Bistability through Triadic Closure

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

Abstract. We propose and analyze a class of evolving network models suitable for describing a dynamic topological structure. Applications include telecommunication, on-line social behavior, and information processing in neuroscience. We model the evolving network as a discrete-time Markov chain and study a very general framework in which edges conditioned on the current state appear or disappear independently at the next time step. We show how to exploit symmetries in the microscopic, localized rules in order to obtain conjugate classes of random graphs that simplify analysis and calibration of a model. Further, we develop a mean field theory for describing network evolution. For a simple but realistic scenario incorporating the triadic closure effect that has been empirically observed by social scientists (friends of friends tend to become friends), the mean field theory predicts bistable dynamics, and computational results confirm this prediction. We also discuss the calibration issue for a set of real cellphone data, and find support for a block model in which individuals are assigned to one of two distinct groups having different within-group and across-group dynamics.

1. Motivation

A diverse range of application areas gives rise to large, complex interaction patterns. In the field of network science, classes of random graphs have been proposed and tested as models to capture the structure of these interactions [Newman 10]. In many cases, the model may be viewed as an iterative procedure that builds a network sequentially by randomly rewiring an existing structure [Newman et al. 00, Wagner 03, Watts and Strogatz 98] or by adding nodes and

links in order to "grow" a network [Barabási and Albert 99]. In these cases, the object of interest, however, is the final, static, network. Our work differs in that we wish to model a network structure that is inherently dynamic, with edges appearing and disappearing between a fixed set of nodes. Such a scenario arises naturally in many modern data-rich applications, for example telecommunication (who phoned whom each day), online social interaction (who interacted with whom in a chat room), online retailing (people who bought this book also bought) [Holme and Saramäki 12]. Attention has recently been paid to the issue of extending traditional graph-theoretic concepts such as paths to the time-dependent setting [Berman 96, Grindrod et al. 11, Holme 05, Kossinets et al. 08, Tang et al. 10, Tang et al. 09] and related computational complexity issues [Avin et al. 08, Clementi et al. 08, Clementi et al. 09]. There has also been interest in the waiting times between link changes [Barabási 05, Zhao et al. 11] and the emergence of communities [Bassett et al. 11, Mucha et al. 10]. Also, *link prediction*—estimating likely new connections a short time ahead—is becoming recognized as an important task [Esfandiar et al. 10, Liben-Nowell and Kleinberg 07, Lu et al. 10]. However, in this work we focus on the fundamental issue of modeling and analyzing such networks directly and from first principles, that is, on prescribing reasonable "laws of motion" and studying the potential behaviors that can arise. Our target applications are digitally generated communication or online social interaction networks, and our interest lies in the changes in the connectivity structure itself. Related work on *adaptive networks* [Gross and Blasius 08] has studied scenarios in which the nodes are involved in their own dynamical system that is coupled to the dynamic topology—for example, in an epidemiological SIS model, susceptible (S) nodes may seek to avoid links with nodes that are currently infected (I).

The main novel contributions in our work are (a) the introduction of the concept of conjugate graphs, which can play an important role in understanding and analyzing a model, (b) the derivation of a mean-field-theory approach to summarizing long-term behavior, (c) the introduction of a simple but realistic nonlinear network-evolution model driven by the concept of triadic closure from social science and to have shown that it admits bistable behavior, and (d) a consideration of the issue of model calibration and a demonstration that the use of a block model improves the fit for a voice call data set.

The presentation is organized as follows. In the next section, we introduce a general stochastic framework and show how modeling and simulation can be simplified by focusing on the dynamics of individual edges. We also illustrate these ideas in the particular case that triadic closure is encouraged—friends of friends tend to become friends. Section 3 then introduces the concept of conjugacy, which can be used to describe inherent symmetries in a model—the most

natural example being the case that the dynamics do not depend on the choice of node labeling. A mean field approach is described in Section 4 and applied to the triadic closure model in Section 5. This model, and a more general block version in which the nodes are partitioned into two distinct groups, are fitted to voice call data in Section 6, and conclusions are given in Section 7.

2. Preliminaries

Throughout this work we focus on undirected graphs on n vertices with no loops, which may be represented by symmetric binary $n \times n$ adjacency matrices. Here $A = (a_{ij}) \in \mathbb{R}^{n \times n}$ has $a_{ij} = a_{ji} = 1$ if there is an edge from node i to node j and has $a_{ij} = a_{ji} = 0$ otherwise, with all a_{ii} equal to 0. Let S_n denote the set of all such adjacency matrices and let $\mathbf{1}$ denote the adjacency matrix for the n-vertex clique. If $A \in S_n$, then $\mathbf{1} - A \in S_n$ is the adjacency matrix for the complementary graph to that represented by A. Let R_n denote the set of symmetric $n \times n$ real matrices with all elements taking values in $[0, 1]$, with zeros on the main diagonal. So $S_n \subset R_n$. For real $n \times n$ matrices M_1 and M_2, the Hadamard (or Schur) product, denoted by $M_1 \circ M_2$, is the matrix obtained by elementwise multiplication, so that

$$(M_1 \circ M_2)_{ij} = (M_1)_{ij}(M_2)_{ij}.$$

If $A_1, A_2 \in S_n$, then $A_1 \circ A_2 \in S_n$ represents the adjacency matrix for the graph of common edges.

Following the treatment in [Grindrod and Higham 10], we use the phrase *evolving network model* to describe a stochastic rule that generates a sequence of networks represented by a sequence of adjacency matrices $\{A_k\}_{k=0}^K$, where each A_k belongs to S_n. We say that A_k represents the state of the evolving network at the kth time step, t_k, where $t_0 < t_1 < \cdots < t_K$ are equally spaced points in time. We emphasize that the number of network nodes, n, is fixed in this approach. However, the effect of a node "joining" the network can be treated by, for example, allowing the node to have connections only after an appropriate time point. Similarly, a node may "leave" the network at some point in time if it is guaranteed never to have any future links.

It is natural and analytically convenient to focus on Markovian models. We therefore consider the case of a *first-order evolving network model* characterized by the conditional probability distribution for A_{k+1} given A_k, denoted by $P(A_{k+1} \mid A_k)$, defined for all pairs $A_{k+1}, A_k \in S_n$. The expected value of A_{k+1}

given A_k will be written as

$$\langle A_{k+1} \mid A_k \rangle := \sum_{A_{k+1} \in S_n} A_{k+1} P(A_{k+1} \mid A_k).$$

By construction, $\langle A_{k+1} \mid A_k \rangle \in R_n$, and the (i,j)th element of $\langle A_{k+1} \mid A_k \rangle$ contains the conditional probability that the corresponding edge is present in A_{k+1}.

We will say that the first-order evolving network model $P(A_{k+1} \mid A_k)$ is *edge independent* if given A_k, information about the existence of any particular edges in A_{k+1} has no effect on the probability that any other edge is in A_{k+1}.

