Internship
Temporal Matching in Link Streams

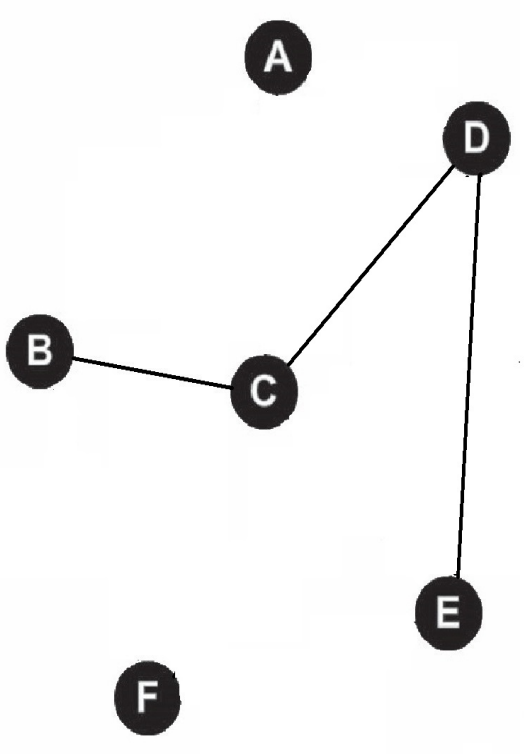
Katia AMICHI

Supervisor: Binh-Minh Bui-Xuan

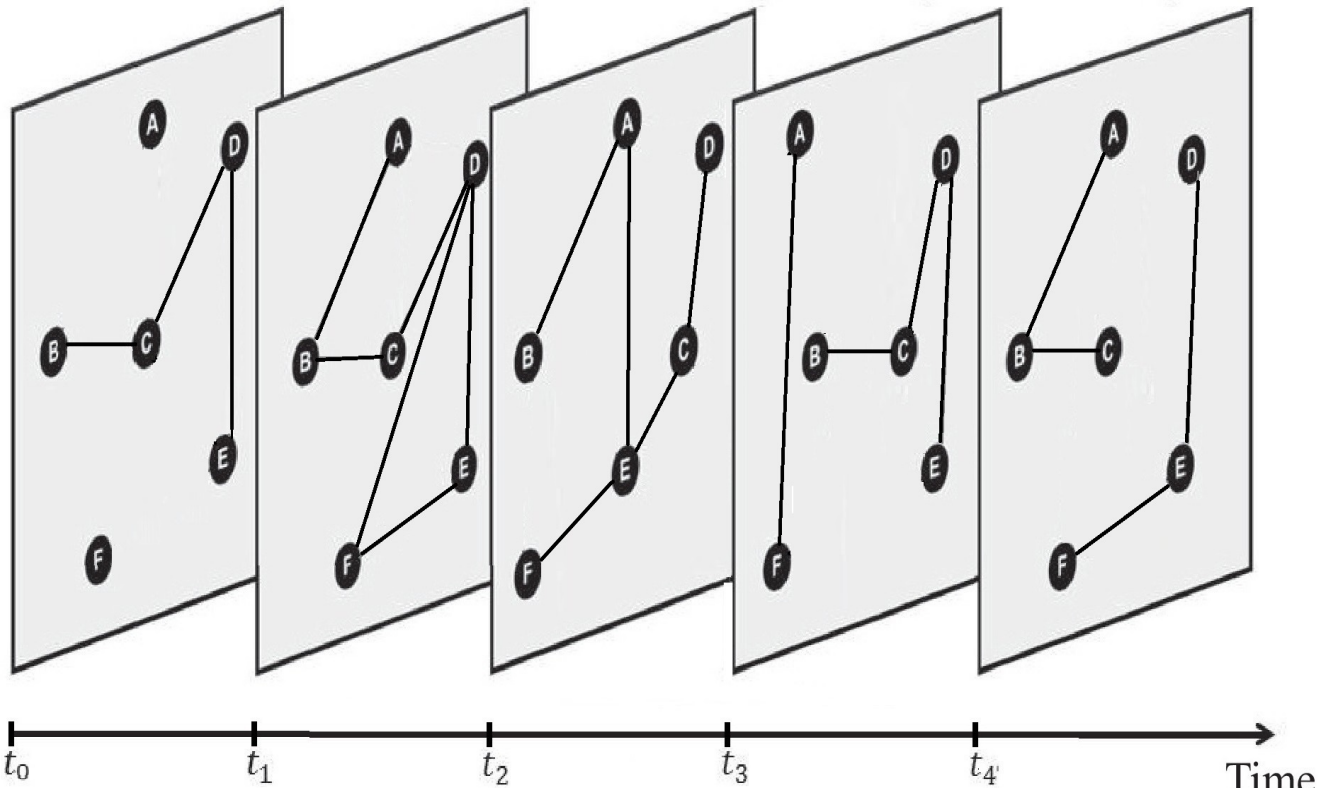
October 28, 2019

Overview

Temporal graph

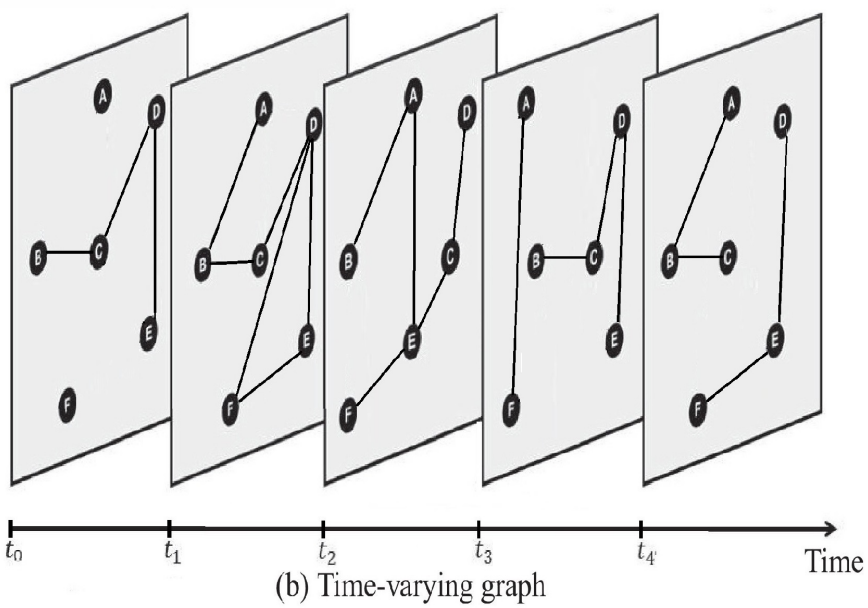


(a) Graph

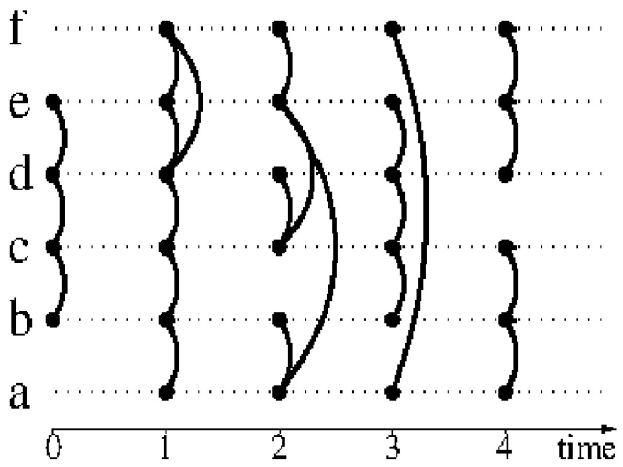


(b) Time-varying graph

Link stream



(b) Time-varying graph



Link stream

Link stream

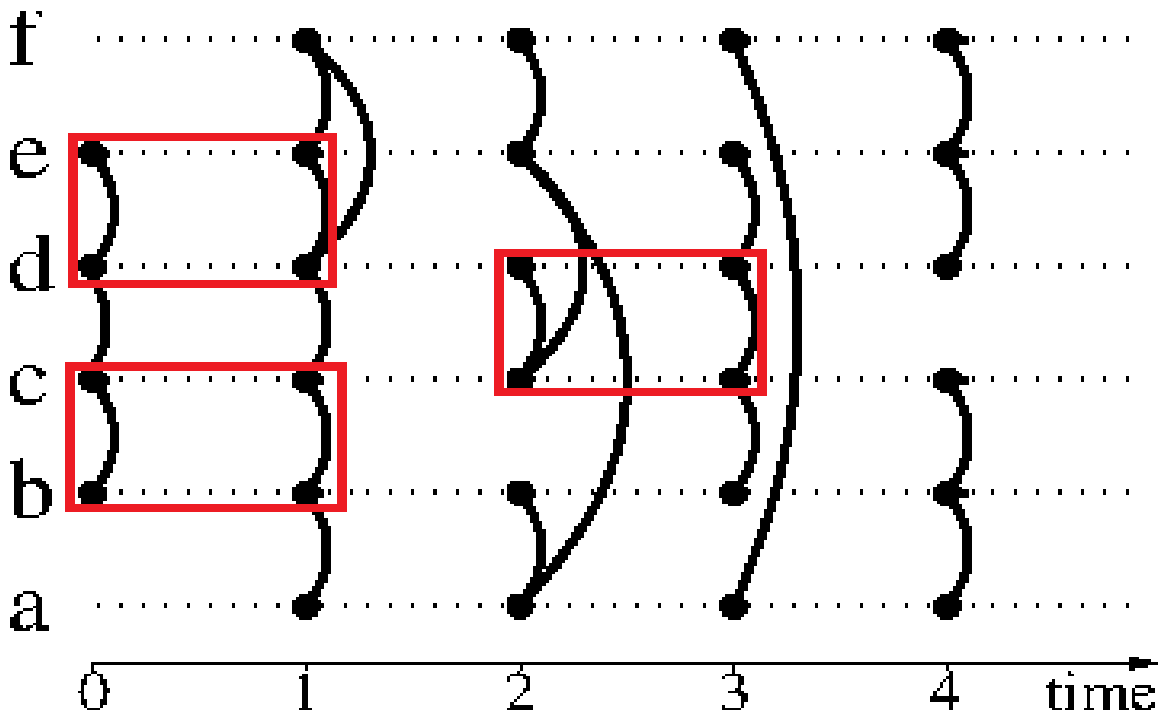
Definition: Let $L = (V,T,E)$:

- V set of vertices in the graph
- T $\subseteq$ N an integer represents a discretized time instant
- E $\subseteq \{(t, uv) : t \in T \wedge u \in V \wedge v \in V \wedge u \neq v\}$

Note :

- A graph G is link stream if $\mid T \mid = 1$
- Link stream is graph if every pair $(t, u, v) \in L$ satisfies t $= t_0$ for same fixed t_0.

Link stream, γ-edges and γ-matching



Temporal matching

Let $L = (V,T,E)$ a link stream and γ an integer
A γ-edges is set of repeated edges between vertices u and v starting at t and ending at t + γ - 1: $\{(t', \{u,v\}) \mid t' \in [t, t+\gamma-1] \}$.

