Googling the Brain: Discovering Hierarchical and Asymmetric Network Structures, with Applications in Neuroscience*

Jonathan J. Crofts† and Desmond J. Higham‡

†*School of Science and Technology, Nottingham Trent University, Nottingham, UK.*

‡*Department of Mathematics and Statistics, University of Strathclyde, Glasgow, UK.*

Abstract

Hierarchical organisation is a common feature of many directed networks arising in nature and technology. For example, a well-defined message-passing framework based on managerial status typically exists in a business organisation. However, in many real-world networks such patterns of hierarchy are unlikely to be quite so transparent. Due to the nature in which empirical data is collated the nodes will often be ordered so as to obscure any underlying structure. In addition, the possibility of even a small number of links violating any overall "chain of command" makes the determination of such structures extremely challenging. Here we address the issue of how to reorder a directed network in order to reveal this type of hierarchy. In doing so we also look at the task of quantifying the level of hierarchy, given a particular node ordering. We look at a variety of approaches. Using ideas from the graph Laplacian literature, we show that a relevant discrete optimization problem leads to a natural hierarchical node ranking. We also show that this ranking arises via a maximum likelihood problem assiocated with a new range-dependent hirearchical random graph model. This random graph insight allows us to compute a likelihood ratio that quantifies the overall tendency for a given network to be hierarchical. We also develop a node ordering algorithm based on the combinatorics of directed walks. In passing, we note that Google's PageRank algorithm tackles a closely related problem, and may also be motivated from a walk-counting viewpoint. We illustrate the performance of the

*This manuscript appears as University of Strathclyde Mathematics and Statistics Research Report 35 (2010). The authors were supported by the Medical Research Council, the Engineering and Physical Sciences Research Council and the Research Councils UK Digital Economy programme.

resulting algorithms on synthetic network data, and on a real-world network from neuroscience where results may be validated biologically.

1 Introduction

Determining hidden structure and substructure within a complex network provides us with a plethora of clues concerning both its functional capabilities and its evolutionary history. This motivates the need for quantitative tools to discover significant topological features, such as well-connected communities, bipartite structures, bottle-necks, motifs, hubs and authorities (5; 15; 16; 21; 31). In this work we focus on a particular type of structure, namely that of a directed hierarchy; a notion that will be defined more precisely as we proceed.

We begin by noting that the phrase *hierarchical* may be used to denote a fractal type of network organisation, whereby smaller modules repeat in a self-similar manner (30; 35). In this work we are using the alternative meaning of hierarchical to denote a directed network structure that supports a well defined message-passing or "chain of command" scenario, such that the processing of information or the exercise of managerial control proceeds sequentially in a top-to-bottom fashion. This is the viewpoint taken in the recent articles (3; 26; 33). Our work differs in that we aim to use first principle arguments based on quantitative measures of network hierarchy in order to derive algorithms that reveal these structures. We also aim to provide a quantitative summary of the overall tendency of a given network to be hierarchical.

We use the following notation. Given an unweighted, directed network consisting of N nodes, we denote by $A \in \mathbb{R}^{N \times N}$ the corresponding adjacency matrix. So A is generally unsymmetric and has $a_{ij} = 1$ if there is an edge from node i to node j, and $a_{ij} = 0$ otherwise. We assume $a_{ii} = 0$ for $i = 1, \dots N$, discounting self-loops. The out and in degree of node k are specified as

$$\deg_k^{\text{out}} := \sum_j a_{kj} \quad \text{and} \quad \deg_k^{\text{in}} := \sum_i a_{ik},$$

respectively. We denote by $\mathcal{P}$ the set of all permutations of the integers $1, 2, \dots, N$, and use p_i to denote the ith component of member $p \in \mathcal{P}$. The Euclidean vector norm is denoted by $\| \cdot \|_2$ and $\mathbf{1} \in \mathbb{R}^N$ represents the vector with all components equal to one.

The layout of the paper is as follows. In Section 2 we recall how the widely used Fiedler vector is relevant as the solution to an optimization problem that emphasizes similarity between nodes. We then introduce a new optimization problem as a means to quantify hierarchy, and show how it can be solved. This results in a simple network reordering algorithm based on the difference between out and in degrees. In Section 3 we exploit ideas from random graph theory that can be used to reorder a network from a maximum likelihood perspective. We show that the algorithm from Section 2 may be viewed as a maximum likelihood reordering under a new, directed, range-dependent random graph hypothesis. In addition to providing an alternative motivation for the algorithm, this viewpoint also allows us to quantify the level of hierarchy in a network by comparing log-likelihood ratios for hierarchical versus non-hierarchical range-dependent structure. We give some illustrative computational results on synthetically generated networks, quantifying the extent to which this approach is tolerant of noise. Section 4 then looks at an alternative, combinatoric approach. We show how the level of hierarchy in a particular node can be conveniently quantified by counting directed walks. We show in section 5 how Google's PageRank algorithm performs a related task. In section 6 we evaluate the new methods on a neuronal network where the hiereachical structure has biological significance.

