



Université de Namur  
Faculté d’Informatique

# SDASM101 – Graph Mining

Professeur : Cyril de Bodt

Projet académique

---

Graphes analysés (base de données SocioPatterns) :

Primary School Temporal Network  
Hospital Ward Dynamic Contact Network

---

Ariane SAUVENIER Certificat Data Science  
Michaël MESTRE Certificat Data Science

Année académique 2025-2026

---

# 1 Introduction

Nous avons choisi deux environnements de contact humain – une école primaire et un service hospitalier - pour étudier comment la structure des interactions sociales influence le risque de propagation d'un agent infectieux.

Nous analysons les deux graphes empiriques issus de la base de données SocioPatterns, et comparons nos conclusions à celles des articles scientifiques associés.

Bien que les graphes aient initialement été construits sous forme pondérée, afin de tenir compte de la durée cumulée des interactions, l'analyse présentée dans ce rapport se concentre principalement sur une représentation non pondérée et statique des réseaux. Ce choix vise à privilégier l'étude des propriétés structurelles des graphes, en adéquation avec les outils et concepts abordés durant le cours, tout en conservant les poids uniquement à des fins descriptives ou illustratives lorsque cela s'avérait pertinent.

Afin de respecter la contrainte de longueur, les figures et tableaux détaillés sont regroupés en annexe. Le texte principal se concentre sur l'interprétation des résultats essentiels.

## 2 Description des données

### 2.1 Hospital Ward Dynamic Contact Network

Ce dataset<sup>1</sup> contient les interactions en face-à-face entre quatre catégories de personnes (médecins, infirmières, personnel administratif et patients) d'un service hospitalier. Un contact est enregistré lorsque deux badges se trouvent à moins d'1,5 m, ce qui permet de construire un graphe pondéré où le poids représente la durée totale des interactions. Le réseau comprend 75 individus répartis en : 27 infirmières (NUR), 11 médecins (MED), 8 administratifs (ADM) et 29 patients (PAT).

### 2.2 Primary School Temporal Network

Le dataset Primary School Temporal Network<sup>2</sup> décrit les interactions entre élèves et enseignants d'une école primaire française, enregistrées par capteurs durant deux journées. Chaque contact correspond à une proximité inférieure à 1,5 m, permettant de construire un graphe pondéré ou temporel. Le réseau comprend 242 individus structurés en plusieurs classes (1A, 1B, 2A, ...) d'une vingtaine d'élèves chacune.

### 2.3 Prétraitement et représentation des graphes

Les données de l'hôpital ont été agrégées en un graphe pondéré non orienté où chaque noeud représente un individu et chaque arête est pondérée par la durée cumulée des contacts entre eux. Les interactions ont d'abord été regroupées par paire ( $i, j$ ), puis sommées pour obtenir un poids unique par arête. Les catégories des individus (MED, NUR,

---

1. <http://www.sociopatterns.org/datasets/hospital-ward-dynamic-contact-network/>  
2. <http://www.sociopatterns.org/datasets/primary-school-cumulative-networks/>

ADM, PAT) ont été conservées comme attributs de noeuds. La figure 1 illustre le graphe pondéré obtenu, avec une représentation des rôles par couleur et une épaisseur d'arête proportionnelle à la durée cumulative des interactions.

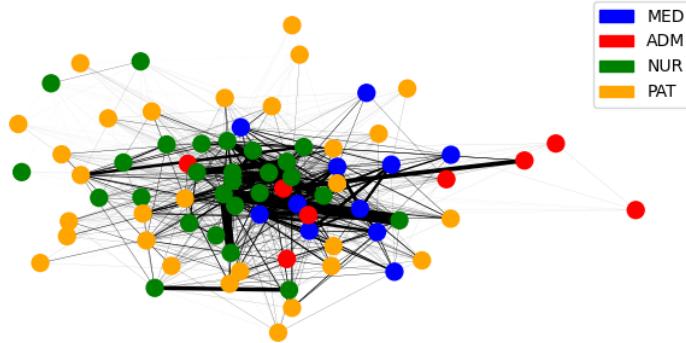


FIGURE 1 – Graphe pondéré du service hospitalier. Les couleurs indiquent les catégories (MED, NUR, ADM, PAT) et l'épaisseur des arêtes est proportionnelle aux temps cumulés de contact entre individus.

Le même travail a été effectué pour le réseau scolaire. Ici, ce sont les classes (1A, 1B, 2A,...) qui ont été conservés comme attributs de noeuds. Nous avons choisi de donner une couleur bien visible aux enseignants étant donné leur rôle différent, même si chacun d'eux est responsable d'une seule classe dans cette école. L'épaisseur des arêtes est également pondérée par la durée cumulée des contacts.

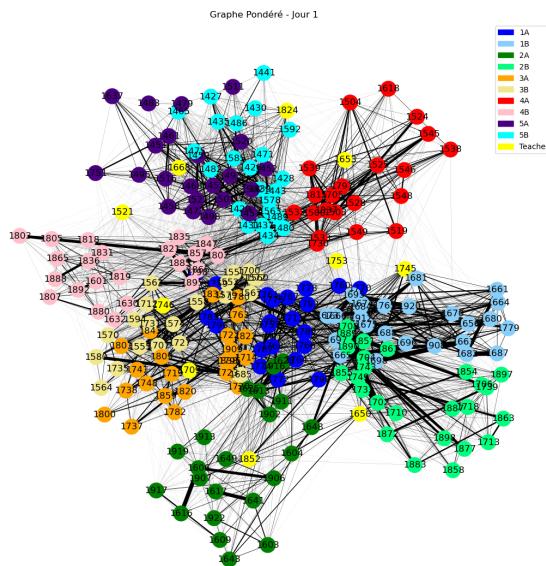


FIGURE 2 – Graphe pondéré de l'école primaire, pour le premier jour de mesures. Les couleurs indiquent les classes, une autre couleur a été choisie pour les enseignants.

---

## 3 Analyse structurelle

### 3.1 Distribution des degrés

Les distributions des degrés pondérés et non pondérés pour les données de l'hôpital ainsi que leur graphe log/log ont été générés (voir Annexe A).