Two γ-edges $\Gamma_1 = (t_1, u, v)$ and $\Gamma_2 = (t_2, u', v')$ are independent if $t_1 > t_2 + \gamma - 1$. $(t_1 < t_2 + \gamma - 1)$ where $u \neq u'$ and $v \neq v'$.

Temporal matching

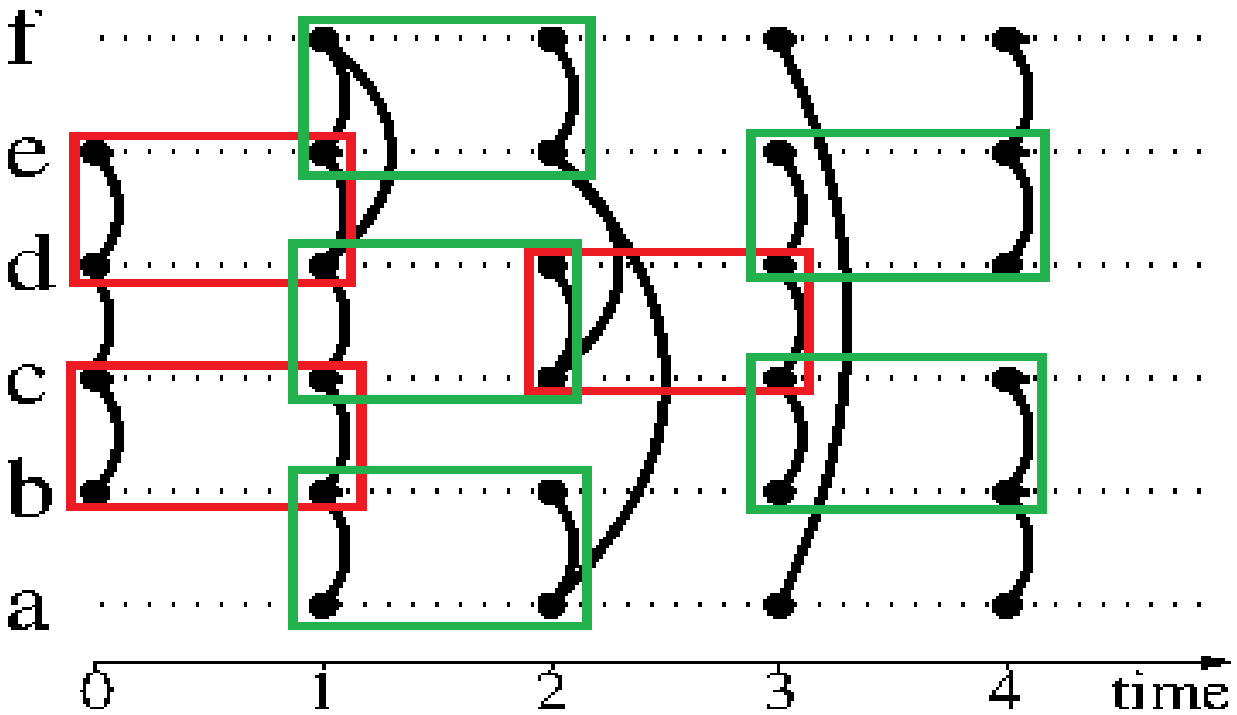
Let $L = (V,T,E)$ a link stream and γ an integer
A γ-edges is set of repeated edges between vertices u and v starting at t and ending at $t + \gamma - 1$: $\{(t', \{u,v\}) \parallel t' \in t, t + \gamma - 1 \}$.

Two γ-edges $\Gamma_1 = (t_1, u, v)$ and $\Gamma_2 = (t_2, u', v')$ are independent if $t_1 > t_2 + \gamma - 1$. $(t_1 < t_2 + \gamma - 1)$ where $u \neq u'$ and $v \neq v'$.

Temporal matching

Definition: A γ-matching is a set of independent γ-edges.

Our question: find maximum γ-matching ?



Our question: find maximum γ-matching ?

Problem γ-matching:

INPUT: link stream L, integer γ
OUTPUT: γ-matching in L of size k

Theorem [Baste et al 2019]

γ-matching is NP-hard for $\gamma > 1$
2 approximation algorithm (BBR19 greedy algorithm)

