



OPEN

# Optimal search strategies on complex multi-linked networks

SUBJECT AREAS:

COMPLEX NETWORKS  
NONLINEAR PHENOMENAFrancesca Di Patti<sup>1,2</sup>, Duccio Fanelli<sup>1,2</sup> & Francesco Piazza<sup>3</sup>

<sup>1</sup>Università degli Studi di Firenze, Dipartimento di Fisica e Astronomia and CSDC, via G. Sansone 1, 50019 Sesto Fiorentino, Firenze, Italia, <sup>2</sup>INFN, Sezione di Firenze, Italia, <sup>3</sup>Université d'Orléans, Centre de Biophysique Moléculaire, CNRS-UPR4301, Rue C. Sadron, 45071, Orléans, France.

Received  
8 January 2015Accepted  
10 March 2015Published  
7 May 2015

In this paper we consider the problem of optimal search strategies on multi-linked networks, i.e. graphs whose nodes are endowed with several independent sets of links. We focus preliminarily on agents randomly hopping along the links of a graph, with the additional possibility of performing non-local hops to randomly chosen nodes with a given probability. We show that an optimal combination of the two jump rules exists that maximises the efficiency of target search, the optimum reflecting the topology of the network. We then generalize our results to multi-linked networks with an arbitrary number of mutually interfering link sets.

Correspondence and requests for materials should be addressed to D.F. (duccio.fanelli@unifi.it)

Complex networks are ubiquitous in nature and play a role of paramount importance in many contexts. Internet and the cyberworld, which permeate our everyday life, are self-organized hierarchical graphs. Urban traffic flows on intricate road networks, which impact both transportation design and epidemic control. In the brain, neurons are cabled through heterogeneous connections, which support the propagation of electric signals. In all these cases, the true challenge is to unveil the mechanisms through which specific dynamical features are modulated by the underlying topology of the network. Here, we study the utterly general problem of optimal search strategies on multi-linked networks, i.e. graphs which have independent sets of edges that connect the same group of nodes.

To this aim, let us start by considering a given agent (e.g. an electric pulse, an excitation, an animal or a human individual, such as a web surfer) located at a node of a network. The agent can hop to a neighbouring node, provided a link exists as specified by the adjacency matrix associated with the graph. The walker wanders on the network through a sequences of steps, that allow for a *local* exploration of its support. In such situations, the efficiency in reaching a specified location may be quantified by the mean first passage time, a robust and widely used measure of transport efficiency on networks in many contexts<sup>5,4</sup>, from biology<sup>19</sup> and ecology<sup>8,10,40</sup> to road network dynamics<sup>17</sup> and quantum systems<sup>16,38</sup>.

However, local moves are not always the best option to reach a target efficiently. For example, facilitated diffusion in the cell nucleus, a mix of one-dimensional gliding along the DNA and three-dimensional jumps to adjacent DNA strands, is believed to account for the efficiency of transcription factors in finding their binding sites<sup>7,32,9</sup>. Analogously, inspired by the behaviour of foraging animals, it has been hypothesised that the *local* exploration of a connected territory might be complemented by intermittent relocation phases in order to optimize the searching strategy<sup>23</sup>. Accordingly, the animal would venture off-track through ballistic runs from time to time, thus sampling larger portions of space. In such examples, the relative duration of the local and relocation stages may control the optimization of the dual-stage strategy<sup>36</sup>.

Walkers on complex networks could in principle rely on similar integrated strategies, possibly tuned to the heterogeneous nature of the underlying support. In the seminal paper by Kleinberg<sup>25</sup>, the navigation on a regular two dimensional lattice distorted with the inclusion of long-range connections was studied and an optimum for the delivery time found, when tuning the probability of off-lattice connections between nodes. Exact results and new insights on the navigability of small-world networks were later reported independently in Refs. 14 and 15. Since the pioneering contribution by Kleinberg, the problem of assisted exploration of a network has been addressed in a number of different contexts<sup>22,26,28,29,30,35,37</sup>, mainly relying on numerical investigations targeted to the problem at hand.

To help intuition let us consider, for example, web surfing. Starting the exploration from an arbitrary web page, one usually proceeds by following the hyperlinks which are therein available. This is a *local* search, which the user abandons when she opens a new tab to look for a different, potentially related topic, eventually landing into another virtual compartment which will be again probed locally for some time. On a different level, the brain displays



multi-layered architectures of connections, that assist the finely orchestrated spatio-temporal patterns underlying brain function<sup>6</sup>. One may then speculate that electric signals can be transmitted across different layers, thus realizing *de facto* non-local jumps in the overall brain connectome between single-layer connected components. Non-local jumps (moves directed towards target sites which are not physically linked to the departure node) may be also loosely referred to as *long-range*, even though the underlying network is not necessarily endowed with a metric.

Building upon these ideas, we start by investigating the conditions for optimal target searches on a generic network of  $N$  nodes. In order to quantify search efficiency on a given network, we shall analytically compute mean first passage times<sup>31,34,24</sup>, which are widely used to gauge search strategies in many contexts<sup>42,2,21,46</sup>. To investigate the combined effect of local and non-local moves, we study a simple stochastic process which accommodates for both local diffusion and non-local relocation to sites which are not connected through the adjacency matrix.

An interesting application of such analysis is random walk on two networks that share the same nodes but are endowed with two different sets of links. Altogether, the two nested graphs sharing the same nodes might be termed a *multi-linked* network. Only local moves are allowed in this case, as prescribed by the two adjacency matrices associated with the two link sets. Moving from this observation, in the second part of the paper we will examine a more general setting where agents can diffuse in multi-linked networks endowed with an arbitrary number of independent link sets. The ensuing sub-networks can be imagined to be assembled as a stratified hierarchy of independent levels, each hosting an identical replica of the nodes, as in the spirit of a multiplex<sup>20,11,33,1</sup>. Indeed, it is straightforward to extend our analysis to the relevant, and more general situation where only a subset of the nodes is present in each nested level.