La distribution des degrés du réseau hospitalier est clairement hétérogène : la plupart des individus ont un nombre de contacts relativement modéré, mais quelques noeuds présentent des degrés beaucoup plus élevés. Le graphique log–log indique qu'il ne s'agit pas d'une loi de puissance.

Lorsque l'on considère le degré pondéré, l'hétérogénéité devient plus marquée : une minorité d'acteurs accumule des durées d'interaction plus élevées. Cela révèle l'existence de super-connecteurs, susceptibles de jouer un rôle central dans l'environnement hospitalier.

A nouveau, nous avons effectué les mêmes analyses pour le réseau scolaire. Ce réseau est assez dense, et ne présente pas de noeud isolé. La distribution des degrés y est aussi relativement hétérogène avec quelques individus qui présentent des degrés plus élevés. Nous ne rencontrons pas non plus, sur ce réseau, une loi de puissance, comme illustré sur le graphique log-log (voir Annexe A, figure 6).

De même, le graphique illustrant le degré pondéré par la durée des contacts remet en évidence cette hétérogénéité du réseau : les degrés pondérés varient beaucoup. La majorité des nœuds a un degré pondéré entre 3000 et 15000, donc un temps total de contact modéré. La distribution étant assez étalée, cela montre que certains individus passent beaucoup plus de temps en contact que d'autres. La distribution des degrés pondérés indique aussi des individus très impliqués dans les contacts prolongés.

### 3.2 Clustering

Le coefficient de clustering moyen du réseau hospitalier est élevé (0.64), ce qui indique que les individus ont tendance à former des triangles d'interaction. L'analyse par catégories montre que les patients et le personnel administratif présentent les plus hauts niveaux de clustering (0.70), ce qui traduit des interactions fortement localisées au sein de petits groupes. Les médecins ont un clustering légèrement inférieur (0.66), tandis que les infirmières affichent la valeur la plus basse (0.55), cohérente avec leur rôle plus mobile et distribuée entre différents patients et collègues.

Le coefficient de clustering moyen du réseau scolaire est modéré (jour 1 : 0.50 et jour 2 : 0.56), ce qui indique que les contacts forment des groupes bien connectés. Cependant, en tenant compte de la durée des contacts (pondération), le clustering chute fortement ( $\approx 0,008$ ), ce qui signifie que les contacts longs sont moins souvent partagés au sein de triangles. Certaines classes comme la 1B montrent un clustering pondéré légèrement plus élevé, suggérant que dans ces groupes les interactions prolongées entre amis communs sont un peu plus fréquentes. Globalement, cela révèle que les relations sociales forment des petits groupes denses, mais que les contacts prolongés s'organisent plutôt en chaînes qu'en triangles fortement connectés.

---

### 3.3 Centralités

Les mesures de centralité du réseau hospitalier (voir annexe B) donnent toute des conclusions similaires : les individus les plus centraux sont essentiellement issus du personnel administratif, des infirmières et, dans une moindre mesure, des médecins. Le nœud 1098 (ADM) occupe systématiquement la première place en degree, closeness, betweenness, Katz et PageRank, ce qui indique un rôle fortement connecté. Les infirmières sont majoritairement présentes dans le top-10 pour l'ensemble des centralités, ce qui reflète leur fonction de contact fréquent et distribué entre patients et personnel soignant. Les médecins, quant à eux, obtiennent des scores élevés mais plus hétérogènes : ils sont moins nombreux mais souvent connectés à plusieurs catégories d'acteurs, ce qui leur confère un rôle de relais secondaires.

En ce qui concerne le réseau scolaire (jour 1), nous avons pu mettre en évidence l'existence d'un "Hub" : l'élève de 3B identifié par le numéro 1551 ! En effet, cet élève est le numéro un de tous les types de centralité, avec notamment un degré valant 98 (pour le jour 1!). En observant le réseau sur Gephi (taille des noeuds pondérée par leur degré), nous constatons que cet élève est en contact avec presque tous les élèves ayant un degré élevé. Sept des élèves qui arrivent dans le top 10 du classement en terme de centralités sont issus des classes de 3ème, A et B ; deux élèves viennent de la classe de 2B et un élève de la classe de 1A en ce qui concerne les calculs de centralité de degré, la Closeness, la centralité de Katz et le PageRank. Nous constatons que la centralité Betweenness est légèrement différente.

### 3.4 Corrélation de Pearson

En ce qui concerne le réseau scolaire, les mesures étant réalisées sur deux journées, nous avons pris le temps, comme les auteurs, de calculer le coefficient de corrélation de Pearson afin de vérifier si le comportement des individus varie fort d'un jour à l'autre. Dans le cas du nombre de contacts par jour pour chaque individu, nous obtenons une valeur élevée de 0.536. En ce qui concerne la corrélation calculée sur le temps de contact de chaque individu chaque jour, le coefficient est élevé aussi et vaut 0.545. Ceci nous permet d'affirmer que le comportement des enfants et de leurs enseignants est cohérent d'un jour à l'autre.

## 4 Modèles de réseaux

Les trois modèles aléatoires présentent des propriétés différentes de celles du réseau hospitalier réel (les données comparatives détaillées sont présentées en annexe C). Le modèle d'Erdős–Rényi (ER), bien qu'il reproduise correctement la densité, ne parvient pas à reproduire le même coefficient de clustering (0.40) que celui du graphe réel (0.64). Cela montre que la présence de triangles dans le service hospitalier ne peut pas être expliquée par un choix aléatoire.

Le modèle Configuration, ne parvient pas non plus à reproduire un graphe similaire

---

au graphe réel. Son clustering (0.42) et sa densité (0.31) demeurent nettement plus faibles que ceux du graphe empirique.

Le modèle Barabási-Albert (BA) génère une structure plus hétérogène que ER, avec un clustering légèrement meilleur (0.44), mais il reste en dessous des valeurs réelles.

Dans l'ensemble, aucun des modèles ne parvient à capturer la combinaison caractéristique du réseau hospitalier : densité élevée, clustering fort et distances courtes, reflétant les interactions propres à cet environnement. Cela indique une structure organisationnelle particulière que les modèles théoriques évalués ne reproduisent pas.