2 Out Minus In Degree

To motivate the work in this section, we begin with the *two-sum* (1; 20)

$$\sum_{i=1}^{N}\sum_{j=1}^{N}(i-j)^2 a_{ij}.$$

This nonnegative quantity is small when the presence of edges is biased towards nodes that have nearby indices. Equivalently, it is small when the nonzeros in the adjacency matrix appear close to the diagonal. Reordering the nodes corresponds to mapping $i \to p_i$ for some permutation $p \in \mathcal{P}$, and the task of finding a minimum two-sum reordering may thus be written

$$\min_{p\in\mathcal{P}} \sum_{i=1}^{N}\sum_{j=1}^{N}(p_i-p_j)^2 a_{ij}. \tag{1}$$

In general, this discrete optimization problem is computationally intractable and it is therefore common to consider a relaxed version where p is allowed to take real values. To avoid redundancies from scaling and shifting, we also impose constraints $\|p\|_2 = 1$ and $p^T\mathbf{1} = 0$. This leads to

$$\min_{p\in\mathbb{R}^N,\, \|p\|_2=1,\, p^T\mathbf{1}=0} \sum_{i=1}^{N}\sum_{j=1}^{N}(p_i - p_j)^2 a_{ij}. \tag{2}$$

We may now introduce the graph Laplacian $L := D - B$, where $B := \frac{1}{2}(A + A^T)$ and the diagonal matrix D has $d_{ii} = \sum_k b_{ik}$. In the case where the graph represented by B is connected, L is symmetric positive semi-definite with a single eigenvalue equal to zero and corresponding eigenvector proportional to $\mathbf{1}$. It follows that (2) is solved by taking p to be the *Fiedler vector*, v; that is, the eigenvector corresponding to the second smallest eigenvalue of L. We refer to (4; 32; 34) for further details, noting that most authors treat the case where A is symmetric. Having obtained the relaxed solution v, we may recover a permutation $p \in \mathcal{P}$ by the natural procedure of ordering the nodes according to their real-valued components. More precisely, we compute $p \in \mathcal{P}$ such that

$$v_i \le v_j \quad \Longleftrightarrow \quad p_i \le p_j, \tag{3}$$

with some rule for treating ties.

To illustrate this idea, the picture on the left in Figure 1 shows the adjacency matrix for a network showing a strong preference for short-range edges (to produce this matrix, we computed an instance of the short-range dRDRG model described in Section 3, with $\beta = 0.015$ and $N = 100$). After applying a random shuffle to the nodes we obtain the adjacency matrix shown in the middle in Figure 1; any evidence of the range dependent structure has been completely destroyed. Finally, the picture on the right in Figure 1 shows the adjacency matrix when reordered via the Fiedler vector. We see that the hidden structure has been recovered. The respective values of the two-sum starting from left to right were 32037, 1283966 and 30155.

We now consider how this approach can be adapted to discover hierarchical structure. We use the convention that nodes are to be ordered in descending order of importance, so node 1 is at the top of the hierarchy and node N at the bottom. For a given network and a given node ordering,

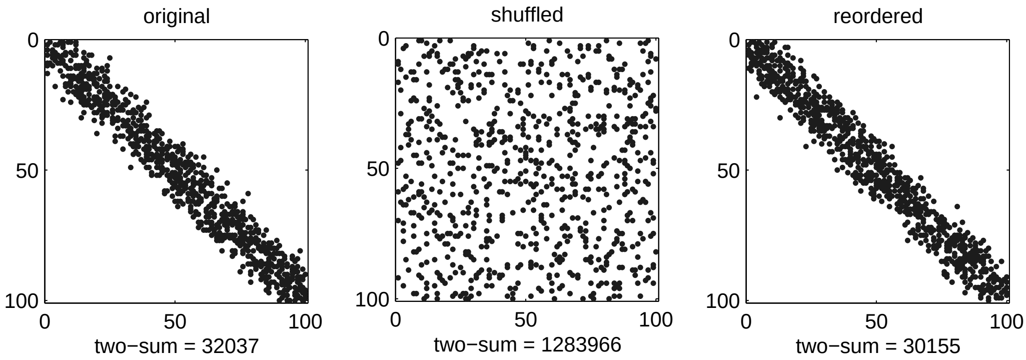

Figure 1: Left: adjacency matrix for a network with short-range connectivity structure. Middle: the connectivity structure is hidden via a random shuffle of the network nodes. Right: structure recovered through a Fiedler vector reordering.

we may then introduce the new concept of a *directed one-sum*

$$\sum_{i=1}^{N}\sum_{j=1}^{N}(i-j)a_{ij},$$

to quantify the level of hierarchy. In this expression, each edge is involved in the overall sum. An edge that respects the hierarchy, $i < j$, contributes a negative amount, with the weight $i - j$ being more negative when i is further up the hierarchy than j. Similarly, an edge that violates the hierarchy, $i > j$, contributes a positive amount, with the weight $i - j$ being more positive when i is further down the hierarchy than j. It follows that a more negative value for the directed one-sum corresponds to a more hierarchical combination of network and node ordering. It is therefore reasonable to consider reordering a given network in order to minimize the directed one-sum:

$$\min_{p\in\mathcal{P}}\sum_{i=1}^{N}\sum_{j=1}^{N}(p_i-p_j)a_{ij}. \tag{4}$$

This is very simply re-written as

$$\min_{p\in\mathcal{P}}\sum_{i=1}^{N}p_i\left(\deg_i^{\mathrm{out}}-\deg_i^{\mathrm{in}}\right),$$

and it follows that the problem (4) is solved by ordering the nodes according to the difference between out and in degrees:

$$\deg_i^{\text{out}} - \deg_i^{\text{in}} \geq \deg_j^{\text{out}} - \deg_j^{\text{in}} \quad \Longleftrightarrow \quad p_i \leq p_j. \tag{5}$$

We emphasize that, unlike in the two-sum case (1)–(3), this ordering solves the discrete formulation of the problem, not just a relaxed version of it.

