A SIR model approach to study a Psychometric Network

A project by Irene Silvestro for the Digital Epidemiology exam

Index:

1) Introduction

What is a Psychometric Network?

- Method to represent **the partial correlation** between **nodes** that are **psychopathological symptoms**

- **Borsboom** & **Epskamp (2010):**
  - **Cross-sectional networks** (tipically trasversal, multiple subjects)
  - **Dynamical networks** (tipically longitudinal and within subject)

- **Personalized network**:
  - **“Temporal” network**
  - **Contemporaeneous network**

Case of study: a personalizad ‘dynamical’ network

- Reference Paper: ***Intraindividual Dynamic Network Analysis– Implications for Clinical Assessment*** (David et al., 2017)
- **Partecipant:** 44-year-old woman diagnosed with **major depressive disorder, persistent depressive disorder (dysthymia), and social anxiety disorder**, assessed using the Structured Clinical Interview for DSM-IV (SCID-IV)
- **Symptoms:**
  - derived from 22 items of the **Individualized Daily Questionnaire (IDQ)**, selected from five main scales: Depression, Anxiety, Mixed, Anhedonia, and Positive Affect
  - 90 daily questionnaires over a period of 122 days
- **Data Analysis:**
  - **multiple linear regression** to estime **lag-1 partial correlations** between symptoms (partial correlations between symptoms at time (t-1) and time (t), controlling for all other symptoms)
  - regularized using **LASSO**
  - **directed graph**, with **19 nodes** (sympthoms)

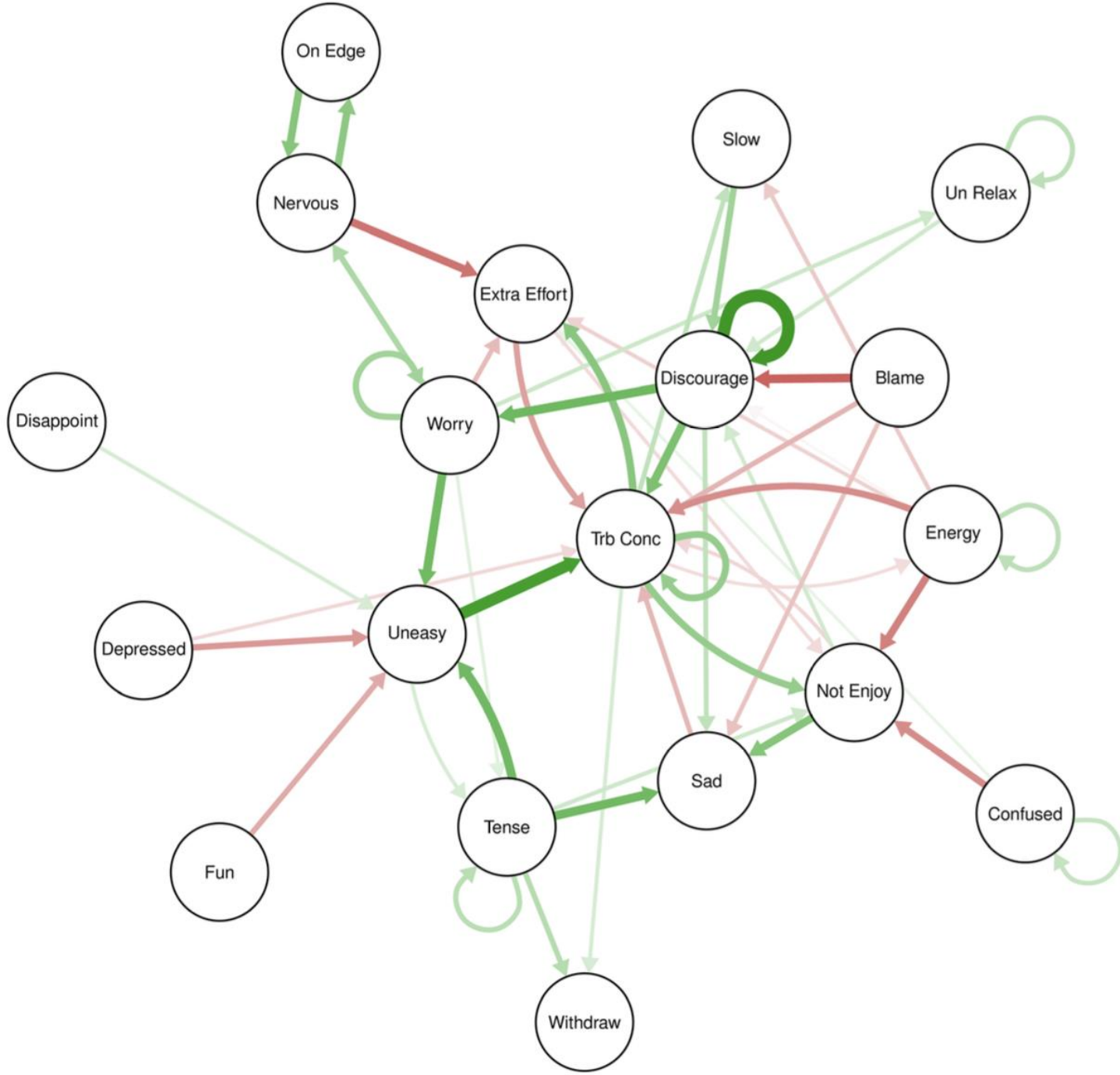


Fig. 1 Regularized Lag-1 Partial Correlation Network

My project

- Constructed sets of **differential equations** inspired by the **compartmental SIR model**, making hypotheses about the dynamics
- The **lag-1 correlation matrix** was used as a **proxy for contagion probability** between nodes, as suggested in the reference paper
- Contagion probability is **heterogeneous** and incorporated into **contact matrices**
- Simulated symptom dynamics by activating different types of nodes according to **various centrality metrics**
- Introducted:
  - **Random node activation** every 10 time steps
  - **Node removal** every 300 time steps, simulating the potential effects of **psychotherapeutic interventions**

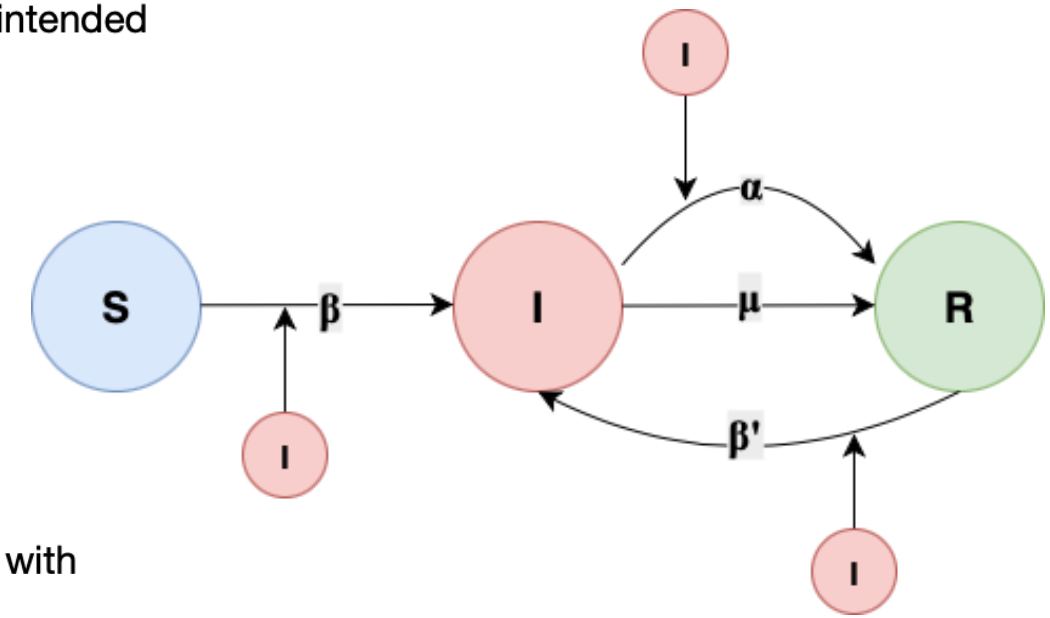
- Explored how the **order of node removal** affects the number of activated nodes over time

Note: These simulations are **for illustrative modeling purposes only**. The assumptions made do **not reflect real-life conditions**.

2) Compartimental Model: studying parameters, endemic state and R0 definition

As a first step, deterministic differential equations were used to represent the intended dynamics.

$$\frac{dS}{dt} = -\beta \frac{I}{N} S$$
$$\frac{dI}{dt} = \beta \frac{I}{N} S - \mu I + \beta_1 \frac{I}{N} R - \alpha \frac{I}{N} I$$
$$\frac{dR}{dt} = \mu I - \beta_1 \frac{I}{N} R + \alpha \frac{I}{N} I$$



> **Note:** A **contagion from *I* → *R*** is also included, representing contact with another infected node.
> This mechanism is intended to capture the effect of **negative partial correlations** between symptoms, which will be discussed in later sections.

For this set of equations, using the **early-stage approximation** for the I equation, we obtain:

$$R_0 = \frac{\beta}{\mu}$$

It can also be shown that the **endemic equilibrium state** (i.e., the equilibrium different from the disease-free state) has an I^* value of:

$$I^* = \frac{(\beta_1 - \mu)}{\alpha + \beta_1} N$$

Finding R_0

$$\frac{dI}{dt} = -\mu I - \alpha \frac{I}{N} I + \beta \frac{I}{N} S + \beta' \frac{I}{N} R$$

Early stage assumption:

$I \approx 0,\ R \approx 0,\ S \approx N$

$$\frac{dI}{dt} \approx -\mu I - \alpha \frac{I}{N} I + \beta \frac{I}{N} N + \beta' \frac{I}{N} R$$

Neglect higher-order terms:

$$\frac{dI}{dt} \approx -\mu I + \beta I$$

$$\approx \mu \left(\frac{\beta}{\mu} - 1 \right) I \longrightarrow R_0 \sim \frac{\beta}{\mu}$$

Finding the Endemic State

Equilibrium condition:

$$\frac{dS}{dt} = \frac{dI}{dt} = \frac{dR}{dt} = 0$$

Set of equations:

$$\begin{aligned}
&\text{(i)} \quad 0 = -\beta \frac{I}{N} S \\
&\text{(ii)} \quad 0 = -\mu I - \alpha \frac{I}{N} I + \beta \frac{I}{N} S + \beta' \frac{I}{N} R \\
&\text{(iii)} \quad 0 = -\beta' \frac{I}{N} R + \mu I + \alpha \frac{I}{N} I
\end{aligned}$$

