Epidemiological processes on networks: Data analysis

5 JAN - 5 APR 2019

Supervisor

Liubov Tupikina

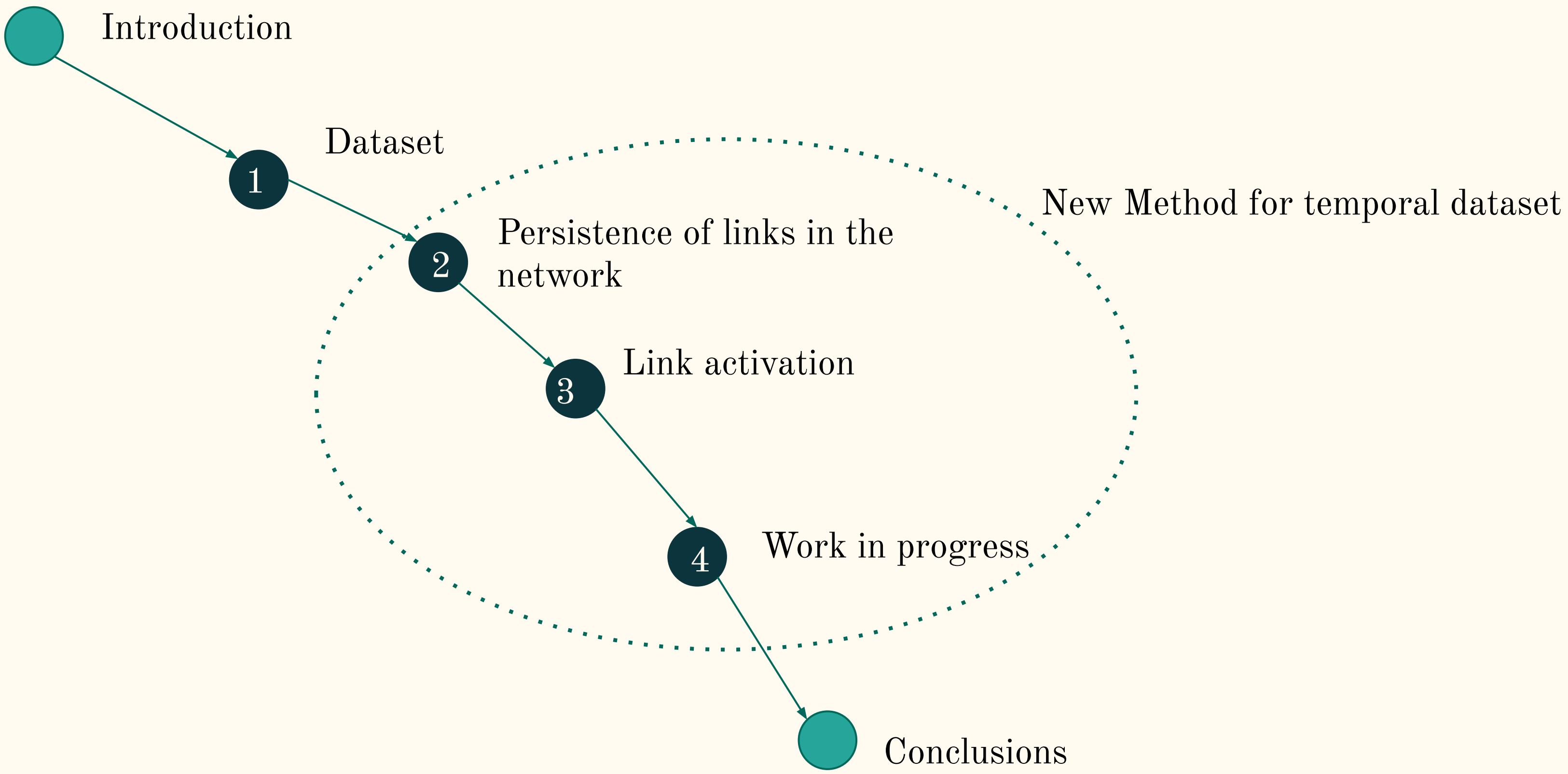
Aurélie Faure de Pebeyre

CRI-PARIS

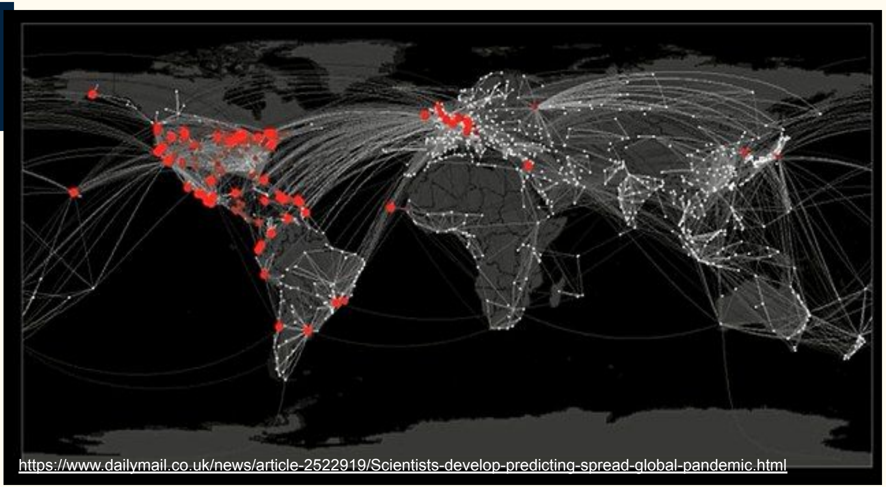
Motivation & Introduction

Image Source: PBS

Outline


Introduction
1
Dataset
New Method for temporal dataset
2
Persistence of links in the network
3
Link activation
4
Work in progress
Conclusions

Introduction



Networks can help represent datas like connection between cities, or animal transport.

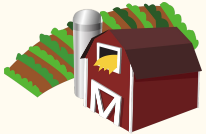
On a epidemic point of view, it is used to see the spreading.

Data collected from different sources: people, animals...
Here we will focus on animal transported from farm to farm.

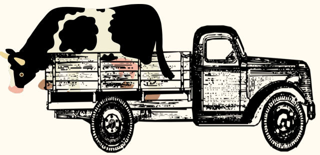