We note that the authors in (26) considered the *ratio* of in to out degree in order to rank a node within a hierarchy, without giving any justification for the measure. Of course, this is equivalent to a ranking based on the log-difference $\log(\deg_i^{\text{out}}) - \log(\deg_i^{\text{in}})$; whereas the ranking that we have derived via the directed one-sum does not use logs. We note that the new ranking derived here is also able to deal with nodes containing no out-going connections.

3 Random Graph Viewpoint

In this section we show that the out-in ordering (5) can also be justified via a random graph argument, and use this as a means to quantify the amount of hierarchy in a given network.

Extending the range-dependent random graph (RDRG) idea of Grindrod (17; 18), we make the following new definition.

Definition 3.1. *For a given function f that maps from $\{1, 2, \ldots, 2N-1)\}$ to $[0,1]$, a* directed range-dependent random graph *(dRDRG) has an edge from node i to node j with independent probability $f(i-j+N)$.*

We then note that a hierarchical structure will arise when f is monotonically decreasing: in that case the edge from node $i = 1$ to node $j = N$ is the most likely and the reverse edge from $i = N$ to $j = 1$ is the least likely.

Given a network and an ordering, the likelihood that this data came from the dRDRG model is given by

$$\prod_{a_{ij}=1} f(i-j+N) \times \prod_{a_{ij}=0} \left(1 - f(i-j+N)\right).$$

We may then ask for the reordering that is most likely under this model:

$$\max_{p \in \mathcal{P}} \prod_{a_{ij}=1} f(p_i - p_j + N) \times \prod_{a_{ij}=0} \left(1 - f(p_i - p_j + N)\right). \tag{6}$$

The next result connects this problem to the directed one-sum minimization.

Result 1. *The one-sum minimization problem (4) is equivalent to the maximum likelihood reordering problem (6) when the directed range dependency takes the form*

$$f(i - j + N) = \frac{e^{-\alpha(i-j+N)}}{1 + e^{-\alpha(i-j+N)}}, \tag{7}$$

for any fixed $\alpha > 0$.

Proof. We may follow the technique used in (17) from the undirected case by re-writing (6) as

$$\max_{p \in \mathcal{P}} \prod_{a_{ij}=1} \frac{f(p_i - p_j + N)}{1 - f(p_i - p_j + N)} \times \prod_{\text{all } i,j} \left(1 - f(p_i - p_j + N)\right).$$

The second factor, the probability of a null graph, is independent of p, so an equivalent problem is

$$\max_{p \in \mathcal{P}} \prod_{a_{ij}=1} \frac{f(p_i - p_j + N)}{1 - f(p_i - p_j + N)}.$$

When f has the form in (7), this may be re-written as

$$\max_{p \in \mathcal{P}} \prod_{a_{ij}=1} e^{-\alpha(p_i - p_j + N)}.$$

Taking logs and negating the objective function leads to

$$\min_{p \in \mathcal{P}} \alpha \sum_{i=1}^{N} \sum_{j=1}^{N} (p_i - p_j + N)\, a_{ij},$$

which is clearly equivalent to (4). □

The picture on the left in Figure 2 shows an instance of the hierarchical dRDRG model, with the nodes ordered according to the range-dependency, in the case where $N = 100$ and $\alpha = 0.0262$. We

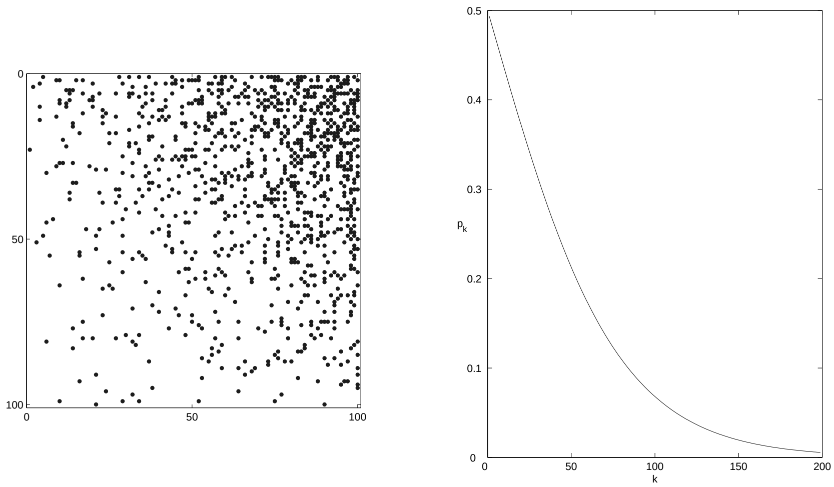

Figure 2: An instance of a hierarchical dRDRG with $N = 100$ and $\alpha = 0.0262$ (left). The model connects node i to node j with probability $f(i - j + N)$, where f has the form shown on the right.

see that the structure is consistent with the aims of the directed one-sum minimization—nonzeros are most prevalent at the upper right hand corner of the matrix, and the density decreases as we move towards the lower left hand corner. In the right of the figure we plot the range-dependency function, f, from (7).

