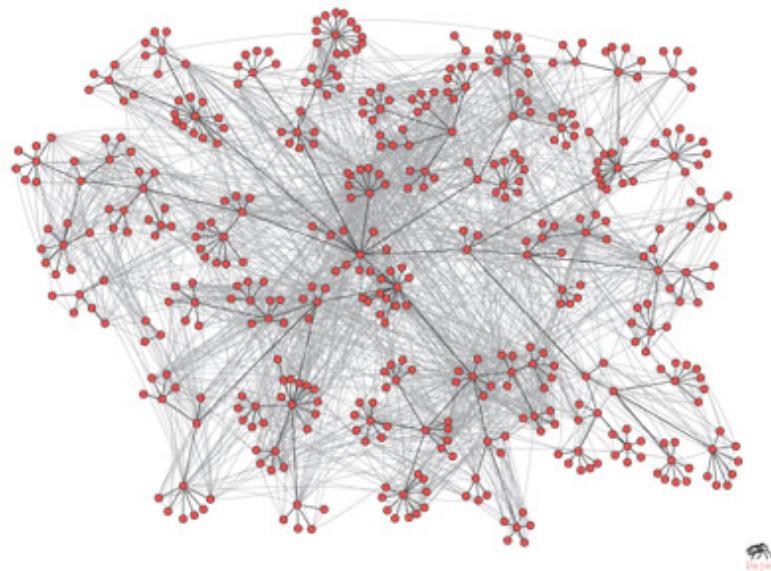


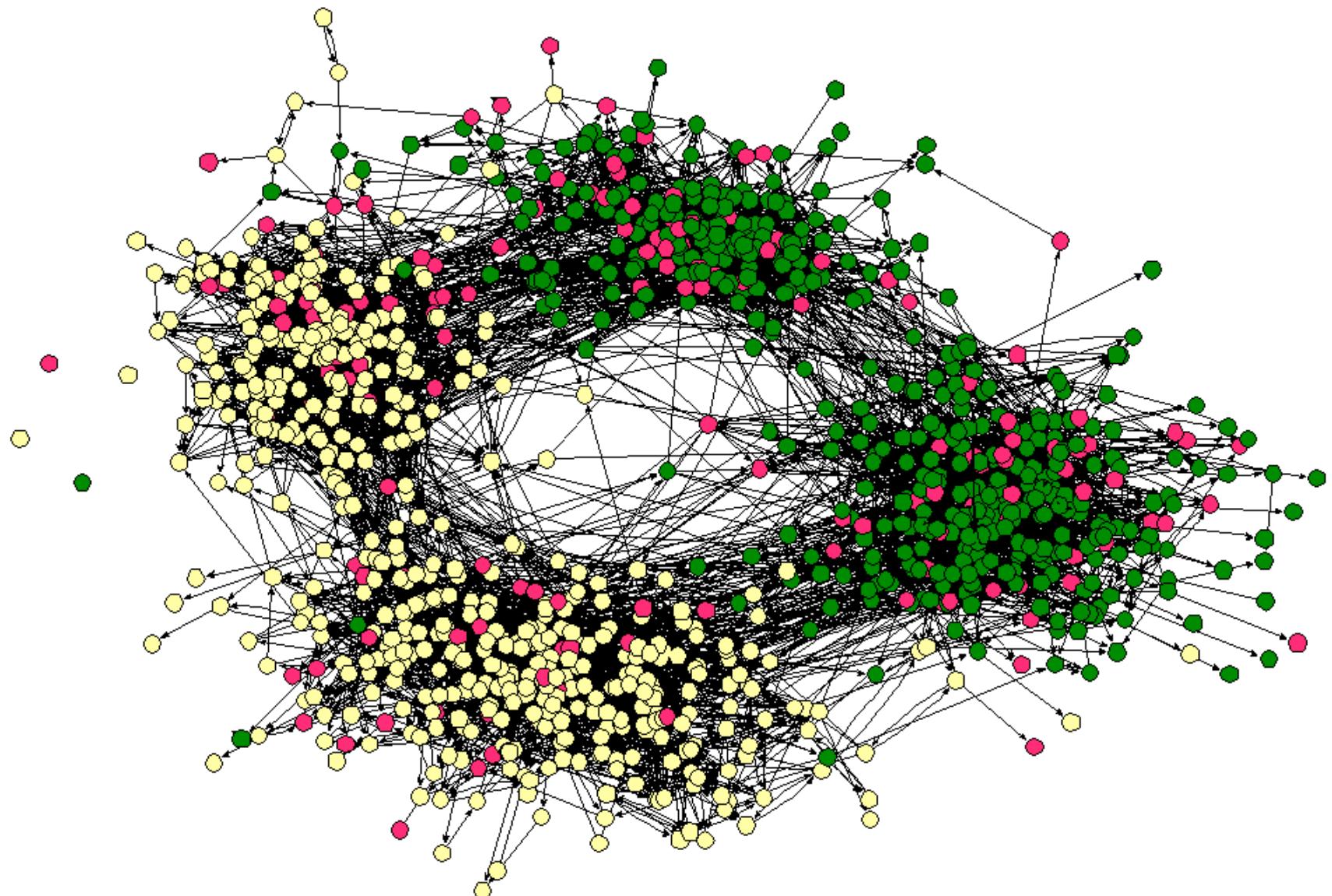
# Temporal Distance Metrics for Social Networks Analysis



John Tang<sup>1</sup>, Mirco Musolesi<sup>1</sup>, Cecilia Mascolo<sup>1</sup> and Vito Latora<sup>2</sup>

Computer Laboratory, University of Cambridge

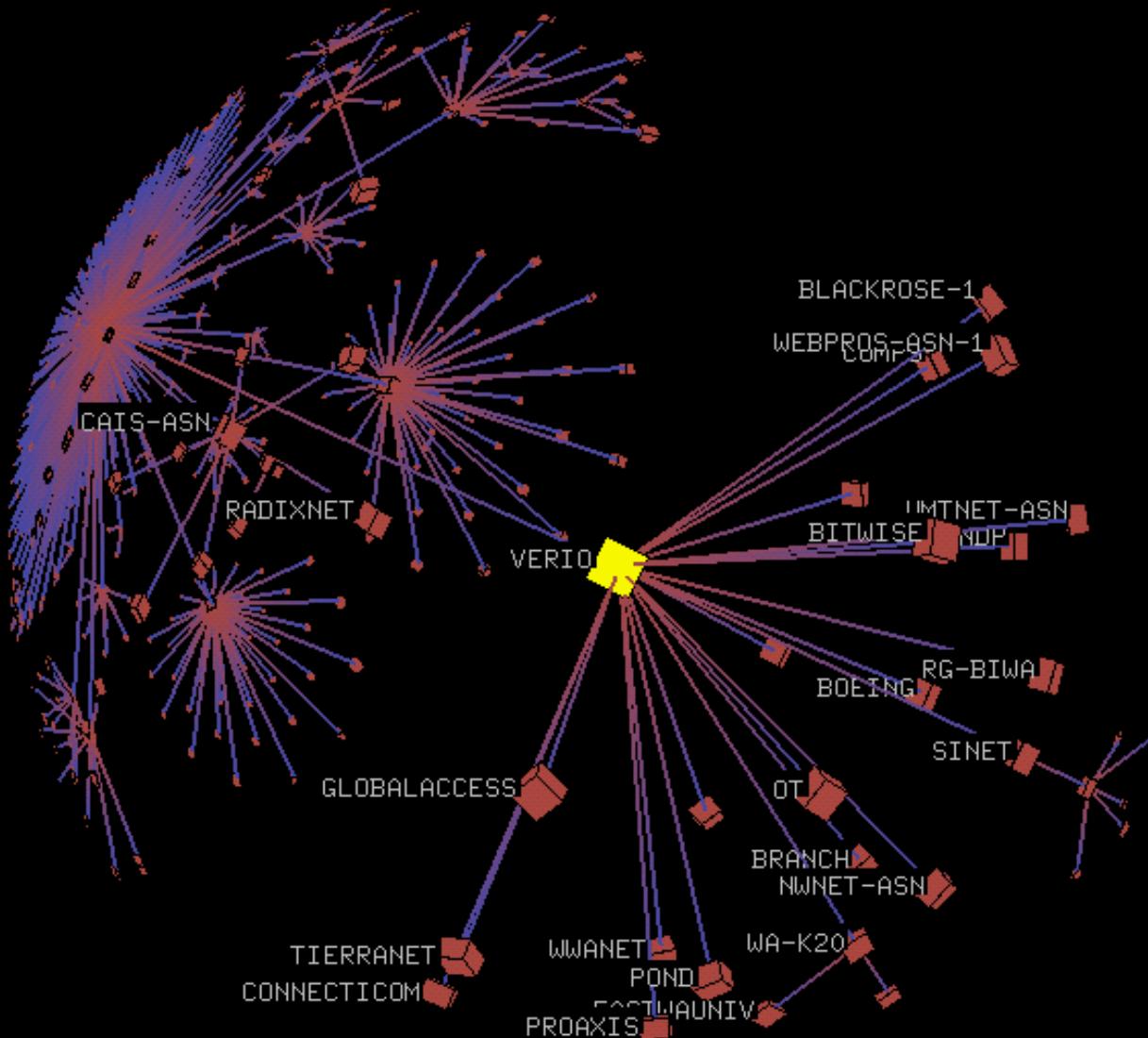
<sup>2</sup>INFN/Dept of Physics, University of Catania



Credit: Mark Newman

[mirco.musolesi@cl.cam.ac.uk](mailto:mirco.musolesi@cl.cam.ac.uk)

2

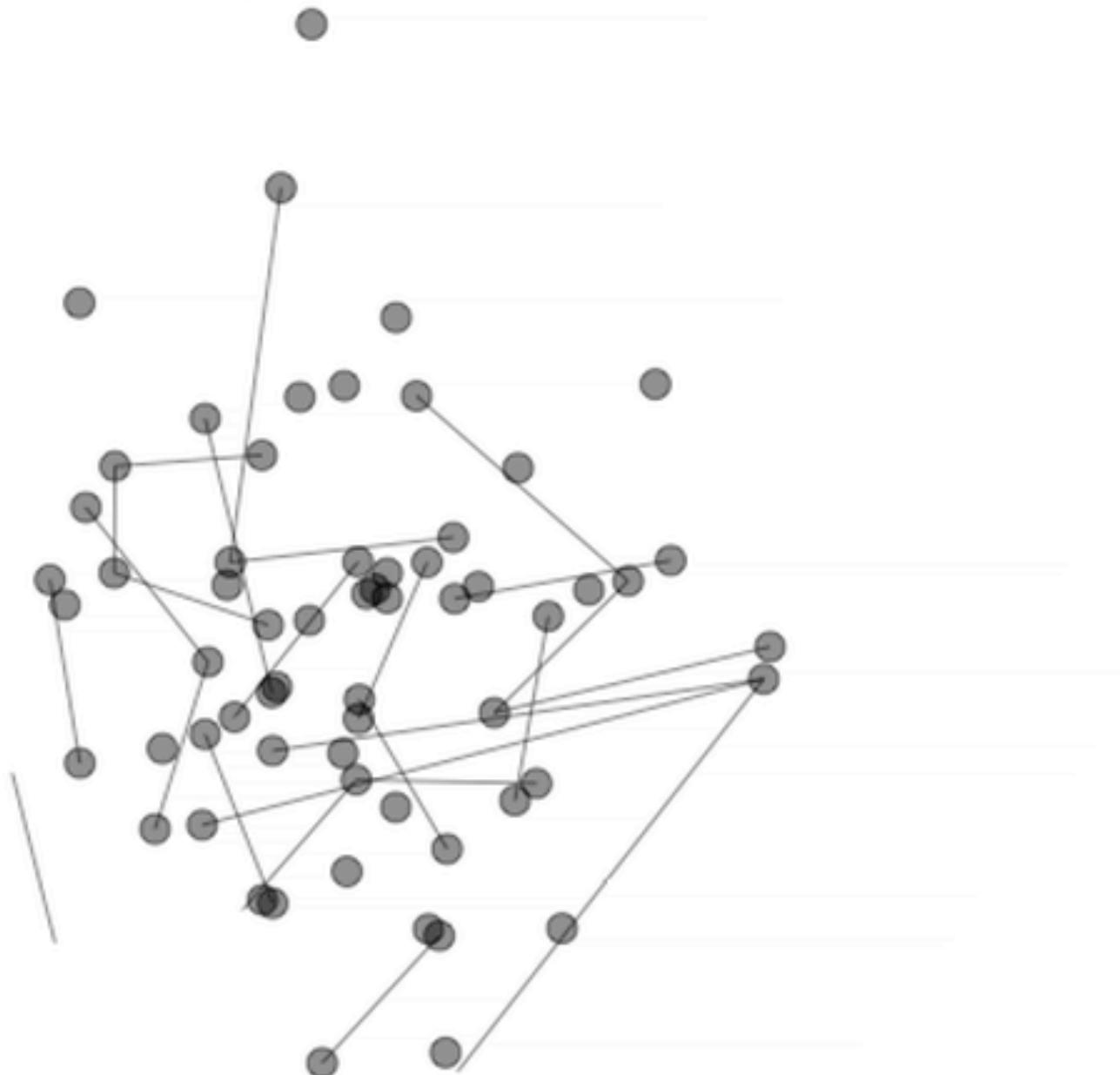


Hyperbolic view of BGP tables

Credit: kc claffy, CAIDA

mirco.musolesi@cl.cam.ac.uk

slice:0 time:20.730-21.730



Reality Mining Dataset

[mirco.musolesi@cl.cam.ac.uk](mailto:mirco.musolesi@cl.cam.ac.uk)

Explore ▾

Bebo is a social media network where friends share their lives and explore great entertainment. Get started by [signing up](#).

