Diagnostic décentralisé à l'aide d'automates cellulaires

Nazim Fatès, Régine Marchand, Irène Marcovici

Univ. de Lorraine, Nancy

Journée scientifique de la Fédération Charles Hermite

Science des Réseaux

Jeudi 7 octobre 2021



























pour simplifier : grille finie $\mathcal{L}=\{1,\ldots,X\}\times\{1,\ldots,Y\}$; simu : espace torique



• neutre ightarrow excité avec prob. p_{T} si au moins une cellule excitée dans son voisinage, neutre sinon

- excité devient réfractaire
- réfractaire devient neutre

ondes qui ne s'atténuent pas ; dynamique d'épidémie (S-I-R), communication cellulaire, hola, etc.

transition de phase Greenberg-Hastings





Évolution de la densité de cellules excitées $d_M(t)$ pour M = 4(A) Un seul échantillon, L = 400. (B) Courbe en échelle log-log avec des moyennes pour 30 échantillons ; L = 800pointillés indiquent percolation orientée : $d_M(t) \sim t^{-0.451}$.





corrélation : taux de défauts dans la grille et seuil critique

question : peut-on inverser le processus ? Peut-on avoir un changement qualitatif de comportement pour un taux de cellules défaillantes donné ?

variété de règles possibles





stage de Master N. Gauville

QuorumD

Diagnostic décentralisé à l'aide d'automates cellulaires Nicolas Gauville, Nazim Fatès, Irène Marcovici Proceedings of JFSMA'19 , Toulouse

(France), 2019, p. 96-105

Grille $\mathbb{Z}^2,$ sur laquelle chaque cellule peut prendre 3 états :

- N : normal,
- D : défaillant (il s'agit d'un état fixe),
- A : alerte.

Objectif

Chercher une règle locale la plus simple possible telle qu'à partir d'une configuration initiale avec des cellules **N** et **D**, l'état A envahit la grille si et seulement la proportion d'états **D** dépasse un certain seuil (qu'on souhaite pouvoir choisir).

Première règle proposée



- $\bullet\,$ Si toutes les voisines sont N, le nouvel état est N.
- Si toutes les voisines sont A ou D, le nouvel état est A.
- Sinon, le nouvel état de la cellule est :

N avec proba.
$$\frac{\exp(\lambda n_N)}{\exp(\lambda n_N) + \exp(\lambda(n_A + n_D))}$$
A avec proba.
$$\frac{\exp(\lambda(n_A + n_D))}{\exp(\lambda n_N) + \exp(\lambda(n_A + n_D))},$$

où n_N , n_A , n_D sont resp. les nb. de voisines **N**, **A**, **D**.





En ajustant la valeur de λ , on peut détecter différent seuils de défaillance.



Même règle, mais sur le voisinage de Toom.



On simplifie encore un peu...

- Si les deux voisines (rouges) sont $\boldsymbol{\mathsf{N}},$ le nouvel état est $\boldsymbol{\mathsf{N}}.$
- Si les deux voisines (rouges) sont A ou D, le nouvel état est A.
- Sinon, on garde l'ancien état avec proba. p, et on le change avec proba. 1 – p.



| Config. init. | Cas $p < 1/2$ | Cas $p > 1/2$ |
|---------------------------------|--|--|
| Nb fini de A et | $\mathbb{P}(survio dos \Lambda) > 0$ | Extinction dos A |
| pas de D | $\mathbb{I}\left(\sup_{n \in \mathbb{N}} \operatorname{des} \mathbf{A} \right) > 0$ | |
| Nb fini de D et A | Chaîne de Markov | Chaîne de Markov |
| | transiente | récurrente |
| Densité ρ de A et | Évolution vers un | $ ho > 1/2$: extinction de ${f N}$ |
| $1- ho$ de ${f N}$ | damier A / N | ho < 1/2 : extinction de $f A$ |
| Densité ρ de D et | | $\rho > \rho_c(p)$: extinction |
| 1- ho de N | | de N |
| | | $ ho < ho_{m{c}}(m{p})$: survie de N |

Nombre fini de cellules en Alerte

UNIVERSITÉ De LORRAINE

- Etat initial: nombre fini de cellules en Alerte, aucune cellule Défaillante.
- Nombre de lignes contenant des cellules A: décroissant.