>**Goal of the internship**: to understand the effects of spreading on temporal networks, to develop theoretical framework for analysis of disease spreading processes.

1 Dataset

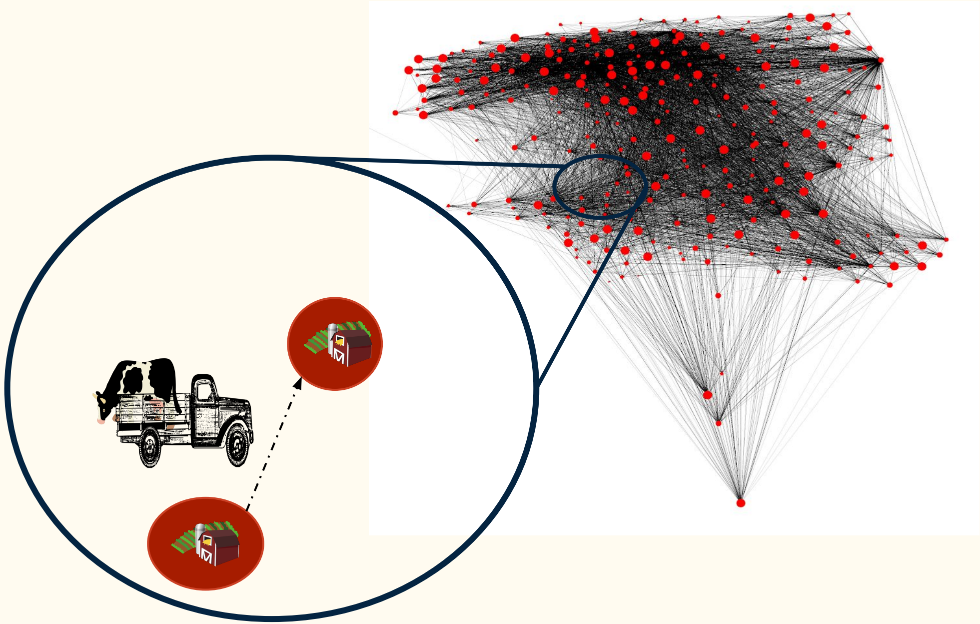
DATASET : cow movement between farms



node = farm (or market)



link = a animal is transported



INRA dataset Elizabetha Vergu

INRA: P.Hoscheit, et al. "Dynamical network models for cattle trade: towards economy-based epidemic risk assessment." (2016)

DATASET : properties

: 5 798 nodes
3 types of nodes
(most of the type=farm,
few of the two other kind)

: 456 420 edges

2005>2009 =5 ans

Dataset properties to study:

1- temporality and directionality
2- heterogeneous properties (in-out degree)

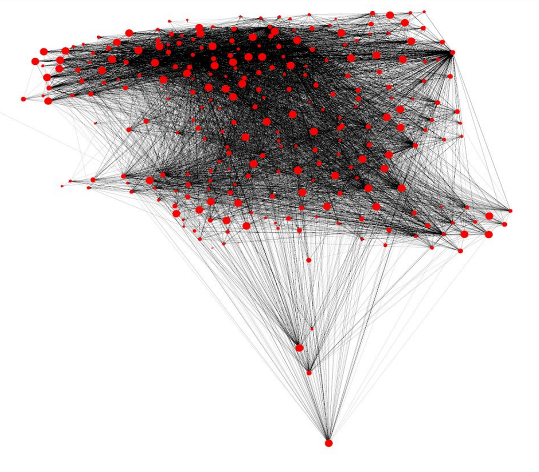
Ideas developed:

