An efficient procedure for mining egocentric temporal motifs

Antonio Longa, Giulia Cencetti, Bruno Lepri and Andrea Passerini

https://antoniolonga.github.io/

TABLE OF CONTENTS

01

NETWORK MOTIFS

NETWORK

Network:
A network **G** is a pair of sets **G**=(**N**,**E**). Where **N** is a set of nodes and **E** is a set of edges (couple of nodes).

NETWORK

Network:

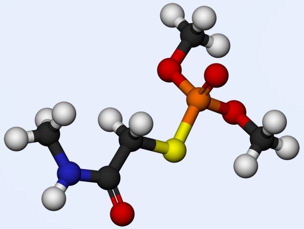
A network **G** is a pair of sets **G**=(**N**,**E**). Where **N** is a set of nodes and **E** is a set of edges (couple of nodes).

Social networks

NETWORK

Network:
A network **G** is a pair of sets **G**=(**N**,**E**). Where **N** is a set of nodes and **E** is a set of edges (couple of nodes).

Social networks

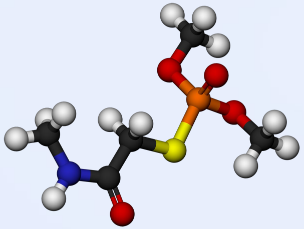


Moleculas

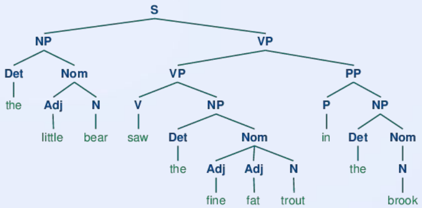
NETWORK

Network:
A network **G** is a pair of sets **G**=(**N**,**E**). Where **N** is a set of nodes and **E** is a set of edges (couple of nodes).

Social networks



Moleculas



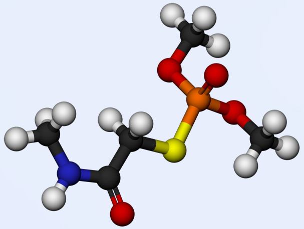
Sentence

NETWORK

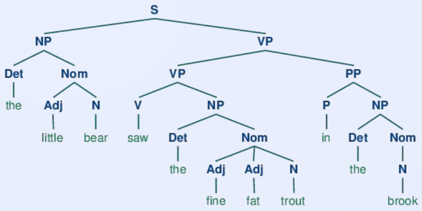
Network:

A network **G** is a pair of sets **G**=(**N**,**E**). Where **N** is a set of nodes and **E** is a set of edges (couple of nodes).

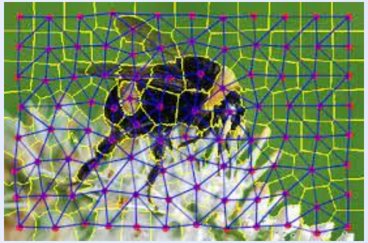
Social networks



Moleculas



Sentence



Images

NETWORK

How can we study networks?

01

Topology

Many properties:
Degree, clustering,
assortativity, connectivity,
...

02

Dynamic simulations

Simulate an epidemic in a network. (R0, infected individuals, ...)

NETWORK MOTIFS [1]

are **subgraphs**, that **appear** in an observed network **significantly more often** than in compatible randomized networks.

[1] Milo, Ron, et al. "Network motifs: simple building blocks of complex networks." *Science* 298.5594 (2002): 824-827.

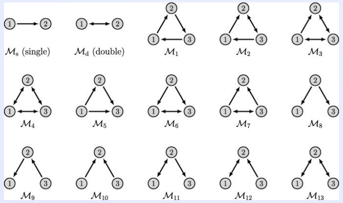
NETWORK MOTIFS [1]

are **subgraphs**, that **appear** in an observed network **significantly more often** than in compatible randomized networks.

Procedure

1) Count all possible substructure of a given network.

Input network



Subgraph counts

[1] Milo, Ron, et al. "Network motifs: simple building blocks of complex networks." *Science* 298.5594 (2002): 824-827.

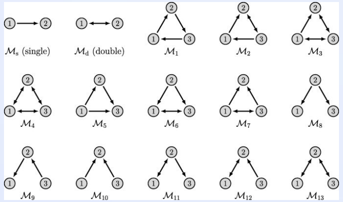
NETWORK MOTIFS [1]

are **subgraphs**, that **appear** in an observed network **significantly more often** than in compatible randomized networks.

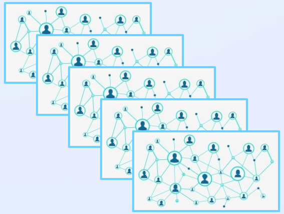
Procedure

1) Count all possible substructure of a given network.
2) Generate networks similar to the input one.

Input network



Subgraph counts



Null model

[1] Milo, Ron, et al. "Network motifs: simple building blocks of complex networks." *Science* 298.5594 (2002): 824-827.

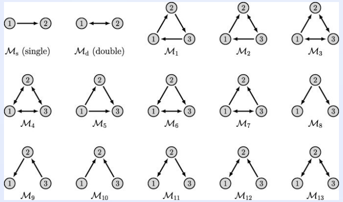
NETWORK MOTIFS [1]

are **subgraphs**, that **appear** in an observed network **significantly more often** than in compatible randomized networks.

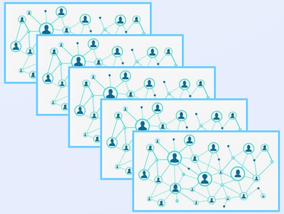
Procedure

1) Count all possible substructure of a given network.
2) Generate networks similar to the input one.
3) Count all possible substructure in the generated networks

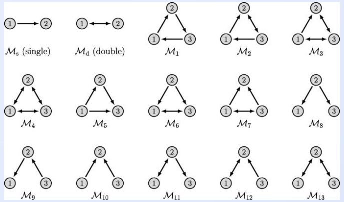
Input network



Subgraph counts



Null model



[1] Milo, Ron, et al. "Network motifs: simple building blocks of complex networks." *Science* 298.5594 (2002): 824-827.

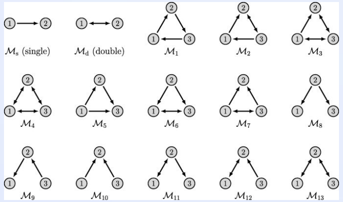
NETWORK MOTIFS [1]

are **subgraphs**, that **appear** in an observed network **significantly more often** than in compatible randomized networks.

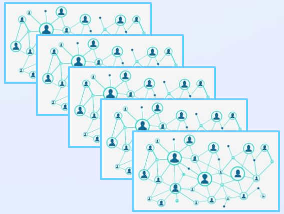
Procedure

1) Count all possible substructure of a given network.
2) Generate networks similar to the input one.
3) Count all possible substructure in the generated networks
4) Check for those substructure that are:
    1. Over-represented
    2. Minimum deviation
    3. Minimum frequency

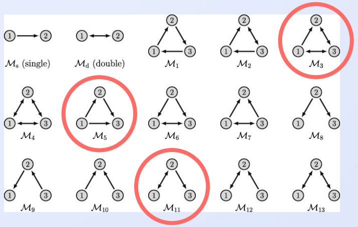
Input network



Subgraph counts



Null model



[1] Milo, Ron, et al. "Network motifs: simple building blocks of complex networks." *Science* 298.5594 (2002): 824-827.

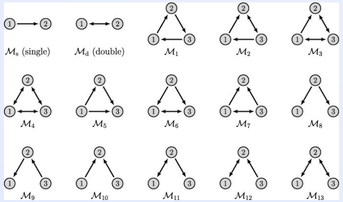
NETWORK MOTIFS [1]

are **subgraphs**, that **appear** in an observed network **significantly more often** than in compatible randomized networks.

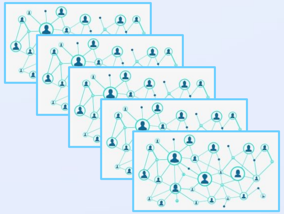
Procedure

1) Count all possible substructure of a given network.
2) Generate networks similar to the input one.
3) Count all possible substructure in the generated networks
4) Check for those substructure that are:
    1. Over-represented
    2. Minimum deviation
    3. Minimum frequency
5) Those structure are the motifs of the network.

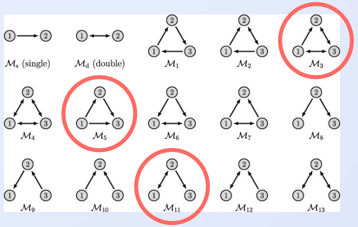
Input network



Subgraph counts



Null model



Network motifs

[1] Milo, Ron, et al. "Network motifs: simple building blocks of complex networks." *Science* 298.5594 (2002): 824-827.

