School of Mathematical
and Physical Sciences

University of Reading

Department of Mathematics and Statistics

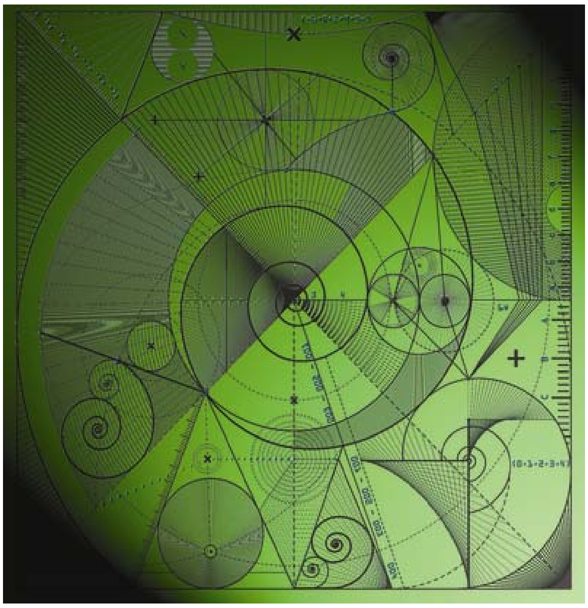
Preprint MPS_2010-32

17 November 2010

Communicability Across Evolving Networks

by

Peter Grindrod, Desmond J. Higham, Mark C. Parsons and Ernesto Estrada



Communicability Across Evolving Networks*

Peter Grindrod† Desmond J. Higham‡ Mark C. Parsons§
Ernesto Estrada¶

Abstract

Many natural and technological applications generate time ordered sequences of networks, defined over a fixed set of nodes; for example time-stamped information about 'who phoned who' or 'who came into contact with who' arise naturally in studies of communication and the spread of disease. Concepts and algorithms for static networks do not immediately carry through to this dynamic setting. For example, suppose A and B interact in the morning, and then B and C interact in the afternoon. Information, or disease, may then pass from A to C, but not vice versa. This subtlety is lost if we simply summarize using the daily aggregate network given by the chain A-B-C. However, using a natural definition of a walk on an evolving network, we show that classic centrality measures from the static setting can be extended in a computationally convenient manner. In particular, communicability indices can be computed to summarize the ability of each node to broadcast and receive information. The computations involve basic operations in linear algebra, and the asymmetry caused by time's arrow is captured naturally through the non-commutativity of matrix-matrix multiplication. Illustrative examples are given for both synthetic and real-world communication data sets. We also discuss the use of the new centrality measures for real-time monitoring and prediction.

---

*This manuscript appears as University of Strathclyde Mathematics Research Report 32 (2010).

†Department of Mathematics, University of Reading, UK

‡Department of Mathematics and Statistics, University of Strathclyde, UK

§Department of Mathematics, University of Reading, UK

¶Department of Mathematics and Statistics & Department of Physics, University of Strathclyde, UK

PACS: 89.75.Fb, 89.75.Hc, 84.40.Ua

1 Motivation

At the heart of network science are the well established mathematical fields of deterministic and random graph theory, with concepts such as connectedness, pathlength, diameter, degree and clique playing key roles [8, 21]. The motivation for this work is that a new type of time-dependent network-based object is emerging from a range of digital technologies that requires a fundamentally different way of thinking.

In Figure 1 we show a simple example of an evolving network, where undirected connections between a fixed set of seven nodes is recorded over three days. If we regard the links as representing communication, for example, by telephone or email, then we see that A may pass a message to C through the links A $\leftrightarrow$ B and B $\leftrightarrow$ G on day 1 and then through the links G $\leftrightarrow$ E and E $\leftrightarrow$ C on day 2. However, there is no way for C to pass a message to A. Analogously, if the links represent physical proximity, then A may pass an infection to C but C cannot cause A to be infected. This asymmetry, which arises even though each individual network is symmetric, is caused by the arrow of time. It is clear that simply aggregating the individual networks would present a very misleading summary. This highlights a fundamental gap between the static and dynamic cases, and points out the need for a theory of evolving networks that

1. deals with the time ordering inherent in the edge lists when considering communication around the network,

2. respects the inherent asymmetry imposed by the arrow of time, even when each individual snapshot consists of an undirected network.

Many application areas give rise to connectivity patterns that change over time in this manner. As well as the traditional context of individual-to-individual contacts in epidemiology [16], the digital revolution is generating novel large scale examples, including

- networks of mobile users with a link denoting current "interaction", i.e., either copresence in a location or logged contact through their mobile devices [25],

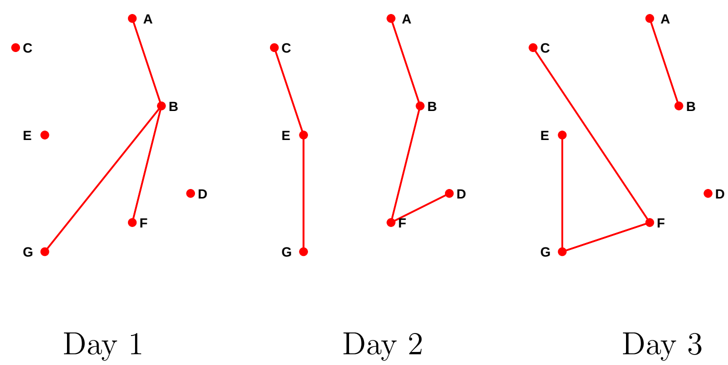


Figure 1: Simple example of an evolving network.

- networks of online social users (e.g., Facebook) interacting through messaging [25] or online chatting systems (e.g. MSN) [19],
- networks of travellers, vehicles or available routes defined over a dynamic transportation infrastructure [14, 18],
- networks describing transient social interactions over cyberspace [23],
- networks describing individuals' attendance at regularly scheduled events over time [1],
- correlated neural activity in response to a functional task [15].