1-temporal network can be represented via aggregated, accessibility network[1]

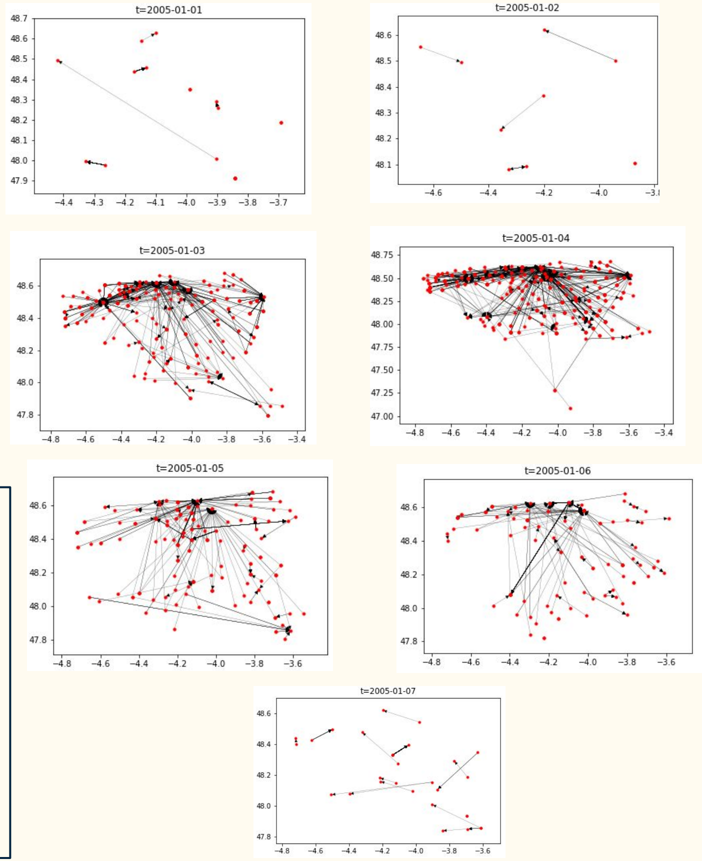
2-apply algorithms for heterogeneous network properties [2]

[1] H.Lentz et al, PRL (2013)
[2] D.Grebenkov et al. PRE (2018)
[3] P.Hoscheit, et al. J. Com. Net. (2016)