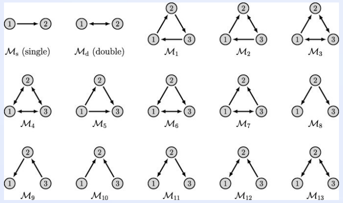
NETWORK MOTIFS [1]

are **subgraphs**, that **appear** in an observed network **significantly more often** than in compatible randomized networks.

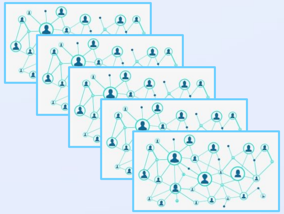
Procedure

1) Count all possible substructure of a given network.
2) Generate networks similar to the input one.
3) **Count all possible substructure in the generated networks**
4) Check for those substructure that are:
    1. Over-represented
    2. Minimum deviation
    3. Minimum frequency
5) Those structure are the motifs of the network.

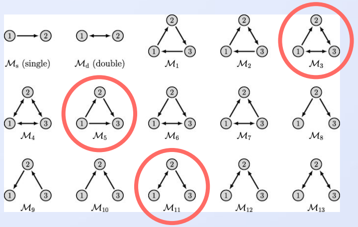
Input network



Subgraph counts



Null model



Network motifs

[1] Milo, Ron, et al. "Network motifs: simple building blocks of complex networks." *Science* 298.5594 (2002): 824-827.

NETWORK MOTIFS

are **subgraphs**, that **appear** in an observed network **significantly more often** than in compatible randomized networks.

Procedure

1) Count all possible substructure of a given network.
2) Generate networks similar to the input one.
3) **Count all possible substructure in the generated networks**

How many substructure are there?

NETWORK MOTIFS

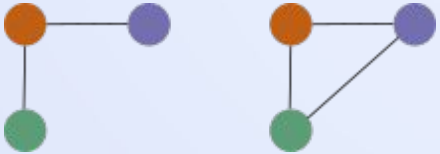
are **subgraphs**, that **appear** in an observed network **significantly more often** than in compatible randomized networks.

Procedure

1) Count all possible substructure of a given network.
2) Generate networks similar to the input one.
3) **Count all possible substructure in the generated networks**

How many substructure are there?

3 nodes



NETWORK MOTIFS

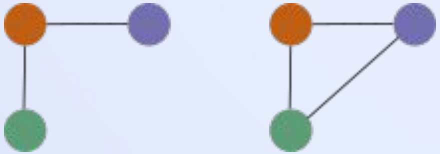
are **subgraphs**, that **appear** in an observed network **significantly more often** than in compatible randomized networks.

Procedure

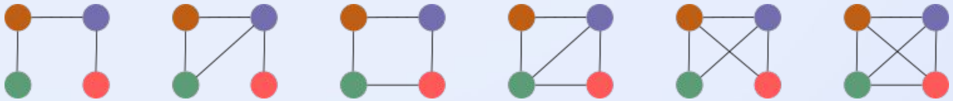
1) Count all possible substructure of a given network.
2) Generate networks similar to the input one.
3) **Count all possible substructure in the generated networks**

How many substructure are there?

3 nodes



4 nodes



NETWORK MOTIFS

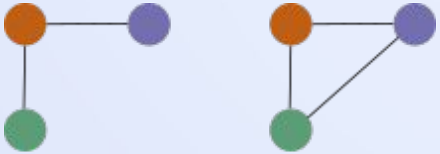
are **subgraphs**, that **appear** in an observed network **significantly more often** than in compatible randomized networks.

Procedure

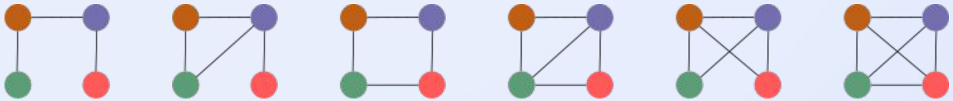
1) Count all possible substructure of a given network.
2) Generate networks similar to the input one.
3) **Count all possible substructure in the generated networks**

How many substructure are there?

3 nodes



4 nodes



5 nodes

21 structures

NETWORK MOTIFS

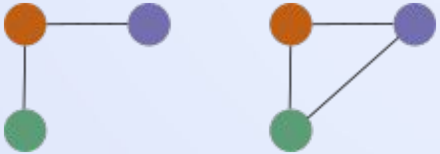
are **subgraphs**, that **appear** in an observed network **significantly more often** than in compatible randomized networks.

Procedure

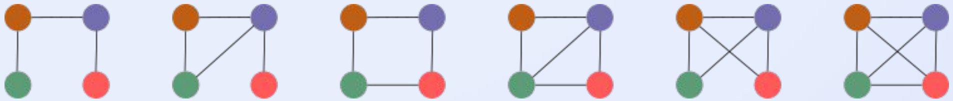
1) Count all possible substructure of a given network.
2) Generate networks similar to the input one.
3) **Count all possible substructure in the generated networks**

How many substructure are there?

3 nodes



4 nodes



5 nodes

21 structures

10 nodes

******

NETWORK MOTIFS

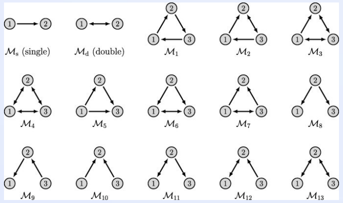
are **subgraphs**, that **appear** in an observed network **significantly more often** than in compatible randomized networks.

Procedure

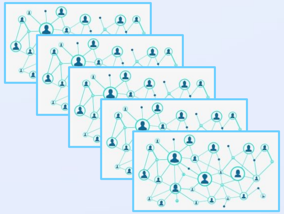
1) Count all possible substructure of a given network.
2) Generate networks similar to the input one.
3) **Count all possible substructure in the generated networks**
4) Check for those substructure that are:
    1. Over-represented
    2. Minimum deviation
    3. Minimum frequency
5) Those structure are the motifs of the network.

Computational expensive

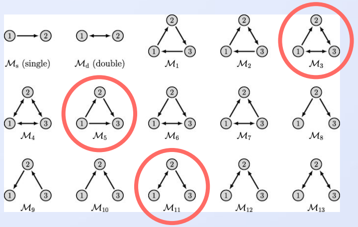
Input network



Subgraph counts



Null model



Network motifs

Temporal network motifs

02

TEMPORAL NETWORK MOTIFS

Many times networks are not enough to represent real world scenarios.

Interactions change over time…
Images could be videos…
Traffic on roads change…

So temporal networks solve this problem.

TEMPORAL NETWORK MOTIFS

Many times networks are not enough to represent real world scenarios.

Interactions change over time...
Images could be videos...
Traffic on roads change...

So temporal networks solve this problem.

Temporal network:

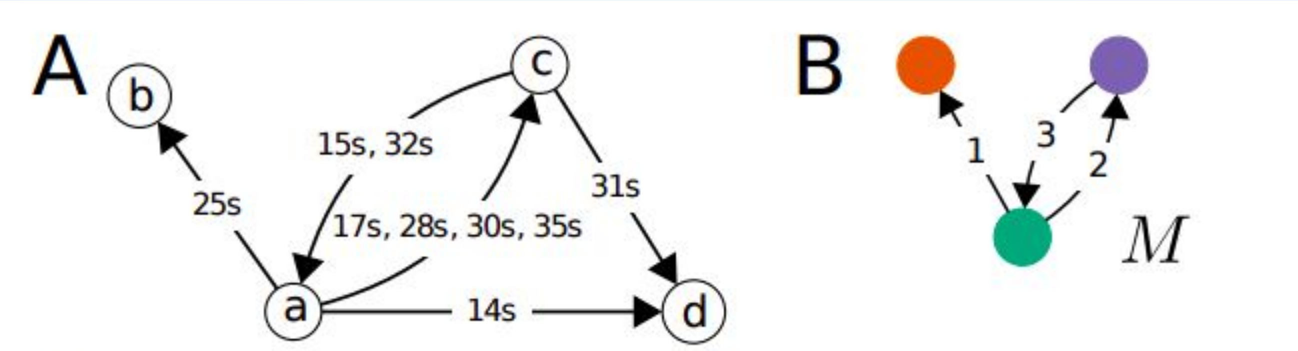
1) **Edges** → interactions among peoples
2) **Nodes** → users in social networks
3) **Attributes** → enemies can become friends

TEMPORAL NETWORK MOTIFS

Obviously, even temporal network has motifs.

