LSCPM: finding communities in Link Streams by Clique Percolation Method

Alexis Baudin*, Lionel Tabourier and Clémence Magnien

June 21st, 2023

Lip6, Sorbonne Université, CNRS, Complex Networks

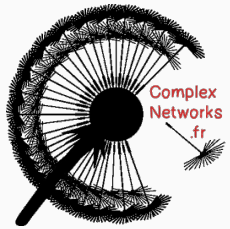
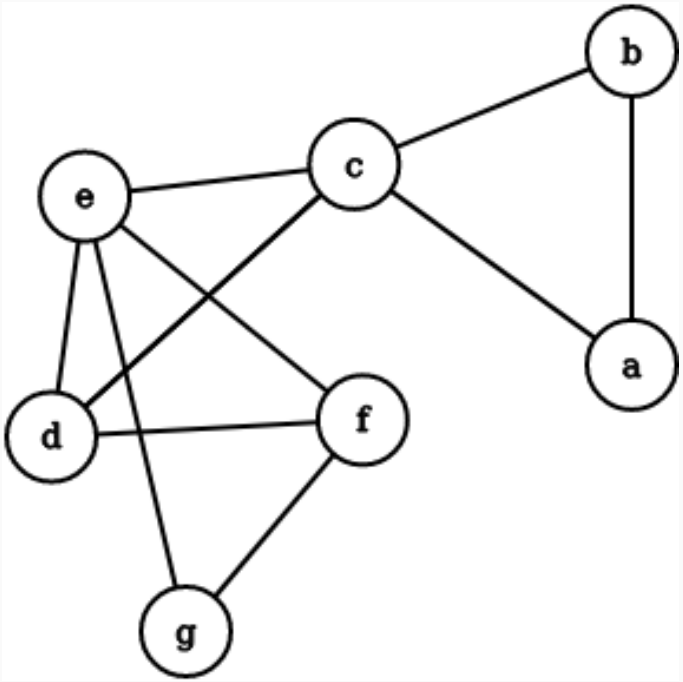


Table of contents

1 - Link stream formalism to model temporal data

> Definition – Graph



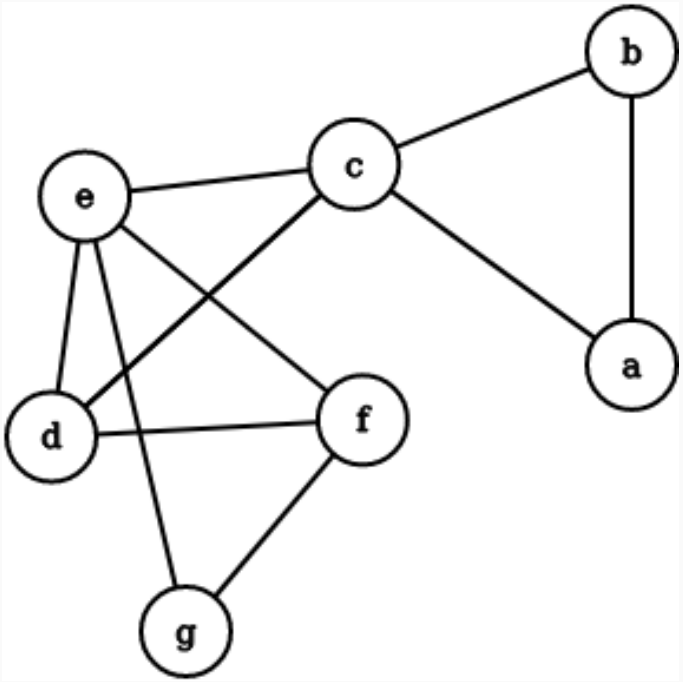
Graph formalism

- Vertices: $a, b, ..., g$
- Interactions: edges

Example

Contacts between people, brain networks, ...

> Definition – Graph



Graph formalism

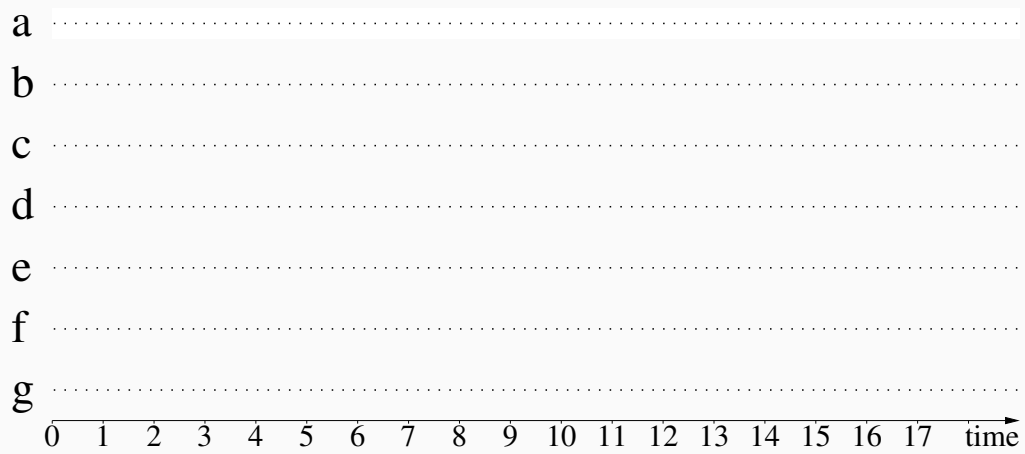
- Vertices: $a, b, ..., g$
- Interactions: edges

Example

Contacts between people, brain networks, ...

$\rightarrow$ What about temporal interaction ?

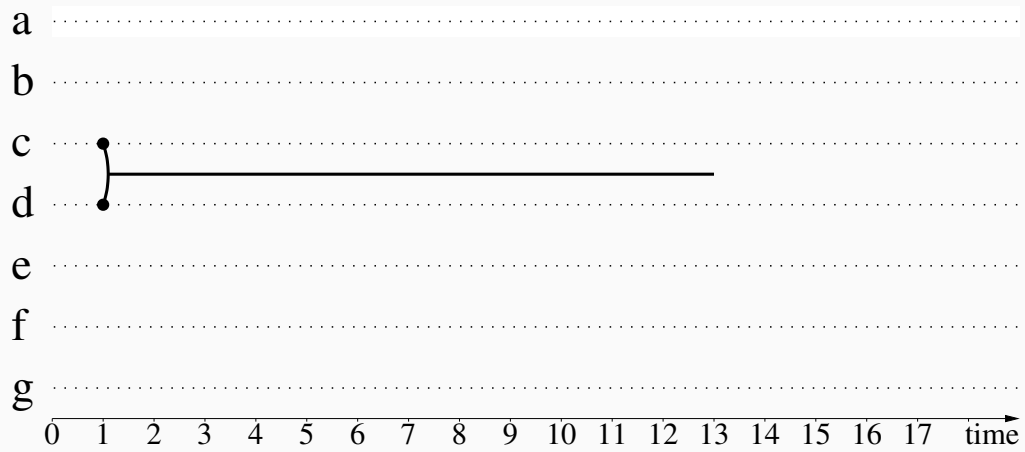
> Definition – Link stream



Link stream formalism

- Vertices: $a, b, ..., g$
- Time period: $[0, 18]$
- Interactions: temporal edges

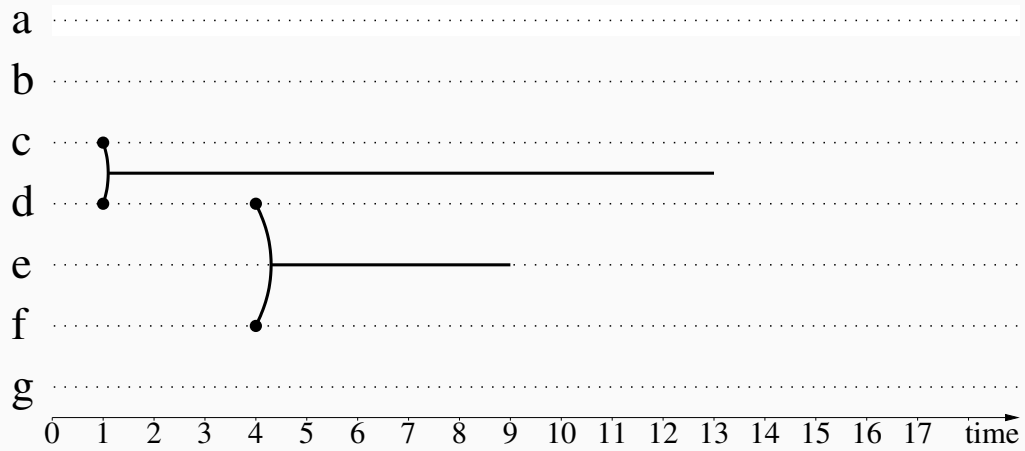
> Definition – Link stream



Link stream formalism

- Vertices: $a, b, ..., g$
- Time period: $[0, 18]$
- Interactions: temporal edges
  - c, d linked over $[1, 13]$

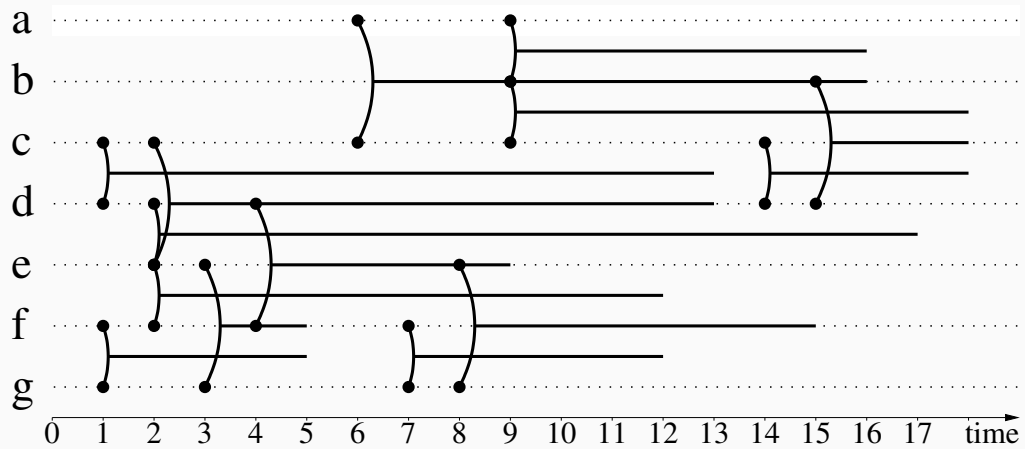
> Definition – Link stream



Link stream formalism

- Vertices: $a, b, ..., g$
- Time period: $[0, 18]$
- Interactions: temporal edges
  - c, d linked over $[1, 13]$
  - d, f linked over $[4, 9]$

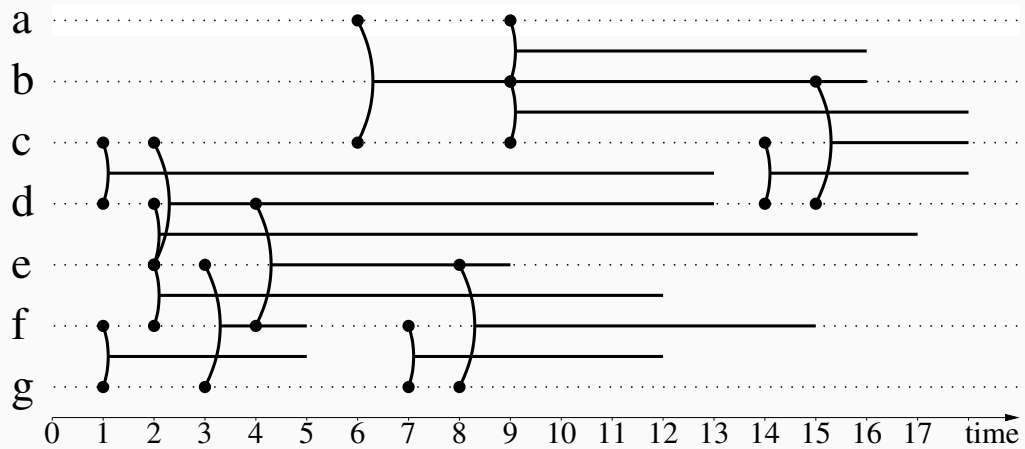
> Definition – Link stream



Link stream formalism

- Vertices: $a, b, ..., g$
- Time period: $[0, 18]$
- Interactions: temporal edges
  - c, d linked over $[1, 13]$
  - d, f linked over $[4, 9]$
  - ...