## Results

**Relocation-assisted diffusion.** Let  $A$  denote the  $N \times N$  adjacency matrix of the network, with  $A_{ij}=1$  if  $i$  and  $j$  are physically connected by a link, and  $A_{ij}=0$  otherwise. The degree of node  $i$  is given by  $k_i^A = \sum_j A_{ij}$ . The probability that a particle sitting at node  $i$  jumps on any other node  $j$  is specified by the following matrix

$$T_{ij} = \alpha \frac{A_{ij}}{k_i^A} + (1 - \alpha) \frac{S_{ij}}{k_i^S} \quad (1)$$

where  $S_{ij} \in \{0,1\}$  are the entries of a random symmetric sparse  $N \times N$  matrix, that controls the relocation via non-local hops (self-loops are, in principle, also allowed for). The density of ones in  $S$  is measured by the parameter  $\delta \in [0,1]$ , so that the average number of nodes that can be reached from any node  $i$  via off-network long-range jumps is  $\langle k_i^S \rangle = \langle \sum_j S_{ij} \rangle = N\delta$ . The parameter  $\alpha \in [0,1]$  tunes the relative strength of the two competing mechanisms, local diffusion and random relocation. When  $\alpha=1$  the walker explores the network according to a purely local rule, while in the opposite limit,  $\alpha=0$ , hopping towards disconnected sites are the only allowed moves. For  $\delta=1$ , the matrix  $S$  is filled with ones and  $T$  becomes the known Google matrix used in the PageRank Algorithm<sup>12,27</sup>.

We define the search time as the time needed by a particle starting at node  $i$  to reach an absorbing trap located at node  $j$ . This satisfies the following relation (see Methods)

$$t_{i \rightarrow j} = \sum_{k=1}^{N-1} \left( \mathbf{Z}_j^{-1} \right)_{ik} \quad (2)$$

where  $\mathbf{Z}_j = \mathbb{I}_{N-1} - \mathbf{T}_j$ . The subscript  $j$  indicates an  $(N-1) \times (N-1)$  submatrix obtained by suppressing the  $j$ -th row and the  $j$ -th column, while  $\mathbb{I}_{N-1}$  denotes the identity matrix of size  $N-1$ . To assess the

overall ability of the walker to find a target, we introduce a global parameter  $\langle t \rangle$  by averaging Eq. (2) over all possible starting nodes ( $i$ ) and trap locations ( $j$ ), that is,

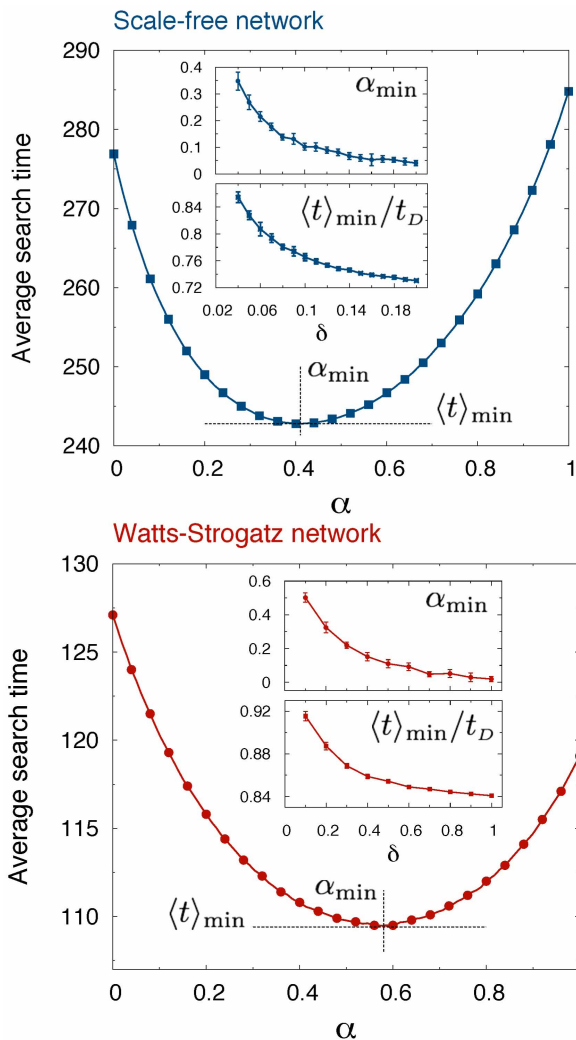
$$\langle t \rangle = \frac{1}{N(N-1)} \sum_{j \neq i} t_{i \rightarrow j} \quad (3)$$

In short,  $\langle t \rangle$  quantifies the ability of the walker to search for targets at the global scale of the network. The shorter  $\langle t \rangle$ , the more efficient the search. The quantity  $\alpha$  acts as a free parameter – it can be adjusted to select the optimal balance between local and non-local hops, with the aim of minimizing the global exploration time.

Figure 1 illustrates how  $\langle t \rangle$  changes as a function of the relative weight of local and non-local moves for two different classes of synthetic undirected networks, the scale-free<sup>3,13</sup> and the small-world<sup>44</sup> networks. The curves display a clear minimum, implying the existence of an optimal value of  $\alpha$  which minimizes the search time. Exactly the same behaviour is displayed by directed networks. The location of the minimum depends on the topology of the network, which defines the backbone for local diffusion, but also on the average number of sites that can be reached through a single long-range hop,  $N\delta$ . Remarkably, the fewer sites are accessible through non-local jumps (*i.e.* the smaller  $\delta$ ), the more pronounced the optimality condition (see upper insets in Fig. 1).

