Social and Technological Network Analysis

Lecture 10: Temporal Social Network Metrics and Applications

John Tang

UNIVERSITY OF CAMBRIDGE

In This Lecture

- In this lecture we will show how the concepts of message dissemination and epidemics have been applied to study properties and processes in real networks.

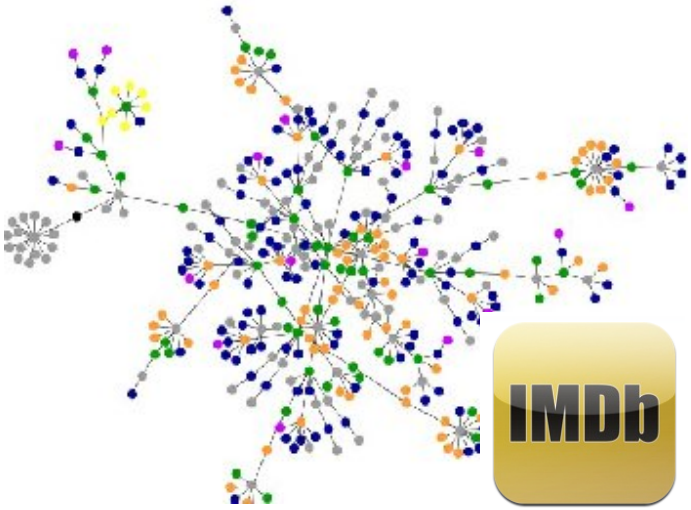
UNIVERSITY OF CAMBRIDGE

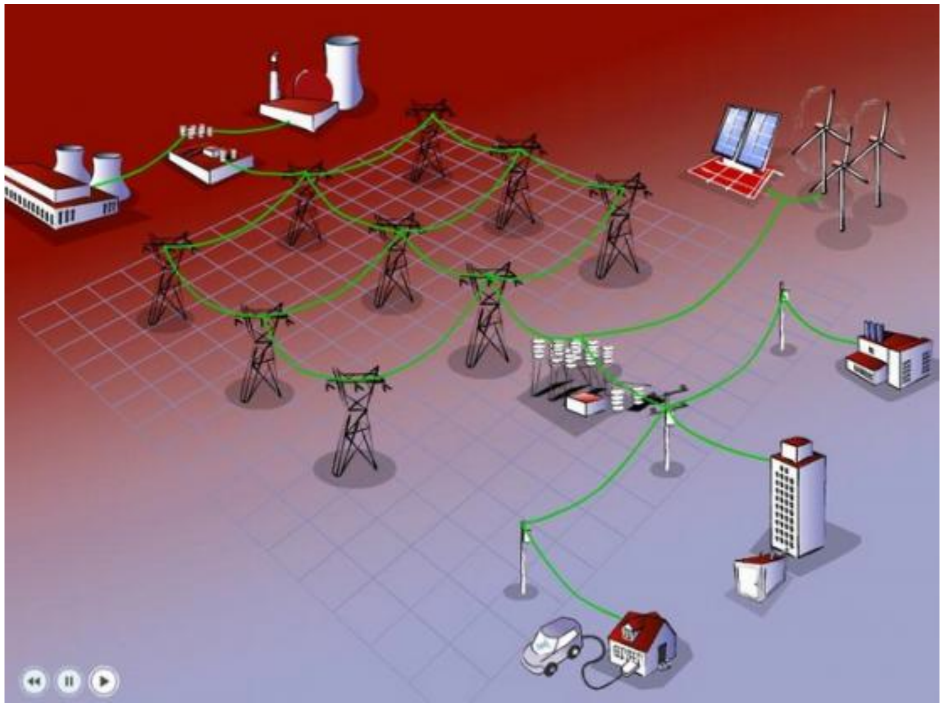
Why Temporal Social Network

- Most of the analysis we have seen has been done on aggregated network graphs
- Time has not been kept into account by the metrics
- Why does this matter?

UNIVERSITY OF CAMBRIDGE

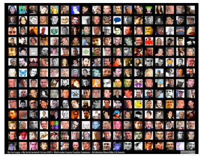
Empirical Networks

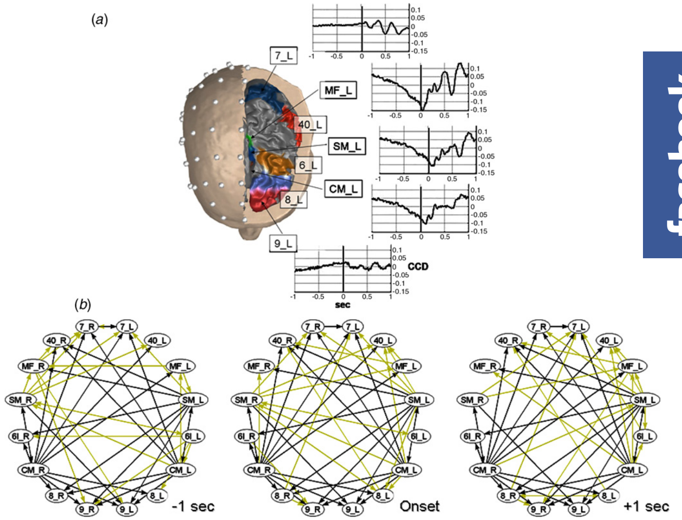




UNIVERSITY OF CAMBRIDGE

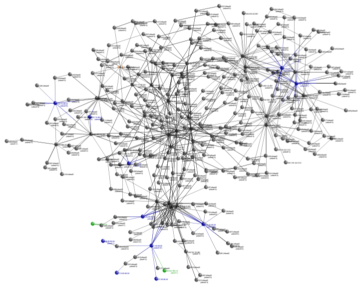
More Empirical Networks



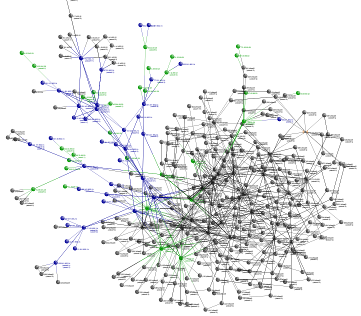


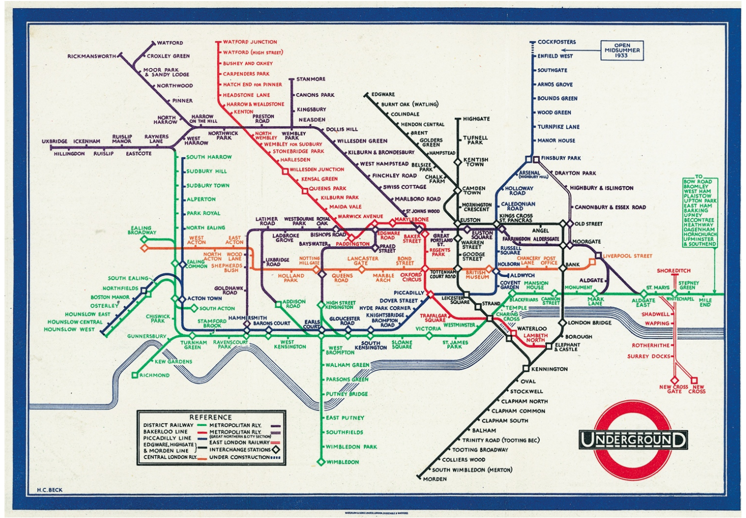
Libya on the Internet

11 Aug 2011



22 Aug 2011

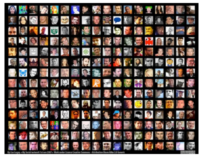




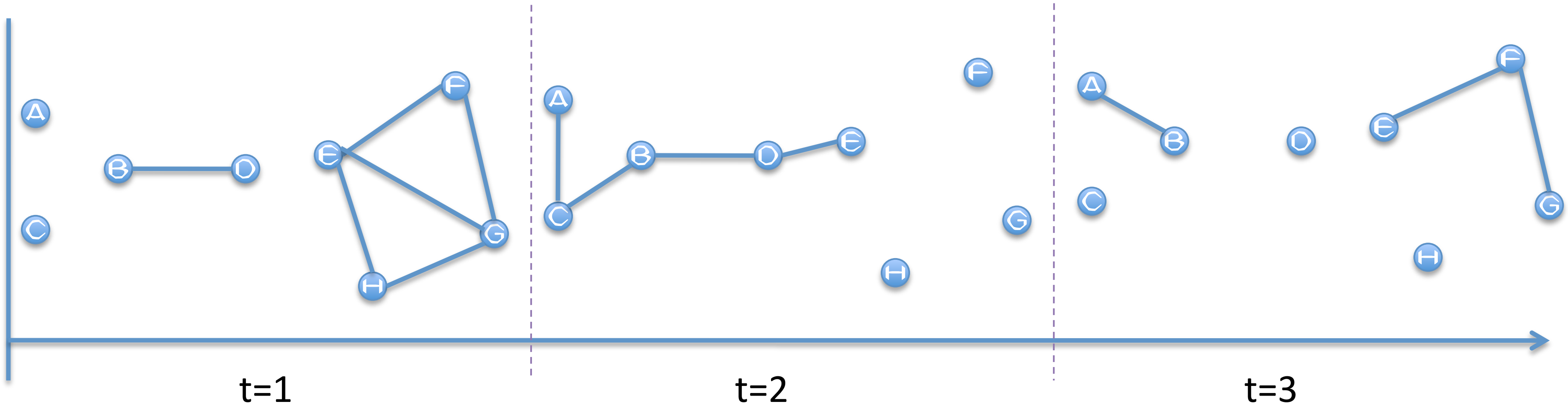
Time in networks

- Timestamps
  - e.g. Facebook: friends added and removed over time
- Duration
  - e.g. Spending time with friends
- Frequency
  - e.g. Friends, colleagues, strangers
- Time-order
  - e.g. Timetables in public transport systems

