Models and Algorithms for Time-Dependent Networks

Des Higham
Department of Mathematics and Statistics
University of Strathclyde

Joint work with
Fiona Ainley (Bloom), Ernesto Estrada (Strathclyde),
Peter Grindrod (Oxford), Peter Laflin (Bloom),
Alex Mantzaris (Strathclyde/Bloom),
Amanda Otley (Bloom), Mark Parsons (Reading)
From Twitter’s homepage:

*Businesses can also use Twitter to listen and gather market intelligence and insights*
Oreo at Superbowl 2013

Power out? No problem.
pic.twitter.com/dnQ7pOgC

YOU CAN STILL DUNK IN THE DARK
Audi did this slightly earlier

Sending some LEDs to the @MBUSA Superdome right now...

9,708 RETWEETS  3,217 FAVORITES
Influence over time

![Graph showing influence over time for Audi, Oreo, Pepsi, and Budweiser. The x-axis represents time from 0100 GMT to 0230 GMT, and the y-axis represents influence ranging from 0 to 200. The graph indicates fluctuations in influence throughout the time period.]
Large-scale, dynamic interaction data arises in, e.g.,

- social media
- on-line business
- text, email, voicemail
- epidemiology
- neuroscience
Some Algorithmic Challenges

- identify **communities**
- discover other types of **structure**
- **categorize** the roles of individuals
- **target** key players
Matrix Computation for Centrality

Unweighted, with \( N \) nodes

Adjacency matrix \( A \)

\[
(A^2)_{ij} := \sum_{p=1}^{N} a_{ip}a_{pj}
\]

counts paths of length two from node \( i \) to node \( j \)

Generally, \((A^k)_{ij}\) counts the number of walks of length \( k \) from node \( i \) to node \( j \). Then

\[
(A + A^2/(2!) + A^3/(3!) + \cdots)_{ij} \quad \text{leads to} \quad (e^A)
\]

and

\[
(A + \alpha^2 A^2 + \alpha^3 A^3 + \cdots)_{ij} \quad \text{leads to} \quad (I - \alpha A)^{-1}
\]

to quantify how well information can flow from \( i \) to \( j \)
Time Ordered Sequence of Networks

Lack of symmetry caused by time's arrow
Aggregation would also overestimate the spread
Lack of symmetry caused by time's arrow

Aggregation would also overestimate the spread
Time Ordered Sequence of Networks

Time 1

Time 2

Time 3

Lack of symmetry caused by time's arrow
Aggregation would also overestimate the spread
Lack of symmetry caused by time’s arrow
Lack of symmetry caused by time’s arrow

Aggregation would also overestimate the spread
Dynamic Walks

Time points $t_0 < t_1 < t_2 < \cdots < t_M$

Adjacency matrices $A^0$, $A^1$, $A^2$, $\ldots$, $A^M$

**Dynamic walk of length** $w$ from node $i_1$ to node $i_{w+1}$: sequence of times $t_{r_1} \leq t_{r_2} \leq \cdots \leq t_{r_w}$ and a sequence of edges $i_1 \leftrightarrow i_2$, $i_2 \leftrightarrow i_3$, $\ldots$, $i_w \leftrightarrow i_{w+1}$, such that $i_m \leftrightarrow i_{m+1}$ exists at time $t_{r_m}$

(Several variations are possible)

Use this to define **centrality** of a node, generalizing Katz (1953)
Key observation: the matrix product


has \( i, j \) element that counts the number of dynamic walks of length \( w \) from node \( i \) to node \( j \), where the \( m \)th step takes place at time \( t_{rm} \)

Keep track of all such walks and discount by \( \alpha^w \)


This is achieved by

\[ Q := (I - \alpha A[0])^{-1} (I - \alpha A[1])^{-1} \ldots (I - \alpha A[M])^{-1} \]

Then \( Q_{ij} \) is our overall summary of how well information can be passed from node \( i \) to node \( j \)
Dynamic Centrality

We will call the row and column sums

\[ \sum_{k=1}^{N} Q_{nk} \quad \& \quad \sum_{k=1}^{N} Q_{kn} \]

the broadcast and receive communicabilities

- generalizes Katz centrality in social networks
- involves sparse linear solves
- captures the asymmetry through non-commutativity of matrix multiplication
Listen to **tweets** containing the phrases city break, cheap holiday, travel, insurance, cheap flight **plus two brand names**