Consider equation (ii) for I:

$$I \left(-\mu - \alpha \frac{I}{N} + \beta \frac{S}{N} + \beta' \frac{R}{N} \right) = 0$$

Excluding the Disease-Free Equilibrium $I = 0$ and using $R = N - S - I$:

$$-\mu - \alpha \frac{I}{N} + \beta \frac{S}{N} + \beta' \left(\frac{N - I - S}{N} \right) = 0$$

$$-\mu - \alpha \frac{I}{N} + \beta \frac{S}{N} + \beta' - \beta' \frac{I}{N} - \beta' \frac{S}{N} = 0$$

Grouping terms:

$$\frac{\beta - \beta'}{N} S = \mu + \alpha \frac{I}{N} - \beta' - \beta' \frac{I}{N}$$

Solution for S^*

$$S^* = \frac{\mu N + \alpha I^* - N\beta' + \beta' I^*}{\beta - \beta'}$$

$$S^* = \frac{(\mu - \beta')N + (\alpha + \beta')I^*}{\beta - \beta'}$$

Substitute in (i):

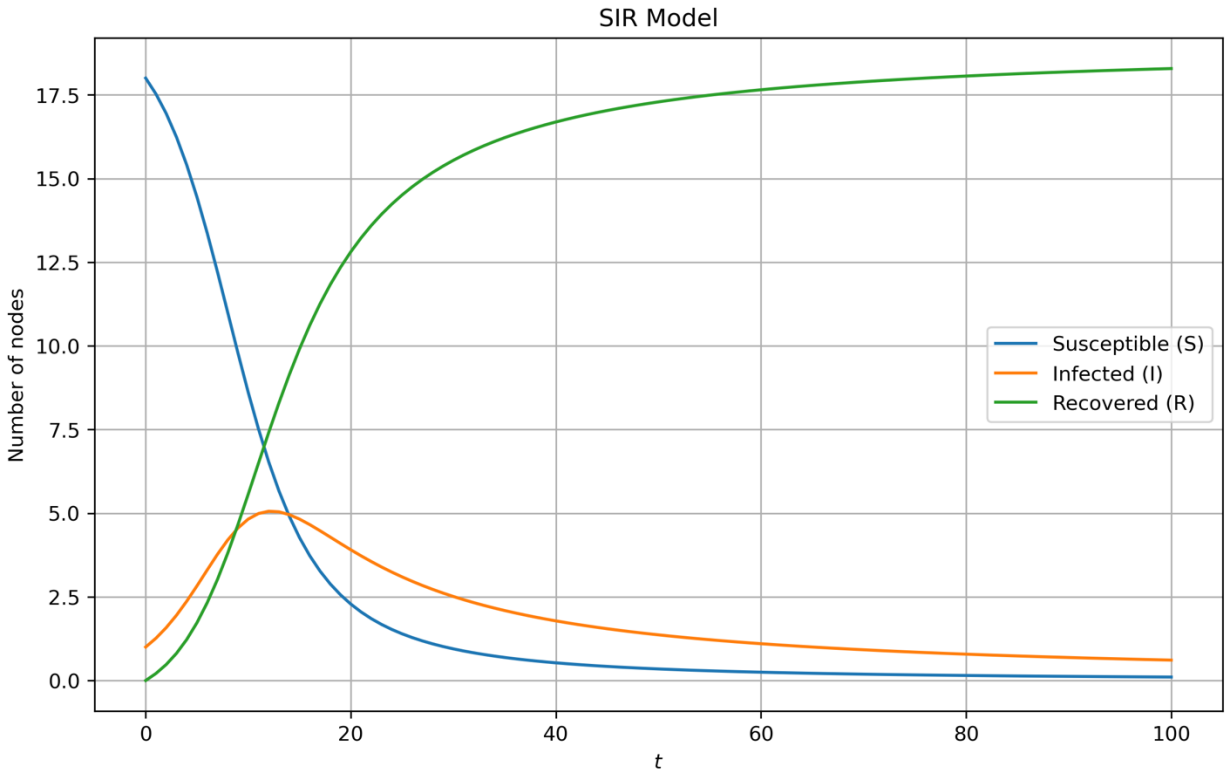
$$-\beta \frac{I}{N} S = 0$$

$$\frac{(\mu - \beta')N + (\alpha + \beta')I^*}{\beta - \beta'} = 0$$

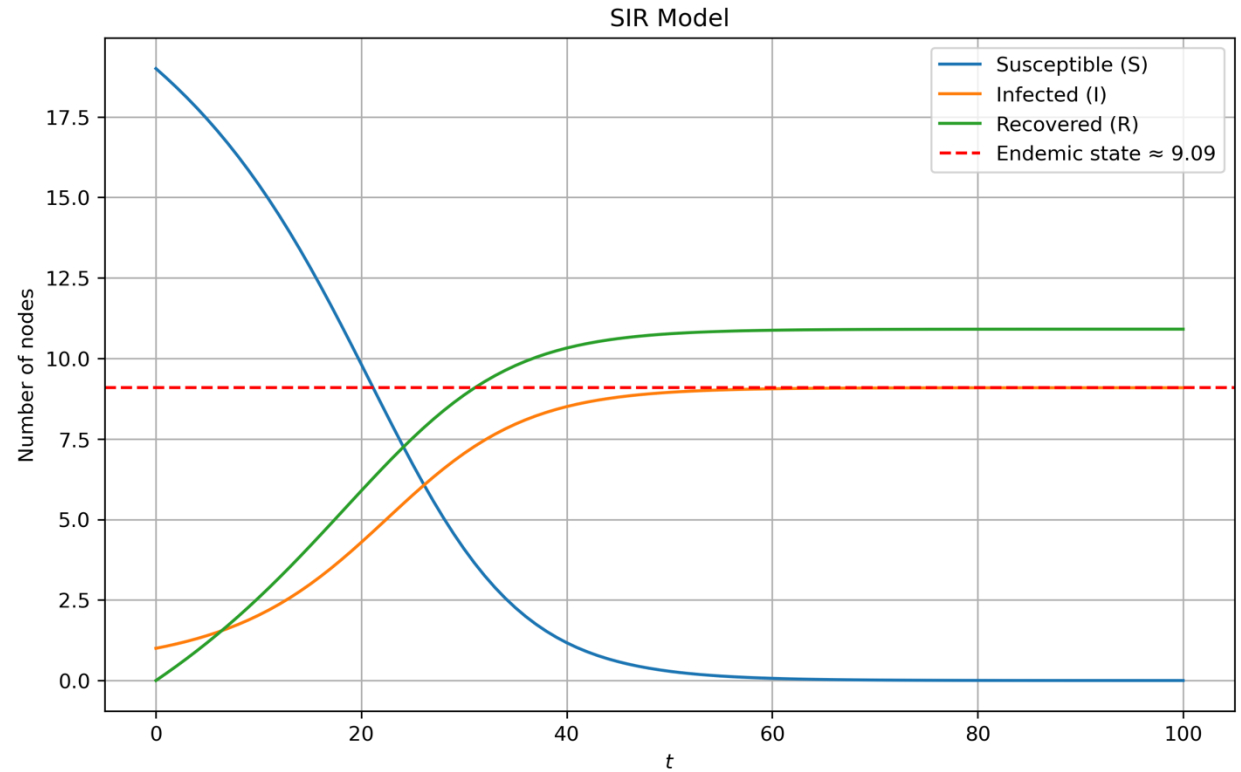
Solution for I^*

$$I^* = \left(\frac{\beta' - \mu}{\alpha + \beta'} \right) N$$

$$\mu = 0.2$$
$$\alpha = 0.2$$
$$\beta = 0.5$$
$$\beta_1 = 0.2$$



$$\mu = 0.2$$
$$\alpha = 0.4$$
$$\beta = 0.3$$
$$\beta_1 = 0.7$$



3) Defining Contact Matrices using the Psychometric Adjacency Matrix

Appendix

Table 2

Item	Sad	Depressed	Discourage	Disappt	Blame	Nervous	Tense	Uneasy	Un relax	On edge	Worry	Trb Conc	Confused	Not enjoy	Withdraw	Extra Effort	Slow	Fun	Energy
Sad	0	0	0	0	0	0	0	0	0	0	0	−0.099	0	0	0	0	0	0	0
Depressed	0	0	0	0	0	0	0	−0.144	0	0	0	−0.054	0	0	0	0	0	0	0
Discourage	0.095	0	0.348	0	0	0	0	0	0	0	0.232	0.206	0	0	0	0	0	0	0
Disappt	0	0	0	0	0	0	0	0.06	0	0	0	0	0	0	0	0	0	0	0
Blame	−0.084	0	−0.222	0	0	0	0	0	0	0	0	−0.104	0	0	0	0	0	0	0
Nervous	0	0	0	0	0	0	0	0	0	0.179	0	0	0	0	0	−0.196	0	0	0
Tense	0.227	0	0	0	0	0	0.107	0.227	0	0	0	0	0	0.075	0.113	0	0	0	0
Uneasy	0	0	0	0	0	0	0.059	0	0	0	0	0.322	0	0	0	0	0	0	0
Un relax	0	0	0.071	0	0	0	0	0	0.101	0	0	0	0	0	0	0	0	0	0
On edge	0	0	0	0	0	0.188	0	0	0	0	0	0	0	0	0	0	0	0	0
Worry	0	0	0	0	0	0.127	0.046	0.226	0.076	0	0.143	0	0	0	0	−0.094	0	0	0
Trb Conc	0	0	0	0	0	0	0	0	0	0	0	0.163	0	0.164	0.062	0.184	0.094	0	−0.048
Confused	0	0	0	0	0	0	0	0	0	0	0	0	0.088	−0.161	0	0.035	0	0	0
Not Enjoy	0.191	0	0.068	0	0	0	0	0	0	0	0	−0.057	0	0	0	0	0	0	0
Withdraw	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Extra effort	0	0	0	0	0	0	0	0	0	0	0	−0.132	0	−0.051	0	0	0	0	0
Slow	0	0	0.144	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Fun	0	0	0	0	0	0	0	−0.117	0	0	0	0	0	0	0	0	0	0	0
Energy	0	0	−0.014	0	0	0	0	0	0	0	0	−0.159	0	−0.178	0	−0.063	−0.074	0	0.101