UNIVERSITY OF CAMBRIDGE

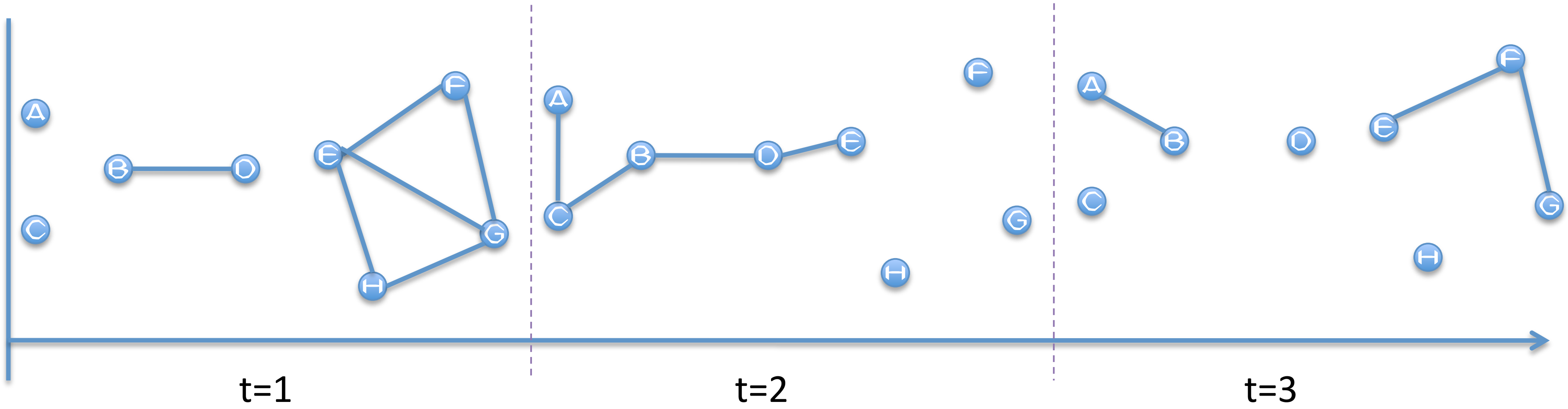


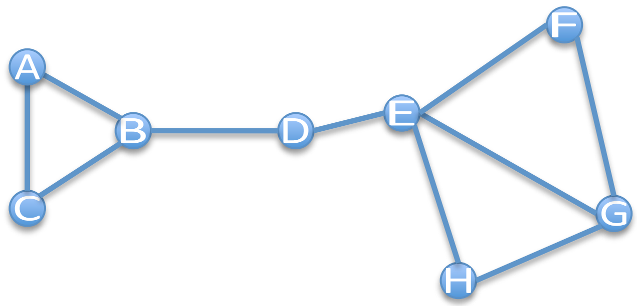
Temporal Graph



UNIVERSITY OF CAMBRIDGE

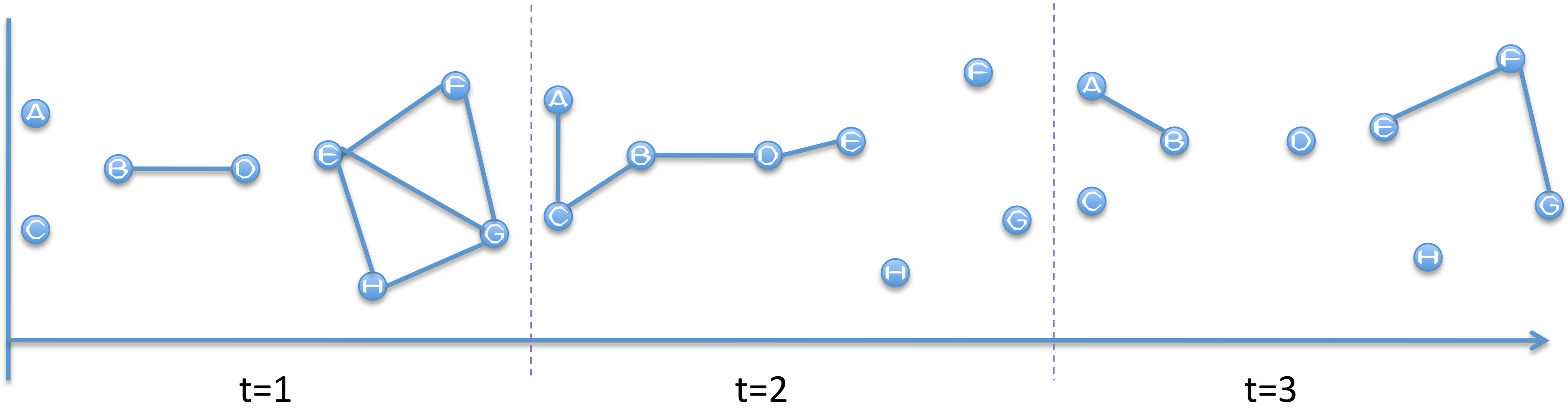
Temporal Graph



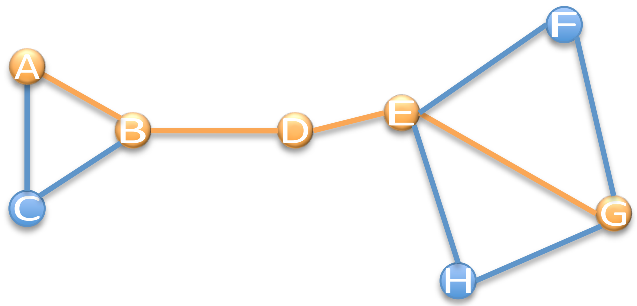


UNIVERSITY OF CAMBRIDGE

Temporal Graph

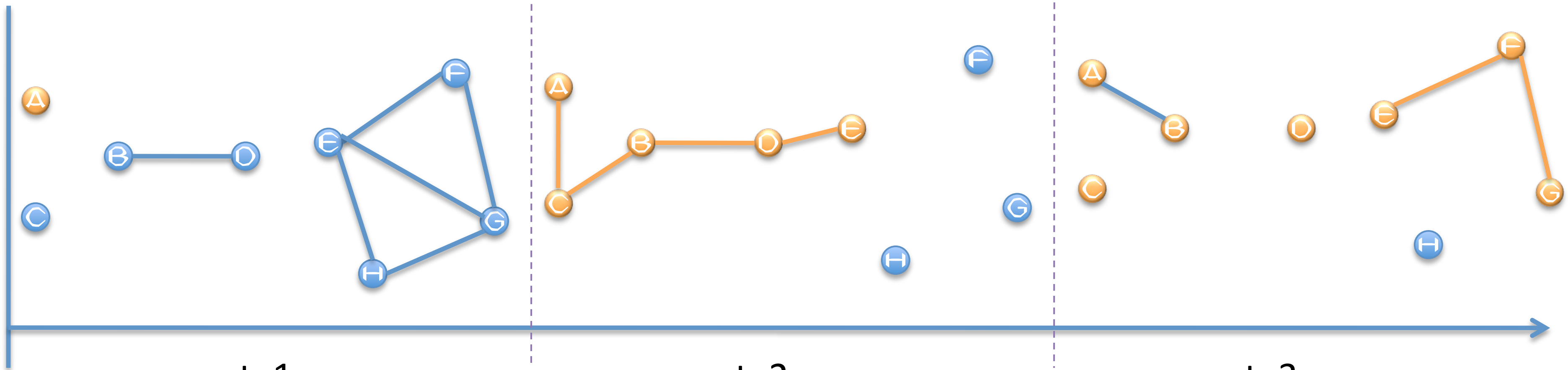


- Static
  - Shortest path (A,G) = [A,B,D,E,G]
  - Shortest path length (A,G) = 4 hops

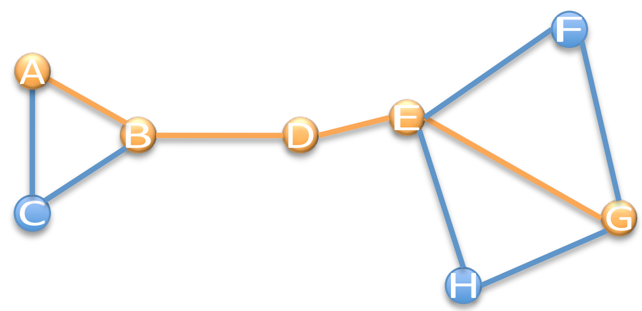


UNIVERSITY OF CAMBRIDGE

Temporal Graph



- Static
  - Shortest path (A,G) = [A,B,D,E,G]
  - Shortest path length (A,G) = 4 hops
- Temporal
  - Shortest path (A,G) = [A,C,B,D,E,F,G]
  - Shortest path length (A,G) = 6 hops
  - Time=3 seconds



UNIVERSITY OF CAMBRIDGE

Temporal Measures

- d_{ij} — Shortest Temporal Path Duration
- d^*_{ij} — Number of Hops in shortest temporal path
- $E_{ij} = \frac{1}{d_{ij}}$ — Temporal Efficiency of communication

UNIVERSITY OF CAMBRIDGE

Temporal Measures

- Average Temporal Path Duration

$$L = \frac{1}{N(N-1)} \sum_{ij} d_{ij}$$

- Average Temporal Path Hops

$$L^* = \frac{1}{N(N-1)} \sum_{ij} d^*_{ij}$$

- Average Temporal Efficiency

$$E_{glob} = \frac{1}{N(N-1)} \sum_{ij} E_{ij}$$

UNIVERSITY OF CAMBRIDGE

Does it really matter?

- Infocom 2005 conference environment
- Bluetooth colocation scans
- 5 Minute Windows
- Measure 24 hours starting 12am

					Static		Temporal		
Day	N	<k>	Activity	Contacts	L	Eglob	L*	L	Eglob
1	37	25.73	6pm-12pm	3668	1.291	0.856	4.090	19h 39m	0.003
2	39	28.31	12am-12pm	8357	1.269	0.870	4.556	9h 6m	0.024
3	38	22.32	12am-12pm	4217	1.420	0.798	4.003	10h 32m	0.018
4	39	21.44	12am-5pm	3024	1.444	0.781	4.705	9h 55m	0.013

UNIVERSITY OF CAMBRIDGE

"Temporal Distance Metrics for Social Network Analysis", ACM WOSN '09

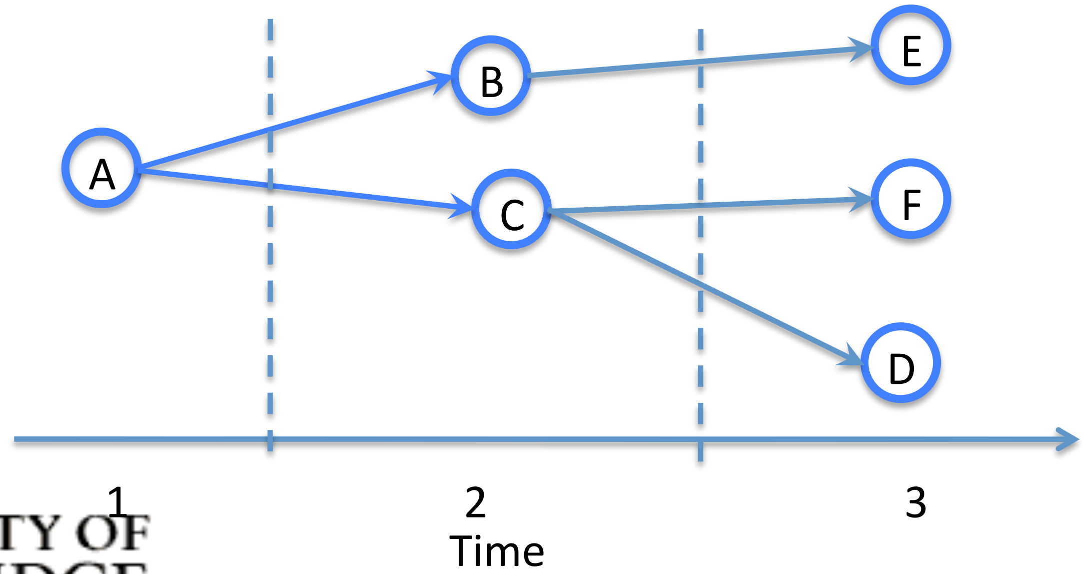
Temporal Centrality Measures

- Static Closeness and Betweenness based on *static shortest paths*
- Reformalise *closeness* and *betweenness* with temporal paths:
  - Duration
  - Time Order
  - Frequency

