Research

Author for correspondence:
Desmond J. Higham
e-mail: d.j.higham@strath.ac.uk

A dynamical systems view of network centrality

Peter Grindrod[1] and Desmond J. Higham[2]

[1]Mathematical Institute, University of Oxford, Oxford OX2 6GG, UK
[2]Department of Mathematics and Statistics, University of Strathclyde, Glasgow G1 1XH, UK

To gain insights about dynamic networks, the dominant paradigm is to study discrete *snapshots*, or *timeslices*, as the interactions evolve. Here, we develop and test a new mathematical framework where network evolution is handled over continuous time, giving an elegant dynamical systems representation for the important concept of node centrality. The resulting system allows us to track the relative influence of each individual. This new setting is natural in many digital applications, offering both conceptual and computational advantages. The novel differential equations approach is convenient for modelling and analysis of network evolution and gives rise to an interesting application of the matrix logarithm function. From a computational perspective, it avoids the awkward up-front compromises between accuracy, efficiency and redundancy required in the prevalent discrete-time setting. Instead, we can rely on state-of-the-art ODE software, where discretization takes place adaptively in response to the prevailing system dynamics. The new centrality system generalizes the widely used Katz measure, and allows us to identify and track, at any resolution, the most influential nodes in terms of broadcasting and receiving information through time-dependent links. In addition to the classical static network notion of attenuation across edges, the new ODE also allows for attenuation over time, as information becomes stale. This allows 'running measures' to be computed, so that networks can be monitored in real time over arbitrarily long intervals. With regard to computational efficiency, we explain why it is cheaper to track good receivers of information than good broadcasters. An important consequence is that the overall broadcast activity in the network can also be monitored efficiently. We use two synthetic examples to validate the relevance of the new measures. We then illustrate the ideas on a large-scale voice call network, where key features are discovered that are not evident from snapshots or aggregates.

1. Motivation

Systems involving transient interactions arise naturally in many areas, including telecommunication, online social networking and neuroscience. Dealing with the inherent time-dependency of edges in a network raises a variety of challenges for modelling and computation [1,2]. Currently, the dominant paradigm is to consider activity over time-windows, and hence to deal with network snapshots over discrete time; for example, hourly or daily summaries of who phoned whom or who texted whom. This approach can be used to study both the evolution of the underlying connectivity and the dynamics of processes over the network [3–5]. The idea of viewing a dynamic network as a time-ordered sequence of sparse adjacency matrices is also extremely convenient for data-driven analysis [6–11]. More generally, abstracting from time to any countable index representing possible levels in a hierarchy (such as gene/mRNA/protein) we arrive at multiplex networks [12], which can also be viewed as multi-dimensional tensors [13,14].

However, owing to the bursty behaviour of human interactions [15] aggregating over pre-chosen time windows has a number of drawbacks. On the one hand, by taking the time window too large we lose the ability to reproduce high-frequency transient behaviour, where an edge switches on and off multiple times in the space of a single window. On the other hand, time points that are too finely spaced can lead to redundant, empty windows that waste computational effort. Very short time windows may also lead to a false impression of accuracy—for remote interactions such as e-mails and tweets, although there is an unambiguous and instantaneous time at which information was sent, the time at which the receiver (i) digests and then, possibly, (ii) acts upon the information is generally not recordable. Hence, although it is intuitively appealing owing to its simplicity, we believe that in many settings, notably the avalanches of human communication arising through digital media, the discrete snapshot view is unnatural and forces us to make unnecessary compromises. Our aim is therefore to overcome these conceptual and practical shortcomings by developing the first model to extract centrality insights directly from an adjacency matrix defined as a function of continuous time. We focus on the task of simulating with real dynamic network data. However, we note that by making available the tools of dynamical systems, this work also allows us to incorporate the concept of centrality into theoretical models of network evolution based, for example, on a law of social balance [16,17] or triadic closure [18]. Hence, while this work focuses on the key step of deriving the continuum limit and confirming its relevance on a real dataset, the ideas also open up the opportunity to study the evolution of centrality under various network 'laws of motion' and to compare predictions from competing models.

In the next section, we summarize some relevant background material in static network centrality measures. In §3, we then set up and derive a differential equation for continuous-time network centrality. Section 4 continues with some observations about the extraction of node-level information, choice of parameters and extensions of the model. We illustrate the relevance of the new measures in §5 by testing on synthetically constructed dynamic networks that contain known hierarchies. Section 6 then applies the methodology to voice call interactions, showing that key features can be extracted. We finish in §7 with some conclusions.

2. Background on network centrality

Using standard notation, we let $A \in \mathbb{R}^{N \times N}$ denote the adjacency matrix for an unweighted, directed, static network of N nodes, so that $A_{ij} = 1$ if and only if there is a link from i to j and $A_{ij} = 0$ otherwise. In many circumstances, it is desirable to quantify the level of importance or influence of each node in the network. This task has been addressed by many authors, with key early contributions emerging in the field of social science [19,20]. We focus here on the widely used concept of Katz centrality [21], which is motivated by the expansion

