Link Streams
for the
Modeling of Interactions over Time

Work in progress...

ANR CONTINT – projet CODDDE ANR-13-CORD-0017-01

Matthieu Latapy, Tiphaine Viard,
Clémence Magnien, Noé Gaumont, ...

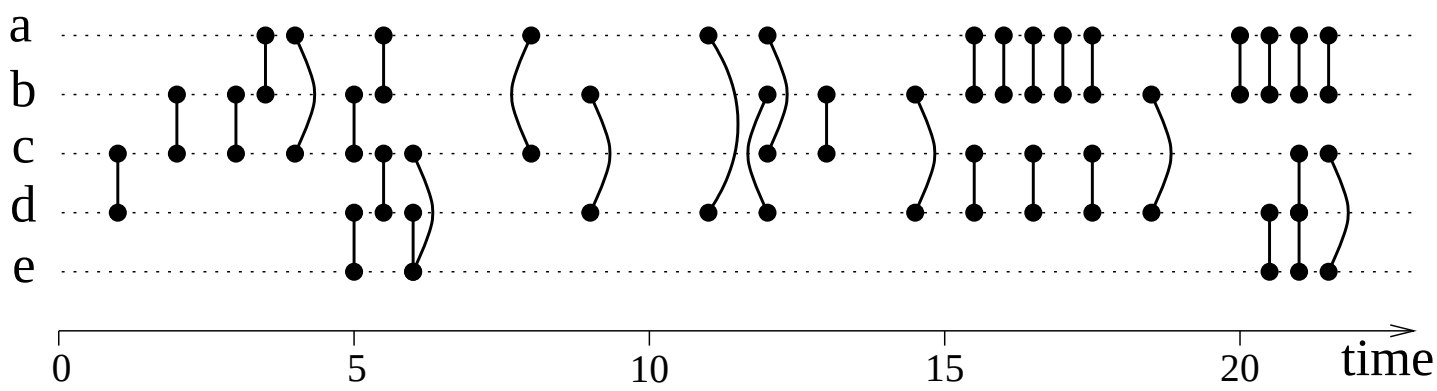
http://complexnetworks.fr

Matthieu.Latapy@lip6.fr

LIP6 – CNRS and UPMC

Our topic: link streams

interactions over time



$$l = (t, u, v)$$

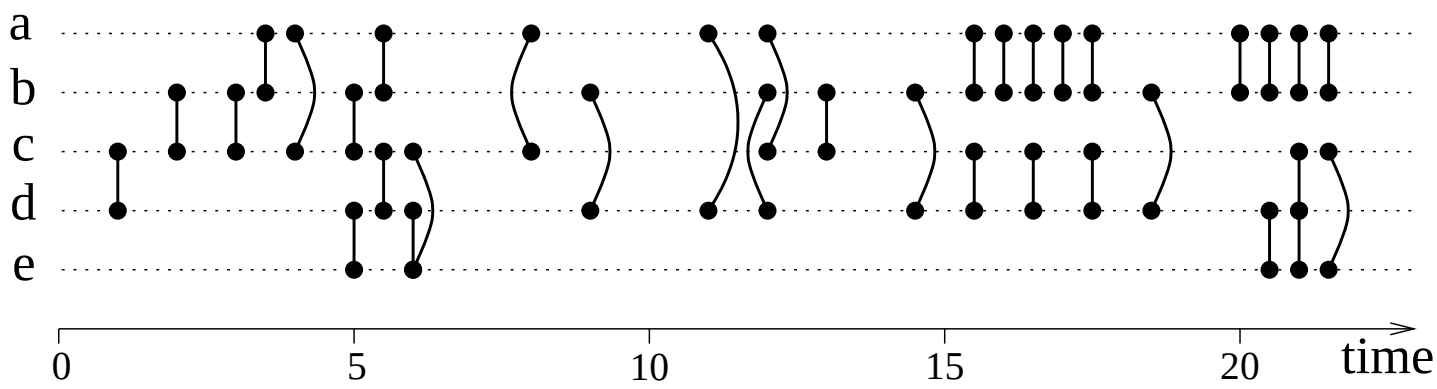
$t \in [\alpha, \omega]$: time
$u, v \in V$: nodes

Our topic: link streams

countless examples

email exchanges, network traffic, payments,
physical contacts, phone calls, web surfing, ...

interactions over time



$$l = (t, u, v)$$

$t \in [\alpha, \omega]$: time
$u, v \in V$: nodes

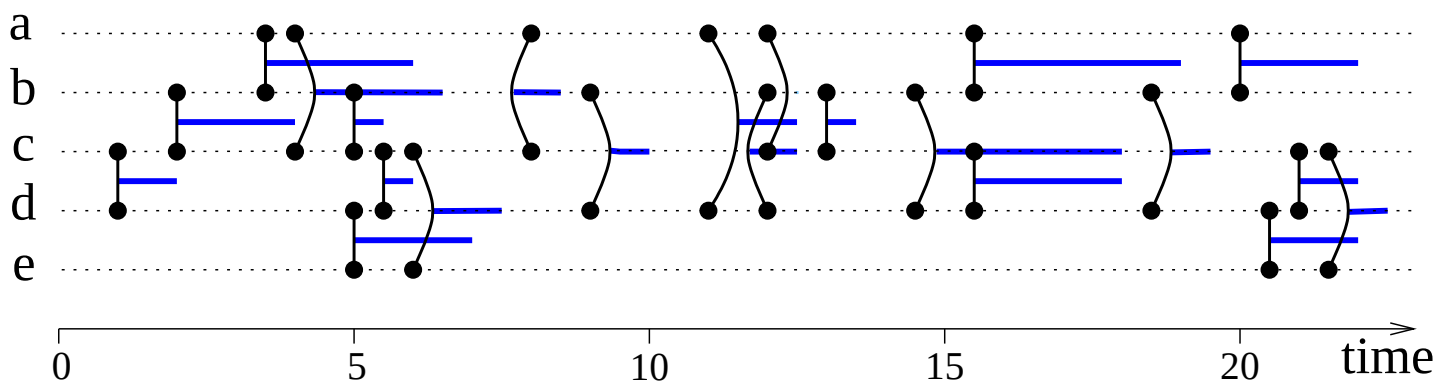
$\hookrightarrow$ already much studied

Our topic: link streams

countless examples

email exchanges, network traffic, payments,
physical contacts, phone calls, web surfing, ...