Aggregated network VS temporal network



One week from Saturday to Friday



Aggregated Network:links are aggregated over all time steps

+: gives an broder view

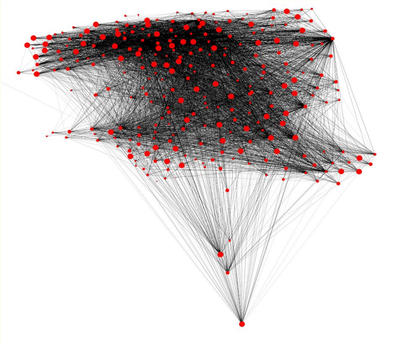
- : can erase some information

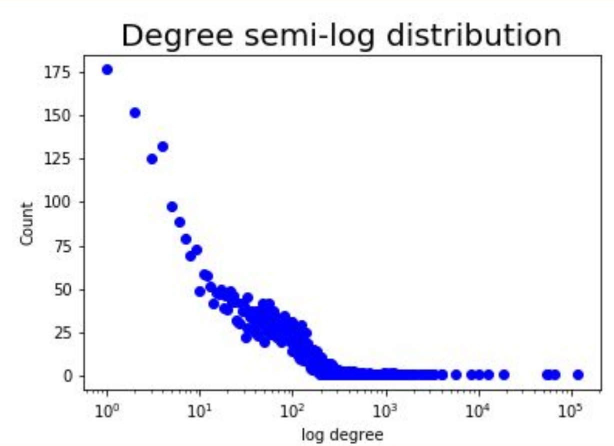
Temporal Network: one view per timestep

+: more precise

- : hard to have a good overview

IN-OUT degree distribution





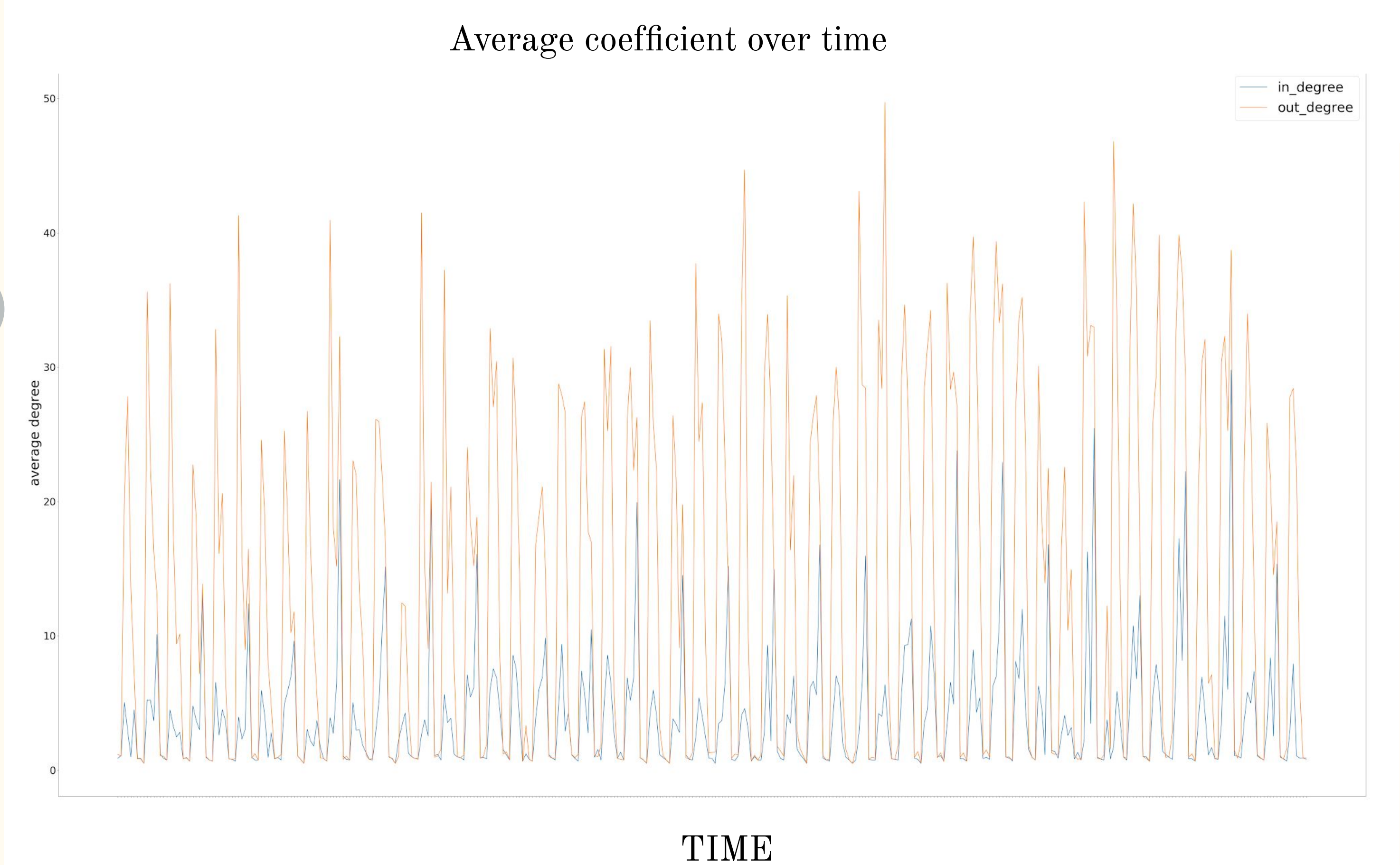
Degree of a node: number of edges connected to that node
here: scale free network

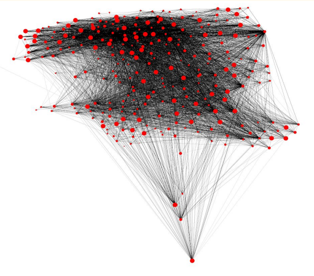
In-degree : number of edges incoming to a node.

Out degree: number of edges outcoming from a node.

Why : we want to study how much the directionality affects the whole network structure

IN-OUT degree





Out degree> In degree.

farms are sending many cows (thus the out degree is high) to different farms (thus in degree low)

Dataset : summary and ideas

What we got so far from the dataset:

Each temporal network can be represented as: **aggregated** or **temporal network**

- based on networks and epidemic dynamic we can choose one of the representations, or combine them.

In-out degree can be very heterogeneous,
-nodes differences can influence spreading differently.

2 mains ideas:

-study the persistence of links
-study the activation of links though time

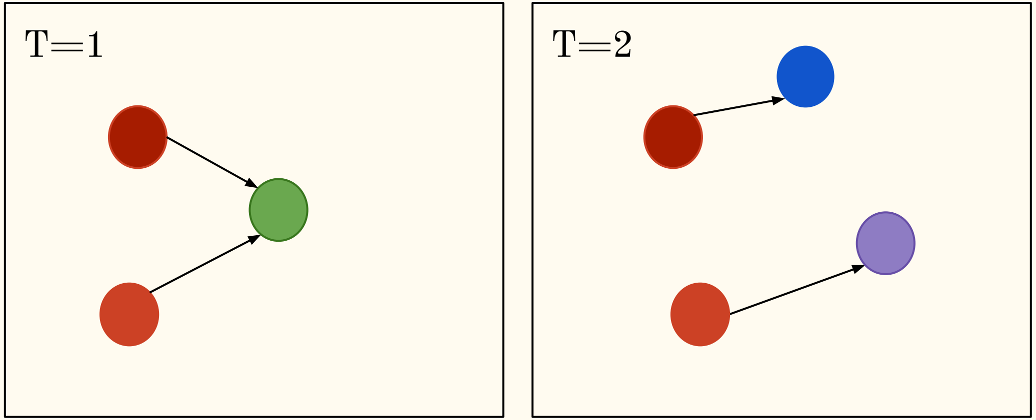
2

Persistence of links in the network

Persistence of links in the network

What: the number of nodes are active during two consecutive intervals.