> Definition – Link stream



Link stream formalism

- Vertices: $a, b, ..., g$
- Time period: $[0, 18]$
- Interactions: temporal edges
  - c, d linked over $[1, 13]$
  - d, f linked over $[4, 9]$
  - ...

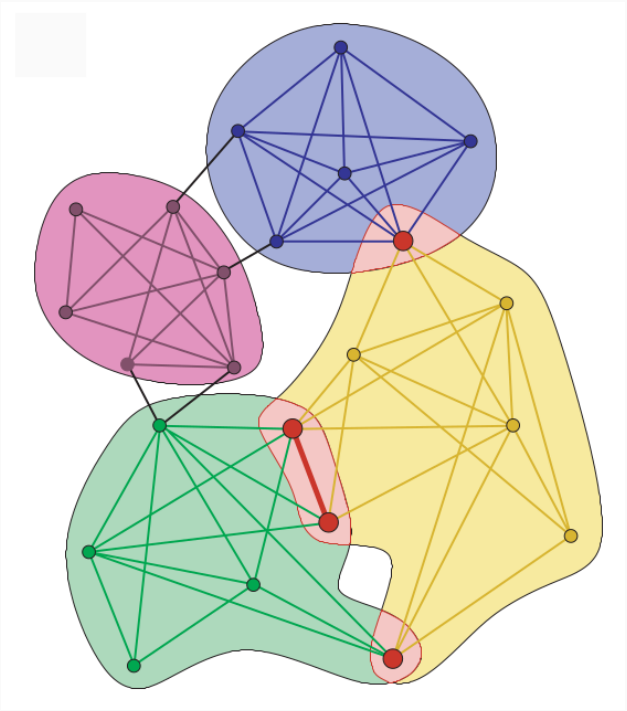
Advantages

- deals directly with the stream of interactions
- no arbitrary choice of time scale
- time is continuous

2 - Describing temporal data by finding communities in link streams

Communities: sets of vertices

- Densely connected inside
- Sparsly connected outside



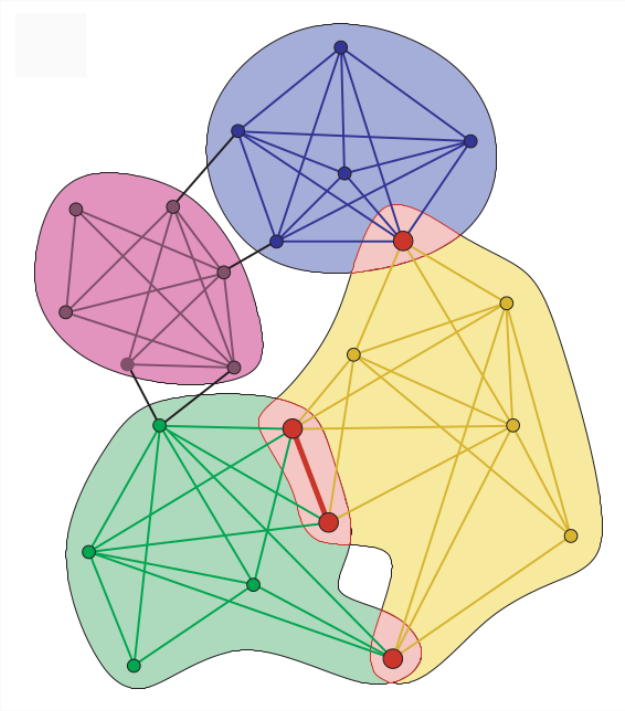
Palla et al. 2005

Communities: sets of vertices

- Densely connected inside
- Sparsly connected outside

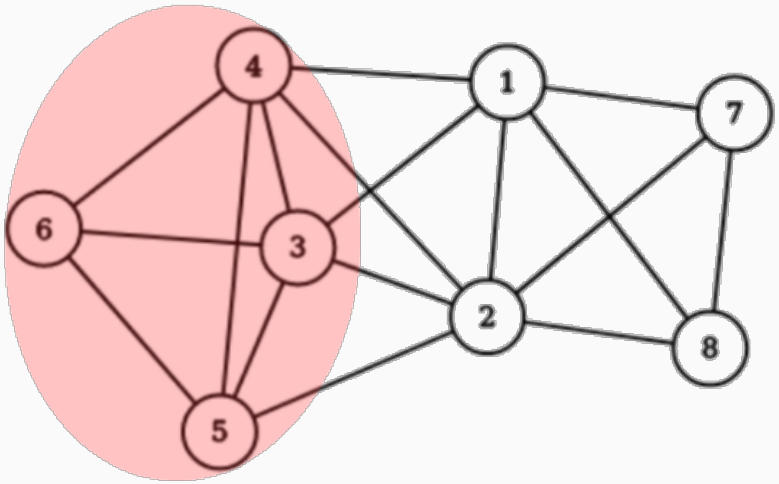
Interest:

- Locate areas of high interaction density
- Understanding the organizational structure of interactions
- Zoom in / out



Palla et al. 2005

> Clique Percolation Method in static networks (CPM)

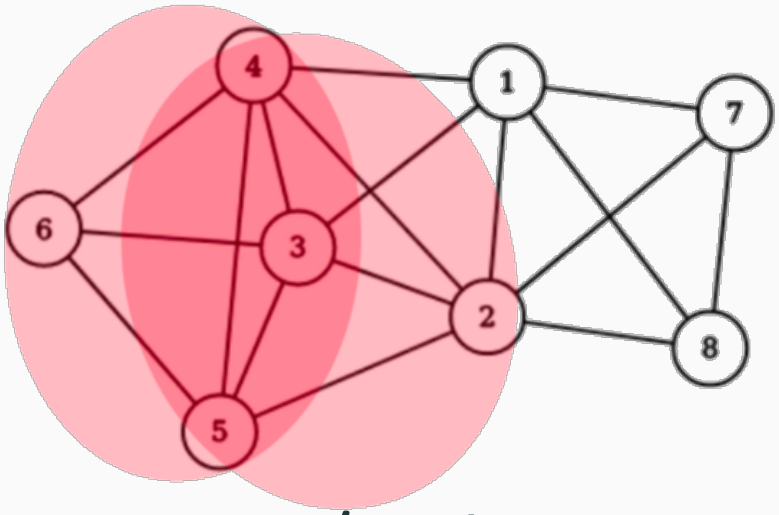


$k = 4$

k-clique

Set of k nodes all connected to each other.

> Clique Percolation Method in static networks (CPM)



$k = 4$

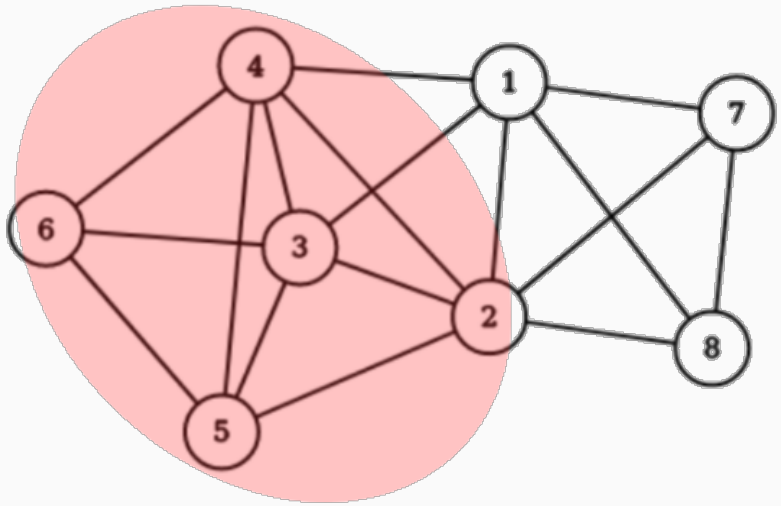
k-clique

Set of k nodes all connected to each other.

Grouping rule

Two k-cliques are adjacent if they share $k - 1$ nodes.

> Clique Percolation Method in static networks (CPM)



$k = 4$

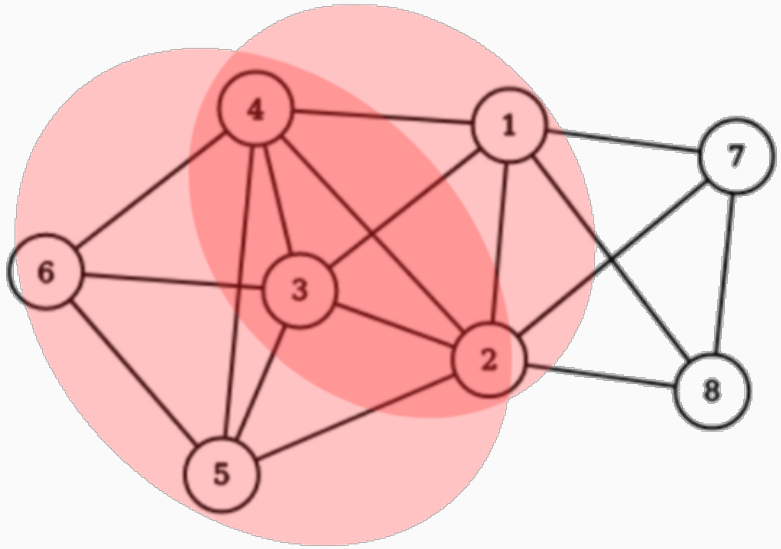
k-clique

Set of k nodes all connected to each other.

Grouping rule

Two k-cliques are adjacent if they share $k - 1$ nodes.

> Clique Percolation Method in static networks (CPM)



$k = 4$

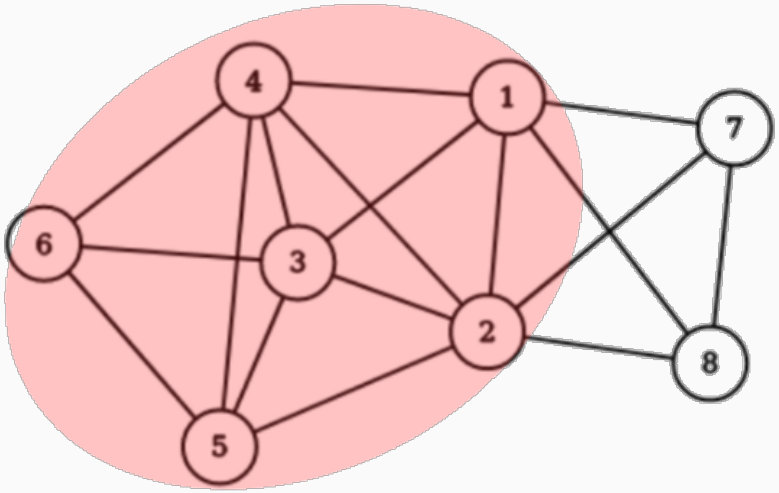
k-clique

Set of k nodes all connected to each other.

Grouping rule

Two k-cliques are adjacent if they share $k - 1$ nodes.

> Clique Percolation Method in static networks (CPM)



$k = 4$

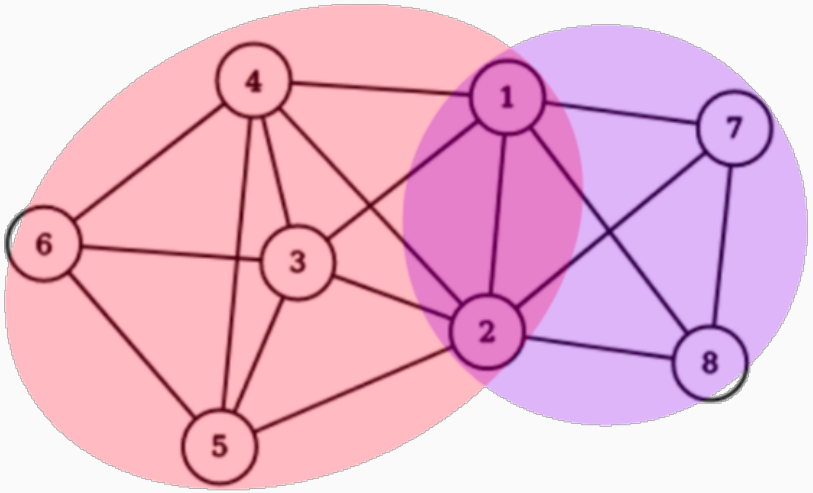
k-clique

Set of k nodes all connected to each other.