Under this assumption of edge independence, every element $W \in R_n$ defines a random graph. The (i,j)th element of W gives the probability that the corresponding edge is present, and we can construct the associated probability distribution over S_n, say $P_W(A)$:

$$P_W(A) = \prod_{i=1, j=i+1}^{n} (W)_{ij}^{(A)_{ij}} \ (1 - (W)_{ij})^{1-(A)_{ij}}.$$

For example, if we have $p \in (0,1)$, then $p\mathbf{1} \in R_n$ represents a classical Erdős–Rényi/Gilbert random graph (usually denoted by $G(n,p)$), where each edge exists with independent probability p [Newman 10]. Let $\mathrm{CL}_{(n,k)} \in S_n$ represent the adjacency matrix for a circular lattice on n vertices, with each vertex connected to its k nearest clockwise and its k nearest counterclockwise neighbors. Then for any choice of constants $0 < q \le p < 1$, $p\mathrm{CL}_{(n,k)}$ is a partial lattice with every edge present with independent probability p, while $p\mathrm{CL}_{(n,k)} + q(\mathbf{1} - \mathrm{CL}_{(n,k)})$ is a partial lattice with uniform "short cuts" in the spirit of the classical Watts–Strogatz small-world model [Newman et al. 00, Watts and Strogatz 98].

By focusing on first-order edge-independent evolving-network models, we may restrict our attention to matrix-valued functions

$$\mathcal{F}: S_n \to R_n. \tag{2.1}$$

Any such mapping $\mathcal{F}$ generates a first-order edge-independent evolving-network model for which

$$\langle A_{k+1} \mid A_k \rangle = \mathcal{F}(A_k). \tag{2.2}$$

A particularly useful form for $\mathcal{F}(A_k)$ is

$$\mathcal{F}(A_k) = (\mathbf{1} - \omega(A_k)) \circ A_k + \alpha(A_k) \circ (\mathbf{1} - A_k), \tag{2.3}$$

where $\alpha(A_k)$ and $\omega(A_k)$ are given mappings $S_n \to R_n$, representing conditional birth rates and death rates respectively, as introduced in [Grindrod and Higham 10]. It is straightforward to compute a path for such a Markov chain, that is, a particular network sequence whose transitions respect the relevant edge

birth and death rates. The following pseudocode summarizes this approach, given an initial adjacency matrix A_0:

for $k = 0, 1, 2, \ldots,$

    compute $\alpha(A_k), \omega(A_k) \in R_n$

    for all disjoint pairs $i \neq j$

      if $(A_k)_{ij} = 0$ then set

        $(A_{k+1})_{ij} = 1$ with probability $\alpha(A_k)_{ij}$ (birth)

        $(A_{k+1})_{ij} = 0$ with probability $1 - \alpha(A_k)_{ij}$ (no change)

      else we have $(A_k)_{ij} = 1$, so set

        $(A_{k+1})_{ij} = 0$ with probability $\omega(A_k)_{ij}$ (death)

        $(A_{k+1})_{ij} = 1$ with probability $1 - \omega(A_k)_{ij}$ (no change)

      end if

    end for all pairs

end for k

In this formulation, $n(n-1)/2$ uniform pseudorandom numbers need to be drawn on each iteration.

We finish this section by introducing a novel example that fits into this framework. We consider the case that the death rate is a constant for all edges,

$$\omega(A_k) \equiv \widetilde{\omega}\mathbf{1}, \quad \text{for } \widetilde{\omega} \in (0, 1), \tag{2.4}$$

while the birth rates for edges not present in A_k are given by

$$\alpha(A_k) = \delta\mathbf{1} + \epsilon\mathbf{1} \circ A_k^2, \tag{2.5}$$

for some constants δ and ϵ. We assume that $0 < \delta \ll 1$, and to guarantee probabilities in the range $[0, 1]$, that $0 < \epsilon(n-2) < 1 - \delta$.

This model reflects a situation whereby the more common the adjacencies between two nonadjacent vertices in A_k, the more likely they are to become adjacent in A_{k+1}. In other words, in the social network case, somebody who is not currently your friend, but who is currently a friend of many of your current friends, has an enhanced chance of becoming your friend at the next step. Forming new associations by the process of triangulating current adjacencies is very natural within a number of applications. In social networks, two peers may be likely to become introduced through common friends; in this context, the
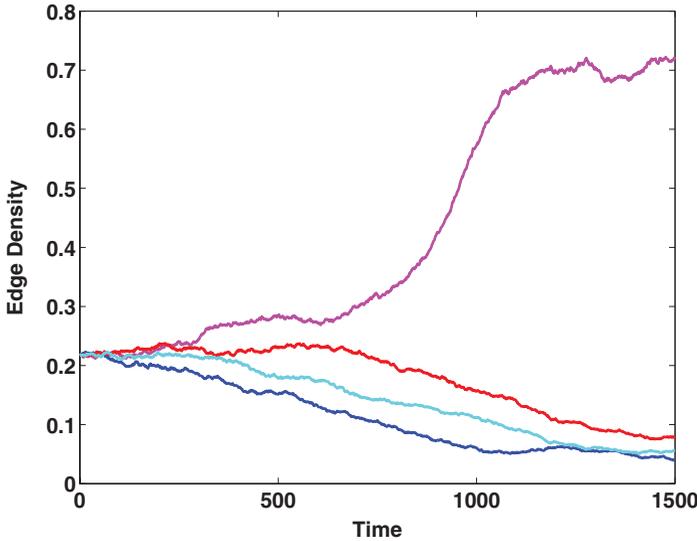
Figure 1. Edge density $\widehat{p}_k$ in (2.7) extracted from four independent paths of the Markov chain using the same model parameters and the same initial condition (color figure available online).

mechanism is often referred to as triadic closure [Goodreau et al. 09, Kossinets and Watts 06, Szell and S. Thurner 10]. In developing cognitive processing capability in the brain, triangulation increases efficiency of communication and resilience, and overabundance of triangles has been observed in both anatomical and functional studies [Bullmore and Sporns 09, He et al. 07, Liu et al. 08].

To create Figure 1, we simulated this model with $n = 100$ nodes and parameter values

$$\widetilde{\omega} = 0.01, \quad \epsilon = 0.0005, \quad \delta = 0.0004. \tag{2.6}$$

The initial network was a sample of an Erdős–Rényi graph with expected edge density $p = 0.3$; each possible edge exists with independent probability p. The four jagged curves correspond to four independent paths of the Markov chain. In each case, they show the evolution of the edge density

$$\widehat{p}_k := \frac{1}{n(n-1)/2}\sum\sum_{i>j}(A_k)_{ij}, \tag{2.7}$$

up to time $k = 1500$.

We see in Figure 1 that with the same model parameters and the same initial condition, the edge density for one path has increased to around 70%, while three

other paths produce an edge density that decreases to around 5%. The analysis that we develop in Sections 3 and 4 will explain this difference in behavior, pointing out for the first time in this context that bistability arises naturally.