**What's Hot?**

Explore the new Bebo  
Video Music Authors Apps

**Video**

- December Give-away
- Orange Film Club
- The Gap Year
- Top Photos
- The hits



Marmos Full Length Episodes, T-Shirts, DVDs & much more!

**Profiles**

- Warner Home Videos UK
- Twilight Movie
- Animal Crossing
- Guitar Hero Guitar

**Bebo Originals**

- The Gap Year
- Meet the Freshers
- Sofia Taylor
- Barn King

**Applications**

- Rockstar Cars
- StyleGlam
- MindJolt Games
- Dumper Stars

**Login**

Forgot Password

Email or Username  Password

Automatically for 2 weeks

[SIGN UP](#)

**Online Gift Card**

Bebo Gift Card - THE gift to GET. Use online anywhere Visa is accepted.

[Find out more!](#)



Order 4 free  
O2 Pay & Go SIMs

**Videos**

MTV's Tom Takes On The Hills | Lost Scenes



[Play Video](#) [Play Indivisuals](#)

facebook

Remember Me  [Forgot your password?](#)

**Facebook email virus!**

[Back to Facebook](#)

[Discussion Board](#) [Topic View](#)

**Topic: Facebook email virus!**

Displaying posts 1 - 30 out of 211 by 182 people. [1](#) [2](#) [3](#) [4](#) [5](#) [Next](#) [Last](#)

 **Tracy wrote** on August 24, 2008 at 2:04pm

How do I contact Facebook about a virus going around in emails I got it today and apparently automatically gets sent to everyone's friends lists.  
It contains a YouTube link with a title that says, "I can see youoo".  
It would be nice if Facebook could send out some sort of alert ASAP.  
I can't seem to email to my large friends list.  
Keep bumping this to the top. Thank you.

 **Niki wrote** on August 24, 2008 at 2:31pm

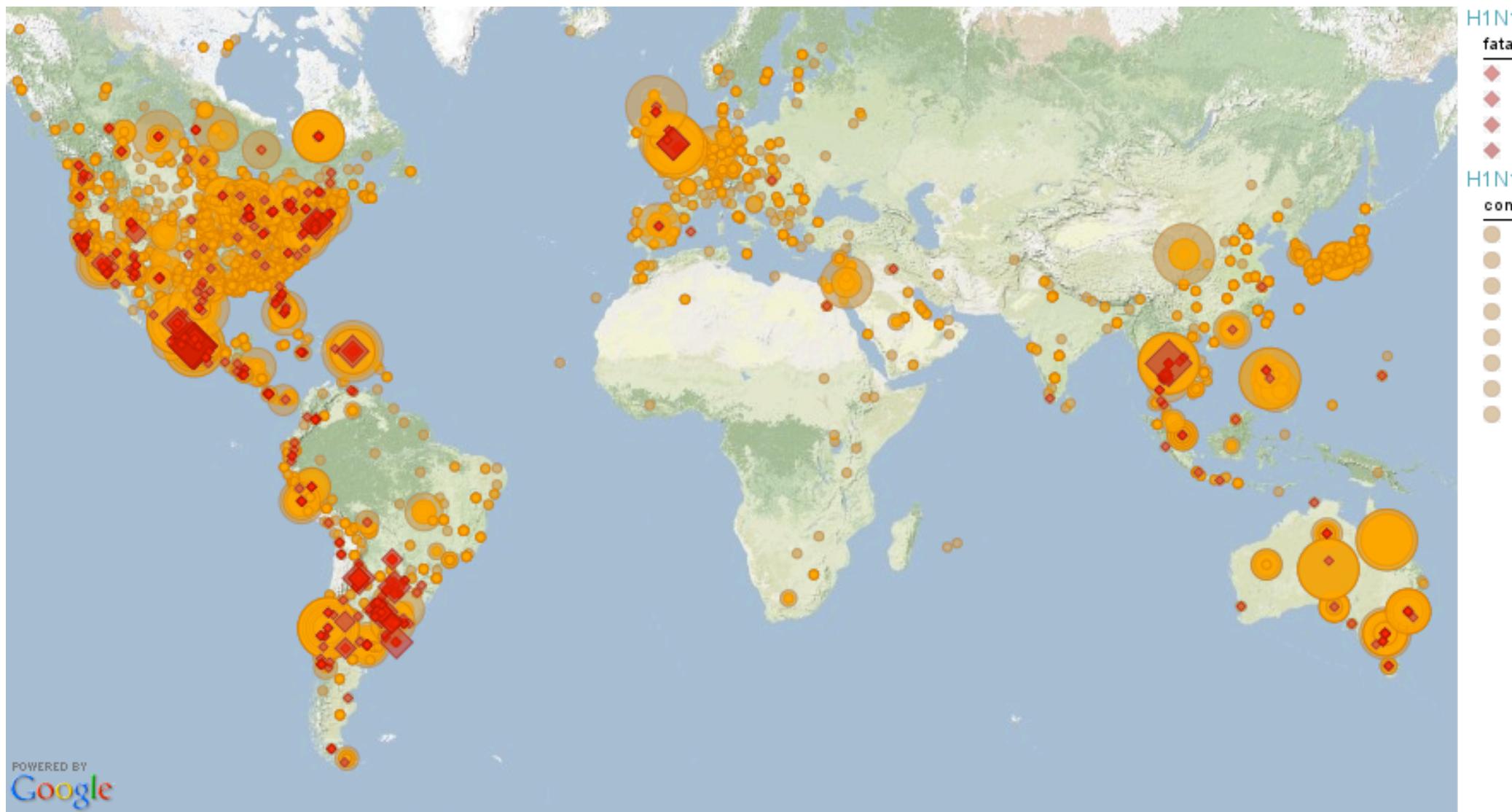
I had the same problem. What can I do? From my email someone send links to everyone. I didn't know, now everyone is in trouble, what can I do?  
HELP!

 **Nikki wrote** on August 24, 2008 at 3:34pm

I had to send the following status:  
Please read my status! NIKI DID NOT SEND THE VIDEO LINK. DO NOT OPEN IT. I BELIEVE IT'S A VIRUS AND SOMEONE IS USING OUR PROFILES TO SEND PHONY LINKS! I DON'T EVEN TALK LIKE THAT!!!!  
Someone sent it to me also. It's a Trojan virus link. Please understand, I wouldn't do this. I looked it up: Win32/Loob.MB (CA)  
Trojan.Zoo (Symantec)  
**MORE INFORMATION:**  
TrojanDownloader.Win32.Renesi.P is a generic detection for a family of trojans that monitor web searches performed by users and send this information to a remote site. They may also visit other websites, and replace search engines' advertising content, in order to collect per-click advertising payments.

 **3H wrote** on August 24, 2008 at 3:44pm

Just an FYI,  
There is a virus being sent out by the Application called "CRUSH ME". It will automatically send out open links to all of the friends on your friends list.  
I spent 2 hours last night going to each of my friends page and deleting the open virus email that was sent. I'm very happy that I don't have more than 350 friend...LOL  
I must say that it was strange to find my viral file with myself with the entire list of my friends that had posts sent from my account. The makers of "CRUSH ME" have definitely misused individuals information.



POWERED BY  
Google

**FluTracker**

H1N1 Incidents 12:24 EDT 22 July 2009

<http://flutracker2.rhizalabs.com/cbi/snapshot/page?concept=~fd000a02514fc4889a9d0322ec393fb59c1f1ba0935a4ed8a99c>

Created by  
Flu Tracker Admin  
on Jul 23, 2009

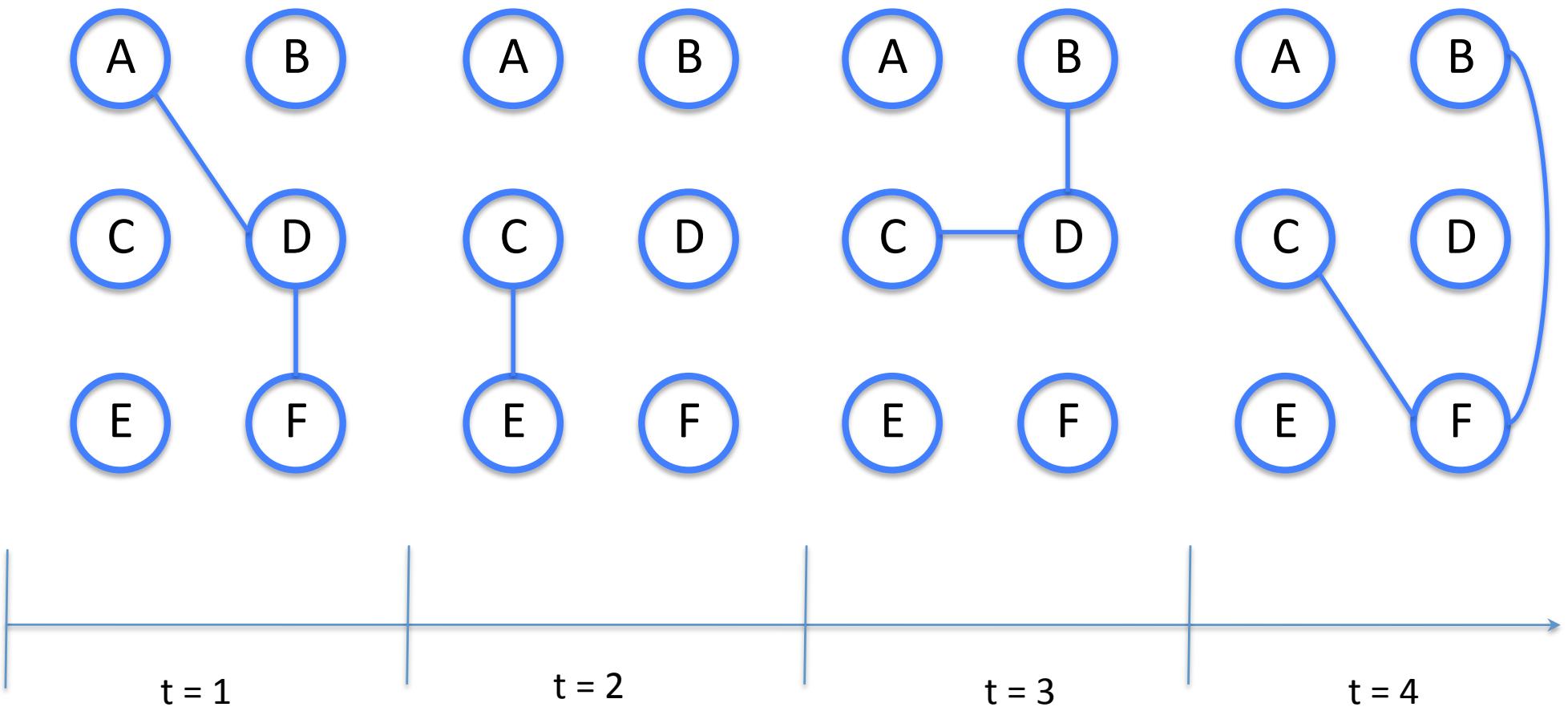
Credit: Flutracker.com

mirco.musolesi@cl.cam.ac.uk

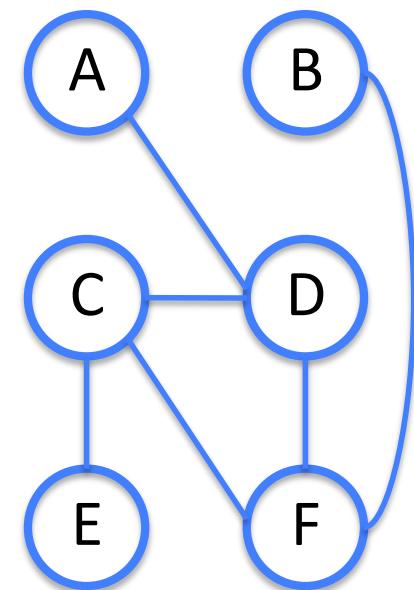
*Problem:* existing metrics do not capture the inherent dynamism of networks over time.