$$(I - aA)^{-1} = I + aA + a^2 A^2 + \cdots + a^k A^k + \cdots . \qquad (2.1)$$

Here, $0 < a < 1$ is a scaling parameter and the expansion is valid for $a < 1/\rho(A)$, where $\rho(\cdot)$ denotes the spectral radius. To interpret (2.1), we note that $(A^k)_{ij}$ counts the number of distinct *walks of length k* from i to j. In this setting, a walk of length k is defined to be a traversal involving exactly k edges, with no constraint on the re-use of edges or nodes along the way. The i, j element in expansion (2.1) therefore has a natural combinatoric interpretation; it counts a weighted sum of the number of walks from i to j, with walks of length k scaled by a^k. As mentioned in the original work of Katz, we can interpret the attenuation factor a as the probability that an edge is successfully traversed. If we consider messages being passed along the directed edges, it follows that the resolvent matrix $(I - aA)^{-1}$ has components that quantify the ability of i to pass a message to j: we are accounting for all possible routes, with longer ones given less importance. The row and column sums

$$(I - aA)^{-1}\mathbf{1} \quad \text{and} \quad (I - aA)^{-T}\mathbf{1}, \tag{2.2}$$

where $\mathbf{1} \in \mathbb{R}^N$ denotes the vector of ones, then have ith components that measure the centrality of node i via its ability to broadcast and receive information, respectively. Loosely, a node with a high Katz broadcast centrality is a good place to start a rumour, and a node with a high Katz receive centrality is a good place to hear the latest rumour.

Let us now consider interactions that come and go over time. The concept of centrality must take account of a new factor—the *follow-on effect* arising from the *time-ordering* of the interactions. If A meets B today and B meets C tomorrow, then a message may pass from A to C, but not from C to A. If we consider only independent snaphots or an overall aggregate of the interactions, then this type of follow-on influence becomes lost, considerably undermining the task of discovering key players. It is reasonable to expect that influential or astute individuals will generate ideas that are subsequently passed around the network by others. Indeed, this effect was observed empirically in [22, p. 7], where *discussion catalysts* in an online community were seen to be 'responsible for the majority of messages that initiate long threads'. Similarly, Huffaker [23, p. 594] identifies *online leaders* who 'trigger feedback, spark conversations within the community, or even shape the way that other members of a group "talk" about a topic'. Related experiments in Mantzaris & Higham [24] on e-mail and voice mail data led to the term *dynamic communicators* to describe individuals with an enhanced ability to disseminate or collect information relative to their centrality based on static or aggregate summaries. Our aim here is to present a framework that allows us to quantify these effects elegantly and efficiently in a continuous-time setting.

3. Continuous-time analysis

We consider the case where a fixed set of N nodes undergo interactions that can appear and disappear over time; i.e. we have time-dependent, directed, unweighted edges. We then let $A(t)$ be the corresponding time-dependent adjacency matrix. We have two fundamentally different settings in mind. For direct engagement, as in the case of voice call data, it is natural for the entries in the adjacency matrix to be discontinuous functions of time, switching on to 1 and off to 0 when the interaction starts and finishes, respectively. So in this case each $A(t) \in \mathbb{R}^{N \times N}$ is binary and symmetric with $A_{ij}(t) = A_{ji}(t) = 1$ if and only if i and j are in communication at time t. (To be concrete, and without loss of generality, we will assume continuity from the right.) For indirect social interactions such as e-mails and tweets, this framework allows us to spike the relevant edge when a message is sent out, and then, as t increases, let the value to decay towards zero—this reflects the fact that there may be a delay before the recipient digests the message, and also that messages lose their impact and visibility over time. In this case, $A(t)$ would generally be unsymmetric, with components taking values between zero and one.

Our aim is now to generate a real-valued, unsymmetric *dynamic communicability matrix* $S(t) \in \mathbb{R}^{N \times N}$ such that $S(t)_{ij}$ quantifies the ability of node i to communicate with node j up to time t. This will be constructed from first principles by counting *dynamic walks* that make use of edges appearing at time t or earlier. In this setting, in addition to downweighting walks according

to the number of edges they use, we will also downweight according to the temporal age of the walk, on the grounds that information typically becomes less relevant over time. On this basis, we propose that $S(t)$ is updated over a small time interval δt according to

$$S(t + \delta t) = (I + e^{-b\delta t}S(t))(I - aA(t + \delta t))^{-\delta t} - I, \tag{3.1}$$

with $S(0) = 0$. Here, $I \in \mathbb{R}^{N \times N}$ denotes the identity matrix, and $a, b > 0$ are parameters. To justify this new iteration and explain the role of the parameters, suppose first that $\delta t = 1$, and we are therefore given access to the adjacency matrix only on a fixed grid of timepoints. In the extreme case of a single timepoint, (3.1) reduces to $S(1) + I = (I - aA(1))^{-1}$. For $i \neq j$, this corresponds to the classical resolvent-based measure of Katz [21] in (2.1). With multiple time points and $b = 0$ in (3.1) we have

$$S(N) + I = (I - aA(1))^{-1}(I - aA(2))^{-1} \cdots (I - aA(N))^{-1}. \tag{3.2}$$

It was shown in [6] that this iteration generalizes the Katz centrality measure to the case of a discrete-time network sequence. Here $S(N)_{ij}$ records an edge-downweighted count of the total number of distinct dynamic walks. We note that the same idea was later used to define the weaker concept of accessibility [9]. The general case of (3.1) with $b > 0$ and $\delta t = 1$ was introduced in [7], where it was shown that the parameter b plays the role of downweighting for time—walks that began k time units ago are scaled by e^{-bk}.