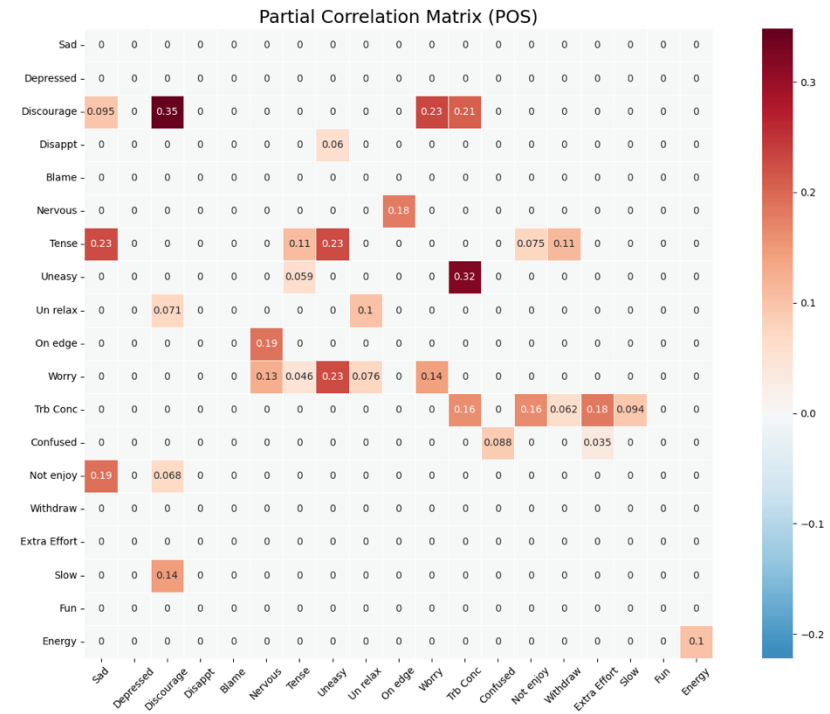
Note. Regularized Lag-1 Partial Correlation Matrix. For each analysis, today's score, the dependent variable, is the top row. Example: The partial correlation of today's (t) Sad score with yesterday's ($t - 1$) Discourage score was 0.095 and with yesterday's Tense score was 0.227, whereas the correlation of today's Tense score with yesterday's Sad score was 0 (after regularization).

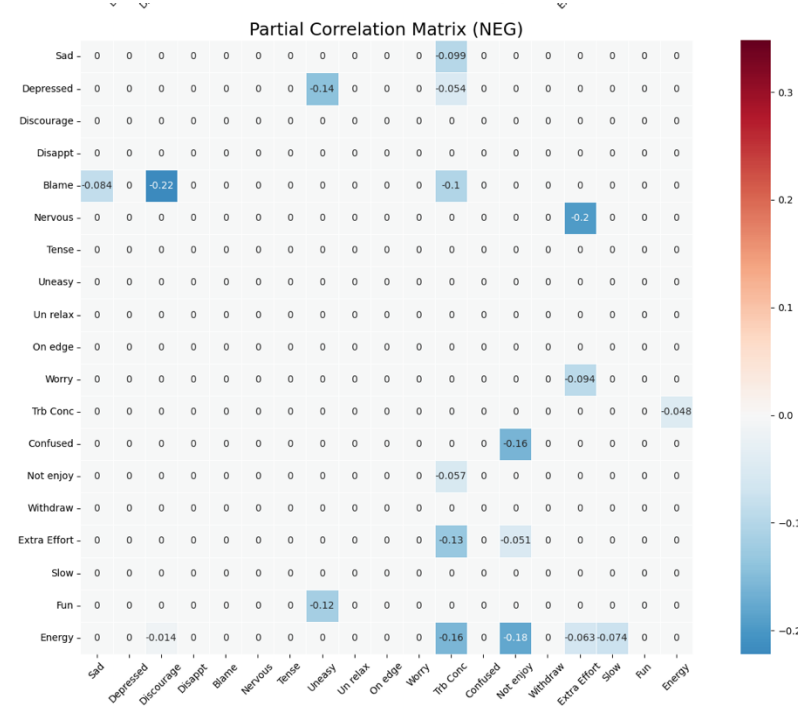
This is the **Regularized Lag-1 Partial Correlation**, from which **two separate contact matrices** were defined

Specifically, the **positive values** were used to construct a matrix of positive contacts, where the weight of each link serves as a proxy for the probability of contagion from **S** → **I** and **R** → **I** (through contact with another infected node). In the model, it was assumed that a recovered node could become infected again.

Additionally, **negative correlations** were used to simulate a “negative contagion”, serving as a proxy for the probability of transition from **I** → **R** (through contact with another infected node) . As explained in the paper, if at time (t-1) a symptom (i) was negatively correlated with another node (j), the latter (if infected at (t-1)) has a higher probability of disappearing (recovering) at time (t).

Example reported from the article : "the arrow pointing from Discouraged to Worry represents the relation between Discouraged at time t– 1 (yesterday) and Worry at time t (today), controlling for all other symptoms, including Worry yesterday."





4) Centrality Metrics

I computed some centrality mesuremets, to see the most central nodes that may impact the dynamics and to define lists of nodes ordinated by this centrality mesurements.

OutDegree

$$\text{Outdegree}_i^{\text{pos}} = \sum_j |C_{ij}| \qquad \text{Outdegree}_i^{\text{neg}} = \sum_j |C_{ij}^{\text{neg}}|$$

InDegree

$$\text{Indegree}_i^{\text{pos}} = \sum_j |C_{ji}| \qquad \text{Indegree}_i^{\text{neg}} = \sum_j |C_{ji}^{\text{neg}}|$$

Betweeness

$$c_B(v) = \sum_{s,t \in V} \frac{\sigma(s,t|v)}{\sigma(s,t)}$$

where V is the set of nodes, $\sigma(s,t)$ is the number of shortest (s,t)-paths, and $\sigma(s,t|v)$ is the number of those paths passing through some node v other than s,t. If $s = t$, $\sigma(s,t) = 1$, and if $v \in s,t$, $\sigma(s,t|v) = 0$ [2].

ClusteringCoefficient

$$c_u = \frac{2T(u)}{deg(u)(deg(u)-1)},$$

where $T(u)$ is the number of triangles through node u and $deg(u)$ is the degree of u.

5) Stochastic Compartimental Model with Contact Matrices: studying Initialization of Infected Nodes and R0 Definition

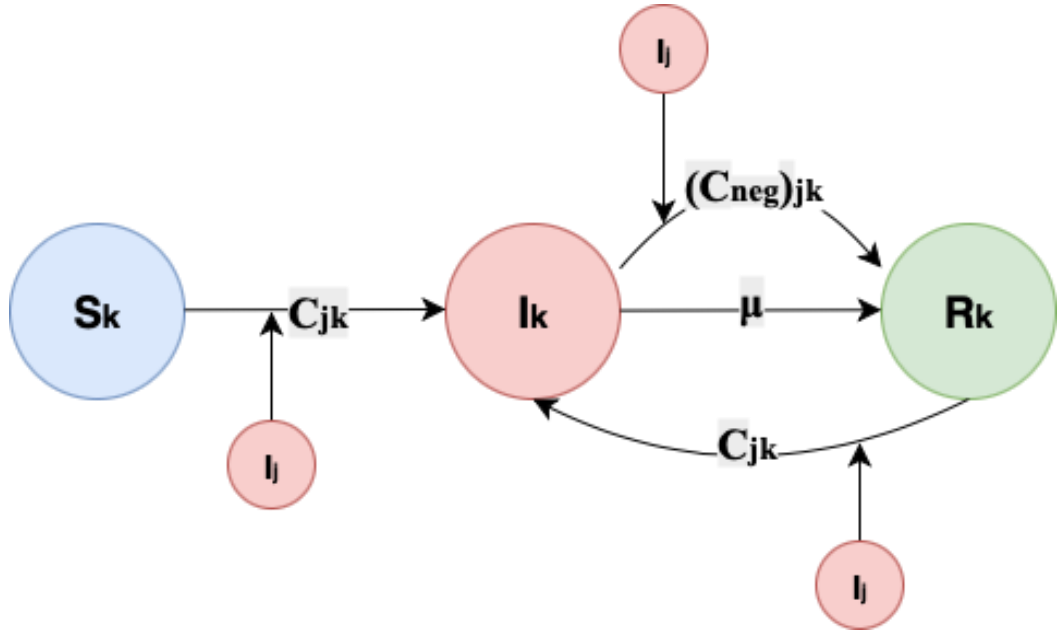
In this section I define the **set of differential equation** that define the dynamics on the graph.

The **stochastic dynamics for each node (k) can be described as follows:**

$$\frac{dS_k}{dt} = -\beta\, S_k \sum_{j=1}^{n_{\text{nodes}}} \frac{C_{jk} I_j}{N_k}$$

$$\frac{dI_k}{dt} = \beta\, S_k \sum_{j=1}^{n_{\text{nodes}}} \frac{C_{jk} I_j}{N_k} - \mu\, I_k + \beta\, R_k \sum_{j=1}^{n_{\text{nodes}}} \frac{C_{jk} I_j}{N_k} - I_k \sum_{j=1}^{n_{\text{nodes}}} \beta \frac{(C_{\text{neg jk}}) I_j}{N_k}$$

$$\frac{dR_k}{dt} = \mu\, I_k - \beta\, R_k \sum_{j=1}^{n_{\text{nodes}}} \frac{C_{jk} I_j}{N_k} + I_k \sum_{j=1}^{n_{\text{nodes}}} \beta \frac{(C_{\text{neg}})_{jk} I_j}{N_k}$$

