

UNIVERSITÉ DE PARIS
UFR MATHÉMATIQUES ET INFORMATIQUE

Finding remarkably dense sequences of contacts in link streams for COVID-19

Master 1 Réseaux et Systèmes Autonomes (RSA)

Idriss KOURBANHOUSSEN – Ounas YHIA – Thinujan THILAHENDRAN

Encadré par Chuan Li – Vincent Gauthier – Hassine Moun gla

Année universitaire 2022 – 2023

Table des matières

1.	INTRODUCTION	3
2.	MÉTHODE ÉTUDIÉE	4
3.	MISE EN PLACE DE LA MÉTHODE SUR UN JEU DE DONNÉES	5
3.1	ÉTAPE 1 : CRÉATION DU GRAPHE	5
3.2	ÉTAPE 2 : DÉTECTION DE COMMUNAUTÉS DANS LE GRAPHE	7
3.3	ÉTAPE 3 : CALCULE DE LA DENSITÉ DE CHAQUE COMMUNAUTÉ ET DÉTECTION DES GROUPES PERTINENTS	8
4.	EXEMPLES D'APPLICATIONS CONCRÈTE DES RÉSULTATS	14
5.	CONCLUSION	15
6.	BIBLIOGRAPHIE	16

1. Introduction

Dans le cadre de notre première année de master informatique, nous avons entrepris un projet tuteuré axé sur la recherche, visant à identifier les groupes de liens structurellement et temporellement denses d'un ensemble de données de contact tracing.

La lutte contre la pandémie de coronavirus a mis la société face à un défi sans précédent, nécessitant des mesures massives telles que les tests à grande échelle, l'interdiction des rassemblements, et plus généralement la mise en place de restrictions importantes de mobilité pour retarder la propagation du virus. La situation a été particulièrement critique dans les grandes villes, où la densité de population et une forte mobilité favorisent la propagation des maladies. Dans ce contexte, la compréhension de la dynamique de propagation est cruciale pour prendre des décisions politiques éclairées de la part des autorités publiques afin de réduire le nombre de cas.

Les informations à grande échelle sur le taux de contact dans la population générale ont été identifiées comme le paramètre le plus pertinente pour caractériser la propagation d'une épidémie telle que la COVID-19. Dans ce projet, nous avons à disposition des données de contact tracing et souhaitons capturer les groupes spécifiques de liens qui sont fortement connectés temporellement et structurellement, à l'aide d'une méthode spécifique que nous allons présenter un peu plus tard dans ce rapport. Ladite méthode est tirée de l'article "*Finding remarkably dense sequences of contacts in link streams*".

Les résultats de l'étude montrent que cette méthode peut être utilisée pour obtenir une description globale du taux de contact au fil du temps mais aussi sa fluctuation en fonction de la mobilité de la population.

2. Méthode étudiée

Notre projet se base sur les recherches et la méthode décrite par Noé Gaumont, Clémence Magnien et Matthieu Latapy dans l'article « *Finding remarkably dense sequences of contacts in lien streams* ». Cet article présente une méthode de détection des groupes denses dans un flux de liens. Les groupes denses sont des ensembles de nœuds qui ont un grand nombre de connexions entre eux.

Cette méthode se divise en trois étapes :

- La première étape consiste à projeter le flux de liens de nos données temporelle de contact tracing sur un graphe non orienté :
 - Chaque lien est représenté par un nœud sur le graphe
 - Deux nœuds du graphe sont reliés si leur lien respectif partage un nœud en commun et que leurs intervalles de temps se chevauchent
- La deuxième étape consiste à appliquer un algorithme de détection de communauté sur la projection obtenue afin d'obtenir une partition des liens.
- Enfin, la partition obtenue est filtrée pour ne conserver que les groupes denses, qui ont une densité plus élevée que leurs voisins et sont suffisamment grands. Nous considérons que les groupes très petits ont un intérêt limité en termes de description du flux de liens. Ces groupes denses sont considérés comme des groupes significatifs dans le flux de liens.