A key step in writing down the new iteration (3.1) was to devise a scaling under which the $\delta t \to 0$ limit is meaningful. Our choice of a δt-dependent power in the matrix products can be explained by focusing on the $b = 0$ case over a short-time interval. If we refine from the single interval $[t, t + \delta t]$ to the pair of intervals $[t, t + \delta t/2]$ and $[t + \delta t/2, t + \delta t]$ and consider the scenario where $A(t)$ does not change, then the simple identity

$$(I - aA(t))^{-(1/2)\delta t}(I - aA(t))^{-(1/2)\delta t} = (I - aA(t))^{-\delta t}$$

shows that our scaling produces consistent results.

For convenience, we will let $U(t) = I + S(t)$ in taking the $\delta t \to 0$ limit. We then write (3.1) in the form

$$U(t + \delta t) = (I + e^{-b\delta t}(U(t) - I))\exp(-\delta t \log(I - aA(t + \delta t))),$$

where exp and log denote the matrix exponential and logarithm, respectively, and we have used the identities $H = e^{\log H}$ and $\log(H^\alpha) = \alpha \log H$ for $-1 \leq \alpha \leq 1$ (e.g. [25]). Expanding and taking the limit $\delta t \to 0$, we then arrive at the matrix ODE

$$U'(t) = -b(U(t) - I) - U(t)\log(I - aA(t)), \tag{3.3}$$

with $U(0) = I$. In summary, dynamical system (3.3) is driven by the continuous-time adjacency matrix $A(t)$ with $U(t)_{ij}$ for $i \neq j$ quantifying the current ability of node i to pass information to node j, with less emphasis given to longer and older dynamic walks.

Generalizing (2.2), in order to focus on nodal properties, we can take row and column sums in the matrix $U(t)$ to define the *dynamic broadcast* and *dynamic receive* vectors

$$\mathbf{b}(t) = U(t)\mathbf{1} \quad \text{and} \quad \mathbf{r}(t) = U(t)^{\mathsf{T}}\mathbf{1}. \tag{3.4}$$

Here, $\mathbf{b}_i(t)$ and $\mathbf{r}_i(t)$ measure the current propensity for node i to have broadcasted and received information, respectively, across the dynamic network, under our assumptions that longer and older walks have less relevance.

4. Remarks on the new dynamical system

It is interesting to note from (3.3) that the receive centrality satisfies its own vector-valued ODE

$$\mathbf{r}'(t) = -b(\mathbf{r}(t) - \mathbf{1}) - (\log(I - aA(t)))^{\mathsf{T}}\mathbf{r}(t), \tag{4.1}$$

with $\mathbf{r}(0) = \mathbf{1}$, which is a factor of N smaller in dimension than (3.3) and hence cheaper to simulate. By contrast, it is not possible to disentangle the broadcast centrality in this way. Intuitively, this

difference arises because the node-based receive vector $\mathbf{r}(t)$ keeps track of the overall level of information flowing into each node, and this can be propagated forward in time as new links become available. However, the broadcast vector $\mathbf{b}(t)$ keeps track of information that has flowed out of each node—it does not record where the information currently resides and hence we cannot update based on $\mathbf{b}(t)$ alone. So, with this methodology, real-time updating of the receive centrality is fundamentally simpler than real-time updating of the broadcast centrality.

We also note that as the sum of the row sums is equivalent to the sum of the column sums, the overall dynamic broadcast centrality of the network, $\sum_{i=1}^{N} \mathbf{b}_i(t)$, can be computed via (4.1) by forming $\mathbf{r}(t)^{\mathrm{T}}\mathbf{1}$. Hence, we can monitor the current broadcast ability of the whole network by simulating an ODE system that is a factor of N smaller than the system required for nodal broadcast information.

So, overall, in terms of computational cost, the dynamic vector of receive centralities, $\mathbf{r}(t)$, or the aggregate network centrality, $\sum_{i=1}^{N} \sum_{j=1}^{N} U_{ij}(t)$, may be computed, via simulation, for very large, sparse networks:

— the storage requirement scales like the number of nodes in the system, N,
— the dominant computational task, evaluating the right-hand side of (4.1), reduces to a small number of products of a sparse matrix with a full vector. In this sense, the cost is $O(N)$ per unit time for a network with $O(1)$ edges per node.

However, to simulate the dynamic broadcast centrality vector, $\mathbf{b}(t)$ in (3.4), requires us to deal with the full matrix $U(t)$ in (3.3), leading to $O(N^2)$ storage and an $O(N^2)$ cost per unit time. It is therefore of considerable interest to develop approximation methods that overcome this barrier; for example, by sparsifying $U(t)$ or, as mentioned later, reducing its dimension.

The choice of downweighting parameters, a and b, clearly plays an important role in (3.3), and should depend to a large extent on the nature of the interactions represented by $A(t)$. The temporal downweighting parameter, b, has the natural interpretation that walks starting t time units ago are downweighted by e^{-bt}, so b allows us summarize the rate at which news goes stale. Hence b could be calibrated by studying the typical half-life of a link [26].

For the edge-attenuation parameter, a, in discrete iteration (3.2), there is a clear upper limit given by the reciprocal of the spectral radius of $A(t)$. This upper limit is bounded below by the reciprocal of the maximum in-degree and maximum out-degree over the nodes. For the continuous-time ODE (3.3), the principal logarithm $\log(I - aA(t))$ is well defined if $1 - a\lambda_i > 0$ for all real eigenvalues λ_i of $A(t)$, [25]. We note that in the case of one-to-one undirected communication, such as voice calls in the absence of tele-conferences, $A(t)$ can always be permuted into a block diagonal structure, with non-trivial blocks of the form

$$\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix},$$

and hence this constraint on a reduces to $a < 1$.