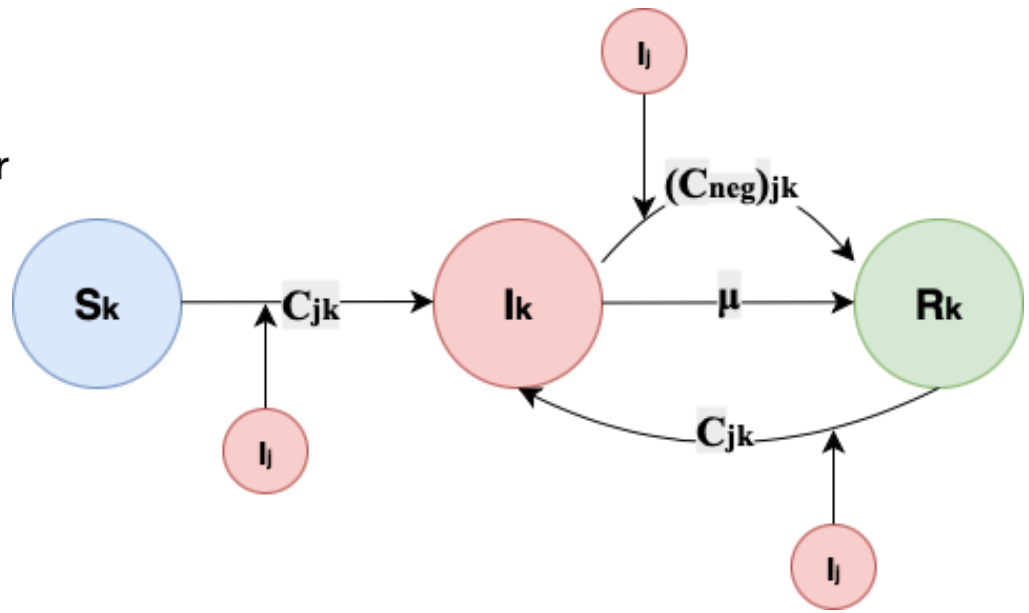


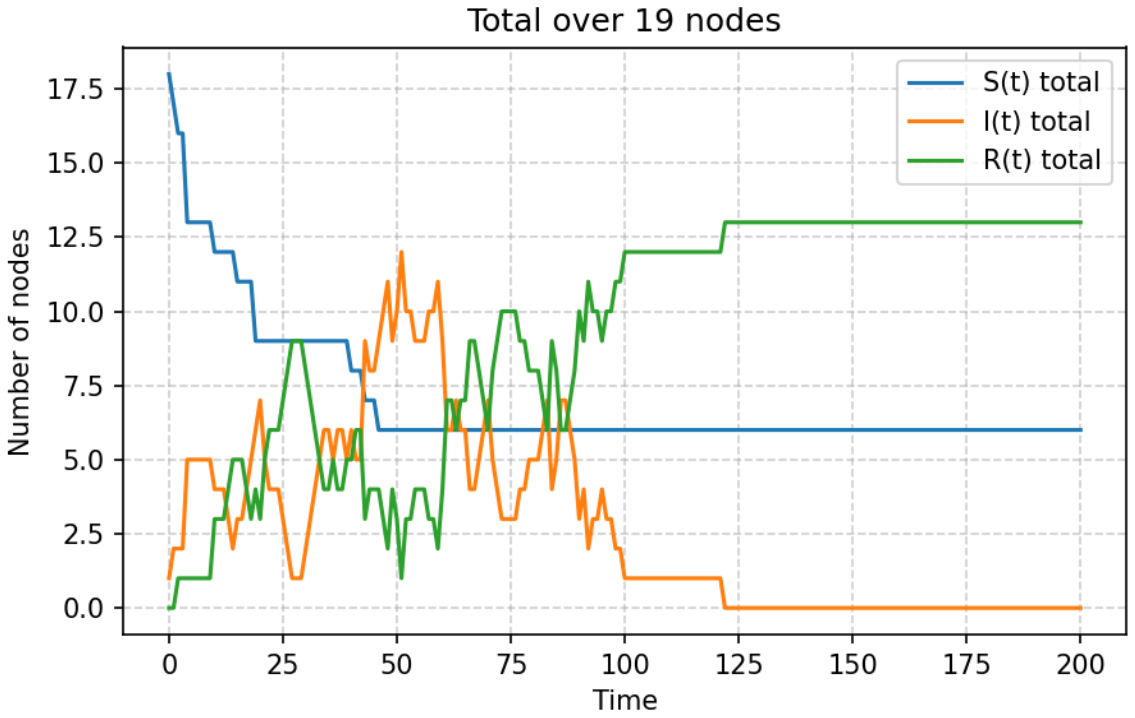
Note that k is the **index of a node** and not a class, as in the classical contact matrix approach. Therefore, each node can only occupy **one compartment at a time**. The dynamics of each node depend both on its current state and on the states of the nodes it can potentially contact. For this reason, in the following simulations, N_k is always set to 1.

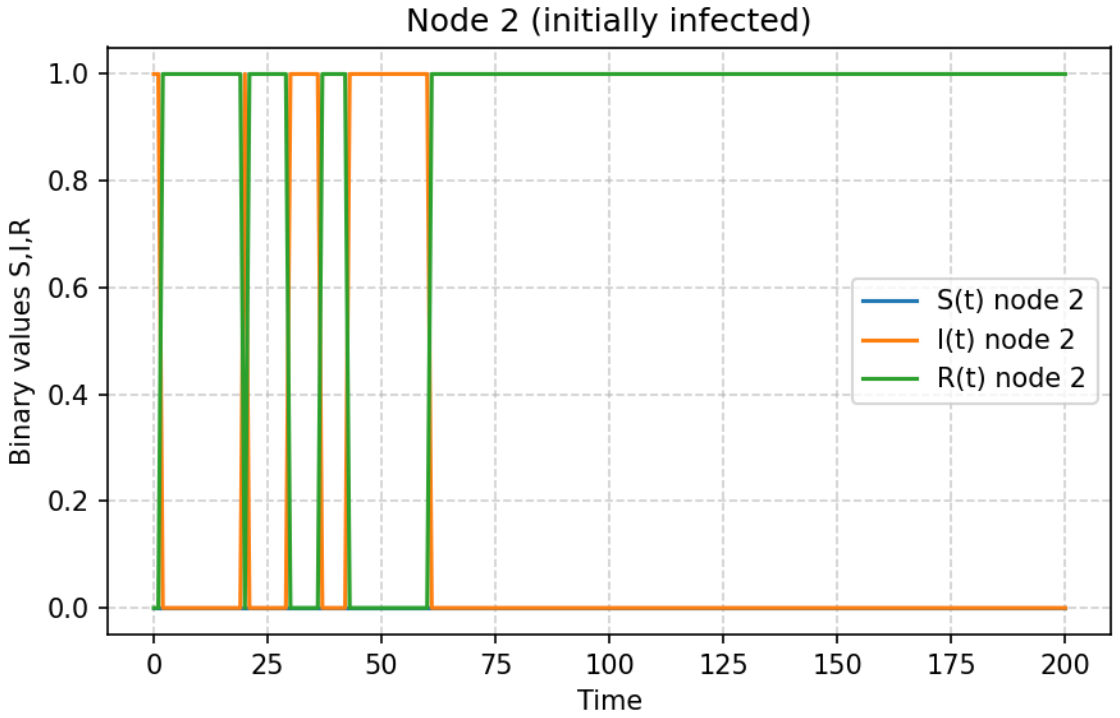
The previously defined matrices C and C_{neg} are used as **contact matrices**, encoding the graph underlying structure, i.e., which nodes can have contact and the probability of contagion in case of contact. This highlights the power of the model: the contagion probability is **heterogeneous** and personalized, thanks to the psychometric network.

For both **positive** (C) and **negative** (C_{neg}) contagion, β is fixed at 1, because the actual probability of contagion is already included in the contact matrix. Moreover, in the code, a minus sign is applied to C_{neg} to correctly account for the negative coefficients in the matrix.

In the dynamics, two pathways to infection are considered: a node can become infected either from the susceptible state ($S \rightarrow I$) or from the recovered state ($R \rightarrow I$) through contact with an infected node. Additionally, beyond the standard recovery process with rate μ, a “negative contagion” mechanism is included, allowing an infected node to transition to recovered ($I \rightarrow R$) if it interacts with another node that can induce negative influence.