Our result

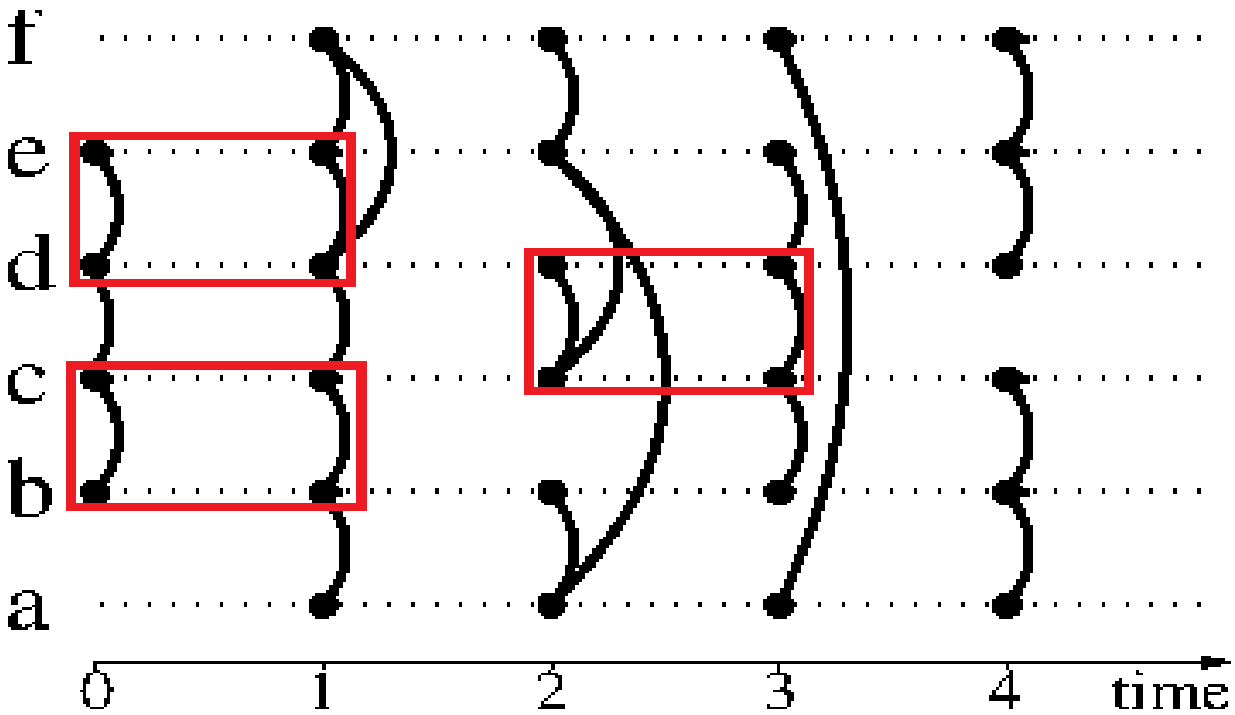
Approximation algorithms:

- Local search heuristic (LS)
- Divide and conquer heuristic (DC)

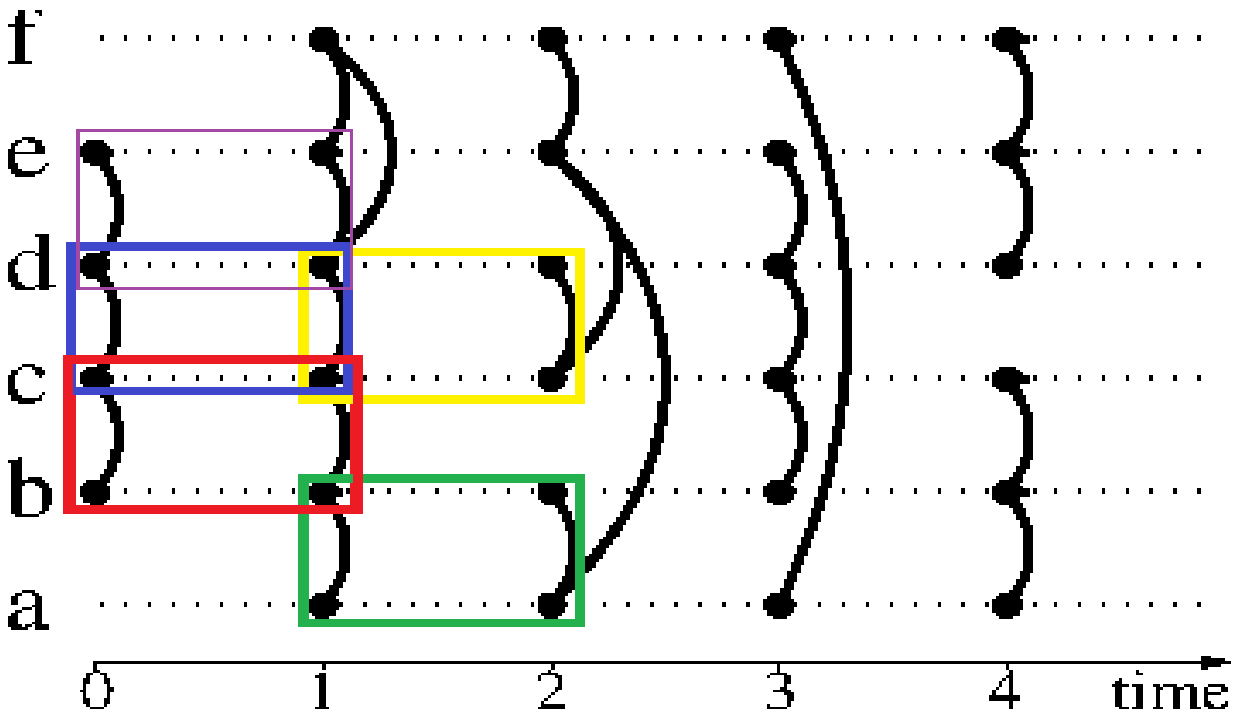
Numerical analysis: approximation

- Generated link stream (3000 tests): moving particles under natural simulation in 2D and 1D euclidean spaces.
- Real world link stream (1000 tests): scenario of Enron dataset.
- Real world linkbstream (1000 tests): scenario of Rollernet (Rollerblading tourning Paris) dataset.