The starting point (3.1) is based on the idea of counting dynamic walks in a very generous sense; any traversal that uses zero, one or more edges per time point is allowed. It is possible to take an alternative approach where other types of dynamic walk are counted, after downweighting. For example, we could start with the iteration

$$S(t + \delta t) = (I + \mathrm{e}^{-b\delta t} S(t)) H(A(t))^{\delta t} - I,$$

where $H(A(t))$ is some matrix function. Appropriate choices of $H(A(t))$ include truncated power series of the form $I + aA(t) + a^2 A(t)^2 + \cdots + a^p A(t)^p$, in which case we are counting only those dynamic walks that use at most p edges per time step in the discrete setting. In this generality, the ODE (3.3) becomes
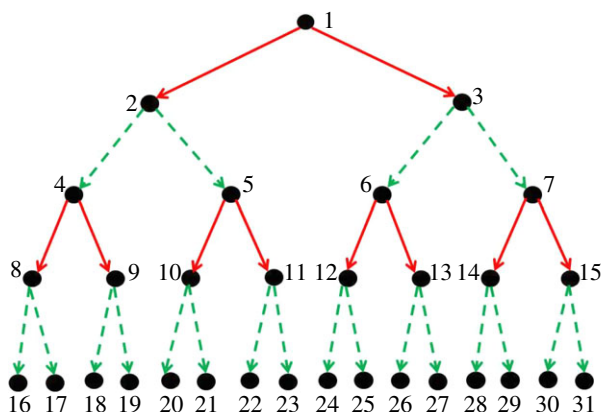
$$U'(t) = -b(U(t) - I) + U(t)\log(H(A(t))).$$

Figure 1. Network structure for the first synthetic experiment. The active links alternate between the solid and dashed sets, with extra noise added. The pattern repeats periodically over 10 cycles. (Online version in colour.)

In future work, it would be natural to generalize further to other types of dynamical system that provide different summaries. For example, given an evolving network $A(t) \in [0,1]^{N \times N}$ where N is extremely large, suppose that a subset of $M \ll N$ nodes are found to be of interest. It may then be worthwhile to study an ODE involving $V(t) \in \mathbb{R}^{M \times M}$ of the form

$$V'(t) = P(V(t)) + Q(V(t))F(A(t)),$$

where P and Q are polynomials or other matrix valued functions, and $F : [0,1]^{N \times N} \to \mathbb{R}^{M \times M}$ is an appropriate matrix valued mapping. This type of system allows us to project the interactions taking place between all N individuals onto the subset of interest, and then calculate a measure of the resulting evolution of behaviour at this lower dimension.

5. Synthetic experiments

To illustrate the use of our dynamical systems approach and give a feel for the role of the parameters, we now present two synthetic examples of networks evolving over continuous time. These examples have been constructed to have an intuitively reasonable hierarchy of influence that is hidden from the static or aggregate view. To aid visualization, the networks are small, and hence we were able to solve the ODE system (3.3) accurately with a highly resolved Euler iteration.

Our first experiment creates cascades through a directed binary tree structure, as illustrated in figure 1. We use a time interval $[0, 20]$, with $A(t)$ constant over each subinterval $[i, i + 1)$. To define $A(t)$ over $i \le t < i + 1$, we use the solid edges when i is even and the dashed edges when i is odd. Furthermore, we add some noise to this structure by including extra directed edges— these are chosen uniformly at random for each subinterval, with an average of five edges added each time. Overall, because of the hierarchy and timing of the edges, information is seen to be flowing from lower indices to higher. In particular, node 1 has the greatest propensity for passing information through the network, even though any single snapshot in time would not reveal this feature.