Temporal network

Temporal network motifs



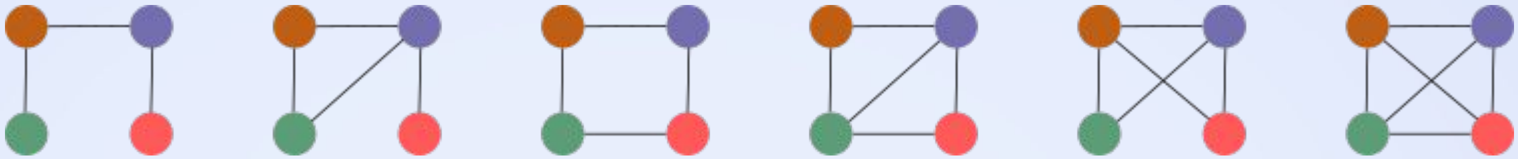
Paranjape, Ashwin, Austin R. Benson, and Jure Leskovec. "Motifs in temporal networks." *Proceedings of the tenth ACM international conference on web search and data mining*. 2017.

TEMPORAL NETWORK MOTIFS

Obviously, even temporal network has motifs.

How many substructure are there?

Network

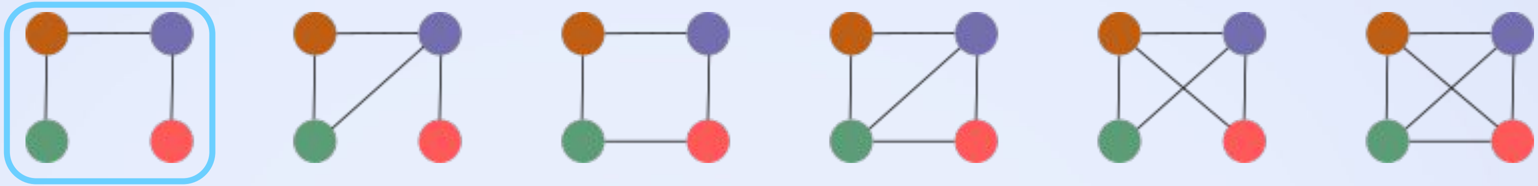


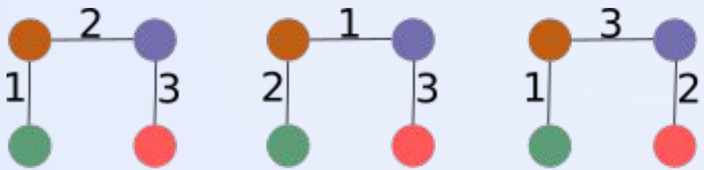
TEMPORAL NETWORK MOTIFS

Obviously, even temporal network has motifs.

How many substructure are there?

Network





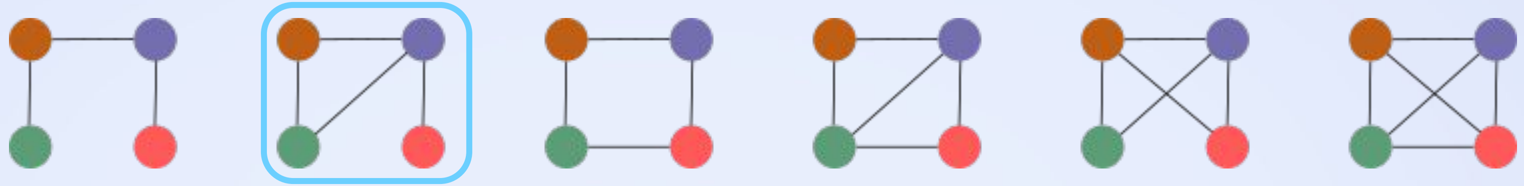
Temporal network

TEMPORAL NETWORK MOTIFS

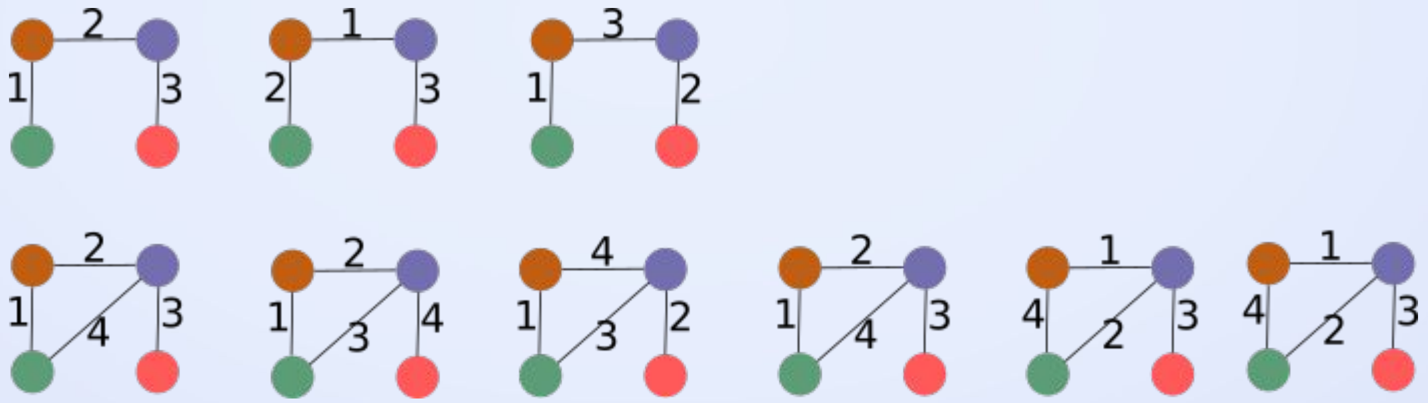
Obviously, even temporal network has motifs.

How many substructure are there?

Network



Temporal network



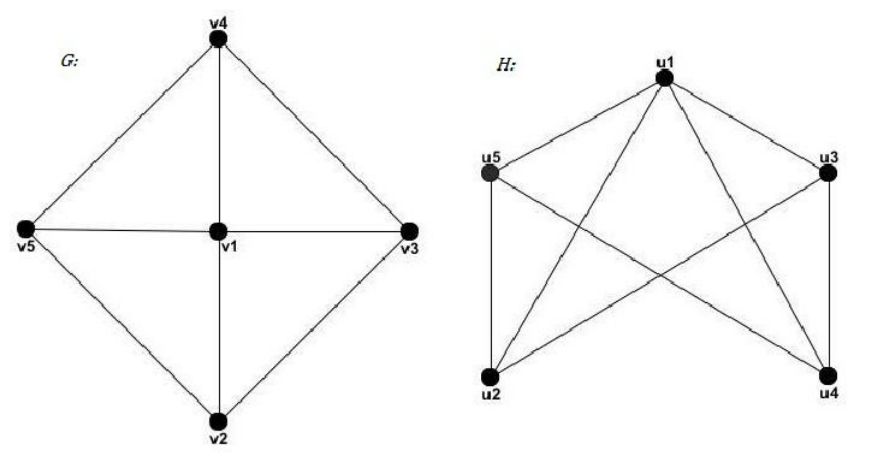
A lot of more

TEMPORAL NETWORK MOTIFS

The **time** required to **count** motifs in **temporal network** is **higher** due to the **complexity** introduced by the **temporal** dimension.

TEMPORAL NETWORK MOTIFS

The **time** required to **count** motifs in **temporal network** is **higher** due to the **complexity** introduced by the **temporal** dimension.



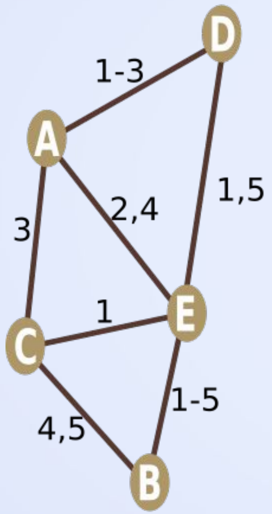
If the size of the sub graph is big, we have to compute an **isomorphism test.**
It requires lot of time!

Egocentric Temporal Motifs

03

EGOCENTRIC TEMPORAL MOTIFS

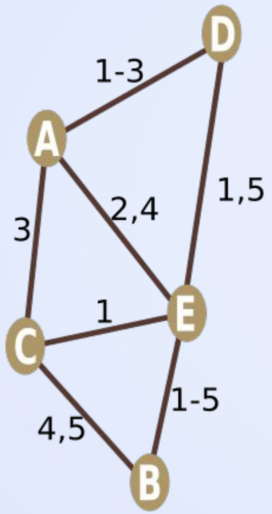
Temporal graph

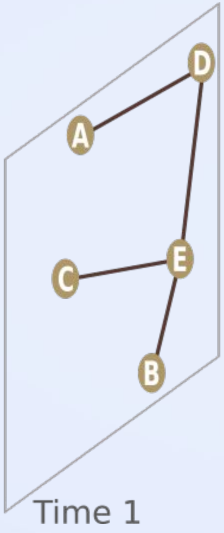


A. Longa, G. Cencetti, B. Lepri, and A. Passerini, “An efficient procedure for mining egocentric temporal motifs,” Data Mining and Knowledge Discovery, pp. 1–24, 2021.