interactions over time



$$l = (b, e, u, v)$$

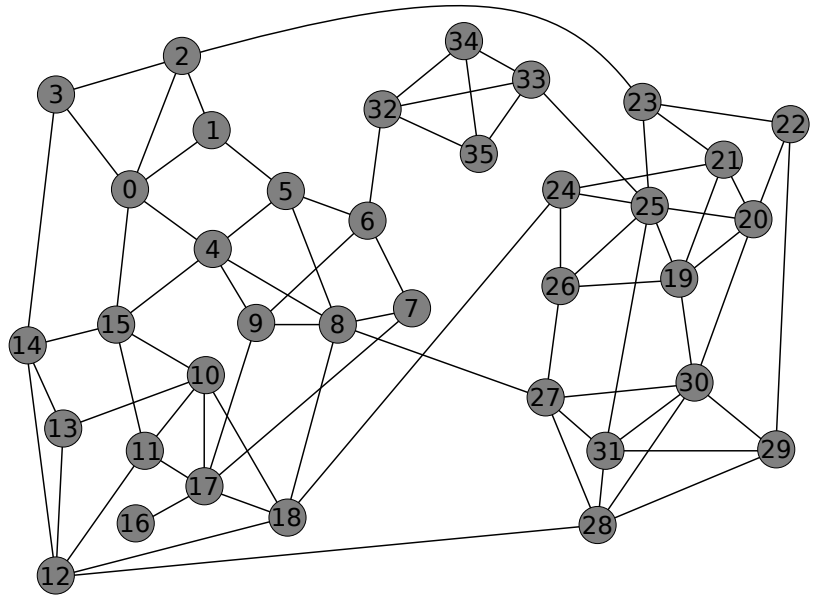
$b, e \in [\alpha, \omega]$: time
$u, v \in V$: nodes

$\hookrightarrow$ already much studied

Current situation (1/3)

focus on links: $\{(a, b)\}$

relations, structure

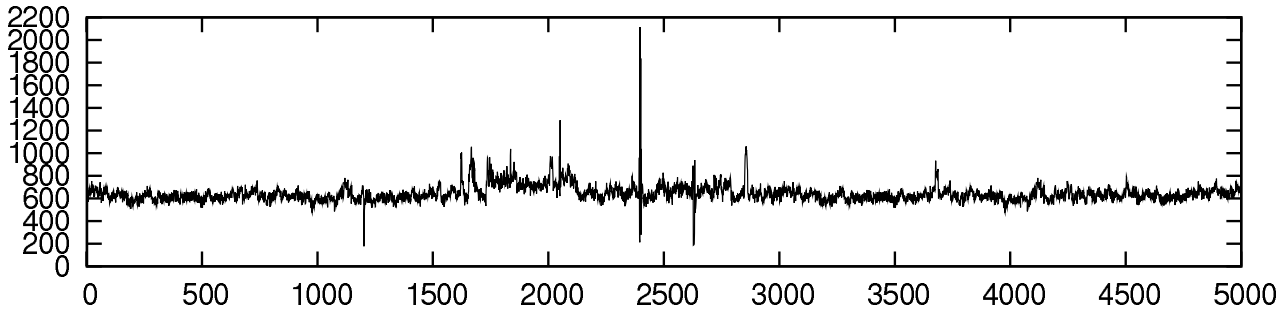


↪ **graph theory / network science**
density, degrees, clustering,
paths, diameter, distances, etc

Current situation (2/3)

focus on time: $\{(t, f(t))\}$

events, time series



$\hookrightarrow$ **signal processing / discrete event theory**
frequency, speed, inter-event times,
acceleration, self-similarity, periodicity, etc

Current situation (3/3)

sequences of graphs

split time into slices
one graph per slice

time-varying graphs (TVG)

graph with labelled edges
labels = times of presence

↪ **upgrades of graph and signal approaches**
many problems

Our proposal

a language for link streams
like graph theory for networks

to deal *directly* with link streams

describe them: what do them look like?

take advantage of their rich structure+time nature

$\hookrightarrow$ understand/detect events (attacks, anomalies), meetings, discussions, epidemies, ...

Wanted features

generalizes graphs and time series

simple and intuitive

bring fundamental *and applied* progress (e.g. event detection)

extensible (to weighted, directed, ...)

This presentation

key notions/intuitions of graphs/networks

translated to link streams

0. Basic notions
1. Density and related notions
2. Paths, distances, ...
3. Clusters and communities
4. Instantaneous links
5. Bipartite and other extensions

Upcoming...

Basic notions

What is it?