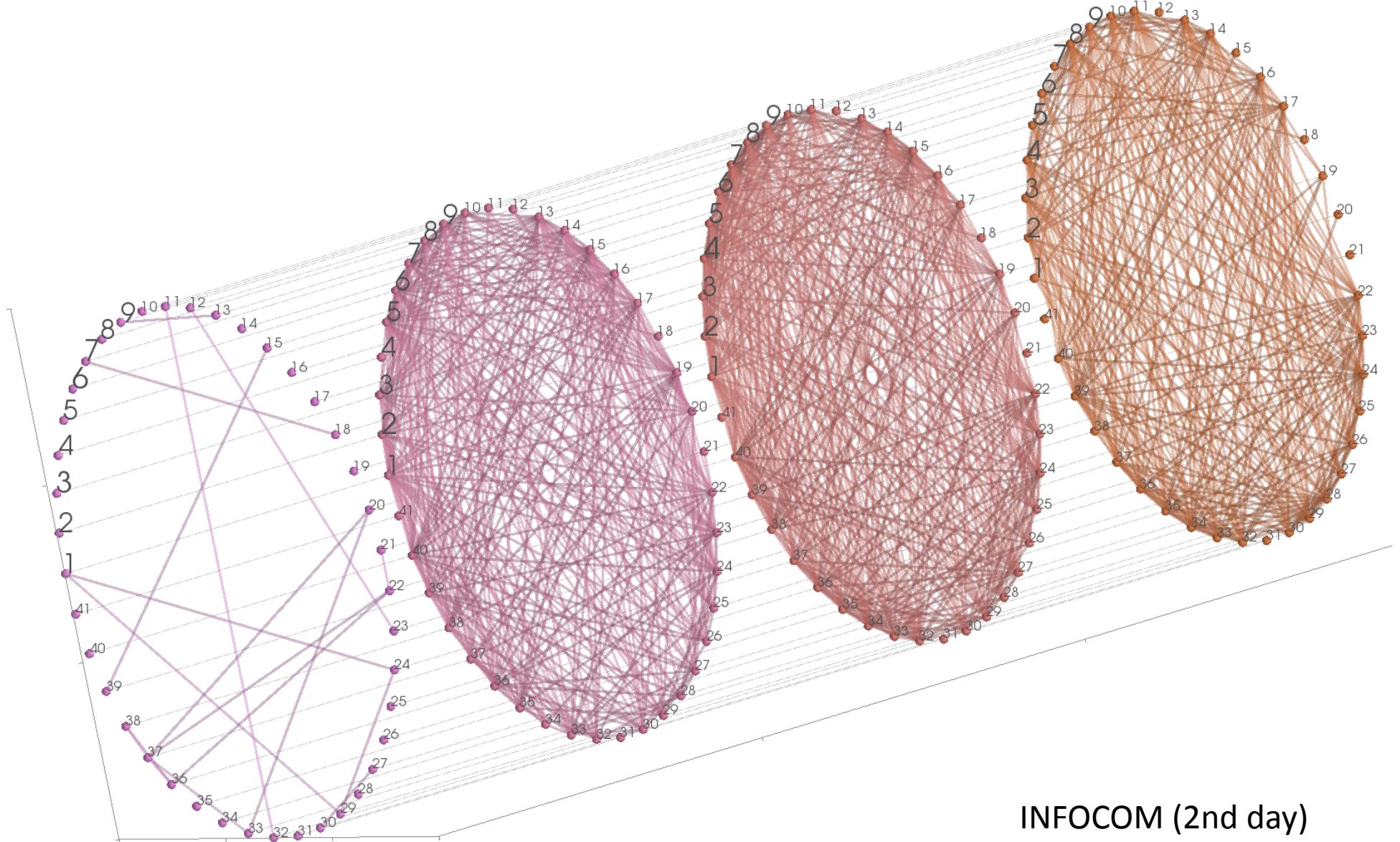
We need new **temporal metrics** defined over **temporal graphs** for studying dynamic processes over these networks.