EGOCENTRIC TEMPORAL MOTIFS

Temporal graph

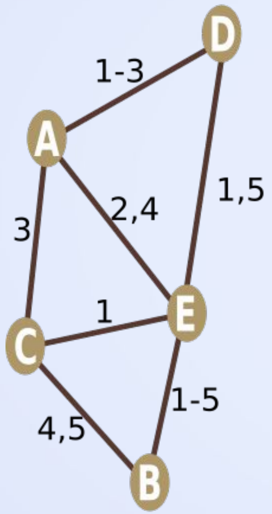


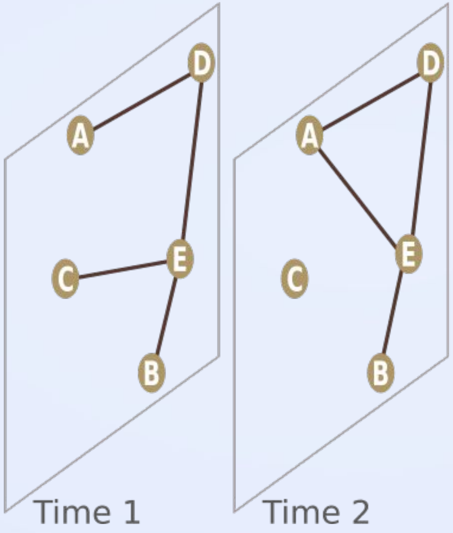


A. Longa, G. Cencetti, B. Lepri, and A. Passerini, "An efficient procedure for mining egocentric temporal motifs," Data Mining and Knowledge Discovery, pp. 1–24, 2021.

EGOCENTRIC TEMPORAL MOTIFS

Temporal graph

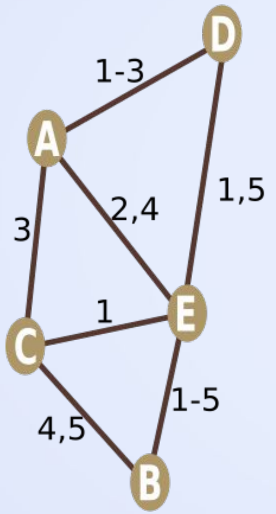


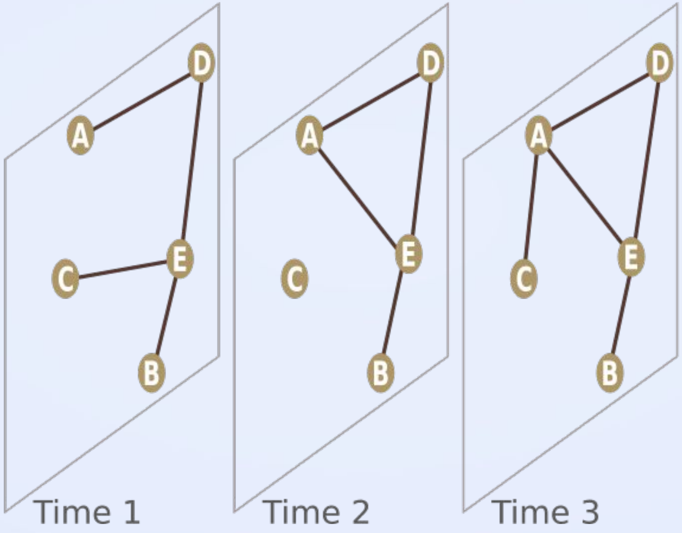


A. Longa, G. Cencetti, B. Lepri, and A. Passerini, "An efficient procedure for mining egocentric temporal motifs," Data Mining and Knowledge Discovery, pp. 1–24, 2021.

EGOCENTRIC TEMPORAL MOTIFS

Temporal graph

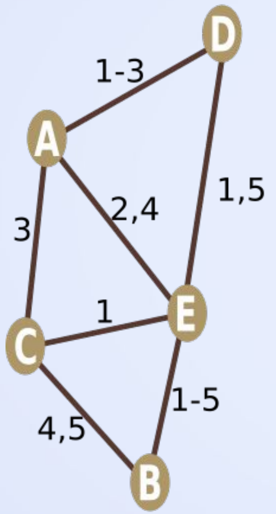




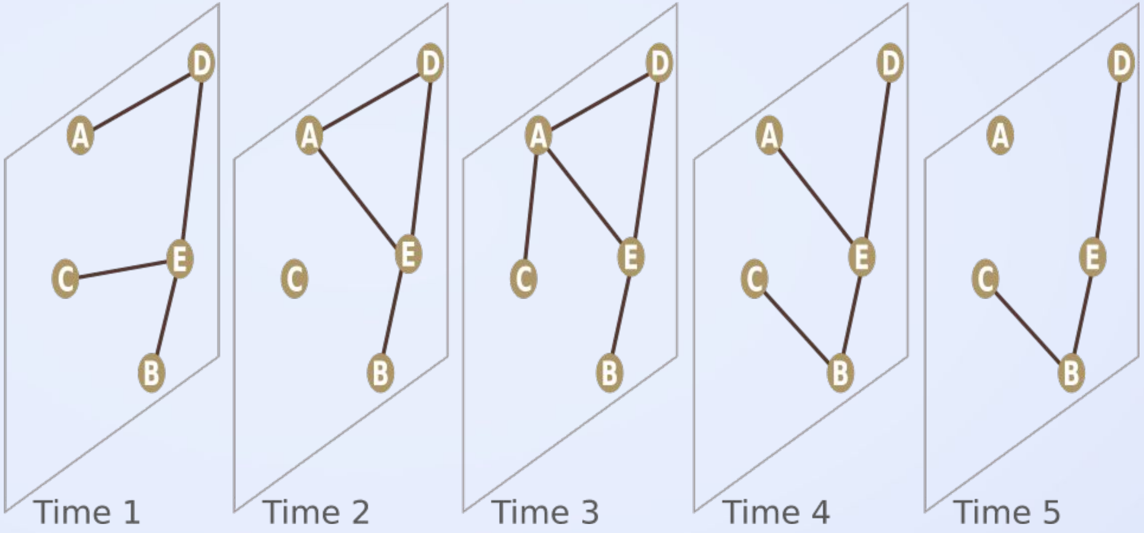
A. Longa, G. Cencetti, B. Lepri, and A. Passerini, “An efficient procedure for mining egocentric temporal motifs,” Data Mining and Knowledge Discovery, pp. 1–24, 2021.

EGOCENTRIC TEMPORAL MOTIFS

Temporal graph



Temporal graph snapshots



A. Longa, G. Cencetti, B. Lepri, and A. Passerini, “An efficient procedure for mining egocentric temporal motifs,” Data Mining and Knowledge Discovery, pp. 1–24, 2021.

EGOCENTRIC TEMPORAL MOTIFS

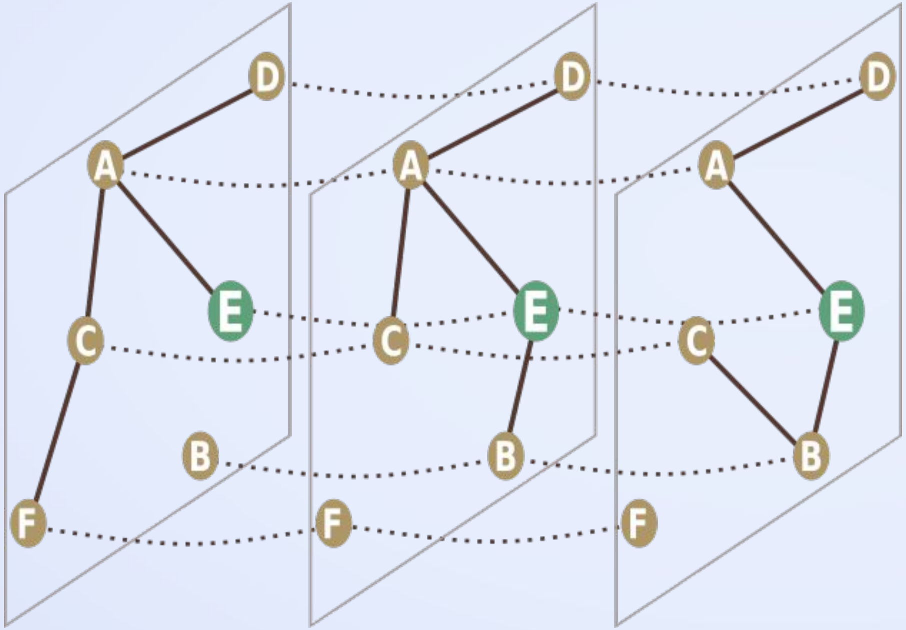
K = 2

A. Longa, G. Cencetti, B. Lepri, and A. Passerini, “An efficient procedure for mining egocentric temporal motifs,” Data Mining and Knowledge Discovery, pp. 1–24, 2021.