Pour illustrer cette méthode, prenons l'exemple des utilisateurs d'un réseau social : le flux de liens représente ici leurs interactions et chaque lien du flux s'apparente à une interaction entre deux de ses utilisateurs dans l'intervalle de temps $[t, t+x]$. Ces liens peuvent être projetés sur un graphe, où chaque lien est représenté par un nœud. Deux nœuds du graphe sont reliés si leur lien respectif partage un utilisateur en commun et que leurs intervalles de temps se chevauchent. Par la suite, on applique l'algorithme de détection de communauté pour identifier les groupes de liens qui sont étroitement liés les uns aux autres. Enfin, les groupes denses peuvent en être filtrés pour ne conserver que les communautés significatives d'utilisateurs du réseau qui interagissent de manière cohérente.

Étant donné la grande quantité de données qui aurait été nécessaire pour la réalisation de ce projet si nous avions voulu traiter des données réelles liées au contact tracing de la Covid-19, il nous a été demandé par nos responsable d'utiliser un jeu de données plus simple (celui d'une école primaire de Lyon), disponible gratuitement sur le lien suivant : <http://www.sociopatterns.org/datasets/primary-school-temporalnetwork-data/>. Ce jeu de données contient un réseau temporel de contacts entre les enfants et les enseignants de cette école. Malgré une quantité de données moindre, le principe et l'application de la méthode décrite dans l'article reste la même.

3. Mise en place de la méthode sur un jeu de données

3.1 Étape 1 : Création du graphe

Le jeu de données que nous avons utilisées est un réseau temporel de contacts entre les enfants et les enseignants d'une école primaire de Lyon. Ce dataset a déjà été utilisé à plusieurs reprises et notamment dans l'étude « *Mitigation of infectious disease at school: targeted class closure vs school closure* » publiée dans le journal BMC Infectious diseases en 2014 ; ce qui en démontre sa qualité.

Chaque ligne de ce dataset correspond à un contact entre deux individus pendant un intervalle de temps de 20 secondes. Les lignes de ce fichier sont de la forme $\{t \mid i \mid j \mid C_i \mid C_j\}$ où i et j sont les IDs anonyme des personnes en contact, C_i et C_j sont leur classe et l'intervalle de temps durant lequel ces contact ont été actif est $[t - 20s, t]$, où t est le temps mesuré en secondes. Si plusieurs contacts ont été actifs pendant le même intervalle, plusieurs lignes commençant par la même valeur de t sont présentes dans le fichier. Au total, on dénombre 125773 lignes dans le dataset.

Avant toute chose, il est important de noter que nous n'avons pas pris en considération les colonnes C_i et C_j de notre dataset car elles sont inutiles pour l'élaboration de notre graphe.

Pour simplifier la réalisation de ce graphe, nous avons fait le choix de créer un dictionnaire rassemblant l'ensemble des informations pour chaque contact du flux de lien ; à savoir : l'heure de début du contact, l'heure de fin, et les deux individus concernés par ce contact. Notre dictionnaire est structuré de la façon suivante : $\{(t_début, i, j) : t_fin, \dots\}$.

Comme illustré dans le tableau ci-après (Fig. 1), deux individus i et j donné peuvent être en contact pendant un certain temps, avant de se séparer et de se retrouver à un autre moment de la journée. Prenons comme exemple l'individu A et B : ils étaient en contact du temps T20 à T60 puis de T100 à T120. Nous aurons donc pour les mêmes individus A et B deux éléments différents dans le dictionnaire (représentant deux contacts différents).

T	I	J
20	A	B
40	A	B
60	A	B
70	B	C
80	E	G
100	A	B
120	A	B

Fig. 1 : Exemple d'un flux de liens.

Une fois notre dictionnaire constitué, nous avons pu faire la projection de notre flux sur un graphe non orienté que nous appellerons le graphe de liens. Ce dernier a pu être construit facilement car chaque élément du dictionnaire correspond à un nœud dans le graphe selon la définition donnée un peu plus tôt.

Enfin, pour achever le graphe, il ne nous restait plus qu'à créer les liens existants entre les nœuds de ce dernier. D'après la définition donnée un peu plus tôt, deux nœuds du graphe sont liés si leur lien correspondant dans le flux de lien partage un nœud en commun et se chevauchent temporellement. On parle alors de chevauchement temporelle et structurelle.