From 17 June 2012 at 14:41 to 18 June at 12:41

0.5 Million Tweeters/Followers
Twitter account with fourth highest out degree is a very poor dynamic broadcaster

Closer inspection ⇒ an automated process

Five social media experts were given the Twitter data and asked to rank the accounts according to importance

We found that dynamic centrality measures are hard to distinguish from human experts
**Motivation:** News goes stale, messages become irrelevant, viruses mutate, ... *old information is less important*

The algorithm can be generalized naturally to

\[
S[k] = (I + e^{-b \Delta t_k} S[k-1]) \left( I - \alpha A[k] \right)^{-1} - I
\]

Here, \((S[k])_{ij}\) counts the number of dynamic walks from \(i\) to \(j\) up to time \(t_k\), scaled by

- a factor \(\alpha^w\) for **dynamic walks of length** \(w\)
- a factor \(e^{-bt}\) for **walks that begin** \(t\) time units ago

We have a new parameter, \(b\):

- \(b = 0\) is the previous algorithm
- \(b = \infty\) is Katz on the current network
Continuous Time


Discretizing in time and binarizing is convenient, but
- $\Delta t$ too large can **overlook** or **smear** events
- $\Delta t$ too small may **introduce redundant computations**
  and give a **false impression** of accuracy

So use $A(t)$ over continuous time

Define communicability by taking $\Delta t \to 0$ limit

**Key idea**: use the scaling $(I - \alpha A(k\Delta t))^{-\Delta t}$

**Justification**: $A$ constant $\Rightarrow$ correct behaviour

\[
(I - \alpha A)^{-\frac{1}{2}\Delta t} (I - \alpha A)^{-\frac{1}{2}\Delta t} = (I - \alpha A)^{-\Delta t}
\]
Matrix ODE for the evolution of pairwise communicability

\[ U'(t) = -b(U(t) - I) - U(t) \log \left( I - \alpha A(t) \right) \]

Here:
- \( U(t)_{ij} \) is the communicability between nodes \( i \) and \( j \)
- \( \log \) is the **matrix logarithm**
Voice calls from a controversial socio-political movement

- 400 IDs
- 10 days
- each call: time stamp (hour/minute) plus duration (sec)
- a key inner circle exists with a single ringleader

Based on published challenge entries

- inner circle of five and ringleader have been identified
- these five people change IDs at end of day 6

We take $A(t)$ to be symmetric and use MATLAB’s ode23
Unsupervised Analysis: up to Day 7

Dynamic broadcast at Day 7

Bandwidth

Dynamic broadcast at Day 7

Bandwidth
850,000 tweets in 4 hours. Up to 24,000 tweets per minute.
Adebayor: volume of tweets

North London Derby: 17 November 2012 - Adebayor Tweets

Tweets per minute vs. Minutes after 1200 GMT
Adebayor: sentiment across time
Adebayor: sentiment weighted by influence

Tweets about adebayor

- adebayor scores
- adebayor sent off
- half time, fans blaming adebayor for the score
- game coming to an end, arsenal fans thanking adebayor
- stat appears that adebayor has been sent off 4 times, 3 of them by Howard Webb
Current Challenge: Spikes in activity

Estimate **who will be busy** during a spike
Explain the **10-20 minute half-life**
Current Challenge: Discovering Hierarchy

In Minus Out Degree Ordering

0 50 100 150 200 250 300 350 400
Part 2 will look at **mathematical models** for dynamic networks
First: static network models

Model: typically a probabilistic rule for joining pairs of nodes.

Useful if we want to know whether a given network has some unusual structure.
Part 2: Network Models

First: static network models

**Model**: typically a **probabilistic rule** for joining pairs of nodes.

Useful if we want to know whether a given network has some **unusual structure**
Influential Random Network Models

Erdös-Rényi, 1959, $\text{ER}(p)$

Watts-Strogatz, 1998

Barabási-Albert, 1999
Geometric, e.g., Penrose 2003
Prob. $f(|i-j|)$
Fig. 1. Spy plots showing nonzero patterns for a $100 \times 100$ sample from each of the nine models.