3.   Classes of Conjugate Random Graphs

In this section, we show how symmetries in a model give rise to a natural definition of conjugacy. Assume that the mapping $\mathcal{F}$ in (2.1) characterizing a first-order edge-independent evolving network can be extended to a mapping $\mathcal{F} : R_n \to R_n$. We will see in Section 4 how this can be rather natural when we wish to replace A_k by its own expected value. Let λ denote a parameter ranging over a domain Λ within some suitable space. Let $W : \Lambda \to R_n$. We will say that $W(\lambda)$ is a *class of conjugate random graphs* for $\mathcal{F}$ if for each $\lambda \in \Lambda$, there exists a unique element $g(\lambda) \in \Lambda$ such that

$$W(g(\lambda)) = \mathcal{F}(W(\lambda)).$$

Hence the parameterized set $\{W(\lambda) \mid \lambda \in \Lambda\} \subset R_n$ is positively invariant under $\mathcal{F}$. Moreover, the action of $\mathcal{F}$ on R_n can be reduced to the action of g on Λ.

Suppose that $\mathcal{F} : R_n \to R_n$ also possesses some symmetries in the sense that there is a subgroup of $n \times n$ permutation matrices $H = \{Q_r\}$ such that $\mathcal{F}$ is invariant under each of these these permutations. More precisely,

$$\mathcal{F}(A) = Q_r \mathcal{F}(Q_r^T A Q_r) Q_r^T, \quad A \in S_n, \quad Q_r \in H.$$

Then $\mathcal{F}$ uses no extra a priori information about the vertices that distinguishes one permutation in H from another. Now let $\mathcal{W} \subset R_n$ denote the subset of random graphs in R_n that are invariant under the symmetries in H, so that

$$Q_r^T W Q_r = W,$$

for all $W \in \mathcal{W}$, $Q_r \in H$. It then follows for $W \in \mathcal{F}$ that

$$Q_r^T \mathcal{F}(W) Q_r = \mathcal{F}(Q_r^T W Q_r) = \mathcal{F}(W).$$

Hence $\mathcal{F}$ maps $\mathcal{W}$ into $\mathcal{W}$. So under a suitable parameterization, with some $\lambda \in \Lambda$, the subset $\mathcal{W}$ is a possible class of conjugate random graphs for $\mathcal{F}$.

In the simple case in which $\mathcal{F}$ is invariant under all possible permutations on n vertices, we know that no subset of nodes is distinguished in any way. For example, this will certainly be so whenever

$$\mathcal{F}(W) = Q_1(W) \circ Q_2(W) \circ \cdots \circ Q_S(W),$$

where each of the $Q_s(W) : R_n \to R_n$ is a polynomial in W. The random graph $W = p\mathbf{1}$ is invariant under all such permutations; $\mathcal{W}$ is the set of such graphs, parameterized by $p \in [0,1]$. This situation represents an egalitarian scenario whereby every node is the same and none is distinguished based on any a priori information. So we treat all edge dynamics according to the same model.

Alternatively, we may have a *block model* in which prior considerations allow us to divide the nodes onto two disjoint subsets. (The generalization to a finer partition is straightforward.) This could arise in a social network between individuals from two sexes for instance, where it was posited that male–male, female–male and female–female interactions follow different birth/death dynamics. Similarly, there may be an elite and nonelite (officers and troops), or cultural, or functional splitting of the vertices. To be concrete, suppose that all edges between all of the first n_1 vertices satisfy a given identical dynamic; all edges between all of the last $n_2 = n - n_1$ vertices satisfy a distinct given identical dynamic; and all edges between all of the first n_1 vertices with all of the last n_2 vertices satisfy a third given identical dynamic. Then $\mathcal{F}$ will be invariant under all permutations Q that permute within the two subsets separately, and do not swap any vertices between them. In that case, W is invariant with respect to all such permutations if it has a symmetric block structure, say

$$(W)_{ij} = \begin{cases} 0 & \text{if } i = j, \\ p & \text{if } i \neq j \text{ and } i,j \leq n_1, \\ q & \text{if } i \neq j \text{ and } i,j > n_1, \\ r & \text{if } \min\{i,j\} \leq n_1 \text{ and } \max\{i,j\} > n_1. \end{cases}$$

Then $\mathcal{W}$ is the set of such graphs, parameterized by three constants: $\lambda = (p,q,r) \in [0,1]^3 = \Lambda$.

For such a block random graph $W \in \mathcal{W}$, the expected number of edges is given by

$$\frac{n_1(n_1 - 1)}{2}p + \frac{n_2(n_2 - 1)}{2}q + n_1 n_2 r.$$

We also recall that the Watts–Strogatz clustering coefficient for a node is defined as the ratio of links between the vertices within its neighborhood to the number of links that could possibly exist between them [Newman 10, Watts and Strogatz 98]. In other words, for node B, it is the frequency with which a path $A - B - C$ is also involved in the closed loop $A - B - C - A$. The expected

Watts–Strogatz clustering coefficient then has the form of the quotient

$$
\left(\frac{n_1!}{2!(n_1-3)!}p^3 + \frac{n_2!}{2!(n_2-3)!}q^3 + \frac{n_1 n_2!}{2!(n_2-2)!}r^2 q + \frac{n_2 n_1!}{2!(n_1-2)!}r^2 p \right.
$$
$$
\left. + n_1(n_1-1)n_2 pr^2 + n_2(n_2-1)n_1 qr^2 \right) \Big/
$$
$$
\left(\frac{n_1!}{2!(n_1-3)!}p^2 + \frac{n_2!}{2!(n_2-3)!}q^2 + \frac{n_1 n_2!}{2!(n_2-2)!}r^2 + \frac{n_2 n_1!}{2!(n_1-2)!}r^2 \right.
$$
$$
\left. + n_1(n_1-1)n_2 pr + n_2(n_2-1)n_1 qr \right).
$$

This expression has six terms in the denominator and six in the numerator, representing respectively the expected number of open jaws (pairs of edges from a *central* vertex to two other distinct vertices) and the expected number of those open jaws that are complete triangles. Six terms arise because the *central* vertex (of the open jaw) may be in either of the subsets, while the other two vertices may be such that none, one, or both lie within the same subset.

In Section 6, we return to the concept of clustering coefficients as a means to compare standard and block models on voice call data. The next two sections focus on mean field theory.

4.   A Mean Field Approximation for Evolving Networks

Our aim in this section is to develop a heuristic approach for analyzing the behavior of an edge-independent first-order evolving network, as defined in (2.2). We begin with the simplifying assumption that A_k can be adequately summarized by its own expected graph, $\langle A_k \mid X \rangle$, given any prior information X, rather than including all of the details of any particular value for A_k. Then we may use this expected value to calculate the consequent expectation, $\langle A_{k+1} \mid \langle A_k \mid X \rangle \rangle$.

In reality, of course, each of the edges in A_k is there or not, taking a binary value. Hence by employing the expectation for A_k, the analysis is only an approximation. It will be particularly poor at times when the probability of edges appearing is relatively small and/or the existence and distribution of just a few edges has a critical effect: they cannot in reality be *smeared out*.