When  $\delta \rightarrow 1$ ,  $\alpha_{min}$  approaches (but never reaches) the limiting solution  $\alpha_{min} = 0$ . In this case, the walker can virtually land on any node with just one jump (the matrix  $S$  is completely filled with ones), and local diffusion contributes modestly to further reduce the average searching time. Although a minimum always exists also for  $\delta = 1$  (the Google Matrix case),  $\langle t \rangle_{min}$  is very close to  $N$ , the time the walker needs to reach an isolated trap when  $\alpha$  is exactly set to zero. Conversely, when  $\delta < 1$ , non-local short-cuts are only available towards a subset of nodes. This is a more plausible situation, bearing in mind the afore-mentioned applications. When surfing the web, from time to time one will abandon a given area of exploration to look for the presumed central node of a new region that she wishes to sample. Similarly, long-range connections in the brain, established through trans-layer channels, are certainly fewer than those accounting for effective bridges among the  $N$  nodes of a given layer.

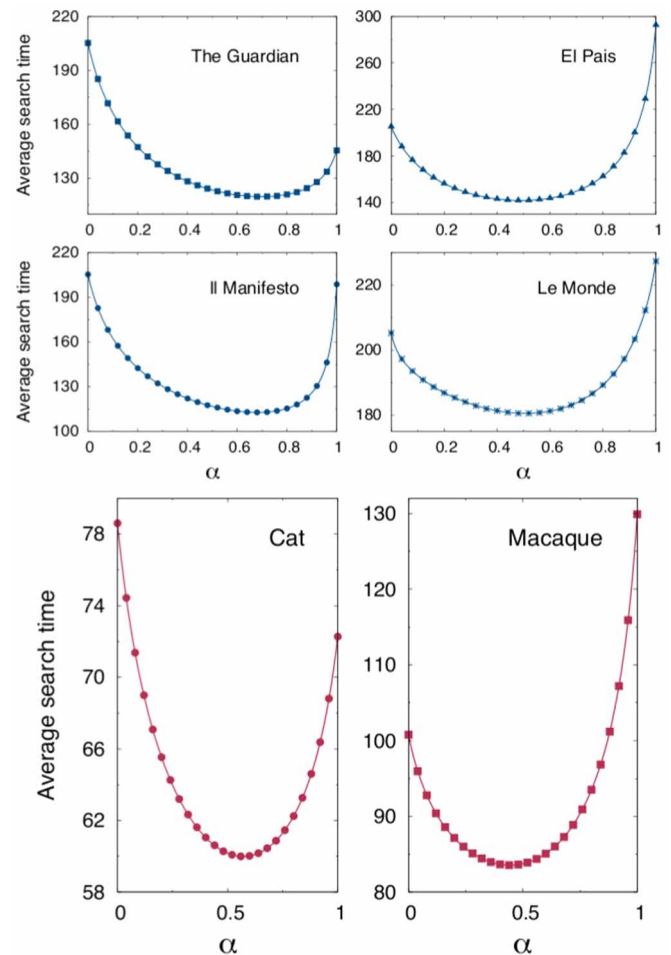
The relocation-assisted search is 10–15% more efficient with respect to the purely local dynamics for intermediate values of the density  $\delta$  of available distant nodes (insets in Fig. 1). The same analysis performed with different values of the average connectivity  $\langle k \rangle$  (scale-free network) and of the rewiring parameter  $p$  (Watts-Strogatz) yields similar results. In particular, upon decreasing  $\langle k \rangle$  one recovers the same qualitative behaviour as obtained when increasing  $\delta$  (data not shown). In this paper, we have chosen to focus on small networks because of the computational cost required to invert the matrix  $\mathbf{Z}_j$ , for any given  $j$ . However, the method is utterly general and may be applied to systems of arbitrary size  $N$ . In the Supplementary Information, we show that  $\alpha_{min}$  progressively decreases as  $N$  increases, while keeping  $\delta$  constant. Conversely, working at fixed size  $N$  one can change the position of the minimum as sought, by acting on  $\delta$ . The counter-intuitive conclusion is that the density of non-local links should be reduced as  $N$  gets larger to keep the optimum value of  $\alpha$  fixed. Interestingly, the advantage brought about by the assisted search over purely local diffusion is preserved if  $N$  and  $\delta$  are simultaneously tuned so to keep  $\alpha_{min}$  fixed. The data reported in the Supplementary Information refer to the Watts-Strogatz setting but the same conclusions hold in general. As a side remark, we recall that the searching time can be expressed as a function of the nodes connectivity, for specific classes of networks, in the large  $N$  limit<sup>43</sup>. Such large networks could share the same characteristics of searchability for the limiting choice  $\alpha=1$ . When introducing the effect of long-range teleportation one can in principle remove such degeneracy and obtain distinct optimal values of  $\alpha$ .



**Figure 1** | The average search time on synthetic networks displays an optimum as a function of the relative weight of local and non-local moves. Upper panel: scale-free network generated with the preferential attachment method<sup>3</sup> with  $N = 200$  and average connectivity  $\langle k \rangle = 20$ . Lower panel: Watts and Strogatz small-world random network with  $p = 0.5^{44}$ ,  $N = 100$  and average connectivity  $\langle k \rangle = 9.5$ . Here the sparse symmetric matrix  $S$  has been generated with  $\delta = 0.04$  (scale-free) and  $\delta = 0.1$  (Watts-Strogatz). The insets show the position of the minimum  $\alpha_{\min}$  and the corresponding shortest average time  $\langle t \rangle_{\min}$  (normalized to the case of a purely local walker,  $t_D \equiv \langle t \rangle_{\alpha=1}$ ) as a function of the average fraction of long-range accessible nodes  $\delta$ . The data are averaged over 10 independent realizations of the random matrices  $S$  and error bars correspond to one standard deviation.