3. MODELS
In this section, we give brief descriptions of the nine models implemented, and show how to use the corresponding MATLAB functions. In each case, the output argument, $A$, is a sparse, symmetric, zero-diagonal matrix of dimension $n$, with $n$ being the first of the input arguments. The remaining input arguments take default
Some modelling challenges...

- understand mechanisms
- **calibrate** parameters and compare models
- characterize **business as usual**
- **forecast** future behaviour
- simulate **what-if** scenarios
Motivated by digital human interaction,

- fixed set of nodes
- edges that may appear and disappear over time

Further, assume that over discrete time points

- dynamics are Markovian
- each edge can be treated independently
Friends of friends become friends

Given $N$ people, “friending” and “unfriending”
Let $A[k]$ be the adjacency matrix at time $k$

Edge death probability is a constant $\omega \in (0, 1)$

Edge birth probability between nodes $i$ and $j$ given by

$$\delta + \epsilon \left( (A[k])^2 \right)_{ij}$$

where $0 < \delta \ll 1$ and $0 < \epsilon(N - 2) < 1 - \delta$

Consider $N = 100, \omega = 0.01, \epsilon = 5 \times 10^{-4}, \delta = 4 \times 10^{-4}$
Triadic closure: start with ER(0.3)

Edge density at time 750 is 0.712
Triadic closure: start with ER(0.15)

Edge density at time 750 is 0.051
Mean field analysis for $\delta + \epsilon \left( (A[k])^2 \right)_{ij}$

**Ergodicity and symmetry** $\Rightarrow$ Erdős-Rényi limit: every edge present with probability $p^*$

Heuristic **mean field** approach: insert the ansatz “$A[k] = \text{ER}(p_k)$” into the model to obtain

$$p_{k+1} = (1 - \omega)p_k + (1 - p_k)(\delta + \epsilon(N - 2)p_k^2)$$

Generically: **three real roots**

**Two are stable, one is unstable**

$N = 100, \omega = 0.01, \epsilon = 5 \times 10^{-4}, \delta = 4 \times 10^{-4}$
$N = 100$, $\omega = 0.01$, $\epsilon = 5 \times 10^{-4}$, $\delta = 4 \times 10^{-4}$

Stable fixed points at $p^* = 0.049$ and $p^* = 0.721$
Unstable fixed point at $p^* = 0.229$
Mean-field vs. simulation from ER(0.4)
Four simulations from ER(0.23)

Stable fixed points: 0.049 & 0.721
Unstable: 0.229

Given model parameters, we can compute the probability of observing the data: **likelihood**

Tests on synthetic data show that we can correctly infer the triadic closure effect

**Wealink** data from Hu and Wang, Phys. Lett. A, 2009. 26 Million time stamps, over 841 days 0.25 Million nodes No edge death
Fig. 11: Density of the Wealink online social network as a function of time in seconds.
Evidence for $\epsilon > 0$ in this dynamic network

Fig. 13: Log likelihood of the triadic closure model as a function of triad closure strength, $\epsilon$. The basal edge birth is fixed at $9.95 \times 10^{-8}$. The x-axis shows the log base 10 values used for $\epsilon$ in the search for the largest loglikelihood.

Although likelihood-based calibration and model comparison is conceptually straightforward for stochastic, Markov chain based, models of the type used here, the fundamental technique remains difficult.
Other Edge Dynamics?

- higher order **motifs**
- **centrality** (rich get richer)
- **hierarchy** (chain of command)
- **gravity** models (relative location)

Could also couple topology with state of node

- **homophily** (like-minded nodes associate)
- **heterophily** (opposite of homophily)
- **social balance** (my enemy’s enemy is my friend)

In principle, **model comparison** can be used to determine which one best fits the data
What’s Next?

Algorithms
- efficient communicability computation
- detection of communities over time

Modelling/Calibration/Prediction
- explain 20 minute half-life for Twitter spike-decay
- compare social science hypotheses (triadic closure, homophily, heterophily, hierarchy)

Thanks to
CAPITA
EPSRC/RCUK Digital Economy programme
Leverhulme Trust
National Physical Laboratory
Royal Society
Wolfson Foundation
EPSRC/Strathclyde Impact Acceleration Account
Bloom Agency