In this work, we show how centrality concepts that have proved useful for determining important nodes in static networks can be extended to this dynamic setting. Our approach is related to that of [22, 23, 24], in the sense that static graph concepts are directly generalized in a manner that respects the time dependency, but we take a walk counting viewpoint and focus on the type of centrality measures that are popular for social networks [17]. We also note that dynamic networks are treated in [20], but the emphasis in that work is to discover communities, and the algorithm does not fully respect the time ordering of the data—for example, backward and forward running clocks would be treated similarly.

Let us emphasize at this stage that unlike in the well-studied 'network growth' context, where new nodes and accompanying edges are accumulated and only the final, aggregate network is of interest [2, 21], we are concerned here with a different time-dependent scenario where the population of nodes remains fixed from the outset, and the graph evolves through the appearance (birth) or the deletion (death) of edges.

2 Dynamic Centralities

To formalize our ideas, given a set of N nodes we consider an ordered sequence $\{G^{[k]}\}$ for $k = 0, 1, 2, \ldots, M$, where each $G^{[k]}$ is an unweighted graph defined over those nodes. We think of a corresponding ordered sequence of time points $t_0 \leq t_1 \leq \cdots \leq t_M$, so that $G^{[k]}$ records the state of the network at time t_k. Each graph may be represented by an N-by-N binary adjacency matrix, $A^{[k]}$, where a nonzero i, j entry records the presence of a link from node i to node j. We allow for $A^{[k]}_{ij} \neq A^{[k]}_{ji}$, so that the adjacency matrix may be unsymmetric.

To address the question of how well information can be passed between pairs of nodes, we generalize the static graph concept of a walk as follows.

Definition 1 *A* dynamic walk of length w *from node* i_1 *to node* i_{w+1} *consists of a sequence of edges* $i_1 \to i_2, i_2 \to i_3, \ldots, i_w \to i_{w+1}$ *and a non-decreasing sequence of times* $t_{r_1} \leq t_{r_2} \leq \ldots \leq t_{r_w}$ *such that* $A^{[r_m]}_{i_m, i_{m+1}} \neq 0$. *We also define the* lifetime *of this walk to be* $t_{r_w} - t_{r_1}$.

We note that an analogous definition of a *dynamic path* can be made by insisting that no node is visited more than once—that concept was developed recently in [25]. In this work we focus on walks, rather than paths, on the grounds that (a) information does not necessarily flow along geodesics [3, 4, 13] and (b) walk counting is more tolerant of errors (missing and spurious edges) than path counting. The walk viewpoint is also in line with the influential work of Katz [17] for the study of static, undirected social networks. The explicit use of a walk based measure of centrality was proposed for the static case in [9], and the idea has been shown to lead to very powerful measures that are useful across a range of application areas [6, 10, 11]. A further benefit of the walk counting approach is that the combinatorics can be conveniently described and implemented in terms of operations in linear algebra, and we will show that this feature can be carried through to the dynamic case.

We emphasize that the sequence of times $t_{r_1}, t_{r_2}, \ldots, t_{r_w}$ in Definition 1 must be nondecreasing, in order to respect the arrow of time, but

- repeated times are allowed: for example, if $r_1 < r_2 = r_3 < r_4$ then precisely two edges are followed at time t_{r_2},
- times are not required to be consecutive: for example, if $r_2 > r_1 + 1$ then the networks corresponding to times in between t_{r_1} and t_{r_2} have not been used during the walk.

Of course, depending on the application area, it may be reasonable to alter these features; forcing at most one edge per time level and/or forcing time levels to be consecutive. The ideas presented here could be adjusted accordingly.

Our key observation, which generalizes a simple result from graph theory (see, for example, [10, Lemma 1.1]) is that the matrix product $A^{[r_1]}A^{[r_2]}\cdots A^{[r_w]}$ has i, j element that counts the number of dynamic walks of length w from node i to node j on which the mth step of the walk takes place at time t_{r_m}.

Now, suppose that we wish to quantify the propensity for node i to communicate, or interact, with node j. For each length $w = 1, 2, \ldots$, we may count the number of dynamic walks from i to j, and this information may then be combined into a single, cumulative total over all w. Allowing for the fact that shorter walks are generally more important (since, for example, the noise or cost of a transmission may increase with length), it makes sense to scale the counts according to the walk length. A particularly attractive choice is to downweight walks of length w by a factor a^w, where $0 < a < 1$. Using our matrix multiplication setting, this leads to the task of summing all products of the form

$$a^w A^{[r_1]}A^{[r_2]}\cdots A^{[r_w]}, \quad \text{where } r_1 \le r_2 \le \cdots \le r_w. \tag{1}$$

Letting $I \in \mathbb{R}^{N\times N}$ denote the identity matrix, and noting that the resolvent $(I - aA)^{-1}$ has the expansion $I + aA + a^2A^2 + \cdots$, these arguments motivate the matrix product

$$\mathcal{Q} := \left(I - aA^{[0]}\right)^{-1}\left(I - aA^{[1]}\right)^{-1}\cdots\left(I - aA^{[M]}\right)^{-1}. \tag{2}$$

The use of the identity matrices in (2) is crucial in our target case of large, sparse networks—it allows a message to 'wait' at a node until a suitable connection appears at a later time.

Overall, as required, the matrix $\mathcal{Q}$ records the sum of all terms of the form (1). We may therefore use $\mathcal{Q}_{ij}$ as our summary of how well information can be passed from node i to node j. The nth row and column sums

$$C_n^{\text{broadcast}} := \sum_{k=1}^{N} \mathcal{Q}_{nk} \quad \text{and} \quad C_n^{\text{receive}} := \sum_{k=1}^{N} \mathcal{Q}_{kn} \tag{3}$$

are centrality measures that quantify how effectively node n can *broadcast* and *receive* messages, respectively[1].