EGOCENTRIC TEMPORAL MOTIFS

K = 2

Decide and EGO Node = E

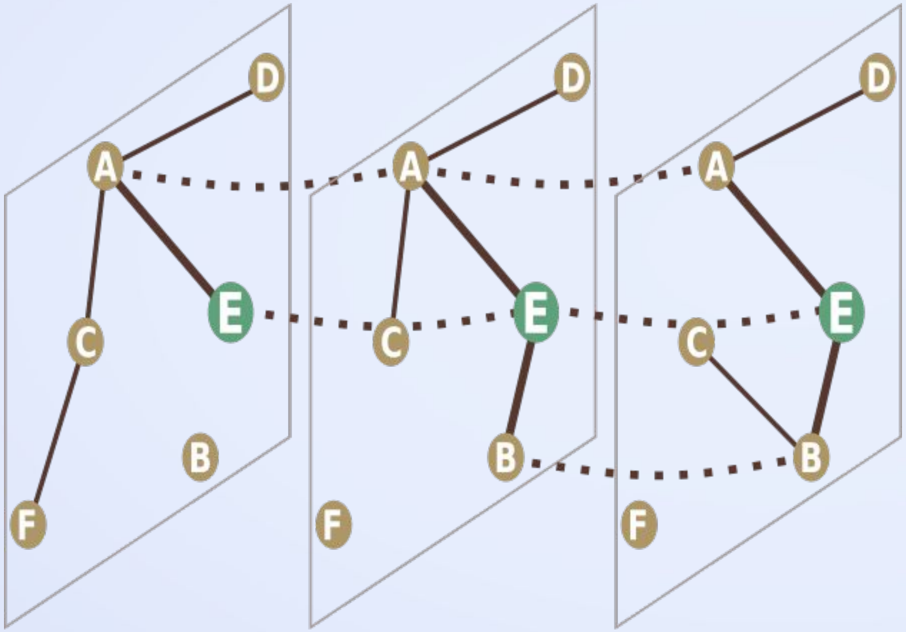


A. Longa, G. Cencetti, B. Lepri, and A. Passerini, "An efficient procedure for mining egocentric temporal motifs," Data Mining and Knowledge Discovery, pp. 1–24, 2021.

EGOCENTRIC TEMPORAL MOTIFS

K = 2

Decide and EGO Node = E

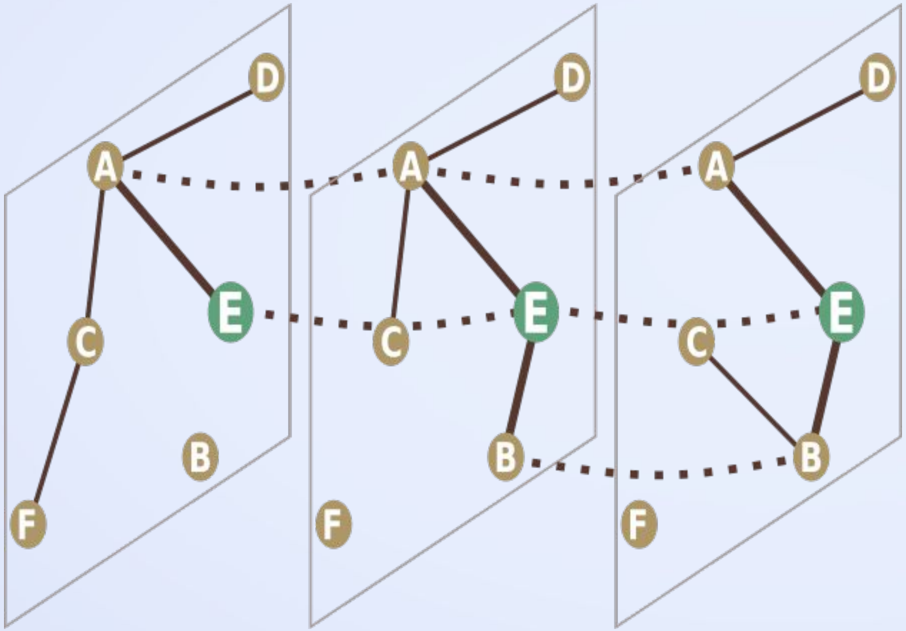


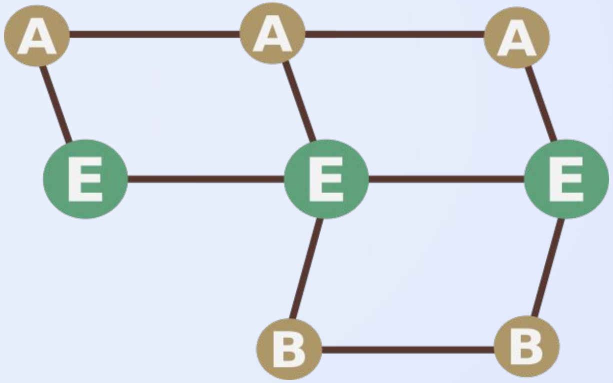
A. Longa, G. Cencetti, B. Lepri, and A. Passerini, "An efficient procedure for mining egocentric temporal motifs," Data Mining and Knowledge Discovery, pp. 1–24, 2021.

EGOCENTRIC TEMPORAL MOTIFS

K = 2

Decide and EGO Node = E

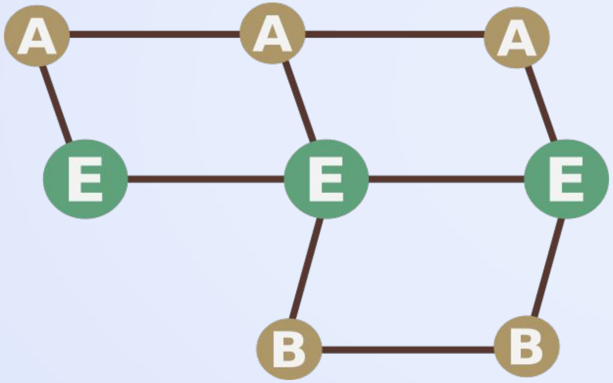




A. Longa, G. Cencetti, B. Lepri, and A. Passerini, “An efficient procedure for mining egocentric temporal motifs,” Data Mining and Knowledge Discovery, pp. 1–24, 2021.

EGOCENTRIC TEMPORAL MOTIFS

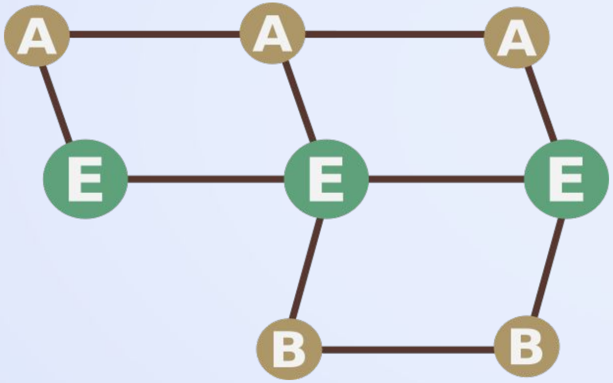
K = 2
Decide and EGO Node = E



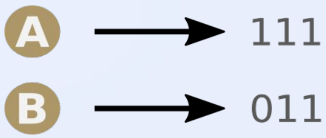
A. Longa, G. Cencetti, B. Lepri, and A. Passerini, “An efficient procedure for mining egocentric temporal motifs,” Data Mining and Knowledge Discovery, pp. 1–24, 2021.

EGOCENTRIC TEMPORAL MOTIFS

K = 2
Decide and EGO Node = E



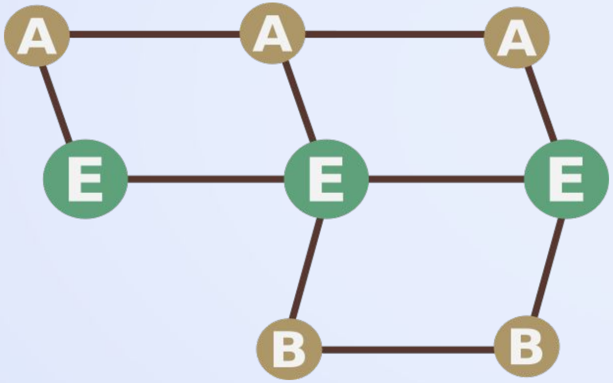
NODE ENCODING

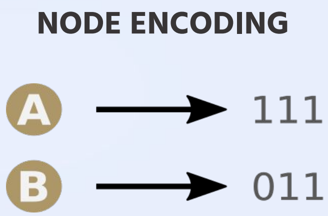


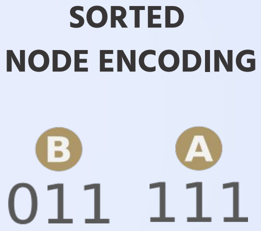
A. Longa, G. Cencetti, B. Lepri, and A. Passerini, “An efficient procedure for mining egocentric temporal motifs,” Data Mining and Knowledge Discovery, pp. 1–24, 2021.