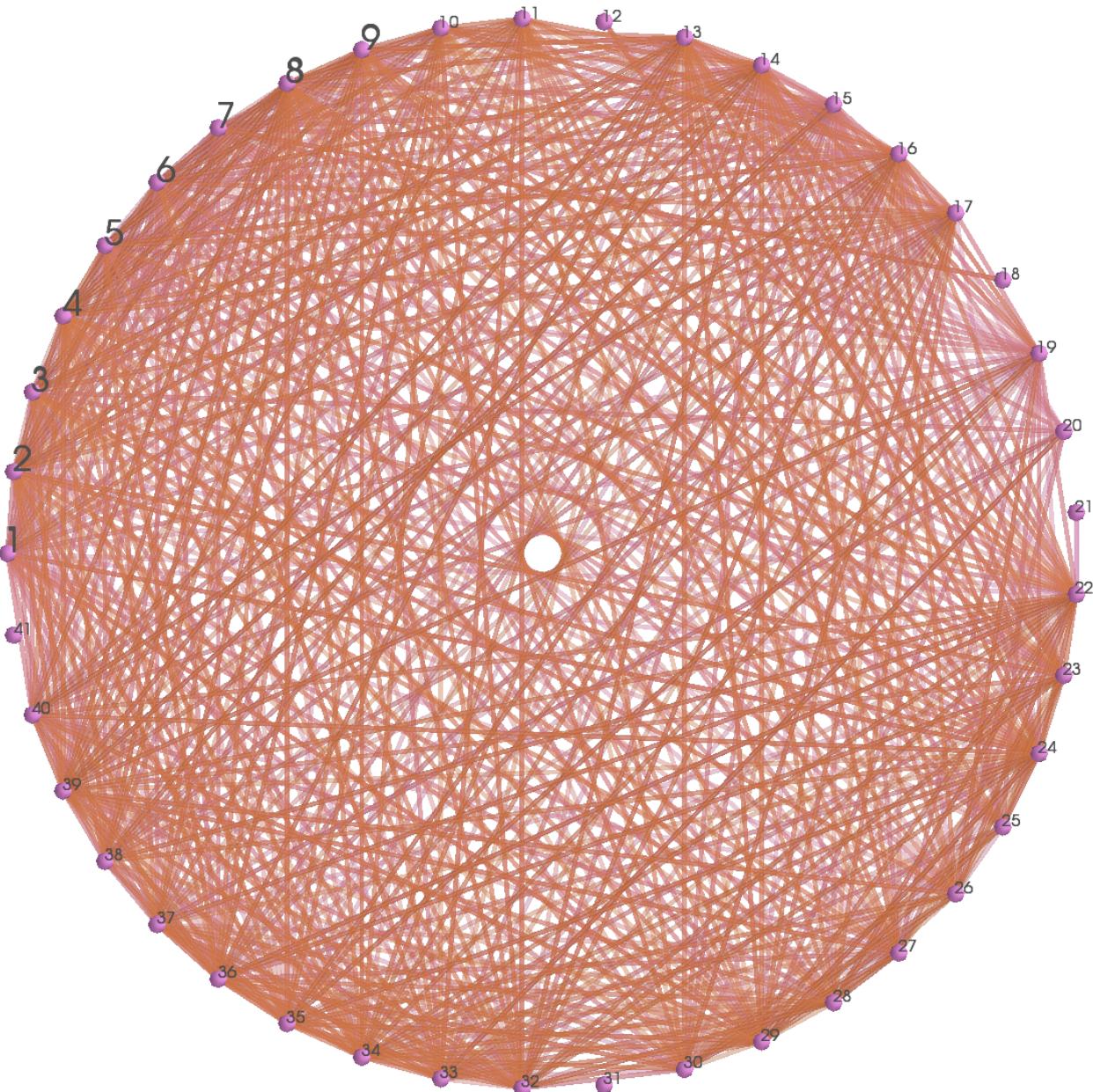
# An Example of Temporal Graph



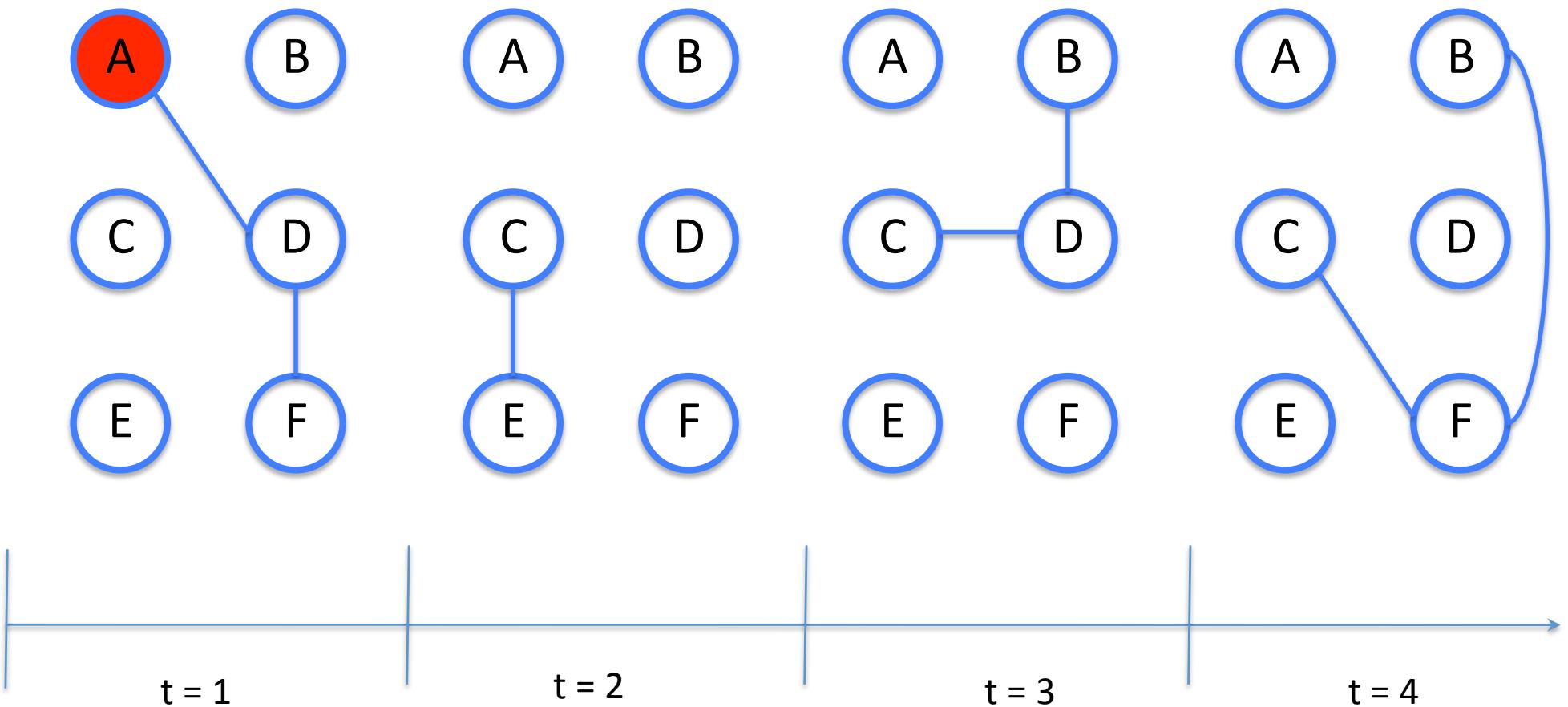
# ...and the Corresponding Static Graph







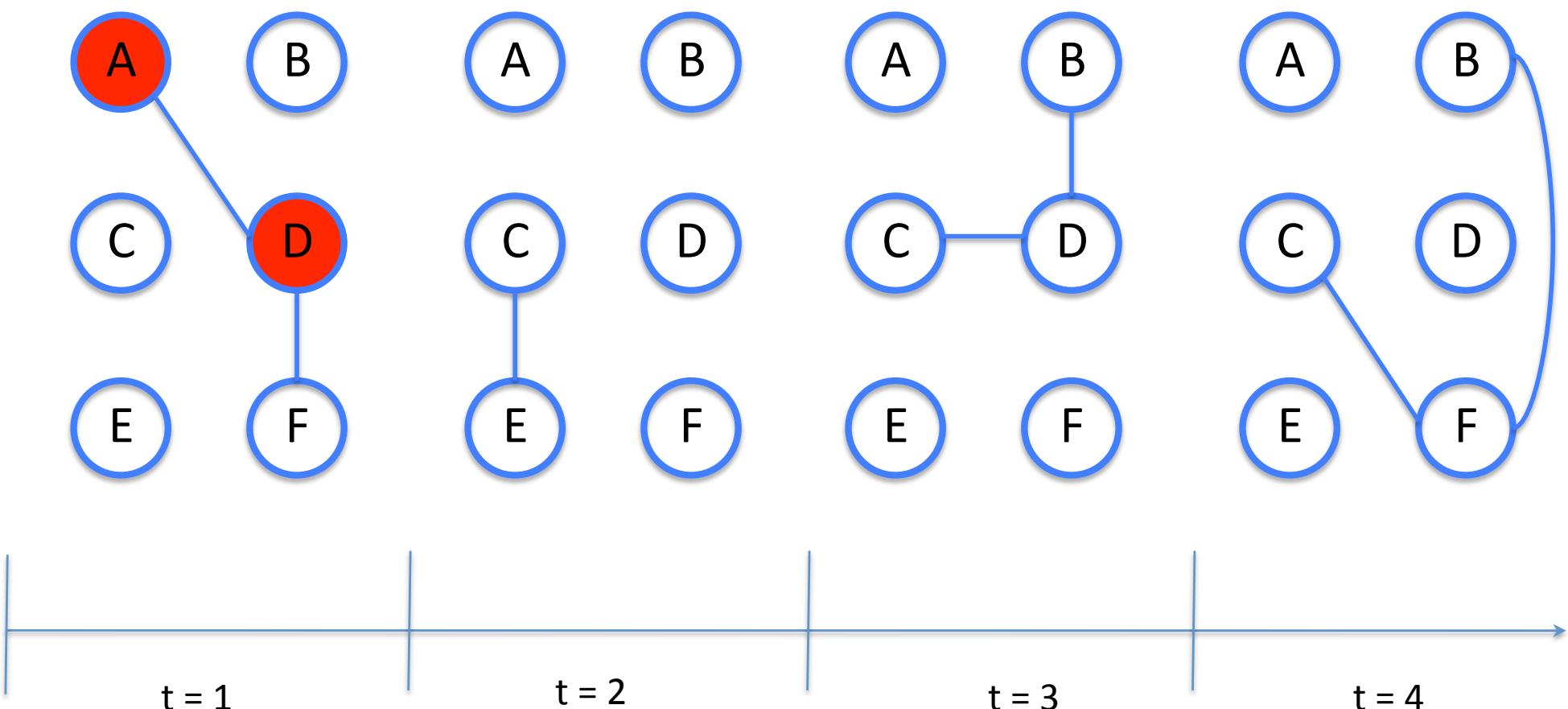
# Calculating the Temporal Distance



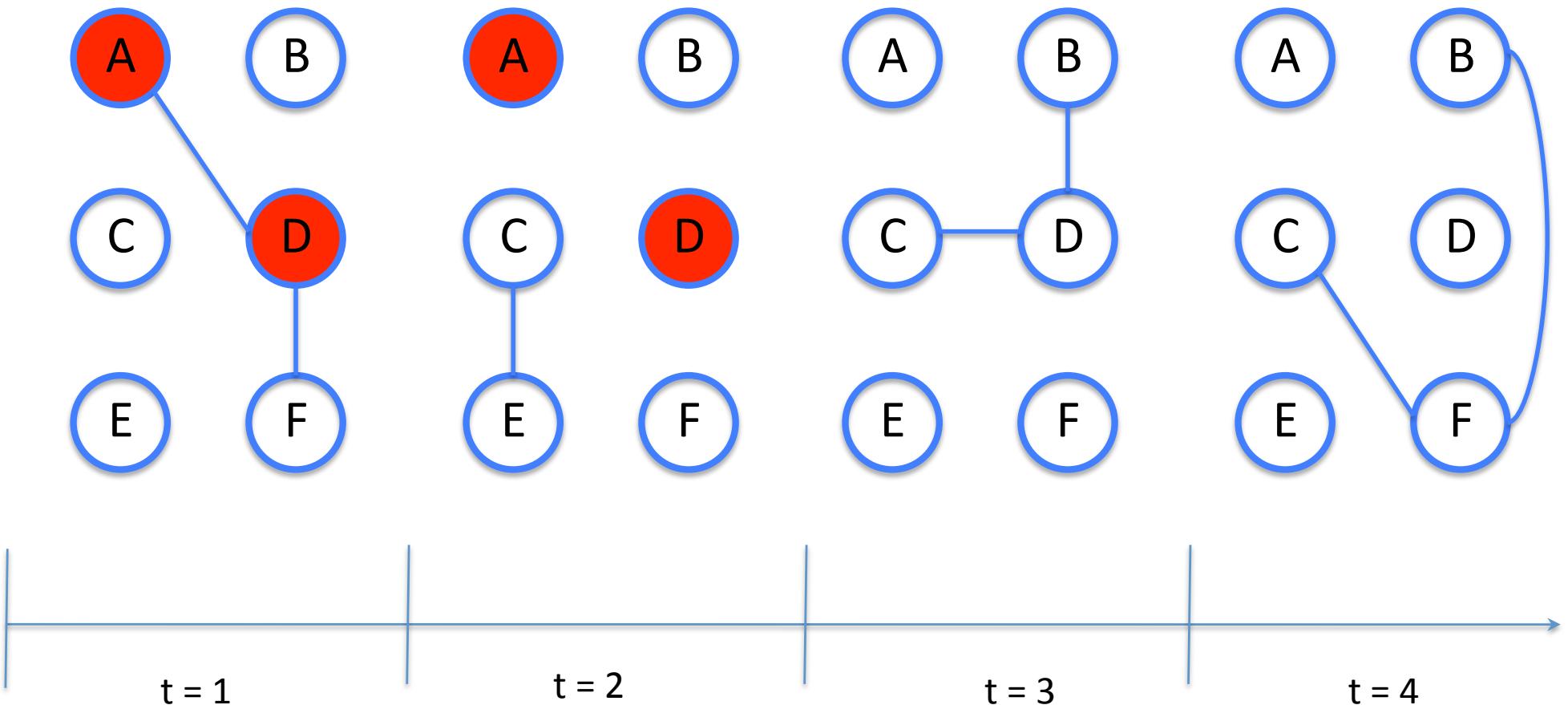
# Calculating the Temporal Distance

( $t = 1$ )