Temporal Closeness

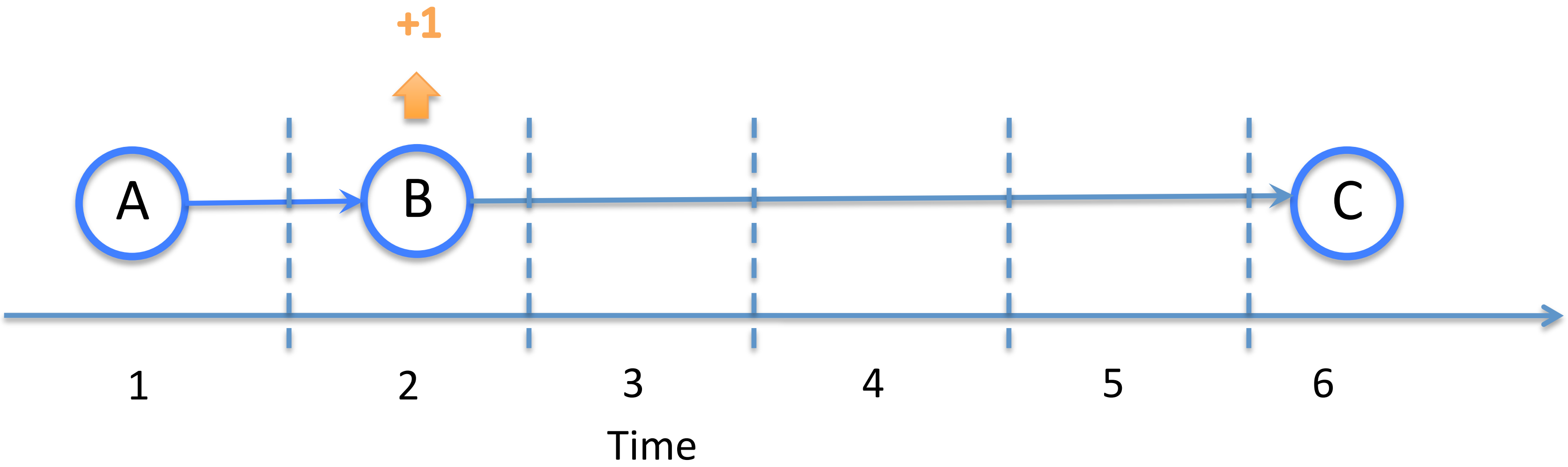
$$C_i = \frac{1}{W(N-1)} \sum_{j \neq i \in V} d_{i,j}$$

$$C_A = \frac{(2+2)+(3+3+3)}{(3*(6-1))} = 0.867$$



Temporal Betweenness

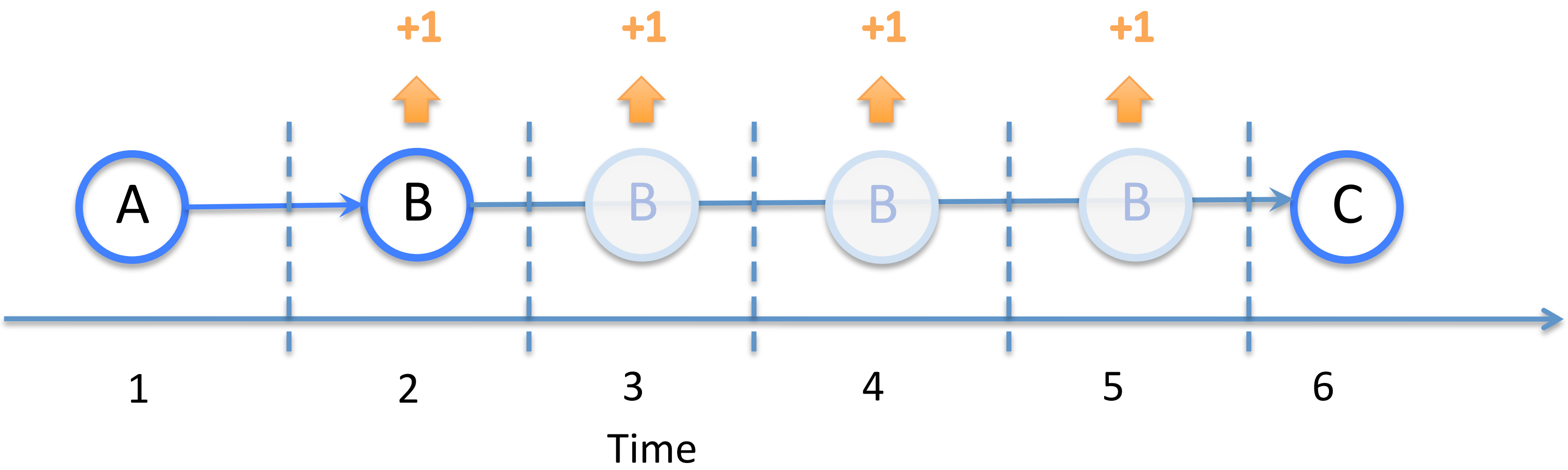
- Using temporal path length



UNIVERSITY OF CAMBRIDGE

Temporal Betweenness

- Take into account **duration**



UNIVERSITY OF CAMBRIDGE

Evaluating Centrality

- Two perspectives:
  - Semantic: known roles of nodes
  - Dynamic Processes: mobile malware containment

Enron in the News

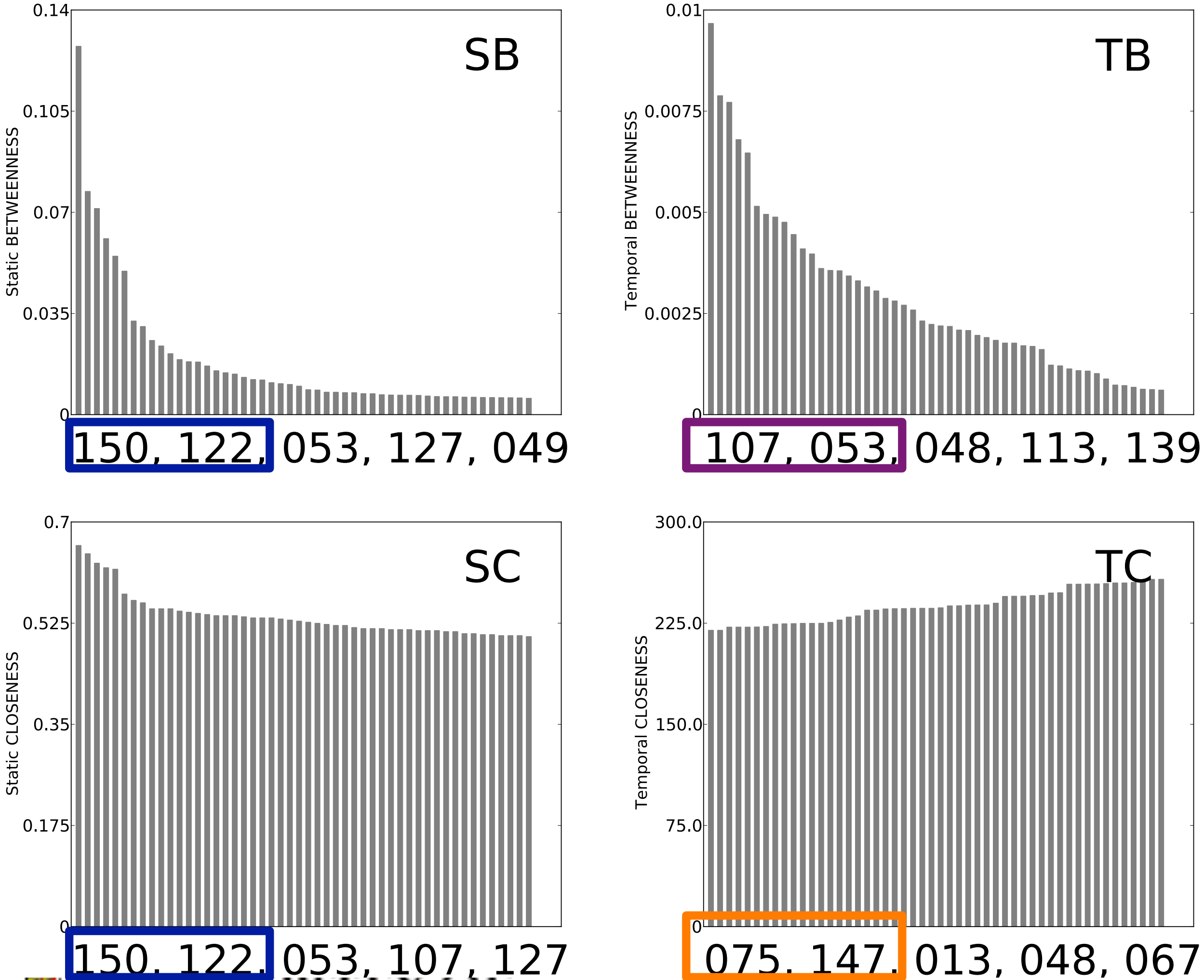
Public Investigation

- Telephone logs
- Documents
- Financials
- Emails

  - 151 user mailboxes
  - May 1999 to Jun 2002
  - 250,000 emails
  - NOT anonymised

UNIVERSITY OF CAMBRIDGE

Semantics



ID	Role
009	(Unknown)
013	Legal
017	Manager
048	Executive
053	Trader
054	President
067	Vice President
073	Trader
075	Director of Trading
107	Trader
122	Managing Director
127	Manager
139	Director
147	Trader
150	Secretary

UNIVERSITY OF CAMBRIDGE

Semantics

ID	Name	Role
9	Stephanie Panus	(Unknown)
13	Marie Heard	Legal
17	Mike Grigsby	Manager
48	Tana Jones	Executive
53	John Lavorato	Trader
54	Greg Whalley	President
67	Sara Shackleton	Vice President
73	Jeff Dasovich	Trader
75	Gerald Nemec	Director of Trading
107	Louise Kitchen	Trader
122	Sally Beck	Managing Director
127	Kenneth Lay	Manager
139	Mary Hain	Director
147	Carol Clair	Trader
150	Liz Taylor	Secretary