Because we are interested in the relative values of the centrality measures across all nodes, rather than their absolute sizes, it makes sense to avoid under or overflow in the computation of $\mathcal{Q}$ using an iteration such as

$$\mathcal{Q}^{k+1} = \frac{\mathcal{Q}^k \left(I - aA^{[k+1]}\right)^{-1}}{\| \mathcal{Q}^k \left(I - aA^{[k+1]}\right)^{-1} \|},$$

where $\| \cdot \|$ denotes any convenient matrix norm. In our computations we use the Euclidean norm.

These new centralities are a natural generalization of the type of measure developed in [17], which has widely influenced the study of static social networks, [3, 4]—$C_n^{\text{broadcast}}$ and C_n^{receive} reduce to Katz's centrality measure when there is a single time point with undirected edges.

Two features of this new approach are immediately apparent.

- The basic computational tasks are linear system solves, which are convenient and efficient for large, sparse networks.

- The inherent asymmetry caused by the dynamics is captured directly through the non-commutativity of matrix multiplication.

To help understand the role of the downweighting parameter, a, we first note that for a fixed collection of network data, in the limit $a \to 0$ the

[1] An alternative is to specify $k \neq n$ in the summations, so that closed walks are not included. However, such closed walks can play an important role as indicators of centrality [12].

centrality measures reduce to multiples of the aggregate out and in degrees, shifted by unity;

$$\lim_{a\to 0^+} \frac{C_n^{\text{broadcast}}-1}{a} = \sum_{k=1}^{N}\left(\sum_{p=0}^{M} A^{[p]}\right)_{nk},$$
$$\lim_{a\to 0^+} \frac{C_n^{\text{receive}}-1}{a} = \sum_{k=1}^{N}\left(\sum_{p=0}^{M} A^{[p]}\right)_{kn}.$$

At the other extreme, to guarantee that each resolvent $\left(I - aA^{[s]}\right)^{-1}$ in (2) exists, we require that

$$a < \frac{1}{\rho(A^{[s]})}, \quad \text{for all } s,$$

where $\rho(\cdot)$ denotes the spectral radius; that is, the largest eigenvalue in modulus. Furthermore, choosing a close to $\max_s \rho(A^{[s]})$ will cause the corresponding time t_s to dominate the overall communicability matrix $\mathcal{Q}$. In practice, a suitable choice of a would be sufficiently below $1/\max_s \rho(A^{[s]})$ that the results are not sensitive to small changes in a and sufficiently above zero that they do not collapse to the shifted aggregate out and in degrees.

Let us consider how the asymmetry of $\mathcal{Q}$ arises, and hence that between the centrality measures, in the case of an evolving undirected graph. Here, all the adjacency matrices are symmetric. For any N-by$-N$ matrix, B, we define $\mathcal{S}(B) := \frac{1}{2}(B + B^T)$ and $\mathcal{AS}(B) := \frac{1}{2}(B - B^T)$ to be the projections of B onto the space of symmetric matrices and the orthogonal space of anti-symmetric matrices, respectively. The anti-symmetric part of $\mathcal{Q}$ governs the differences between the column and row sums of $\mathcal{Q}$, since

$$2\mathcal{AS}(\mathcal{Q}).\mathbf{1} = \mathbf{C}^{\text{broadcast}} - \mathbf{C}^{\text{receive}},$$

where $\mathbf{1} = (1, 1, ..., 1)^T$, and $\mathbf{C}^{\text{broadcast}} = (C_1^{\text{broadcast}}, ..., C_N^{\text{broadcast}})^T$ are N-vectors, with $\mathbf{C}^{\text{receive}}$ defined analogously.

Working with the non-normalized version of $\mathcal{Q}$ in (2), we have

$$\mathcal{Q} = I + a\sum_{p=0}^{M} A^{[p]} + a^2 \sum_{p=0}^{M}\sum_{p'=p}^{M} A^{[p]}A^{[p']} + O(a^3).$$

It follows that

$$\mathcal{S}(\mathcal{Q}) = I + a \sum_{p=0}^{M} A^{[p]} + O(a^2), \tag{4}$$

and

$$2\mathcal{A}\mathcal{S}(\mathcal{Q}) = a^2 \sum_{p=0}^{M} \sum_{p'=p+1}^{M} [A^{[p]}, A^{[p']}] + O(a^3), \tag{5}$$

where $[A, B] := AB - BA$ denotes the commutator of matrices A and B. Since each separate graph is undirected, this shows that the leading anti-symmetric terms arise only from interactions over distinct pair of time steps.

We also mention that this work focuses on the case of a fixed set of data. In some applications, $A^{[k]}$ will itself be an aggregate of activity over a time window; for example, in sections 3.2 and 3.3 we consider daily telephone and email communication. If we reduce the time window down to hours, minutes, seconds, ..., then, eventually, the communicability matrix $\mathcal{Q}$ would not change—when there is at most one link per time period then, intuitively, no further walks can be created through refinement and in the linear algebra setting of (2) no new non-indentity factors will arise. In this ultra-high-frequency regime, if the edges are undirected then it would be natural to replace (2) with the "at most one link per time window" version

$$\mathcal{Q} := \left(I + aA^{[0]}\right)\left(I + aA^{[1]}\right)\cdots\left(I + aA^{[M]}\right),$$

so that simple, inter-window closed walks such as $i \mapsto j \mapsto i$ are avoided. However, the idea that the finest time level gives the most accurate picture must be treated with caution—in the email context the order in which messages are read or acted upon does not necessarily reflect the order of arrival.