De même, pour le réseau scolaire (jour 1), les modèles théoriques présentent des propriétés différentes de celles du réseau réel. Le modèle d'Erdos-Rényi reproduit une densité très proche mais un coefficient de clustering valant moins de la moitié de celui du réseau réel (réseau réel : 0.50, ER : 0.213). Pour le modèle de Configuration et celui de Barabási-Albert, tant la densité que le coefficient de clustering restent éloignés des valeurs du réseau réel.

## 5 Détection de communautés

L'algorithme de Louvain appliqué au réseau hospitalier identifie quatre communautés de tailles comparables, mais la modularité obtenue est faible ( $Q = 0.099$ ). Cette valeur indique que les regroupements du graphe sont peu marqués : les interactions sont fortement entrelacées entre les catégories, ce qui reflète un environnement de travail dense où les rôles interagissent continuellement.

La composition des communautés (voir en Annexe D) confirme cette faible séparation structurelle : chaque groupe contient un mélange des différentes catégories.

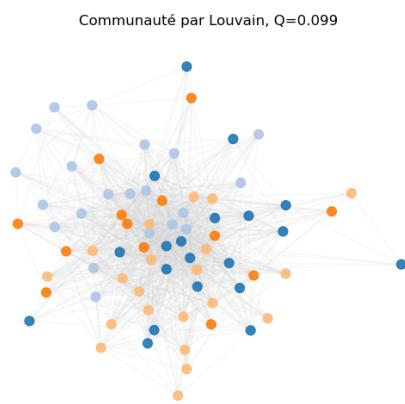


FIGURE 3 – Partition Louvain du réseau hospitalier. Les quatre couleurs indiquent les communautés détectées ( $Q = 0.099$ ).

L'algorithme de Louvain appliqué à l'école primaire (jour 1 et jour 2) identifie 5 communautés mettant en évidence la structure du réseau. La modularité est bonne pour

les deux jours : ( $Q_1 = 0.375$  et  $Q_2 = 0.386$ ), ce qui indique que les regroupements trouvés par Louvain sont significatifs. Le réseau possède des communautés distinctes avec plus de liens internes qu'attendus aléatoirement.

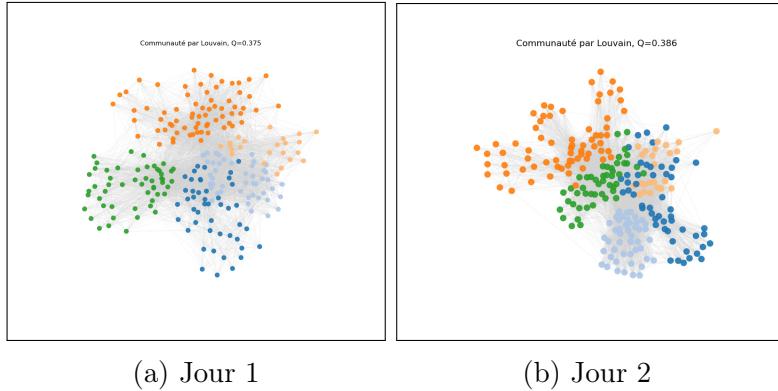


FIGURE 4 – Détection de communautés par l'algorithme de Louvain

La table en Annexe D montre la comparaison entre les classes réelles de l'école et les communautés détectées par Louvain. Nous remarquons que trois des communautés regroupent deux classes et qu'une communauté en regroupe trois. Cela pourrait s'expliquer par la réalité du terrain : les enfants de ces classes ont probablement des temps de récréation et de cantine communs. Nous constatons enfin qu'aucun élève n'est "exclu" de sa classe par l'algorithme. Les temps de contact entre les élèves d'un même groupe classe sont suffisamment importants.

Pour le deuxième jour, nous avons des résultats similaires : 5 communautés détectées par Louvain et le nombre de classe par communauté reste le même. Néanmoins, les regroupements sont différents. Prenons l'exemple de la communauté 4 : cette fois-ci, c'est la classe de 5A qui se retrouve dans cette communauté avec la classe de 1B.

## 6 Dynamique

La dynamique de consensus a été étudiée pour le graphe de l'hôpital. Cependant, étant donné la faible modularité et la densité élevée, aucun pré-alignement n'est observé et le consensus global est atteint très rapidement.

Pour le réseau scolaire, nous avons attribué une variable d'état initial, générée aléatoirement, à chaque noeud. Les valeurs initiales des noeuds sont aussi différenciées en fonction de leur classe. Le consensus est rapidement atteint, surtout le premier jour, et sans passer par un pré-alignement par classe, comme nous pouvons le constater sur les figures en Annexe E.

---

## 7 Marches aléatoires

### 7.1 Marche aléatoire stationnaire

Les probabilités stationnaires (voir annexe F.1) les plus élevées observées sur le graphe de l'hôpital correspondent principalement aux infirmières et à certains médecins, qui présentent les degrés les plus importants. Le nœud le plus visité est un membre du personnel administratif (1098), suivi par plusieurs infirmières très centrales (1193, 1115, 1164). À l'inverse, les patients occupent des positions plus périphériques.

En ce qui concerne le réseau scolaire, les probabilités stationnaires les plus élevées observées confirment également les observations précédentes. Nous retrouvons en effet les mêmes élèves dans le top 10 des mesures de centralité. Ces élèves sont structurellement les noeuds les plus centraux, les plus fréquemment visités et donc de potentiels super-diffuseurs.

### 7.2 Mean first passage

Les temps de premier passage moyen entre catégories ont été estimés en simulant des marches aléatoires sur les deux graphes. La heatmap (voir annexe F.2) présente le nombre moyen d'étapes nécessaires pour atteindre un nœud d'une catégorie cible à partir d'un noeud d'une catégorie source. Pour le réseau hospitalier, les infirmières sont atteintes le plus rapidement depuis toutes les catégories, reflétant leur rôle central et leur forte connectivité. Les médecins montrent des valeurs plus élevées, ce qui confirme leur position plus périphérique dans la dynamique de diffusion. Viennent ensuite les patients et le personnel administratif.

Pour le réseau scolaire, les classes de 1A, 3A et 3B sont celles contenant les élèves atteints le plus rapidement depuis les autres classes. Comme nous l'avons observé plus haut, notamment en étudiant les mesures de centralité, sept des dix élèves ayant le degré le plus élevé sont justement issus de ces classes de 3ème.