Graphs:

$G = (V, E)$, $E \subseteq V \times V$

links : (u, v)

u and v are linked together

Link streams :

$L = (T, V, E)$, $E \subseteq T \times T \times V \times V$

$l = (b, e, u, v)$

u and v are in interaction from b to e

simple, no overlap, undirected, etc
+ extensions

Sub-graphs and sub-streams

Graphs $G = (V, E)$ and $G' = (V', E')$:

G' **sub-graph** of G iff

$V' \subseteq V$ and $E' \subseteq E$

Links $l = (b, e, u, v)$ and $l' = (b', e', u', v')$:

l' **sub-link** of l iff

$u' = u$, $v' = v$, $[b', e'] \subseteq [b, e]$

Link streams $L = (T, V, E)$ and $L' = (T', V', E')$:

L' **sub-stream** of L iff

$V' \subseteq V$, $T' \subseteq T$,

and all links of L' are sub-links of links in L

Sub-graphs and sub-streams

Graphs $G = (V, E)$ and $G' = (V', E')$:

G' **sub-graph** of G iff

$V' \subseteq V$ and $E' \subseteq E$

Links $l = (b, e, u, v)$ and $l' = (b', e', u', v')$:

l' **sub-link** of l iff

$u' = u$, $v' = v$, $[b', e'] \subseteq [b, e]$

Link streams $L = (T, V, E)$ and $L' = (T', V', E')$:

L' **sub-stream** of L iff

$V' \subseteq V$, $T' \subseteq T$,

and all links of L' are sub-links of links in L

Sub-graphs and sub-streams

Graphs $G = (V, E)$ and $G' = (V', E')$:

G' **sub-graph** of G iff

$V' \subseteq V$ and $E' \subseteq E$

Links $l = (b, e, u, v)$ and $l' = (b', e', u', v')$:

l' **sub-link** of l iff

$u' = u$, $v' = v$, $[b', e'] \subseteq [b, e]$

Link streams $L = (T, V, E)$ and $L' = (T', V', E')$:

L' **sub-stream** of L iff

$V' \subseteq V$, $T' \subseteq T$,

and all links of L' are sub-links of links in L

Induced streams and graphs

Graph induced by a set of nodes or a set of links.

Link stream induced by a set of nodes, a time interval, or a set of (sub-)links.

+link stream induced by a pair of nodes and by a node.

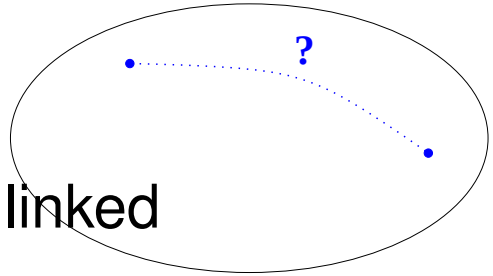
Graph induced by a link stream.

$\hookrightarrow$ *Sequence of graphs over time-windows of duration* Δ*:* $G(L_{t..t+\Delta})$

Upcoming...

Density
and
related notions

Density



Graphs:
proba two random nodes are linked

$$\delta(G) = \frac{2 \cdot m}{n \cdot (n-1)}$$

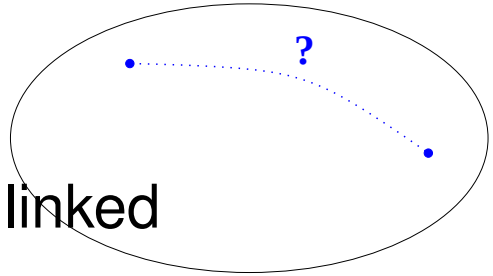
Link streams:
proba two random nodes are linked
at a random time instant

$$\delta(L) = \frac{2 \cdot \sum_l \bar{l}}{n \cdot (n-1) \cdot (\omega - \alpha)}$$

$\bar{l}$: duration of link l

Note: if $\bar{l} = \omega - \alpha$ for all l, then graph density

Density



Graphs:
proba two random nodes are linked

$$\delta(G) = \frac{2 \cdot m}{n \cdot (n-1)}$$

Link streams:
proba two random nodes are linked
at a random time instant

$$\delta(L) = \frac{2 \cdot \sum_l \bar{l}}{n \cdot (n-1) \cdot (\omega - \alpha)}$$

$\bar{l}$: duration of link l

Note: if $\bar{l} = \omega - \alpha$ for all l, then graph density

Degree

Graphs: size of the neighborhood

$$d(v) = |N(v)|$$

Link streams: what neighborhood?

each neighbor weighted by its link duration :

$$d(v) = \sum_{l \in L(v)} \frac{\bar{l}}{\omega - \alpha}$$

In graphs *and* in link streams : $\delta = \frac{d}{n-1}$

Degree

Graphs: size of the neighborhood

$$d(v) = |N(v)|$$

Link streams: what neighborhood?

each neighbor weighted by its link duration :

$$d(v) = \sum_{l \in L(v)} \frac{\bar{l}}{\omega - \alpha}$$

In graphs *and* in link streams : $\delta = \frac{d}{n-1}$

Degree

Graphs: size of the neighborhood

$$d(v) = |N(v)|$$

Link streams: what neighborhood?

each neighbor weighted by its link duration :

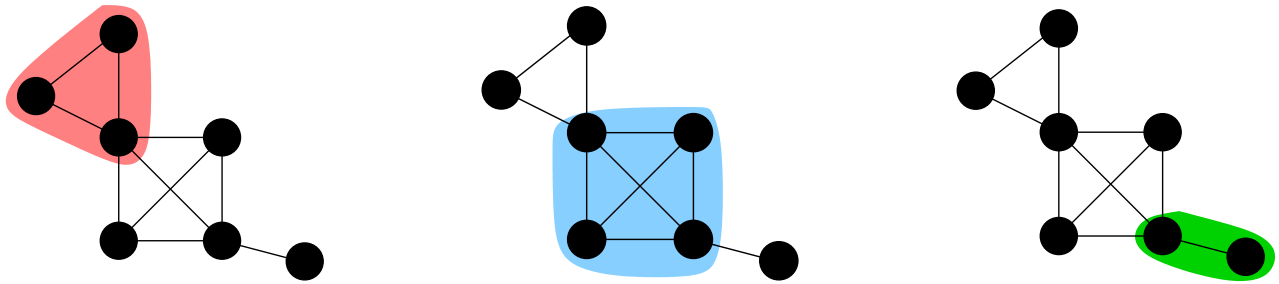
$$d(v) = \sum_{l \in L(v)} \frac{\bar{l}}{\omega - \alpha}$$