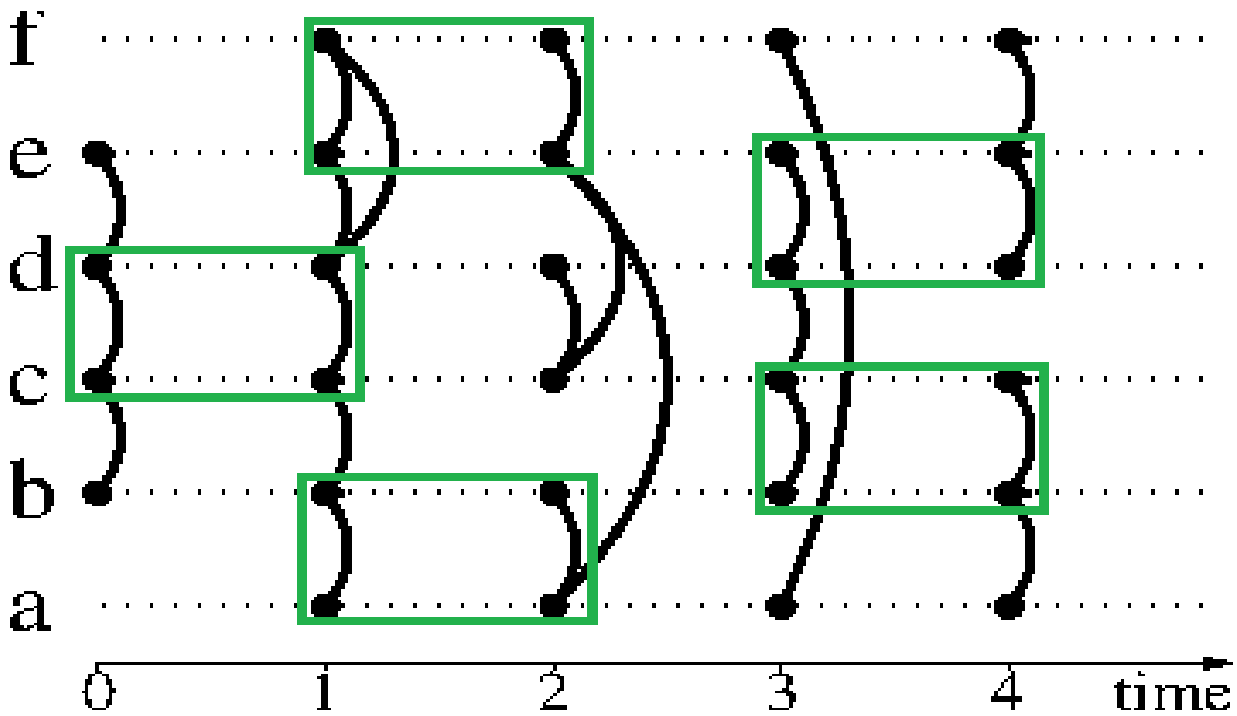
Greedy algorithm (BBR19)



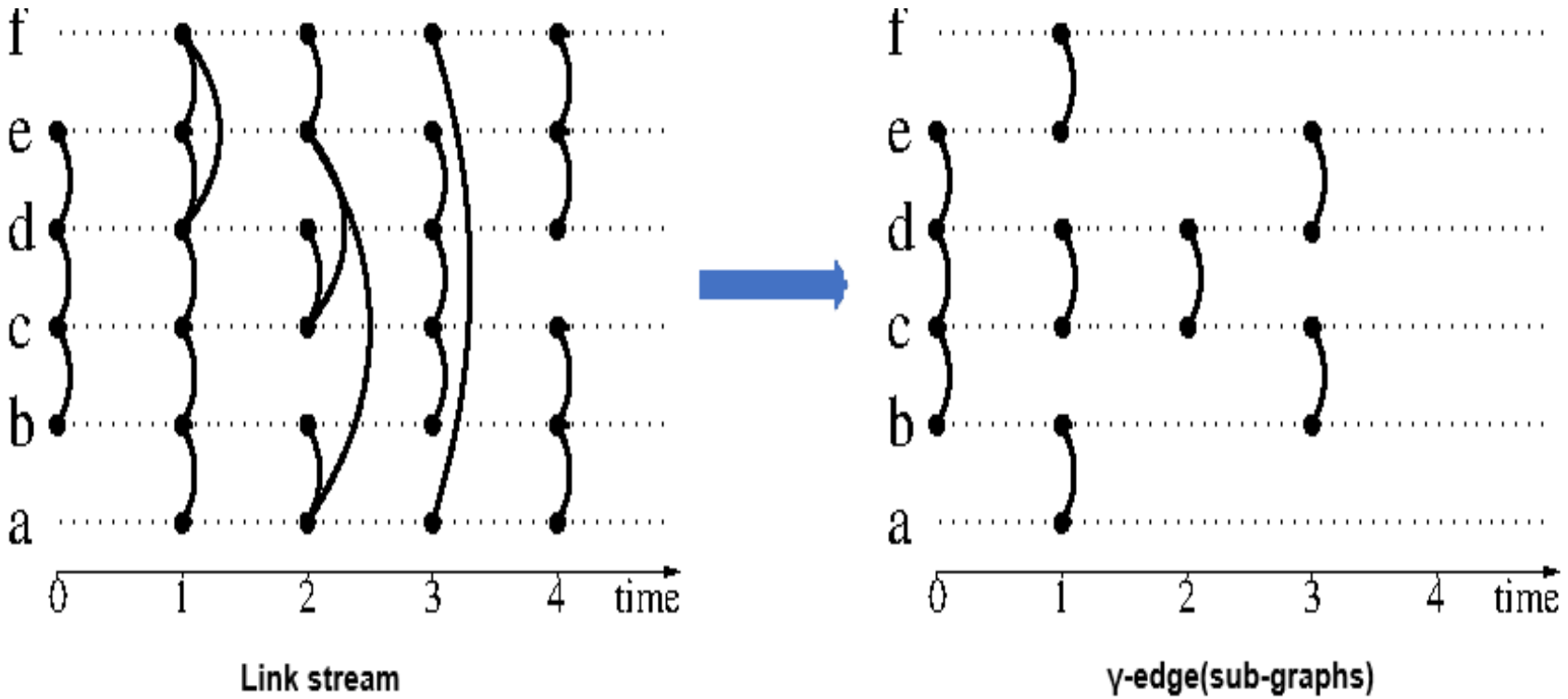
Local search heuristic (LS)