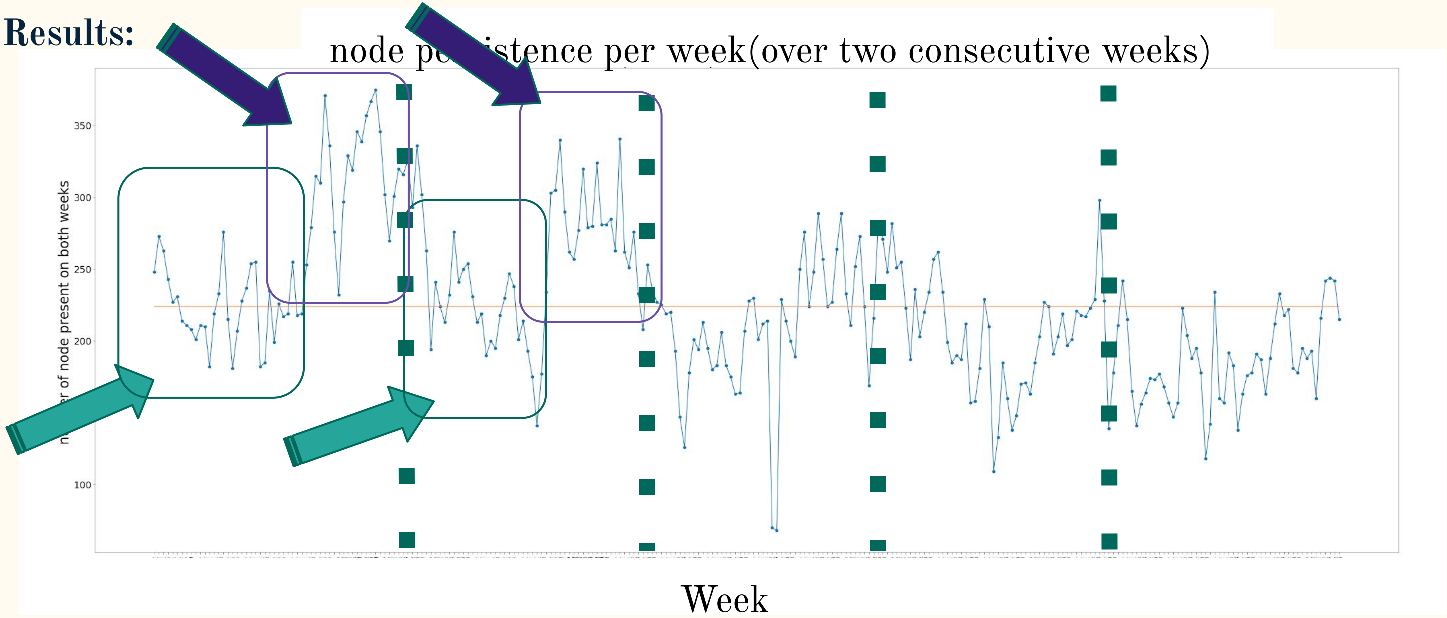
Why: to capture changes in the network to **identify important nodes**.



2 persistant nodes (active both on T1 and T2):

Persistence of links in the network

Results:



Interpretation: beginning : the same pattern, then it seems to always oscillate for the last 3 years.
mean : 223 nodes present over both week on average.

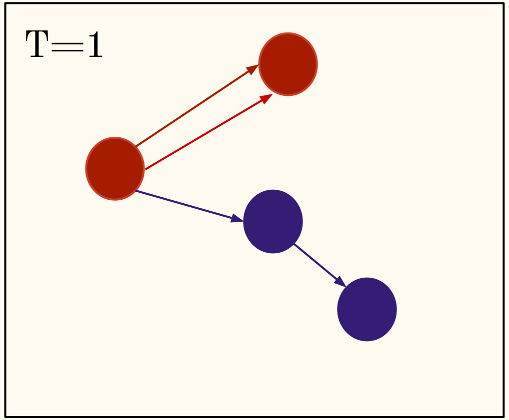
To Do: compare with the activity , see if correlation with spreading.