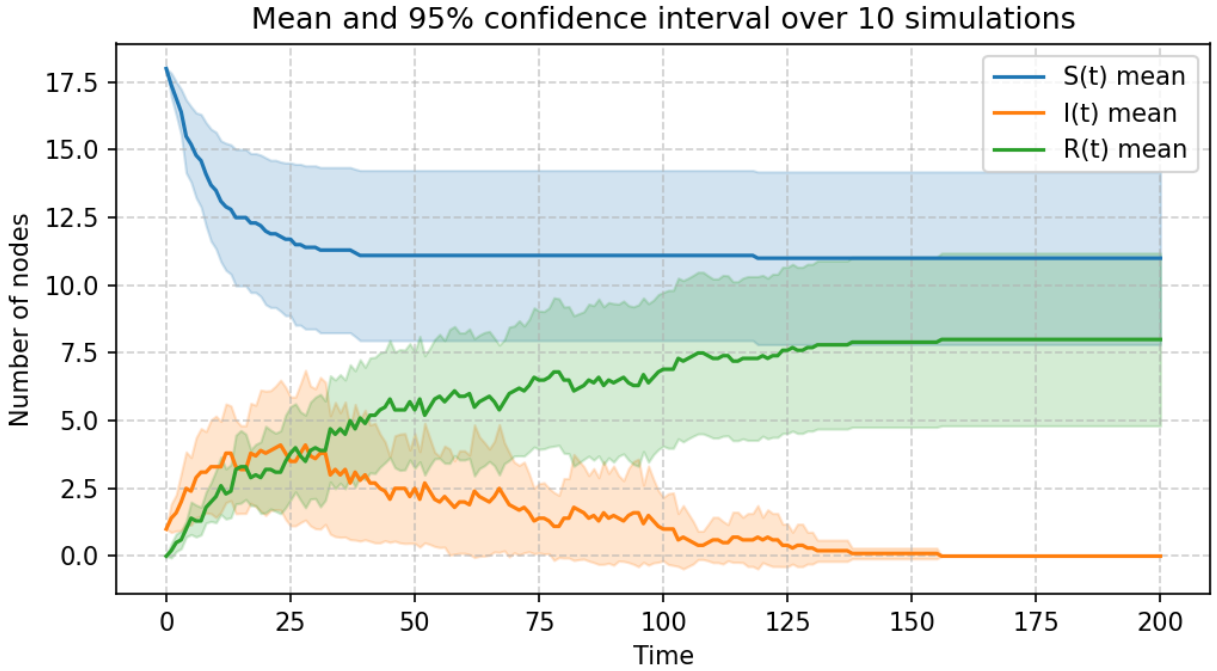


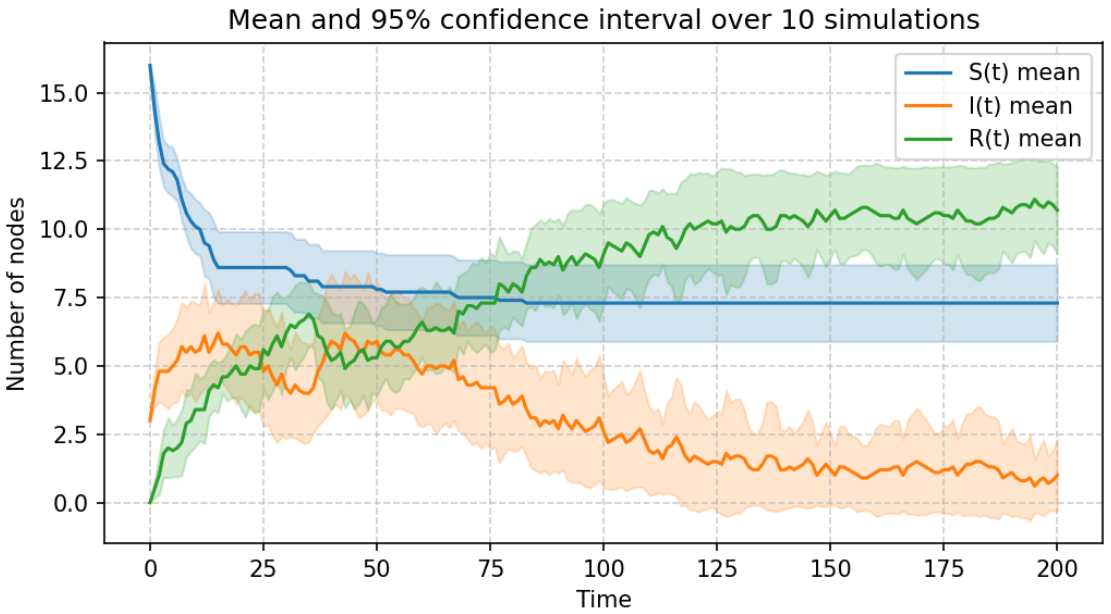




$\mu = 0.1$
$\beta = 1$

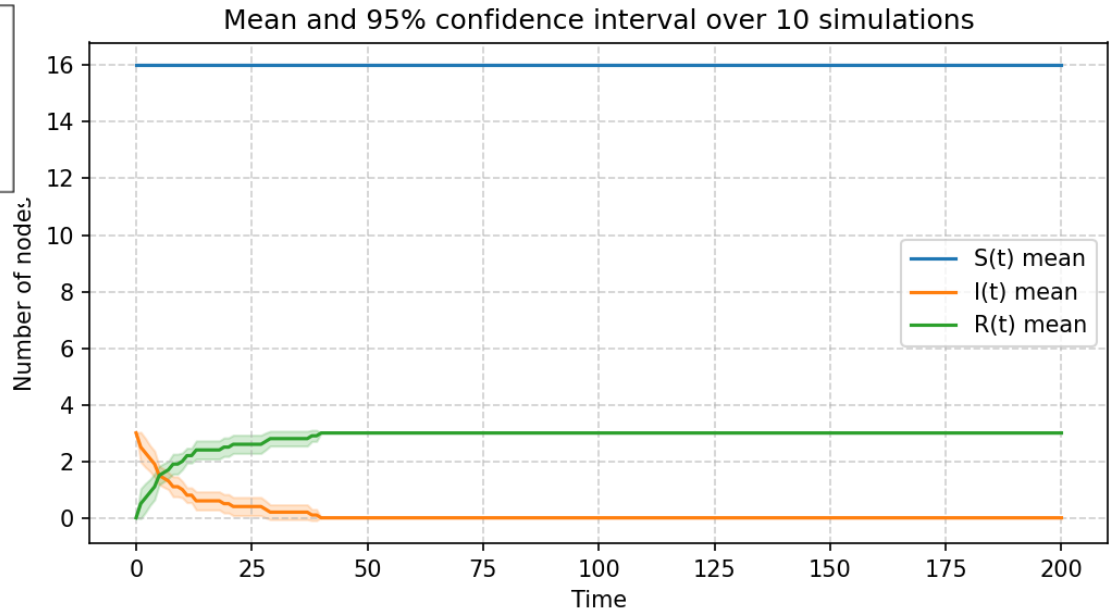
infected = [2]





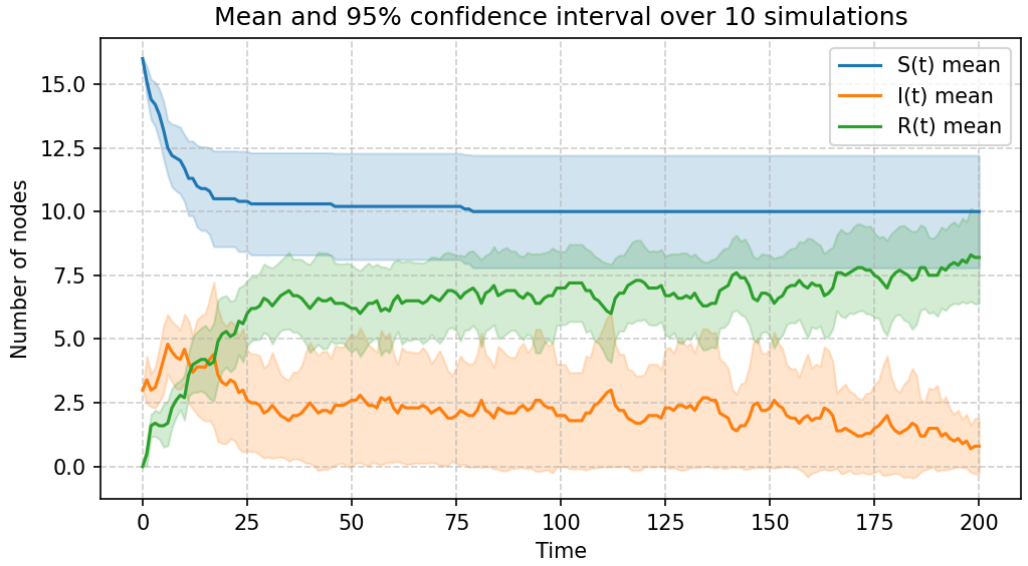
$\mu = 0.1$

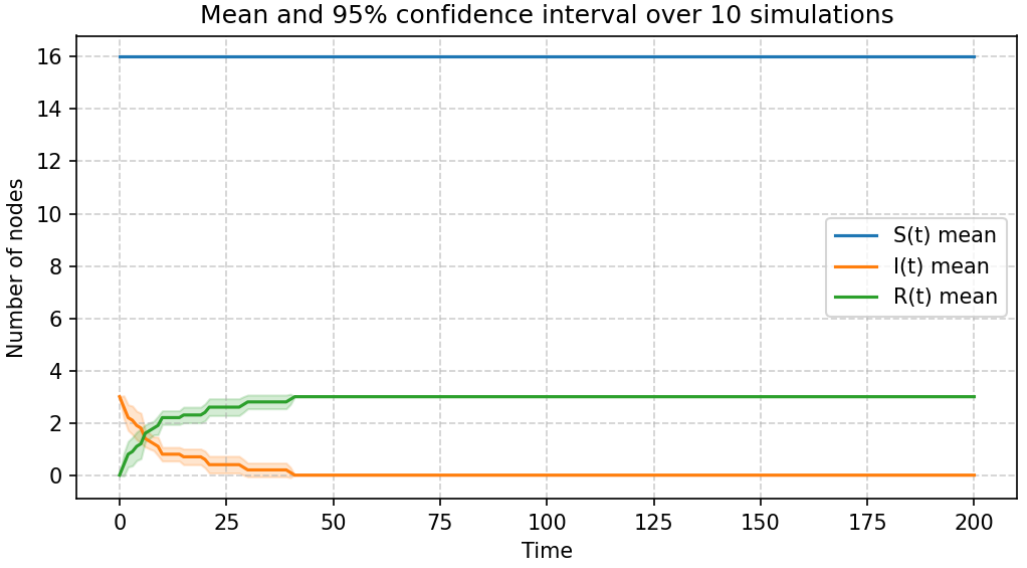
$\beta = 1$



infected = [2,6,11] #nodes with **highest positive outdegree**

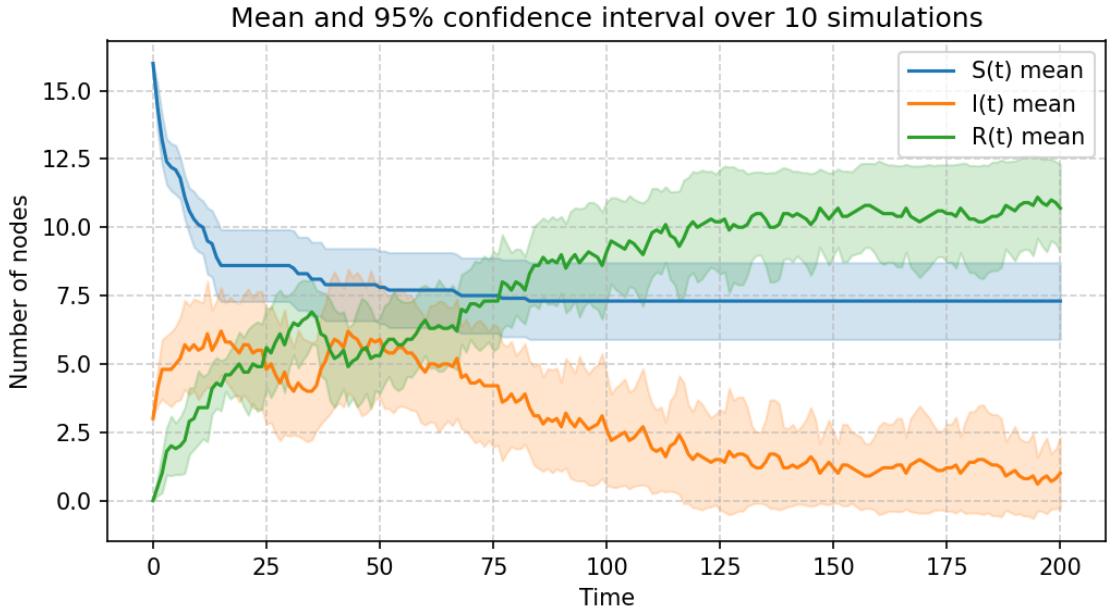
infected = [0,1,4] #nodes with **lower positive outdegree**