To confirm the existence of an optimal searching strategy on real data sets, we have extracted the adjacency graph of small portions of the web, starting from the home-pages of four main European newspapers (see Methods). The top panel of Fig. 2 shows that the general picture described above for synthetic data sets is valid for real networks too. This has nothing to do with the peculiar structure of the Web, for the same analysis performed on small neural networks of different animals confirms the existence of a clear minimum in the average search time (bottom panel in Fig. 2).

In all the cases examined,  $\langle t \rangle$  appears to be a convex smooth function of  $\alpha$  with a clear minimum. One may ask whether this is a widespread feature of many graphs. More generally, it would be helpful to work out a quantitative criterion to predict whether an optimal search strategy exists at all, possibly also identifying the



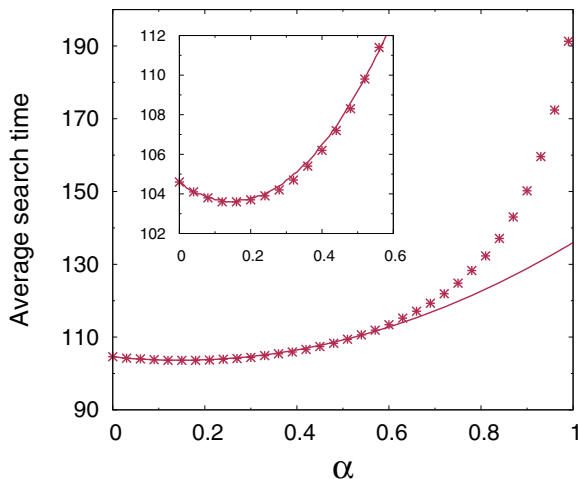
**Figure 2** | The average search time on real data sets displays an optimum as a function of the relative weight of local and non-local moves. Top: average time  $\langle t \rangle$  as a function of  $\alpha$  for four real Web subgraphs. The  $100 \times 100$  adjacency matrices have been mapped out by a Web crawler starting from the web sites of four major European newspapers (see Methods). The sparse symmetric matrix  $S$  has been generated with  $\delta = 0.04$ . Bottom: search time in two neuronal networks: cortical connectivity network of cats (52 nodes, left<sup>39</sup>) and macaques (71 nodes, right<sup>45</sup>). In both cases we have used  $\delta = 0.1$ .

optimal balance between local and non-local moves required to place oneself in such regime. Unfortunately, the exact dependence of  $\langle t \rangle$  on  $\alpha$  is hidden in the inverse of the matrix  $Z$  which, in general, cannot be computed explicitly. However, a principle of this sort can be formulated by resorting to a perturbative approach. If we assume that the stationary point is located at sufficiently small values of  $\alpha$ , then we may consider a search time of the form

$$\langle t \rangle \simeq c_0 - c_1 \alpha + c_2 \alpha^2 \quad (4)$$

In this case, the coefficients  $c_0$ ,  $c_1$  and  $c_2$ , which depend on the topology of the network, can be computed analytically (see Methods). A necessary and sufficient condition for a meaningful minimum to occur is  $c_1 > 0$  and  $c_2 > 0$  with  $c_1 < 2c_2$ , which ensures that  $\alpha_{\min} < 1$ . This provides a handy rule to enquire about the existence of an optimality condition for any given network.

In all the cases that we examined, the coefficients  $c_k$  turn out to be positive. Therefore a minimum is always predicted to exist under the quadratic approximation, and closed expressions for both  $\alpha_{\min}$  and  $\langle t \rangle_{\min}$  can be readily obtained. These match well the exact data computed through Eq. (3). The agreement is of course better when the minimum is found close to  $\alpha = 0$  (Fig. 2–4). Explicit analytical



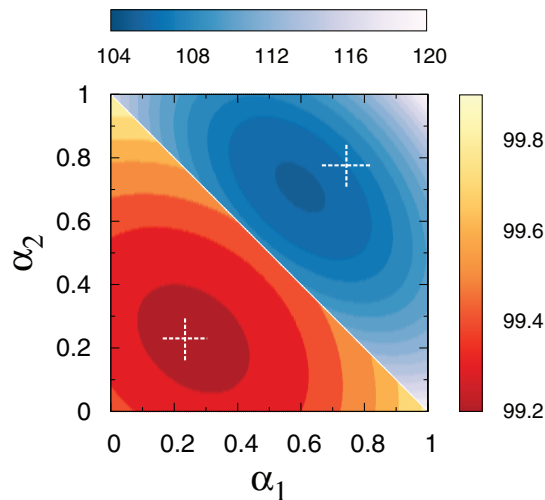
**Figure 3** | Formula (4) provides a convenient tool to enquire about the existence of an optimality criterion for the search time in a given network. The average search time in a small random network of  $N = 100$  nodes (symbols) computed from Eq. (3) is compared to the approximated quadratic profile described by Eq. (4) (solid line). The inset shows a close-up of the region around the minimum. Other parameters are:  $p = 0.1$ ,  $\delta = 0.29$ .

expressions can be obtained in some limiting cases. When  $\delta = 1$  one recovers the Google matrix and the transition rate from node  $i$  to node  $j$  reads  $T_{ij} = \alpha A_{ij}/k_i^A + (1 - \alpha)/N$ . In this case it is not difficult to show that  $c_0 = N$ ,  $c_1 = N/(N - 1)$  and  $c_2 = N/(N - 1) \left[ \sum_{j=1}^N \gamma_j (\gamma_j + 1/k_j^A) - (N + 1) \right]$ , where  $\gamma_j = \sum_{i=1}^N A_{ij}/k_i^A$  (see Supplementary Information for the full derivation). In the case of a regular lattice of connectivity  $\kappa$ , one immediately finds  $\alpha_{min} = \kappa/(2N)$  and  $\langle t \rangle_{min} = N - c/(4N)$ . The more links per node are added, the larger the value of  $\alpha_{min}$  ( $\in [1/N, 0.5]$ ), and the deeper the minimum for  $\langle t \rangle$  vs  $\alpha$ . Although  $\langle t \rangle_{min}$  is shorter than the search time  $t_D$  obtained for  $\alpha = 0$ , the associated correction is just  $O(1/N)$ .