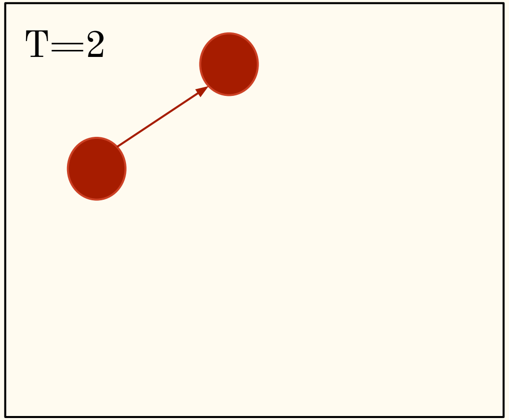
3 Link activation

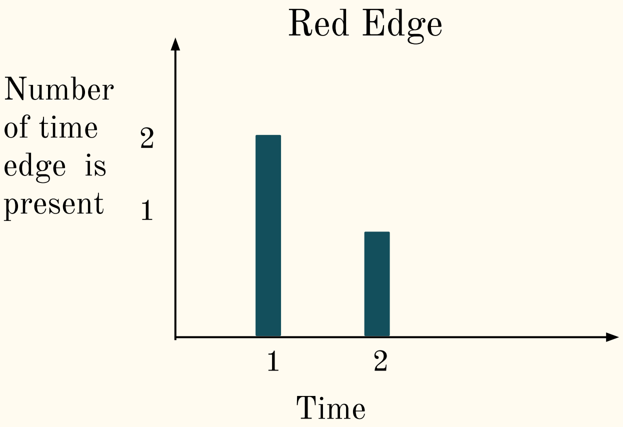
Link activation

What: number of time-steps an edge is activated during the time window.
For each link we will measure how often it is **active**.

Why: -to study the **activation** of links
-to know if the activation of the all links is the same or different.
-to use those properties to better simulate the network and/or the spreading.

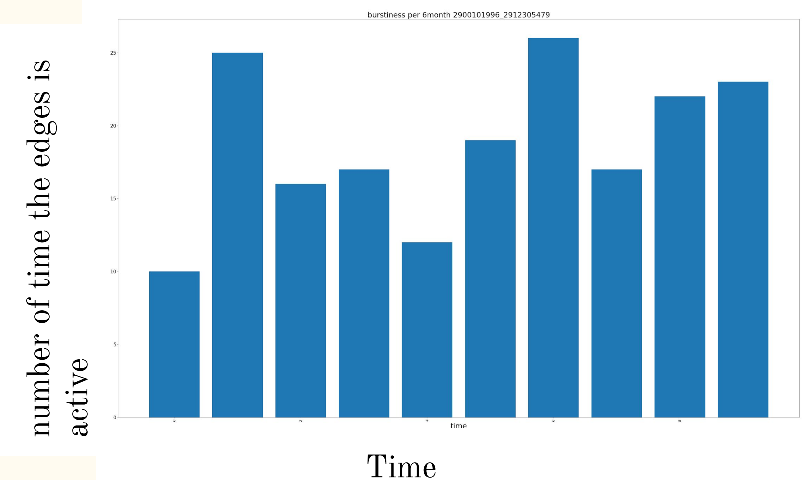


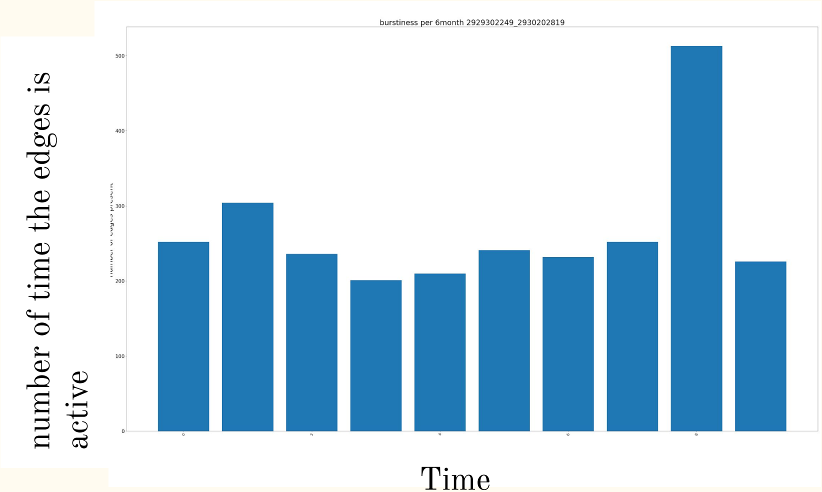




Link activation

Results for two different edges:





Interpretation: really heterogeneous: presence max = 25 compare to presence max = 400.