Grouping rule

Two k-cliques are adjacent if they share $k - 1$ nodes.

> Clique Percolation Method in static networks (CPM)



$k = 4$

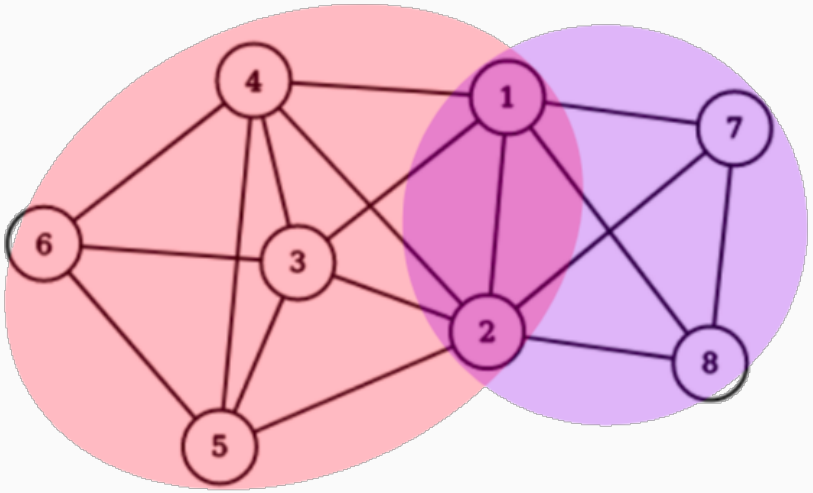
k-clique

Set of k nodes all connected to each other.

Grouping rule

Two k-cliques are adjacent if they share $k - 1$ nodes.

> Clique Percolation Method in static networks (CPM)



$k = 4$

k-clique

Set of k nodes all connected to each other.

Grouping rule

Two k-cliques are adjacent if they share $k - 1$ nodes.

Advantages of the definition

- Definition local
- Deterministic; no need of heuristic function
- Communities can overlap

> First extension of CPM to Temporal Graphs (DCPM)

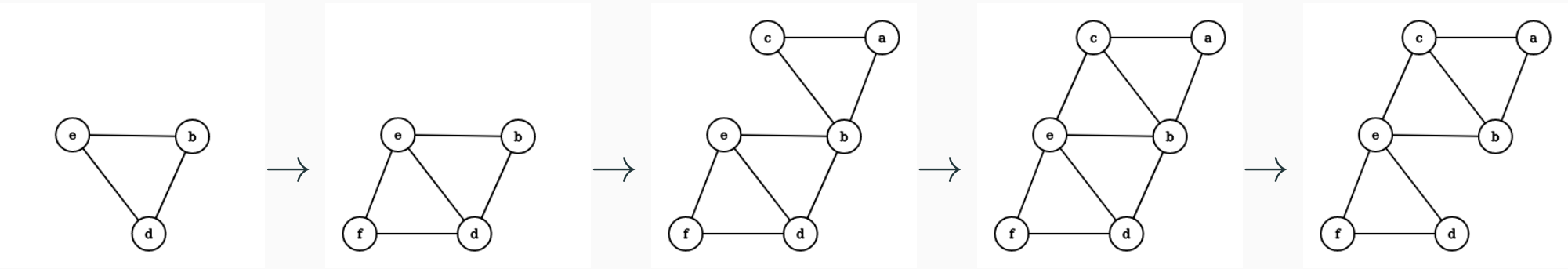
Question

How to extend CPM communities to temporal networks ?

CPM communities in temporal graphs: *Palla et. al 2007*

$\Rightarrow$ Communities that evolve from one time to the next

$k = 3$:



$t = 1$ $t = 2$ $t = 3$ $t = 4$ $t = 5$

> First extension of CPM to Temporal Graphs (DCPM)

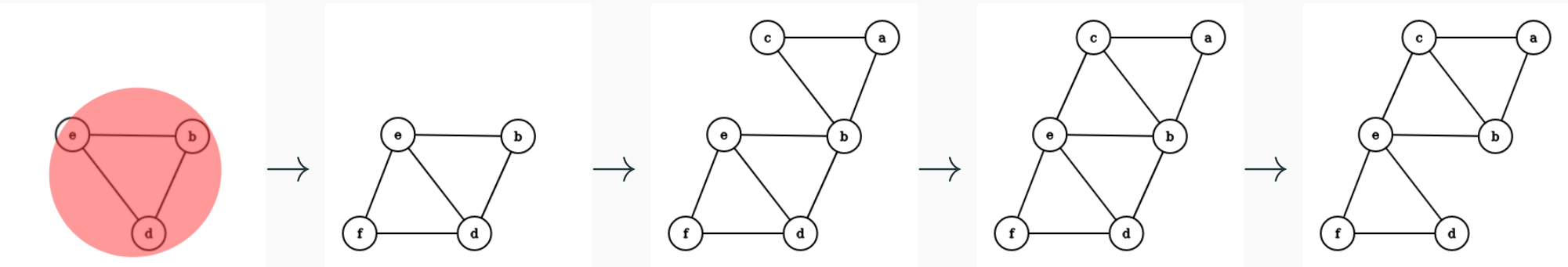
Question

How to extend CPM communities to temporal networks ?

CPM communities in temporal graphs: *Palla et. al 2007*

$\Rightarrow$ Communities that evolve from one time to the next

$k = 3$:



$t = 1$ $t = 2$ $t = 3$ $t = 4$ $t = 5$

> First extension of CPM to Temporal Graphs (DCPM)

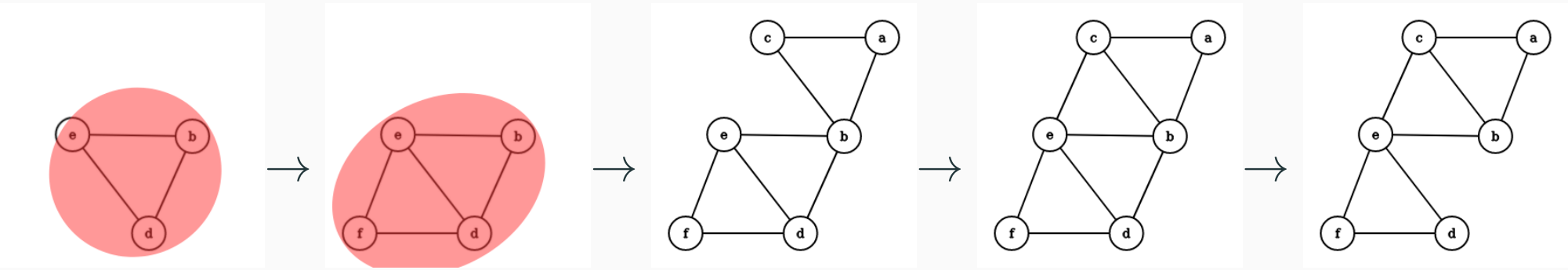
Question

How to extend CPM communities to temporal networks ?

CPM communities in temporal graphs: *Palla et. al 2007*

$\Rightarrow$ Communities that evolve from one time to the next

$k = 3$:



$t = 1$ $t = 2$ $t = 3$ $t = 4$ $t = 5$

> First extension of CPM to Temporal Graphs (DCPM)

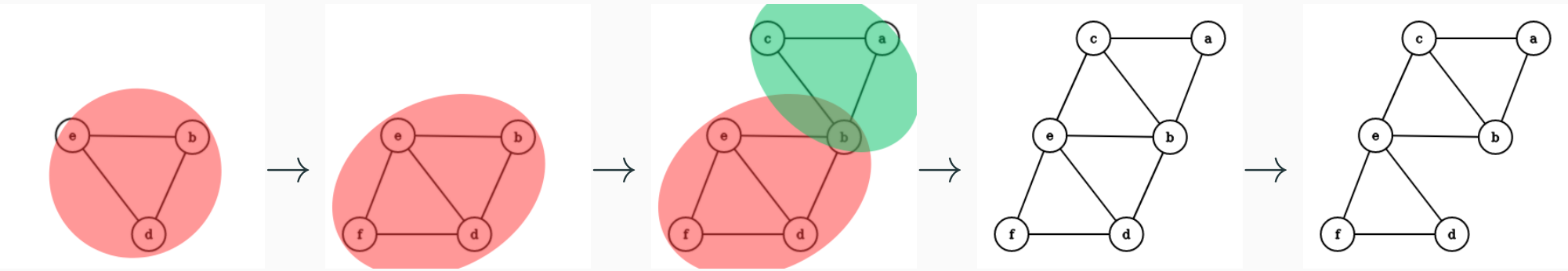
Question

How to extend CPM communities to temporal networks ?

CPM communities in temporal graphs: *Palla et. al 2007*

$\Rightarrow$ Communities that evolve from one time to the next

$k = 3$:



$t = 1$ $t = 2$ $t = 3$ $t = 4$ $t = 5$

> First extension of CPM to Temporal Graphs (DCPM)

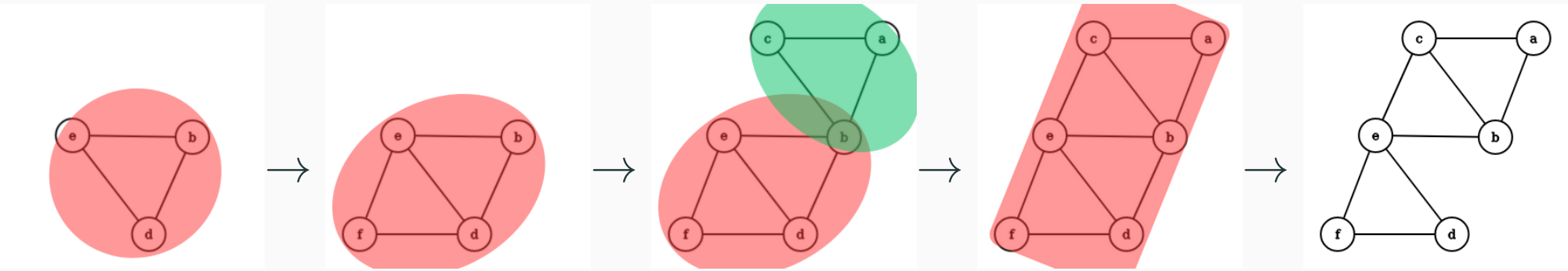
Question

How to extend CPM communities to temporal networks ?

CPM communities in temporal graphs: *Palla et. al 2007*

$\Rightarrow$ Communities that evolve from one time to the next

$k = 3$:



$t = 1$ $t = 2$ $t = 3$ $t = 4$ $t = 5$

> First extension of CPM to Temporal Graphs (DCPM)

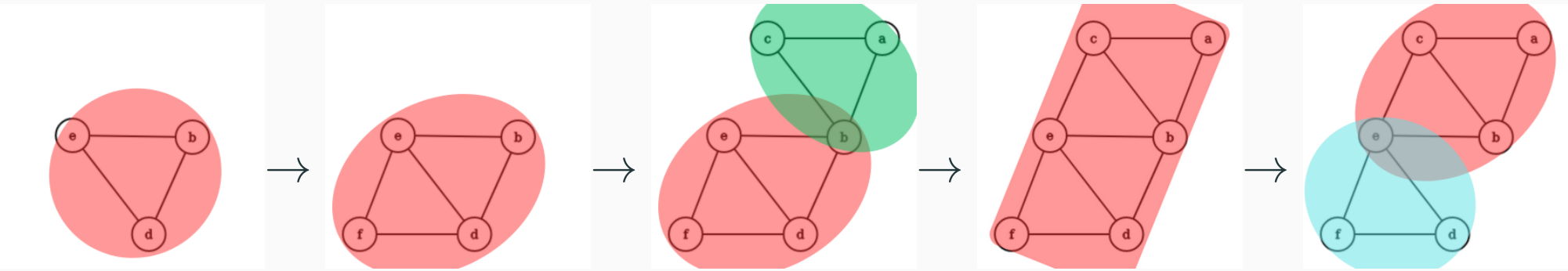
Question

How to extend CPM communities to temporal networks ?