• Longueur de la ligne frontale: domination stochastique par une marche aléatoire de pas

$$(1-p)^2\delta_1 + 2p(1-p)\delta_0 + p^2\delta_{-1}$$

• Conclusion:

si p > 1/2, extinction des **A** en un temps T intégrable, si p < 1/2, persistence des **A** avec probabilité strictement positive.





On regarde donc les (T_i)_i: création de lignes vides en 0.
Les (T_{i+1} - T_i)_i sont iid et admettent un moment exponentiel.
On appelle **bulle numéro** i la partie du processus qui évolue entre les lignes créés en T_i et T_{i+1}.
Les bulles évoluent de façon indépendante.

Une seule cellule Défaillante (2)



La bulle numéro k nait à l'instant T_k , vit un temps V_k , puis disparait.

Description par une file d'attente $(R_t)_{t\geq 1}$: $R_0 = 0$ et

$$\begin{cases} \text{ si } t = T_k \text{ alors } & R_{t+1} = \max(R_t - 1, V_k), \\ & \text{ sinon } & R_{t+1} = \max(R_t - 1, 0). \end{cases}$$
(1)

Lemme

<

Si le temps de vie V_1 d'une bulle est intégrable, alors la file d'attente est récurrente positive, et le nombre de **A** est infiniment souvent nul.

Démonstration. Critère de Foster.

Théorème

Si p > 1/2, alors le temps de vie V_1 d'une bulle est intégrable, et le nombre de **A** est infiniment souvent nul.

Temps de vie d'une bulle (p > 1/2)



- Le nombre N de lignes admet des moments exponentiels
- longueur de la ligne $1 \leq marche aléatoire biaisée: disparition à l'instant <math>H_1$ intégrable.
- Tant que la ligne 1 existe, elle favorise la persistence de la ligne 2.

 $r_n^2 - r_n^1 \preceq$ marche aléatoire symétrique,

donc contrôle de la longueur de la ligne 2 à l'instant H_1 :

 $\mathbb{E}(\ell_2) \leq A\mathbb{E}(\sqrt{H_1}) + B.$

Ensuite, elle évolue comme la ligne 1, et disparait à l'instant H_2 :

$$\mathbb{E}(H_2) - \mathbb{E}(H_1) \leq A' \mathbb{E}(\sqrt{H_1}) + B'.$$

- récurrence: $\mathbb{E}(H_i) \leq A'' i^2$.
- $\mathbb{E}(V_1) = \mathbb{E}(H_N) = \sum_i \mathbb{E}(H_i)\mathbb{P}(T=i) < +\infty$ Gagné !

Network science and dynamic graphs

Sebastian Miron¹, Konstantin Usevich ¹, Marianne Clausel²

¹ Centre de Recherche en Automatique ² Institut Élie Cartan de Lorraine



Journée FCH « Sciences des réseaux », 7 octobre 2021

Motivation

Dynamic networks are ubiquituous...



Example 1 : fMRI and brain connectivity [Achard et al. 2007]¹

^{1.} Achard S, Bullmore E. Efficiency and cost of economical brain functional networks. PLoS computational biology. 2007 Feb;3(2):e17.

Motivation

Dynamic networks are ubiquituous...

Contacts in primary school



Example 2 : dynamic networks and socio patterns [Gauvin et al., 2014]²

^{2.} http://www.sociopatterns.org/gallery/dynamical-contact-patterns-in-a-primary-school/

Outline

Challenges Tensor approach for dynamic graphs Multivariate time series approach for dynamic graphs

Challenges

Static networks representations



Modeling dynamic networks?

Graph vs. multivariate time series

How can we take into account evolution with time?

- the value of each node
- the dependence between the nodes
- the topology of the network can evolve with time

Modeling dynamic networks?



Static graph with dynamic signal

Dynamic graph with static signal

Dynamic graph with dynamic signal

Scientific questions

Dynamic networks for practitioners

Practical issues

- Can we group the nodes into relevant categories?
- Structural changes
- Novelty and anomaly detection

Tensor approach for dynamic graphs

Tensors : a natural model for dynamic network



Tensors : a natural model for dynamic network

Tensors and dynamic networks analysis

- Classical static graph analysis is based on the study of spectral properties of the adjacency matrix/Laplacian
- Taking into account time needs more complex objects : tensors



- Many challenges with tensors ^{3 4}
 - storage
 - no straightforward extension of classical spectral analysis!