EGOCENTRIC TEMPORAL MOTIFS

K = 2
Decide and EGO Node = E



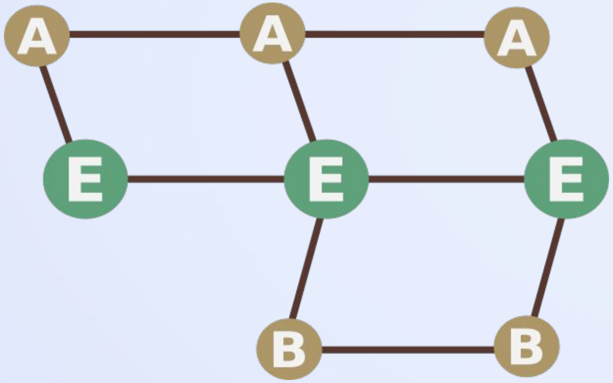




A. Longa, G. Cencetti, B. Lepri, and A. Passerini, "An efficient procedure for mining egocentric temporal motifs," Data Mining and Knowledge Discovery, pp. 1–24, 2021.

EGOCENTRIC TEMPORAL MOTIFS

K = 2
Decide and EGO Node = E



NODE ENCODING

A → 111
B → 011

SORTED
NODE ENCODING

B A
011 111

Egocentric Temporal Neighbourhood Signature
ETNS

011 111

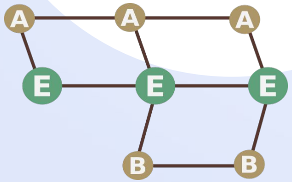
A. Longa, G. Cencetti, B. Lepri, and A. Passerini, “An efficient procedure for mining egocentric temporal motifs,” Data Mining and Knowledge Discovery, pp. 1–24, 2021.

IN SHORT

01

ETN

Egocentric
Temporal
Neighbourhood.
(a sub structure)

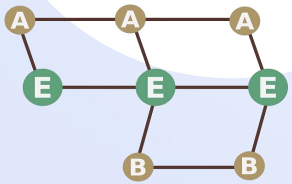


IN SHORT

01

ETN

Egocentric
Temporal
Neighbourhood.
(a sub structure)



02

ETNS

Egocentric Temporal
Neighbourhood Signature.
(a string representing a sub
structure)

011 111

IN SHORT

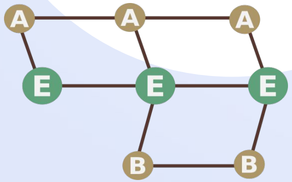
03

Fast way to compute if two sub structures are identical

01

ETN

Egocentric Temporal Neighbourhood.
(a sub structure)



02

ETNS

Egocentric Temporal Neighbourhood Signature.
(a string representing a sub structure)

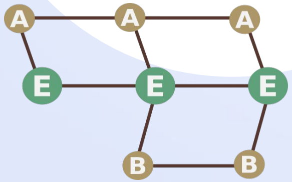
011 111

IN SHORT

01

ETN

Egocentric
Temporal
Neighbourhood.
(a sub structure)



02

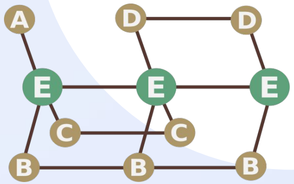
ETNS

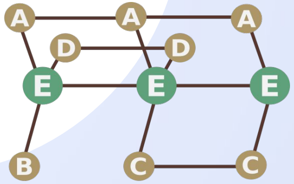
Egocentric Temporal
Neighbourhood Signature.
(a string representing a sub
structure)

011 111

03

Fast way to compute if two
sub structures are identical



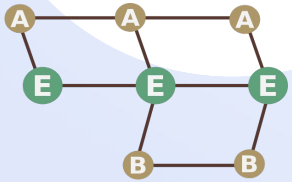


IN SHORT

01

ETN

Egocentric Temporal Neighbourhood. (a sub structure)



02

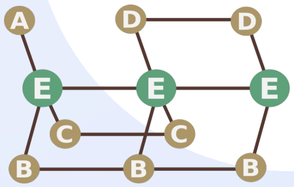
ETNS

Egocentric Temporal Neighbourhood Signature. (a string representing a sub structure)

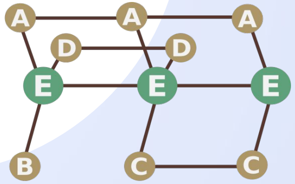
011 111

03

Fast way to compute if two sub structures are identical



011 100 110 111



011 100 110 111

Egocentric temporal motifs

Procedure

1) Count all possible **egocentric** substructure of a given network.
2) Generate networks similar to the input one.
3) **Count all possible ego substructure in the generated networks**
4) Check for those **egocentric** substructure that are:
    1. Over-represented
    2. Minimum deviation
    3. Minimum frequency
5) Those egocentric structure are the EGOCENTRIC TEMPORAL MOTIFS

Now it is fast

04

APPLICATIONS

Cool! How can we use those structures?

APPLICATIONS

COMPUTE DISTANCES

APPLICATIONS

COMPUTE DISTANCES

$$dist_M(\mathcal{G}_1, \mathcal{G}_2) = 1 - \frac{EMB_M(\mathcal{G}_1) \cdot EMB_M(\mathcal{G}_2)}{||EMB_M(\mathcal{G}_1)|| \; ||EMB_M(\mathcal{G}_2)||}$$

APPLICATIONS

COMPUTE DISTANCES

$$dist_M\boxed{(\mathcal{G}_1, \mathcal{G}_2)} = 1 - \frac{EMB_M(\mathcal{G}_1) \cdot EMB_M(\mathcal{G}_2)}{||EMB_M(\mathcal{G}_1)||\ ||EMB_M(\mathcal{G}_2)||}$$

Input graphs

APPLICATIONS

COMPUTE DISTANCES

Cosine similarity

$$dist_M\boxed{(\mathcal{G}_1, \mathcal{G}_2)} = 1 - \boxed{\frac{EMB_M(\mathcal{G}_1) \cdot EMB_M(\mathcal{G}_2)}{||EMB_M(\mathcal{G}_1)|| \; ||EMB_M(\mathcal{G}_2)||}}$$

Input graphs

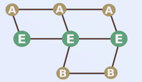
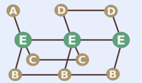
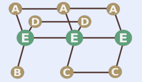
APPLICATIONS

COMPUTE DISTANCES

Cosine similarity

$$dist_M\boxed{(\mathcal{G}_1, \mathcal{G}_2)} = 1 - \boxed{\frac{\boxed{EMB_M(\mathcal{G}_1)} \cdot EMB_M(\mathcal{G}_2)}{||EMB_M(\mathcal{G}_1)||\ ||EMB_M(\mathcal{G}_2)||}}$$

Input graphs

ETM			
Counts	11561	35815	85112

COMPUTE DISTANCES

RESULTS

Sociopatter data, face to face interactions

Workplace
Hospital
High School 11
High School 12
High School 13
Primary school
University

COMPUTE DISTANCES

RESULTS

				ETMM-DIST			
	VS13	LH10	HS11	HS12	HS13	PS	DTU
Workplace	0	0.07	0.29	0.22	0.29	0.67	0.47
Hospital		0	0.29	0.22	0.30	0.66	0.45
High School 11			0	0.04	0.04	0.59	0.06
High School 12				0	0.02	0.61	0.13
High School 13					0	0.62	0.08
Primary school						0	0.62
University							0

COMPUTE DISTANCES
RESULTS

				ETMM-DIST			
	VS13	LH10	HS11	HS12	HS13	PS	DTU
Workplace	0	0.07	0.29	0.22	0.29	0.67	0.47
Hospital		0	0.29	0.22	0.30	0.66	0.45
High School 11			0	0.04	0.04	0.59	0.06
High School 12				0	0.02	0.61	0.13
High School 13					0	0.62	0.08
Primary school						0	0.62
University							0