In graphs *and* in link streams : $\delta = \frac{d}{n-1}$

(Maximal) cliques in graphs

Graphs: (maximal) sub-graph of density 1

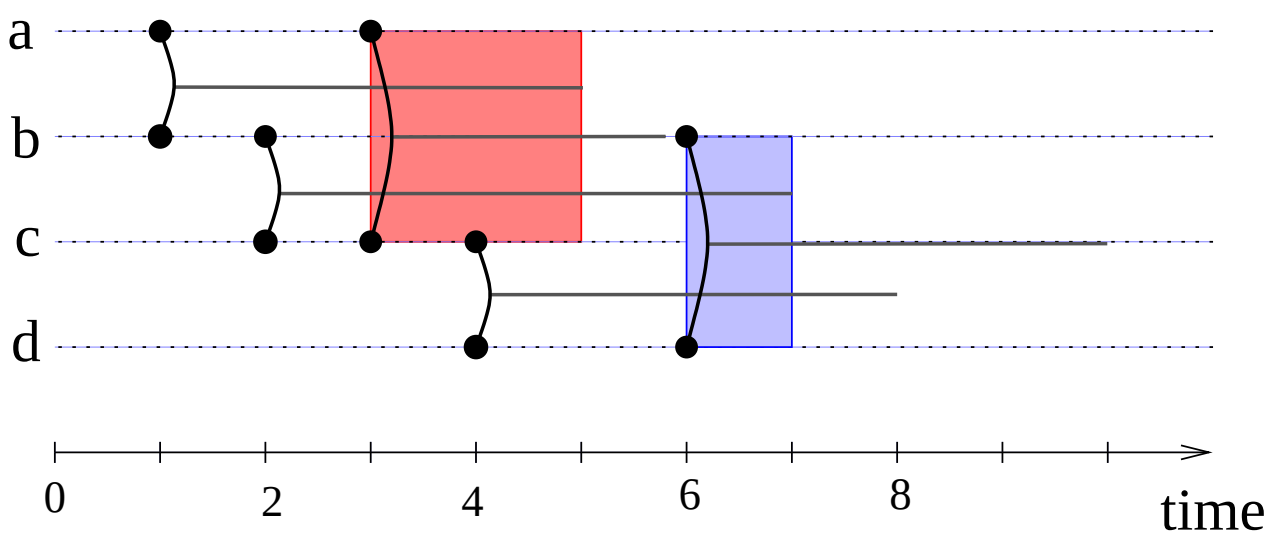
all nodes are linked together



(Maximal) cliques **in link streams**

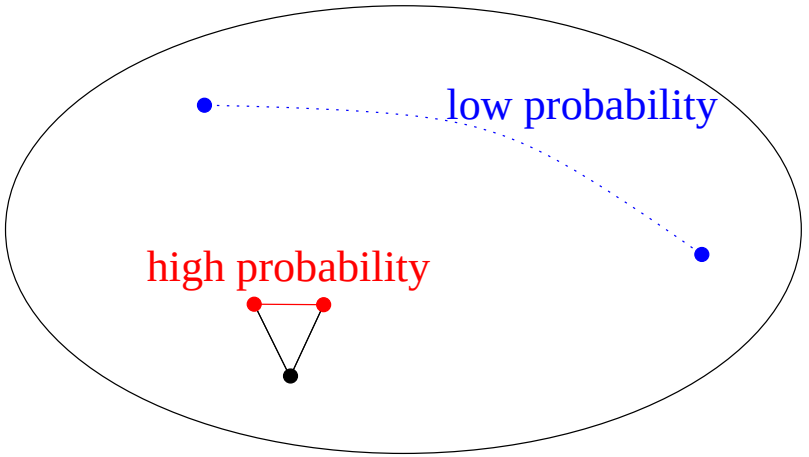
the same: (maximal) sub-stream of density 1

all nodes interact all the time



Clustering coefficient in graphs

intuition: “my friends are friends with each other”
- low global density
- high local density



***clustering coefficient*:**
density of neighborhood

to what point all neighbors are linked together

Clustering coefficient **in link streams**

the same?

density of neighborhood

to what point all neighbors interact all the time

each neighbor weighted by its link duration

Upcoming...

Paths,

distances,

centralities,

...

Paths

Graphs: sequences of links (u_i, v_i) such that $u_i = v_{i-1}$

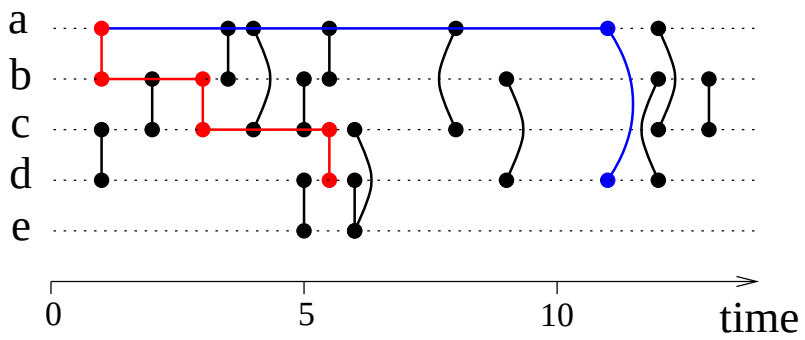
Link streams: sequences of triplets (t_i, u_i, v_i) such that $u_i = v_{i-1}$ and $t_i \geq t_{i-1}$

Links with duration: sequences of sub-links $(t_i, t_i + \gamma, u_i, v_i)$ such that $u_i = v_{i-1}$ and $t_i \geq t_{i-1} + \gamma$