Grindrod's original RDRG model (17) also places nodes in successive positions on an integer lattice. In this case nodes i and j are connected with an independent probability that depends on their *undirected range*, $|i - j|$. More precisely, for a given decay function g, we have $a_{ij} = 1$ with probability $g(|i-j|)$ and $a_{ij} = 0$ otherwise. It was shown in (20) that the corresponding maximum likelihood reordering problem

$$\max_{p \in \mathcal{P}} \prod_{a_{ij}=1} g(|p_i - p_j|) \times \prod_{a_{ij}=0} (1 - g(|p_i - p_j|)) \tag{8}$$

is equivalent to two-sum minimization (1) for the particular range dependency

$$g(|i-j|) = \frac{e^{-\beta(i-j)^2}}{1 + e^{-\beta(i-j)^2}}, \tag{9}$$

where we have introduced a free parameter $\beta > 0$ to be compatible with the general case specified in Result 1.

This insight puts us in a position to compare the likelihood that a given network arose from each of these two model classes. We may summarize such an algorithm as follows.

Step 1 Fit the parameters α and β in the two models. This can be done by matching the number of edges in the given network to the expected number of edges in the random graph model. The latter is a monotonic function of the free parameter in both cases, so a simple bisection root-finding method is guaranteed to succeed.

Step 2 Compute the maximum likelihood from the hierarchical model, say $\mathcal{L}_{\rm h}$, which is given by inserting the minimum directed one-sum ordering into the objective function in (6) with f defined in (7). Compute the approximate[1] maximum likelihood from the non-hierarchical model, say $\mathcal{L}_{\rm nh}$, which is given by inserting the Fiedler vector ordering into the objective function in (8) with g defined in (9).

Step 3 Compute the normalized log likelihood ratio

$$L := \frac{2}{n(n-1)} \log \left(\frac{\mathcal{L}_{\rm nh}}{\mathcal{L}_{\rm h}} \right). \tag{10}$$

A negative ratio indicates that the network data supports the hierarchical range-dependency over the non-hierarchical alternative.

As a proof-of-principle, we give results for a large-scale experiment using synthetically generated networks. In each test, we began with an instance of a network from one of the two random classes under consideration. To assess robustness to noise we then added spurious links at random. For each network instance, we placed a new link between all possible pairs of nodes with small independent probability, q. If the likelihood ratio correctly identified the structure of interest, we increased the noise parameter q and repeated the analysis. Tables 1 and 2 record the smallest value of q (averaged over an ensemble of 1000 random graphs) for which the likelihood ratio incorrectly classified the structure. We used graphs of order $N = 50, 100, 200, 500, 1000$, which had on average $(N^2 - N) \times s$ edges; here $s \in \{0.1, 0.2, 0.3, 0.4, 0.5\}$ denotes the proportion of edges present, on average, for each network instance.

[1] This is approximate because the Fiedler vector solves a relaxed version of the minimum two-sum problem.

	N				
s	50	100	200	500	1000
0.1	0.1150	0.1147	0.1149	0.1122	0.1151
0.2	0.1360	0.1354	0.1370	0.1361	0.1359
0.3	0.1472	0.1472	0.1479	0.1471	0.1478
0.4	0.1548	0.1552	0.1550	0.1547	0.1546
0.5	0.1572	0.1573	0.1577	0.1571	0.1572

Table 1: Proportion of spurious links tolerated when discovering a hierarchical dRDRG network.

	N				
s	50	100	200	500	1000
0.1	0.1290	0.1259	0.1256	0.1271	0.1274
0.2	0.1356	0.1339	0.1352	0.1353	0.1348
0.3	0.1366	0.1360	0.1368	0.1363	0.1357
0.4	0.1362	0.1362	0.1360	0.1365	0.1359
0.5	0.1359	0.1358	0.1358	0.1357	0.1355

Table 2: Proportion of spurious links tolerated when discovering a non-hierarchical RDRG network.

From Tables 1 and 2, we see that the likelihood approach continues to classify the network models correctly even after some 10%–15% of additional links have been added. Note that robustness generally improves as we increase the proportion of connections present in the initial instance; a trend that is particularly evident for the hierarchical model. This observation is consistent with the idea that the presence of additional links translates into additional information concerning network architecture. To match the case of typical real networks, we have focussed on sparse structures. If we tuned α and β to allow for increasingly dense connectivity patterns, then, intuitively, we would expect to see the test become less reliable—as both networks models approach the complete graph, they become indistinguishable.

4 Directed Walks and Matrix Functions

The out-in ordering studied in the previous two sections is computationally convenient and can be justified from first principles. However, it has the possible drawback that, being based on integer-valued quantities nodes may frequently be tied. In particular, for the simple case of a directed binary tree, all nodes except the root have an equal out-in score.

A more global strategy that looks beyond the degree structure was considered in (26), where the concept of *attraction basin hierarchy* was introduced. The resulting algorithm is based upon counting the number of shortest paths to/from a particular node from/to all other nodes, which is a hard combinatorial problem in general. Here, we propose an alternative measure that uses directed walks rather than paths[2]. Recent work has shown that walk-based measures provide an effective tool for determining other types of connectivity pattern (6; 10; 12), and there are at least four specific arguments in their favour:

- information does not generally flow along geodesics (2; 27),
- walk counts are less sensitive than pathlength counts to spurious or missing edges (19),
- walk counts can be computed conveniently using basic operations in linear algebra (11),
- we will show below that Google's successful PageRank algorithm has a walk-based interpretation.