To Do: -see if those can be correlated with the type of nodes.
-can we simplify the network to weight the links to their availability
-see how and if epidemic depends on availability [5, 6]

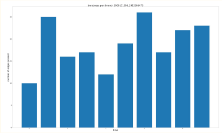
[5] R. Lambiotte, L.Tabourier, J.-C. Delvenne. "Burstiness and spreading on temporal networks", arxiv (2017)
[6] D. Grebenkov, L. Tupikina "Heterogeneous CTRW" PRE, (2018)

4 Work in progress

Work in progress:

Use mathematical framework for the data

Estimated link availability distribution $\Psi_{ij}(t)$



$+$ calculs $\longrightarrow$ $\Psi_{ij}(t)$

Mathematical framework using matrix $\Psi(t)$ from elements $\Psi_{ij}(t)$, $r_i(t)$

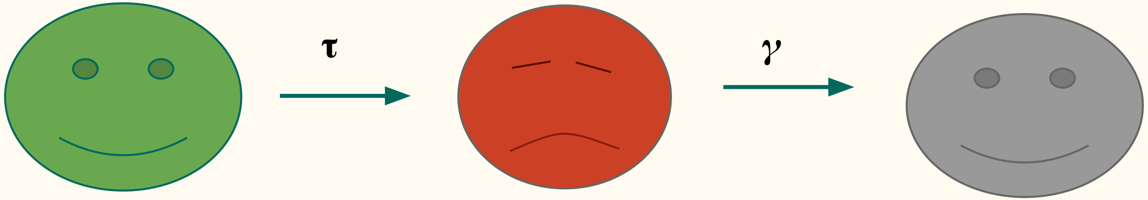
1	0	0	0
0	1	$\Psi_{32}(t)$	$\Psi_{42}(t)$
$\Psi_{13}(t)$	0	1	0
$\Psi_{14}(t)$	0	0	1

Estimation of **probability of infection** to arrive at point j at time t

$$P_{ij}(t) = \psi_{ij}(t) \int_t^{\infty} r_i(t')dt'$$

[5] R. Lambiotte, L.Tabourier, J.-C. Delvenne. “Burstiness and spreading on temporal networks”, European Physical Journal (2013)
[6] D. Grebenkov, L. Tupikina “Heterogeneous Continuous Time Random Walk”. PRE, (2018)

SIR epidemiological model



S susceptible — can become — **I** Infected — then should — **R** Recover

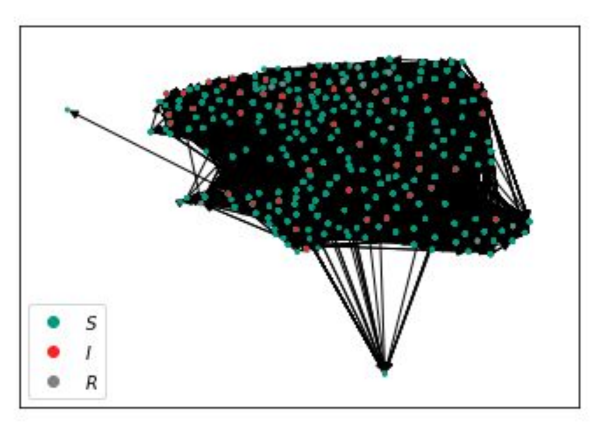
γ= recovery rate (per node)
$\boldsymbol{\tau}$= transmission rate (per edge)

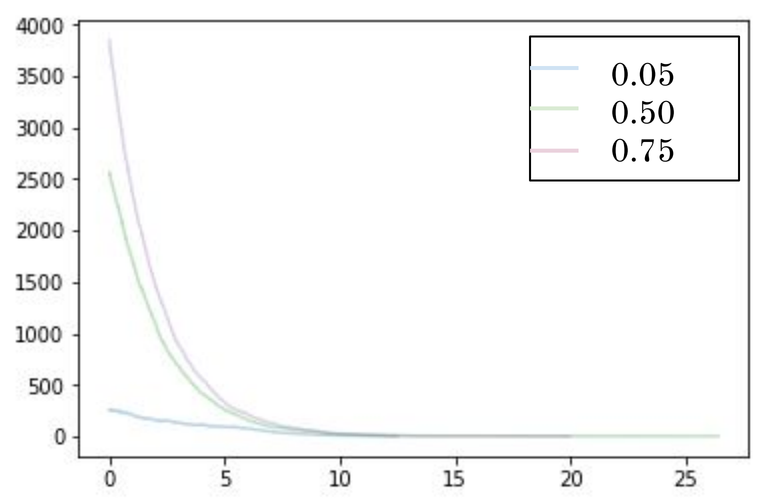
Goal:
-to identify epidemics properties on temporal networks
-to use the persistence measures in order to analyse spreading