Paths

Graphs: sequences of links (u_i, v_i) such that $u_i = v_{i-1}$

Link streams: sequences of triplets (t_i, u_i, v_i) such that $u_i = v_{i-1}$ and $t_i \geq t_{i-1}$

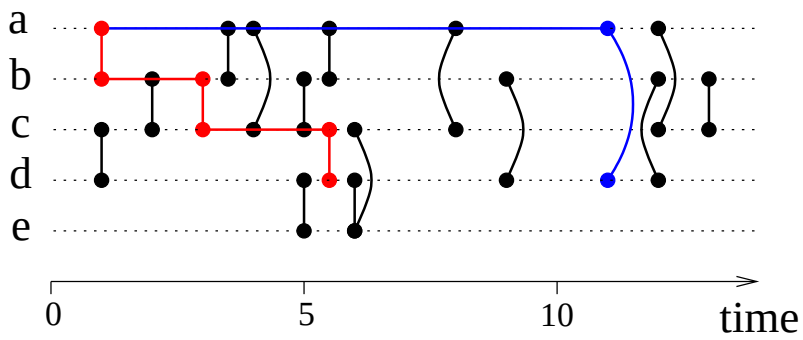


Links with duration: sequences of sub-links $(t_i, t_i + \gamma, u_i, v_i)$ such that $u_i = v_{i-1}$ and $t_i \geq t_{i-1} + \gamma$

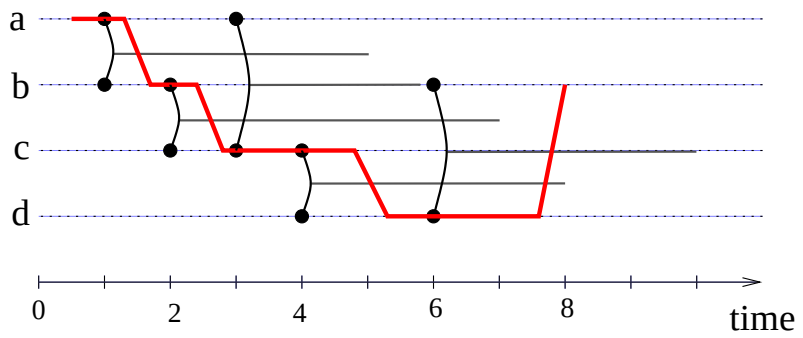
Paths

Graphs: sequences of links (u_i, v_i) such that $u_i = v_{i-1}$

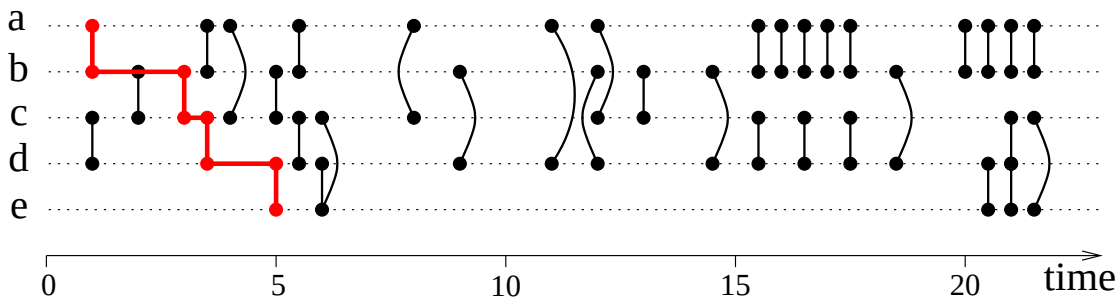
Link streams: sequences of triplets (t_i, u_i, v_i) such that $u_i = v_{i-1}$ and $t_i \geq t_{i-1}$



Links with duration: sequences of sub-links $(t_i, t_i + \gamma, u_i, v_i)$ such that $u_i = v_{i-1}$ and $t_i \geq t_{i-1} + \gamma$