CNN.com/LAWCENTER

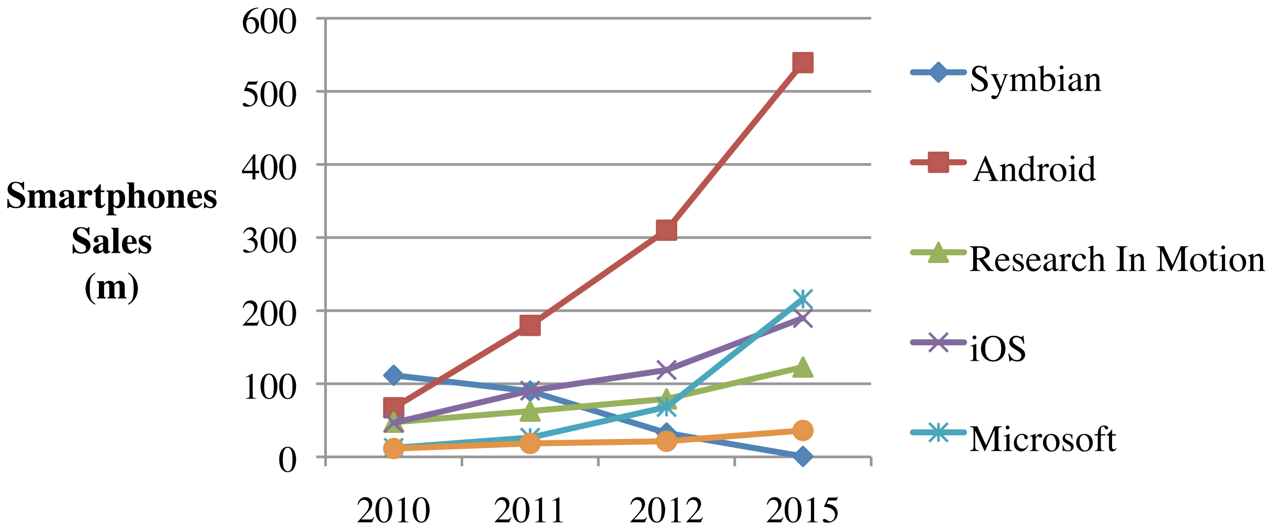
Top bonuses awarded

John Lavorato: $5 million
Louise Kitchen: $2 million
Jeffrey McMahon: $1.5 million
James Fallon: $1.5 million
Raymond Bowen Jr.: $750,000
Mark Haedicke: $750,000
Gary Hickerson: $700,000
Wesley Colwell: $600,000
Richard Dimichele: $600,000

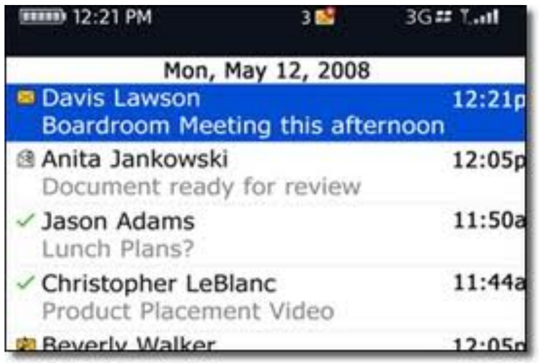
- Big bonuses linked with information mediators

UNIVERSITY OF CAMBRIDGE

On the rise

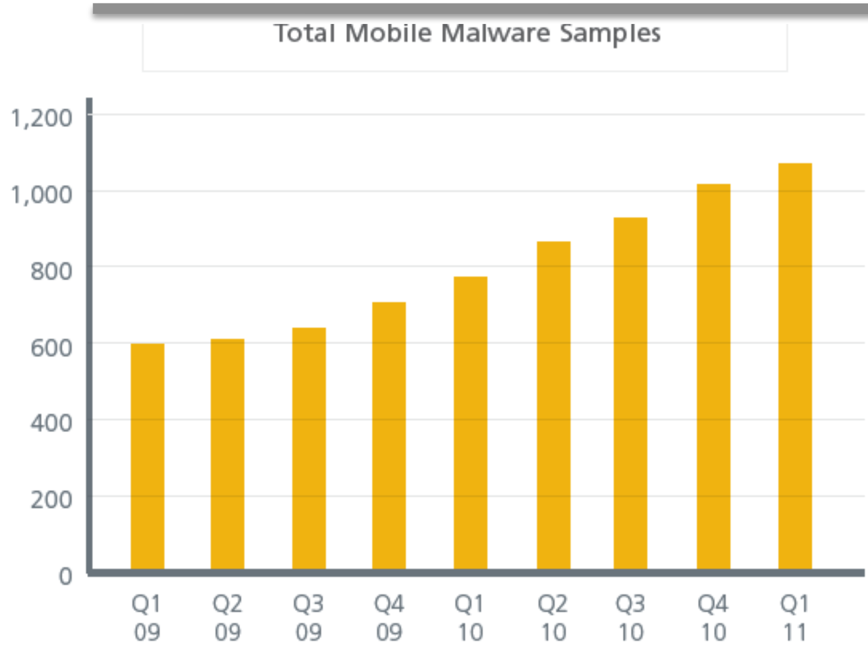


Source: Gartner 2011

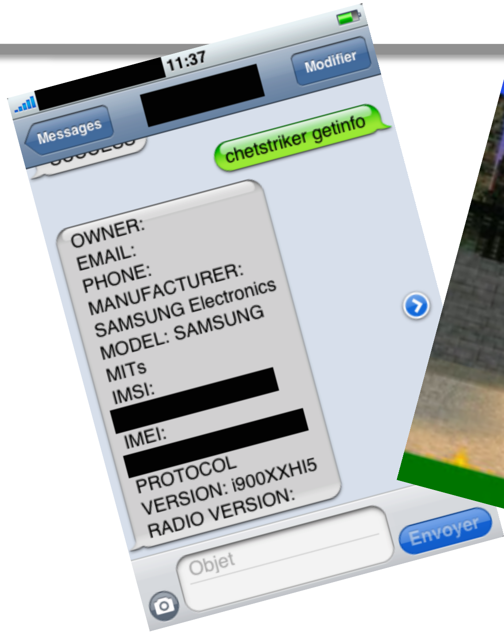


UNIVERSITY OF CAMBRIDGE

On the rise



Source: McAfee Q1 2011



UNIVERSITY OF CAMBRIDGE

Mobile Malware Propagation

- Long Range
  - Sms, mms, email
  - Can be filtered by central service provider
- Short Range
  - Bluetooth, wifi
  - Evades central service provider

UNIVERSITY OF CAMBRIDGE

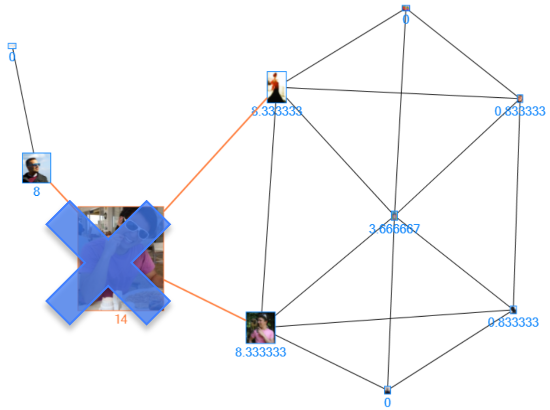
Limitations

- Devices
  - Resource constrained
- Infrastructure
  - Limited bandwidth
- Prioritise Devices using SNA
  - Patch individual devices via nodes with high **Betweenness**
  - Flood patch via nodes with high **Closeness**