3 Computational Tests

3.1 Synthetic Data

Figure 2 shows a proof-of-principle test of the ideas behind (2). Here we used $N = 1001$ nodes and simulated networks at 31 time points; that is, a month of daily data. At each time point, for nodes 1 to 1000 we constructed, independently, a classical, undirected, Erdos/Renyi random graph—each was chosen uniformly from the collection of all graphs with 1000 nodes and 1000

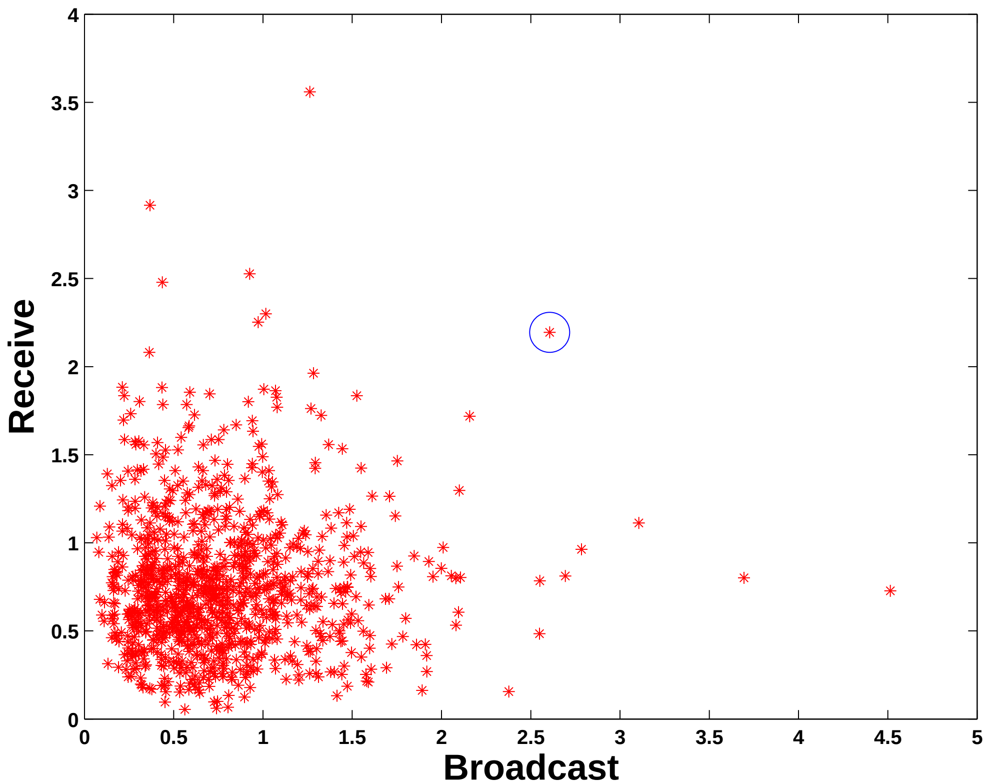


Figure 2: Scatter plot of broadcast and receive centralities (3) for a synthetically generated network of 1000 nodes, where node 1001 (circled) is designed to have average activity, as measured by the aggregate degree, but enjoys high quality connections.

edges. Then at each time the final node 1001 was connected to the two nodes with largest degree. In this way, node number 1001 is distinguished only by the time-sensitive 'quality' of its links—at each time t_k it has a degree that matches that of the average node, and it will never be among the highest degree nodes at any time; so any static or aggregative measure is likely to fail to indentify this node as being special. In the figure we have scattered the (normalized) broadcast and receive centralities (3) for each node, with node 1001 identifed by a circle. We see that the new measures correctly identify the fact that this node can communicate well, despite never enjoying a high degree.

3.2 Telecommunication Data

We now consider telecommunication data from [7]. We have daily "who phoned who" information between 106 individuals based at M.I.T. over 365 days, with starting date 20th July 2004. Because phone conversations are bi-directional, we have symmetrized the data, so $A_{ij}^{[k]} = 1$ if individuals i and j had at least one interaction on day k. Figure 3 shows a summary of the adjacency matrices aggregated into 28 day intervals (day 365 omitted). We notice a decrease in activity outside the traditional academic teaching periods.

The upper left picture in Figure 4 shows the daily edge count. For this data the maximum spectral radius is $\max_s \rho(A^{[s]}) = 8.23$, giving an upper limit of 0.12 for a, and the figure shows centrality results for $a = 0.1$. In the upper right picture we scatter plot on a log-log scale the broadcast and receive centralities (3). Here, and in all other scatter plots, the correlation coefficient for a pair of raw (not log transformed) centralities is quoted to two decimal places in the figure caption. We see that even though the individual adjacency matrices are symmetric, there is no strong correlation between the two centralities. The lower pictures scatter plot the broadcast and receive centralities for each node against the total degree; that is, the sum of the node's degrees over all days. This makes it clear that the new centralities are not simply repeating the degree information. The figure captions also quantify the overlap between the sets of nodes ranked among the top twenty. In Figure 4, the top twenty nodes ordered from twentieth place to first place in terms of the three measures are

$$\begin{aligned}
\text{broadcast} &: 27, 32, 38, 44, 47, 7, 45, 6, 2, 4, 10, 3, 30, 49, 26, 1, 46, 8, 5, 102 \\
\text{receive} &: 94, 58, 76, 95, 15, 20, 12, 89, 93, 30, 19, 49, 6, 35, 39, 52, 42, 8, 13, 53, \\
\text{totaldegree} &: 21, 9, 93, 100, 32, 10, 57, 22, 49, 25, 53, 23, 6, 40, 20, 3, 2, 4, 8, 5.
\end{aligned}$$

In this case the overlaps between broadcast & receive, broadcast & total degree and receive & total degree contain 4, 9 and 6 nodes, respectively. Only one node appears in all three top twenty lists.

Figure 5 examines the sensitivity of the results to the parameter a. The upper pictures show how the centralities change from $a = 0.1$ to $a = 0.05$. The top twenty broadcast lists have 14 nodes in common and for the receiver lists the overlap is 16. The lower pictures show the change from $a = 0.05$ to $a = 0.01$, and in this case the top twenty overlap counts are 16 and 11.