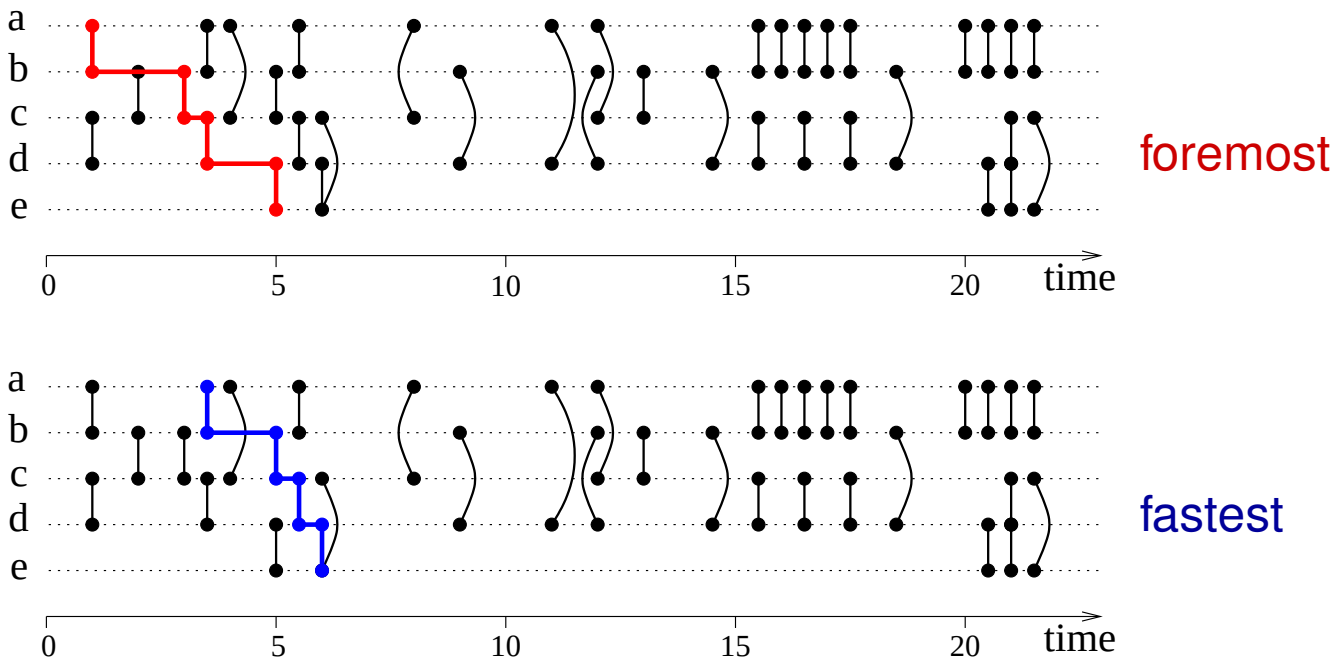


Distances in link streams

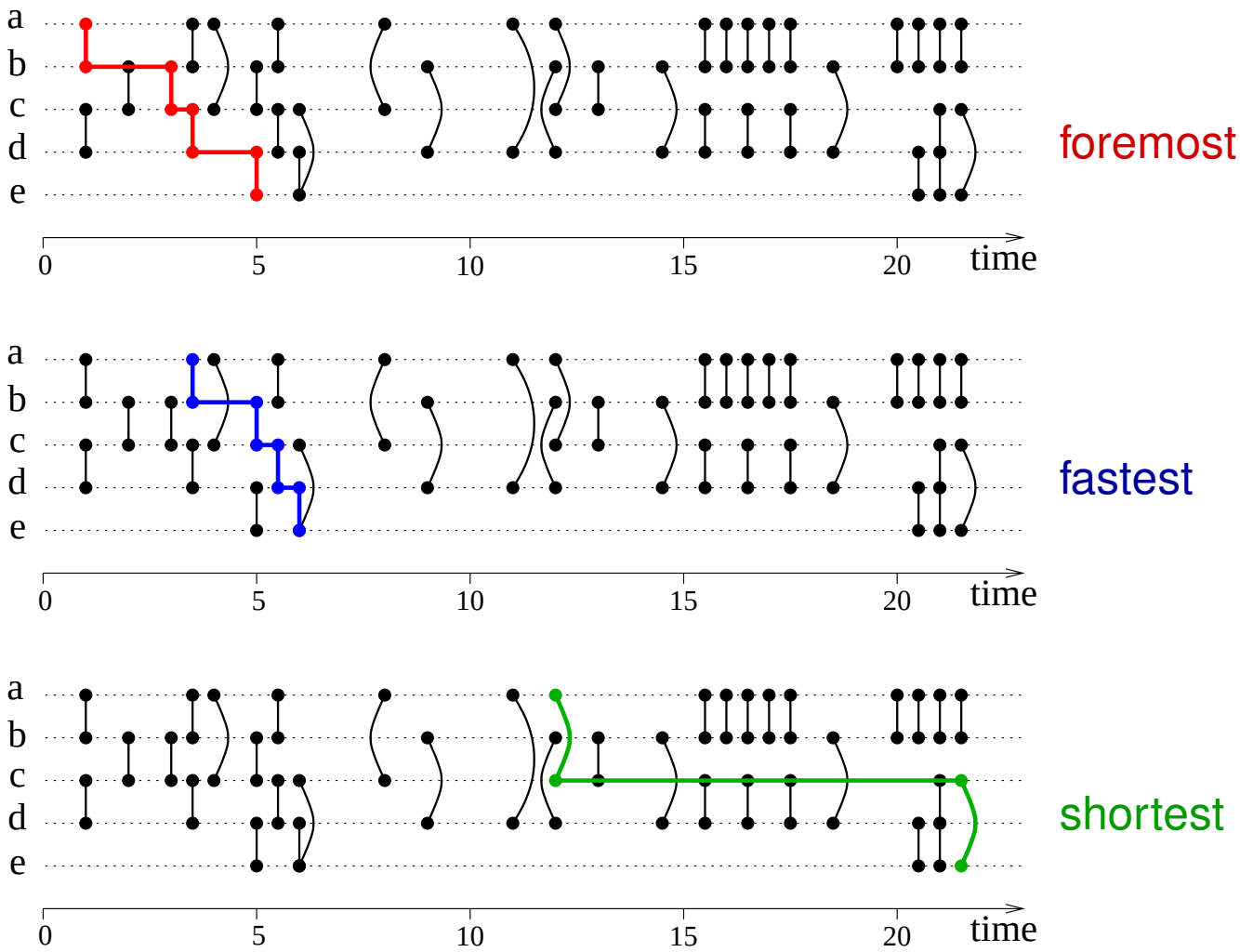


foremost

Distances in link streams



Distances in link streams



Centralities

Graphs: closeness, betweeness, ...

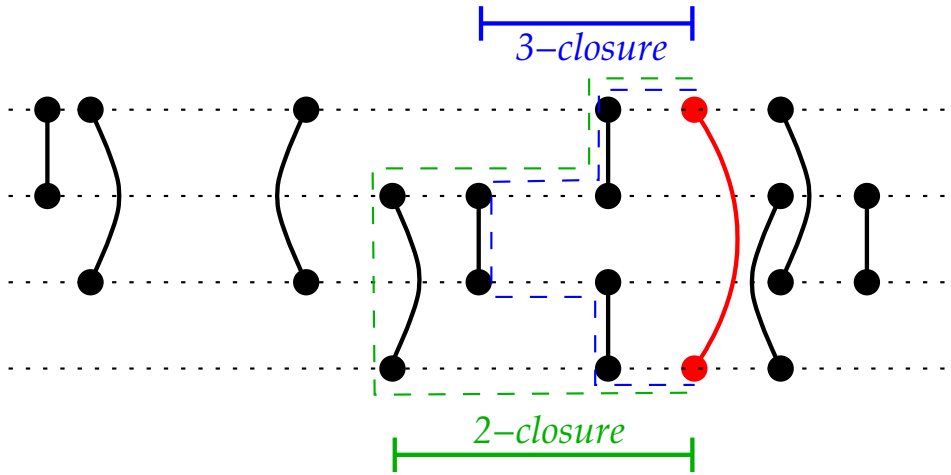
Link streams: **centrality of node v at time t**; centrality of v? of time t?