### 7.3 Markov Stability

Nous avons évalué la robustesse de la partition Louvain sur les données de l'hôpital via la stabilité de Markov en observant la rétention d'un marcheur aléatoire dans les communautés au fil du temps (annexe F.3). La stabilité est élevée pour les temps courts puis décroît de manière régulière, ce qui indique que les communautés détectées sont cohérentes localement mais se mélangent rapidement à plus grande échelle. Ce comportement est attendu compte tenu de la densité du réseau hospitalier.

Pour le réseau scolaire, le constat est un peu différent : la partition détectée par Louvain est très cohérente à petite échelle, ce qui montre encore que les communautés sont bien définies localement. La décroissance est moins rapide que pour le réseau hospitalier, ce qui rappelle encore que la structure communautaire de l'école est plus marquée que celle de l'hôpital. Cette décroissance de la stabilité montre tout de même que la structure communautaire du réseau n'est pas robuste aux plus grandes échelles.

---

## 7.4 Infomap

Nous avons appliqué l'algorithme Infomap sur les données de l'hôpital. Contrairement à Louvain, Infomap a produit une seule communauté rassemblant l'ensemble des nœuds. Ce résultat indique que le réseau hospitalier est trop dense et trop transversal pour contenir des sous-structures dynamiques distinctes : un marcheur aléatoire se propage rapidement dans tout le réseau, sans rester confiné dans un sous-groupe particulier. Cette absence de modularité dynamique est cohérente avec la faible modularité globale et la forte densité observées précédemment.

En revanche, l'application de cet algorithme au réseau scolaire (jour 1) donne des résultats très proches de ceux produits par l'algorithme de Louvain. En effet, Infomap génère 6 communautés qui sont, pour 4 d'entre-elles, identiques aux communautés de Louvain. La différence réside dans la séparation de deux classes, qui étaient regroupées avec Louvain (les classes de 1A et 2A).

## 7.5 Walktrap

L'algorithme Walktrap, fondé sur la similarité des trajectoires issues de random walks courts, identifie principalement une seule grande communauté contenant 67 nœuds sur 75. Deux petits groupes résiduels (6 nœuds et 2 nœuds) apparaissent, mais ils correspondent surtout à des sous-ensembles faiblement connectés ou très localement liés, sans réelle signification structurelle. Ce résultat est cohérent avec celui obtenu via Infomap : la densité très élevée du réseau hospitalier empêche la formation de communautés bien séparées.

Pour le réseau scolaire, le premier jour, Walktrap identifie 3 communautés, présentant une certaine cohérence avec celles identifiées par Louvain. La communauté regroupant les classes de 4A, 5A et 5B reste la même qu'avec Louvain et Infomap. Walktrap fusionne trois communautés qui étaient distinctes pour les algorithmes testés précédemment : les communautés 0, 1 et 3 de Louvain (qui sont les communautés 0, 1, 3 et 5 de Infomap). La communauté formée par les classes 1B et 2B reste la même pour les 3 algorithmes.

L'application de l'algorithme Walktrap au deuxième jour de mesures donne des résultats très différents : il propose cette fois 6 communautés. Pour deux des classes, les élèves sont partagés entre 2 communautés différentes.

## 7.6 Core-periphery

La décomposition en k-cores met en évidence une structure cœur-périmétrie très nette dans le réseau hospitalier. Le  $k$ -core maximal ( $k = 22$ ) regroupe 39 nœuds, majoritairement des infirmières (21) et des médecins (8), ce qui souligne leur rôle central dans la structure d'interaction. Quelques administratifs (4) et patients (6) apparaissent également dans ce noyau dense, reflétant leur forte connectivité locale. La périphérie est constituée de patients et de quelques membres du personnel moins connectés, répartis dans des k-cores de plus faible ordre. Cette organisation est cohérente avec le fonctionnement d'un

---

service hospitalier, où le personnel soignant constitue le cœur opérationnel du réseau.

Pour le réseau scolaire, le  $k$ -core maximale ( $k = 34$ ) regroupe 104 élèves venant de toutes les classes et de tous les niveaux. Ce coeur forme un noyau très stable et dense, socialement mixte et dans lequel toutes les classes sont représentées. L'intégration entre les différentes classes est forte et le réseau est principalement porté par les élèves (aucun enseignant ne se retrouve dans le coeur). La périphérie est composée de deux sous-structures intermédiaires en termes de taille (36 et 35 individus) et de toute une série de petits groupes allant de 3 à 11 individus.

## 8 Processus épidémiques

Pour le réseau hospitalier, nous considérons un modèle de type SI, dans lequel un individu infecté peut transmettre l'infection à ses voisins au cours du temps.

Nous avons simulé un processus de propagation sur le réseau hospitalier afin d'illustrer l'impact de la structure du graphe sur la diffusion d'une infection (voir annexe G). Les simulations montrent une propagation rapide, en particulier lorsque l'infection débute à partir d'un nœud central du réseau. À l'inverse, une infection initiée depuis un nœud périphérique se diffuse plus lentement. Ces résultats sont cohérents avec la forte densité du réseau et la présence d'un noyau central identifié précédemment.

Pour étudier la propagation d'une infection sur le réseau scolaire, nous avons simulé les modèles SIS et SIR avec  $\beta = 0.3$  et  $\mu = 0.1$ , en démarrant avec cinq individus infectés, choisis soit dans le core, soit en périphérie. Le modèle SIS étudie l'évolution d'une maladie dans laquelle un individu infecté revient ensuite dans l'état susceptible, et dans le modèle SIR, un individu infecté atteint ensuite l'état rétabli.

La forte densité du réseau entraîne une diffusion très rapide : presque tous les individus sont infectés après deux pas de temps si l'infection démarre dans le core, et après trois pas de temps en périphérie. Dans le modèle SIR, l'ensemble de la population est rétabli peu après 40 pas de temps, tandis que dans le modèle SIS, le nombre d'infectés se stabilise autour de 210 individus dès trois pas de temps.

La localisation initiale a peu d'effet à moyen et long terme, ce qui s'explique par la structure compacte du réseau : diamètre de 3, distance moyenne des plus courts chemins de 1,86 et degré moyen très élevé ( $\langle k \rangle \approx 50$ ). Cette forte connectivité homogénéise la propagation, et le rapport  $\beta/\mu = 3$  place le réseau largement au-dessus du seuil épidémique, garantissant un plateau élevé d'infectés dans le modèle SIS, indépendamment du point de départ.