The computational convenience arises because the element $(A^n)_{ij}$ counts the number of directed walks of length n that start at node i and terminate at node j. We may then introduce coefficients $c_0, c_1, c_2, c_3 \ldots$ and consider the expansion

$$F(A) = c_0 I + c_1 A + c_2 A^2 + c_3 A^3 + \cdots . \tag{11}$$

Because longer walks are generally (a) more numerous, and (b) less important than shorter walks, it is reasonable to choose a decreasing sequence of coefficients. Estrada and Rodríguez-Velázquez (14) suggested $c_k = 1/(k!)$, so that $F(A) = \exp(A)$, and this choice has proved to be very useful in many scenarios (6; 7; 9; 10; 12; 13; 11). The resolvent function $F(A) = (I - \delta A)^{-1}$, corresponding to $c_k = \delta^k$, was considered in (11) and can be traced back to earlier work on node centrality indices in social network analysis (23; 28).

We may now classify the hiererchical rank of a node by measuring

- $\sum_{j \neq i} (F(A))_{ij}$, which quantifies how effectively node i is able to pass information to or exert control over the other nodes in the network, and

[2]A walk differs from a path in that nodes and edges may be re-used during the traversal.

- $\sum_{j\neq i}(F(A))_{ji}$, which quantifies how effectively the other nodes in the network are able to pass information to or exert control over node i.

We therefore propose that node i be assigned the hierarchical ranking

$$r_i = \sum_{j\neq i}(F(A))_{ij} - \sum_{j\neq i}(F(A))_{ji}, \tag{12}$$

leading to a reordering in descending rank order; that is

$$r_i \geq r_j \iff p_i \leq p_j.$$

Note that the walk based hierarchy measure of (12) can be considered as a natural generalisation of the degree based algorithm introduced earlier, in the sense that by choosing $F(A) = A$ we recover the out-in degree ordering.

5 Links to Google's PageRank

The well-known PageRank algorithm is used by Google to quantify the importance of web pages based on the hyperlink topology of the WWW (24; 29). Letting A represent the web adjacency matrix, so $a_{ij} = 1$ means that page i has a hyperlink to page j, this algorithm assigns rank according to the vector

$$(I - \theta A^T D^{\text{out}^{-1}})^{-1}\mathbf{1}, \tag{13}$$

with a higher value in component i denoting more importance for page i. Here $\theta \in (0,1)$ is a free parameter and D^{out} is the diagonal out degree matrix, so $D^{\text{out}}_{ii} = \deg^{\text{out}}_i$. We assume for the moment that $\deg^{\text{out}}_i \geq 1$ for all i. Writing $\widehat{A} := A^T D^{\text{out}^{-1}}$ we see that, for sufficiently small θ, the ranking in (13) may be expanded as

$$\left(I + \theta\widehat{A} + \theta^2\widehat{A}^2 + \theta^3\widehat{A}^3 + \cdots\right)\mathbf{1}.$$

This shows that PageRank may be interpreted as a directed walk counting algorithm, with a few twists.

Link direction is reversed and edges are scaled. The new matrix $\widehat{A}$ has $\widehat{a}_{ij} \neq 0$ if and only if there is a hyperlink from page j to page i. This is intuitively reasonable—the $j \mapsto i$ hyperlink may be interpreted as page j deferring to, or handing control over to, page i. (In the WWW context it is also extremely pertinent that page i has no direct influence over the creation of a $j \mapsto i$ hyperlink. This makes it difficult to boost artificially your own page's ranking.) Supposing that the hyperlinks $i_{k-1} \mapsto i_{k-2}$, $i_{k-2} \mapsto i_{k-3}$, $\ldots$, $i_2 \mapsto i_1$ exist on the WWW. Then the walk of length k given by $i_1 \to i_2 \to \cdots i_{k-1}$ exists in the reversed network, but in addition to the θ^k scaling that penalizes long walks, there is also a scaling

$$\frac{1}{\deg_2^{\text{out}}} \frac{1}{\deg_3^{\text{out}}} \cdots \frac{1}{\deg_{k-1}^{\text{out}}}$$

penalizing walks involving "promiscuous" nodes that make themselves available for many other such walks. This method for nullifying the influence of overactive nodes that are likely to be lowly ranked is clearly relevant in the WWW setting. However, in the context of this work, where we are seeking to discover a hierarchical "chain of command," it seems less appropriate to normalize in this way. If node i gives orders to (i.e. has a hyperlink from) a node that also gets orders from (i.e. has hyperlinks to) many other nodes, then this could be regarded as strong evidence for placing node i near the top of the hierarchy.

Closed walks are considered. Node i is given a ranking based on the sum over all j of a weighted count of directed walks that begin at i and end at j, including the closed walk case $j = i$.

6 Local C. elegans Network

In his PhD thesis in 1987, Durbin (8) went about the task of sorting the neurons in the nerve ring of *C. elegans*[3] vertically, in such a way that as many of the synapses as possible pointed downwards. Durbin used an ad hoc combinatoric algorithm to create an ordering. Here, we propose to repeat Durbin's analysis using the automated algorithms that we have derived. We consider a local subnetwork of 130 frontal neurons and 964 chemical synapses of the neuronal network for *C. elegans* (22). We exclude gap junctions from the analysis since experimental techniques are unable

[3] A 1mm long, transparent, roundworm.