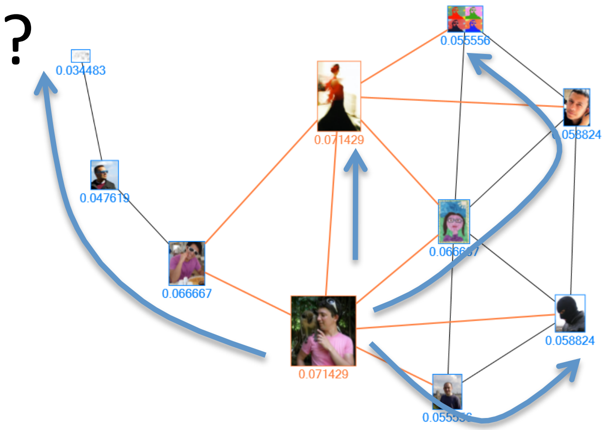
UNIVERSITY OF CAMBRIDGE

Priority Patching Schemes

1. Traditional Patching
   → Can we block path of malware?
   → **Betweenness**

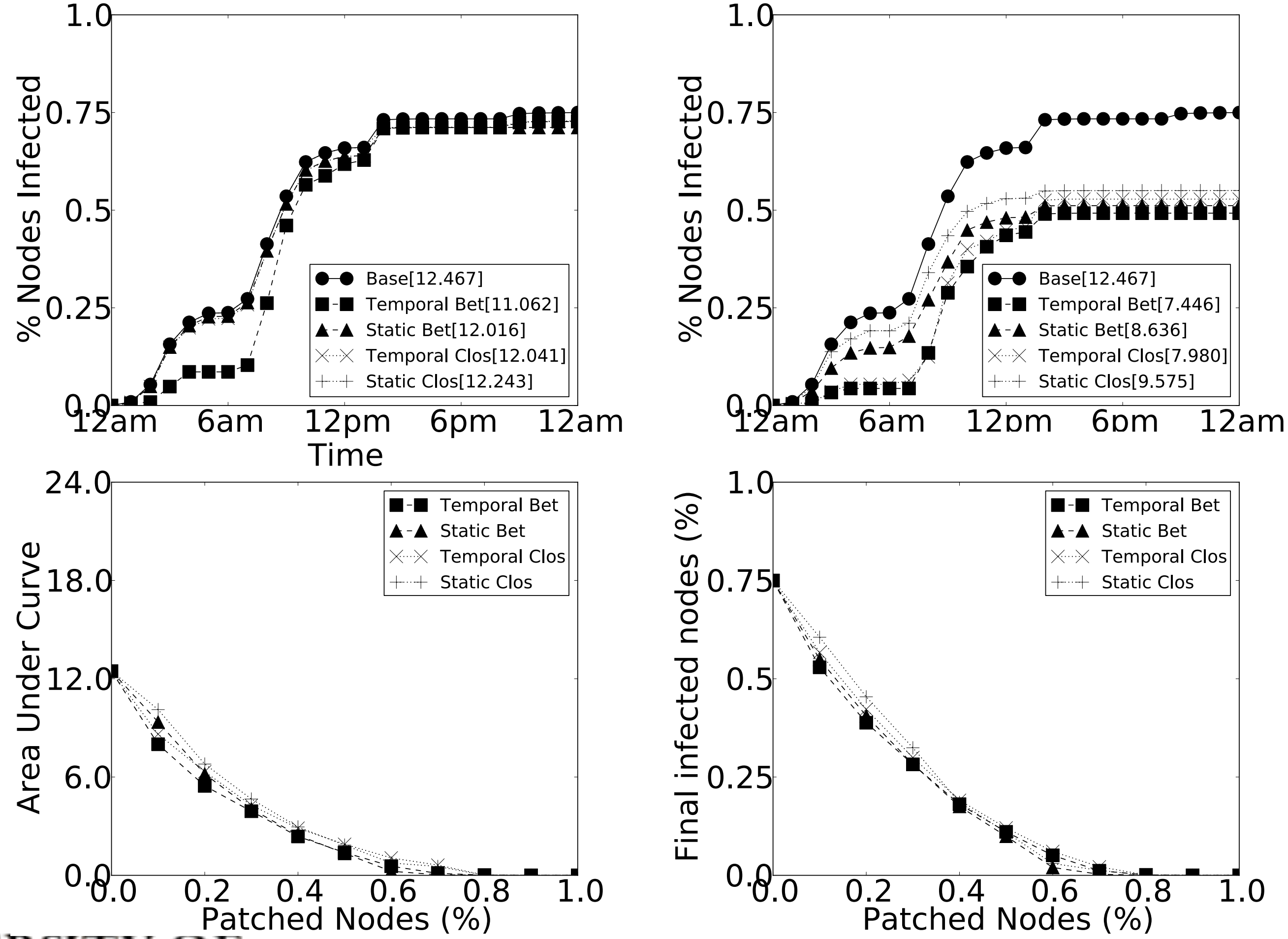


2. Opportunistic Patching
   → Can we compete with malware?
   → **Closeness**

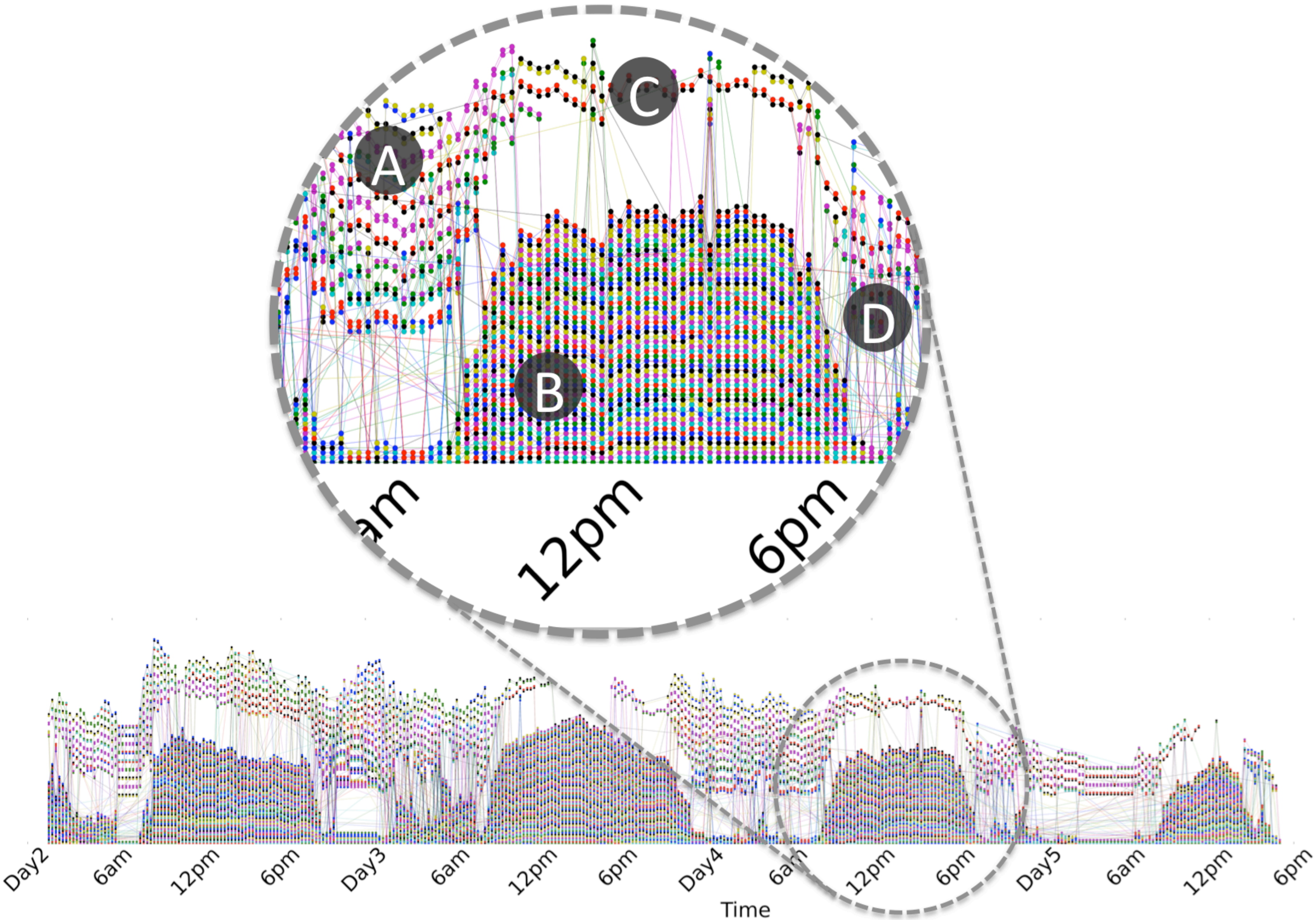


UNIVERSITY OF CAMBRIDGE

Patching Individual Nodes

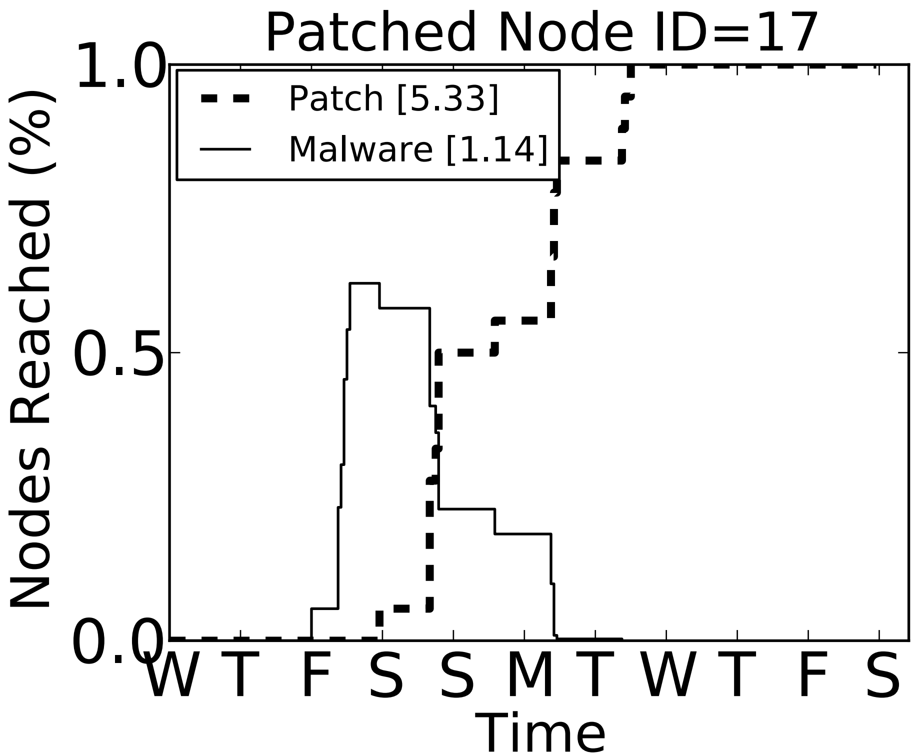


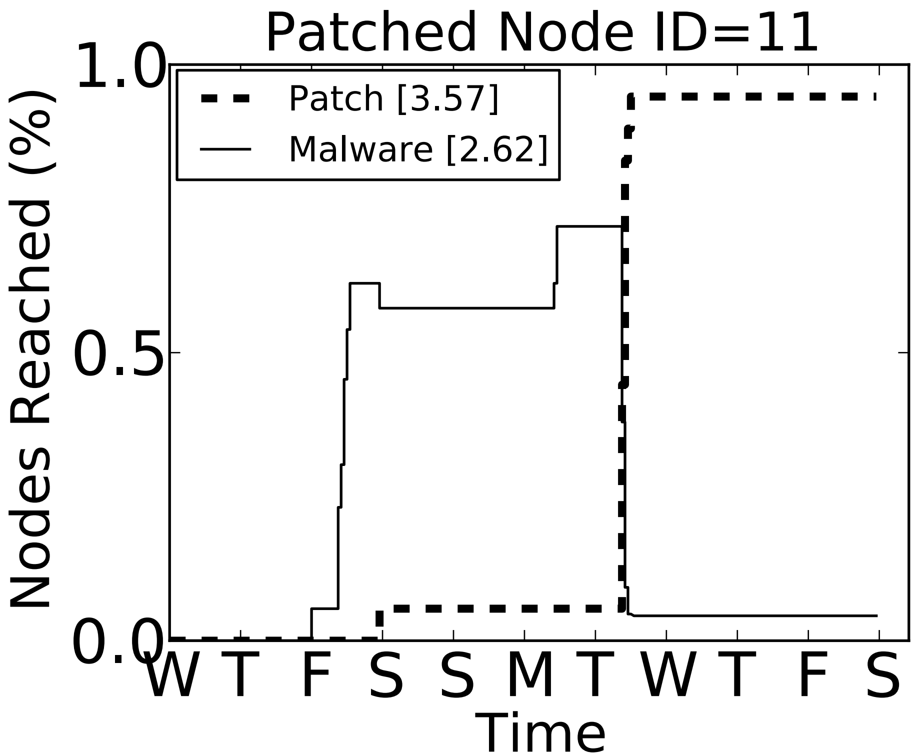
UNIVERSITY OF CAMBRIDGE



A
B
C
D
am
12pm
6pm
Day2
6am
12pm
6pm
Day3
6am
12pm
6pm
Day4
6am
12pm
6pm
Day5
6am
12pm
6pm
Time

Flood Network with Patch

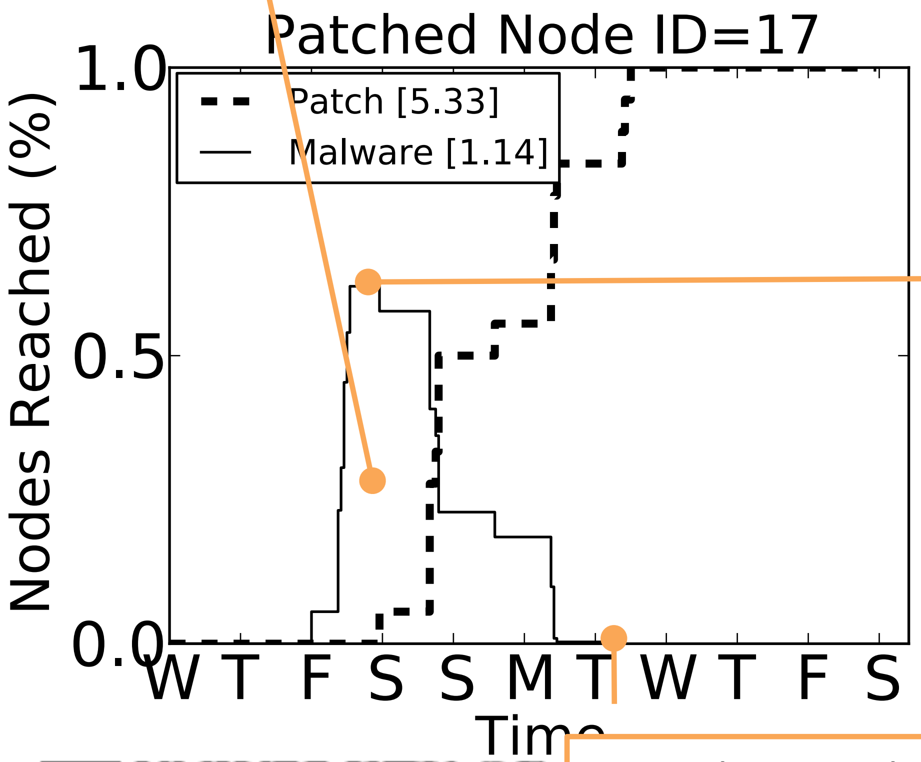




UNIVERSITY OF CAMBRIDGE

Flood Network with Patch

Area under Curve (AUC)