CPM communities in temporal graphs: *Palla et. al 2007*

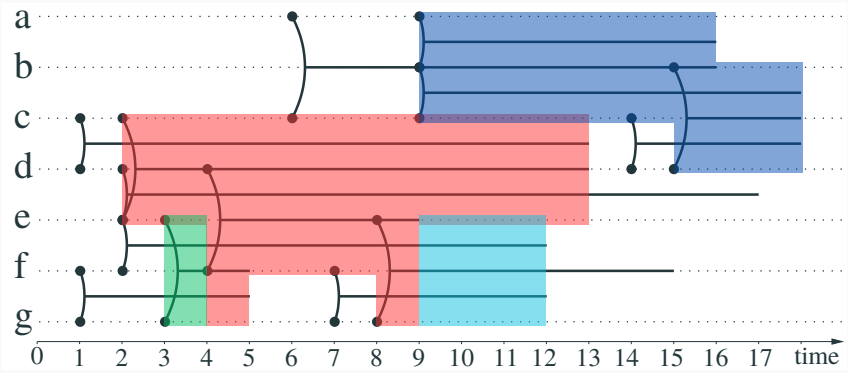
$\Rightarrow$ Communities that evolve from one time to the next

$k = 3$:



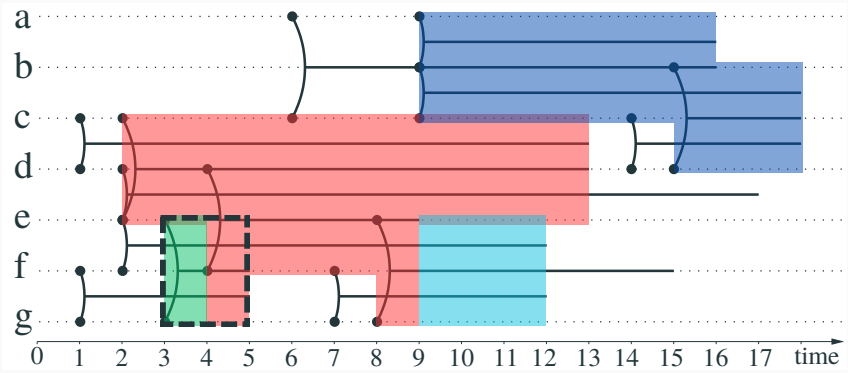
$t = 1$ $t = 2$ $t = 3$ $t = 4$ $t = 5$

CPM in temporal graphs



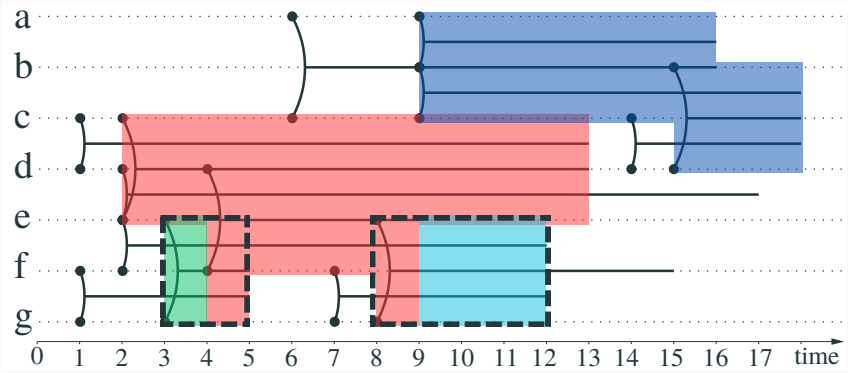
- Computing communities at each time step: time consuming;
- Some temporal data expected to be grouped are not.

CPM in temporal graphs



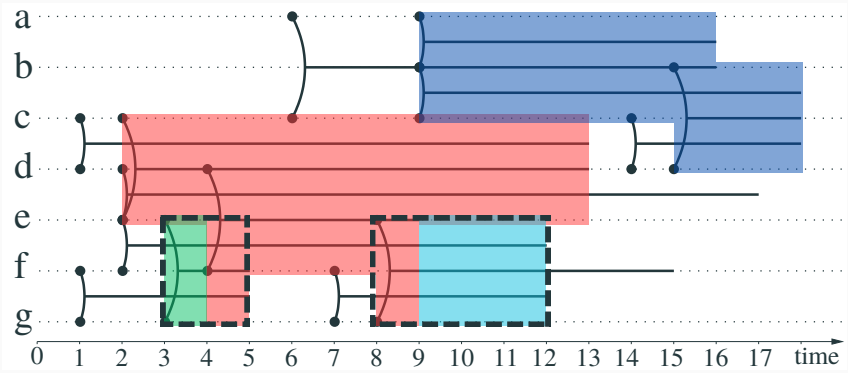
- Computing communities at each time step: time consuming;
- Some temporal data expected to be grouped are not.

CPM in temporal graphs



- Computing communities at each time step: time consuming;
- Some temporal data expected to be grouped are not.

CPM in temporal graphs



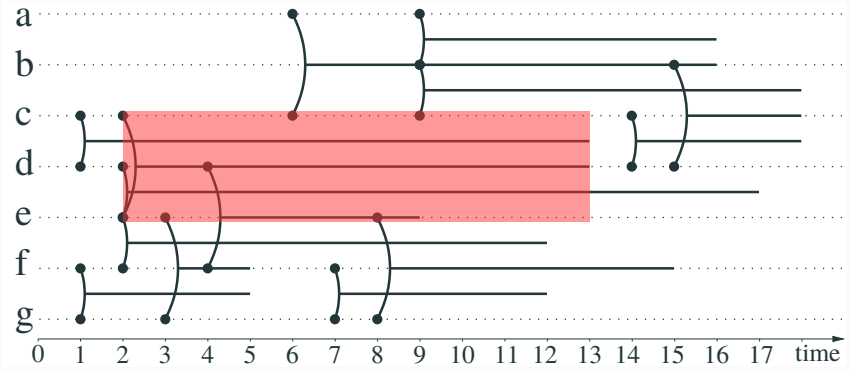
- Computing communities at each time step: time consuming;
- Some temporal data expected to be grouped are not.

CPM in link stream (LSCPM)

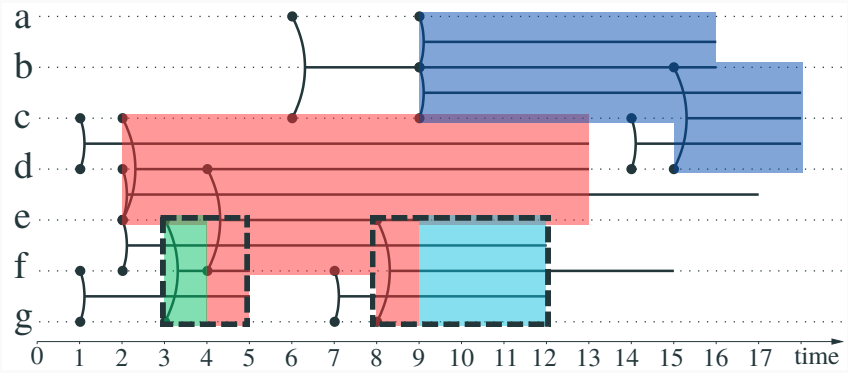
k-clique in link stream

k nodes and $[t_0, t_1]$ such that all nodes are connected to each other over $[t_0, t_1]$.

$\Rightarrow$ *Grouping rule with $k = 3$:*



CPM in temporal graphs



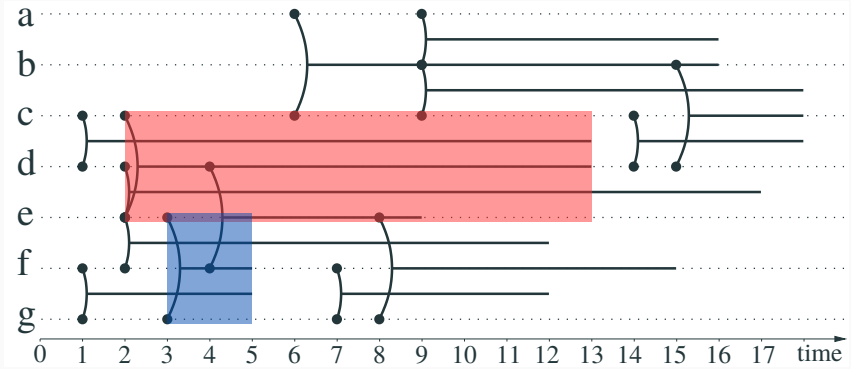
- Computing communities at each time step: time consuming;
- Some temporal data expected to be grouped are not.

CPM in link stream (LSCPM)

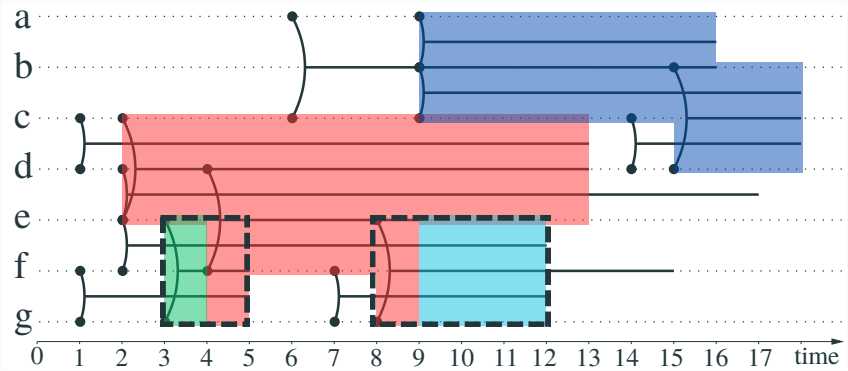
k-clique in link stream

k nodes and $[t_0, t_1]$ such that all nodes are connected to each other over $[t_0, t_1]$.

$\Rightarrow$ *Grouping rule with $k = 3$:*



CPM in temporal graphs



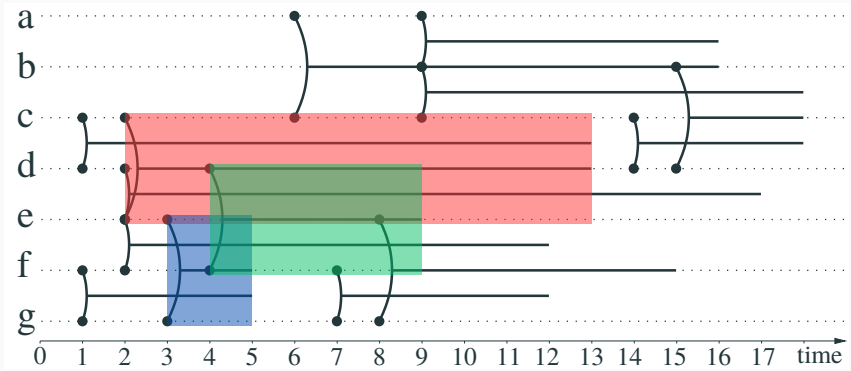
- Computing communities at each time step: time consuming;
- Some temporal data expected to be grouped are not.

CPM in link stream (LSCPM)

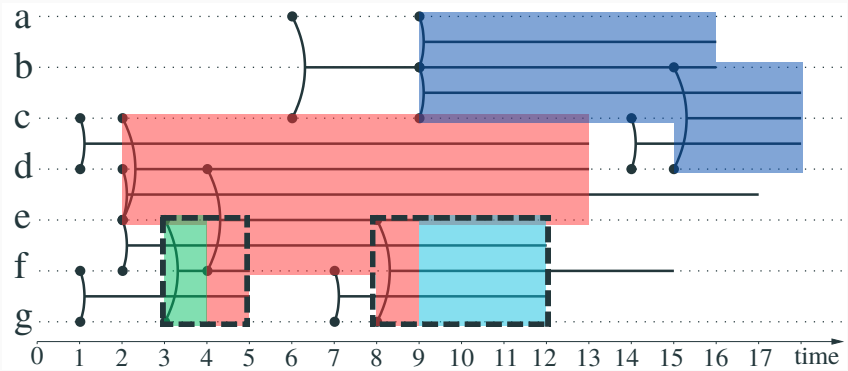
k-clique in link stream

k nodes and $[t_0, t_1]$ such that all nodes are connected to each other over $[t_0, t_1]$.