Alternatively, we may be given only $\langle A_k \mid X \rangle \in R_n$ as a random graph itself, conditional on any previous information X, and we may wish to use our evolving graph model to calculate an estimate for $\langle A_{k+1} \mid X \rangle$, and so on. We should calculate

$$
\langle A_{k+1} \mid X \rangle = \sum_{A_k \in S_n} \mathcal{F}(A_k) P(A_k \mid X).
$$

Instead, we might calculate the approximation

$$\langle A_{k+1} \mid X \rangle \approx \mathcal{F} \left(\sum_{A_k \in S_n} A_k P(A_k \mid X) \right) = \mathcal{F}(\langle A_k \mid X \rangle). \qquad (4.1)$$

Remark 4.1. There is equality in (4.1) if the only nonlinearities in $\mathcal{F}$ involve the multiplication of independent stochastic variables. In most cases, we will have to consider the expected number(s) of some combinations of edges being present. Here the edge-independence assumption is exactly what we need. For example, the expected value of the number of mutual adjacencies (for any given pair of vertices) involves a sum over all pairs of edges connecting to the possible mutual adjacent vertex. These are mutually independent, and of course, the two necessary edges within each term are mutually independent. Hence for this type of $\mathcal{F}$, with each term in the range involving only sums over independent events each of which itself is a product over individual edges, (4.1) is exact. We refer to (4.1) as a *mean field approximation* for the evolving graph.

Suppose that we may represent $\langle A_{k_0} \mid X \rangle$ by some random graph, say $W_{k_0} \in R_n$. Then using the mean field approximation, we simply iterate with $\mathcal{F}$:

$$W_{k+1} = \mathcal{F}(W_k), \quad k = k_0, k_0 + 1, k_0 + 2, \ldots,$$

to obtain $\langle A_k \mid X \rangle = W_k$ for all $k = k_0, k_0 + 1, k_0 + 2, \ldots$. This iteration generates a sequence of expected values for the evolving network at all future time steps, given the approximation for A_{k_0}, but using the mean field approximation.

Now let us assume that $W(\lambda)$ is a conjugate random graph for $\mathcal{F}$. If we have $W_{k_0} = W(\lambda_{k_0})$ for some $\lambda_{k_0} \in \Lambda$, then we can iterate with g to produce a sequence

$$\lambda_{k+1} = g(\lambda_k), \quad k = k_0, k_0 + 1, k_0 + 2, \ldots,$$

and it follows that

$$\langle A_k \mid X \rangle = W_k = W(\lambda_k), \quad k = k_0, k_0 + 1, k_0 + 2, \ldots.$$

Hence under the mean field approximation, a conjugate random graph is particularly useful, since it descends from the mean field iteration over R_n to one over Λ.

5.   Bistability through Triadic Closure

We now apply this mean field theory to the triadic closure model (2.4)–(2.5), where

$$\mathcal{F}(A_k) = (1 - \widetilde{\omega})A_k + (\mathbf{1} - A_k) \circ \left(\delta \mathbf{1} + \epsilon A_k^2\right). \tag{5.1}$$

It is easy to see by symmetry of the model (there being neither distinguished vertices nor differences in the way vertices and edges are treated) that the Erdős–Rényi graphs are possible conjugate random graphs for this mapping. Substituting $\langle A_k \mid X \rangle = p_k \mathbf{1}$ for A_k as the heuristic mean field approximation, we obtain the iteration $p_{k+1} = g(p_k)$, where

$$g(p) = (1 - \widetilde{\omega})p + (1 - p)\left(\delta + \epsilon(n - 2)p^2\right). \tag{5.2}$$

At equilibrium, we have $p_{k+1} = p_k \equiv p^\star$, where

$$p^\star = \frac{(1 - p^\star)\left(\delta + \epsilon(n - 2)p^{\star 2}\right)}{\widetilde{\omega}}. \tag{5.3}$$

In the limit $\delta \to 0$, there are three real roots $p^\star$ if and only if $\widetilde{\omega} < \epsilon(n - 2)/4$. The smallest is $\delta/(\delta + \widetilde{\omega}) + O(\delta^2)$, representing a sparse graph with almost no triangulation, where the random birth rate δ equilibrates with death rate $\widetilde{\omega}$, and the larger roots are at

$$\frac{1}{2} \pm \sqrt{\frac{1}{4} - \frac{\widetilde{\omega}}{\epsilon(n - 2)}} + O(\delta),$$

where the nonlinear triangulation term equilibrates with the death rate $\widetilde{\omega}$. Moreover, $g'(p) \geq 0$ for all $p \in [0, 1]$ and $g(0) = \delta$. Hence in this regime, the two outer steady states are stable for this iteration, while the middle root is unstable. Intuitively, with a low initial edge density, the triangulation rule cannot get started, and the network remains sparse, whereas for a sufficiently high initial edge density, the network evolves into an ϵ-dependent state. Figure 2 illustrates the case of $n = 100$, using model parameters given by (2.6). Here, we find stable fixed points at 0.049 and 0.721, surrounding an unstable fixed point at $p^\star = 0.229$. These values are consistent with the results summarized in Figure 1; in that experiment, we deliberately chose an initial edge density very close to the unstable level. Three paths have evolved toward the lower stable level of 0.049, and one has been attracted to the upper stable level of 0.721. The mean field analysis has correctly predicted that two broad classes of behavior are possible over this timescale, and because we have started near the unstable level, the microscale detail determines which of the two regimes dominates for any particular path.
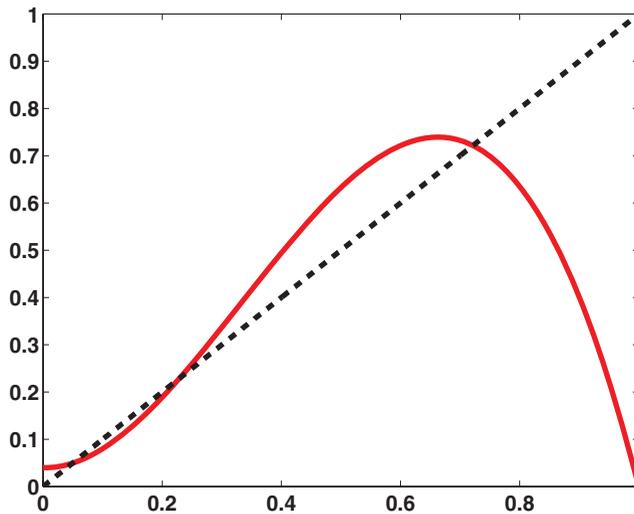
Figure 2. Graph of p (dashed) and $(1-p)(\delta + \epsilon(n-2)p^2)/\widetilde{\omega}$ (solid): the fixed points are at 0.049, 0.229, and 0.721. (Here $n = 100$ and parameter values are taken from (2.6)) (color figure available online).