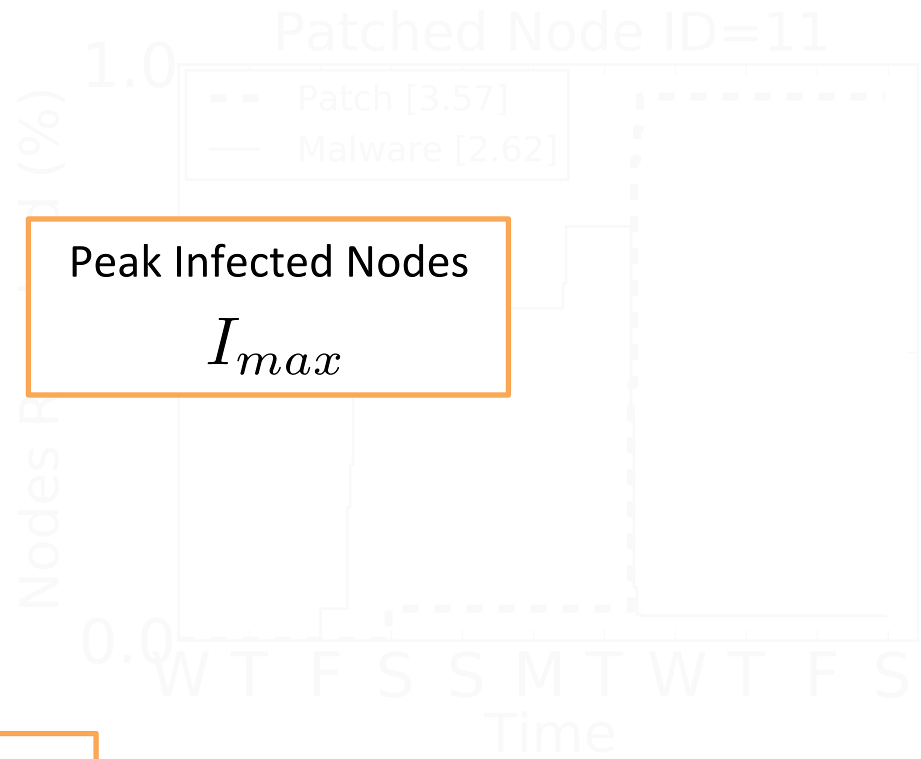


Peak Infected Nodes

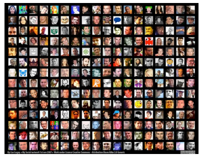
$$I_{max}$$

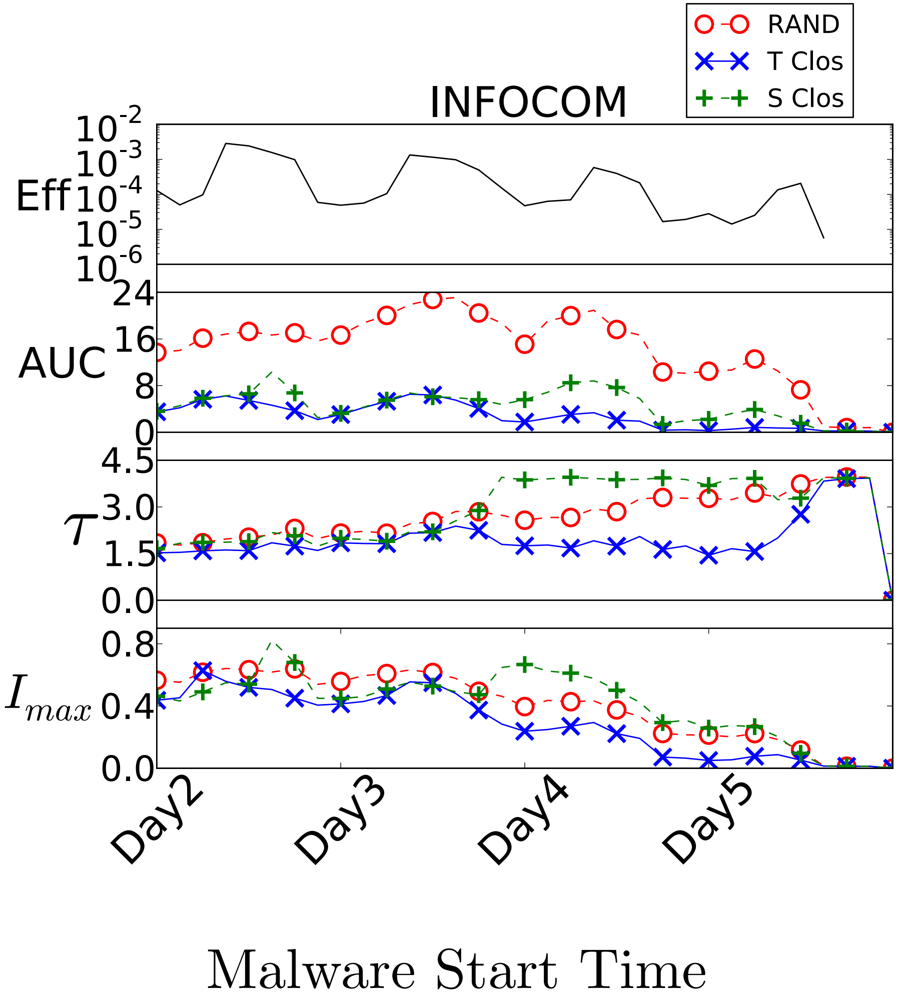
Complete patch time

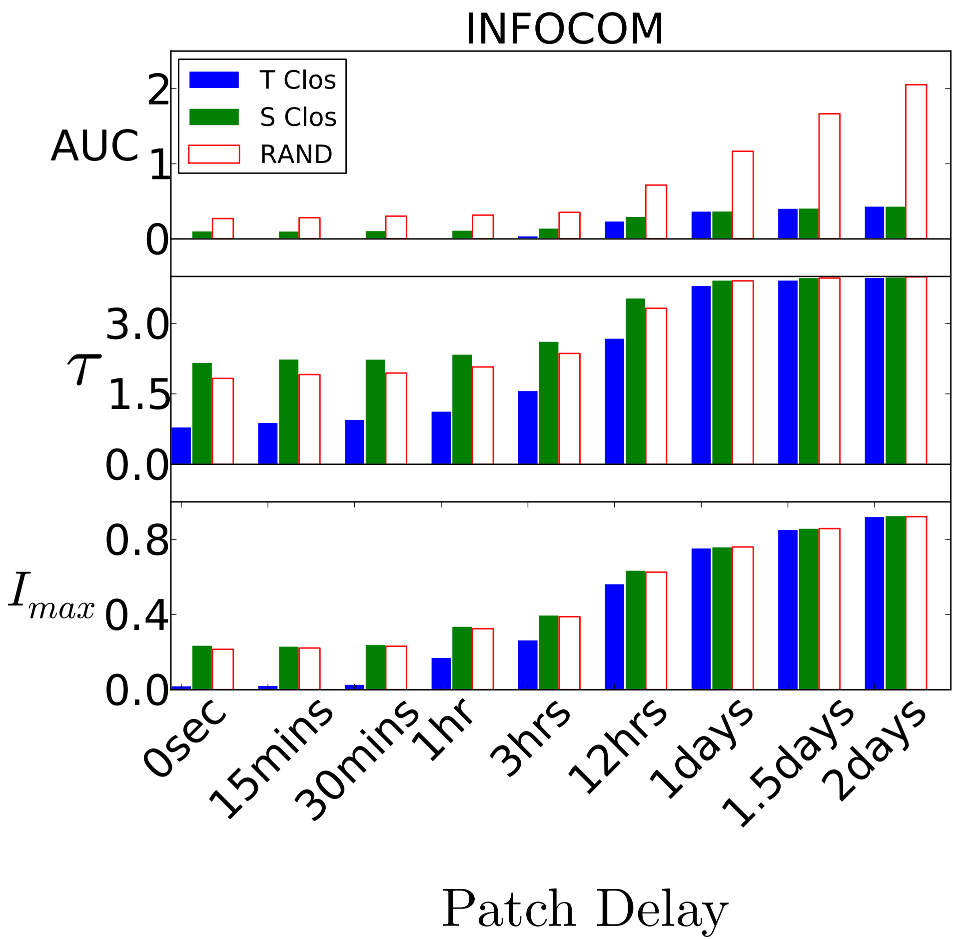
$$\tau$$



2. Opportunistic Patching







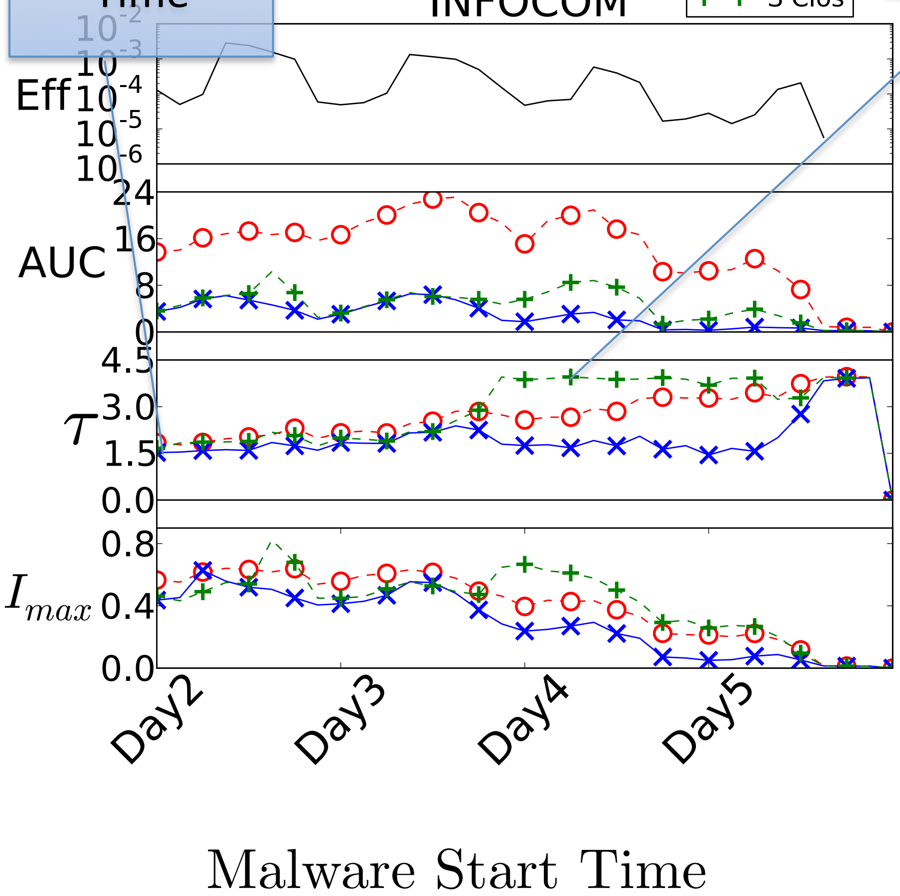
UNIVERSITY OF CAMBRIDGE

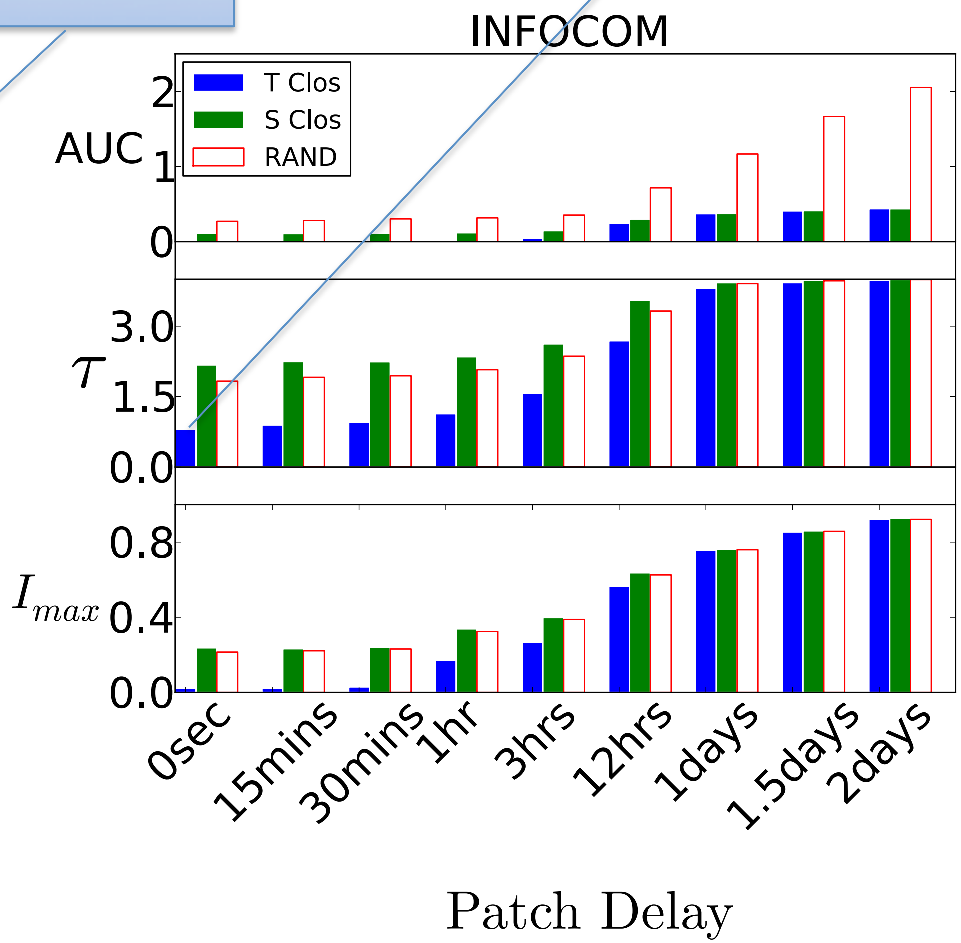
2. Opportunistic Patching

1. Finite Time

2. Static is Poor

3. Temporal is Best



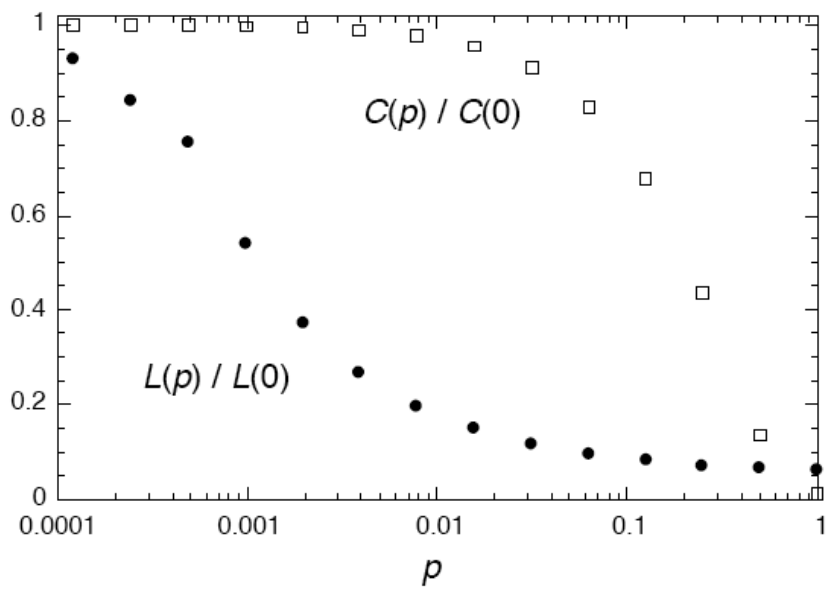


UNIVERSITY OF CAMBRIDGE

"Exploiting Temporal Complex Network Metrics in Mobile Malware Containment", *IEEE WoWMoM 2011*

Static Small World

- Graphs which both are locally clustered but with small average delay
  - High local clustering => Lattice
  - Small average delay => Random



Small World