$\Rightarrow$ *Grouping rule with $k = 3$:*



CPM in temporal graphs



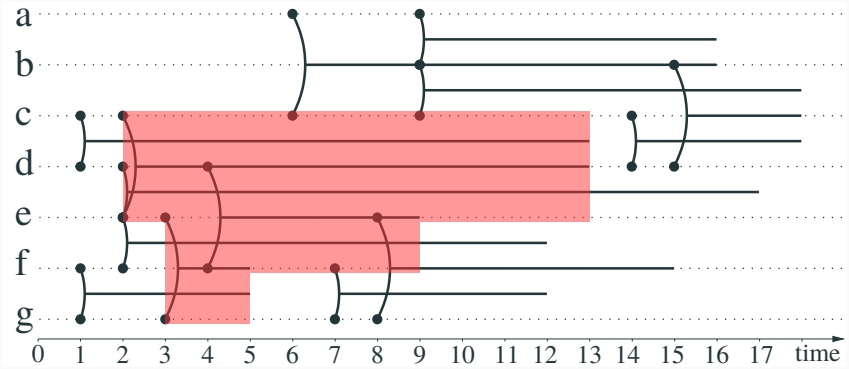
- Computing communities at each time step: time consuming;
- Some temporal data expected to be grouped are not.

CPM in link stream (LSCPM)

k-clique in link stream

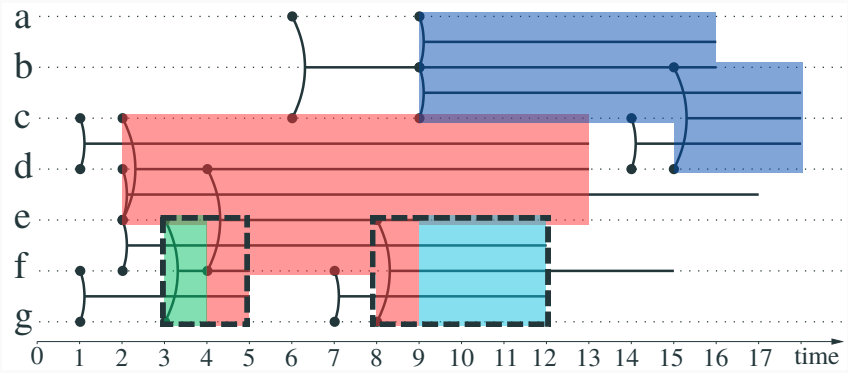
k nodes and $[t_0, t_1]$ such that all nodes are connected to each other over $[t_0, t_1]$.

$\Rightarrow$ *Grouping rule with $k = 3$:*



> LSCPM: CPM in Link Streams

CPM in temporal graphs



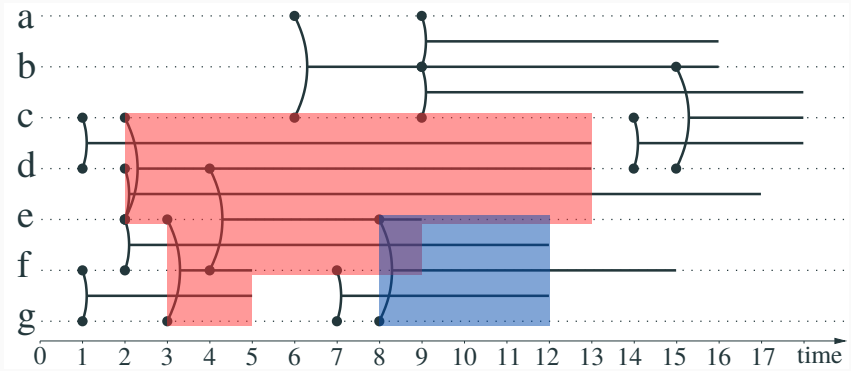
- Computing communities at each time step: time consuming;
- Some temporal data expected to be grouped are not.

CPM in link stream (LSCPM)

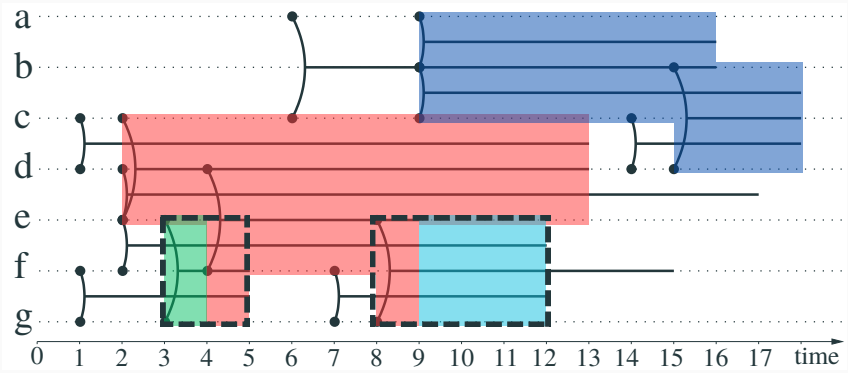
k-clique in link stream

k nodes and $[t_0, t_1]$ such that all nodes are connected to each other over $[t_0, t_1]$.

$\Rightarrow$ *Grouping rule with $k = 3$:*



CPM in temporal graphs



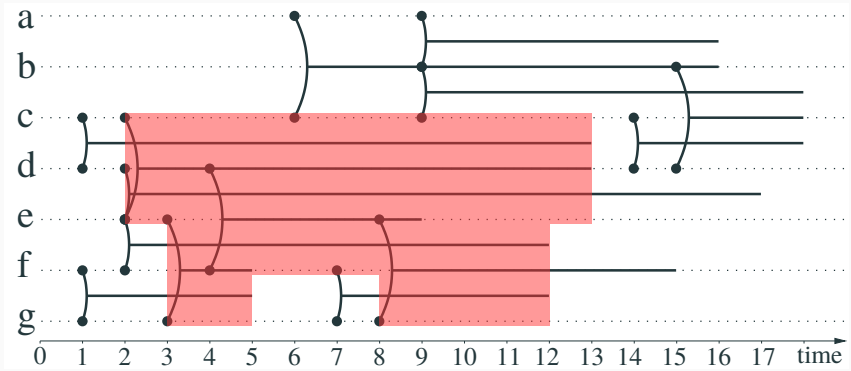
- Computing communities at each time step: time consuming;
- Some temporal data expected to be grouped are not.

CPM in link stream (LSCPM)

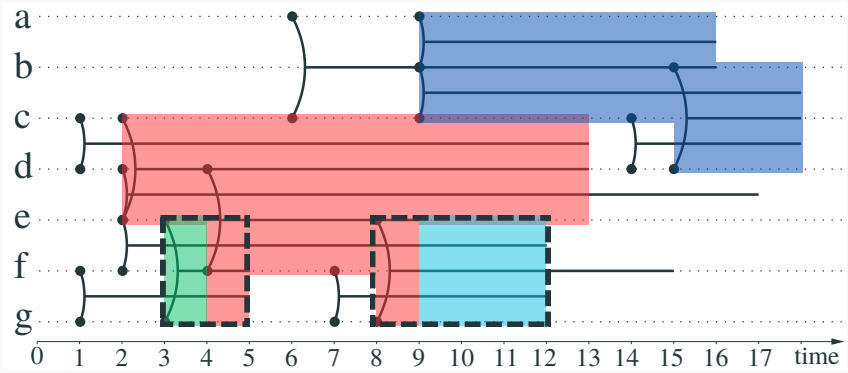
k-clique in link stream

k nodes and $[t_0, t_1]$ such that all nodes are connected to each other over $[t_0, t_1]$.

$\Rightarrow$ *Grouping rule with $k = 3$:*



CPM in temporal graphs



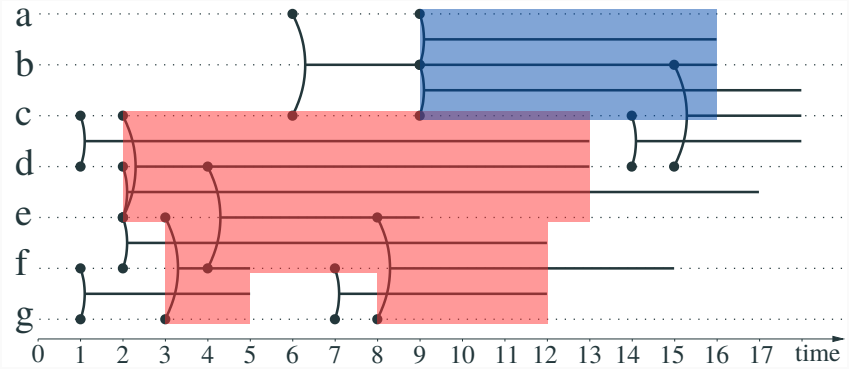
- Computing communities at each time step: time consuming;
- Some temporal data expected to be grouped are not.

CPM in link stream (LSCPM)

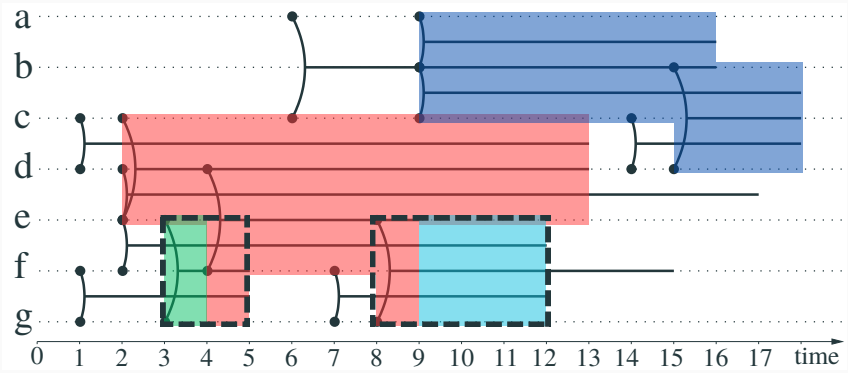
k-clique in link stream

k nodes and $[t_0, t_1]$ such that all nodes are connected to each other over $[t_0, t_1]$.

$\Rightarrow$ *Grouping rule with $k = 3$:*



CPM in temporal graphs



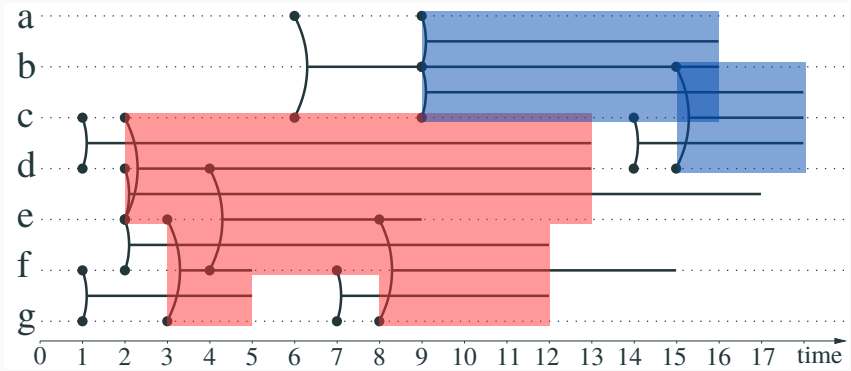
- Computing communities at each time step: time consuming;
- Some temporal data expected to be grouped are not.

CPM in link stream (LSCPM)

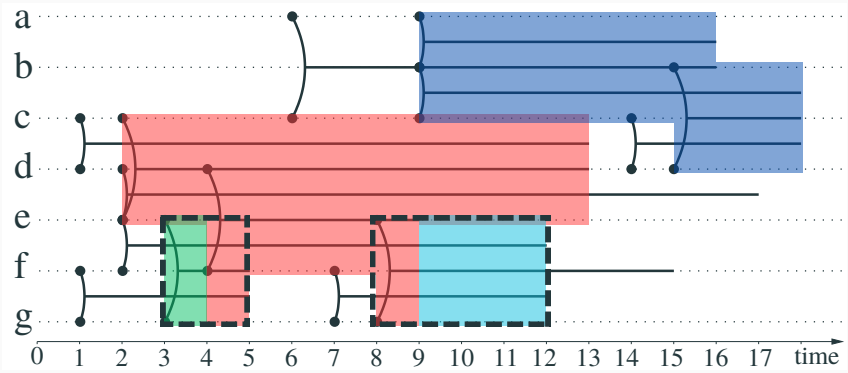
k-clique in link stream

k nodes and $[t_0, t_1]$ such that all nodes are connected to each other over $[t_0, t_1]$.