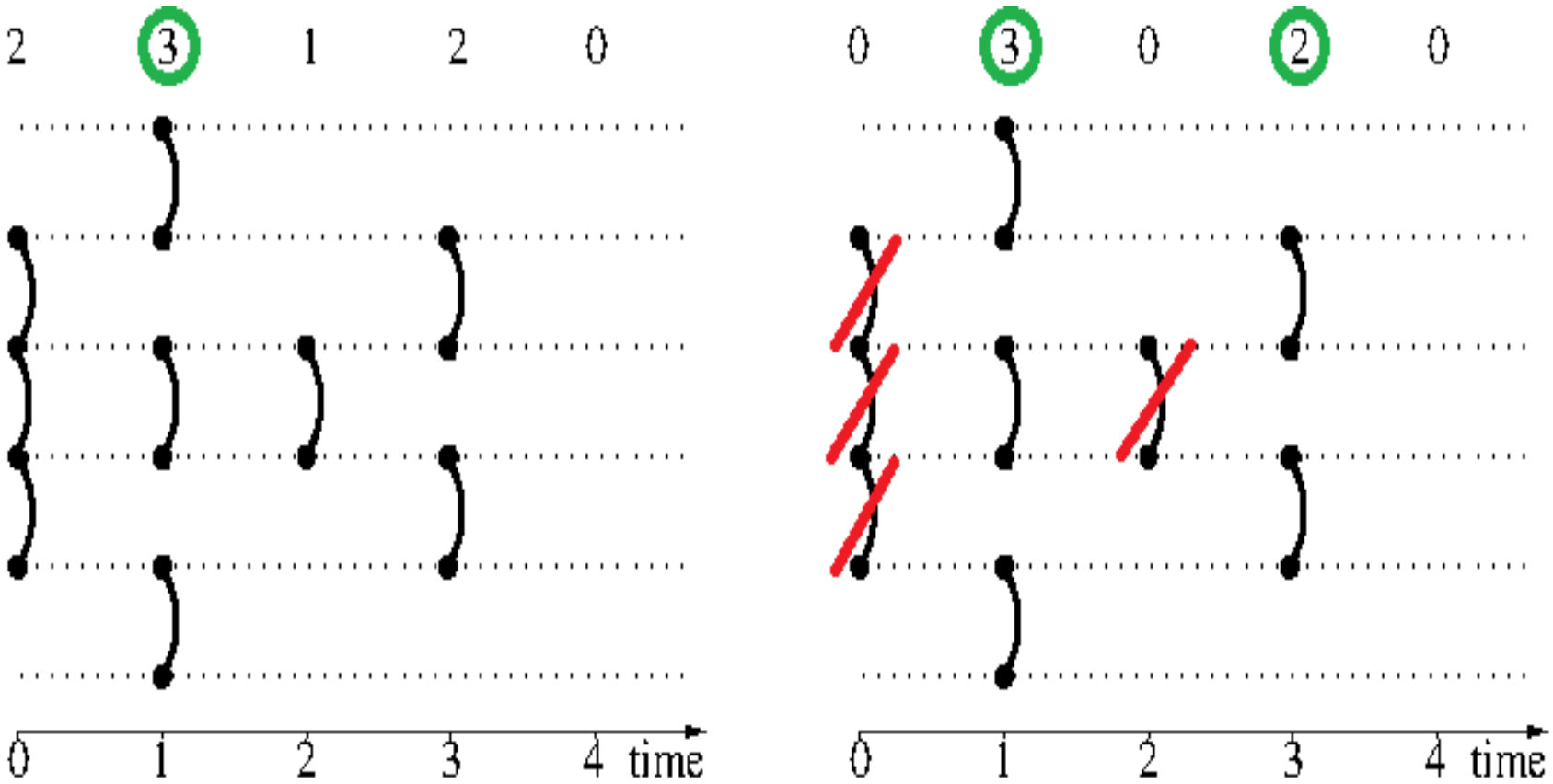
Local search heuristic (LS)