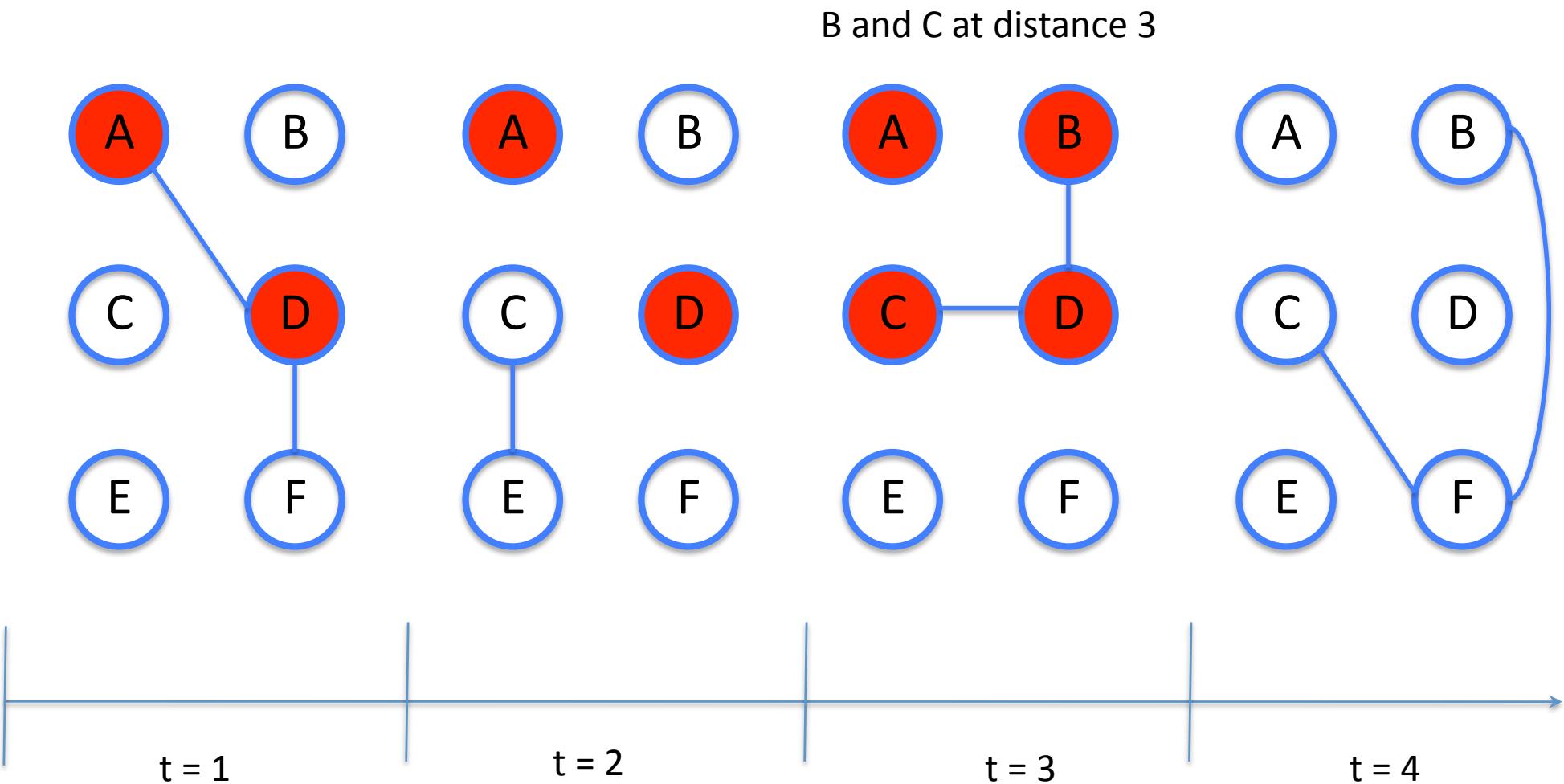
D at distance 1



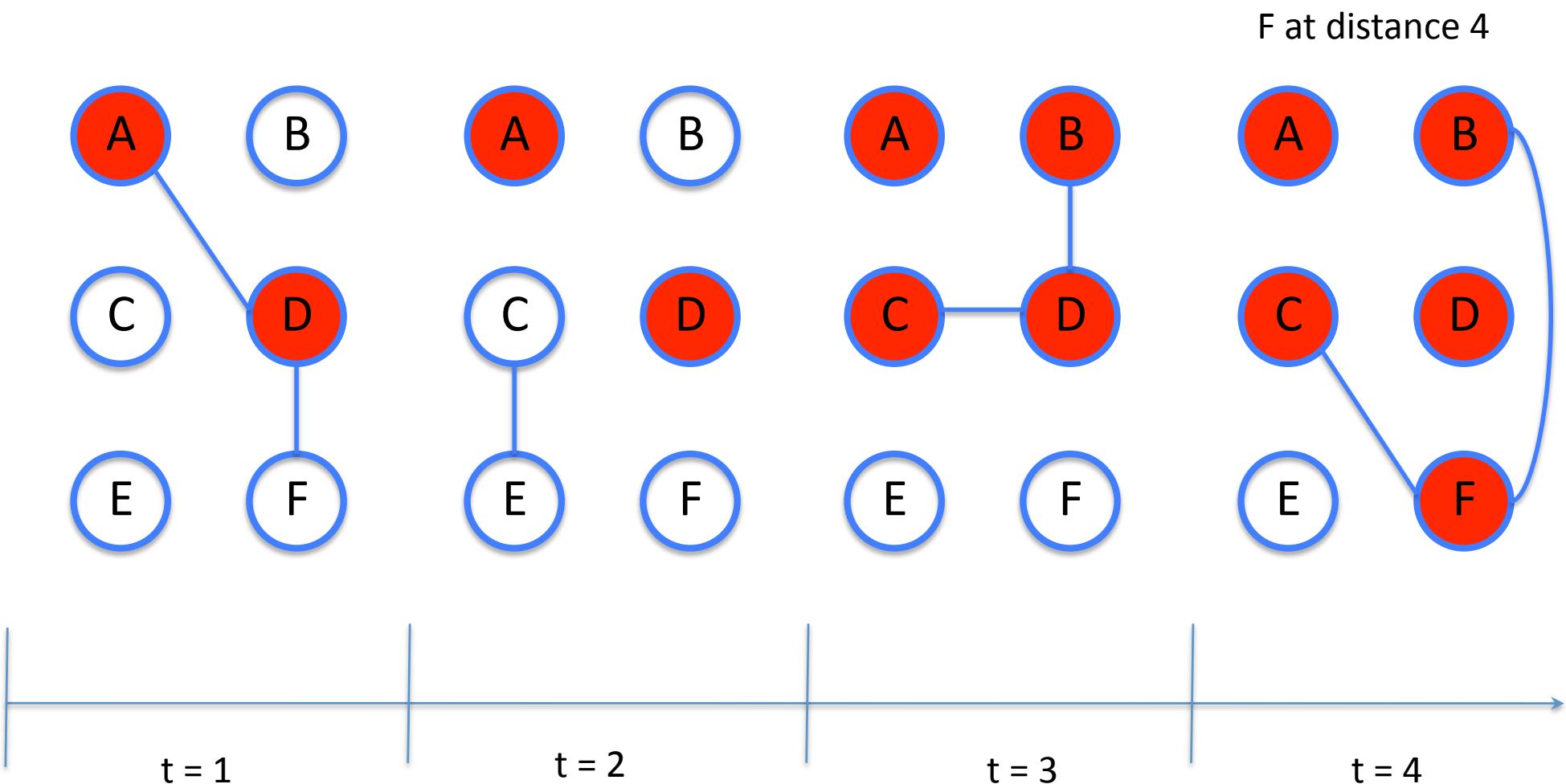
# Calculating the Temporal Distance ( $t = 2$ )



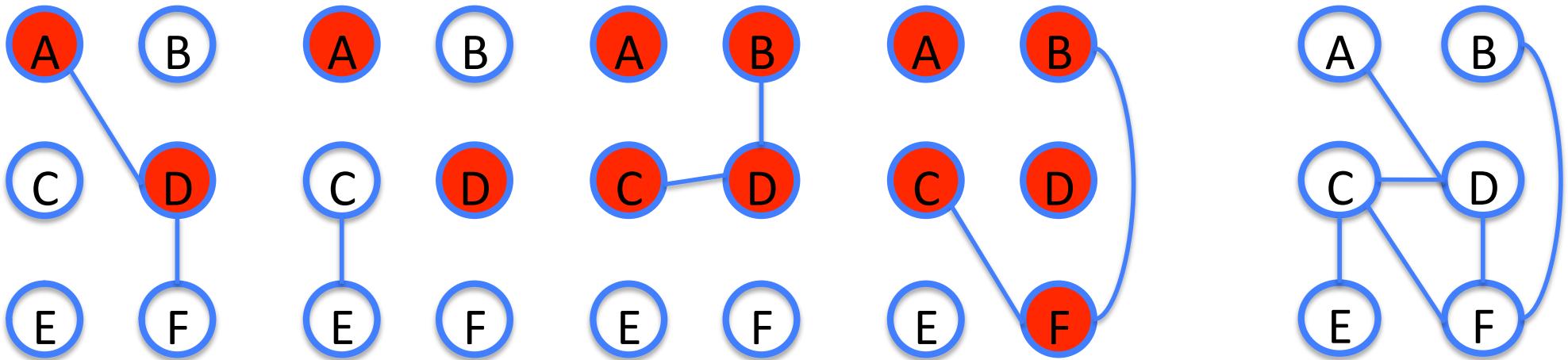
# Calculating the Temporal Distance ( $t = 3$ )



# Calculating the Temporal Distance ( $t = 4$ )



# what about the Static Distance?

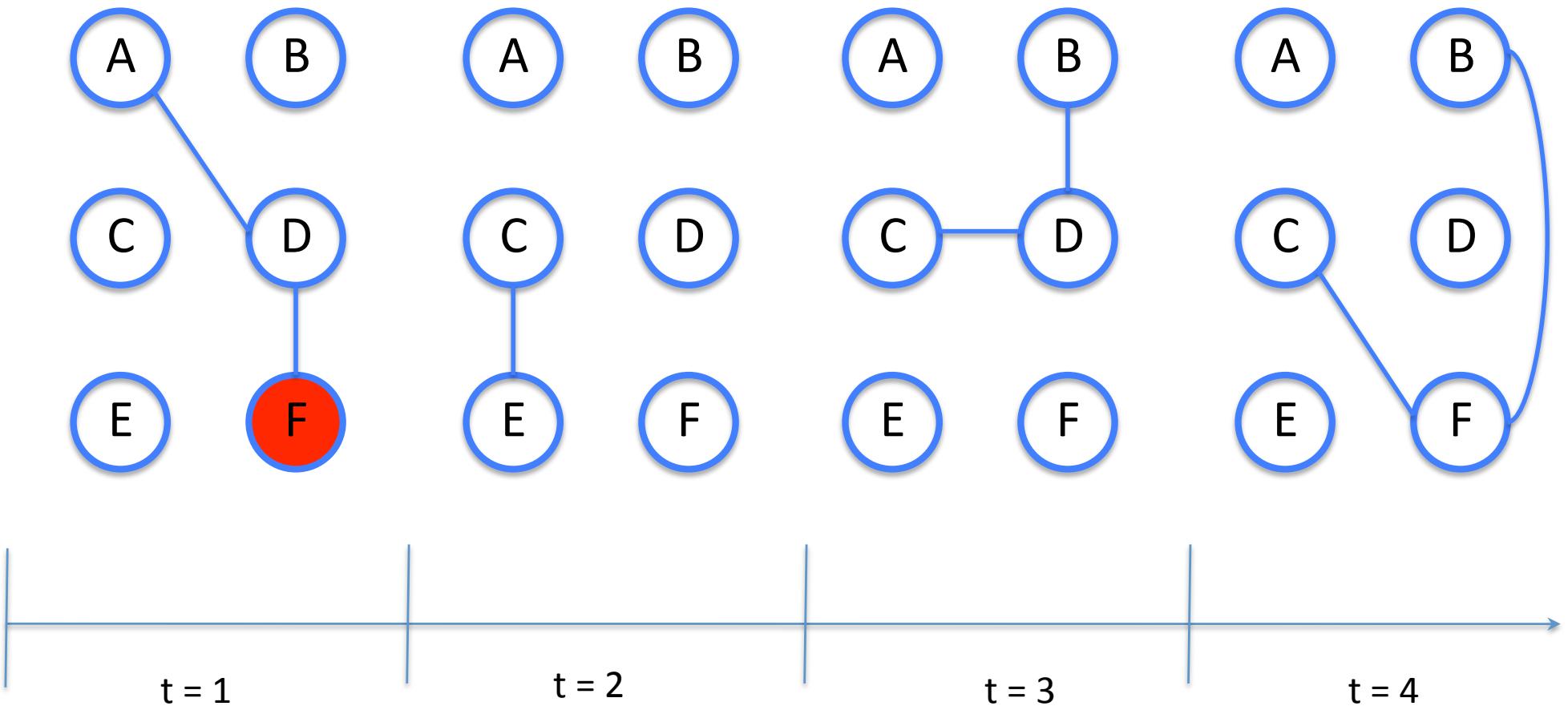


E is *statically* reachable but in reality it is not *dynamically* reachable!

A->F requires 2 transmissions (hops), but in reality it requires 3

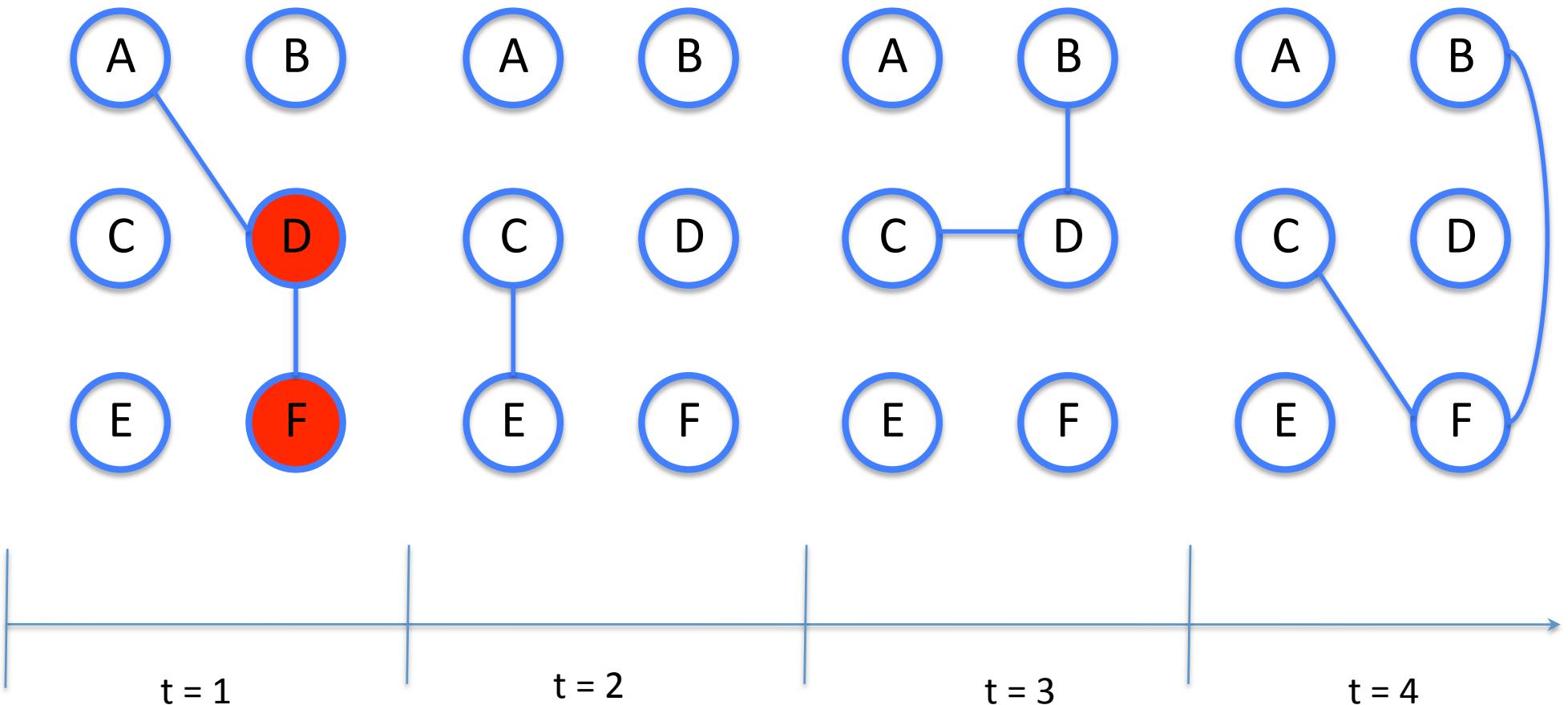
No information about the duration of the process

