when $\epsilon = 10^{-4}$

	t_{rel}	p_{min}^*	t_{mix}
Random	37.0	1.6×10^{-4}	(302,653)
MERW	1.6	7.0×10^{-13}	(5,60)
Core	5.2	3.1×10^{-7}	(35,124)
Degree	5.5	1.4×10^{-7}	(36,133)
Av. Neigh	2.3	2.7×10^{-9}	(11,66)

time for the networks studied here and table 2 shows an example of the t_{mix} bounds obtained for the Airports network. In general we noticed that for all the networks considered here, the MERW mixing time is shorter than the other methods, and the random walk is the largest.

The final example shows a particular characteristic of the MERW. In some networks the random walker can spend a very long time visiting a small selection of nodes compared with a non-biased walk. To get an idea if this is the case in the Airports network we evaluated the return time τ_i . This quantity refers to the time that takes the random walker to return to the starting node i and can be evaluated from the stationary probability as $\tau_i = 1/p_i^*$. If this time is large, that is p_i is small, the walker must be spending a large amount of time somewhere else in the network before returning to the starting node. Fig. 3 compares the return time for the core and degree biased walks against the MERW for the Airports network. The figure shows the return time for every node in the network. In the figure the solid line shows when the return time for the biased walks and the MERW are identical. Fig 3(a) show that the core biased random walk approximates better the MERW than the degree biased walk (Fig. 3(b)), as in the degree biased, the dispersion from the diagonal line is larger. The upper right hand corner of the figures show that the MERW return time for some nodes can be up to six order of magnitude larger than the core or degree biased walks. This is an example where, in the MERW, the walker would spend a long time somewhere else in the network. Notice that the core and degree method have similar return times for all nodes (Fig. 3(c)).

The dataset for the networks CondMat, NetSci, Football, HepTh, Power and LesMis are available from M. Newman’s web page (<http://www-personal.umich.edu/~mejn/netdata/>). The AS-Internet from <https://snap.stanford.edu> and the airport networks from <https://toreopsahl.com/datasets/>. The Erdos-Renyi, Power law and Regular random were generated with igraph.

5 Conclusions

Commonly, the maximal rate entropy random walk is approximated by considering that all walks of a given length are visited with equal probability. In here

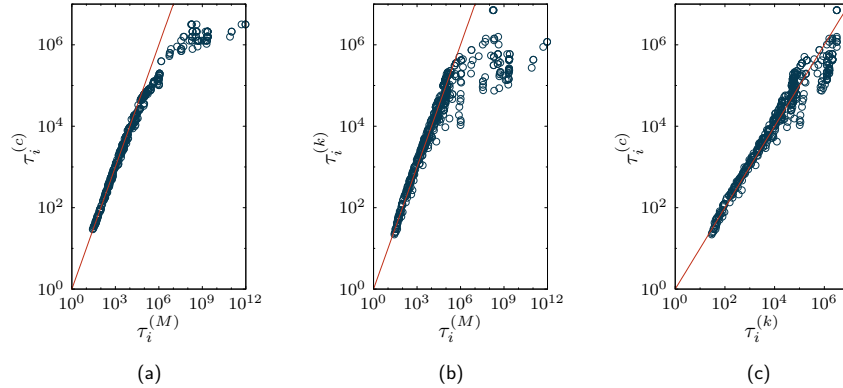


Figure 3: Comparison of the return time of (a) the core biased walks return time $\tau_i^{(c)}$ against the MERW $\tau_i^{(M)}$, where the subindex i labels the starting node. (b) Comparison between the return time of the degree biased walks $\tau_i^{(k)}$ against the MREW and (c) comparison between the core and degree biased return times.

we explored a different approach, the approximation is based on the properties of the well connected nodes. By considering the well connected nodes as the core of the network, the random walker prefers to visit walks that are in the core or are linked to the core. The motivation of using this approach comes from the properties of the MERW jump probability where this probability from one node to a neighbouring node is related to the entries of the eigenvector centrality corresponding to the node and its neighbour. The ranking of the nodes in decreasing order of their degree can give an approximation to the eigenvector centrality as it has been observed that in many networks the degree and eigenvector centralities are correlated. What we have used here is that, once the nodes are ranked, the approximation to the eigenvector centrality can be improved by considering only the connectivity between the nodes that form the core. This approach was motivated by a previous result relating a bound of the largest eigenvalue of the adjacency matrix with the connectivity of the core.