## 9 Comparaison des deux graphes

L'analyse comparative des réseaux hospitalier et scolaire met en évidence deux structures sociales différentes.

Le réseau hospitalier apparaît comme très dense mais faiblement modulaire. Les interactions y sont fortement entremêlées entre les différentes catégories, ce qui se traduit

---

par une modularité faible, une absence de communautés bien séparées et une diffusion très rapide des processus dynamiques. La structure cœur–périphérie y est marquée, avec un noyau central dominé par le personnel soignant.

À l'inverse, le réseau scolaire présente une organisation fortement modulaire, structurée autour des classes. Les communautés détectées correspondent largement aux groupes scolaires réels, et cette modularité se retrouve à dans l'analyse structurelle. Malgré une densité élevée, la présence de communautés bien définies est confirmée par différents algorithmes.

Ces différences illustrent l'influence majeure du contexte social sur la structure des réseaux de contacts. Là où l'hôpital favorise des interactions transversales et fonctionnelles, l'école conserve une organisation plus compartimentée, reflétant des groupes sociaux plus stables. Ces contrastes ont des implications directes sur la modélisation des processus de diffusion, en particulier pour les phénomènes épidémiques.

## 10 Comparaison avec les conclusions des articles

L'étude associée au jeu de données hospitaliers [1] se concentre principalement sur l'analyse de processus épidémiques et sur la réduction optimale du graphe en exploitant le caractère temporel et pondéré des contacts. À l'inverse, notre analyse adopte une approche graph mining structurelle, en travaillant sur un graphe agrégé et non pondéré.

Malgré ces différences, les conclusions sont cohérentes. Le réseau hospitalier apparaît comme dense, faiblement modulaire, dominé par un noyau d'individus fortement connectés, principalement issus du personnel soignant et administratif. Ceci est compatible avec les résultats de l'article mettant en évidence une diffusion rapide des processus épidémiques.

L'article concernant le réseau scolaire [2] s'attache principalement à décrire et comprendre les nombres et durées de contacts par classe ainsi que les interactions entre les différentes classes et l'évolution de ces mesures au cours de chaque journée. Notre approche, utilisant les outils de graph mining, permet d'appréhender la densité du réseau scolaire ainsi que sa structure en communautés. Nous avons aussi mis en évidence les individus qui pourraient jouer un rôle déterminant dans la propagation d'une épidémie ainsi que le nombre élevé d'élèves formant un Core au sein de l'école. Nos résultats permettent de réaliser aussi, d'une manière différente de l'article, la rapidité de propagation d'une infection sur un réseau aussi dense.

---

## Références

- [1] Philippe VANHEMS, Alain BARRAT, Ciro CATTUTO, Jean-François PINTON, Naggam KHANAFER, Corinne RÉGIS, Byeul-a KIM, Brigitte COMTE et Nicolas VOIRIN. “Estimating Potential Infection Transmission Routes in Hospital Wards Using Wearable Proximity Sensors”. In : *PLoS ONE* 8.9 (sept. 2013), e73970. DOI : 10.1371/journal.pone.0073970. URL : <http://dx.doi.org/10.1371%2Fjournal.pone.0073970>.
- [2] Juliette STEHLÉ et al. “High-Resolution Measurements of Face-to-Face Contact Patterns in a Primary School”. In : *PLOS ONE* 6.8 (août 2011), e23176. DOI : 10.1371/journal.pone.0023176. URL : <http://dx.doi.org/10.1371/journal.pone.0023176>.

## A Distributions de degrés détaillées

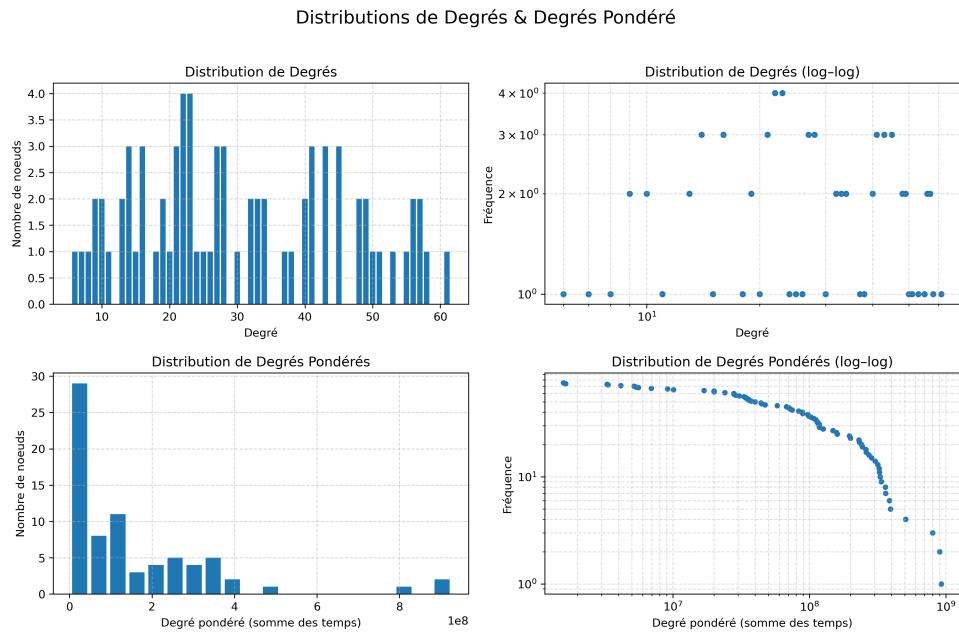


FIGURE 5 – Distributions détaillées des degrés et degrés pondérés du réseau hospitalier.