# What about the Symmetric Distance (F to A)?

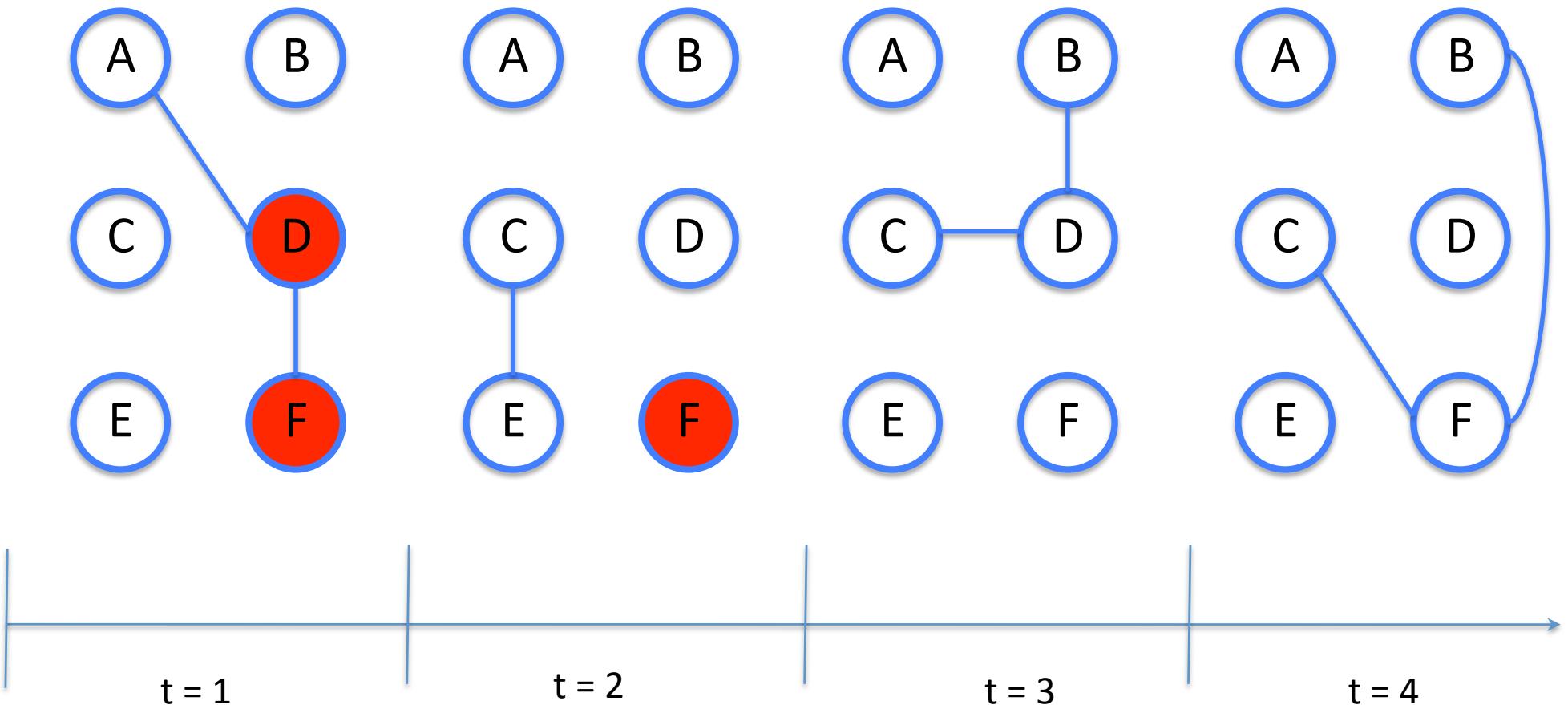


# Calculating the Inverse Temporal Distance

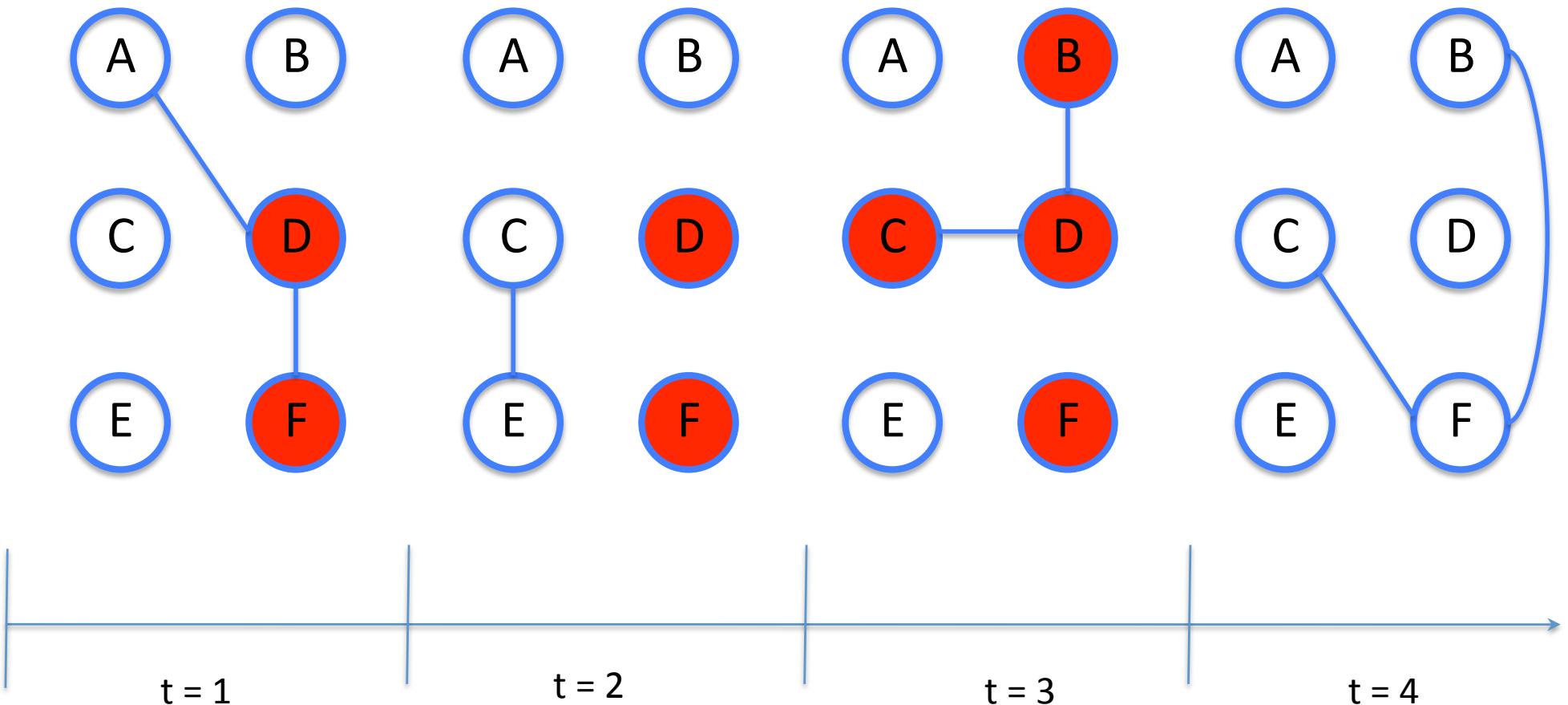
( $t = 1$ )



# Calculating the Inverse Temporal Distance ( $t = 2$ )

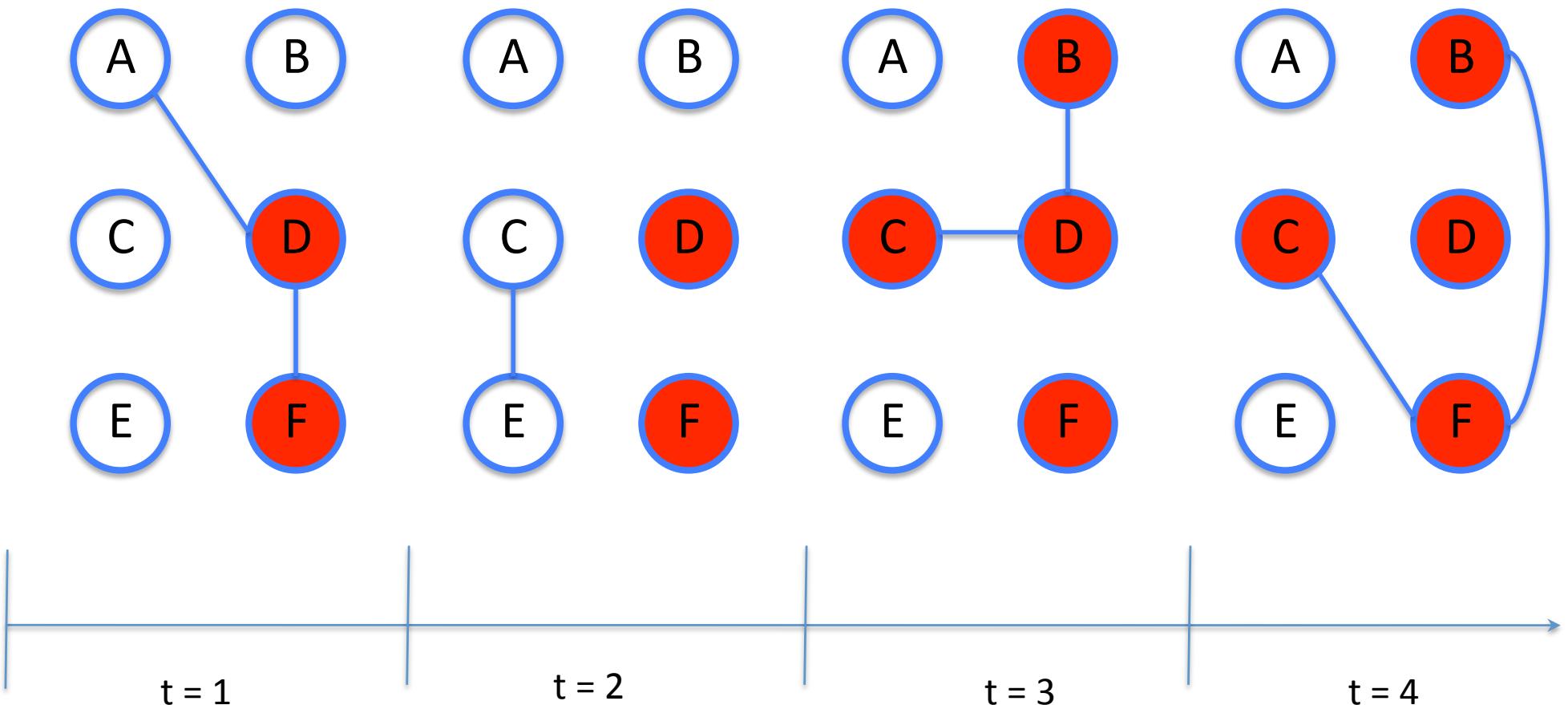


# Calculating the Inverse Temporal Distance ( $t = 3$ )



# Calculating the Inverse Temporal Distance ( $t = 4$ )

A is not reachable  
[infinite distance]



# Let's Get a Bit More Formal...

- Characteristic temporal path length:

$$L^h(t_{min}, t_{max}) = \frac{1}{N(N-1)} \sum_{ij} d_{ij}^h(t_{min}, t_{max})$$

- Defined considering the horizon of the infection
- Possible problem due to the potential divergence due to pairs of nodes that are not temporally connected