Workplace and
Hospital are similar

COMPUTE DISTANCES

RESULTS

				ETMM-DIST			
	VS13	LH10	HS11	HS12	HS13	PS	DTU
Workplace	0	0.07	0.29	0.22	0.29	0.67	0.47
Hospital		0	0.29	0.22	0.30	0.66	0.45
High School 11			0	0.04	0.04	0.59	0.06
High School 12				0	0.02	0.61	0.13
High School 13					0	0.62	0.08
Primary school						0	0.62
University							0

Workplace and Hospital are similar

High Schools are similar

COMPUTE DISTANCES

RESULTS

				ETMM-DIST			
	VS13	LH10	HS11	HS12	HS13	PS	DTU
Workplace	0	0.07	0.29	0.22	0.29	0.67	0.47
Hospital		0	0.29	0.22	0.30	0.66	0.45
High School 11			0	0.04	0.04	0.59	0.06
High School 12				0	0.02	0.61	0.13
High School 13					0	0.62	0.08
Primary school						0	0.62
University							0

Workplace and Hospital are similar

High Schools are similar

Primary school is different from all the networks

COMPUTE DISTANCES
RESULTS

				ETMM-DIST			
	VS13	LH10	HS11	HS12	HS13	PS	DTU
Workplace	0	0.07	0.29	0.22	0.29	0.67	0.47
Hospital		0	0.29	0.22	0.30	0.66	0.45
High School 11			0	0.04	0.04	0.59	0.06
High School 12				0	0.02	0.61	0.13
High School 13					0	0.62	0.08
Primary school						0	0.62
University							0

Quite similar to High Schools

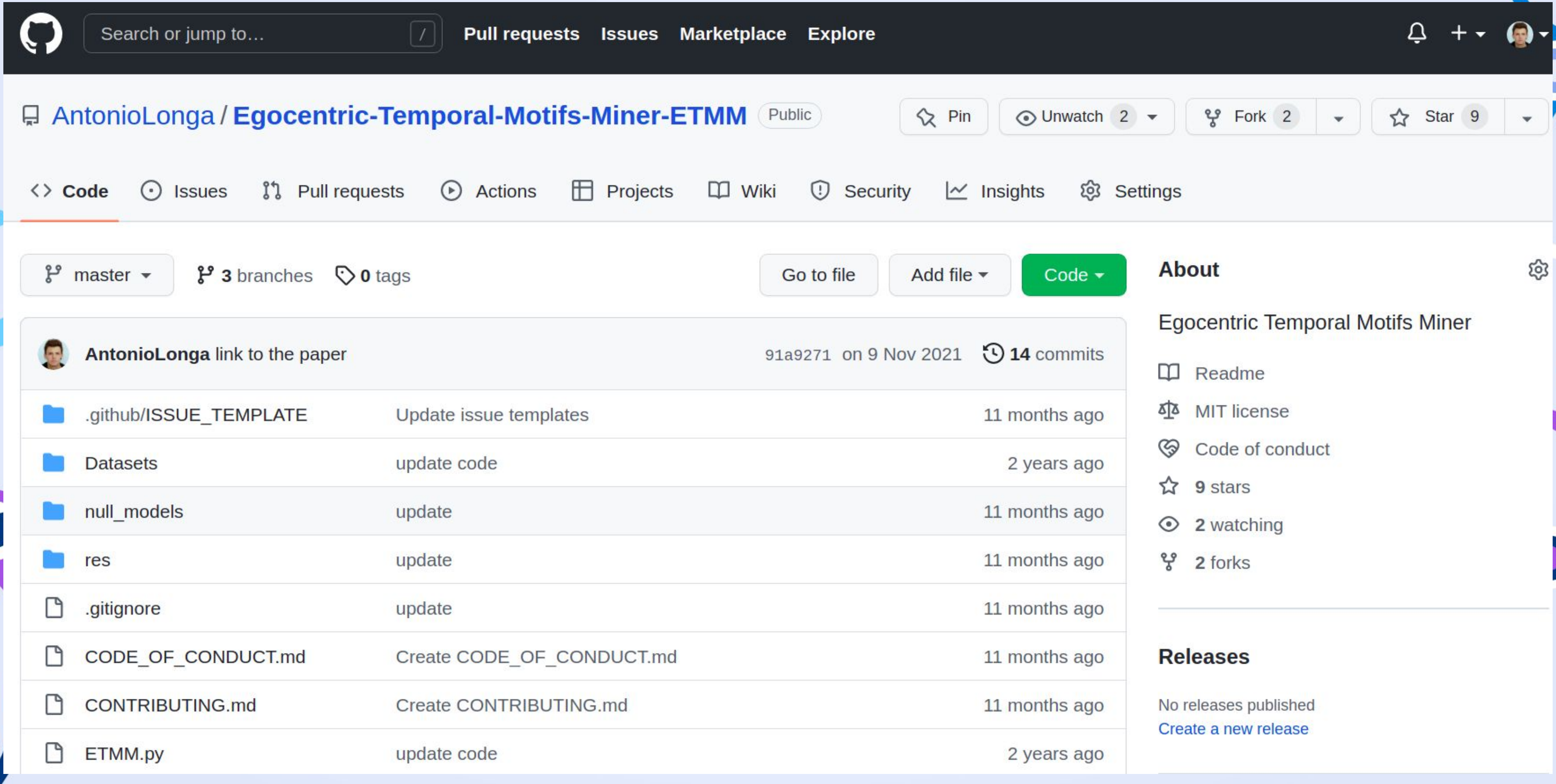
COMPUTE DISTANCES

RESULTS

				ETMM-DIST			
	VS13	LH10	HS11	HS12	HS13	PS	DTU
Workplace	0	0.07	0.29	0.22	0.29	0.67	0.47
Hospital		0	0.29	0.22	0.30	0.66	0.45
High School 11			0	0.04	0.04	0.59	0.06
High School 12				0	0.02	0.61	0.13
High School 13					0	0.62	0.08
Primary school						0	0.62
University							0

Quite similar to High Schools

Different from primary school


Search or jump to…
Pull requests
Issues
Marketplace
Explore
AntonioLonga / Egocentric-Temporal-Motifs-Miner-ETMM
Public
Pin
Unwatch 2
Fork 2
Star 9
Code
Issues
Pull requests
Actions
Projects
Wiki
Security
Insights
Settings
master
3 branches
0 tags
Go to file
Add file
Code
AntonioLonga link to the paper
91a9271 on 9 Nov 2021
14 commits
.github/ISSUE_TEMPLATE
Update issue templates
11 months ago
Datasets
update code
2 years ago
null_models
update
11 months ago
res
update
11 months ago
.gitignore
update
11 months ago
CODE_OF_CONDUCT.md
Create CODE_OF_CONDUCT.md
11 months ago
CONTRIBUTING.md
Create CONTRIBUTING.md
11 months ago
ETMM.py
update code
2 years ago
About
Egocentric Temporal Motifs Miner
Readme
MIT license
Code of conduct
9 stars
2 watching
2 forks
Releases
No releases published
Create a new release

05

Future directions

ETN as building block for temporal network generation

Neighborhood matching creates realistic surrogate temporal networks

Antonio Longa[1,2] Giulia Cencetti[1], Sune Lehmann[3,4], Andrea Passerini[2] and Bruno Lepri[1*]