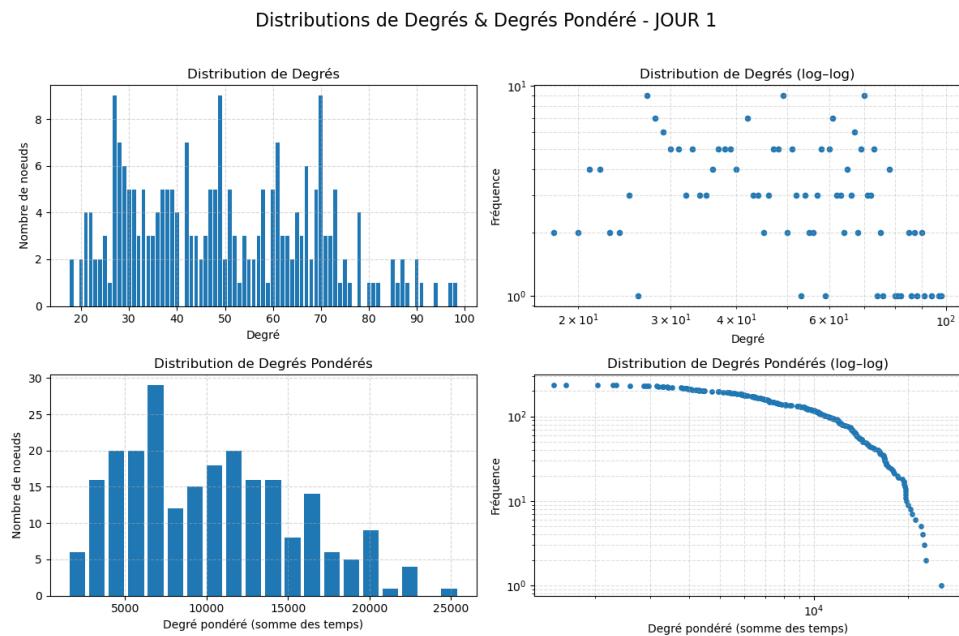


FIGURE 6 – Distributions des degrés et degrés pondérés du réseau scolaire - Jour 1

## B Centralités

Rang	Mesures de centralité									
	Degree		Betweenness		Closeness		Katz		PageRank	
	Nœud	Valeur	Nœud	Valeur	Nœud	Valeur	Nœud	Valeur	Nœud	Valeur
1	1098 (ADM)	61	1098 (ADM)	0.0404	1098 (ADM)	0.8506	1098 (ADM)	0.1326	1098 (ADM)	0.0254
2	1193 (NUR)	58	1193 (NUR)	0.0352	1193 (NUR)	0.8222	1193 (NUR)	0.1309	1193 (NUR)	0.0241
3	1115 (NUR)	57	1164 (NUR)	0.0313	1115 (NUR)	0.8132	1164 (NUR)	0.1304	1164 (NUR)	0.0237
4	1164 (NUR)	57	1115 (NUR)	0.0287	1164 (NUR)	0.8132	1115 (NUR)	0.1304	1115 (NUR)	0.0236
5	1295 (NUR)	56	1109 (NUR)	0.0278	1295 (NUR)	0.8043	1210 (NUR)	0.1299	1295 (NUR)	0.0232
6	1210 (NUR)	56	1295 (NUR)	0.0268	1210 (NUR)	0.8043	1295 (NUR)	0.1299	1210 (NUR)	0.0232
7	1109 (NUR)	55	1210 (NUR)	0.0265	1109 (NUR)	0.7957	1109 (NUR)	0.1293	1109 (NUR)	0.0228
8	1157 (MED)	53	1658 (ADM)	0.0227	1157 (MED)	0.7789	1157 (MED)	0.1283	1157 (MED)	0.0219
9	1658 (ADM)	51	1190 (NUR)	0.0222	1658 (ADM)	0.7629	1658 (ADM)	0.1271	1658 (ADM)	0.0214
10	1144 (MED)	50	1157 (MED)	0.0217	1144 (MED)	0.7551	1144 (MED)	0.1266	1144 (MED)	0.0208

TABLE 1 – Top 10 des nœuds selon différentes mesures de centralité pour le graphe de l'hôpital

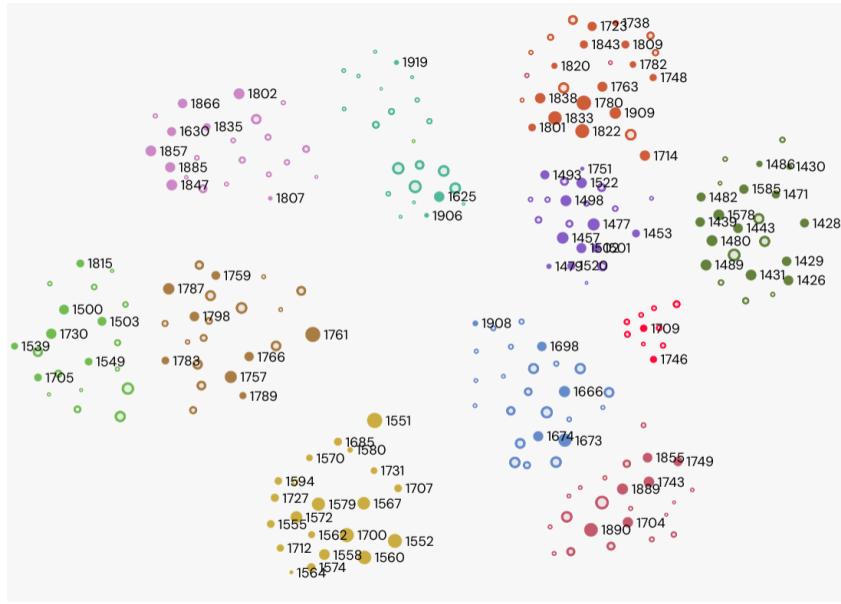


FIGURE 7 – Les voisins du Hub 1551 : les tailles des noeuds sont proportionnelles à leur degré, et ici les voisins de 1551 sont représentés par des boules pleines - Jour 1

## C Modèles de réseau

Indicateur	Graphe réel	ER	Configuration	BA
Nombre de noeuds	75	75	75	75
Nombre d'arêtes	1139	1158	837	900
Densité	0.410	0.417	0.302	0.324
Clustering moyen	0.640	0.424	0.375	0.448
Longueur moyenne des plus courts chemins	1.598	1.583	1.711	1.677

TABLE 2 – Comparaison des propriétés du graphe hospitalier réel et des modèles aléatoires