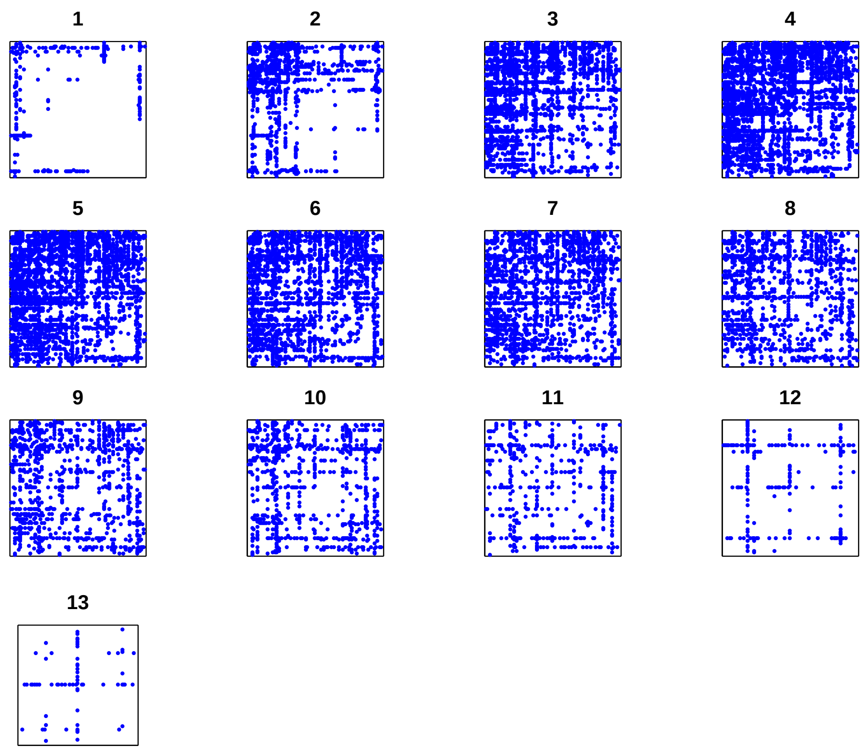


Figure 3: Adjacency matrices for the M.I.T. telecommunication data, symmetrized and aggregated into 13 sets of 28 day windows.

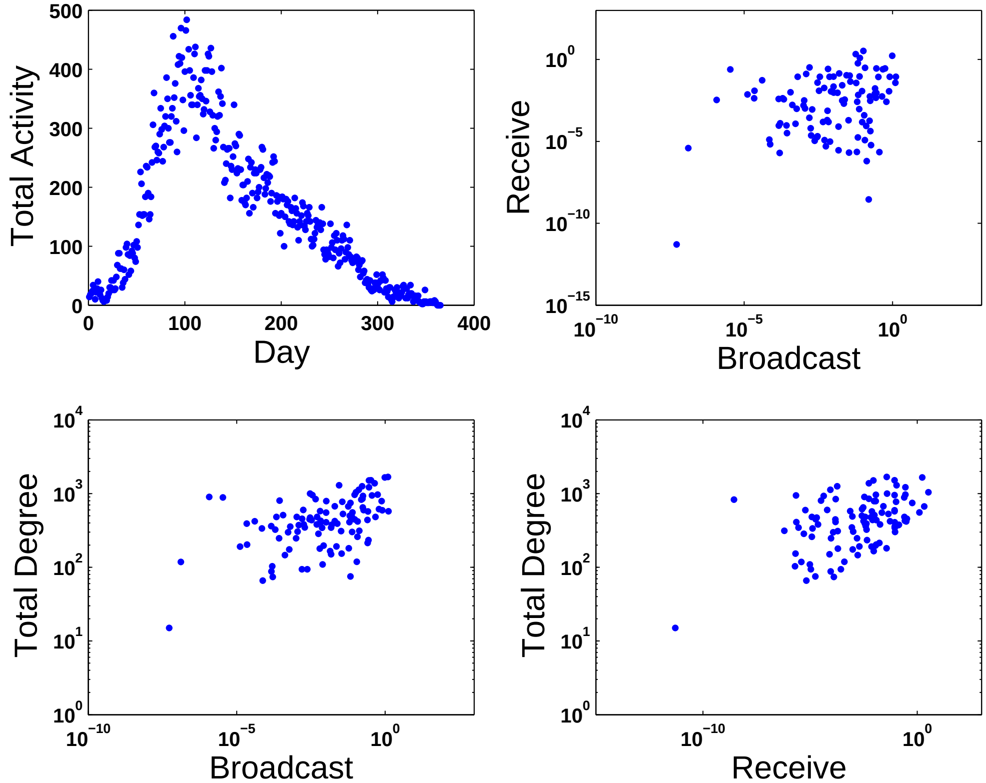


Figure 4: Daily M.I.T. telecommunication data. Upper right: total activity per day. Upper left: Broadcast versus receive centrality; correlation 0.14, top twenty overlap size 4. Lower left: broadcast centrality versus total degree; correlation 0.50, top twenty overlap size 9. Lower right: Receive centrality versus total degree; correlation 0.28, top twenty overlap size 6.

Overall, the experiments indicate that the two new measures deliver distinct information that is different from a raw degree count, and remains consistent over a range of a values.

3.3 Email Data

We now consider a public domain data set concerning email activities of Enron employees. In [5] the static, aggregate network was analysed, but here we treat it as an evolving network. We constructed daily information representing emails between 151 Enron employees, including `to`, `cc` or `bcc`. So $A_{ij}^{[k]} = 1$ if employee i sent at least one message to employee j on day k, but because this type of communication is unidirectional, we do not automatically add the $j \mapsto i$ link. We have data over 1138 days, starting on 11th May 1999. Many of the adjacency matrices are empty, stressing the importance of the identity matrices in (2) for analysing sparse data. The upper left plot in Figure 6 shows the daily edge count.

Using $a = 0.2$, in the upper right of Figure 6 we scatter plot broadcast versus receive centralities, and in the lower plots we show broadcast versus total out degree and receive versus total in degree. In this case the maximum spectral radius is $\max_s \rho(A^{[s]}) = 4.17$, giving an upper limit of $a = 0.24$. As in the previous test, we see that the two new centrality measures are distinct; in particular, only two nodes appear in the overlap of top twenty broadcast and receive and it is clear that some top receivers are very poor broadcasters. The top twenty overlap between broadcast and total out degree is 11 and between receive and total in degree in 6, showing that the new measures do not simply reflect aggregate connectivity.