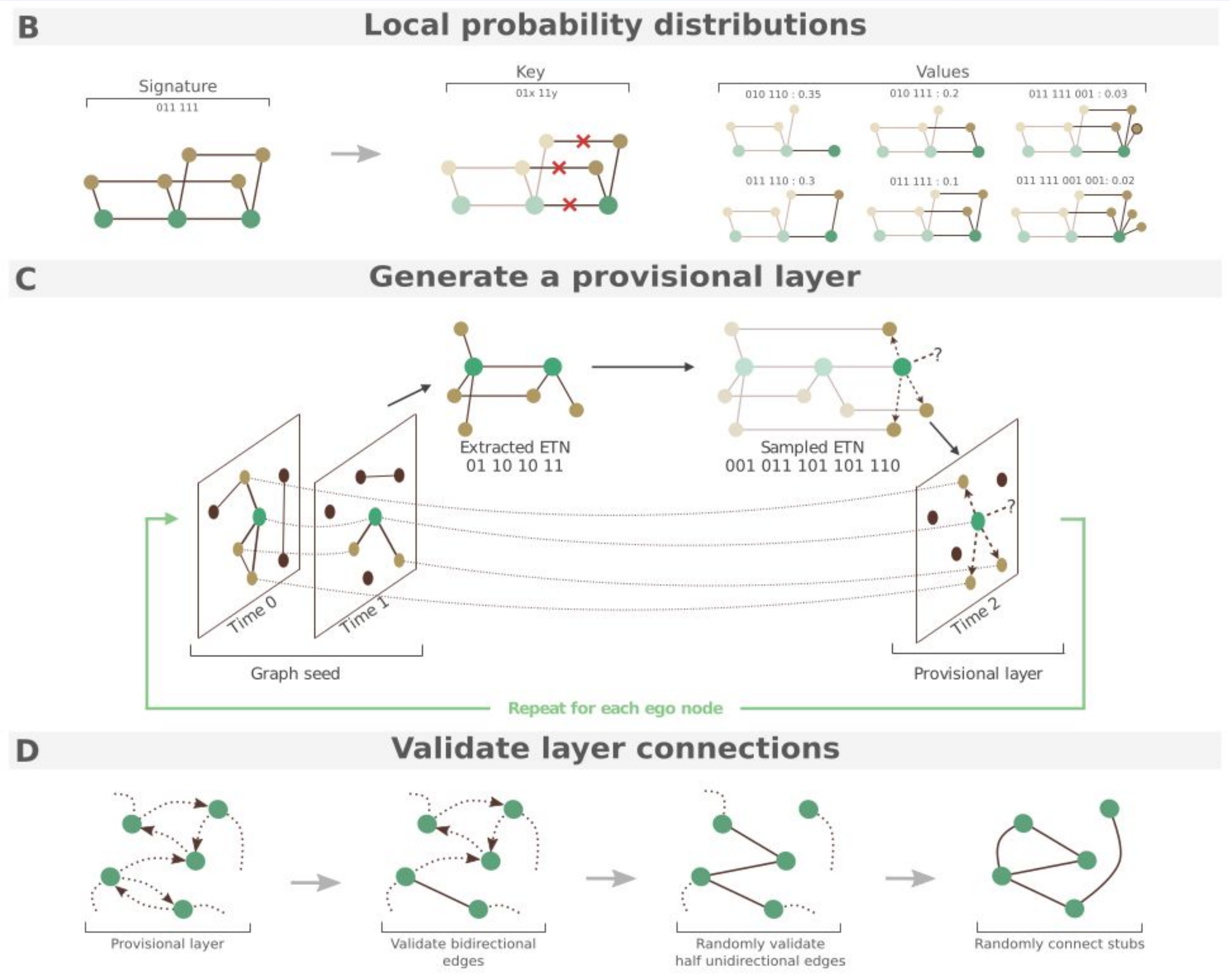
[1]Fondazione Bruno Kessler, Trento, Italy

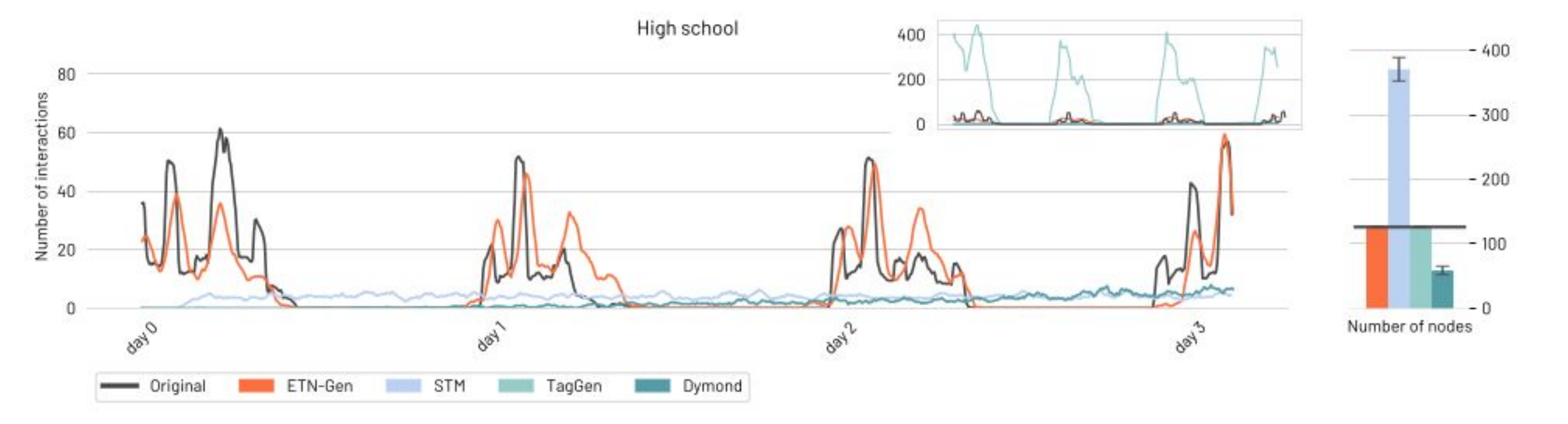
[1]University of Trento, Trento, Italy

[1]Tecnical University of Denmark, Kongens Lyngby, Denmark

[2]Copenhagen Center for Social Data Science, Copenhagen, Denmark

ETN as building block for temporal network generation




High school
Number of interactions
80
60
40
20
0
day 0
day 1
day 2
day 3
400
200
0
400
300
200
100
0
Number of nodes
Original
ETN-Gen
STM
TagGen
Dymond

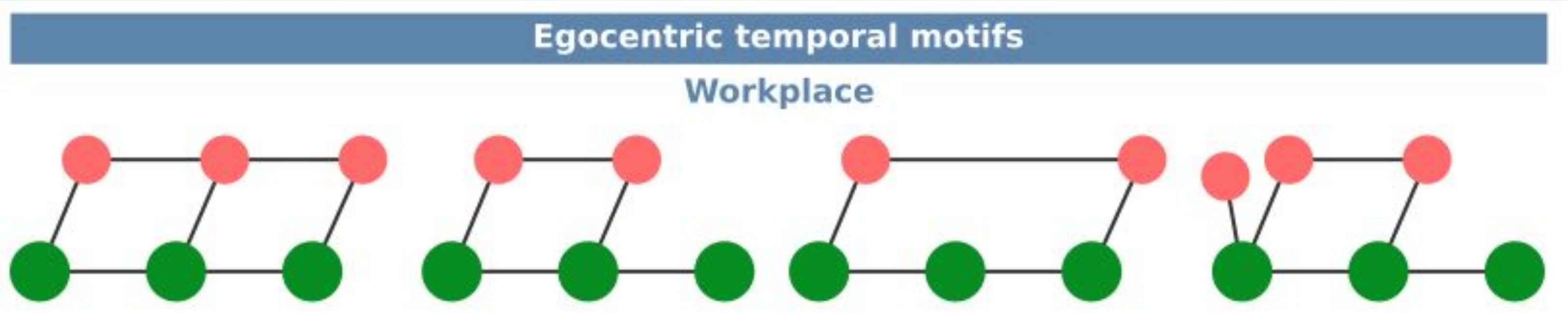
THANKS

Do you have any questions?

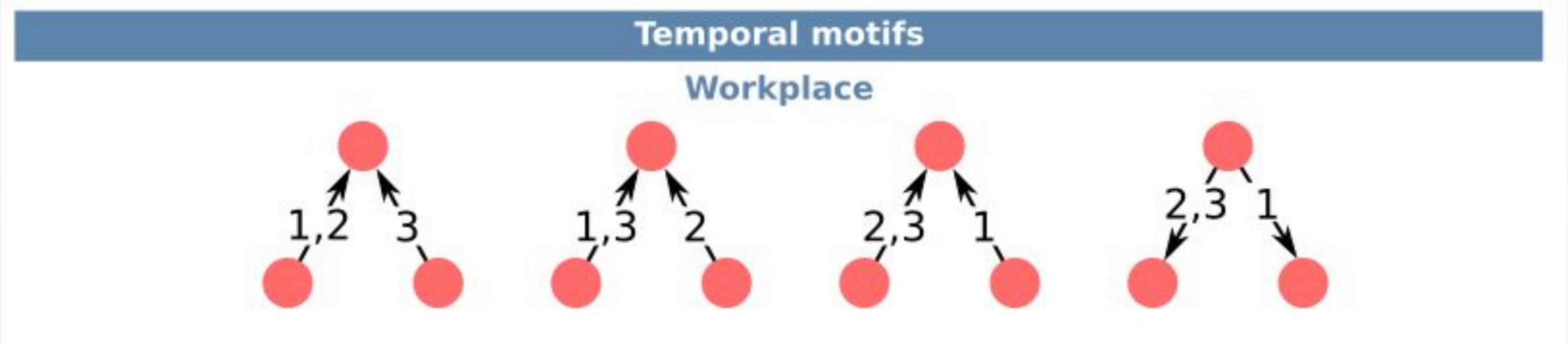
alonga@fbk.eu
https://antoniolonga.github.io/

CREDITS: This presentation template was created by **Slidesgo**, including icons by **Flaticon**, and infographics & images by **Freepik**

WORKPLACE



Egocentric temporal motifs
Workplace



Temporal motifs
Workplace
1,2
3
1,3
2
2,3
1
2,3
1