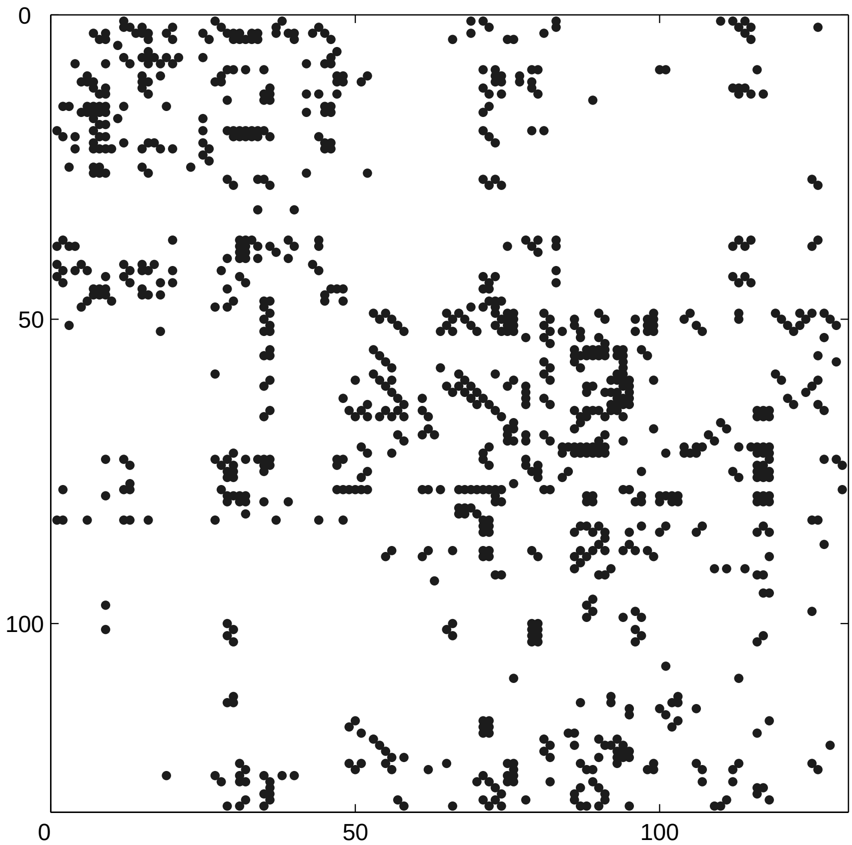

Figure 3: Adjacency matrix for the local network of 130 frontal neurons for *C. elegans*.

to infer directionality in that case. The adjacency matrix is shown in Figure 3.

Figure 4 shows the results of reordering the local *C. elegans* network based upon (i) out-in degree, (ii) the walk based measure (12) with $F(A) = \exp(A)$, (iii) the walk based measure (12) with $F(A) = (I - \delta A)^{-1}$ and $\delta = 0.05$, and (iv) Google's PageRank. In the latter case, dangling nodes, where $\deg_i^{\text{out}} = 0$, require a slight modification of the Google matrix in (13) (24). We used the widely quoted teleporting parameter value $\theta = 0.85$.

Table 3 quantifies the ability of each method to discover hierarchical structure. The first row gives the directed one-sum; that is, the objective function in (4), for each ordering. We note from Result 1 that the out-in degree ordering minimizes this quantity. From Table 3 we see that Equation 4 attains a minimum value of -36247; next came the walk based algorithms, with scores of

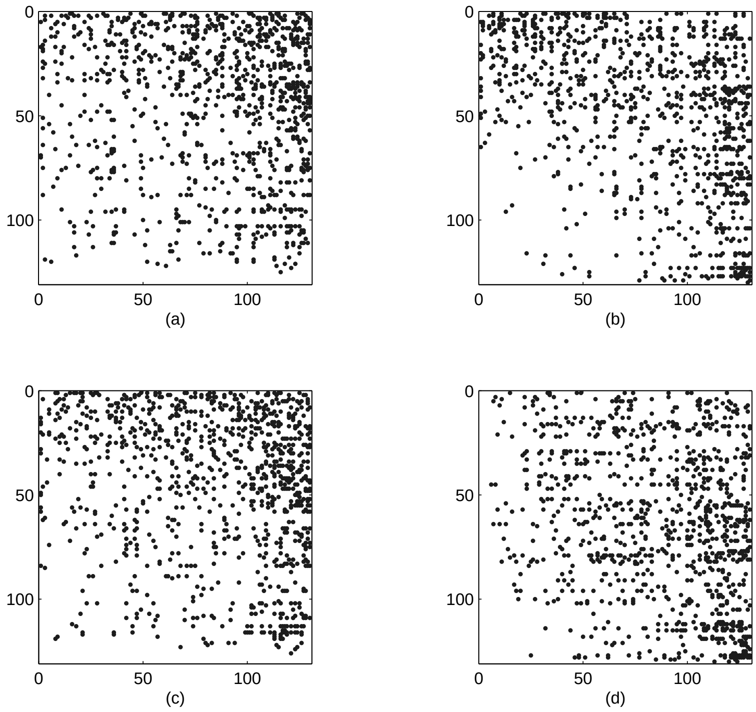

Figure 4: Results for the *C. elegans* neuronal network: (a) reordered using out-in degree; (b) reordered using the walk based measure with $F(A) = \exp(A)$; (c) reordered using the walk based measure with $F(A) = (I - \delta A)^{-1}$ and $\delta = 0.05$; and (d) reordered using Google's PageRank algorithm.