The upper plots of Figure 7 show how the new centralities change when a is reduced from 0.2 to 0.1, indicating robustness in this parameter regime. The lower plots show the effect of symmetrizing the data, so that $j \mapsto i$ whenever $i \mapsto j$, in the case $a = 0.1$. We then have $\max_s \rho(A^{[s]}) = 8.57$, giving a an upper limit of 0.12. We see that the new dynamic centralities are relatively insensitive to this transformation of the data, suggesting that the dominant asymmetry is caused by time.

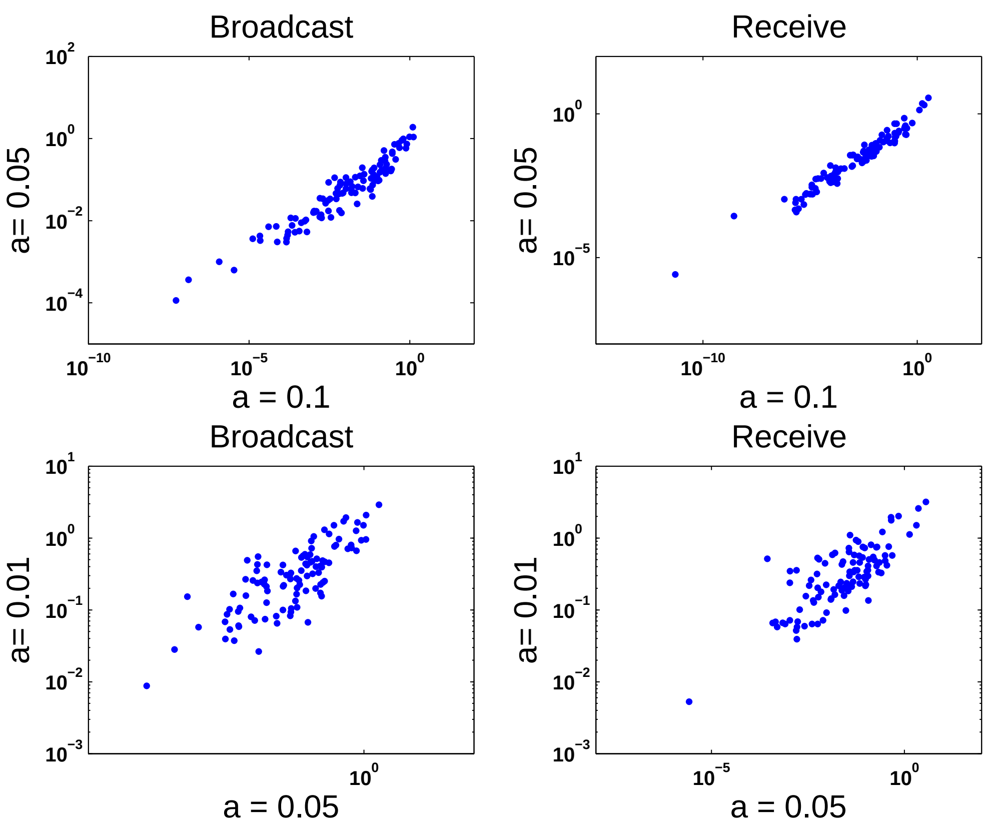


Figure 5: Daily M.I.T. telecommunication data. Upper: $a = 0.1$ versus $a = 0.05$; correlations are 0.93 for broadcast and 0.98 for receive, respective top twenty overlap sizes are 14 and 16. Lower: $a = 0.05$ versus $a = 0.01$; correlations are 0.82 for broadcast and 0.81 for receive, respective top twenty overlap sizes are 16 and 11.

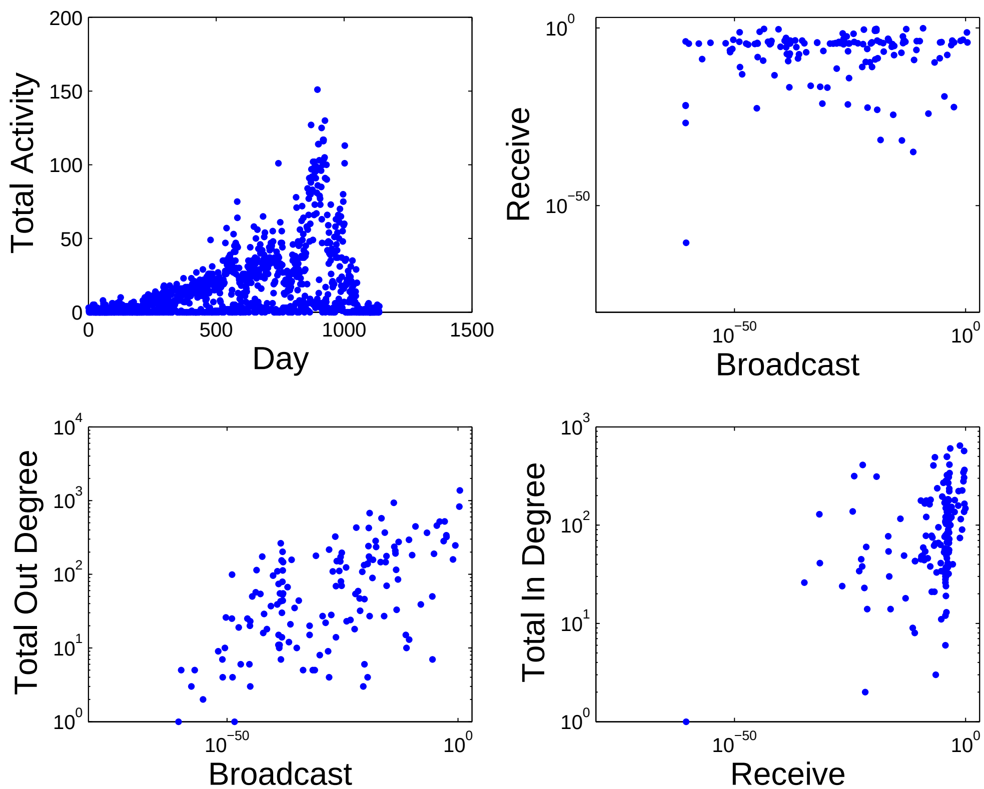


Figure 6: Results for Enron email data. Upper left: total number of edges per day. Upper right: Scatter plot of broadcast and receive centralities; correlation 0.00, top twenty overlap size 2. Lower left: Scatter plot of broadcast centrality and total out degree; correlation 0.62, top twenty overlap size 11 Lower right: Scatter plot of receive centrality and total in degree; correlation 0.28, top twenty overlap size 6.