Indicateur	Graphe réel	ER	Configuration	BA
Nombre de noeuds	236	236	236	236
Nombre d'arêtes	5899	5868	5162	5088
Densité	0.213	0.212	0.186	0.183
Clustering moyen	0.502	0.213	0.226	0.283
Longueur moyenne des plus courts chemins	1.860	1.788	1.818	1.819

TABLE 3 – Comparaison des propriétés du graphe scolaire réel et des modèles aléatoires - Jour 1

## D Communautés

	ADM	MED	NUR	PAT	Total
1	0	0	10	11	21
3	3	0	10	8	21
0	3	11	1	5	20
2	2	0	6	5	13

TABLE 4 – Comparaison de la composition des différentes communautés détectées par l’algorithme de Louvain

	1A	1B	2A	2B	3A	3B	4A	4B	5A	5B	Teachers	Total
2	0	0	0	0	0	0	21	0	22	23	3	69
4	0	25	0	25	0	0	0	0	0	0	2	52
0	22	0	22	0	0	0	0	0	0	0	2	46
1	0	0	0	0	23	21	0	0	0	0	2	46
3	0	0	0	0	0	0	0	22	0	0	1	23

TABLE 5 – Comparaison de la composition des différentes communautés détectées par l’algorithme de Louvain - Jour 1

## E Dynamiques

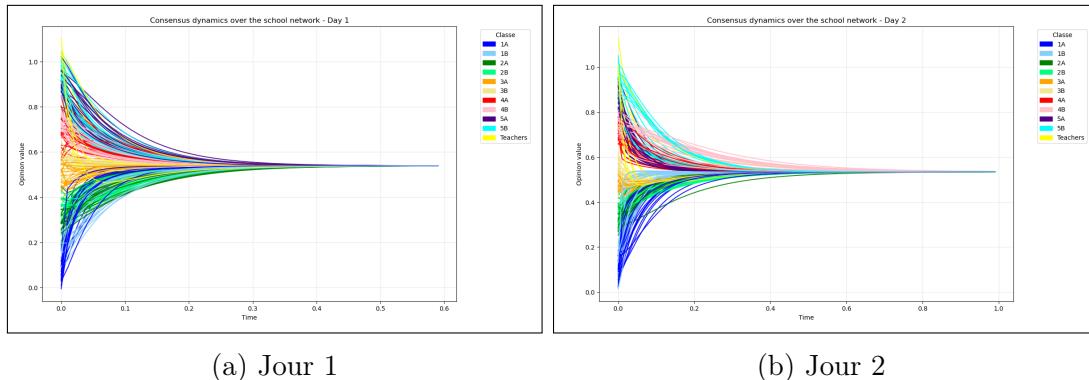


FIGURE 8 – Dynamique de Consensus - Réseau scolaire

## F Marches aléatoire

### F.1 Marche aléatoire stationnaire

Rang	Nœud	Catégorie	Probabilité stationnaire
1	1098	ADM	0.02678
2	1193	NUR	0.02546
3	1115	NUR	0.02502
4	1164	NUR	0.02502
5	1295	NUR	0.02458
6	1210	NUR	0.02458
7	1109	NUR	0.02414
8	1157	MED	0.02327
9	1658	ADM	0.02239
10	1144	MED	0.02195

TABLE 6 – Top 10 des noeuds selon la distribution stationnaire de la marche aléatoire - Réseau hospitalier

Rang	Nœud	Catégorie	Probabilité stationnaire
1	1551	3B	0.008306
2	1761	1A	0.008222
3	1780	3A	0.007967
4	1700	3B	0.007713
5	1822	3A	0.007628
6	1552	3B	0.007628
7	1890	2B	0.007459
8	1560	3B	0.007374
9	1833	3A	0.007374
10	1708	2B	0.007289

TABLE 7 – Top 10 des noeuds selon la distribution stationnaire de la marche aléatoire - Réseau scolaire - Jour 1