In figure 2a, we show, for each of the 31 nodes, the dynamic broadcast score from (3.4) at $t = 20$. In this case, the maximum spectral radius of $A(t)$ over $[0, 20]$ is one, and we used $a = 0.7$ and $b = 0.1$. We see that node 1 is strongly favoured by the centrality measure, and centrality generally decreases with index. Nodes 2 and 5 are ranked above node 3, suggesting that the symmetry breaking effect of the additional noise has slightly favoured this part of the network. In figure 2b, we show the same results when b is decreased to 0.01. In this case, we are allowing older walks to make a greater contribution: a walk starting at time zero is downweighted by
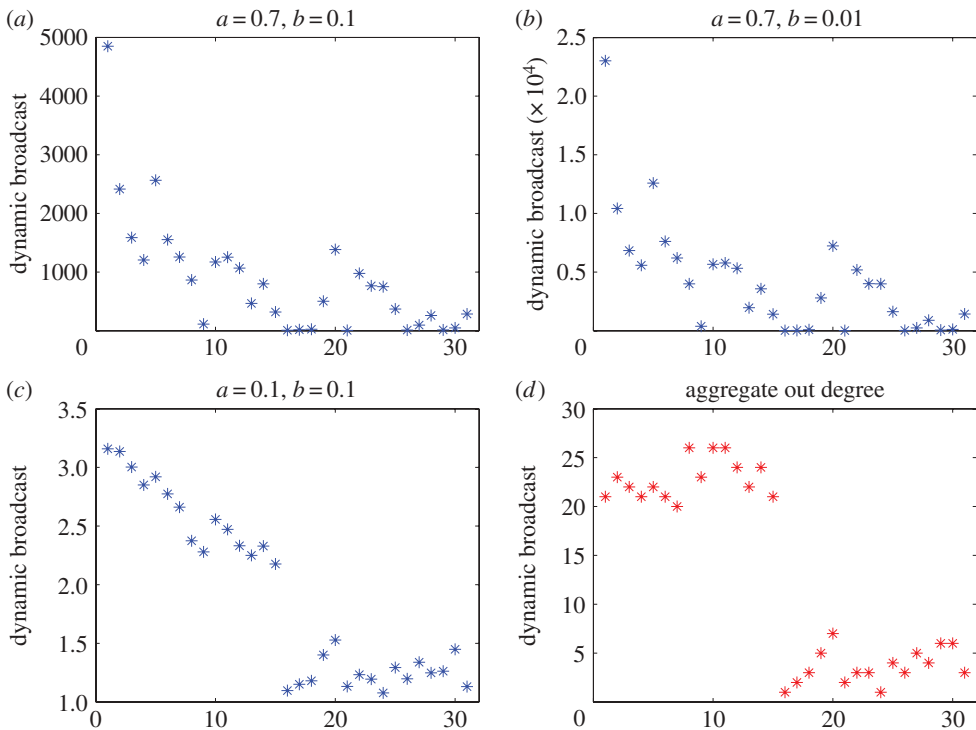
Figure 2. Results from the dynamic network in figure 1. (*a–c*) Broadcast centrality for each of the 31 nodes. (*d*) Shows the aggregate out degree for each node. (Online version in colour.)

$\exp(-20 \times 0.01) \approx 0.8$ rather than $\exp(-20 \times 0.1) \approx 0.1$. Because of the underlying periodicity in the network dynamics, this change has very little effect on the relative node rankings, but generates much larger absolute values. In figure 2*c*, we return to $b = 0.1$ and this time change to $a = 0.1$, so that walks using more edges are more strongly penalized. In this case, although node 1 remains the most central, the differences are less marked because the ability of node 1 to instigate many dynamic walks of length 4 to nodes in the bottom of the hierarchy carries less weight. Figure 2*d* shows the aggregate degree of each node, that is, the out degree integrated over time. For the underlying binary tree structure, this value is equal to 20 for nodes 1 to 15, and relative differences are simply a reflection of the added noise. This picture makes it clear that overall bandwidth can be a poor indicator of influence.

Our second experiment also uses a binary tree structure. In this case, the edges are undirected and every node maintains at most one edge at any given time, mimicking the voice call setting. Figure 3 illustrates the network. Here, labels have been assigned to the edges, indicating when they are active. These time intervals are ordered and non-overlapping, with $t_i := [(i - 1)\tau, (i - 1 + 0.9)\tau)$, for $i = 0, 1, \ldots, 7$, and $\tau = 0.1$. The dynamic network is designed so that node A is a more successful propagator than node B; for example, node A may have a higher status in a business or social setting, or may have access to a more timely item of news. We see that node A talks to node C in time-interval t_1, and this interaction is followed by a cascade of interaction down the hierarchy. By contrast, node B does not contact node C until time interval t_4, and there is no knock-on effect. From the aggregate perspective, nodes A and B have identical behaviour, both contacting nodes C and D for the same length of time. In our experiment, we let the dynamic network repeat over five cycles, so nodes A and B send out 10 messages in total.

For $a = 0.9$ and $b = 0.01$, figure 4 shows the evolution over time of the broadcast centrality for node A (solid) and node B (dashed). We see that this dynamic measure is able to capture the knock-on effect enjoyed by node A.
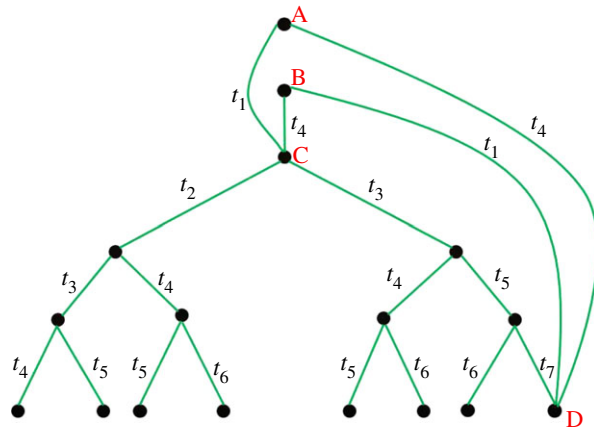
Figure 3. Network structure for the second synthetic experiment. Links are active over non-overlapping time intervals t_i, as indicated. The pattern repeats periodically over five cycles. (Online version in colour.)
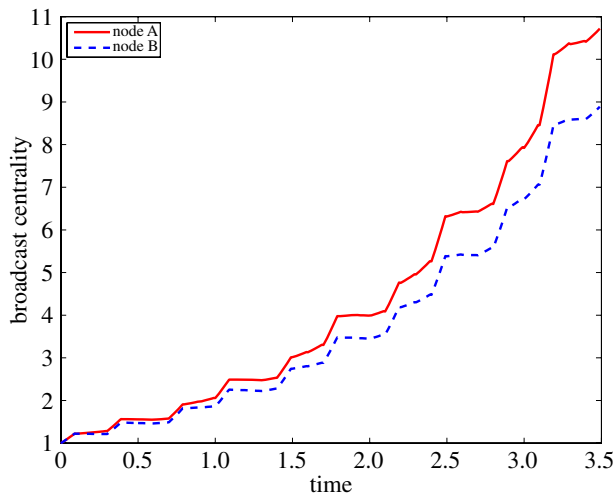

Figure 4. Dynamic broadcast centrality over time for node A (solid) and node B (dashed) in the network of figure 3. (Online version in colour.)