As remarked at the beginning, the analysis carried out above can be interpreted as diffusion on graphs characterized by a single set of nodes and two independent sets of links, represented by two different adjacency matrices  $A$  and  $S$ . This kind of network could be termed *double-linked*. Building on this observation, we discuss in the following the generalization of our approach to arbitrary *multi-linked* networks, where an arbitrary number of independent link sets is imposed upon the same ensemble of nodes.

**The general case of  $M$  networks.** Let us consider agents hopping on a given set of nodes, that are joined together through  $M$  independent sets of links, represented by as many different adjacency matrices. We term such a structure a *multi-linked* network, *i.e.* a graph where links attached to a given node have an *identity*. For example, in a transportation network two nodes (cities) could be linked through a plane connection (air link-set) and by train (railway link-set). With reference to this setting, we set out to derive the optimal coupling that maximizes the transport efficiency across the system. Again with reference to mobility applications, one can consider the effective design of an integrated urban transportation network<sup>18</sup>. Underground, cars and bus service define three different networks (three different link systems on the same set of nodes): their relative load in terms of users and a smart planning of the respective infrastructures could be studied via a systemic approach that builds on the formalism developed here.

In a multi-linked graph the probability of transition is specified by the following stochastic matrix



**Figure 4** | Existence of an optimum for the average search time in a multi-linked network with  $M = 3$ . Density plot of the average search time for networks with  $N = 100$  nodes and  $M = 3$ . Data are obtained from formula (3), via numerical inversion of matrix (5), for every choice of  $\alpha_1$  and  $\alpha_2$ . In both cases here depicted, the three matrices are generated according to the Watts-Strogatz algorithm, for different choices of the probability  $p_i$  and the initial connectivity  $\kappa_i$ , ( $i = 1, 2, 3$ ):  $p_1 = p_2 = p_3 = 0.1$ ,  $\kappa_1 = 70$ ,  $\kappa_2 = 80$ ,  $\kappa_3 = 90$  (lower triangle),  $p_1 = 0.3$ ,  $p_2 = 0.5$ ,  $p_3 = 0.7$ ,  $\kappa_1 = \kappa_2 = \kappa_3 = 10$  (upper triangle). The dashed crosses identify the optimal values for  $\alpha_1$  and  $\alpha_2$  as predicted by the approximated numerical expression (8).

$$T = \sum_{l=1}^{M-1} \alpha_l \mathbf{K}_l^{-1} \mathbf{A}_l + \left( 1 - \sum_{l=1}^{M-1} \alpha_l \right) \mathbf{K}_M^{-1} \mathbf{A}_M$$

where  $\mathbf{A}_l$  are  $M$  matrices of size  $N \times N$ , and  $\mathbf{K}_l$  are their associated diagonal matrices of connectivities, namely  $\mathbf{K}_l = \text{diag}(k_1^A, \dots, k_N^A)$ . In this case the reduced matrix  $\mathbf{Z}_j$  reads:

$$\mathbf{Z}_j = \mathbb{I}_{N-1} - \left( \mathbf{K}_M^{-1} \mathbf{A}_M \right)_j - \sum_{l=1}^{M-1} \alpha_l \left( \mathbf{K}_l^{-1} \mathbf{A}_l \right)_j + \sum_{l=1}^{M-1} \alpha_l \left( \mathbf{K}_M^{-1} \mathbf{A}_M \right)_j \tag{5}$$

where  $j$  flags the position of the absorbing trap. The time it takes to reach  $j$  starting from  $i$  can be computed again by means of expression (2) and the average trapping time  $\langle t \rangle$  estimated through eq. (3). Following a straightforward generalization of the calculation carried out in the preceding paragraph, one can derive an approximate expression for  $\langle t \rangle$  (see Methods)

$$\langle t \rangle \simeq \frac{1}{N(N-1)} \left[ \sum_{l=1}^{M-1} a_l \alpha_l^2 + \sum_{l=1}^{M-1} \sum_{\substack{k=1 \\ k \neq l}}^{M-1} b_{lk} \alpha_l \alpha_k - \sum_{l=1}^{M-1} c_l \alpha_l + d \right] \tag{6}$$

To compute the stationary point of  $\langle t \rangle$ , one needs to impose the condition  $\partial_{\alpha_1} \langle t \rangle = \partial_{\alpha_2} \langle t \rangle = \dots = \partial_{\alpha_{M-1}} \langle t \rangle = 0$ , which is equivalent to setting

$$\begin{pmatrix} 2a_1 & b_{12} & \dots & \dots & b_{1,M-1} \\ b_{21} & 2a_2 & b_{23} & \dots & b_{2,M-1} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \vdots & \vdots & \vdots & \ddots & b_{M-2,M-1} \\ b_{M-1,1} & \dots & \dots & b_{M-1,M-2} & 2a_{M-1} \end{pmatrix} \begin{pmatrix} \alpha_1 \\ \vdots \\ \vdots \\ \vdots \\ \alpha_{M-1} \end{pmatrix} = \begin{pmatrix} c_1 \\ \vdots \\ \vdots \\ \vdots \\ c_{M-1} \end{pmatrix} \tag{7}$$



The stationary point of (6) is a global minimum if the associated Hessian matrix is positive defined.