SIR epidemiological on Dataset

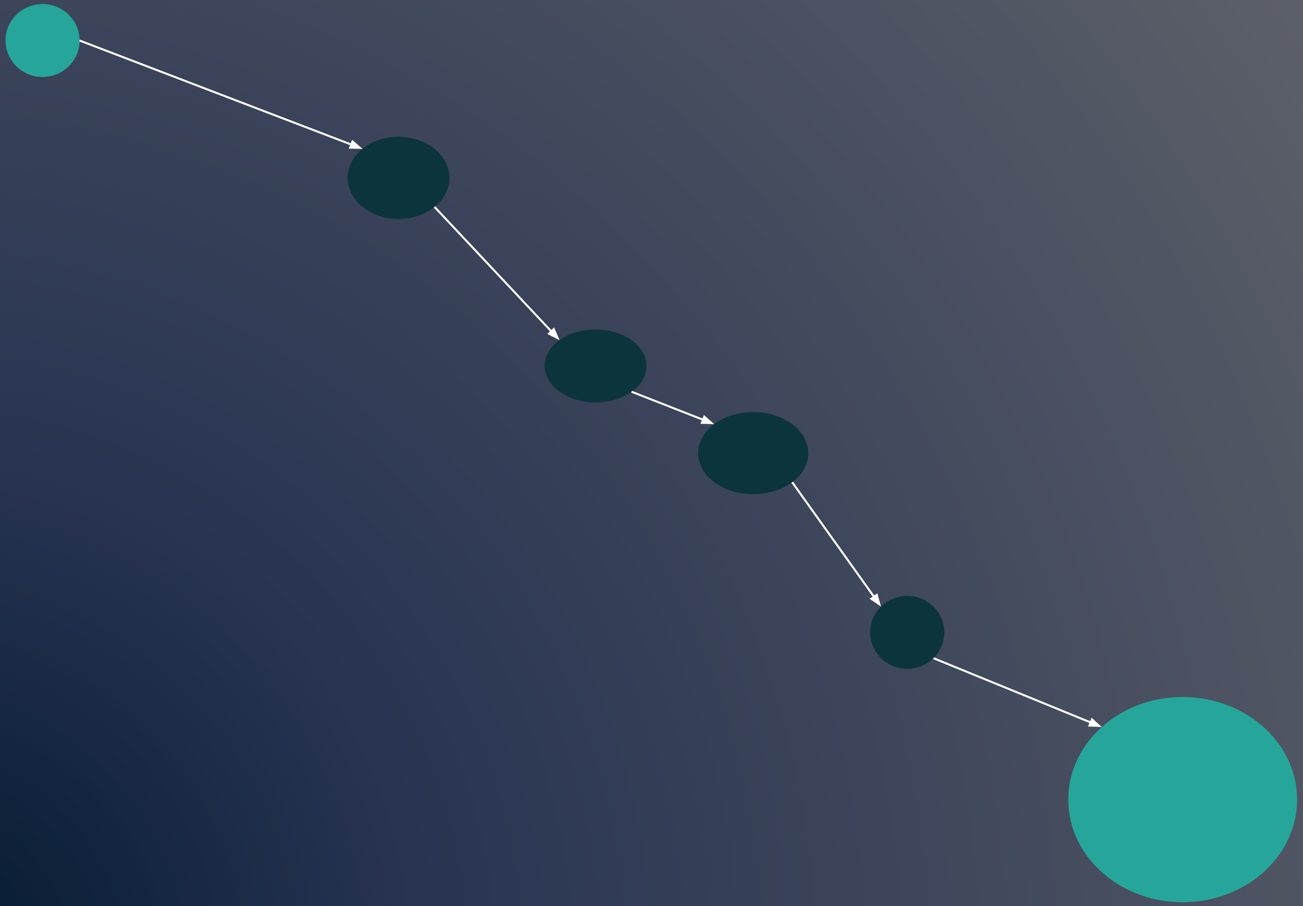
Test on the dataset:
3 differents initial infected fraction (ρ):
-0.05 (very few, more realistic)
-0.50 (average)
-0.75 (high)

ρ=0.05 T=6



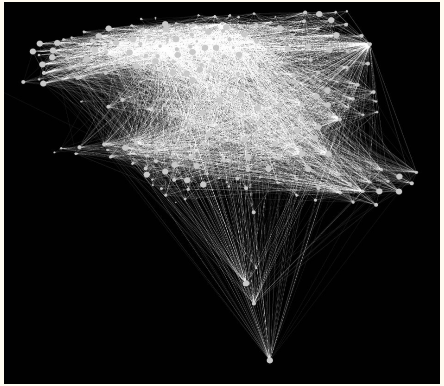


In process:
to choose nodes with a high out-degree instead of random nodes in order to characterize the drastic effects of spreading



Conclusions

Conclusions



We know that people/animals can become sick and infected the whole network.
it is important to better understand epidemic spreading to tackle those problem.

>**Done during the internship**:

- understand how temporal structure of the network, to better tackle the epidemic problem
- have some preliminary calculation that are available on github.

> **work in progress and future work**:

> directed accessibility graph to analyse spreading (persistence, burstiness of links)
> analysis of spreading using analytical framework of heterogeneous random walks


many thanks
to
Supervisor
Dataset
Liubov Tupikina
CRI
Elizabetha Vergu