6. Voice call experiment

We now illustrate the use of this new modelling framework on a set of voice call interactions supplied as part of the IEEE VAST 2008 Challenge [27]. These realistic data were designed 'with the purpose of pushing the forefront of visual analytics tools using benchmark data sets;' further details can currently be found at http://www.cs.umd.edu/hcil/VASTchallenge08/awards.html.

The pattern of voice calls reflects a fictitious, controversial socio-political movement and incorporates some unusual dynamic activity. Across 400 cell phone users, over a 10 day period, the data relevant to our experiment consist of a complete set of time-stamped calls. For each of the 9834 calls, we have IDs for the send and receive nodes, a start time in hours/minutes and a duration in seconds. We will refer to the bandwidth of a node as the aggregate number of seconds for which the ID is active as a sender or receiver. Among the extra information supplied by the competition designers was the strong suggestion that the node with ID of 200 was the 'ringleader' in a key community. Based on analyses submitted by challenge teams, we believe
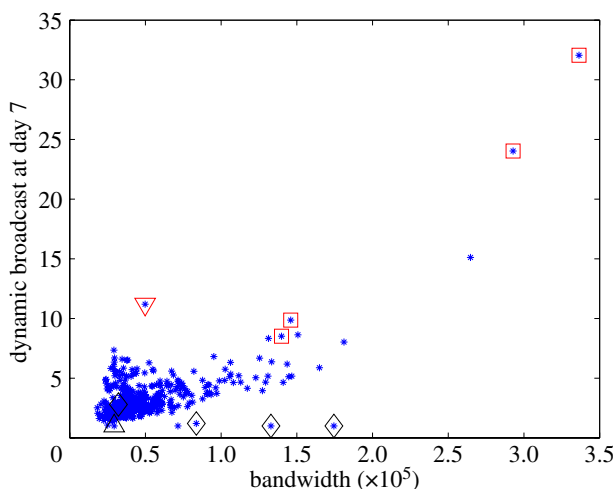
Figure 5. Voice call data, end of day 6: for each node we show dynamic broadcast measure on vertical axis and bandwidth on horizontal axis. The ringleader is marked with a downward pointing triangle and the other four inner circle nodes are marked with squares. After day 6, these individuals are believed to change phone, the ringleader's subsequent ID is marked with an upward pointing triangle, and the other inner circle IDs with diamonds. (Online version in colour.)

that this ringleader controls a tightly knit subnetwork involving nodes with IDs 1, 2, 3 and 5. However, from day 7 onwards, these individuals appear to change their phones: ID 200 changes to 300, and the others change to 306, 309, 360 and 392. Our aim is to show how the ODE (3.3) is useful for dealing with this type of dynamic network.

In our test, we took $A(t)$ to be symmetric, so $A_{ij}(t) = A_{ji}(t) = 1$ if nodes i and j are conversing at time t, measured in seconds. We used $b = 1/(60 \times 60 \times 24) \approx 1.2 \times 10^{-5}$, corresponding to a time downweighting of e^{-1} per day, and a comparable value of $a = 10^{-4}$ for the edge attenuation parameter. We solved the ODE (3.3) numerically with Matlab's `ode23` code, which discretizes the time interval transparently, based on its built-in error control algorithm. We specified absolute and relative error tolerances of 10^{-4}. More precisely, to improve efficiency, we replaced the matrix logarithm $\log(I - aA(t))$ in (3.3) with the expansion $aA(t) - a^2 A(t)^2/2 + a^3 A(t)^3/3 - a^4 A(t)^4/4 + a^5 A(t)^5/5$. Visually unchanged results were found when we increased the number of terms in the expansion to 6 and 7.

Two key findings from our tests were that

   (A) without using any information about the existence of an inner circle, the dynamic broadcast/receive measures (3.4) flag up these key nodes as being highly influential, even when they are not active in terms of overall bandwidth,
   (B) given the IDs of the inner circle, the running centrality measures reveal the change in the nature of the network that occurs during day 7.

To illustrate point (A), figure 5 takes the data up to the end of day 6, and scatter plots bandwidth against dynamic broadcast, **b**. (Results for dynamic receive, **r**, are very similar.) Here the ring leader, 200, is marked with a downward pointing triangle and the related nodes, 1, 2, 3 and 5 are marked with squares. We have also marked the ringleader's follow-on ID, 300, with an upward pointing triangle, and those for the other members, 306, 309, 360 and 392, with diamonds.