As a further test, the jagged curves in Figure 3 record the edge density along three independent paths with the same model parameters (2.6), using initial networks given by sampling an Erdős–Rényi random graph with edge probability $p = 0.4$. In this denser regime, all three paths have edge densities that evolve toward the 0.721 level. The solid curve represents the mean field recurrence from (5.2), with $p_0 = 0.4$. We see that there is very good agreement when this macroscopic quantity is computed directly from the full microscopic simulation and from the mean field approximation. In other tests with initial networks that were not Erdős–Rényi, we found that the mean field iteration gave a good approximation if we started after a suitable "burn-in" period.

It is also informative to calculate, at each t_k, the average Watts–Strogatz clustering coefficient C_k, defined as the clustering coefficient averaged over all nodes. We may then compare its evolution with that of the edge density $\widehat{p}_k$. Since the initial network and long-term networks are Erdős–Rényi graphs, we will have have $C_k = \widehat{p}_k$ there because $\widehat{p}_k$ represents the probability for triangle closure. For the simulation in Figure 4, we see that C_k increases slightly ahead of $\widehat{p}_k$, since initially random clusters strengthen before they infill at the higher density. The individual, vertexwise, values of the clustering coefficient increase in variance during the phase of rapid growth; see Figure 5.
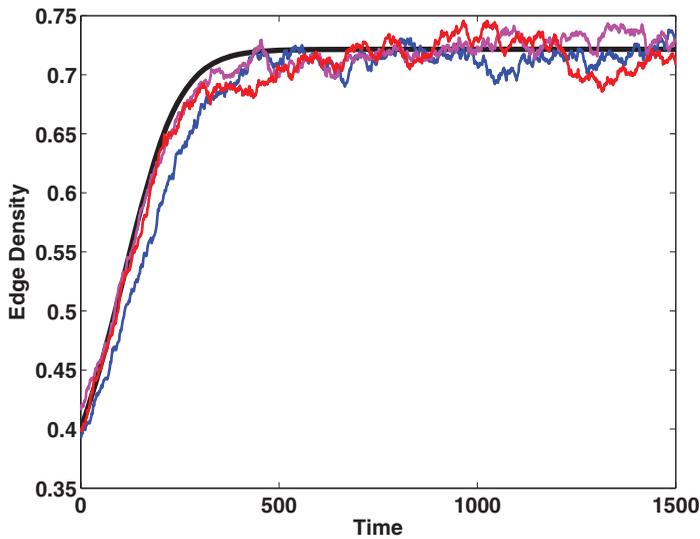
Figure 3. Jagged curves: edge density $\widehat{p}_k$ in (2.7) extracted from three independent paths of the Markov chain using the same model parameters. Smooth curve: mean field approximation from (5.2) (color figure available online).

To finish this section, we emphasize our belief that in order to understand network formation and forecast network behavior, it is essential to have a class of realistic dynamical models capable of replicating behavior seen in evolving network structures. Indeed, if we merely observe evolving social and communication networks, then, as with any dynamical system, we see only some subset of stable behavior. The bistability effect shown in this section confirms that a single realization can give a misleading picture, and also opens up the possibility of *timely and targeted intervention*—for example, a service provider may wish to stimulate early activity to move the edge density into a region where the network will then self-organize into a profitable, well-connected regime.

6. Application to Voice Call Data

In this section we consider an evolving network data set of pairwise mobile phone communications during a period of the year when connections were on the increase. We compare the unstructured (homogeneous) model and a block model introduced in Section 3 by first calibrating them via the evolution of the
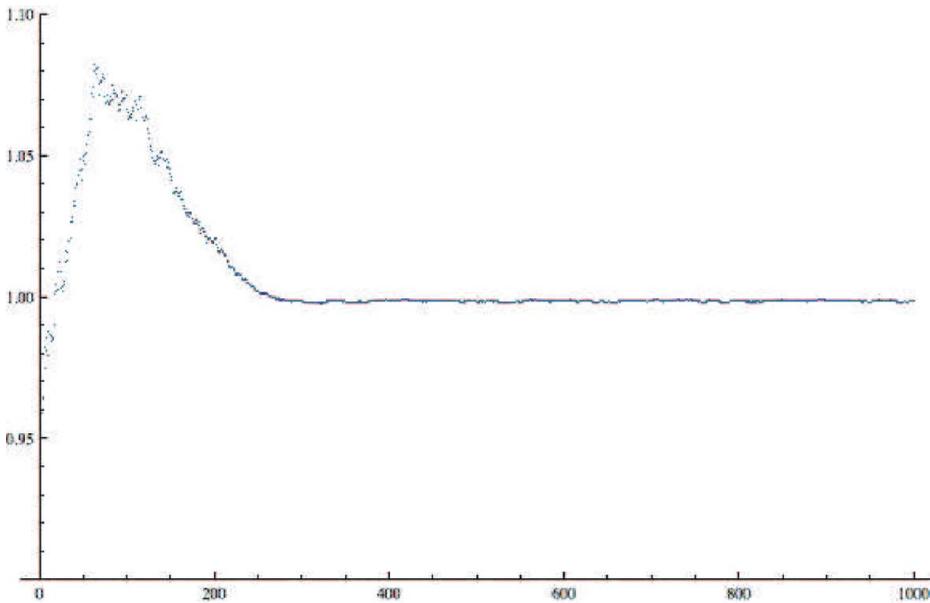
Figure 4. Ratio of average clustering coefficient to edge density: $C_k/\widehat{p}_k$ (color figure available online).

appropriate edge densities, and then examining how they predict the corresponding possible evolution of the average Watts–Strogatz clustering coefficient.

Suppose that we observe an evolving random graph. By making an assumption about $\mathcal{F}$ and selecting an appropriate subgroup of permutations Q, we may derive the mean field equations over $\mathcal{W}$, which will involve the dynamical parameters from $\mathcal{F}$. From data, we can calculate the evolution of the coordinates λ describing $\mathcal{W}$. By fitting these to the mean field model, we can estimate the unknown
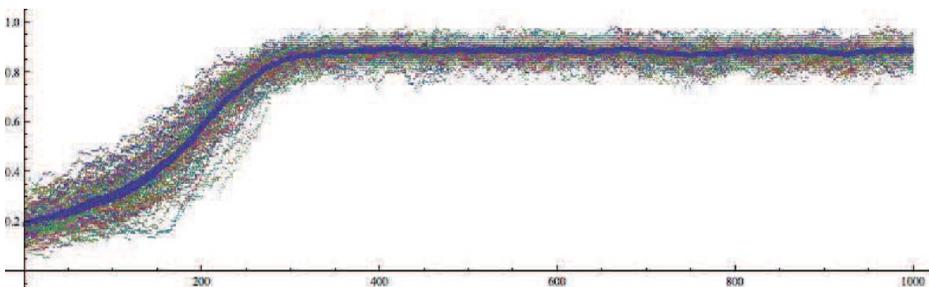

Figure 5. Vertex-level clustering coefficients as the network in Figure 4 evolves (color figure available online).

dynamical parameters. Then the choice of model may be validated or invalidated by checking other evolving network metrics not used to do the calibration.

As an example, we study weeks 8 through 15 from the reality mining data set given in [Eagle et al. 09], showing voice calls among around 150 people over time. We consider the weekly call networks, summarizing which pairs of people communicated during successive weeks.