closeness: easy; betweeness: number of fastest paths?

k-closure

k-closure of (t, a, b):

time until a and b at distance $\leq k$



Notes:

- $k = 1 \longrightarrow$ inter-contact times
- $k = 2 \longrightarrow$ clustering coefficient

mix of time and structure

Going further

trees, spreading

(strong) connectedness, connected components, connecting components, ...

reachability is not symmetric

monsters: connected parts

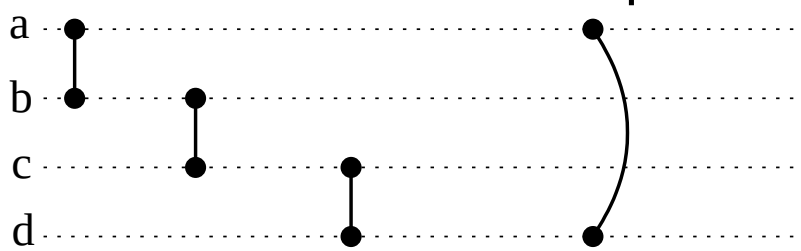
Going further

trees, spreading

(strong) connectedness, connected components, connecting components, ...

reachability is not symmetric

monsters: connected parts

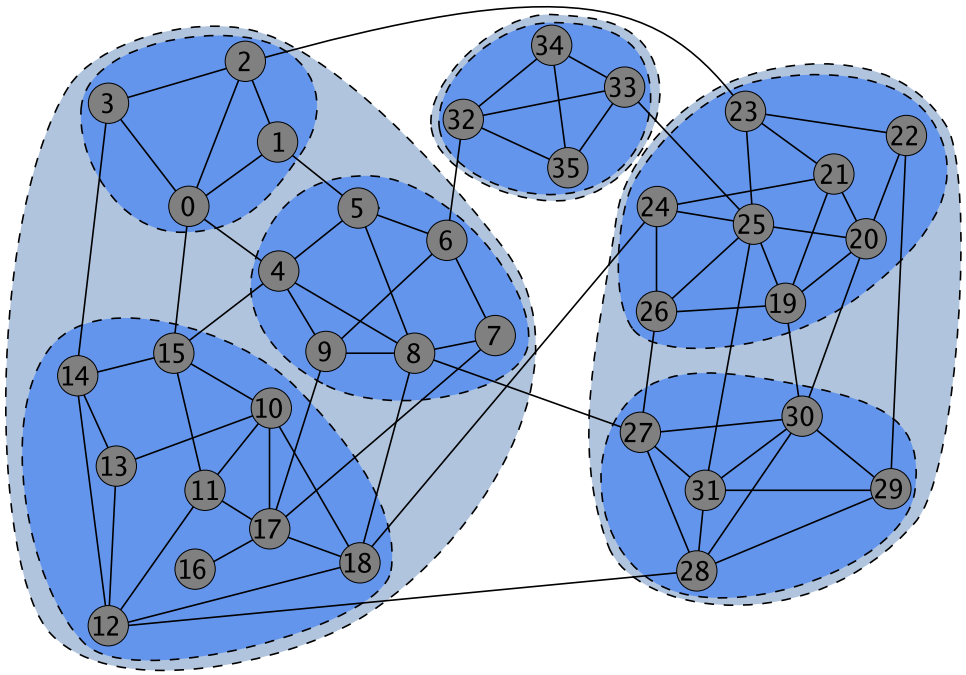


Upcoming...