To clarify the interest of the proposed generalization, we fix  $M = 3$  and investigate the average trapping time as a function of the two independent parameters  $\alpha_1$  and  $\alpha_2$ . In Fig. 4 we show a density map in the  $(\alpha_1, \alpha_2)$  plane of the exact search time  $\langle t \rangle$  computed for two different multi-linked networks following the prescription (3). A minimum is clearly found in both cases, signaling the existence of an optimal choice of the coupling strengths that maximizes the transport efficiency. In the context of the above example, this knowledge could be exploited for the design of efficiently integrated facilities in urban transportation planning, or to favour a smart distribution of users through existing infrastructures.

Approximate analytical expressions for the optimal values of  $\alpha_1$  and  $\alpha_2$  can be readily obtained from system (7), which gives

$$\begin{aligned}\alpha_1 &= \frac{2a_2c_1 - b_{1,2}c_2}{4a_1a_2 - b_{1,2}^2} \\ \alpha_2 &= \frac{2a_1c_2 - b_{1,2}c_1}{4a_1a_2 - b_{1,2}^2}\end{aligned}\quad (8)$$

The white crosses in Fig. 4 refer to the above solution. As it happens in the case of  $M = 2$  analyzed in the previous section, the approximate solution is a better estimate of the true minimum when this is found close to  $\alpha_1 = \alpha_2 = 0$ .

## Discussion

Summarising, in this paper we have addressed the problem of search on networks. To this end, we have preliminarily studied the trapping problem for a modified random walk that combines local hops along the links of the graph with non-local relocation jumps, *i.e.* not along the links contained in the adjacency matrix. We have shown, both for artificial and real datasets, that an optimal balance between local and non-local moves exists which minimizes the average time required to reach a trap. Furthermore, we have derived closed analytical expressions, that enable one to predict the optimal combination as a function of the network topology.

In the second part of the paper we have focused on a random walker diffusing on a arbitrary number of interconnected networks, which share the same set of nodes but are characterized by different adjacency matrices. We have termed a complex structure of this kind a *multi-linked* network. Also in this generalized setting an optimum for the average transport time is found, which can be controlled by tuning the different coupling strengths among the individual sub-networks. In conclusion, the optimality criterion seems to be a universal dynamical mechanism, which might have exerted a critical pressure in the evolutionary selection of many naturally occurring network architectures of the multi-linked kind and that might equally well be exploited in the optimization of man-made technological solutions.

## Methods

**Computation of first passage times.** The mean first passage time  $t_i$ , namely the time it takes for a walker starting at site  $i$  to get to any one of  $N_T$  randomly placed traps, can be computed by extending to the case of a network the standard argument used in the continuum limit for a random walk on a line. Let us consider the interval  $[0, x_0]$  on the real axis and a random walk with two absorbing boundaries located at  $x = 0$  and  $x = x_0$ . The time interval between two jumps is  $\Delta t$  and the lattice spacing is  $\Delta x$ . The exit time  $t(x)$  obeys to  $t(x) = 1/2[t(x + \Delta x) + \Delta t] + 1/2[t(x - \Delta x) + \Delta t]$  meaning that the walker can be regarded as starting one step in the future with equal probability from either  $x + \Delta x$  or  $x - \Delta x$ . The generalization of this equation for a random walk on a network is simply given by  $t_i = \sum_j T_{ij} [t_j + \Delta t]$ , a formulation which proves particularly convenient to investigate the trapping problem. Indeed, re-labelling the nodes of the network so as to have non-trap nodes going from 1 to  $N - N_T$  and all traps located at nodes  $N - N_T + 1$  to  $N$ , one obtains a matrix  $T$  with the last  $N_T$  rows equal to zero. Rearranging correspondingly the array  $t_i$  and recalling that  $t_i = 0$  for  $N - N_T + 1 \leq i \leq N$ , one finds that the exit times are solution of the linear system

$\sum_{j=1}^{N-N_T} Z_{ij} t_j = 1$ , where we have denoted by  $Z$  the upper-left  $(N - N_T) \times (N - N_T)$  block square sub-matrix of  $\mathbb{I}_N - T$ . Eq. (2) is the formal solution of this last equation.

**Adjacency matrices of sub-networks from the web.** To gather real data from the Web we have used the Web crawler *surfer.m* (<http://www.mathworks.com/>). Starting from a selected URL, the crawler identifies all the hyperlinks in the page and adds them to the list of URLs to visit. Once all these URLs are visited, the procedure is repeated recursively for each URL in the list until the assigned number of websites is reached. The outcome of the algorithm is stored in an adjacency matrix where nodes represent the visited pages: the entries of the matrix are 1 if two pages are connected through a hyperlink, 0 otherwise. The matrix is then symmetrized.