# Let's Get a Bit More Formal...

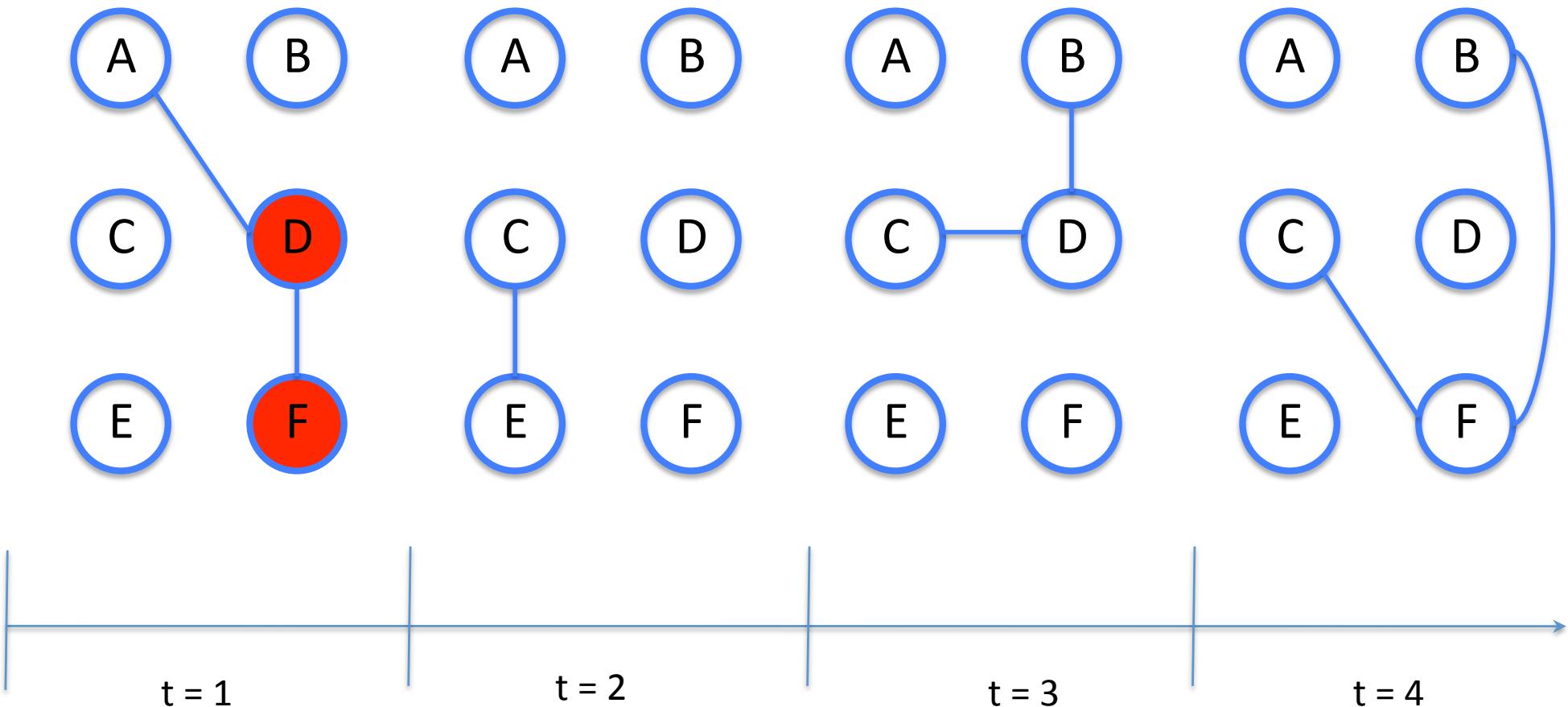
- Characteristic temporal path length:

$$L^h(t_{min}, t_{max}) = \frac{1}{N(N-1)} \sum_{ij} d_{ij}^h(t_{min}, t_{max})$$

- **Defined considering the horizon of the infection**
- Possible problem due to the potential divergence due to pairs of nodes that are not temporally connected

# Impact of the Horizon Parameter

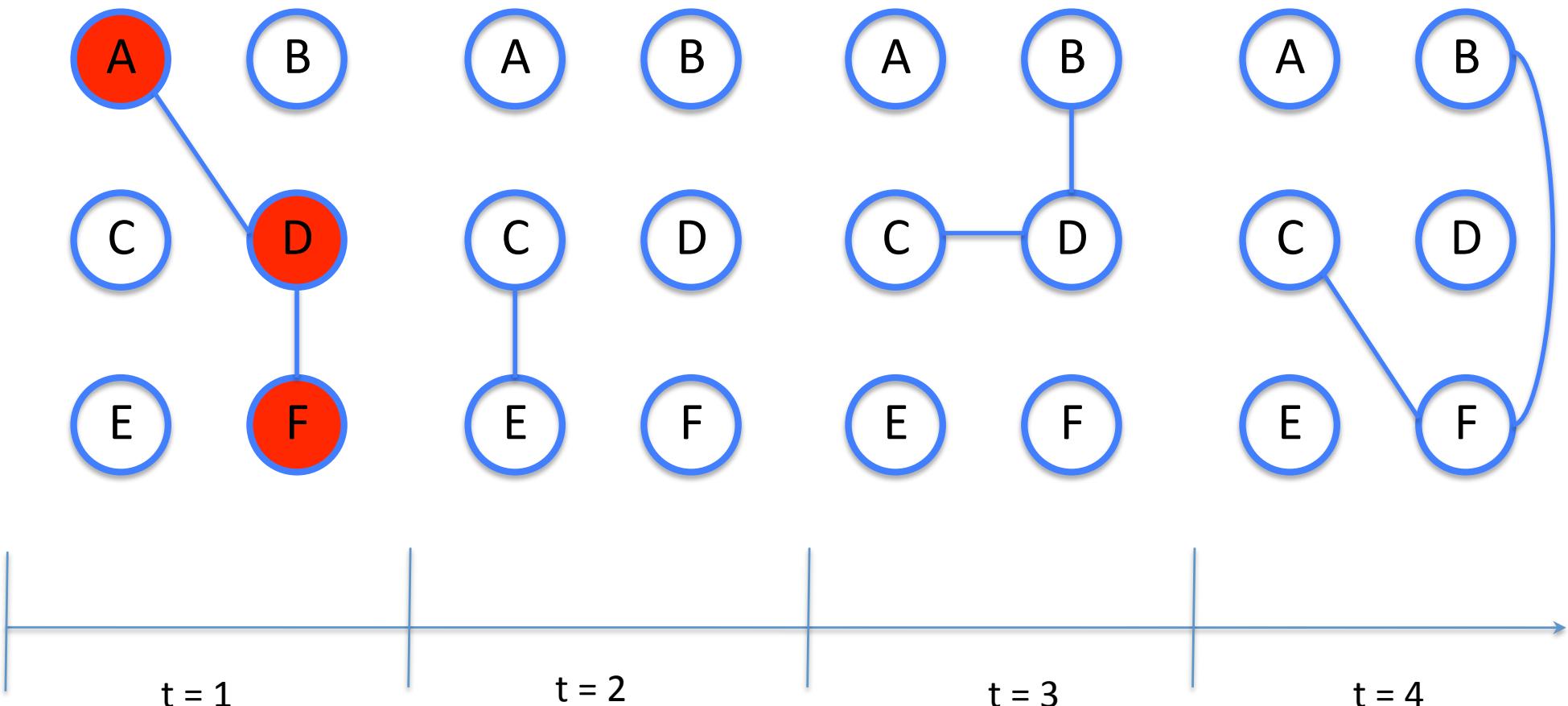
(F  $\rightarrow$  A,  $h = 1$ )



# Impact of the Horizon Parameter

## (F $\rightarrow$ A, $h = 2$ )

A was not reachable at all with  $h = 1$  (in 4 time windows), but with  $h = 2$  it is a distance 1!



# Let's Get a Bit More Formal...

- Characteristic temporal path length:

$$L^h(t_{min}, t_{max}) = \frac{1}{N(N - 1)} \sum_{ij} d_{ij}^h(t_{min}, t_{max})$$

- Defined considering the horizon of the infection
- **Possible problem related to the potential divergence due to pairs of nodes that are not temporally connected**

# Let's Get a Bit More Formal...

- Solution: definition of temporal efficiency:

$$E_{T_{ij}}^h(t_{min}, t_{max}) = \frac{1}{d_{ij}^h(t_{min}, t_{max})}$$

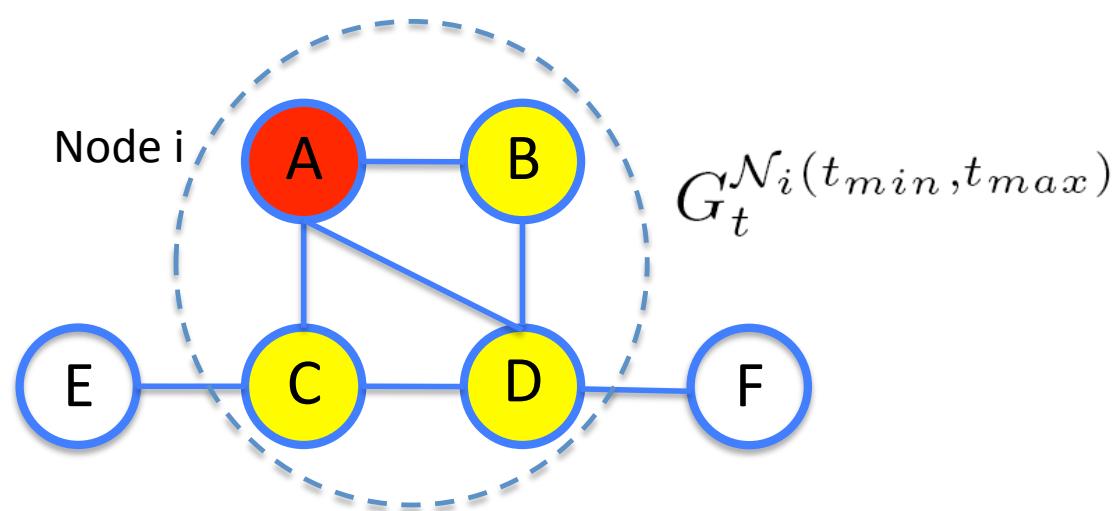
$$E_{glob}^h(t_{min}, t_{max}) = \frac{1}{N(N - 1)} \sum_{ij} E_{T_{ij}}^h(t_{min}, t_{max})$$

- High value of E (low value of L) means that the nodes of the graphs can communicate efficiently

# Local Temporal Efficiency

$$E_{loc_i}(t_{min}, t_{max}) = E_T\{G_t^{\mathcal{N}_i(t_{min}, t_{max})} \quad t \in [t_{min}, t_{max}]\}$$

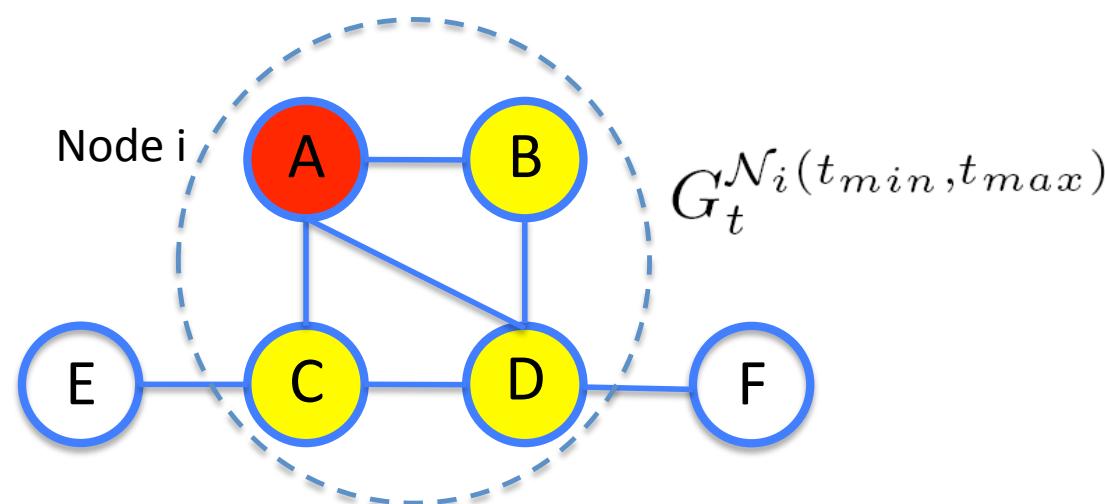
$$E_{loc}(t_{min}, t_{max}) = 1/N \sum_i E_{loc_i}(t_{min}, t_{max})$$



# Temporal Clustering Coefficient

$$C_i(t_{min}, t_{max}) = \frac{\sum_{t=t_{min}}^{t_{max}} \# \text{ of edges in } G_t^{\mathcal{N}_i(t_{min}, t_{max})}}{\tau \cdot \frac{k_i(t_{min}, t_{max})(k_i(t_{min}, t_{max}) - 1)}{2}}$$

$$C(t_{min}, t_{max}) = 1/N \sum_i C_i(t_{min}, t_{max})$$



# INFOCOM Dataset: Static vs Temporal Metrics

			Static				Temporal		
Day	N	$\langle k \rangle$	C	L	$C_{rand}$	$L_{rand}$	C	$L^*$	Disc
1	37	25.7	0.818	1.291	0.764	1.336	0.033	4.090	0.28
2	39	28.3	0.845	1.269	0.824	1.263	0.110	4.556	0.13
3	38	22.3	0.744	1.420	0.644	1.405	0.077	4.003	0.19
4	39	21.4	0.722	1.444	0.541	1.474	0.052	4.705	0.14