Assuming initially that no individuals are distinctive in any way, we first fit the simple three-parameter unstructured model in (5.1). Assuming a homogeneous population, since $\mathcal{F}$ is invariant under all permutations, the mean field dynamics act over $\mathcal{W} = \{p\mathbf{1} \mid p \in [0,1]\}$. So, as described in Section 5, we have $p_{k+1} = g(p_k)$ for g in (5.2). From the data we can estimate p_k via $\widehat{p}_k$ in (2.7), the density of edges present. We have

$$\widehat{p}_{k+1} = (1 - \widetilde{\omega})\widehat{p}_k + (1 - \widehat{p}_k)(\delta + \epsilon(n-2)\widehat{p}_k^2) + \mathrm{err}_k,$$

where the errors, err_k, arise as an average over $n(n-1)/2$ independent edge processes, each of which must take binary values. Thus a Gaussian approximation to the structure of the errors is reasonable, and the parameters may be fitted with simple least squares. We note that the mean field approximation involves the same parameters as the full microscale model. Because the mean field iteration is scalar, this type of calibration is extremely inexpensive. Once the $\widehat{p}_k$ values have been computed, we solve an overdetermined $(K+1) \times 3$ linear system. This approach results in estimates $(\delta, \epsilon, \widetilde{\omega}) = (0.02170, 0.00868, 0.18399)$. In Figure 6, we show $\widehat{p}_k$ as well as the 5th and 95th percentiles arising from 200 simulations with the full model, each starting from A_8 and using the estimated dynamical parameters. We see that the calibrated model provides an ensemble of simulations about the actual data.

To test the accuracy of the approach in a controlled setting, we then simulated data from the model and attempted to reverse engineer known parameter values. We generated multiple synthetic data sets with 106 vertices over eight weeks, with actual parameters $(\delta, \epsilon, \widetilde{\omega}) = (0.02170, 0.00868, 0.18399)$. Typically, the least squares approach produced errors of approximately 20% in all three parameters.

Now let us examine the performance of the model by considering the evolution of the average Watts–Strogatz clustering coefficient C_k from week to week. For an Erdős–Rényi graph, we have $C_k = p_k$ in expectation. In Figure 7, we see that the model performs extremely poorly, since the observed values for C_k are much higher than those achievable under the conditioned Erdős–Rényi mean field evolution.

One explanation for this poor fit is that the population is not homogeneous, which motivates us to consider a block random graph model of the type introduced in Section 3 in order to increase the clustering within some subgroup and thus increase the values for C_k overall while keeping the overall edge density low.
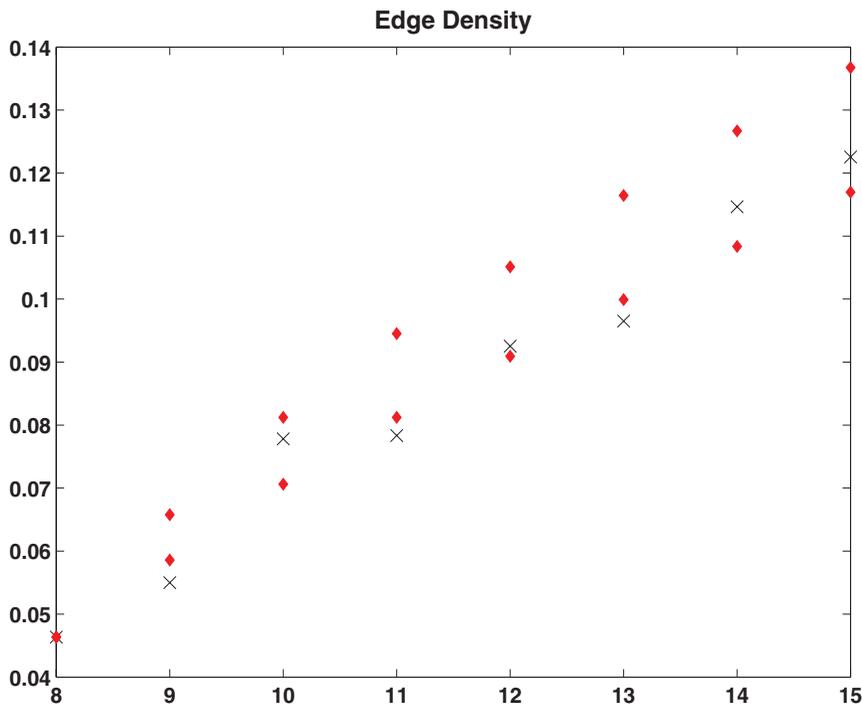
Figure 6. Evolution of edge density by week from data (crosses), and also the 5th and 95th percentile values (diamonds) achieved via an ensemble of model simulations, each using the fitted dynamical parameter values, and each starting from A_8 (color figure available online).

We partitioned the vertices into two sets: one of size $n_1 = 92$, and one of size $n_2 = 14$. This was done by summing the adjacency matrices into a single similarity matrix and adopting a spectral clustering approach [Strang 08]; using the Fiedler vector, we split across the greatest jump where both partitions contain at least 10% of the nodes. In other applications this task might be done a priori on grounds such as gender, functional role, or responsibility of the individuals, and of course, other clustering algorithms could be used.

Assuming an identical birth and death dynamic for edges within each vertex subset and between the subsets, then in the (p, q, r) notation of Section 3, the mean field dynamic at the $(k + 1)$th time step becomes

$$p_{k+1} = (1 - \widetilde{\omega}_p)p_k + (1 - p_k)(\delta_p + \epsilon_p((n_1 - 2)p_k^2 + n_2 r_k^2)),$$
$$r_{k+1} = (1 - \widetilde{\omega}_r)r_k + (1 - r_k)(\delta_r + \epsilon_r((n_1 - 1)p_k r_k + (n_2 - 1)q_k r_k)),$$
$$q_{k+1} = (1 - \widetilde{\omega}_q)q_k + (1 - q_k)(\delta_q + \epsilon_q((n_2 - 2)q_k^2 + n_1 r^2)),$$
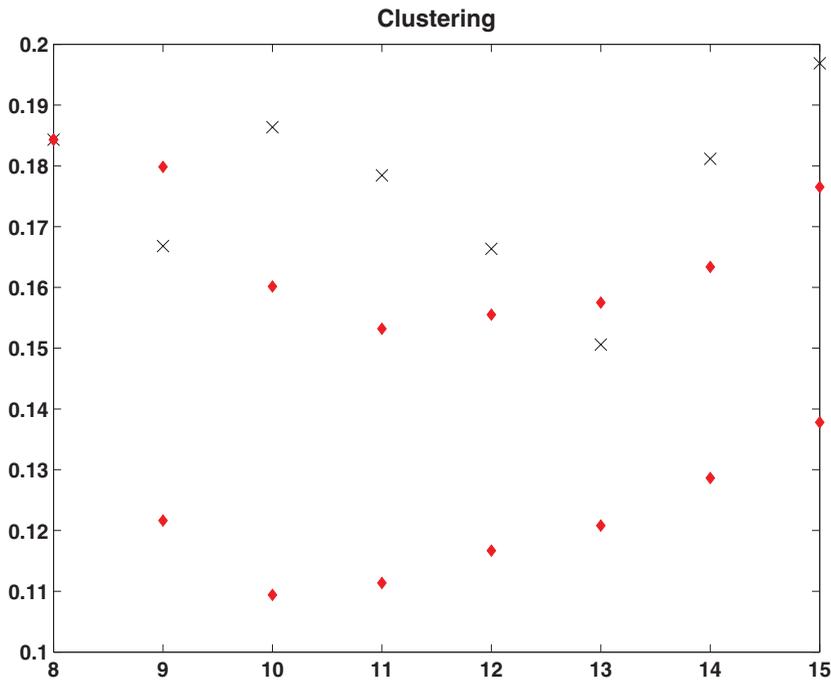
Figure 7. Evolution of C_k by week from data (crosses), and also the the 5th and 95th percentile values (diamonds) achieved via an ensemble of model simulations, each using the fitted dynamical parameter values, and each starting from A_8 (color figure available online).