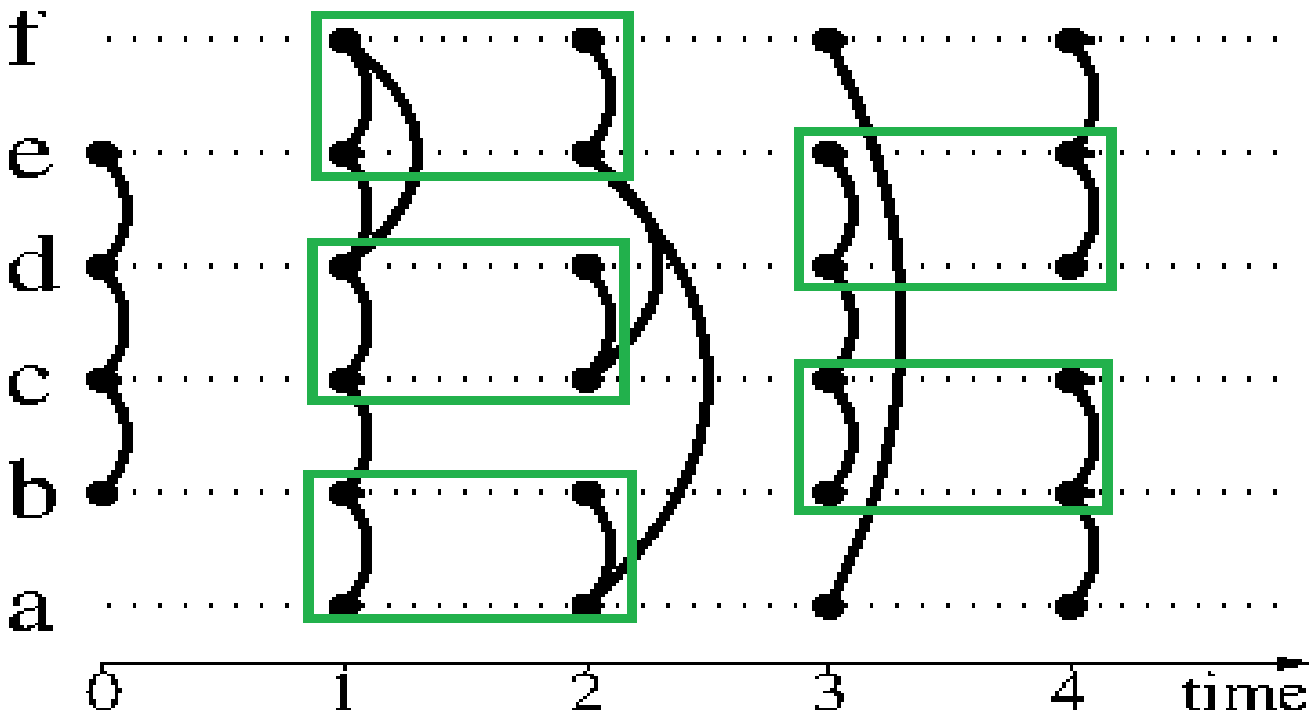
Divide and conquer heuristic (DC)



Divide and conquer heuristic (DC)



Divide and conquer heuristic (DC)



Discussion

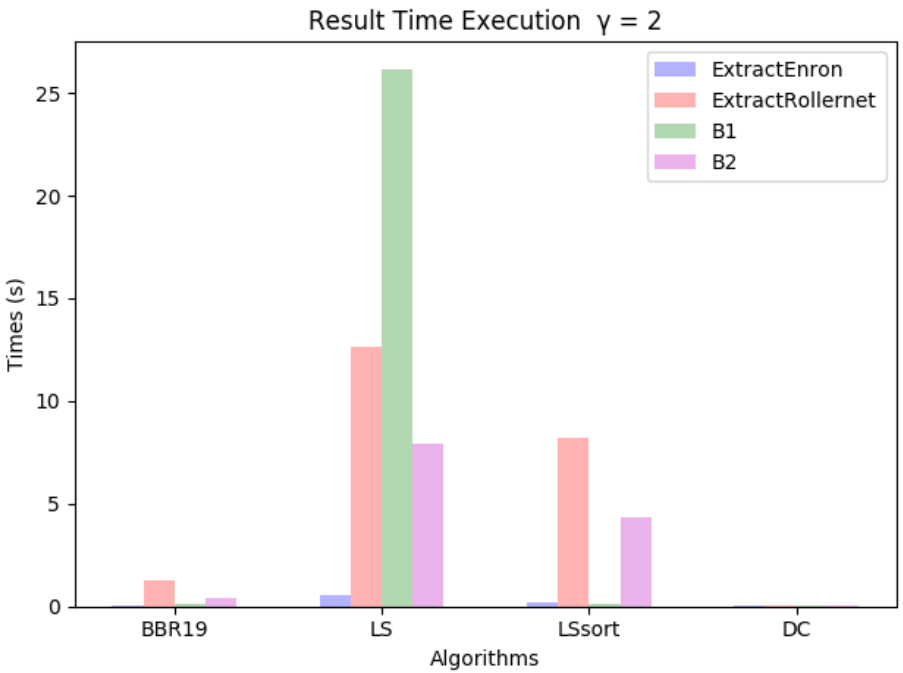
Algorithms

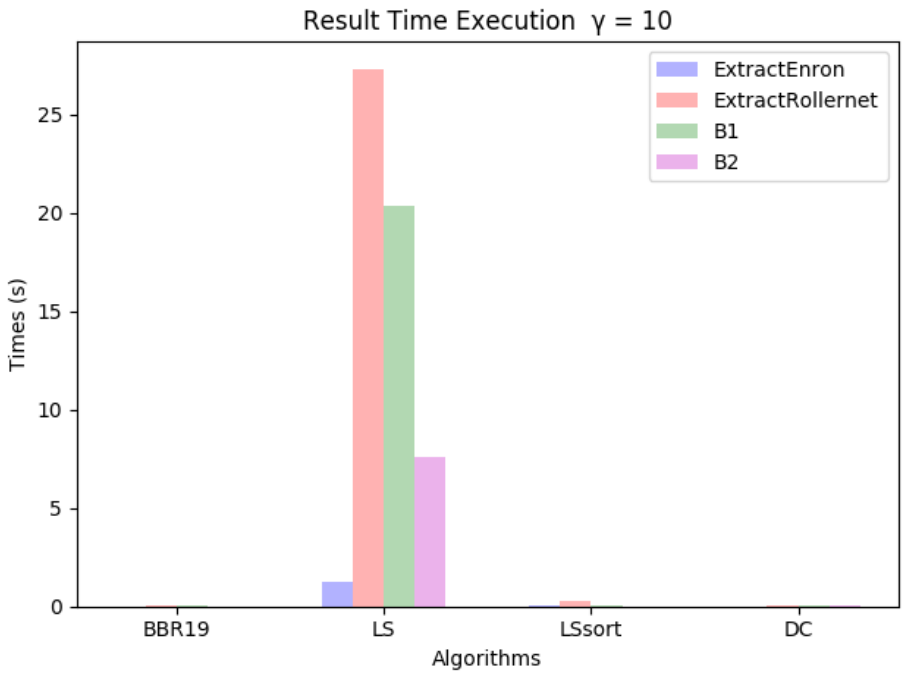
- BBR19 (J. Baste et al) in time $O(n\tau + m)$, where $\tau = |T|$, n $= |V|$ and m $= |E|$.
- LS (Local search heuristic) is very long.
- DC (Divide and conquer heuristic) on $O(m\ log\ m + (\text{m-}\tau')^2)$, where $1 < \tau' < \tau$