- Small world = possible to go from one node to another through a ***small number*** of nodes
- Does this hold in time-varying graphs?
  - Temporal small world = possible to go from one node to another ***quickly***
  - *Problem*: Intuition is that slowly evolving graphs should be **slow** for data communication

UNIVERSITY OF CAMBRIDGE

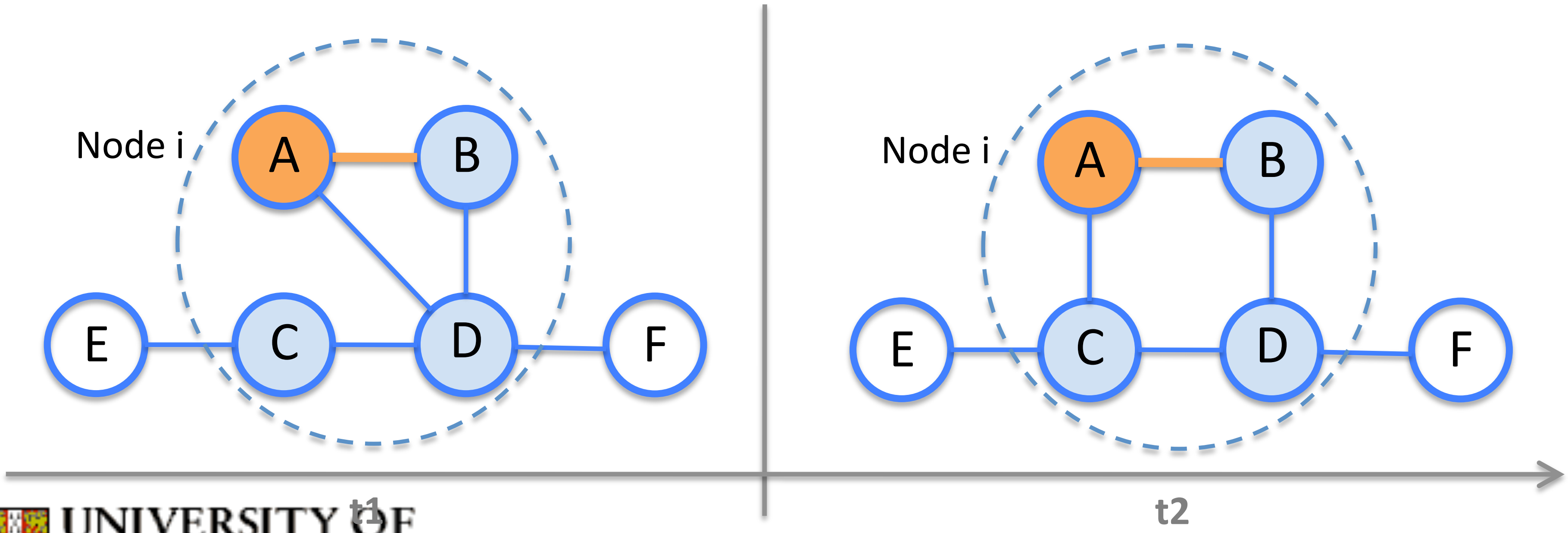
Testing for Temporally Small World

- Measure
  - communication efficiency
    - Temporal shortest path length
  - speed of change
    - Temporal correlation coefficient
    - Measure persistence of links
- Model
  - Recreate a slowly changing and quickly changing temporal graph
  - Brownian motion with prob(jump)

Coefficient of Temporal Clustering

$$C = \frac{\sum_i C_i}{N} \quad C_i = \frac{1}{T-1} \sum_{t=1}^{T-1} \frac{\sum_j a_{ij}(t) a_{ij}(t+1)}{\sqrt{[\sum_j a_{ij}(t)][\sum_j a_{ij}(t+1)]}}$$