$h_{min} = 12\text{am}$ ,  $h_{max} = 12\text{pm}$ ,  $w = 5\text{ min}$

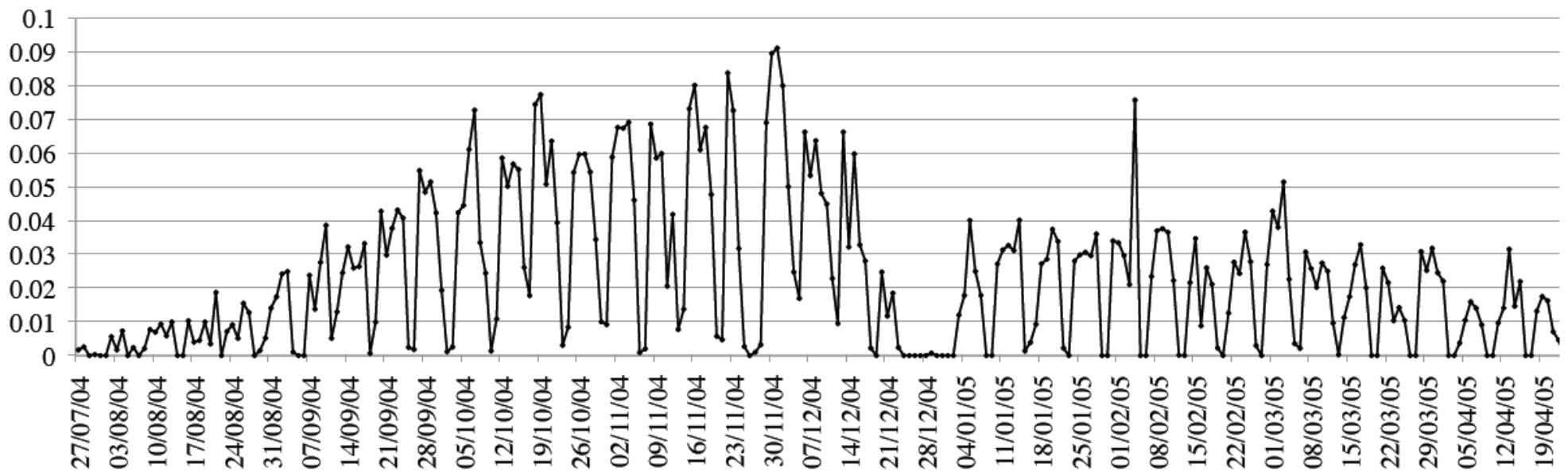
Static metrics underestimate L

# Reality Mining Dataset

Temporal Metrics					Reshuffled		
Date	C	Eloc	L	Eglob	$E_{loc}$	L	$E_{glob}$
08 Sep	0.014	0.000	23h 15m	0.000	0.003	21h 58m	0.010
15 Sep	0.060	0.000	22h 47m	0.001	0.007	19h 55m	0.024
22 Sep	0.061	0.000	22h 53m	0.001	0.007	20h 42m	0.019
29 Sep	0.060	0.001	22h 20m	0.001	0.009	17h 44m	0.037
06 Oct	0.026	0.000	22h 14m	0.001	0.011	16h 23m	0.041
13 Oct	0.038	0.000	21h 37m	0.004	0.013	14h 57m	0.055
20 Oct	0.067	0.001	21h 45m	0.003	0.007	17h 4m	0.031
27 Oct	0.050	0.002	22h 1m	0.001	0.013	15h 19m	0.050
03 Nov	0.051	0.001	21h 6m	0.004	0.012	16h 17m	0.043
10 Nov	0.051	0.000	20h 5m	0.004	0.015	14h 25m	0.061

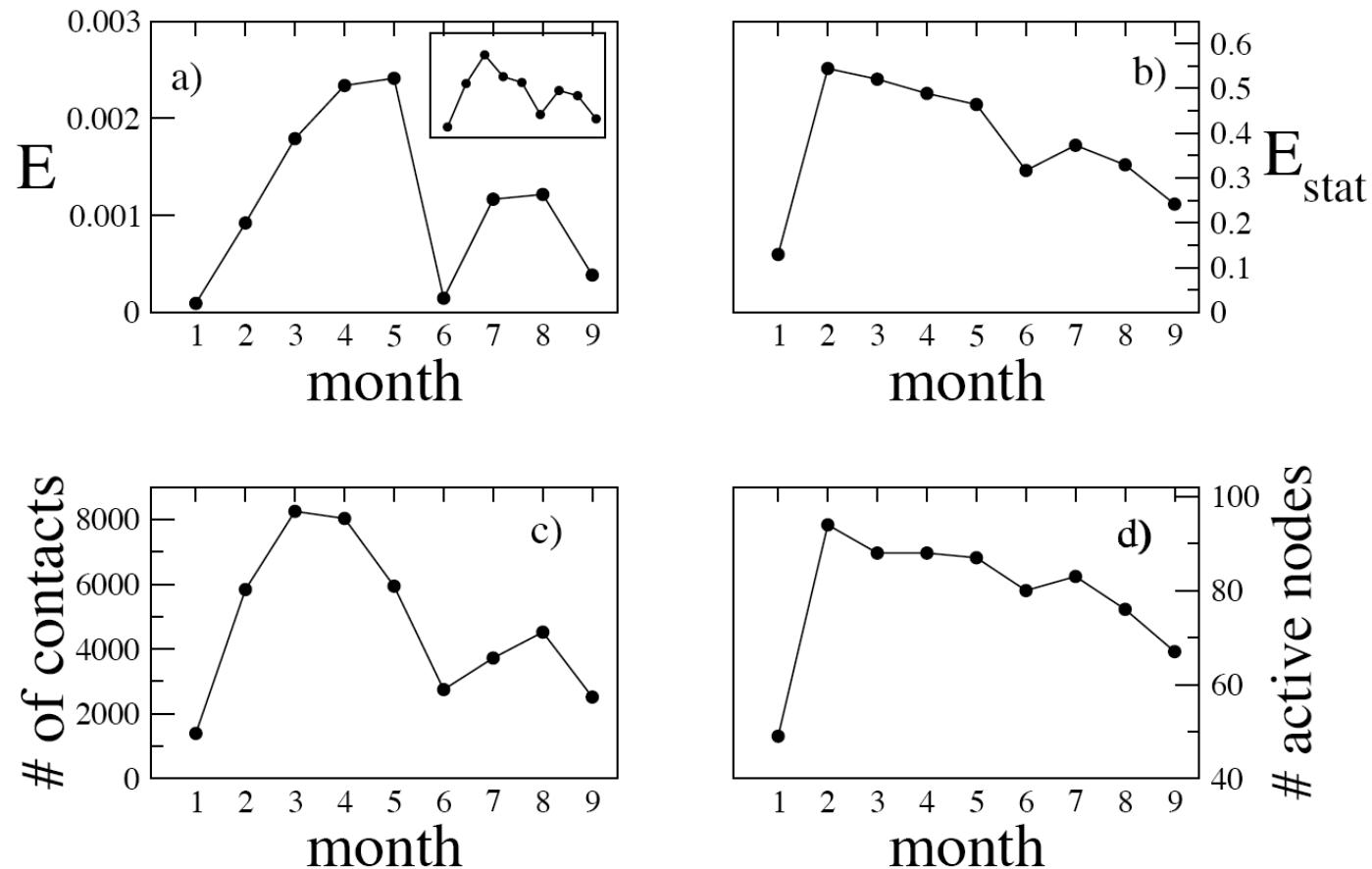
$$h = 1, t_{min} = 12 \text{ am}, t_{max} = 12 \text{ pm}, w = 5 \text{ min}$$

# Variation of the Clustering Coefficient over Time



Reality Mining Dataset ( $t_{\min} = 12 \text{ am}$ ,  $t_{\max} = 12 \text{ pm}$ ,  $w = 5 \text{ min}$ )

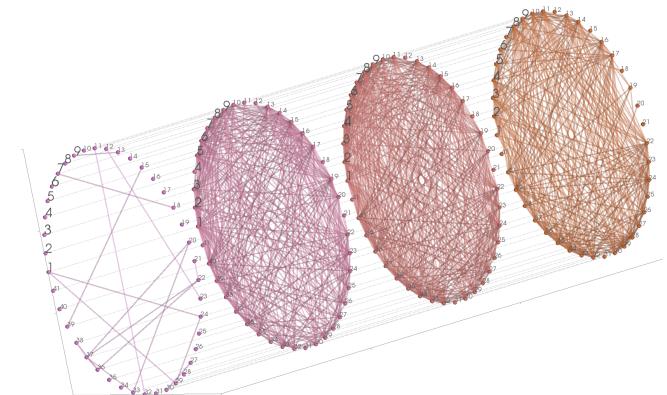
# Temporal Efficiency vs Static Efficiency



MIT Dataset

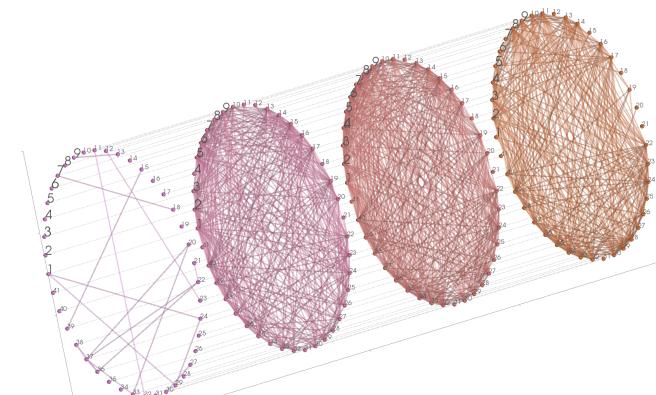
# Current Research Agenda

- Centrality measures
- Study of dynamic processes over the networks:
  - Message dissemination
  - Epidemics
  - Information propagation
- Analysis of new and larger datasets
- Small-world behavior in temporal networks



# Summary of the Talk

- New temporal metrics for studying dynamic processes over dynamic networks
  - Temporal distance
  - Temporal efficiency
  - Temporal clustering
- Analysis using real datasets



# Questions?

**Mirco Musolesi**

`mirco.musolesi@cl.cam.ac.uk`

[soon to be: `mirco@cs.st-andrews.ac.uk`]

`http://www.cl.cam.ac.uk/~mm753`

# INFOCOM Dataset: What Happens if We Reshuffle the Sequence?

Temporal Metrics					Reshuffled		
Day	C	$E_{loc}$	L	$E_{glob}$	$E_{loc}$	L	$E_{glob}$
1	0.033	0.003	19h 39m	0.003	0.077	5h 29m	0.100
2	0.110	0.020	9h 6m	0.024	0.194	2h 45m	0.239
3	0.077	0.013	10h 32m	0.018	0.114	4h 6m	0.167
4	0.052	0.009	9h 55m	0.013	0.104	3h 3m	0.165

$h = 1$ ,  $t_{min} = 12$  am,  $t_{max} = 12$  pm, no runs = 50

# Email Dataset

Temporal Metrics					Reshuffled			
Date	C	Eloc	L	Eglob	$E_{loc}$	L	$E_{glob}$	
27Oct	0	$3.1E^{-8}$	86397.94s	$9.3E^{-7}$	$7.7E^{-8}$	86396.91s	$1.6E^{-6}$	
28Oct	$3.5E^{-7}$	$4.0E^{-8}$	86399.78s	$1.4E^{-7}$	$4.1E^{-8}$	86399.71s	$1.5E^{-7}$	
29Oct	$2.5E^{-7}$	$3.9E^{-8}$	86399.03s	$3.9E^{-7}$	$7.2E^{-8}$	86398.59s	$7.3E^{-7}$	
30Oct	0	$5.8E^{-8}$	86398.76s	$5.5E^{-7}$	$6.9E^{-8}$	86398.48s	$7.5E^{-7}$	
31Oct	0	$4.7E^{-8}$	86398.92s	$4.9E^{-7}$	$6.5E^{-8}$	86398.64s	$6.9E^{-7}$	
01Nov	0	$5.8E^{-8}$	86399.03s	$4.9E^{-7}$	$6.6E^{-8}$	86398.85s	$6.0E^{-7}$	
02Nov	0	$4.3E^{-8}$	86398.68s	$5.4E^{-7}$	$6.5E^{-8}$	86398.67s	$6.8E^{-7}$	

$h = 1$ ,  $t_{min} = 12 \text{ am}$ ,  $t_{max} = 12\text{pm}$ ,  $w = 5 \text{ min}$