Note

The algorithms BBR19, LS and DC are greedy algorithms therefore 2-approximation algorithms)

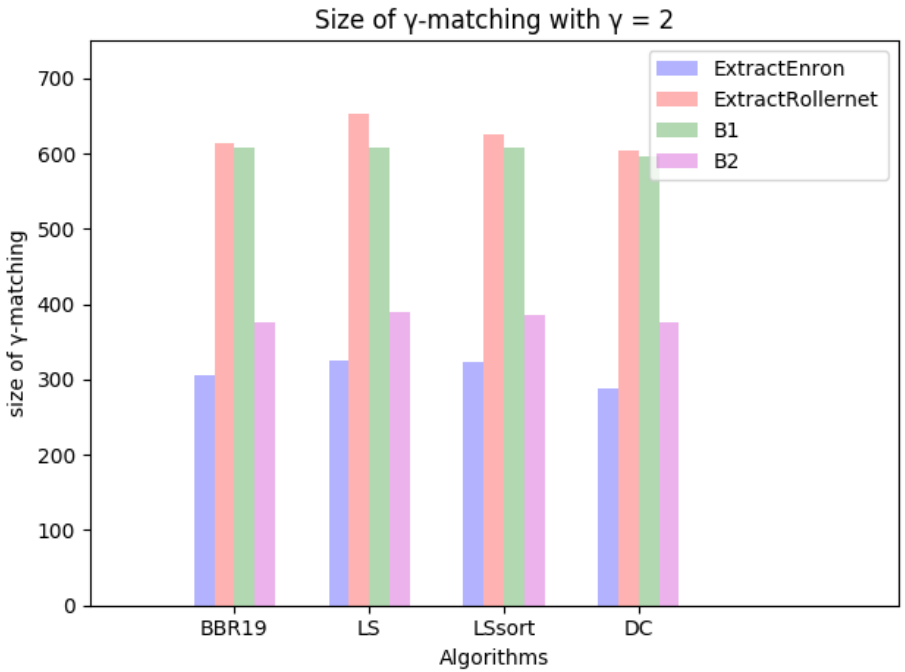
Discussion

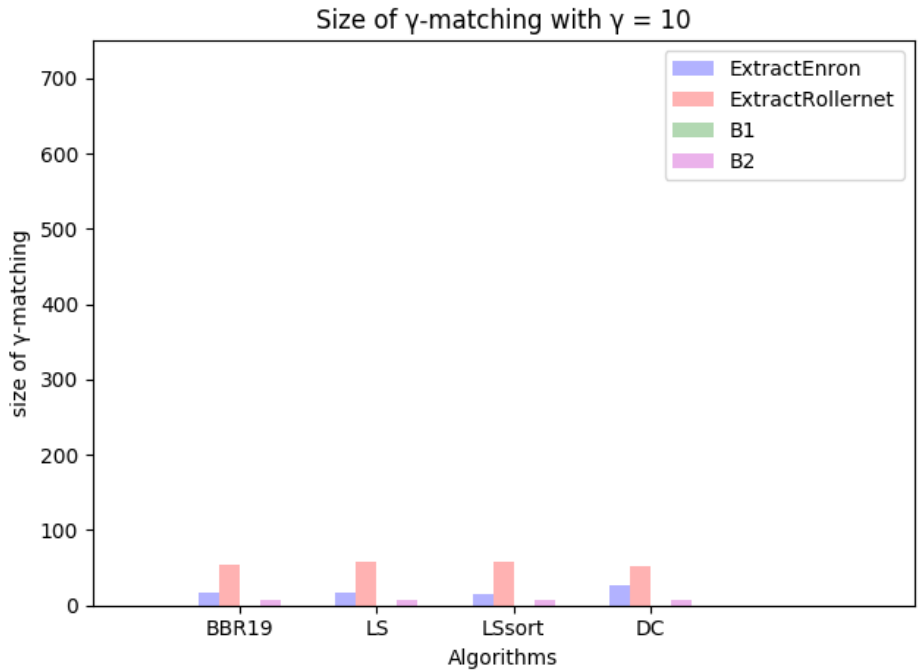




- BBR19 depends mainly on $|E|$ and $|\gamma\text{-edges}|$.
- LS depends on $|E|$ and $|\gamma\text{-edges}|$ but mainly on the density of the sub-graphs for each instance.
- DC depends on $|T|$ and the position of the maximum matching for each iteration.

Discussion





- The order of the edges in the link stream has an impact on the size of γ-matching
- The more one increases the value of γ the smaller the size of γ-matching
- For γ=10 the size of γ-matching=0.

Conclusion

Our result:

- Implementation of two heuristic (LS and DC).
- Generation of 3000 tests.
- # Lines of code $\approx$ 2400 (Python, bash).

Analysis:

- The order of the edges in the link stream has an impact on the results.
- The execution time does not always depend on the size of γ-edges.

`https://github.com/KatiaAMICHI/Temporal-Matching-in-Link-Streams`

Thank you,Questions?