T. G. Kolda, B. W. Bader. Tensor Decompositions and Applications. SIAM Review, 2009.
 P. Comon, Tenseurs en Sciences de Données (Encyclopédie des Techniques de l'Ingénieurs), 2021.

Tensors : a natural model for dynamic network

Main challenges

- compression of information
- dynamic clustering of nodes
- tensor decompositions :
 - possibility to take into account specific constraints (positivity, low rank) related to the application
 - existence of theoretical guarantees
 - efficient optimisation algorithms for decomposition

Tensor decompositions



Canonical Polyadic (CP) Decomposition

Tucker Decomposition

LL1 (block-term) decomposition

Example : contacts in a school



- [Gauvin et al., 2014] : nonnegative CP decomposition
- PhD thesis of M. Diop (CRAN, 2018) : binary tensor decomposition

Multivariate time series approach for dynamic graphs

Multivariate time series

More about dependence modeling

- First approach to model dynamic dependence between nodes : adjacency/weights matrix
- Alternative approach : nodes are time series
- Several notion of dependence are involved
- Time-dependence and dependence between components can be combined
- Several way to model dependence between time series : empirical correlation, long-run correlation, wavelet correlation, signature of order 2...

Multivariate time series



An example of dependence measure : wavelet correlation⁵

^{5.} S. Achard et al., A Resilient, Low-Frequency, Small-World Human Brain Functional Network with Highly Connected Association Cortical Hubs

Multivariate time series



An example of dependence measure : signature coefficients⁶

^{6.} N. Sugiura, Machine Learning Technique Using the Signature Method for Automated Quality Control of Argo Profiles
Conclusion

Perspectives

- Dynamic networks : interdisciplinary field with challenging questions !
- Beyond graphs
 - Model dependence with more than two objects?
 - Analyze information transfer and causality in oriented graphs

Laplacian Spectra of Graphs and Cyber-Insurance Protection

Nabil Kazi-Tani IECL, Université de Lorraine En collaboration avec Thierry Cohignac (CCR, Paris)

Journée Science des Réseaux - 7 octobre 2021

- A given graph has to be protected against communication interruption, via prevention measures, or via insurance.
- Graph = Company's computer network, susceptible to exterior attacks.

▲□▶ ▲□▶ ▲ 三▶ ★ 三▶ 三三 - のへぐ

- A given graph has to be protected against communication interruption, via prevention measures, or via insurance.
- Graph = Company's computer network, susceptible to exterior attacks.
- ► Or... Interconnected data, in the form of a graph → susceptible to data breaches, which can be insured.

- A given graph has to be protected against communication interruption, via prevention measures, or via insurance.
- Graph = Company's computer network, susceptible to exterior attacks.
- ► Or... Interconnected data, in the form of a graph → susceptible to data breaches, which can be insured.

Question: impact of the topology of the network on the chosen protection?

- Nodes can represent: computers, servers, softwares, routers, data sets, etc.
- Edges can be: internet connections, distant access to servers, links between softwares etc.



(a) A graph on n = 7 vertices, with 8 edges.



(b) A 3-regular graph on n = 12 vertices.

Inhomogeneous Suseptible-Infected-Susceptible (SIS) model on Edges.

State at time t of edge i is Bernoulli: X_i(t) = 0 when edge i is healthy and X_i(t) = 1 when edge i is infected.

Inhomogeneous Suseptible-Infected-Susceptible (SIS) model on Edges.

- State at time t of edge i is Bernoulli: X_i(t) = 0 when edge i is healthy and X_i(t) = 1 when edge i is infected.
- When infected, edge *i* can infect each edge *j* sharing a common vertex with rate β_{ij} > 0.

Inhomogeneous Suseptible-Infected-Susceptible (SIS) model on Edges.

- State at time t of edge i is Bernoulli: X_i(t) = 0 when edge i is healthy and X_i(t) = 1 when edge i is infected.
- When infected, edge *i* can infect each edge *j* sharing a common vertex with rate β_{ij} > 0.
- Each edge i can be cured, i.e. its state X_i jumps from 1 to 0 at rate δ_i > 0.
- The Poisson infection and curing processes are assumed independent.