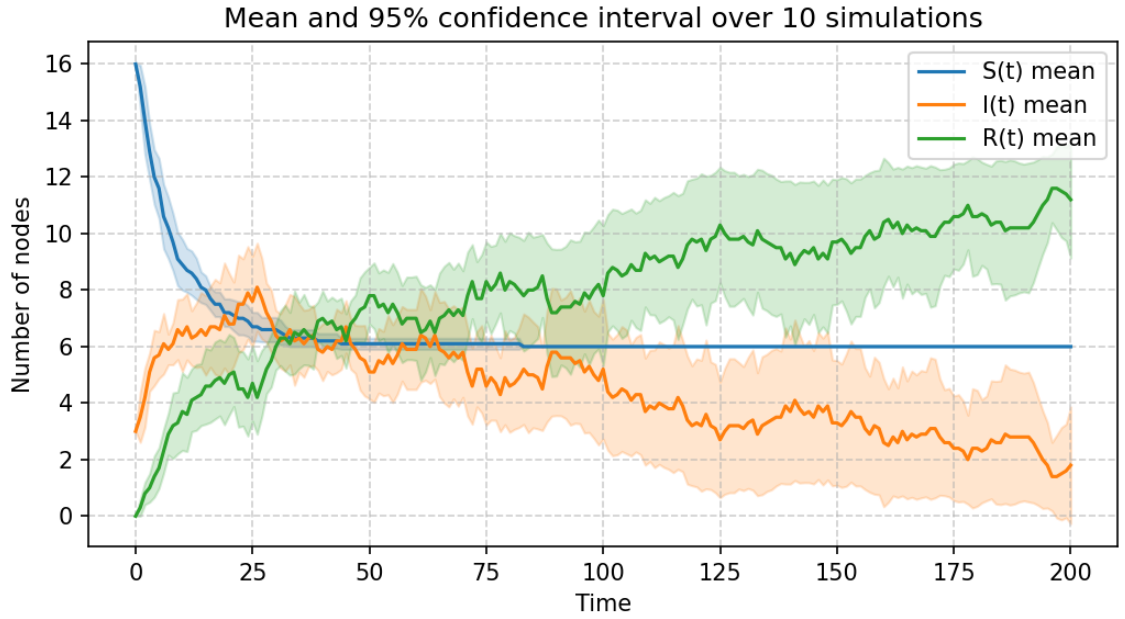
infected = [11,2,0] #nodes with **highest positive indegree**

infected = [18,4,1] #nodes with **highest negative outdegree**



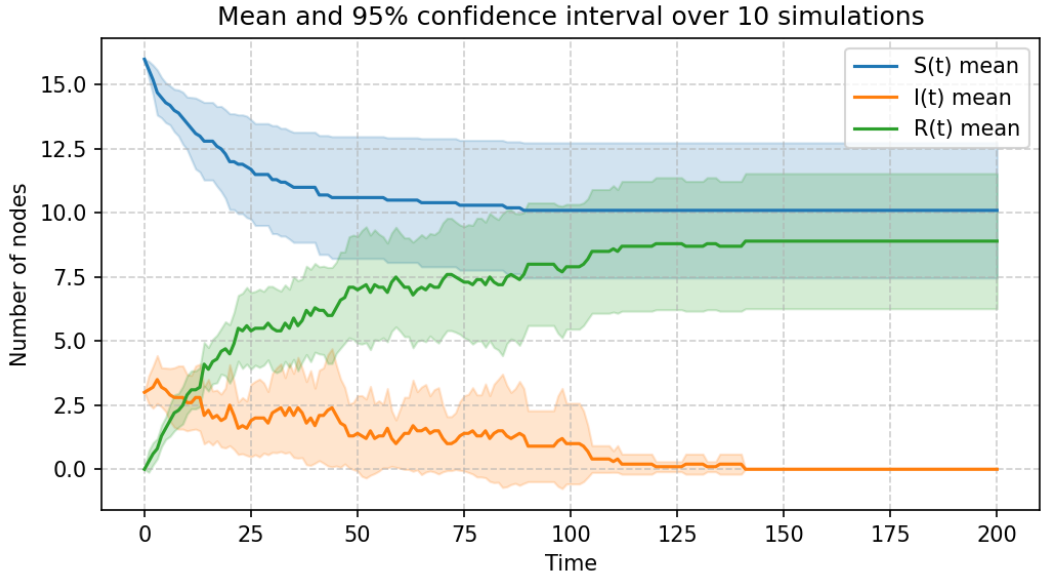
$\mu = 0.1$

$\beta = 1$



infected = [2,6,11] #nodes with **highest positive outdegree**

infected = [11,10,2] #nodes with **highest betweeness**



infected = [8,16,0] #nodes with **highest clustering coefficient**

Finding R_0 with Next Generetion Matrix Aprroach

$$\frac{dI_i}{dt} = -\mu I_i + \beta \sum_{j=1}^{N} C_{ij} \frac{I_j}{N_j} S_i - \beta \sum_{j=1}^{N} C_{ij}^{neg} \frac{I_j}{N_j} I_i + \beta \sum_{j=1}^{N} C_{ij} \frac{I_j}{N_j} R_i \qquad (1)$$

$$\begin{bmatrix} \frac{dI_1}{dt} \\ \frac{dI_2}{dt} \\ \vdots \\ \frac{dI_i}{dt} \\ \vdots \\ \frac{dI_N}{dt} \end{bmatrix} = \begin{bmatrix} +\beta \sum_{j=1}^{N} C_{1j} \frac{I_j}{N_j} S_1 - \beta \sum_{j=1}^{N} C_{1j}^{neg} \frac{I_j}{N_j} I_1 + \beta \sum_{j=1}^{N} C_{1j} \frac{I_j}{N_j} R_1 \\ +\beta \sum_{j=1}^{N} C_{2j} \frac{I_j}{N_j} S_2 - \beta \sum_{j=1}^{N} C_{2j}^{neg} \frac{I_j}{N_j} I_2 + \beta \sum_{j=1}^{N} C_{2j} \frac{I_j}{N_j} R_2 \\ \vdots \\ +\beta \sum_{j=1}^{N} C_{ij} \frac{I_j}{N_j} S_i - \beta \sum_{j=1}^{N} C_{ij}^{neg} \frac{I_j}{N_j} I_i + \beta \sum_{j=1}^{N} C_{ij} \frac{I_j}{N_j} R_i \\ \vdots \\ +\beta \sum_{j=1}^{N} C_{Nj} \frac{I_j}{N_j} S_N - \beta \sum_{j=1}^{N} C_{Nj}^{neg} \frac{I_j}{N_j} I_N + \beta \sum_{j=1}^{N} C_{Nj} \frac{I_j}{N_j} R_N \end{bmatrix} - \begin{bmatrix} \mu I_1 \\ \mu I_2 \\ \vdots \\ \mu I_i \\ \vdots \\ \mu I_N \end{bmatrix}$$

We define the matrices F and V as follows:

$$\begin{bmatrix} \vdots \\ \frac{d\theta_i}{dt} \\ \vdots \end{bmatrix} = \begin{bmatrix} \vdots \\ F_i \\ \vdots \end{bmatrix} - \begin{bmatrix} \vdots \\ V_i \\ \vdots \end{bmatrix}$$

Our inital condition is:

$$(S_k(0), I_k(0), R_k(0)) = (1, 0, 0) \vee (0, 1, 0) \qquad (2)$$

where no node is initialized as recovered, but it can be initialized either as susceptible or infected.

$$F_{ij} = \left.\frac{\partial F_i}{\partial \theta_j}\right|_{t=0}$$

$$F_{ij} = \beta C_{ij} \left.\frac{S_i + R_i}{N_i}\right|_{t=0} + C_{ij}^{neg} \left.\frac{I_i}{N_i}\right|_{t=0}$$

Note that R_i is always equal to zero.

$$F_{ij} = \beta C_{ij} S_i|_{t=0} + C_{ij}^{neg} I_i|_{t=0}$$

$$\mathbf{F} = \begin{bmatrix} \cdots & F_{ij} & \cdots \\ \cdots & \cdots & \cdots \\ \cdots & \cdots & \cdots \end{bmatrix}$$

$$V_{ij} = \left.\frac{\partial V_i}{\partial \theta_j}\right|_{t=0}$$

$$V_{ij}|_{t=0} = \begin{cases} \mu & \text{if } i = j \\ 0 & \text{if } i \neq j \end{cases}$$

$$\mathbf{V} = \begin{bmatrix} \mu & 0 & \cdots \\ 0 & \ddots & 0 \\ \vdots & 0 & \mu \end{bmatrix} = \mu \mathbf{I}$$

$$\begin{aligned} R_0 &\equiv \rho(\mathbf{F}\mathbf{V}^{-1}) \\ &= \rho\left(\mathbf{F}(\mu\mathbf{I})^{-1}\right) \\ &= \frac{1}{\mu}\rho(\mathbf{F}) \end{aligned}$$

where ρ is the largest eigenvalue.

With:

- $\mu = 0.1$
- $\beta = 1$
- infected $= [2]$

The resulting R_0 is:

```
R0: 1.8344481459011046
```

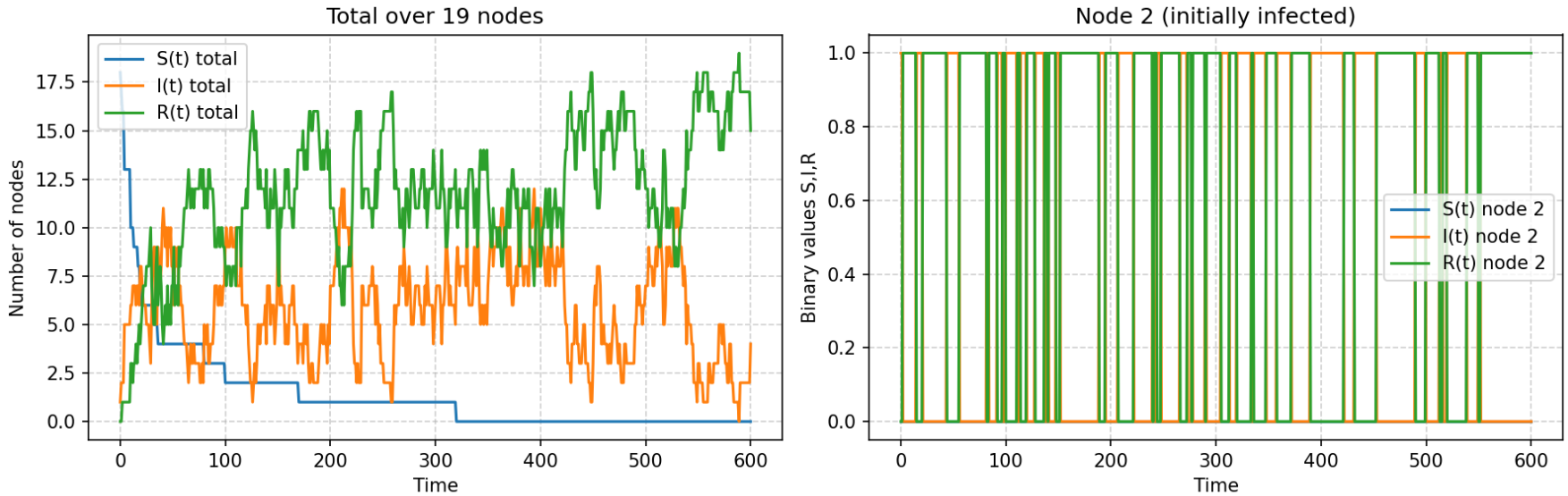
6) Stochastic Compartimental Model with Contact Matrices and Random Infections over time