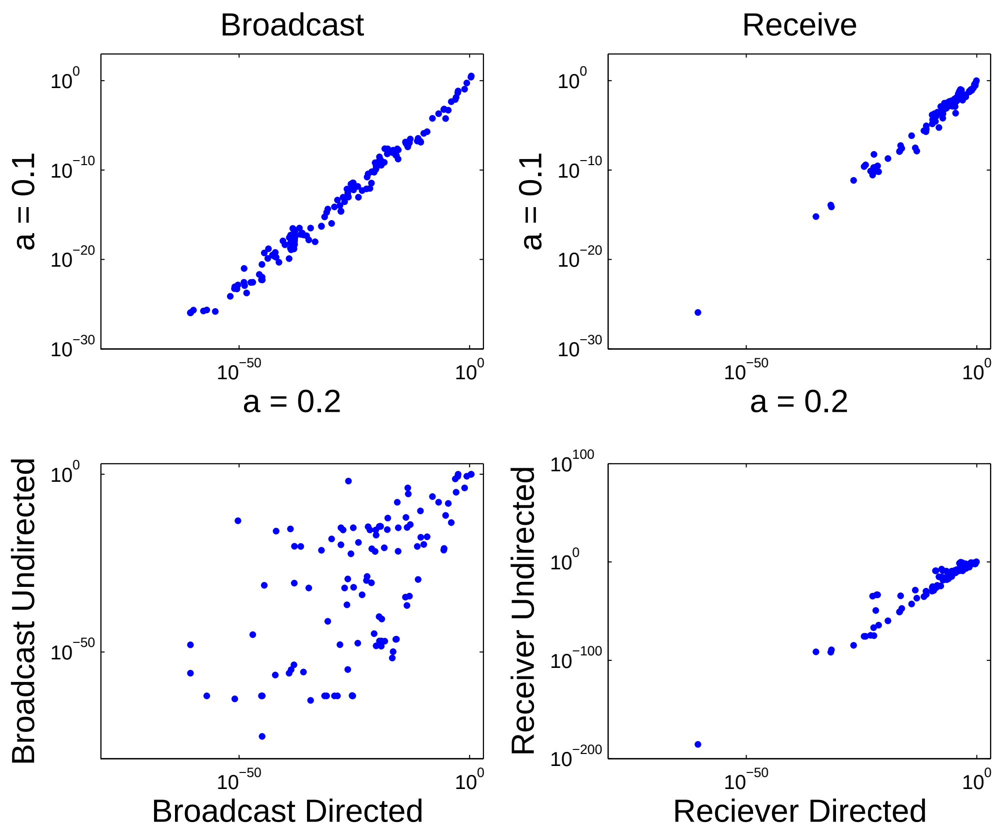


Figure 7: Results for Enron email data. Upper left: Scatter plot of broadcast centralities for $a = 0.2$ and $a = 0.1$; correlation 1.00, top twenty overlap size 18. Upper right: Scatter plot of receive centralities for $a = 0.2$ and $a = 0.1$; correlation 0.97, top twenty overlap size 14. Lower left: Scatter plot of broadcast centralities for directed and undirected networks for $a = 0.1$; correlation 0.82, top twenty overlap size 12. Lower right: Scatter plot of receive centralities for directed and undirected networks for $a = 0.1$; correlation 0.76, top twenty overlap size 13.

4 Discussion

The new centrality measures introduced here can be computed at any point in time, and hence they may be used to monitor network behaviour dynamically. A practical problem with evolving networks is that of an observer who may be able make some kind of intervention; for example by injecting some information (marketing content, rumours, propaganda, misinformation) at key nodes at some instant, or by isolating or even removing a node. This raises the issue of predicting future network behavior. We will briefly discuss an approach based on the observer's expectation of the future communicability: an estimate of $\mathcal{Q}$ going forwards.

Suppose we have a stochastic model for the evolution of the network based on historical data and some specific knowledge. More precisely, suppose we have $P(A^{[p+1]}|H_p)$, the conditional distribution for the adjacency matrix at the next time step given its entire history up to and including step p, so $H_p = \{A^{[p]}, A^{[p-1]}, A^{[p-2]}, \ldots\}$. Then applying this model iteratively we obtain the conditional distribution for $A^{[p']}$ for any $p' > p$; that is, $P(A^{[p']}|H_p)$.

Let us write $E(A^{[p']}|H_p)$ to denote the corresponding expected value of the future adjacency matrix, given H_p.