Comme on peut le voir sur l'image ci-après (Fig.2), le lien vert (a, b) et le lien jaune (b, c) sont reliés graphiquement par une arête sur le graphe car ils ont un nœud en commun qui est le nœud « b » et se chevauchent dans le temps. Ce qui n'est pas le cas pour les liens jaune (b, c) et bleu (c, d), car même s'ils partagent un nœud en commun ils ne se chevauchent pas temporellement.

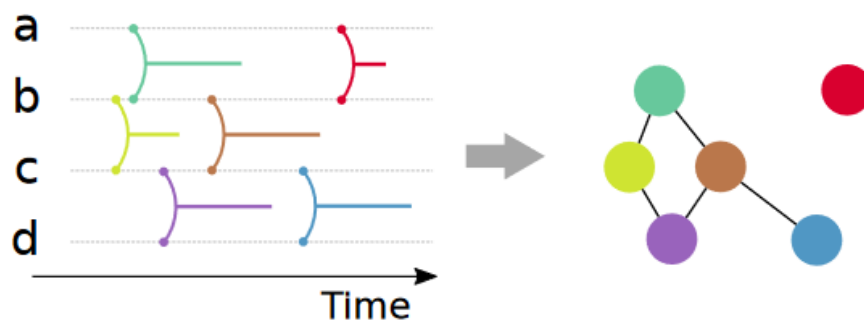


Fig. 2 : Projection d'un flux de lien possédant 4 nœuds sur un graphe

Nous avons implémenté cette méthode sur notre dataset, et avons obtenu un graphe possédant 93581 nœuds et 59211 liens.

3.2 Étape 2 : Détection de communautés dans le graphe

Cette deuxième étape consiste à appliquer un algorithme de détection de communauté sur le graphe obtenu dans l'étape 1. Il existe de nombreuses méthodes de détection de communauté mais l'une des plus connues et qui était recommandée par l'article est la méthode dite de Louvain.

Nous avons appliqué cette méthode à notre graphe et avons obtenu une partition des liens en 51198 communautés.

La méthode de Louvain est l'une des techniques les plus utilisées pour la détection de communautés dans les graphes non orientés. Elle vise à regrouper les nœuds du graphe en communautés denses qui ont des connexions internes fortes et des connexions plus faibles avec le reste du graphe.

Voyons comment la méthode de Louvain fonctionne pour la détection de communauté dans un graphe non orienté :

- **Initialisation** : Chaque nœud du graphe est initialement considéré comme une communauté distincte.
- **Phase ascendante** : Pour chaque nœud, on examine ses voisins et calcule une mesure appelée "modularité" qui représente la qualité de la partition actuelle du graphe. La modularité mesure à quel point les nœuds d'une communauté sont étroitement liés les uns aux autres par rapport aux liens pris au hasard.
- **Optimisation de la modularité** : On répète sur chaque nœud et on essaie de déplacer ce nœud dans une autre communauté pour maximiser la modularité globale. On calcule la variation de modularité que le déplacement de ce nœud entraînerait s'il était déplacé dans la communauté de l'un de ses voisins. Le nœud est ensuite placé dans la communauté qui donne la plus grande augmentation de modularité. Ce processus est répété pour tous les nœuds jusqu'à ce qu'aucun déplacement ne conduise à une augmentation de la modularité.
- **Phase descendante** : Ces 3 dernières étapes sont répétées encore deux ou trois fois en considérant une communauté comme un nœud à part entière. Cela va nous permettre d'avoir une meilleure partition du graphe.

3.3 Étape 3 : Calcul de la densité de chaque communauté et détection des groupes pertinents

CALCUL DE LA DENSITE DE CHAQUE COMMUNAUTE

La densité d'un groupe de liens est utilisée pour déterminer dans quelle mesure tous les couples de nœuds de ce groupe sont connectés tout le temps. La mesure de la densité a été adaptée aux flux de liens où chaque lien est représenté par un quadruplet (b, e, u, v), indiquant qu'il existe un lien entre les nœuds u et v dans l'intervalle de temps [b, e].

La formule de densité pour un ensemble de nœuds V' dans un intervalle de temps $[t, t + \delta]$ est la suivante :

$$d(V', t, \delta) = \frac{1}{|V'| \cdot (|V'| - 1)} \sum_{(u,v) \in V' \times V', u \neq v} \frac{\tau_{t,\delta}(u, v)}{\delta},$$

- $|V'|$ est le nombre de nœuds dans V' ,
- $\sum_{(u,v) \in V' \times V', u \neq v} \tau_{t,\delta}(u, v)$ est la somme des durées de chevauchement des liens entre u et v dans l'intervalle de temps $[t, t + \delta]$,
- δ est la durée de l'intervalle de temps $[t, t + \delta]$.