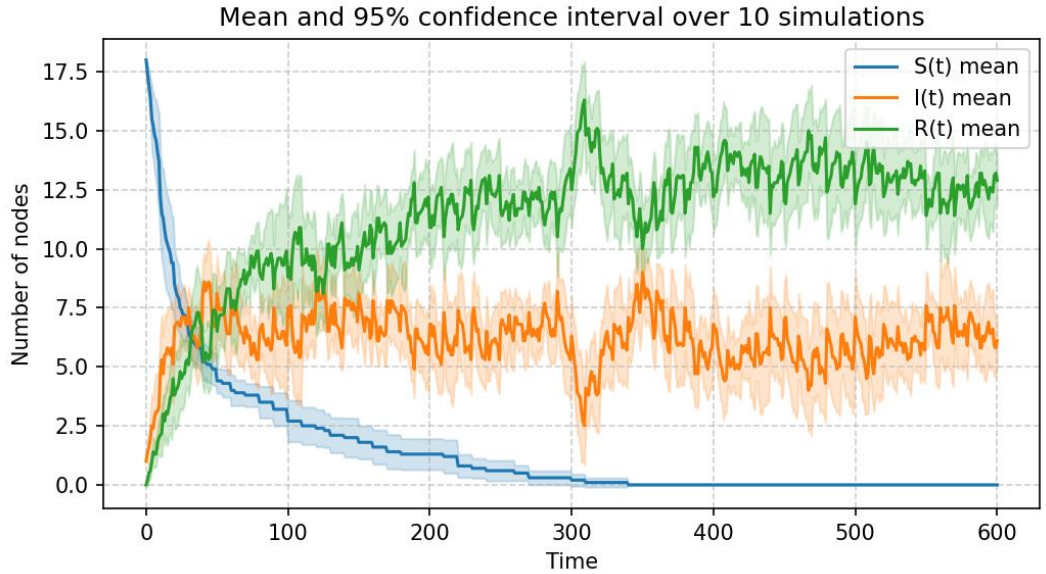
Now, **random infections over time** are implemented.

The aim is to simulate the **emergence of new infections as time progresses**, reflecting the **ongoing dynamics of the individual.** Assuming dt represents one day, two new nodes are introduced as infected every 10 dt (days).

Results for infected = [2]

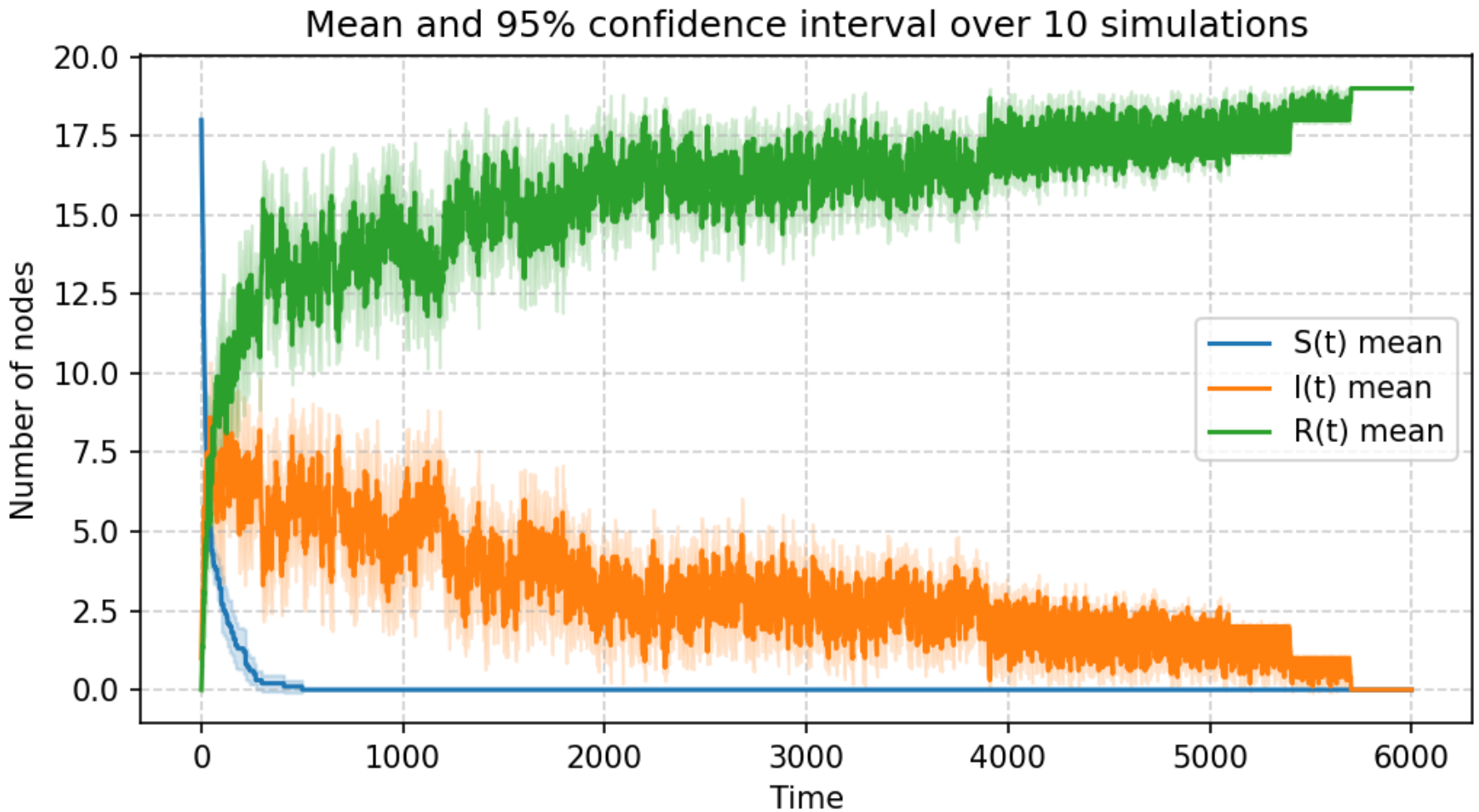
$\mu = 0.1$
$\beta = 1$





7) **Stochastic Compartimental Model with Contact Matrices, Random Infections and Permanent Node Recovery over Time**: exploring Strategies of "Vaccination" through Centrality Metrics

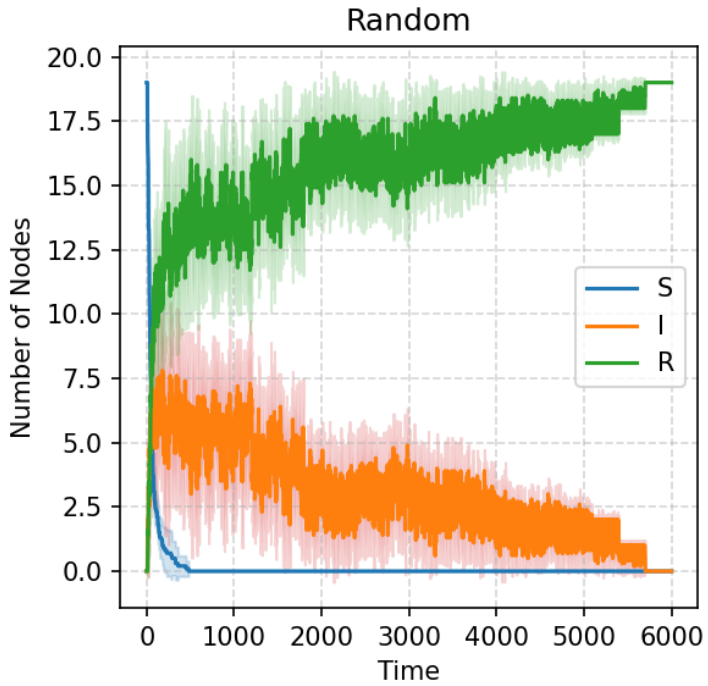
Finally, we introduce **node vaccination**. The idea is to simulate the possibility for an individual to access interventions such as psychotherapy or medication, which can attenuate symptoms. In the model, vaccination is implemented by **permanently setting one node to the recovered state (R) every 300** dt, meaning the node no longer participates in the dynamics.

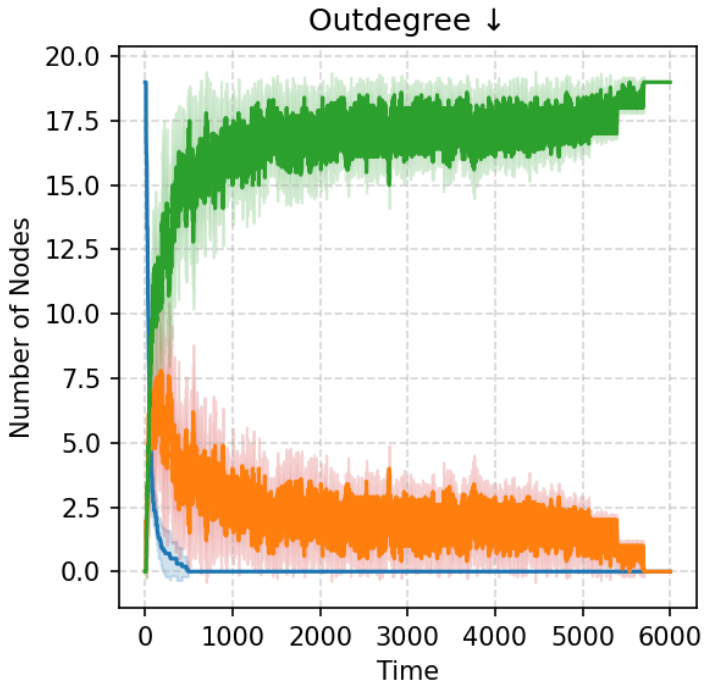


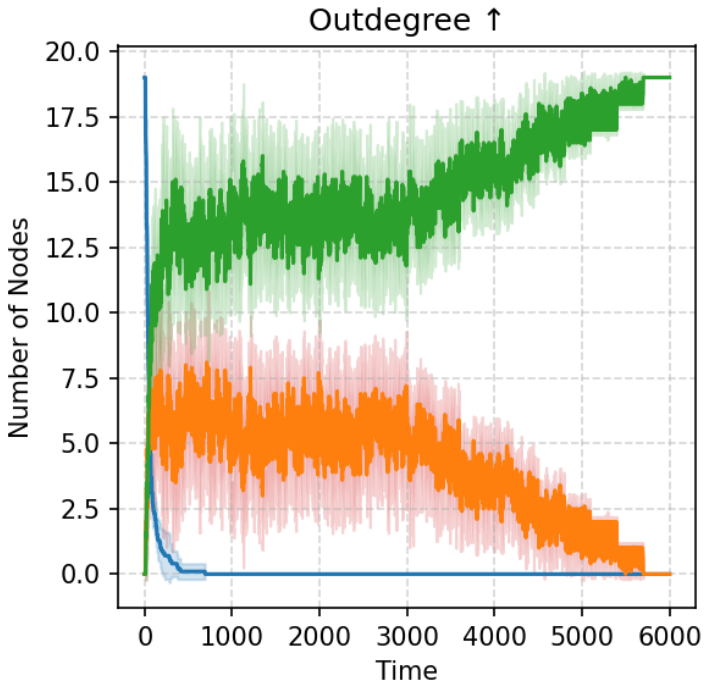
Now the aim is to investigate how the epidemic can be “turned off” in the network by vaccinating nodes according to different ordered lists, and to determine **which strategy is the most effective**. The order in which nodes are vaccinated follows lists of indices sorted according to the following criteria (different **vaccination strategies):**