$\Rightarrow$ *Grouping rule with $k = 3$:*



CPM in temporal graphs



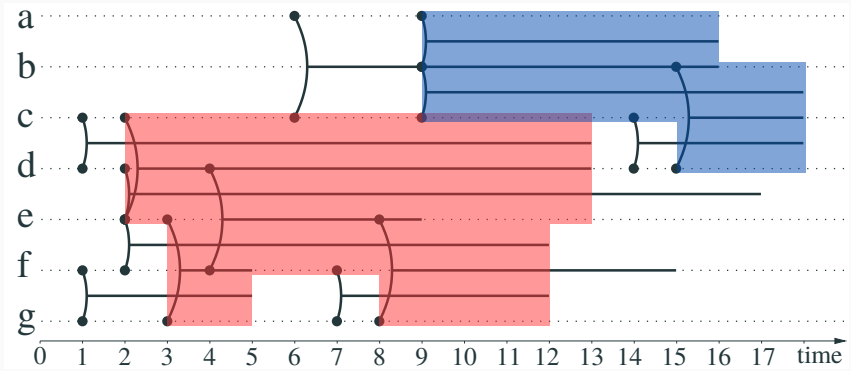
- Computing communities at each time step: time consuming;
- Some temporal data expected to be grouped are not.

CPM in link stream (LSCPM)

k-clique in link stream

k nodes and $[t_0, t_1]$ such that all nodes are connected to each other over $[t_0, t_1]$.

$\Rightarrow$ *Grouping rule with $k = 3$:*



3 - Experiments on real datasets

Efficiency – computation time

		$k = 3$		$k = 4$	
Link stream	**Nb links**	**S. of art**	**LSCPM**	**S. of art**	**LSCPM**
Households	2,136	1.5s	0.1s	1.0s	0.1s
Highschool	5,528	3.6s	0.1s	1.9s	0.1s
Infectious	44,658	10min49s	1.4s	6min12s	3.3s
Foursquare	268,472	3h01min	9.2s	2h28min	43s
Wikipedia	39,953,380	-	13min44s	-	15min29s

Efficiency – computation time

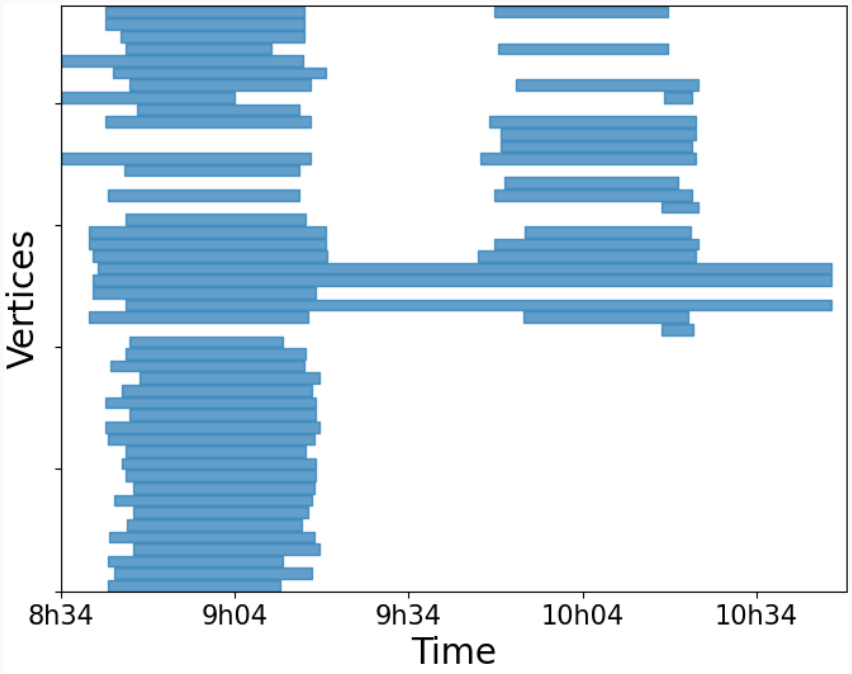
		$k = 3$		$k = 4$	
Link stream	**Nb links**	**S. of art**	**LSCPM**	**S. of art**	**LSCPM**
Households	2,136	1.5s	0.1s	1.0s	0.1s
Highschool	5,528	3.6s	0.1s	1.9s	0.1s
Infectious	44,658	10min49s	1.4s	6min12s	3.3s
Foursquare	268,472	3h01min	9.2s	2h28min	43s
Wikipedia	39,953,380	-	13min44s	-	15min29s

Consistency with metadata

Highschool: 70% of communities are on one class, 23% on two classes, 6% on three classes, 1% on four classes.

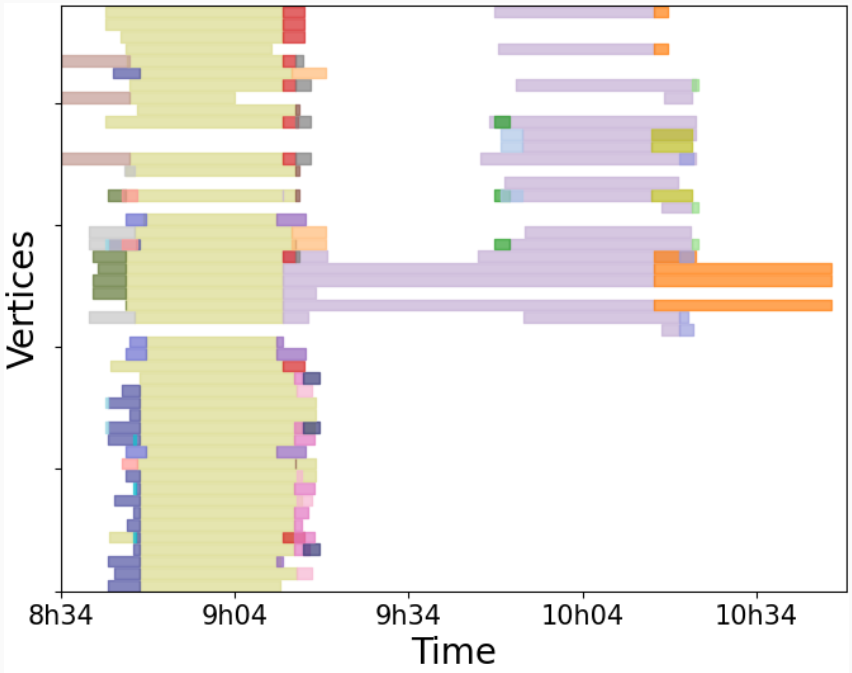
A highschool link stream community

LSCPM community



A highschool link stream community

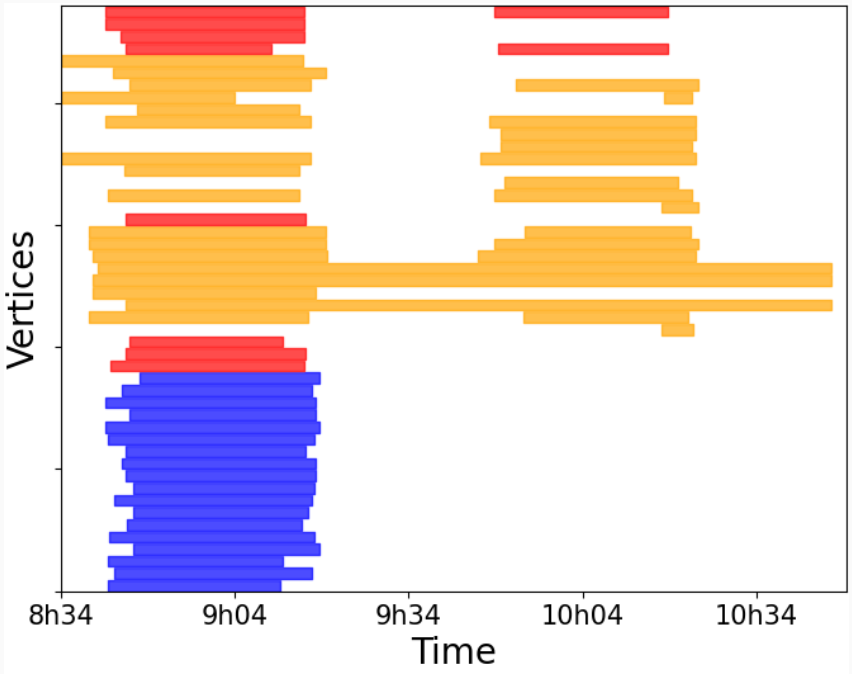
State-of-the-art communities



→ Gather more information over time

A highschool link stream community

Metadata (classes)



→ Gather more information over time

→ Relate metadata information

4 - Conclusion: links with BCI ?

4 - Conclusion: links with BCI ?

→ **Communities in link streams**

- (Re)organization of interactions over time
- Target communities that provide temporal information
- Target vertices that play a central role (or not)

4 - Conclusion: links with BCI ?

→ **Communities in link streams**

- (Re)organization of interactions over time
- Target communities that provide temporal information
- Target vertices that play a central role (or not)

→ **Link stream**

- Study at different time scales
- Online interactions
- Multilayer link stream ?...

LSCPM: finding communities in link streams

Thanks for your attention! Any questions?

Code available at:

`https://gitlab.lip6.fr/baudin`

Alexis Baudin – `alexis.baudin@lip6.fr`

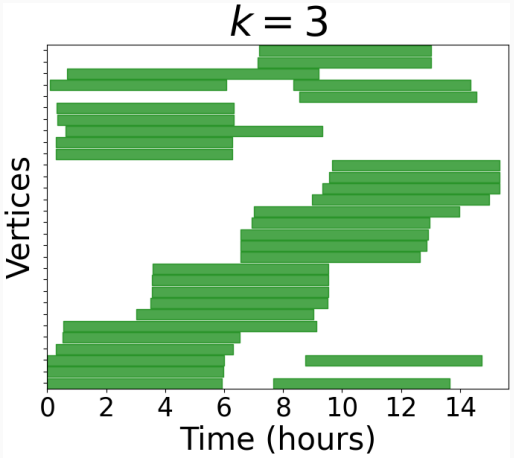
Supplementary material

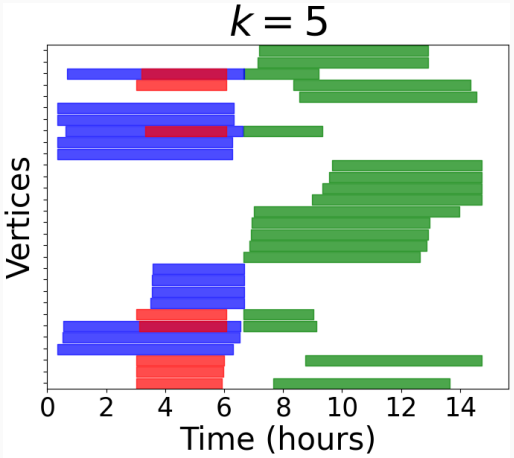
Supplementary material

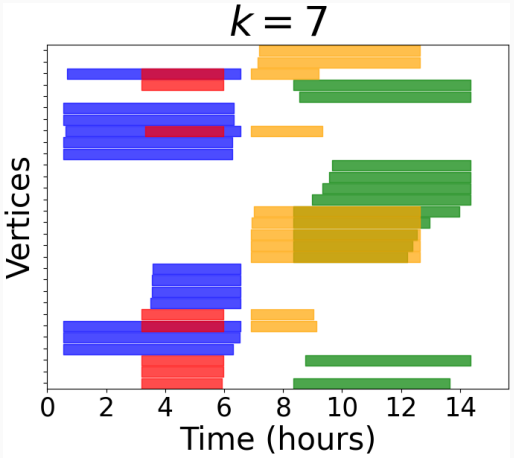
Observation

As k increases, communities split.