**Perturbative expansion of a sum of two matrices.** Let  $C$  and  $B$  be two arbitrary non-singular square matrices of the same dimension and let us introduce the operator  $\Theta$ , that returns the sum of all the elements of a given square matrix. Starting from the relation  $(C + \varepsilon B)^{-1} = (\mathbb{I}_N + \varepsilon C^{-1}B)^{-1} C^{-1}$ , and expressing  $\mathbb{I}_N + \varepsilon C^{-1}B$  by a Neumann series<sup>41</sup>, it follows  $(C + \varepsilon B)^{-1} = C^{-1} - \varepsilon C^{-1}BC^{-1} + \varepsilon^2 C^{-1}BC^{-1}BC^{-1} + \dots$ . To apply this approximation to Eq. (3), we introduce two diagonal matrices associated with  $A$  and  $S$ , namely  $K_A = \text{diag}(k_1^A, \dots, k_N^A)$  and  $K_S = \text{diag}(k_1^S, \dots, k_N^S)$ . In this way,  $T$  takes the form  $\alpha K_A^{-1}A + (1 - \alpha)K_S^{-1}S$ . Consequently, by denoting again by  $j$  the position of the trap, the terms of the reduced matrix  $Z_j$  can be easily rearranged by collecting together those proportional to  $\alpha$ . In formulae:  $Z_j = \mathbb{I}_{N-1} - (K_S^{-1})_j S_j + \alpha [(K_S^{-1})_j S_j - (K_A^{-1})_j A_j]$ . Setting  $C_j = \mathbb{I}_{N-1} - (K_S^{-1})_j S_j$ ,  $B_j = (K_S^{-1})_j S_j - (K_A^{-1})_j A_j$  and  $\varepsilon = \alpha$ , and applying the operator  $\Theta(\cdot)$  to  $(C_j + \varepsilon B_j)^{-1}$ , we recover Eq. (4) with  $c_0 = \sum_j \Theta(C_j^{-1})/N(N-1)$ ,  $c_1 = \sum_j \Theta(C_j^{-1}B_j C_j^{-1})/N(N-1)$  and  $c_2 = \sum_j \Theta(C_j^{-1}B_j C_j^{-1}B_j C_j^{-1})/N(N-1)$ .

**Generalization to the case of  $M$  matrices.** Following the same perturbative scheme detailed above, one can expand the inverse of  $Z_j$  defined in eq. (5) as:  $Z_j^{-1} \simeq C_j^{-1} - \sum_{i=1}^{M-1} \alpha_i C_j^{-1}(B_i)_j C_j^{-1} + \sum_{i=1}^{M-1} \alpha_i^2 C_j^{-1}(B_i)_j C_j^{-1}(B_i)_j C_j^{-1} + \sum_{l=1}^{M-1} \sum_{k=1}^{M-1} \alpha_l \alpha_k C_j^{-1}(B_l)_j C_j^{-1}(B_k)_j C_j^{-1}$ .

Applying the operator  $\Theta(\cdot)$  and summing over all trap locations, one obtains expression (6) for  $\langle t \rangle$ , where:  $a_l = \sum_{j=1}^N \Theta(C_j^{-1}(B_l)_j C_j^{-1}(B_l)_j C_j^{-1})$ ,  $b_{lk} = \sum_{j=1}^N \Theta(C_j^{-1}(B_l)_j C_j^{-1}(B_k)_j C_j^{-1})$ ,  $c_l = \sum_{j=1}^N \Theta(C_j^{-1}(B_l)_j C_j^{-1})$  and  $d = \sum_{j=1}^N \Theta(C_j^{-1})$ .

- Asllani, M., Busiello, D., Carletti, T., Fanelli, D. & Planchon, G. Turing patterns in multiplex networks. *Phys. Rev. E* **90**, 042814 (2014).
- Agliari, E. & Burioni, R. Random walks on deterministic scale-free networks: Exact results. *Phys. Rev. E* **80**, 031125 (2009).
- Barabási, A. L. & Albert, R. Emergence of scaling in random networks. *Science* **286**, 509–512 (1999).
- Baronchelli, A. & Loreto, V. Ring structures and mean first passage time in networks. *Phys. Rev. E* **73**, 026103 (2006).
- Barrat, A., Barthélemy, M. & Vespignani, A. *Dynamical processes in complex networks*. (Cambridge University Press, New-York, 2008).
- Bassett, D. S. & Lynall, M. E. [Network methods to characterize brain structure and function] *Cognitive Neurosciences: the Biology of the Mind* [Gazzaniga, G. R. M. M. & Ivry, R. B. (ed.)] (W. W. Norton, 2014).
- Bauer, M. & Metzler, R. Generalized facilitated diffusion model for dna-binding proteins with search and recognition states. *Biophys. J.* **102**, 2321–2330 (2014).
- Bénichou, O., Coppey, M., Moreau, M., Suet, P. H. & Voituriez, R. Optimal search strategies for hidden targets. *Phys. Rev. Lett.* **94**, 198101 (2005).
- Bénichou, O., Chevalier, C., Meyer, B. & Voituriez, R. Facilitated Diffusion of Proteins on Chromatin. *Phys. Rev. Lett.* **106**, 038102 (2011).
- Bénichou, O., Loverdo, C., Moreau, M. & Voituriez, R. Intermittent search strategies. *Rev. Mod. Phys.* **83**, 81–129 (2011).
- Bianconi, G. Statistical mechanics of multiplex networks: Entropy and overlap. *Phys. Rev. E* **87**, 062806 (2013).
- Brin, S. & Page, L. [The anatomy of a large-scale hypertextual web search engine] *Seventh International World-Wide Web Conference (WWW 1998)* (1998).
- Caldarelli, G. *Scale-Free Networks: Complex Webs in Nature and Technology* (Oxford University Press, 2007).
- Caretta Cortozo, C. and De Los Rios, P. Extended Navigability of Small World Networks: Exact Results and New Insights. *Phys. Rev. Lett.* **102**, 238703 (2009).
- Carmi, S., Carter, S., Sun, J. and ben-Avraham, D. Asymptotic Behavior of the Kleinberg Model. *Phys. Rev. Lett.* **102**, 238702 (2009).
- Caruso, F., Huelga, S. F. & Plenio, M. B. Noise-enhanced classical and quantum capacities in communication networks. *Phys. Rev. Lett.* **105**, 90501 (2010).
- Crisostomi, E., Kirkland, S. & Shorten, R. A google-like model of road network dynamics and its application to regulation and control. *Int. J. Control* **84**, 633–651 (2011).