Now suppose we have observed the network up to and including some time step, say $p = 0$ for convenience. Then from (4) and (5) we can calculate estimates for the expectation of the communicability over the current and future time steps. We have the small a approximations

$$E(\mathcal{S}(\mathcal{Q})|H_0) = I + a\sum_{p=0}^{M} E(A^{[p]}|H_0) + O(a^2),$$

and

$$2E(\mathcal{A}\mathcal{S}(\mathcal{Q})|H_0) = a^2 \sum_{p=0}^{M} \sum_{p'=p+1}^{M} E([A^{[p]}, A^{[p']}])|H_0) + O(a^3).$$

These estimates may be accessible in practice, depending on the complexity and memory dependence of the model. For example, suppose we make the dramatically simplifying assumption that our model is a symmetric, edge independent Markov process. Letting α_{ij} and ω_{ij} denote the stepwise birth and death rates for the evolution of the (i,j)th edge, we have $A^{[p]} \to A^{[\infty]}$ as $p \to \infty$, where $A^{[\infty]}_{ij} = \alpha_{ij}/(\alpha_{ij} + \omega_{ij})$. In this Markovian case we can also replace the history, H_p, by the single previous step $A^{[p]}$. Then considering time steps 0 up to M we have

$$E(\mathcal{Q}|A^{[0]}) = I + a\left(R_M \circ (A^{[0]} - A^{[\infty]}) + (M+1)A^{[\infty]}\right) + O(a^2),$$

where R_M is the symmetric matrix given by $(R_p)_{ij} = (1-(1-\alpha_{ij}+\omega_{ij})^{M+1})/(\alpha_{ij}+\omega_{ij})$, and $\circ$ denotes componentwise multiplication. This quantifies the relative contributions to $\mathcal{Q}$ made by the initial condition and the long term expected *equilibrium* value for each edge. So if the observer wishes to intervene based on the dominance of some large row or column sums of $\mathcal{Q}$, we can see that this may require such action sooner or later depending on the current state of the network and the longer term expectation.

So, overall, we believe that the new class of walk-based centrality measures introduced here offers great potential as a computationally and analytically attractive means to treat time-stamped network sequences, both for summarizing existing data sets and real-time actioning.

Acknowldegement This work was supported by the Engineering and Physical Sciences Research Council and the Research Councils UK Digital Economy programme.

References

[1] E. Acar, D. M. Dunlavy, and T. G. Kolda, *Link prediction on evolving data using matrix and tensor factorizations*, in LDMTA'09: Proceeding of the ICDM'09 Workshop on Large Scale Data Mining Theory and Applications, IEEE Computer Society Press, December 2009, pp. 262–269.

[2] A.-L. Barabási and R. Albert, *Emergence of scaling in random networks*, Science, 286 (1999), pp. 509–12.

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

[4] S. P. Borgatti and M. G. Everrett, *A graph theoretic perspective on centrality*, Social Networks, 28 (2006), pp. 466–484.

[5] A. Chapanond, M. Krishnamoorthy, and B. Yener, *Graph theoretic and spectral analysis of Enron email data*, Computational & Mathematical Organization Theory, 11 (2005), pp. 265–281.

[6] J. J. Crofts and D. J. Higham, *A weighted communicability measure applied to complex brain networks*, J. Royal Society Interface, 33 (2009).

[7] N. Eagle, A. S. Pentland, and D. Lazer, *Inferring friendship network structure by using mobile phone data*, Proceedings of the National Academy of Sciences, 106 (2009), pp. 15274–15278.

[8] E. Estrada, M. Fox, D. J. Higham, and G.-L. Oppo, eds., *Network Science: Complexity in Nature and Technology*, Springer, Berlin, to appear, 2010.

[9] E. Estrada and N. Hatano, *Communicability in complex networks*, Phys. Rev. E, 77 (2008), p. 036111.

[10] E. Estrada and D. J. Higham, *Network propeties revealed through matrix functions*, SIAM Review, 52 (2010), pp. 696–671.

[11] E. Estrada, D. J. Higham, and N. Hatano, *Communicability betweenness in complex networks*, Physica A, 388 (2009), pp. 764–774.

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

[13] L. C. Freeman, S. P. Borgatti, and D. R. White, *Centrality in valued graphs: A measure of betweenness based on network flow*, Social Networks, 13 (1991), pp. 141–154.

[14] A. Gautreau, A. Barrat, and M. Barthelemy, *Microdynamics in stationary complex networks*, Proc. Nat. Acad. Sci., 106 (2009), pp. 8847–8852.

[15] P. Grindrod and D. J. Higham, *Evolving graphs: Dynamical models, inverse problems and propagation*, Proceedings of the Royal Society, Series A, 466 (2010), pp. 753–770.

[16] R. R. Kao, *Networks and models with heterogeneous population structure in epidemiology*, in Network Science: Complexity in Nature and Technology, E. Estrada, D. J. Higham, M. Fox, and G.-L. Oppo, eds., Springer, 2010, pp. 51–84.

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

[18] L. McNamara, C. Mascolo and L. Capra, *Media Sharing based on Colocation Prediction in Urban Transport*, in Proc. of ACM 14th International Conference on Mobile Computing and Networking (Mobicom08), San Francisco, CA, September 2008, pp. 58–69.

[19] J. Leskovec and E. Horvitz, *Planetary-scale views on a large instant-messaging network*, in Proceeding of the 17th international conference on World Wide Web, Beijing, China, 2008, ACM, pp. 915–924.

[20] P. J. Mucha, T. Richardson, K. Macon, M. A. Porter, and J.-P. Onnela, *Community structure in time-dependent, multiscale, and multiplex networks*, Science, 328 (2010), pp. 876–878.

[21] M. E. J. Newman, *The structure and function of complex networks*, SIAM Review, 45 (2003), pp. 167–256.

[22] J. Tang, M. Musolesi, C. Mascolo, and V. Latora, *Temporal distance metrics for social network analysis*, in Proceedings of the 2nd ACM SIGCOMM Workshop on Online Social Networks (WOSN09), Barcelona, 2009.

[23] J. Tang, M. Musolesi, C. Mascolo, and V. Latora, *Characterising temporal distance and reachability in mobile and online social networks*, SIGCOMM Comput. Commun. Rev., 40 (2010), pp. 118–124.

[24] J. Tang, M. Musolesi, C. Mascolo, V. Latora, and V. Nicosia, *Analysing information flows and key mediators through temporal centrality metrics*, in SNS '10: Proceedings of the 3rd Workshop on Social Network Systems, New York, NY, USA, 2010, ACM, pp. 1–6.

[25] J. Tang, S. Scellato, M. Musolesi, C. Mascolo, and V. Latora, *Small-world behavior in time-varying graphs*, Physical Review E, 81 (2010), p. 05510.