involving nine parameters. These nine degrees of freedom can be fitted using the estimates for (p_k, q_k, r_k). We obtain the overall edge density evolution shown in Figure 8.

Now let us examine the performance of the block model by considering the evolution of the Watts–Strogatz clustering coefficient C_k. In Figure 9, we see that the model performs far better than the unstructured version. This indicates that there may well be some hierarchical stratification (such as the one proposed here on the basis of a priori clustering) among the subjects. It would of course be of interest to use further network properties in judging the calibration and model fit, and, where feasible, to use the full microscale-based likelihood.

7. Discussion

The motivation for this work was to develop a framework for modeling and analyzing dynamic connectivity structures. A successful model offers the potential
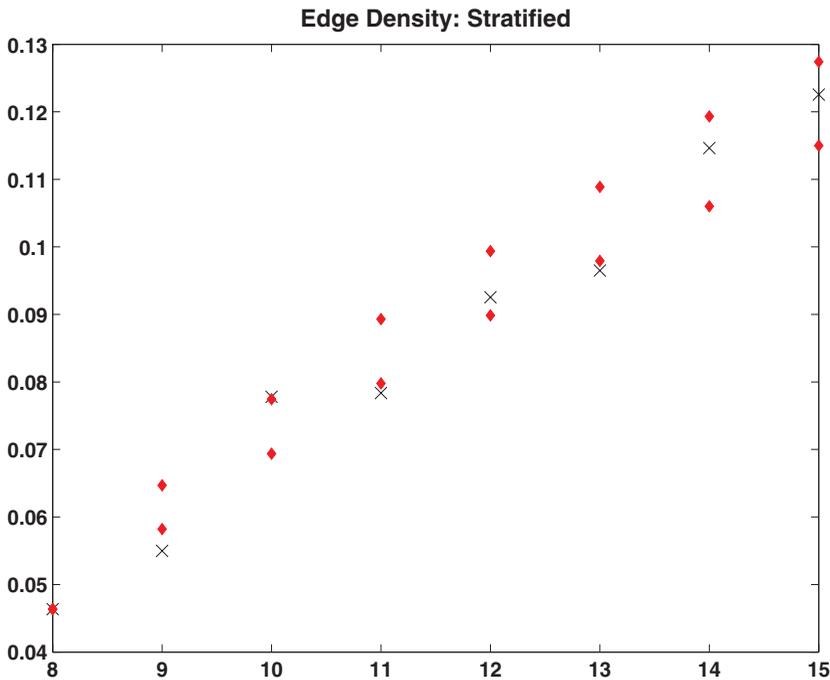
Edge Density: Stratified

Figure 8. Block model: evolution of edge density by week from data (crosses), and also the 5th and 95th percentile values (diamonds) achieved via an ensemble of block model simulations, each using the fitted dynamical parameter values, and each starting from A_8 (color figure available online).

to illustrate the range of possible behaviors and also to allow predictions under various "what if" scenarios, such as spiking information (rumors, marketing, stimulus) or making direct perturbations (disabling or enhancing specific vertices or edges).

By introducing the concept of conjugate graphs and developing a mean field theory, we have opened up the potential to approximate and calibrate the dynamics of observed networks. In particular, this can help us to identify when a network is close to an unstable rest point, where stochastic details will be important.

We illustrated the ideas on a simple but realistic nonlinear model based on a small set of parameters that is amenable both to analysis and to calibration. In this model, new edges are more likely to appear through successive triangulation (in addition to random births and deaths). This is a natural dynamic that is likely to enhance communication efficiency and resilience. However, our analysis showed that given a lack of early stimulus, the networks will remain at a relatively
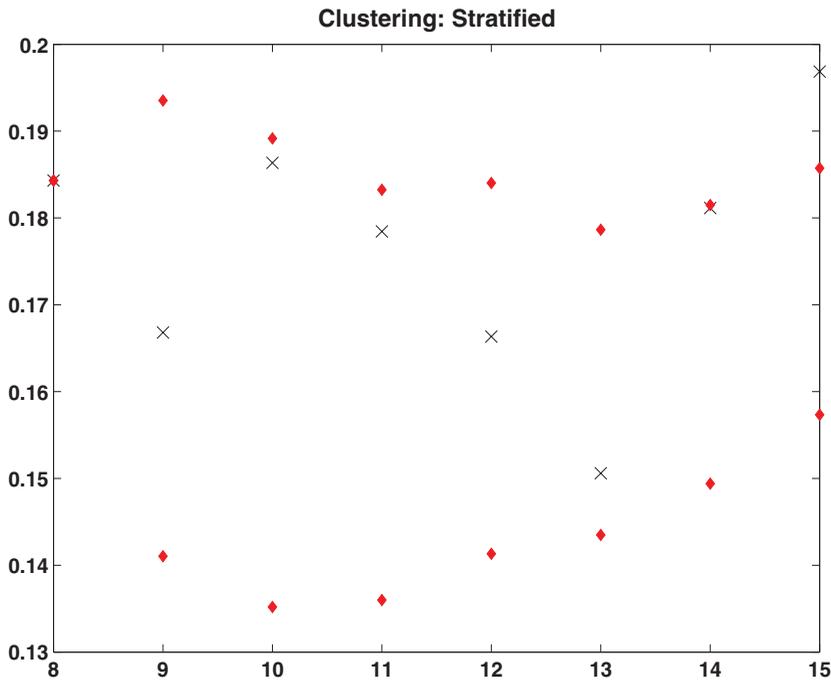
Figure 9. Block model: Evolution of C_k by week from data (crosses), and also the 5th and 95th percentile values (diamonds) achieved via an ensemble of block model simulations, each using the fitted dynamical parameter values, and each starting from A_8 (color figure available online).

sparse state—a sufficiently dense initial state is needed to produce well-clustered long-time networks. There is, of course, great scope for many other classes of evolving network mechanisms to be proposed, analyzed, and calibrated using the ideas and tools developed here.