The core biased random walk gives a better approximation to the MERW than the degree biased and for some networks better than the average neighbours degree biased method. The core biased requires the full knowledge of the network connectivity which compares unfavourably with degree and average neighbour degree which only need local knowledge of the network. However, it is possible to construct a core biased method based in local knowledge by guessing the minimal degree of the nodes that form the core. This crude approach can give a better approximation of the MERW than the degree biased method. The core biased random walk, similarly to the degree and average neighbours degree methods, has a fast mixing time which is an important characteristic to have if the random walk is used in the discovery of network structure.

References

- [1] BATTISTON, F., NICOSIA, V. & LATORA, V. (2016) Efficient exploration of multiplex networks. *New Journal of Physics*, **18**(4), 043035.
- [2] BERESTYCKI, N. (2014) Lectures on Mixing Times. *Cambridge University*.
- [3] BONAVENTURA, M., NICOSIA, V. & LATORA, V. (2014) Characteristic times of biased random walks on complex networks. *Physical Review E*, **89**(1), 012803.
- [4] BURDA, Z., DUDA, J., LUCK, J.-M. & WACLAW, B. (2009) Localization of the maximal entropy random walk. *Physical review letters*, **102**(16), 160602.
- [5] COVER, T. M. & THOMAS, J. A. (2012) *Elements of information theory*. John Wiley & Sons.
- [6] FIOL, M. A. & GARRIGA, E. (2009) Number of walks and degree powers in a graph. *Discrete Mathematics*, **309**(8), 2613–2614.
- [7] FRONCZAK, A. & FRONCZAK, P. (2009) Biased random walks in complex networks: The role of local navigation rules. *Physical Review E*, **80**(1), 016107.
- [8] GÓMEZ-GARDEÑES, J. & LATORA, V. (2008) Entropy rate of diffusion processes on complex networks. *Physical Review E*, **78**(6), 065102.
- [9] LEIBNITZ, K., SHIMOKAWA, T., PEPPER, F. & MURATA, M. (2013) Maximum entropy based randomized routing in data-centric networks. in *Network Operations and Management Symposium (APNOMS), 2013 15th Asia-Pacific*, pp. 1–6. IEEE.
- [10] LI, R.-H., YU, J. X. & LIU, J. (2011) Link prediction: the power of maximal entropy random walk. in *Proceedings of the 20th ACM international conference on Information and knowledge management*, pp. 1147–1156. ACM.
- [11] LIU, W. & LÜ, L. (2010) Link prediction based on local random walk. *EPL (Europhysics Letters)*, **89**(5), 58007.
- [12] MONDRAGÓN, R. J. (2016) Network partition via a bound of the spectral radius. *Journal of Complex Networks*, p. cnw029.
- [13] NEWMAN, M. E. (2002) Assortative mixing in networks. *Physical review letters*, **89**(20), 208701.
- [14] OCHAB, J. K. & BURDA, Z. (2013) Maximal entropy random walk in community detection. *The European Physical Journal Special Topics*, **216**(1), 73–81.

- [15] PU, C., LI, S. & YANG, J. (2015) Epidemic spreading driven by biased random walks. *Physica A: Statistical Mechanics and its Applications*, **432**, 230–239.
- [16] ROSVALL, M. & BERGSTROM, C. T. (2008) Maps of random walks on complex networks reveal community structure. *Proceedings of the National Academy of Sciences*, **105**(4), 1118–1123.
- [17] SINATRA, R., GÓMEZ-GARDENES, J., LAMBIOTTE, R., NICOSIA, V. & LATORA, V. (2011) Maximal-entropy random walks in complex networks with limited information. *Physical Review E*, **83**(3), 030103.
- [18] SINCLAIR, A. (1992) Improved bounds for mixing rates of Markov chains and multicommodity flow. *Combinatorics, probability and Computing*, **1**(4), 351–370.
- [19] VAN MIEGHEM, P. (2010) *Graph spectra for complex networks*. Cambridge University Press.
- [20] YOON, S.-H., YUN, Y.-K. & KIM, Y. (2010) Biased random walk sampling on assortative networks. *Journal of the Korean Physical Society*, **56**(3), 990–993.
- [21] YOON, S., LEE, S., YOON, S.-H. & KIM, Y. (2007) Statistical properties of sampled networks by random walks. *Physical Review E*, **75**(4), 046114.
- [22] YU, J.-G., ZHAO, J., TIAN, J. & TAN, Y. (2014) Maximal entropy random walk for region-based visual saliency. *IEEE transactions on cybernetics*, **44**(9), 1661–1672.