Inhomogeneous Suseptible-Infected-Susceptible (SIS) model.



No edge is infected.

▲□▶ ▲□▶ ▲□▶ ▲□▶ = 三 のへで

Inhomogeneous Suseptible-Infected-Susceptible (SIS) model.



Red edge *i* is infected, and can be cured at rate $\delta_i > 0$.

▲□▶ ▲□▶ ▲□▶ ▲□▶ = 三 のへで

Inhomogeneous Suseptible-Infected-Susceptible (SIS) model.



Edge *i* infects each blue edge *j* at rate $\beta_{ij} > 0$.

▲□▶ ▲□▶ ▲□▶ ▲□▶ = 三 のへで

► We use 2 simplifying approximations:

- We use 2 simplifying approximations:
- *N*-Intertwined Mean Field Approximation (Kooij, Omic, Van Mieghem 2008): upper bounds the infection probabilities.

▲□▶ ▲□▶ ▲ 三▶ ★ 三▶ 三三 - のへぐ

- We use 2 simplifying approximations:
- N-Intertwined Mean Field Approximation (Kooij, Omic, Van Mieghem 2008): upper bounds the infection probabilities.
- Long time steady state: which is accurate when infection probabilities are high, corresponding to cascading infections.

Let \tilde{a} = edge adjacency matrix and m = number of edges.

$$rac{X_i(t+\Delta t)-X_i(t)}{\Delta t}=(1-X_i(t))\sum_{j=1}^m ilde{a}_{ij}eta_{ij}X_j(t)-\delta_iX_i(t).$$

Let $\tilde{a} =$ edge adjacency matrix and m = number of edges.

$$rac{X_i(t+\Delta t)-X_i(t)}{\Delta t}=(1-X_i(t))\sum_{j=1}^m ilde{a}_{ij}eta_{ij}X_j(t)-\delta_iX_i(t).$$

leads to

$$rac{d \mathsf{v}_i(t)}{dt} = \sum_{j=1}^m ilde{a}_{ij} eta_{ij} \mathsf{v}_j(t) - \sum_{j=1}^m ilde{a}_{ij} eta_{ij} \mathbb{E}[X_i(t) X_j(t)] - \delta_i \mathsf{v}_i(t),$$

▲□▶ ▲□▶ ▲□▶ ▲□▶ ▲□ ● ● ●

where $v_i(t) = \mathbb{P}(X_i(t) = 1)$.

N-Intertwined Mean Field Approximation:

$$rac{d \mathsf{v}_i(t)}{dt} = \sum_{j=1}^m ilde{a}_{ij} eta_{ij} \mathsf{v}_j(t) - \sum_{j=1}^m ilde{a}_{ij} eta_{ij} \mathbb{E}[\mathsf{X}_i(t)] \mathbb{E}[\mathsf{X}_j(t)] - \delta_i \mathsf{v}_i(t),$$

N-Intertwined Mean Field Approximation:

$$\frac{dv_i(t)}{dt} = \sum_{j=1}^m \tilde{a}_{ij}\beta_{ij}v_j(t) - \sum_{j=1}^m \tilde{a}_{ij}\beta_{ij}\mathbb{E}[X_i(t)]\mathbb{E}[X_j(t)] - \delta_i v_i(t),$$

i.e

$$\frac{d\mathsf{v}_i(t)}{dt} = \left(\sum_{j=1}^m \tilde{\mathsf{a}}_{ij}\beta_{ij}\mathsf{v}_j(t)\right)(1-\mathsf{v}_i(t)) - \delta_i\mathsf{v}_i(t), \quad i=1,\ldots,m.$$

N-Intertwined Mean Field Approximation and Long time steady state:

$$\frac{dv_i(t)}{dt} = 0 = \left(\sum_{j=1}^m \tilde{a}_{ij}\beta_{ij}v_j(t)\right)(1-v_i(t)) - \delta_i v_i(t), \quad i=1,\ldots,m.$$

(ロ)、(型)、(E)、(E)、 E) のQ(()