Example:



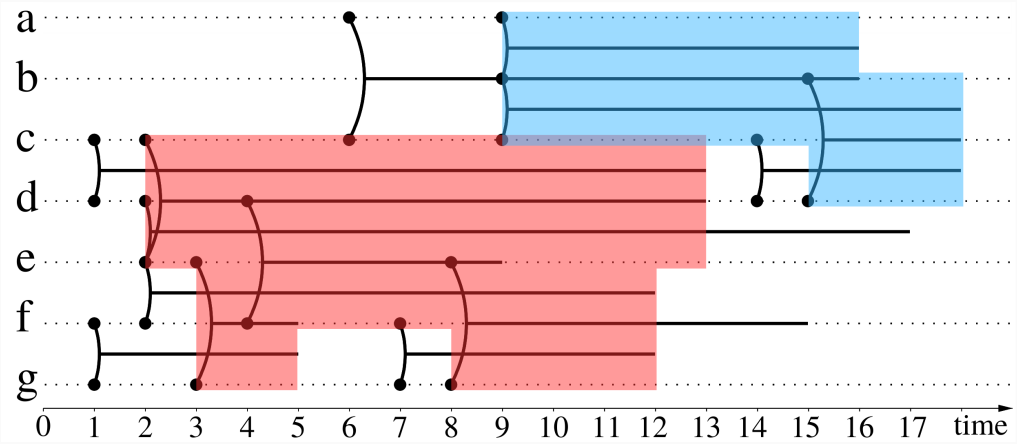




$\Rightarrow$ Refining communities; "core"

$\Rightarrow$ Sub-categorizing data

Supplementary material



Question

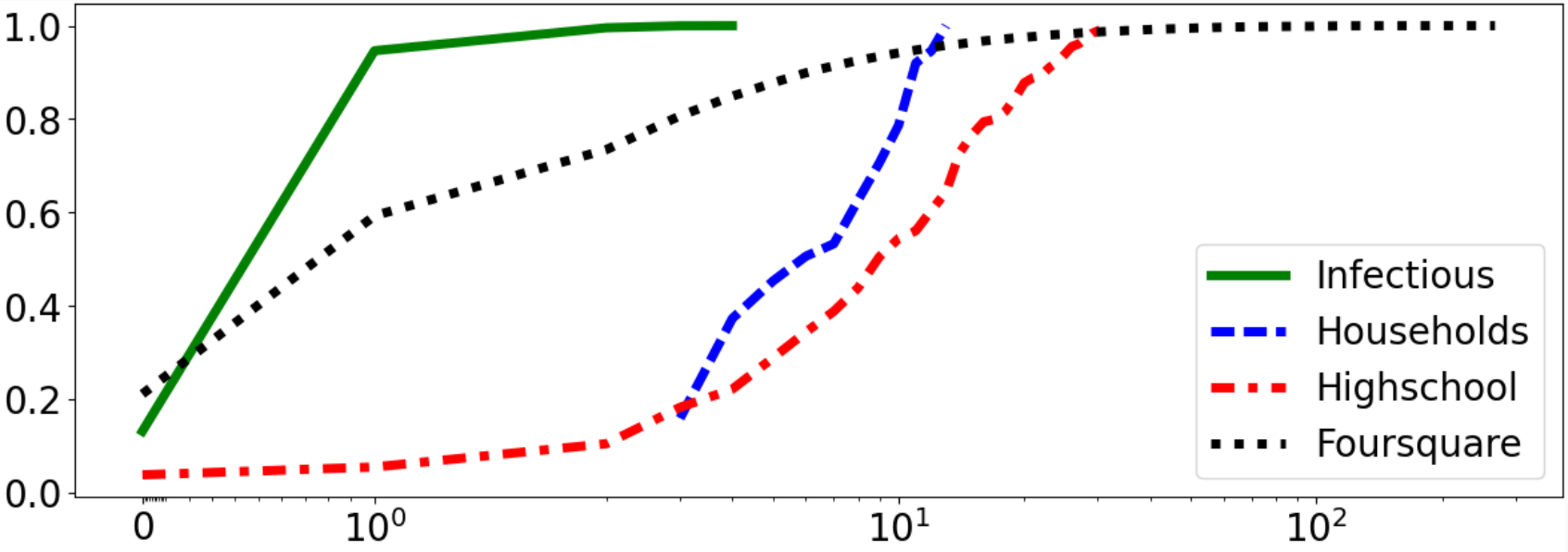
What is a community in a link streams ?

→ Sets of temporal nodes that are

- Densely connected inside
- Sparsly connected outside

Supplementary material

Importance of vertices



Number of communities each vertex belongs to.

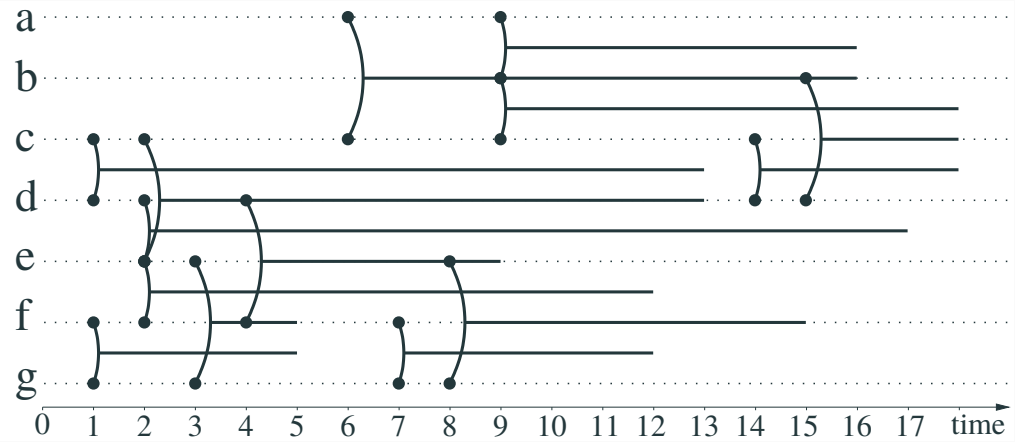
Foursquare dataset

→ Nodes very central (Pennsylvania train station, ...)

→ Nodes not in any community (offices, ...)