-35748 using the resolvent function and -29610 using the matrix exponential; finally, the ordering obtained via the PageRank algorithm performed poorest according to this measure with a score of -21894. The second row in Table 3 records the proportion of nonzeros in the upper triangle of the adjacency matrix. With this measure, there is no extra benefit/penalty from violating/exploiting the hierarchy with a 'long-range' link that spans many intermediate nodes. Instead we simply count the proportion of links that respect the hierarchy. We see from the table that the out-in ordering is marginally better than that given by the walk based algorithms according to this measure; all methods scoring between 75% to 80%, whilst PageRank achieved around 70%.

To justify the choice of $\delta = 0.05$ in the resolvent-based reorderings, Figure 5 shows how the proportion of non-zeros in the upper triangle (left) and directed one-sum (right) vary as a function of δ—both meaures start to degrade beyond this level.

	out minus in	exponential	resolvent ($\delta = 0.05$)	PageRank
Directed one-sum	-36247	-29610	-35748	-21894
Proportion of non-zeros in upper triangle	0.7832	0.7667	0.7832	0.6929

Table 3: Hierarchy scores for the *C. elegans* network.

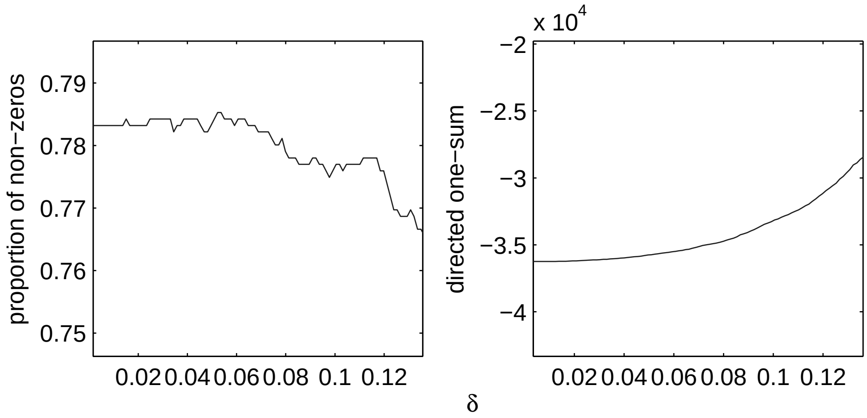

Figure 5: Plots of both hierarchy scores versus δ when using the resolvent function, $F(A) = (I - \delta A)^{-1}$, to reorder the *C. elegans* network.

The normalized log likelihood ratio in (10) was $L = -0.9964$, giving further support for the visually compelling evidence in Figure 4 that this network has a strong hierarchical element.

Let us now focus on biological significance of these results in the particular case of the walk based measure with $F(A) = (I - \delta A)^{-1}$ and $\delta = 0.05$. Table 4 summarizes the results for the reordered *C.elegans* data, where we report those neuronal classes represented by nodes in the top and bottom 10% of the reordered network. Perhaps most noteworthy is the fact that neuronal classes representing sensory neurons are highly prevalent at the top of the hierarchy (some 70%), whilst the foot of the hierarchy consists wholly of a mixture of motor neurons and command interneurons. In general, we found that the ordering returned was in good agreement with that of Durbin's, with sensory neurons placed at the top, motor neurons at the bottom, and interneurons placed in between.

The command interneuron classes AVA, AVB and AVE appearing at the bottom of the hierarchy in Table 4 have been identified in previous studies (25) as being highly connected to both sensory and motor neurons. This suggests that they should lie in the middle of the global hierarchy. However, we note that the majority of postsynaptic connections made by these neurons are with

Top 10%		Bottom 10%	
Neuronal Class	Description	Neuronal Class	Description
RIH	Ring interneuron	SMD	Ring motor
ADL	Amphid sensory neuron	AVB	Command internueron
AIN	Amphid sensory neuron	RMD	Ring motor
AVH	Interneuron	RME	Ring motor
OLL	Head sensory neuron	AVA	Command interneuron
IL2	Head sensory neuron	AVE	Command interneuron
ASI	Amphid sensory neuron		
CEP	Head sensory neuron		
RIR	Ring interneuron		
URB	Head sensory neuron		

Table 4: Neuronal class and type for those nodes contained within the top and bottom 10% after reordering the *C. elegans* network based upon the walk based measure with $F(A) = (I - \delta A)^{-1}$ and $\delta = 0.05$.

motor neurons outside of the local subnetwork studied here, which justifies their placement within this subnetwork

7 Discussion

The new reordering approaches that we derived here from first principles use extremely simple combinatorics, and hence they scale favourably for large, sparse networks. If the expansion (11) is truncated after a finite number, K, of terms (for example, $K = 10$) then the walk-based ranking (12) requires only K sparse matrix-vector multiplications. This gives an overall complexity proportional to the number of edges. The degree-based ordering is comparable with the use of a single term in (11) and also has complexity proportional to the number of edges.

Overall, we believe that is a very promising methodology for discovering and quantifying hierarchy in large, complex, directed networks.

References

[1] S. T. Barnard, A. Pothen, and H. D. Simon, *A spectral algorithm for envelope reduction in sparse matrices*, Numerical Linear Algebra, 2 (1995), pp. 317–334.

[2] S. P. Borgatti, *Centrality and network flow*, Social Networks, 27 (2005), pp. 55–71.