We see that the key nodes for days 1–6 are much more dominant in terms of dynamic broadcast than overall bandwidth. In particular, the ringleader node has a very modest bandwidth but ranks fourth out of 400 for broadcast communicability.
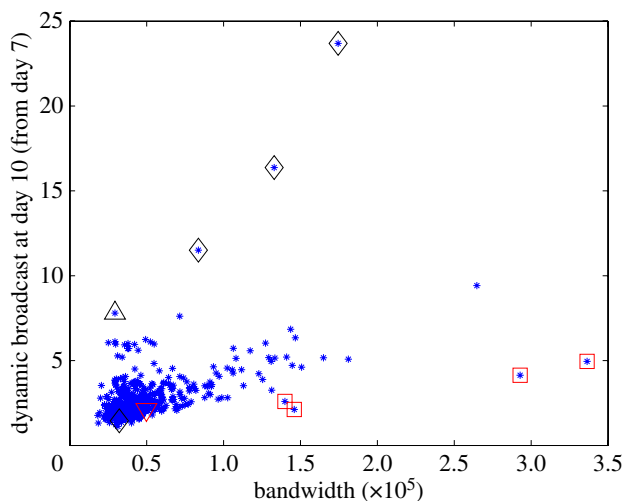
Figure 6. As for figure 5, using data from day 7 to day 10. (Online version in colour.)
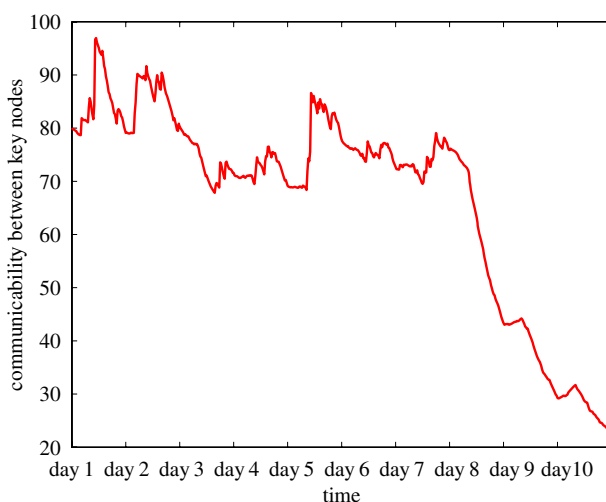

Figure 7. Voice call data: dynamic communicability between the five key nodes (vertical axis) as a function of time (horizontal axis). (Online version in colour.)

Figure 6 shows the same information for the data arising from days 7 to 10. We see again that the dynamic broadcast score does a much better job than bandwidth of revealing the inner circle. In particular, the new ID of the ringleader, marked with an upward pointing triangle, has a very low overall bandwidth but ranks the fifth highest for broadcast centrality. The 'old' IDs from the inner circle, marked with squares, continue to have high bandwidth, but the dynamic broadcast score indicates that they are no longer central players.

To illustrate point (B), figure 7 shows the dynamic communicability between the original IDs of the five key players, 200, 1, 2, 3 and 5, as a function of time. Here, at each time point, we show the average pairwise broadcast plus receive communicability between each pair of nodes in this group scaled by the average pairwise communicability between all pairs of nodes. We see that this running measure is able to track the change in the nature of the network around day 7, when these individuals switch to different IDs.

7. Conclusion

In summary, this work has presented what we believe to be the first attempt to pass to the continuum limit in dynamic network centrality. This dramatically enhances the existing snapshot-based paradigm in terms of both data-driven simulation and theoretical analysis. Our framework fits naturally into the context of online or digital recording of human interactions. The ODE setting conveniently avoids the need to discretize the network data into pre-defined time windows—an approach that can introduce inaccuracies and computational inefficiencies. By defining a continuous-time dynamical system, we can simulate with off-the-shelf, state-of-the-art, adaptive numerical ODE solvers, so that time discretization is performed 'under the hood' and in a manner that automatically handles considerations of accuracy and efficiency. In particular, in this way, we can deal adaptively with dramatic changes in network behaviour. We also showed how to downweight information over time in the new ODE framework, so that a running summary of centrality can be updated in real time without the need to store, or take account of, all previous interaction history. This has immediate applications to the issue of ranking nodes [28], detecting virality [29] and making time-sensitive strategic decisions [30]. Furthermore, the ODEs derived here can feed into the development of new network models where the evolution of centrality is coupled to the evolution of topology, and hence they can contribute to modelling, prediction and hypothesis testing.