N-Intertwined Mean Field Approximation and Long time steady state:

$$\frac{dv_i(t)}{dt} = 0 = \left(\sum_{j=1}^m \tilde{a}_{ij}\beta_{ij}v_j(t)\right)(1-v_i(t)) - \delta_i v_i(t), \quad i=1,\ldots,m.$$

i.e.

$$\mathbf{v}_i = \frac{\beta \sum_{j=1}^m \tilde{a}_{ij} \mathbf{v}_j}{\delta_i + \beta \sum_{j=1}^m \tilde{a}_{ij} \mathbf{v}_j}, \quad i = 1, \dots m,$$

where $\beta_{ij} = \beta$.

Control the vector of curing rates (δ₁,..., δ_m) to make the graph as connected as possible, in a situation of cascading infections.

- Control the vector of curing rates (δ₁,..., δ_m) to make the graph as connected as possible, in a situation of cascading infections.
- Stand Alone Cyber Contracts: clauses of repairing actions, through the intervention of a specified cyber security company.

Prevention or cyber hygiene measures: back-ups, users training, network mapping and inventory etc.

- Control the vector of curing rates (δ₁,..., δ_m) to make the graph as connected as possible, in a situation of cascading infections.
- Stand Alone Cyber Contracts: clauses of repairing actions, through the intervention of a specified cyber security company.
- Prevention or cyber hygiene measures: back-ups, users training, network mapping and inventory etc.
- ▶ How should we measure the "connectedness" of a given graph G?

► Given a graph G with adjacency matrix A and degree matrix D, the Laplacian matrix of G is given by L := D – A.

- ► Given a graph G with adjacency matrix A and degree matrix D, the Laplacian matrix of G is given by L := D A.
- Algebraic connectivity λ₂(G)= lowest strictly positive eigenvalue of the Laplacian matrix.
- ▶ Known to describe the "connectedness" of a graph (Fiedler 1973).



Figure: The lower graph \widetilde{G} is "more connected" than the upper graph G.



Figure: Complete graph K_{10} with $\lambda_2(K_{10}) = 10$.

◆□▶ ◆□▶ ◆三▶ ◆三▶ 三三 - のへぐ

1. Consider a SIS epidemic model on the edges of a graph *G*, for a long time.

1. Consider a SIS epidemic model on the edges of a graph *G*, for a long time.

・ロト・日本・ヨト・ヨト・日・ つへぐ

2. We obtain a random graph \widetilde{G}_{∞} , with less edges.

1. Consider a SIS epidemic model on the edges of a graph *G*, for a long time.

・
 ・
 ・
 ・
 ・
 ・
 ・
 ・
 ・
 ・
 ・
 ・
 ・
 ・
 ・
 ・
 ・
 ・
 ・
 ・
 ・
 ・
 ・
 ・
 ・
 ・
 ・
 ・
 ・
 ・
 ・
 ・
 ・
 ・
 ・
 ・
 ・
 ・
 ・
 ・
 ・
 ・
 ・
 ・
 ・
 ・
 ・
 ・
 ・
 ・
 ・
 ・
 ・
 ・
 ・
 ・
 ・
 ・
 ・
 ・
 ・
 ・
 ・
 ・
 ・
 ・
 ・
 ・
 ・
 ・
 ・
 ・
 ・
 ・
 ・
 ・
 ・
 ・
 ・
 ・
 ・
 ・
 ・
 ・
 ・
 ・
 ・
 ・
 ・
 ・
 ・
 ・
 ・
 ・
 ・
 ・
 ・
 ・
 ・
 ・
 ・
 ・
 ・
 ・
 ・
 ・
 ・
 ・
 ・
 ・
 ・
 ・
 ・
 ・
 ・
 ・
 ・
 ・
 ・
 ・
 ・
 ・
 ・
 ・
 ・
 ・
 ・
 ・
 ・
 ・
 ・
 ・
 ・
 ・
 ・
 ・

- 2. We obtain a random graph \widetilde{G}_{∞} , with less edges.
- Consider the associated random Laplacian matrix L_∞ and its average L := E[L_∞].