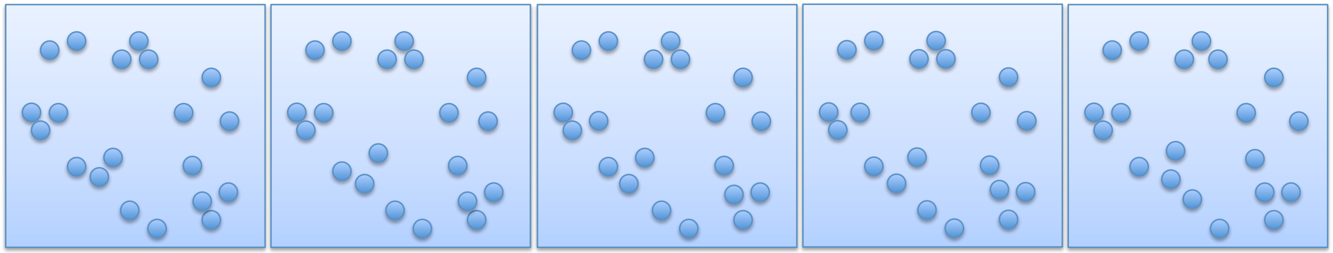
$$C_A = 1/2$$



Temporal SW Model

- N Random Walkers with Prob Jumping P_j

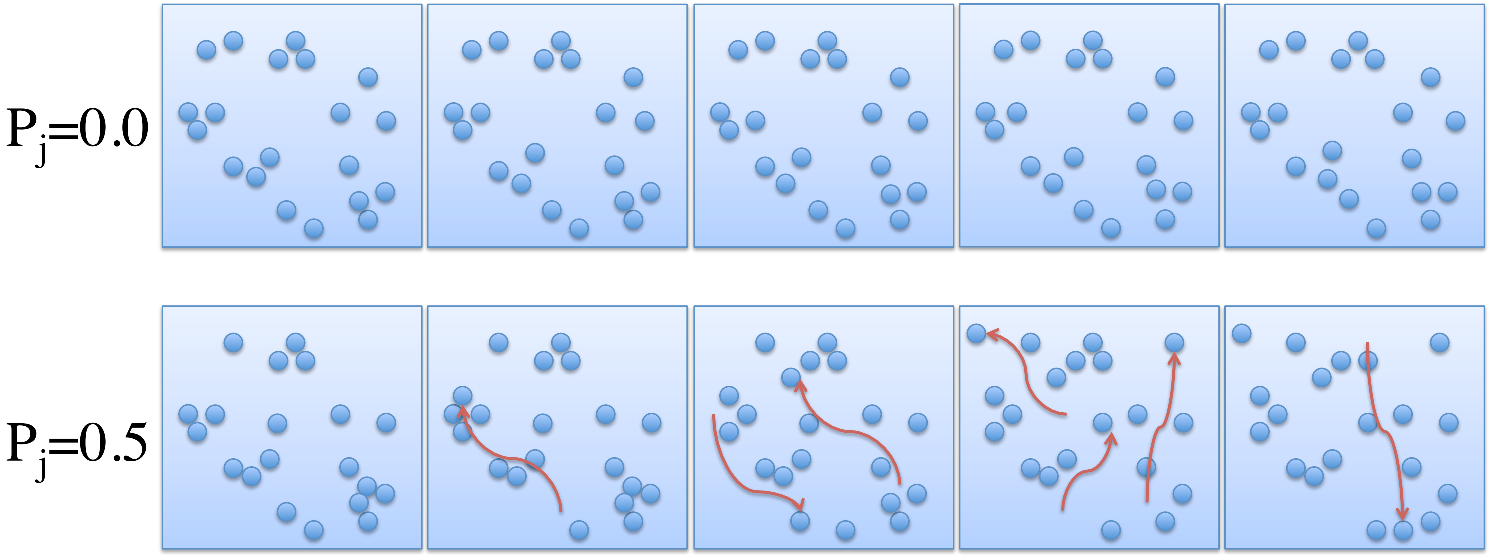
P_j=0.0



UNIVERSITY OF CAMBRIDGE

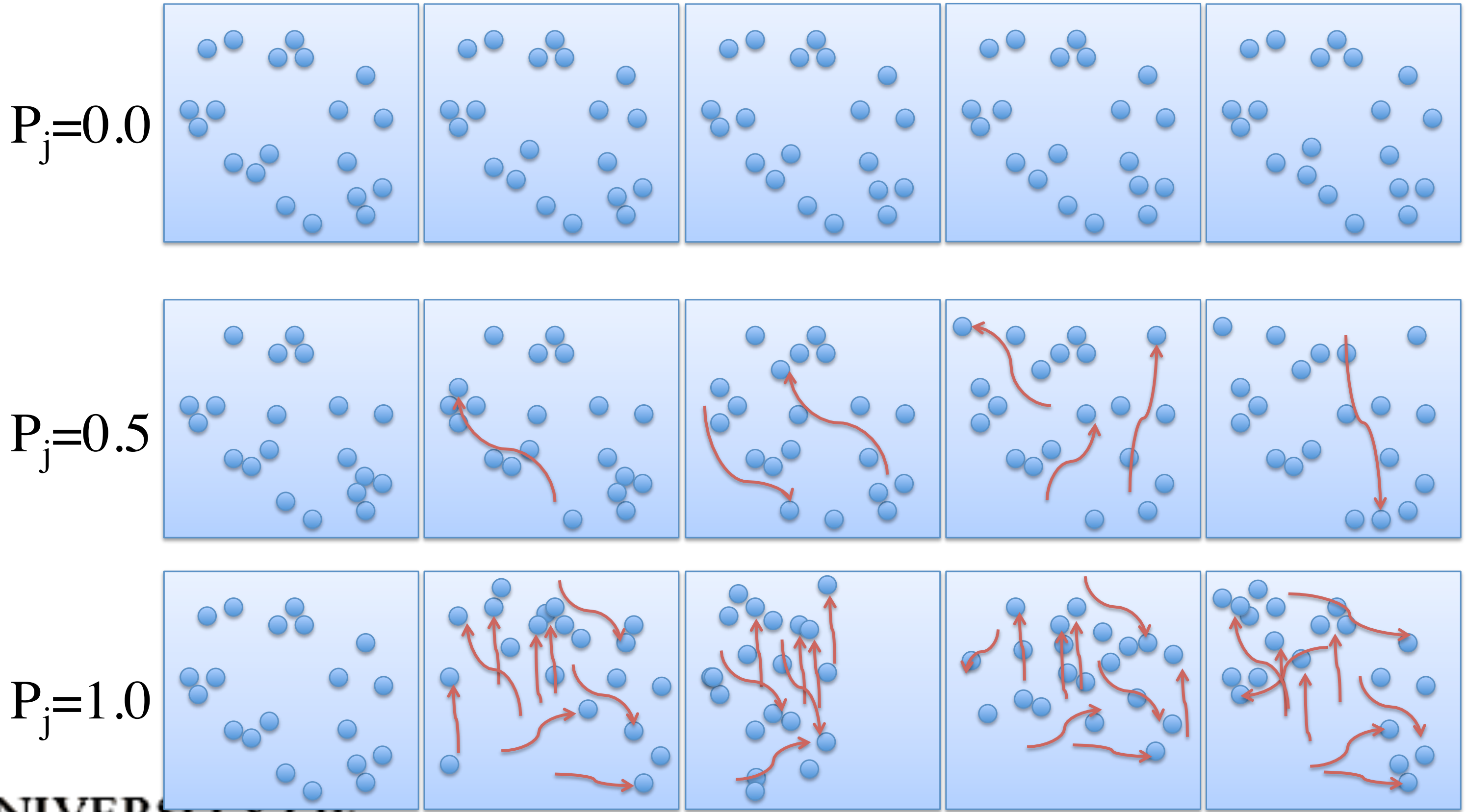
Temporal SW Model

- N Random Walkers with Prob Jumping P_j



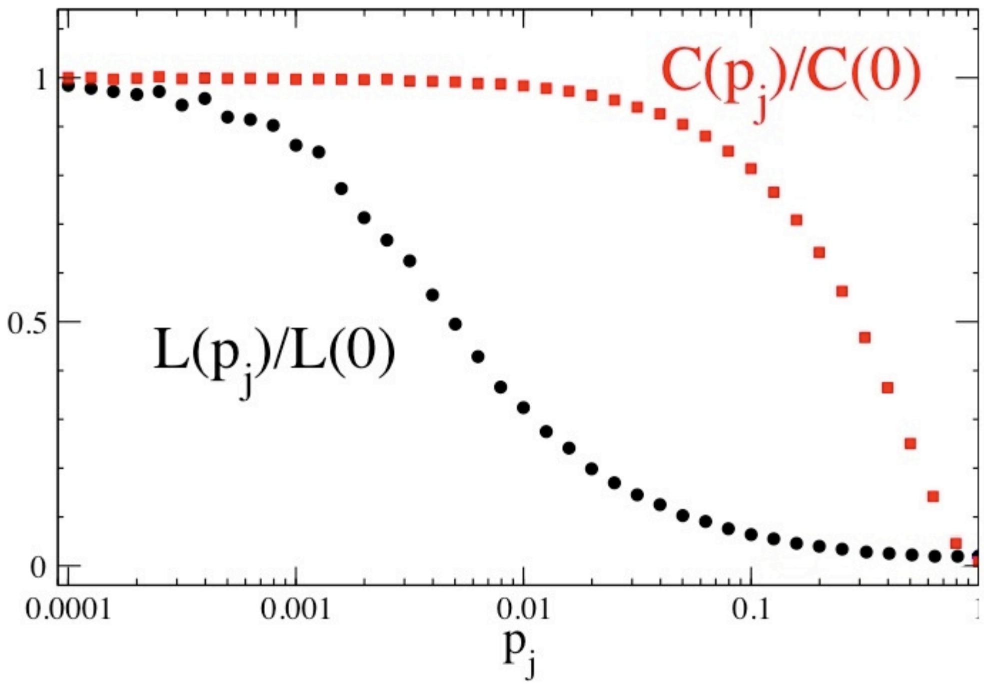
Temporal SW Model

- N Random Walkers with Prob Jumping P_j

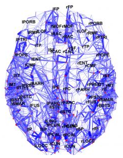


Temporal Small World

- Graphs which evolve slowly over time can still exhibit high communication efficiency
  – Highly temporal-clustering => non-jumping model
  – Low temporal-delay => fully-jumping model



Small-world Behaviour in Real Data



Dataset		C	C^{rand}	L	L^{rand}	E	E^{rand}
Brain network	α	0.44	0.18	3.9 (100%)	4.2 (98%)	0.50	0.48
	β	0.40	0.17	6.0 (94%)	3.6 (92%)	0.41	0.45
	γ	0.48	0.13	12.2 (86%)	8.7 (89%)	0.39	0.37
	δ	0.44	0.17	2.2 (100%)	2.4 (92%)	0.57	0.56
Bluetooth contacts (INFOCOM'06)	d1	0.80	0.44	8.84 (61%)	6.00 (65%)	0.192	0.209
	d2	0.78	0.35	5.04 (87%)	4.01 (88%)	0.293	0.298
	d3	0.81	0.38	9.06 (57%)	6.76 (59%)	0.134	0.141
	d4	0.83	0.39	21.42 (15%)	15.55(22%)	0.019	0.028
facebook (London network)	Mar	0.044	0.007	456	451	0.000183	0.000210
	Jun	0.046	0.006	380	361	0.000047	0.000057
	Sep	0.046	0.006	414	415	0.000058	0.000074
	Dec	0.049	0.006	403	395	0.000047	0.000059

UNIVERSITY OF CAMBRIDGE

"Small world behavior in time varying graphs", *PRE 2011*

Summary

- We have introduced metrics for time varying social networks
- We have shown examples of use on real networks

UNIVERSITY OF CAMBRIDGE

References

- J. Tang, S. Scellato, M. Musolesi, C. Mascolo and V. Latora. **Small-world behavior in time-varying graph** In Physical Review E. Vol. 81 (5), 055101. May 2010.
- J. Tang, M. Musolesi, C. Mascolo, V. Latora, V. Nicosia. **Analysing Information Flows and Key Mediators through Temporal Centrality Metrics**. In Proc. of the 3rd Workshop on Social Network Systems (SNS 2010). Apr 2010.
- J. Tang, M. Musolesi, C. Mascolo and V. Latora. **Temporal Distance Metrics for Social Network Analysis**. In Proc, of the 2nd ACM SIGCOMM Workshop on Online Social Networks (WOSN09). Aug 2009.
- J. Tang, C. Mascolo, M. Musolesi, V. Latora. **Exploiting Temporal Complex Network Metrics in Mobile Malware Containment**. In Proc. of the IEEE 12th International Symposium on a World of Wireless, Mobile and Multimedia Networks (WoWMoM2011). Jun 2011.
- V. Nicosia, J. Tang, M. Musolesi, G. Russo, C. Mascolo, V. Latora. **Components in time-varying graphs** Submitted for Publication. **arXiv:1106.2134**

UNIVERSITY OF CAMBRIDGE