- **Random**
- **Outdegree (positive) descending**
- **Outdegree (positive) ascending**
- **Betweenness descending**
- **Clustering coefficient descending**

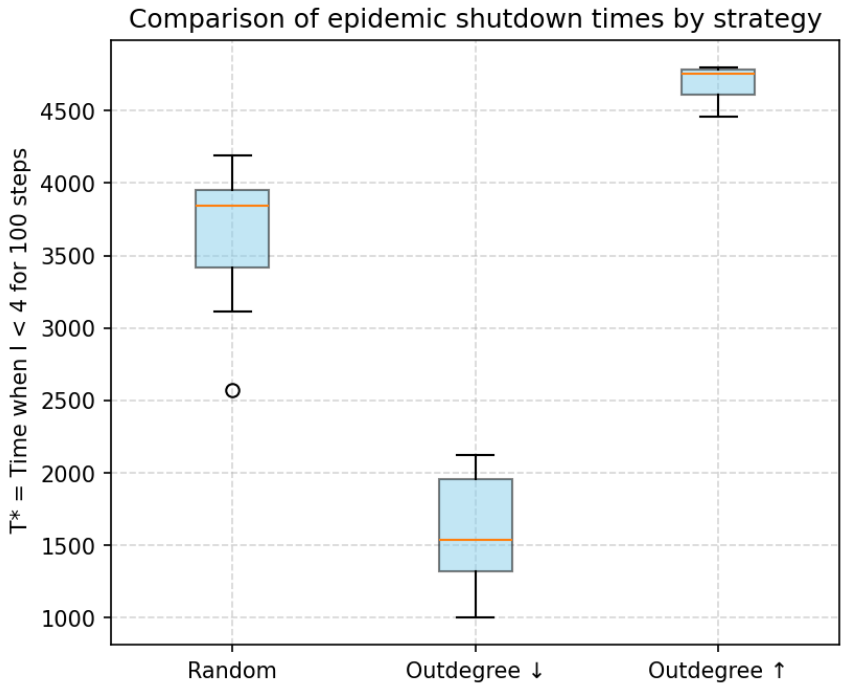
To measure differences between vaccination strategies, a time T^* is defined as the first moment when the number of infected nodes remains **below 4 for 100 consecutive timesteps**, thereby avoiding stochastic fluctuations. By varying the random seed, multiple T^* values are computed for each strategy, and their distributions are visualized using **box plots**.

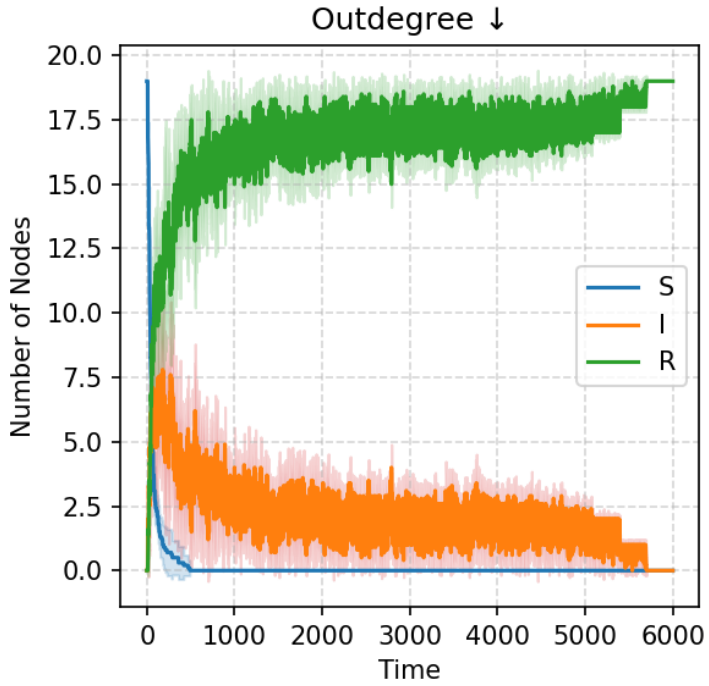

Random
Number of Nodes
Time
S
I
R

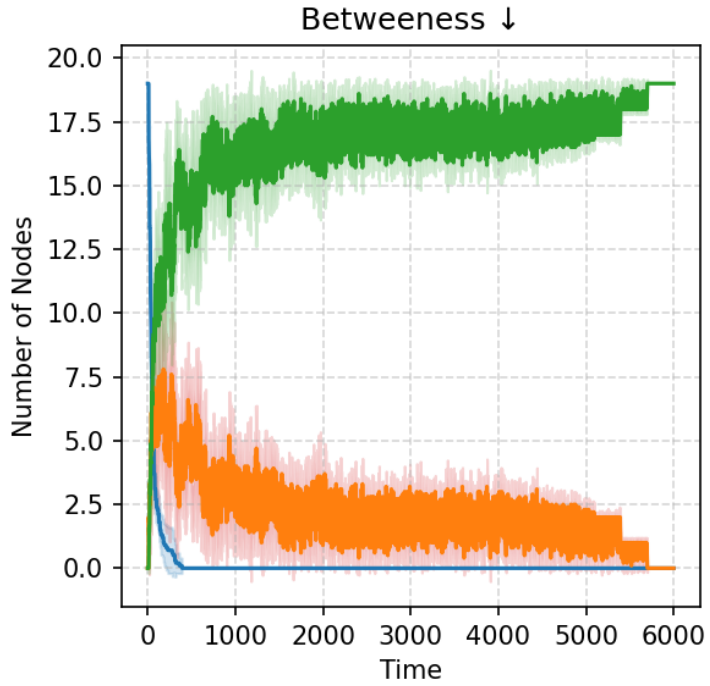

Outdegree ↓
Number of Nodes
Time

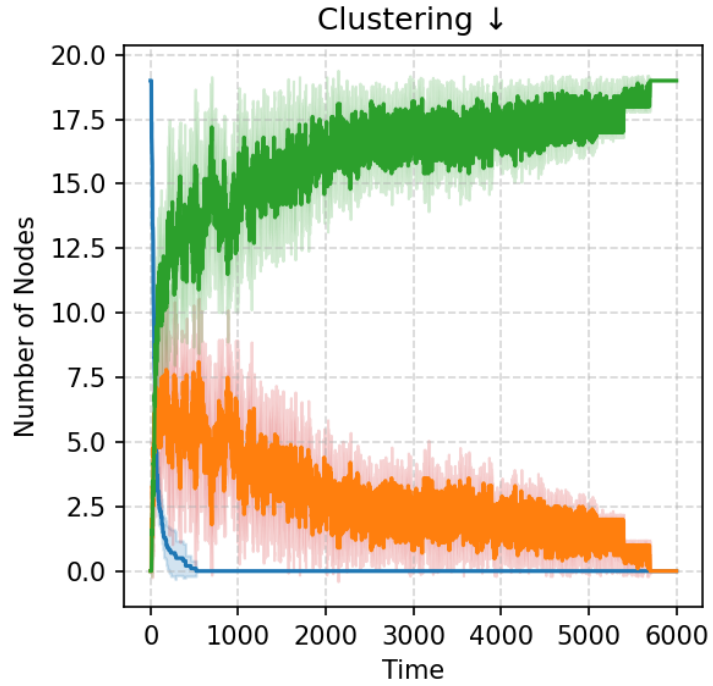

Outdegree ↑
Number of Nodes
Time

$\mu = 0.1$
$\beta = 1$

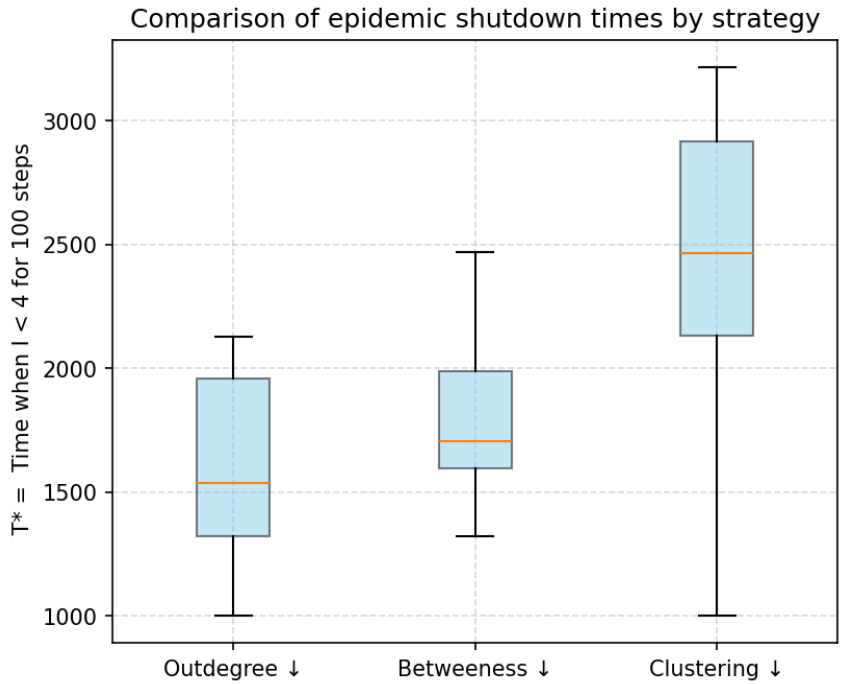

Comparison of epidemic shutdown times by strategy
T* = Time when I < 4 for 100 steps
Random
Outdegree ↓
Outdegree ↑


Outdegree ↓
Number of Nodes
Time
S
I
R


Betweeness ↓
Number of Nodes
Time


Clustering ↓
Number of Nodes
Time

$\mu = 0.1$
$\beta = 1$


Comparison of epidemic shutdown times by strategy
T* = Time when I < 4 for 100 steps
Outdegree ↓
Betweeness ↓
Clustering ↓