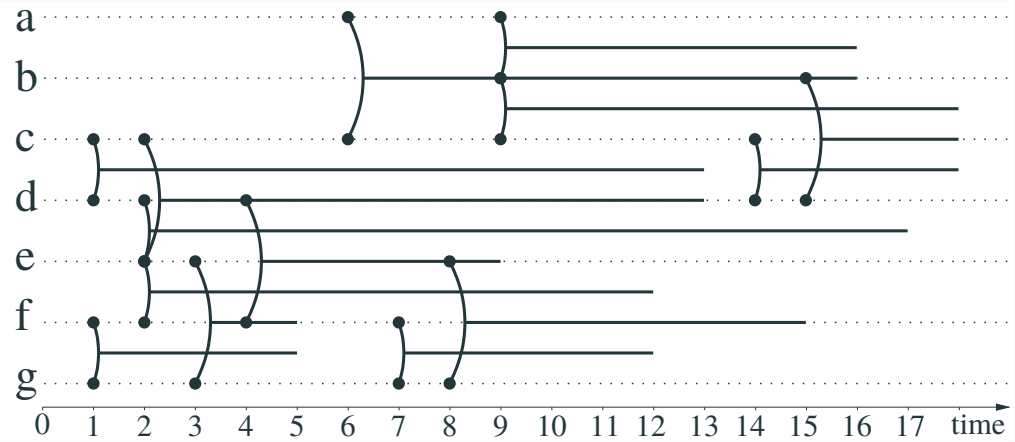
Supplementary material

Link stream

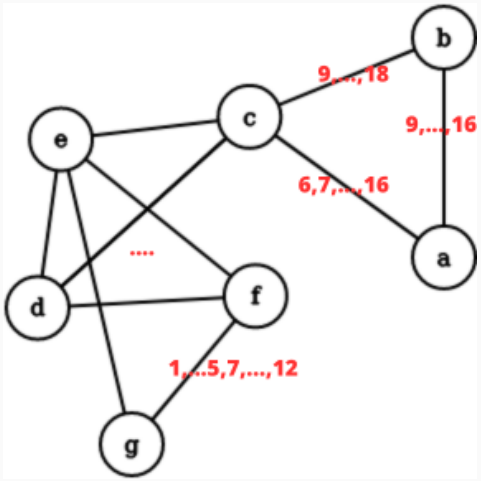


Supplementary material

Link stream

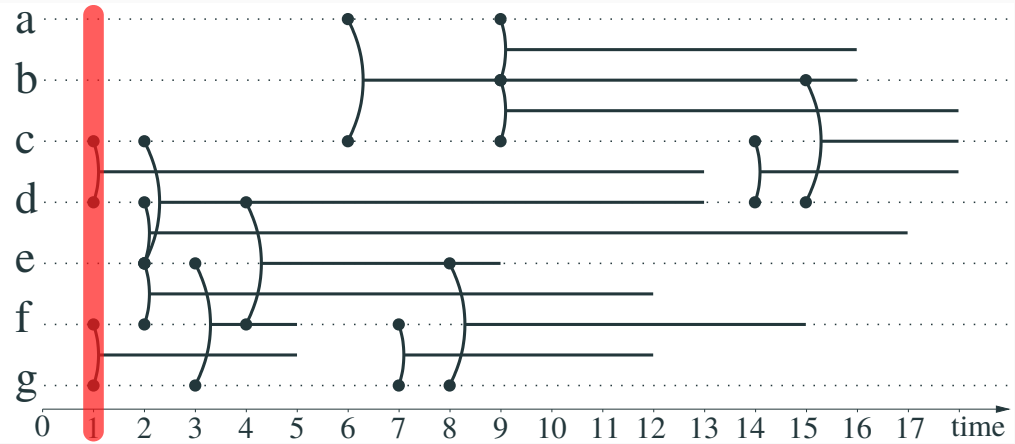


Time varying graphs

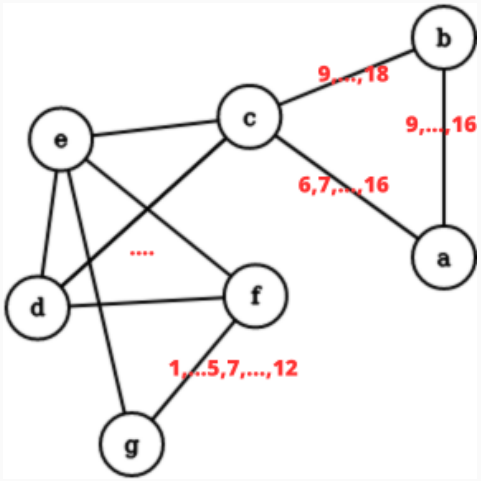


Supplementary material

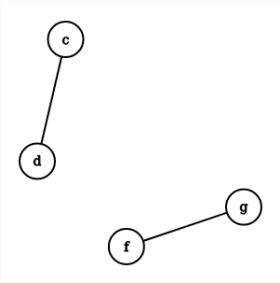
Link stream



Time varying graphs



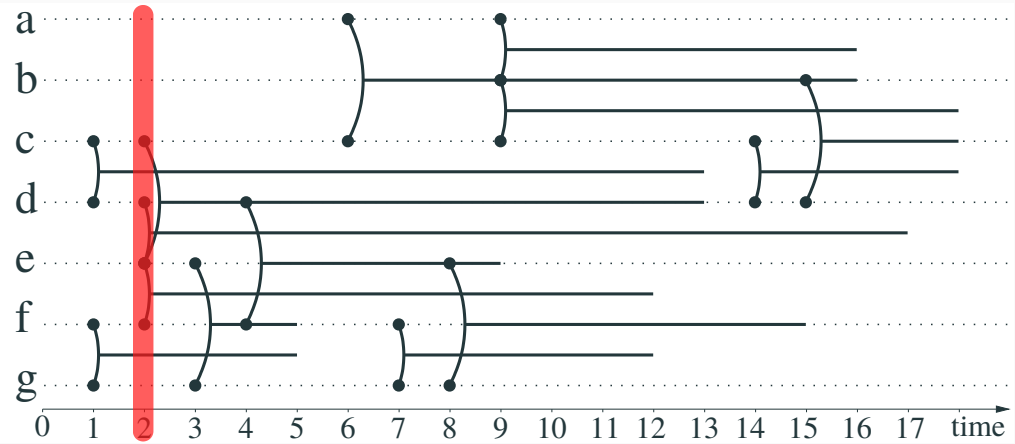
Temporal graphs



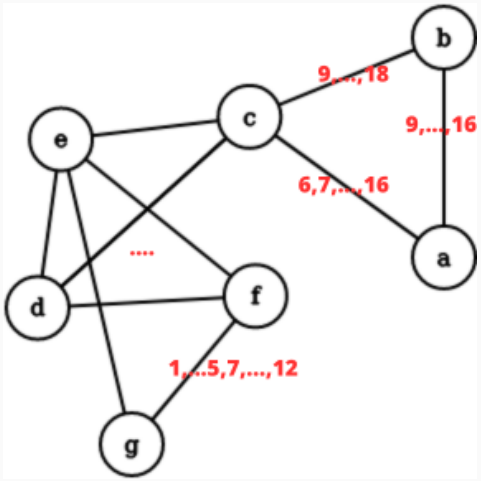
$t = 1$

Supplementary material

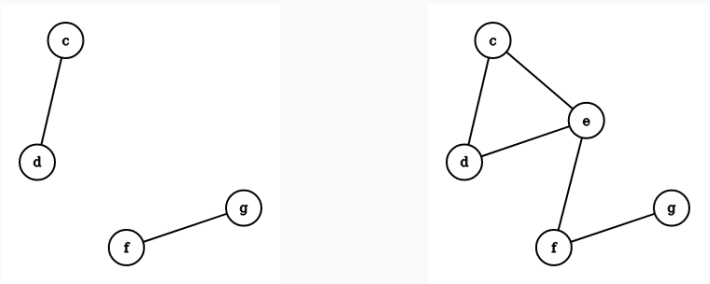
Link stream



Time varying graphs



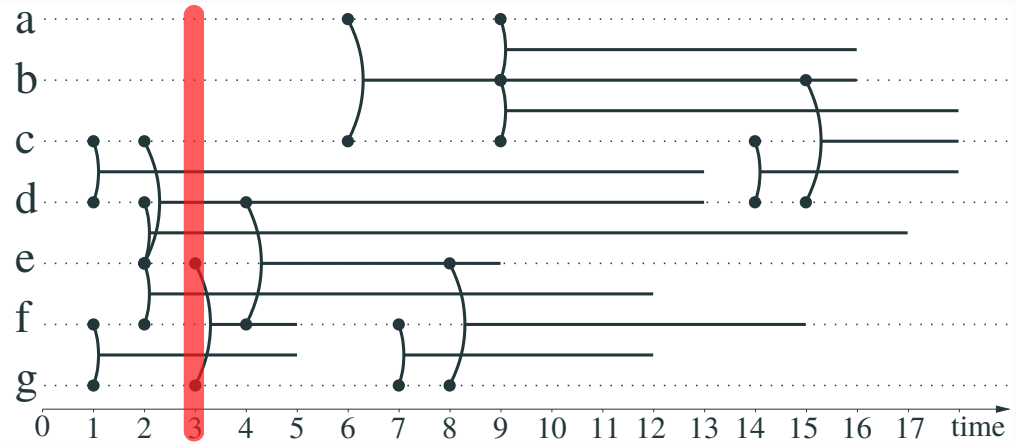
Temporal graphs



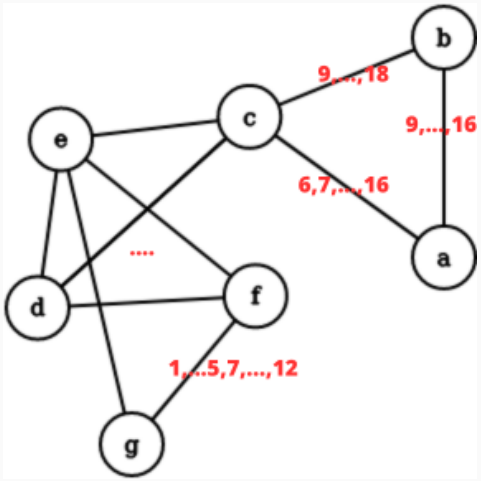
$t = 1 \longrightarrow t = 2$

Supplementary material

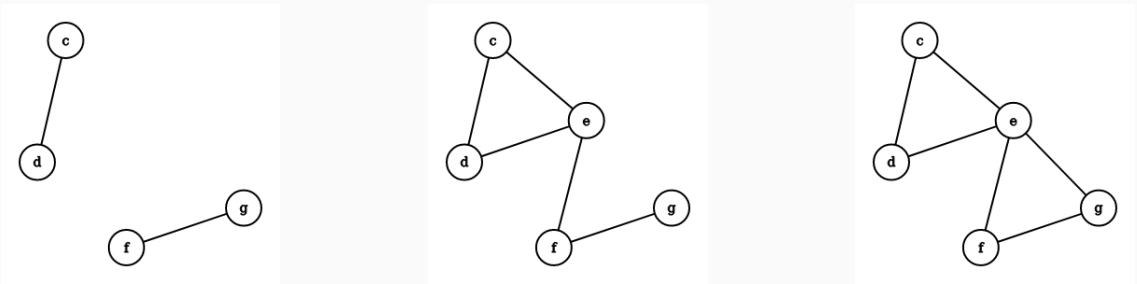
Link stream



Time varying graphs



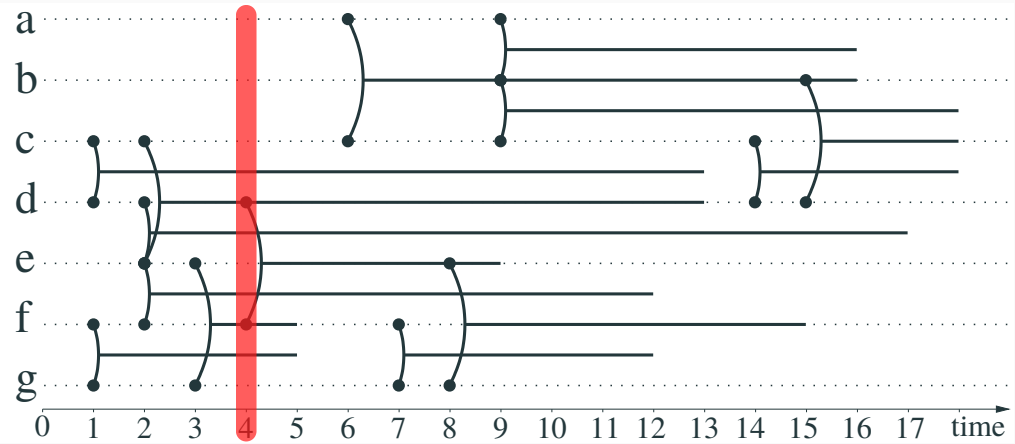
Temporal graphs



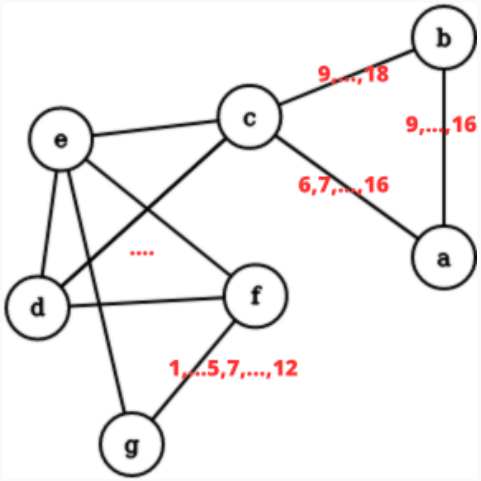
$t = 1 \longrightarrow t = 2 \longrightarrow t = 3$

Supplementary material

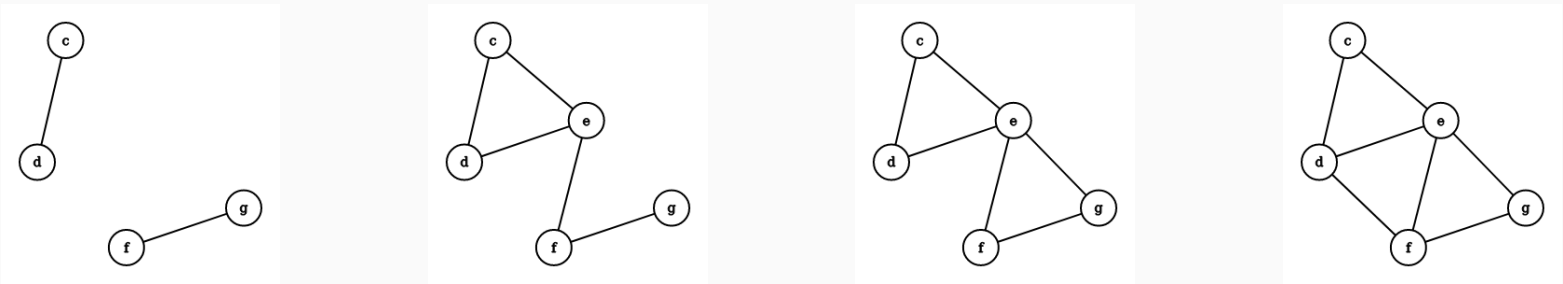
Link stream



Time varying graphs



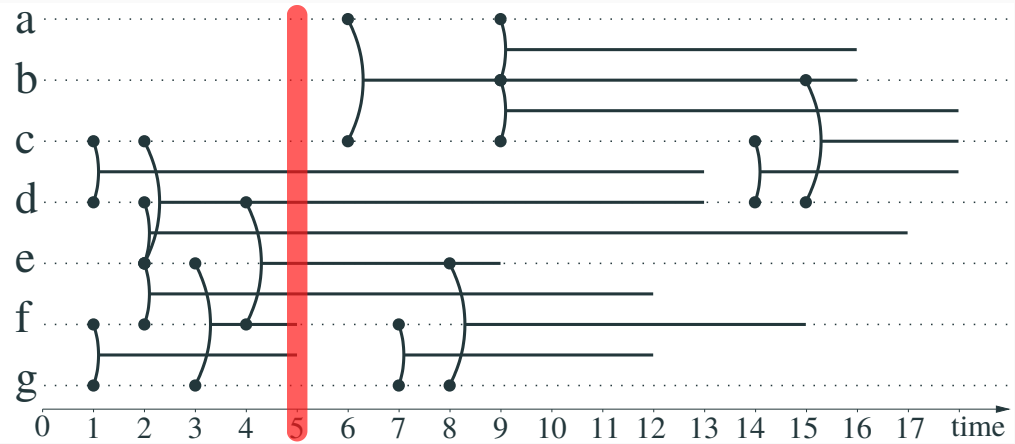
Temporal graphs



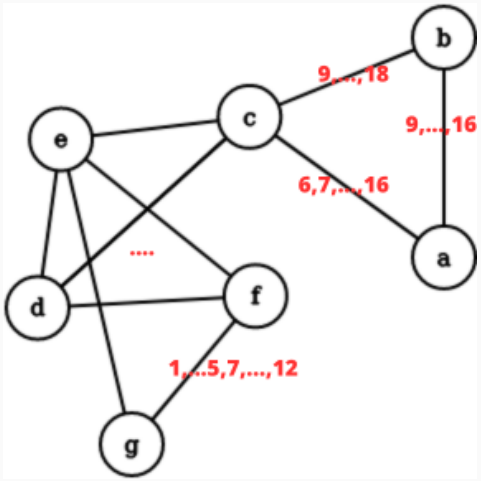
$t = 1 \longrightarrow t = 2 \longrightarrow t = 3 \longrightarrow t = 4$

Supplementary material

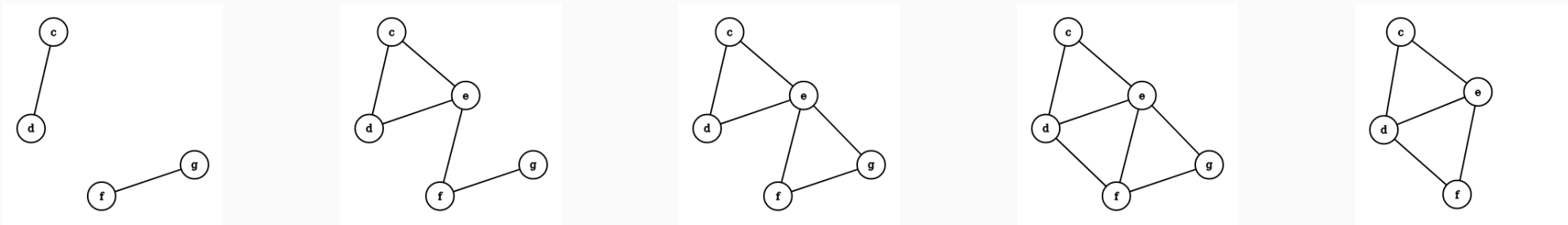
Link stream



Time varying graphs



Temporal graphs



$t = 1 \longrightarrow t = 2 \longrightarrow t = 3 \longrightarrow t = 4 \longrightarrow t = 5$