### F.2 Mean first passage - heatmap

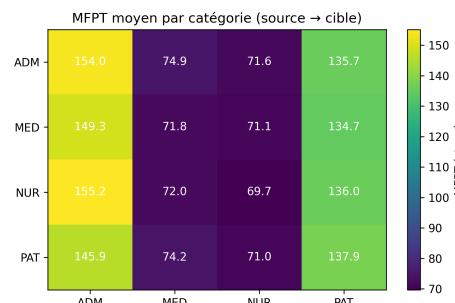


FIGURE 9 – Moyenne de premier passage par catégorie - Réseau hospitalier

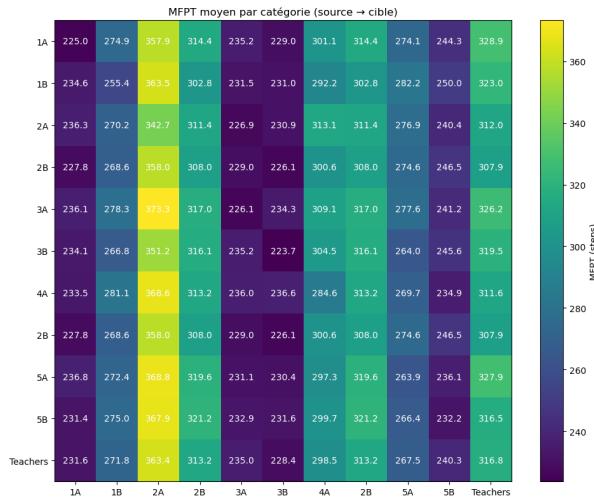


FIGURE 10 – Moyenne du premier passage par catégorie - Réseau scolaire - Jour 1

### F.3 Markov Stability

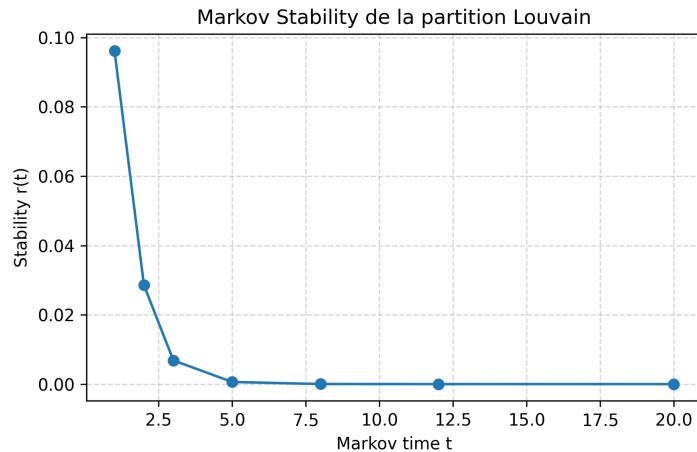


FIGURE 11 – Stabilité de Markov - Réseau hospitalier

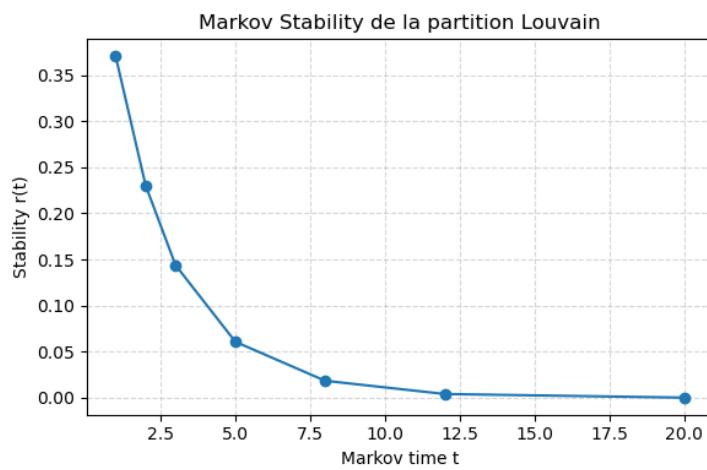


FIGURE 12 – Stabilité de Markov - Réseau scolaire - Jour 1

## F.4 Walktrap

	1A	1B	2A	2B	3A	3B	4A	4B	5A	5B	Teachers	Total
0	22	0	22	0	23	21	0	22	0	0	5	115
1	0	0	0	0	0	0	21	0	22	23	3	69
2	0	25	0	25	0	0	0	0	0	0	2	52

TABLE 8 – Comparaison de la composition des différentes communautés détectées par Walktrap - Réseau scolaire -Jour 1

	1A	1B	2A	2B	3A	3B	4A	4B	5A	5B	Teachers	Total
0	23	0	23	0	0	0	0	0	0	0	2	48
1	0	0	0	0	23	21	0	0	0	0	2	46
2	0	25	0	0	0	0	21	7	21	13	3	90
3	0	0	0	0	0	0	0	15	0	0	1	16
4	0	0	0	26	0	0	0	0	0	0	1	27
5	0	0	0	0	0	0	0	0	0	10	1	11

TABLE 9 – Comparaison de la composition des différentes communautés détectées par Walktrap - Réseau scolaire - Jour 2

## G Processus épidémique

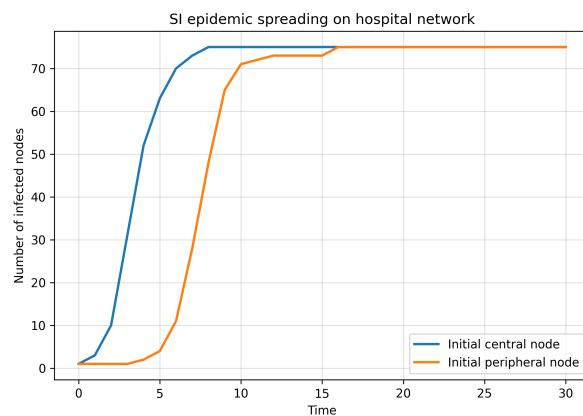


FIGURE 13 – Modèle SI - Réseau hospitalier

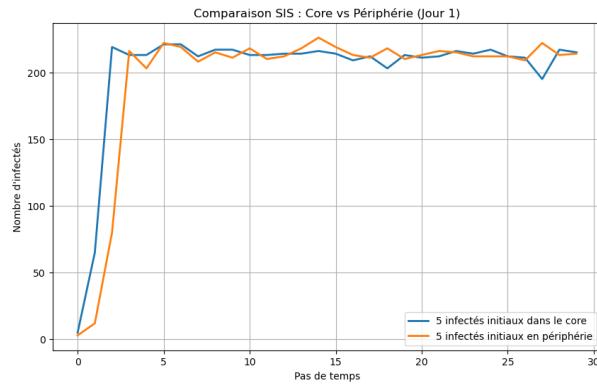


FIGURE 14 – Modèle SIS - Réseau scolaire - Jour 1 : comparaison Core-Periphery

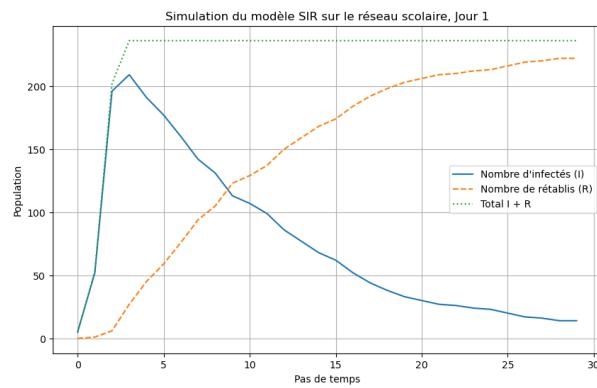


FIGURE 15 – Modèle SIR -Réseau scolaire - Jour 1 : infectés initiaux aléatoires

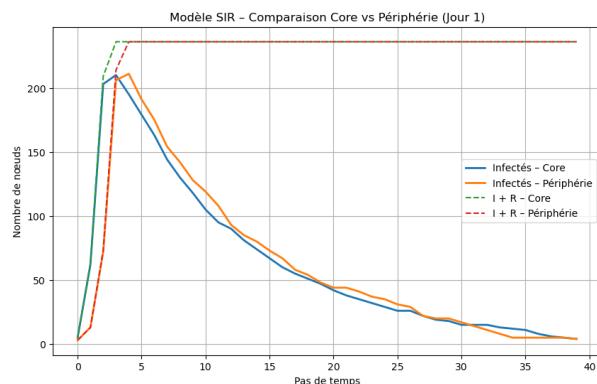


FIGURE 16 – Modèle SIR, réseau scolaire : comparaison Core-Periphery