- 1. Consider a SIS epidemic model on the edges of a graph *G*, for a long time.
- 2. We obtain a random graph \widetilde{G}_{∞} , with less edges.
- Consider the associated random Laplacian matrix L_∞ and its average L := E[L_∞].
- 4. Maximize the algebraic connectivity of L:

where c is a cost function, B a given budget constraint, and A an admissible set of curing rates.

- 1. Consider a SIS epidemic model on the edges of a graph *G*, for a long time.
- 2. We obtain a random graph \widetilde{G}_{∞} , with less edges.
- Consider the associated random Laplacian matrix L_∞ and its average L := E[L_∞].
- 4. Maximize the algebraic connectivity of L:

 $\begin{array}{ll} \underset{\delta \in \mathcal{A}}{\text{maximize}} & \lambda_2(L) \\ \text{subject to} & c(\delta) \leq B, \end{array}$

where c is a cost function, B a given budget constraint, and A an admissible set of curing rates.

5. (Edge *i* interrupted \rightarrow loss Z_i with $\mathbb{E}[Z_i] = c_i$.)

A Solution

Theorem

An optimal solution is given by

$$\delta^{\star}_i = \left(eta \sum_{j=1}^m ilde{a}_{ij} extsf{v}^{\star}_j
ight) rac{1- extsf{v}^{\star}_i}{ extsf{v}^{\star}_i},$$

where $(1 - v_1^*, \dots, 1 - v_m^*)$ is the *w* component of a solution to the following convex problem in $\gamma \in \mathbb{R}$, $\mu \in \mathbb{R}$ and $w \in \mathbb{R}^m$:

A Solution

Theorem

An optimal solution is given by

$$\delta_i^{\star} = \left(\beta \sum_{j=1}^m \tilde{a}_{ij} v_j^{\star}\right) \frac{1 - v_i^{\star}}{v_i^{\star}},$$

where $(1 - v_1^*, \dots, 1 - v_m^*)$ is the *w* component of a solution to the following convex problem in $\gamma \in \mathbb{R}$, $\mu \in \mathbb{R}$ and $w \in \mathbb{R}^m$:

sup
$$\gamma$$

subject to $L(w) - \gamma I + \mu 11^t \succcurlyeq 0$
 $0 \le w \le 1$
 $\sum_{i=1}^m c_j w_j \le B,$

where $1^t := (1, ..., 1) \in \mathbb{R}^m$, L(w) denotes the Laplacian matrix of the weighted graph (V, E_0, w) , $M \succeq 0$ means that M is positive semidefinite and $(\tilde{a}_{ij})_{a \le i,j \le m}$ are the entries of the edge-adjacency matrix.

Examples



Figure: Erdös-Rényi Random Graph with 102 edges. The dotted edges have optimal δ equal to 0.


Figure: Erdös-Rényi Random Graph with 102 edges. The dotted edges have optimal δ equal to 0.

▲□▶ ▲□▶ ▲□▶ ▲□▶ ▲□ ● のへで

Solved with the CVX Matlab Package (Boyd, Grant).



Figure: Each edge width is proportional to the optimal δ .



Figure: Erdös-Rényi Random Graph with 1004 edges. Red edges have null δ^{\star} values.

◆□ → ◆檀 → ◆臣 → ◆臣 → □臣 □



Figure: Each edge width is proportional to the optimal δ .

(ロ)、



Figure: δ^* values, for a graph with 24978 edges.

・ロト ・ 日 ト ・ モ ト ・ モ ト

æ

Perspectives

 Finite horizon problem and dynamic optimization: stochastic control problem! (Avoids both NIMFA and Steady state approximations)

(ロ)、(型)、(E)、(E)、 E) のQ()

Perspectives

 Finite horizon problem and dynamic optimization: stochastic control problem! (Avoids both NIMFA and Steady state approximations)

▲□▶ ▲□▶ ▲□▶ ▲□▶ ▲□ ● ● ●

 Inhomogeneous fast mixing Markov chains (Boyd, Diaconis, Sun, Xiao).

Perspectives

- Finite horizon problem and dynamic optimization: stochastic control problem! (Avoids both NIMFA and Steady state approximations)
- Inhomogeneous fast mixing Markov chains (Boyd, Diaconis, Sun, Xiao).
- Simulation and (optimal) importance sampling for Markov chains.

Optimal stochastic control and games on networks.

Merci de votre attention