References

[Avin et al. 08] C. Avin, M. Koucký, and Z. Lotker. "How to Explore a Fast-Changing World (Cover Time of a Simple Random Walk on Evolving Graphs)." In *ICALP*

'08: *Proceedings of the 35th International Colloquium on Automata, Languages and Programming*, part I, pp. 121–132. Springer, 2008.

[Barabási 05] A.-L. Barabási. "The Origin of Bursts and Heavy Tails in Human Dynamics." *Nature* 435 (2005), 207–211.

[Barabási and Albert 99] A.-L. Barabási and R. Albert. "Emergence of Scaling in Random Networks." *Science* 286 (1999), 509–512.

[Bassett et al. 11] D. S. Bassett, N. F. Wymbs, M. A. Porter, P. J. Mucha, J. M. Carlson, and S. T. Grafton. "Dynamic Reconfiguration of Human Brain Networks during Learning." *Proc. Nat. Acad. Sci.* 108 (2011), doi: 10.1073/pnas.1018985108.

[Berman 96] K. Berman. "Vulnerability of Scheduled Networks and a Generalization of Menger's Theorem." *Networks* 28 (1996), 125–134.

[Bullmore and Sporns 09] E. Bullmore and O. Sporns. "Complex Brain Networks: Graph Theoretical Analysis of Structural and Functional Systems." *Nature Reviews Neuroscience* 10 (2009), 186–198.

[Clementi et al. 08] A. E. Clementi, C. Macci, A. Monti, F. Pasquale, and R. Silvestri. "Flooding Time in Edge-Markovian Dynamic Graphs." In *Proceedings of the 27th Annual ACM SIGACT-SIGOPS Symposium on Principles of Distributed Computing (PODC'08)*, pp. 213–222. ACM Press, 2008.

[Clementi et al. 09] A. E. F. Clementi, F. Pasquale, A. Monti, and R. Silvestri. "Information Spreading in Stationary Markovian Evolving Graphs." In *Proceedings of the 2009 IEEE International Symposium on Parallel & Distributed Processing, Washington, DC, USA*, pp. 1–12. IEEE Computer Society, 2009.

[Eagle et al. 09] 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), 15274–15278.

[Esfandiar et al. 10] P. Esfandiar, F. Bonchi, D. Gleich, C. Greif, L. Lakshmanan, and B.-W. On. "Fast Katz and Commuters: Efficient Estimation of Social Relatedness in Large Networks." In *Algorithms and Models for the Web-Graph*, edited by R. Kumar and D. Sivakumar, Lecture Notes in Computer Science 6516, pp. 132–145. Springer, 2010.

[Goodreau et al. 09] S. Goodreau, J. A. Kitts, and M. Morris. "Birds of a Feather or Friend of a Friend? Using Exponential Random Graph Models to Investigate Adolescent Friendship Networks." *Demography* 46 (2009), 103–126.

[Grindrod and Higham 10] P. Grindrod and D. J. Higham. "Evolving Graphs: Dynamical Models, Inverse Problems and Propagation." *Proceedings of the Royal Society, Series A* 466 (2010), 753–770.

[Grindrod et al. 11] P. Grindrod, M. C. Parsons, D. J. Higham, and E. Estrada. "Communicability across Evolving Networks." *Phys. Rev. E* 83 (2011), 046120.

[Gross and Blasius 08] T. Gross and B. Blasius. "Adaptive Coevolutionary Networks: A Review." *J. Royal Society Interface* 5 (2008), 259–271.

[He et al. 07] Y. He, Z. J. Chen, and A. C. Evans. "Small-World Anatomical Networks in the Human Brain Revealed by Cortical Thickness from MRI." *Cerebral Cortex* 17 (2007), 2407–2419.

[Holme 05] P. Holme. "Network Reachability of Real-World Contact Sequences." *Physical Review E* 71 (2005), 046119.

[Holme and Saramäki 12] P. Holme and J. Saramäki. "Temporal Networks." *Physics Reports* 519 (2012), 97–125.

[Kossinets and Watts 06] G. Kossinets and D. J. Watts. "Empirical Analysis of an Evolving Social Network." *Science* 311 (2006), 88–90.

[Kossinets et al. 08] G. Kossinets, J. Kleinberg, and D. Watts. "The Structure of Information Pathways in a Social Communication Network." In *Proceedings of the 14th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, KDD '08, New York, NY, USA*, pp. 435–443. ACM, 2008.

[Liben-Nowell and Kleinberg 07] D. Liben-Nowell and J. Kleinberg. "The Link-Prediction Problem for Social Networks." *Journal of the American Society for Information Science and Technology* 58 (2007), 1019–1031.

[Liu et al. 08] Y. Liu, M. Liang, Y. Zhou, Y. He, Y. Hao, M. Song, C. Yu, H. Liu, Z. Liu, and T. Jiang. "Disrupted Small-World Networks in Schizophrenia." *Brain* 131 (2008), 945–961.

[Lu et al. 10] Z. Lu, B. Savas, W. Tang, and I. Dhillon. "Supervised Link Prediction Using Multiple Sources." In *IEEE 10th International Conference on Data Mining (ICDM), December 2010*, pp. 923–928, 2010.

[Mucha et al. 10] 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), 876–878.

[Newman 10] M. E. J. Newman. *Networks: An Introduction.* Oxford University Press, 2010.

[Newman et al. 00] M. E. J. Newman, C. Moore, and D. J. Watts. "Mean-Field Solution of the Small-World Network Model." *Physical Review Letters* 84 (2000), 3201–3204.

[Strang 08] G. Strang. *Computational Science and Engineering.* Wellesley–Cambridge Press, 2008.

[Szell and S. Thurner 10] M. Szell and S. Thurner. "Measuring Social Dynamics in a Massive Multiplayer Online Game." *Social Networks* 39 (2010), 313–329.

[Tang et al. 09] J. Tang, S. Scellato, M. Musolesi, C. Mascolo, and V. Latora. "Small-World Behavior in Time-Varying Graphs." *Physical Review E* 81 (2010), 05510.

[Tang et al. 10] 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*, pp. 1–6. ACM, 2010.

[Wagner 03] A. Wagner. "How the Global Structure of Protein Interaction Networks Evolves." *Proceedings of the Royal Society of London. Series B, Biological Sciences* 270 (2003), 457–466.

[Watts and Strogatz 98] D. J. Watts and S. H. Strogatz. "Collective Dynamics of 'Small-World' Networks." *Nature* 393 (1998), 440–442.

[Zhao et al. 11] K. Zhao, J. Stehlé, G. Bianconi, and A. Barrat. "Social Network Dynamics of Face-to-Face Interactions." *Physical Review E* 83 (2011), 056109+.

Peter Grindrod, Department of Mathematics, University of Reading (p.grindrod@reading.ac.uk)

Desmond J. Higham, Department of Mathematics and Statistics, University of Strathclyde (d.j.higham@strath.ac.uk)

Mark C. Parsons, Department of Mathematics, University of Reading (m.c.parsons@pgr.reading.ac.uk)