18. De Domenico, M., Solé-Ribalta, A., Gómez, S. & Arenas, A. Navigability of interconnected networks under random failures. *Proc. Natl. Acad. Sci. U.S.A.* **111**, 8351–8356 (2014).
19. Gallos, L. K., Song, C., Havlin, S. & Makse, H. A. Scaling theory of transport in complex biological networks. *Proc. Natl. Acad. Sci. U.S.A.* **104**, 7746–7751 (2007).
20. Gómez-Gardeñes, J., Reinares, I., Arenas, A. & Floria, L. M. Evolution of cooperation in multiplex networks *Sci. Rep.* **2**, 620 (2012).
21. Haynes, C. P. & Roberts, A. P. Global first-passage times of fractal lattices. *Phys. Rev. E* **78**, 041111 (2008).
22. Lee, S. H., Holme, P. Exploring Maps with Greedy Navigators *Phys. Rev. Lett.* **108**, 128701 (2012).
23. John O'Brien, W., Evans, B. & Browman, H. Flexible search tactics and efficient foraging in saltatory searching animals. *Oecologia* **80**, 100–110 (1989).
24. Kittas, A., Carmi, S., Havlin, S. & Argyrakis, P. Trapping in complex networks. *Europhys. Lett.* **84**, 40008 (2008).
25. Kleinberg, J. M. Navigation in a small world. *Nature* **406**, 854 (2000).
26. Komidis, K., Havlin, S. & Bunde, A. Structural properties of spatially embedded networks. *Europhys. Lett.* **82**, 48005 (2008).
27. Langville, A. M. & Meyer, C. D. *Google's PageRank and Beyond: The Science of Search Engine Rankings* (Princeton University Press, Princeton, 2006).
28. Lee, S. H. & Holme, P. Exploring Maps with Greedy Navigators. *Phys. Rev. Lett.* **108**, 128701 (2012).
29. Li, G. *et al.* Towards Design Principles for Optimal Transport Networks. *Phys. Rev. Lett.* **104**, 018701 (2010).
30. Li, G. *et al.* Optimal transport exponent in spatially embedded networks. *Phys. Rev. E* **87**, 042810 (2013).
31. Lin, Y. & Zhang, Z. Mean first-passage time for maximal-entropy random walks in complex networks. *Sci. Rep.* **4**, 06 (2014).
32. Mirny, L. *et al.* How a protein searches for its site on DNA: the mechanism of facilitated diffusion. *J. Phys. A: Math. Theor.* **42**, 434013 (2009).
33. Nicosia, V., Bianconi, G., Latora, V. & Barthelemy M. Growing Multiplex Networks. *Phys. Rev. Lett.* **111**, 058701 (2013).
34. Noh, J. D. & Rieger, H. Random walks on complex networks. *Phys. Rev. Lett.* **92**, 118701 (2004).
35. Oliveira, C. L. N., Morais, P. A., Moreira, A. A & Andrade, J. S. Enhanced Flow in Small-World Networks. *Phys. Rev. Lett.* **112**, 148701 (2014).
36. Oshanin, G., Lindenberg, K., Wio, H. S. & Burlatsky, S. Efficient search by optimized intermittent random walks. *J. Phys. A: Math. Theor.* **42**, 434008 (2009).
37. Ramezanzpour, A. Intermittent exploration on a scale-free network. *Europhys. Lett.*, **77**, 60004 (2007).
38. Sánchez-Burillo, E., Duch, J., Gómez-Gardeñes, J. & Zueco, D. Quantum navigation and ranking in complex networks. *Sci. Rep.* **2**, 08 (2012).
39. Scannell, J. W., Burns, G. A. P. C., Hilgetag, C. C., O'Neil, M. A & Young, M. P. The connectional organization of the cortico-thalamic system of the cat. *Cerebral Cortex* **9**, 277–299 (1999).
40. Shlesinger, M. F. Mathematical physics: Search research. *Nature* **443**, 281–282 (2006).
41. Stewart, G. W. [Matrix Algorithms: Volume 1: Basic Decompositions] *Other Titles in Applied Mathematics* (SIAM, 1998).
42. Tejedor, V., Bénichou, O & Voituriez, R. Global mean first-passage times of random walks on complex networks. *Phys. Rev. E* **80**, 065104 (2009).
43. von Luxburg U., Radl A. & Hein M., Hitting and Commute times in Large Random Neighbourhood Graphs, *JMLR* **15**, 1751–1798 (2014).
44. Watts, D. J. & Strogatz, S. H. Collective dynamics of small-world networks. *Nature* **393**, 440–442 (1998).
45. Young, M. P. The organization of neural systems in the primate cerebral cortex. *Proc. Biol. Sci.* **252**, 13–18 (1993).
46. Zhang, Z., Lin, Y. & Ma, Y. Effect of trap position on the efficiency of trapping in treelike scale-free networks. *J. Phys. A: Math. Theor.*, **44**, 075102 (2011).

## Acknowledgments

The authors would like to thank Levon Chakhmakhchyan and Filippo Caruso for inspiring discussions and Alessio Cardillo for his critical reading of the manuscript. This work has been partially supported by Ente Cassa di Risparmio di Firenze and program PRIN 2012 founded by the Italian Ministero dell'Istruzione, dell'Università e della Ricerca (MIUR).

## Author Contributions

F. D. P., D.F. and F. P. designed the study and carried out the analysis. F. D. P. carried out the numerical calculations. All authors contributed to the writing of the manuscript.

## Additional information

**Competing financial interests:** The authors declare no competing financial interests.

**How to cite this article:** Di Patti, F., Fanelli, D. & Piazza, F. Optimal search strategies on complex multi-linked networks. *Sci. Rep.* **5**, 9869; DOI:10.1038/srep09869 (2015).



This work is licensed under a Creative Commons Attribution 4.0 International License. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in the credit line; if the material is not included under the Creative Commons license, users will need to obtain permission from the license holder in order to reproduce the material. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>