References

1. Grindrod P, Higham DJ. 2010 Evolving graphs: dynamical models, inverse problems and propagation. *Proc. R. Soc. A* **466**, 753–770. (doi:10.1098/rspa.2009.0456)
2. Holme P, Saramäki J. 2012 Temporal networks. *Phys. Rep.* **519**, 97–125. (doi:10.1016/j.physrep.2012.03.001)
3. Barrat A, Fernandez B, Lin KK, Young L-S. 2013 Modeling temporal networks using random itineraries. *Phys. Rev. Lett.* **110**, 158702. (doi:10.1103/PhysRevLett.110.158702)
4. Lambiotte R, Sinatra R, Delvenne JC, Evans TS, Barahona M, Latora V. 2011 Flow graphs: interweaving dynamics and structure. *Phys. Rev. E* **84**, 017102. (doi:10.1103/PhysRevE.84.017102)
5. Stehlé J *et al*. 2011 Simulation of an SEIR infectious disease model on the dynamic contact network of conference attendees. *BMC Med.* **9**, 87. (doi:10.1186/1741-7015-9-87)
6. Grindrod P, Higham DJ, Parsons MC, Estrada E. 2011 Communicability across evolving networks. *Phys. Rev. E* **83**, 046120. (doi:10.1103/PhysRevE.83.046120)
7. Grindrod P, Higham DJ. 2013 A matrix iteration for dynamic network summaries. *SIAM Rev.* **55**, 118–128. (doi:10.1137/110855715)
8. Tang J, Musolesi M, Mascolo C, Latora V, Nicosia V. 2010 Analysing information flows and key mediators through temporal centrality metrics. In *SNS '10: Proc. 3rd Workshop on Social Network Systems, Paris, France, April 2010*, pp. 1–6. New York, NY: ACM.
9. Lentz HHK, Selhorst T, Sokolov IM. 2013 Unfolding accessibility provides a macroscopic approach to temporal networks. *Phys. Rev. Lett.* **110**, 118701. (doi:10.1103/PhysRevLett.110.118701)
10. Mucha PJ, Richardson T, Macon K, Porter MA, Onnela J-P. 2010 Community structure in time-dependent, multiscale, and multiplex networks. *Science* **328**, 876–878. (doi:10.1126/science.1184819)
11. Tang J, Scellato S, Musolesi M, Mascolo C, Latora V. 2010 Small-world behavior in time-varying graphs. *Phys. Rev. E* **81**, 05510. (doi:10.1103/PhysRevE.81.055101)
12. Gómez S, Diaz-Guilera A, Gómez-Gardeñes J, Pérez-Vicente CJ, Moreno Y, Arenas A. 2013 Diffusion dynamics on multiplex networks. *Phys. Rev. Lett.* **110**, 028701. (doi:10.1103/PhysRevLett.110.028701)

13. De Domenico M, Solé-Ribalta A, Cozzo E, Kivelä M, Moreno Y, Porter MA, Gómez S, Arenas A. 2013 Mathematical formulation of multilayer networks. *Phys. Rev. X* **3**, 041022. (doi:10.1103/PhysRevLett.110.028701)

14. Dunlavy DM, Kolda TG, Acar E. 2011 Temporal link prediction using matrix and tensor factorizations. *ACM Transactions on Knowledge Discovery from Data* **5**, 10. (doi:10.1145/1921632.1921636)

15. Barabási A-L. 2005 The origin of bursts and heavy tails in human dynamics. *Nature* **435**, 207–211. (doi:10.1038/nature03459)

16. Marvel SA, Kleinberg J, Kleinberg RD, Strogatz SH. 2011 Continuous-time model of structural balance. *Proc. Natl Acad. Sci.* USA **108**, 1771–1776. (doi:10.1073/pnas.1013213108)

17. Traag VA, Van Dooren P, De Leenheer P. 2013 Dynamical models explaining social balance and evolution of cooperation *PLoS ONE* **8**, e60063. (doi:10.1371/journal.pone.0060063)

18. Grindrod P, Higham DJ, Parsons MC. 2012 Bistability through triadic closure. *Internet Math.* **8**, 402–423. (doi:10.1080/15427951.2012.714718)

19. Newman MEJ. 2010 *Networks: an introduction*. Oxford, UK: Oxford University Press.

20. Wasserman S, Faust K. 1994 *Social network analysis: methods and applications*. Cambridge, UK: Cambridge University Press.

21. Katz L. 1953 A new index derived from sociometric data analysis. *Psychometrika* **18**, 39–43. (doi:10.1007/BF02289026)

22. Gleave E, Welser HT, Lento TM, Smith MA. 2009 A conceptual and operational definition of 'social role' in online community. In *Proc. 42nd Hawaii Int. Conf. on System Sciences, Waikoloa Village, HI, January 2009*, pp. 1–11. Los Alamitos, CA: IEEE Computer Society.

23. Huffaker D. 2010 Dimensions of leadership and social influence in online communities. *Hum. Commun. Res.* **36**, 593–617. (doi:10.1111/j.1468-2958.2010.01390.x)

24. Mantzaris AV, Higham DJ. 2012 A model for dynamic communicators. *Eur. J. Appl. Math.* **23**, 659–668. (doi:10.1017/S0956792512000186)

25. Higham NJ. 2008 *Functions of matrices: theory and computation*. Philadelphia, PA: Society for Industrial and Applied Mathematics.

26. Bilton N. 2011 The lifespan of a link. *The New York Times*. http://bits.blogs.nytimes.com/2011/09/07/the-lifespan-of-a-link/?_php=true&_type=blogs&_r=0 (accessed 27 February 2014)

27. Grinstein GG, Plaisant C, Laskowski SJ, O'Connell T, Scholtz J, Whiting MA. Vast 2008 challenge: introducing mini-challenges. In *IEEE VAST, IEEE*, vol. 2008, pp. 195–196. New York, NY: IEEE.

28. Blumm N, Ghoshal G, Forró Z, Schich M, Bianconi G, Bouchaud JP, Barabási AL. 2012 Dynamics of ranking processes in complex systems. *Phys. Rev. Lett.* **109**, 128701. (doi:10.1103/PhysRevLett.109.128701)

29. Iribarren JL, Moro E. 2011 Branching dynamics of viral information spreading. *Phys. Rev. E* **84**, 046116. (doi:10.1103/PhysRevE.84.046116)

30. Borge-Holthoefer J, Baňos RA, Gonzáez-Bailón S, Moreno Y. 2013 Cascading behaviour in complex socio-technical networks. *J. Complex Netw.* **1**, 3–24. (doi:10.1093/comnet/cnt006)