[3] F. Chen, Z. Chen, Z. Liu, L. Xiang, and Z. Yuan, *Finding and evaluating the hierarchical structure in complex networks*, Journal of Physics A: Mathematical and Theoretical, 40 (2007), pp. 5013–5023.

[4] F. Chung, *Spectral Graph Theory*, American Mathematical Society, Providence, RI, 1997.

[5] L. d. F. Costa, F. A. Rodrigues, G. Travieso, and P. R. V. Boas, *Characterization of complex networks: A survey of measurements*, Advances in Physics, 56 (2007), pp. 167–242.

[6] J. J. Crofts, E. Estrada, D. J. Higham, and A. Taylor, *Mapping directed networks*, Electronic Transactions in Numerical Analysis, 37 (2010), pp. 337–350.

[7] J. J. Crofts and D. J. Higham, *A weighted communicability measure applied to complex brain networks*, Journal of the Royal Society Interface, 6 (2009), pp. 411–414.

[8] R. M. Durbin, *Studies on the Development and Organisation of the Nervous System of Caenorhabditis Elegans*, PhD thesis, University of Cambridge, 1987.

[9] E. Estrada and N. Hatano, *Statistical-mechanical approach to subgraph centrality in complex networks*, Chemical Physics Letters, 439 (2007), pp. 247–251.

[10] ——, *Communicability in complex networks*, Physical Review E, 77 (2008), p. 036111.

[11] E. Estrada and D. J. Higham, *Network properties revealed through matrix functions*, SIAM Review, 52 (2010), pp. 696–714.

[12] E. Estrada, D. J. Higham, and N. Hatano, *Communicability and multipartite structure in complex networks at absolute temperatures*, Physical Review E, 78 (2008), p. 026102.

[13] ——, *Communicability betweenness in complex networks*, Physica A, 388 (2009), pp. 764–774.

[14] E. Estrada and J. A. Rodríguez-Velázquez, *Subgraph centrality in complex networks*, Phys. Rev. E, 71 (2005), p. 056103.

[15] S. Fortunato, *Community detection in graphs*, Physics Reports, 486 (2010), pp. 75–174.

[16] M. Girvan and M. E. Newman, *Community structure in social and biological networks*, Proceedings of the National Academy of Sciences, 99 (2002), pp. 7821–6.

[17] P. Grindrod, *Range-dependent random graphs and their application to modeling large small-world proteome datasets*, Physical Review E, 66 (2002), p. 066702.

[18] ——, *Modeling proteome networks with range-dependent random graphs*, American Journal of PharmacoGenomics, 3 (2003), pp. 1–4.

[19] P. Grindrod, D. J. Higham, M. C. Parsons, and E. Estrada, *Communicability across evolving networks*, Tech. Rep. 20, University of Strathclyde, Department of Mathematics and Statistics, 2010.

[20] D. J. Higham, *Unravelling small world networks*, Journal of Computational and Applied Mathematics, 158 (2003), pp. 61–74.

[21] P. Holme, F. Liljeros, C. R. Edling, and B. J. Kim, *Network bipartivity*, Physical Review E, 68 (2003), p. 056107.

[22] M. Kaiser and C. C. Hilgetag, *Nonoptimal component placement, but short processing paths, due to long-distance projections in neural systems*, PLoS Computational Biology, 2 (2006), p. e95.

[23] L. Katz, *A new index derived from sociometric data analysis*, Psychometrika, 18 (1953), pp. 39–43.

[24] A. N. Langville and C. D. Meyer, *Google's PageRank and Beyond*, Princeton University Press, Princeton, New Jersey, 2006.

[25] S. Morita, K. Oshio, Y. Osana, Y. Funabashi, K. Oka, and K. Kawamura, *Geometrical stucture of the neuronal Caenorhabditis elegans*, Physica A, 298 (2001), pp. 553–561.

[26] L. Muchnik, R. Itzhak, S. Solomon, and Y. Louzoun, *Self-emergence of knowledge trees: extraction of wikipedia hierarchies*, Physical Review E, 76 (2007), p. 016106.

[27] M. E. J. Newman, *A measure of betweeness centrality based on random walks*, Social Networks, 27 (2005), pp. 39–54.

[28] ——, *Networks: An Introduction*, Oxford University Press, 2010.

[29] L. Page, S. Brin, R. Motwani, and T. Winograd, *The PageRank citation ranking: Bringing order to the web*, tech. rep., Stanford Digital Library Technologies Project, 1998.

[30] E. Ravasz and A. L. Barabási, *Hierarchical organization in complex networks*, Physical Review E, 67 (2003), p. 026112.

[31] O. Sporns and R. Kotter, *Motifs in brain networks*, PLOS Biology, 2 (2004), pp. 1910–1918.

[32] G. Strang, *Computational Science and Engineering*, Wellesley-Cambridge Press, 2008.

[33] A. Trusina, S. Maslov, P. Minnhagen, and K. Sneppin, *Hierarchy measures in complex networks*, Physical Review Letters, 92 (2004), p. 178702.

[34] R. Van Driessche and D. Roose, *An improved spectral bisection algorithm and its application to dynamic load balancing*, Parallel Computing, 21 (1995), pp. 29–48.

[35] C. Zhou, L. Zemanová, G. Zamora, C. C. Hilgetag, and J. Kurths, *Hierarchical organization unveiled by functional connectivity in complex brain networks*, Physical Review Letters, 97 (2006), p. 238103.