Communities

Communities in graphs

dense sub-graphs poorly interconnected

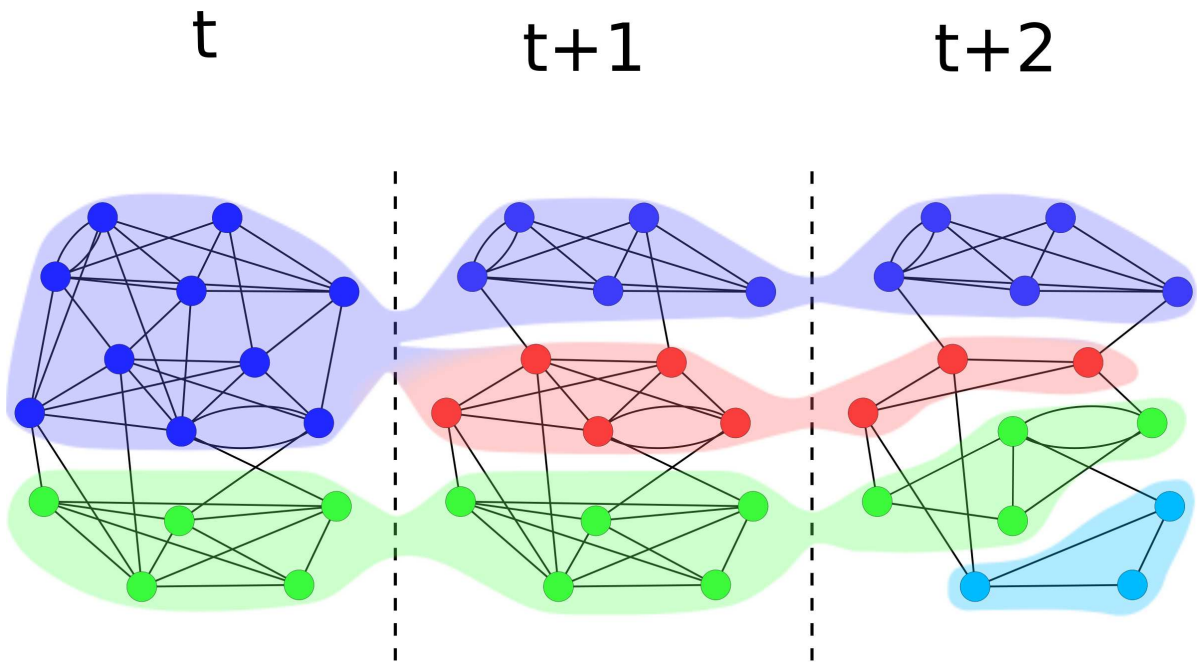


ex: groups of friends, of computers, of products, ...

how to define them? detect them?
hierarchies? overlaps? ...

Communities in *dynamic* graphs

evolution of graph communities

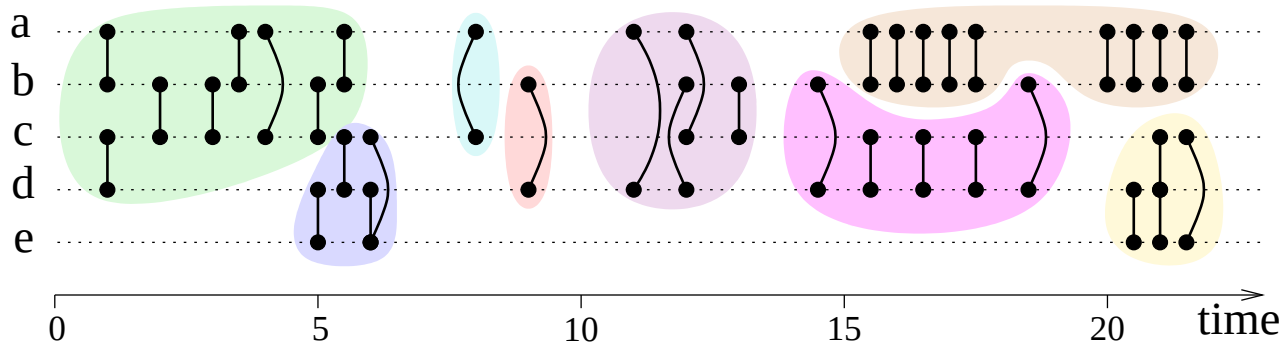


ex: groups of friends evolving over time

Communities in link streams

dense sub-streams poorly interconnected

i.e. temporally and structurally dense series of interactions



ex: discussions, meetings, sessions, ...

link streams $\neq$ dynamic graphs

Going further...

intra-cluster density
inter-cluster density

quotient link stream

quality functions
modularity

algorithms

line stream

Instantaneous link streams

discrete time instants?
not relevant

needs a Δ
$\hookrightarrow$ Δ-analysis of link streams
(ex: Δ-density)

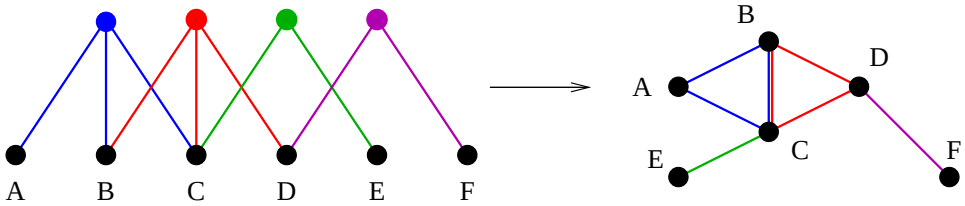
equivalent to links with duration Δ

+ Δ may vary with time, nodes, and more complex features

Bipartite link streams

two kinds of nodes
links only between nodes of different kinds
(client-product, author-paper, actor-movie, ...)

projection of a bipartite graph :

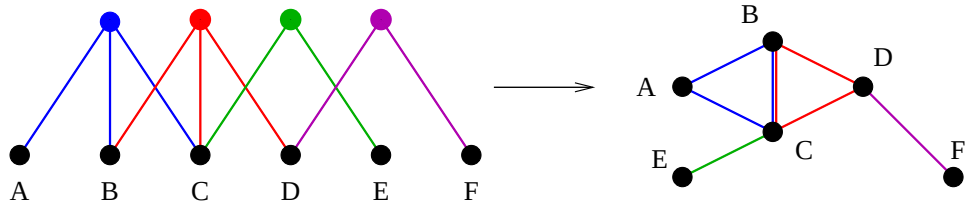


Projection of a bipartite link stream...
into a link stream.

Bipartite link streams

two kinds of nodes
links only between nodes of different kinds
(client-product, author-paper, actor-movie, ...)

projection of a bipartite graph :



Projection of a bipartite link stream...
into a link stream.

A bit of philosophy

graph/networks = relations
(like friendship)

dynamic graphs/networks = evolution of relations
(like new friends)

link streams = interactions
(like phone calls)

interactions = traces/realization of relations?

link streams = traces of graphs/networks?

Conclusion

link streams model interactions over time
link streams $\neq$ dynamic graphs

a language for link streams
simple? intuitive? general? powerful? ...

- **In progress:** actual communities, event and community detection, relations with TVG
- **Case studies:** mailing-lists (Debian), phone calls (D4D), network traffic (Mawi, companies), mobility/contacts (crawdad, sociopatterns), financial transactions (bitcoins, on-line shopping), etc
- **Extensions:** strength, direction, etc of interactions $\rightarrow$ weighted, bipartite, directed, etc link streams