Cette formule calcule la densité en considérant toutes les paires de nœuds dans l'ensemble V' et en prenant en compte la durée de contact entre ces paires dans l'intervalle de temps spécifié.

Plus concrètement, pour calculer la densité de nos groupes de liens, nous avons suivis les étapes suivantes :

1. Créer un dictionnaire ("group_links") qui répertorie l'ensemble des nœuds pour chaque groupe de la partition. Ce dictionnaire sera de la forme { groupe_id : nœud(s) du groupe }
2. Parcourir ce dictionnaire et pour chaque groupe, initialiser une variable de densité à zéro et compter le nombre de nœud présents dans le groupe.
3. Pour chaque paire de nœuds dans un groupe, initialiser des variables pour stocker la somme des durées de chevauchement et le nombre de chevauchements entre les liens (correspondant au nœud) dans le flux de liens.
4. Une fois que toutes les paires de nœuds ont été traitées, vérifier si le nombre de chevauchements est supérieur à zéro. Si c'est le cas, calculer la densité en

divisant la somme des durées de chevauchement par le nombre de chevauchements.

5. Enfin, pour tenir compte de la taille du groupe, le code divise la densité par le nombre total de paires de liens possibles dans le groupe ($\text{link_count} * (\text{link_count} - 1)$).

→ La densité calculée pour chaque groupe est ensuite stockée dans un dictionnaire appelé "densities", où la clé est l'ID du groupe.

En résumé, notre code calcule la densité de chaque groupe de la partition en déterminant la quantité de chevauchements entre les liens présents dans le groupe et en ajustant cette mesure en fonction de la taille du groupe.

Pour finir, afin d'avoir une densité qui soit comprise entre 0 et 1, nous avons implémenté une fonction qui normalise l'ensemble des densités par rapport aux densités des autres groupes de communautés.

DETECTION DES GROUPES PERTINENTS

Arrivé à cette étape, nous avons capturé un grand nombre de groupes (exactement 51198) et possédons pour chacun d'entre eux la valeur de leur densité. Comme évoqué en amont, il est nécessaire maintenant de sélectionner les groupes que l'on considère les plus pertinents. En effet, les différents groupes découverts par la méthode de Louvain ne sont pas tous pertinents. Par exemple, certains d'entre eux sont de très petite taille (1 nœud seulement), et n'ont donc pas d'intérêt analytiquement parlant d'après les auteurs de l'article. Par conséquent, il est nécessaire de mettre en place une méthode pour valider ou rejeter les groupes de notre partition. Dans l'ensemble des groupes découvert au moment du partitionnement, que l'on appellera maintenant des candidats, la première chose qu'il peut être judicieux de faire est de rejeter tous ceux qui ne possèdent qu'un seul nœud. Dans notre code, nous avons fait en sorte que tous les candidats possédant un unique nœud se voit affecté une densité de 0. Ainsi le filtrage est largement facilité.

Ensuite, l'article présente une méthode précise pour déterminer les candidats pertinents. En effet, la densité seul de ceux-ci ne suffit pas. L'idée derrière cette méthode est de pouvoir définir un seuil au-delà duquel les candidats seraient automatiquement considérés comme pertinents. Pour faire cela, l'article explique qu'il faut comparer la densité de chaque candidat avec celle de ses voisins dans le flux de lien. Les « voisins » sont par définition les groupes similaires au candidat comparé, mais différents seulement dans un aspect (Node set : p_{node} , start time : p_t ou duration : $p\delta$). Une densité plus élevée que celle des voisins indique un groupe plus dense et potentiellement plus pertinent.

Plus couramment, on parle de score et on peut les définir ainsi :

- Score du voisinage des nœuds (p_{node}) : Il mesure la densité d'un groupe de liens candidat par rapport à la densité de ses voisins dans le graphe de liens. Un score plus élevé indique que le groupe est plus dense que ses voisins.
- Score du voisinage du temps de début (p_t) : Il évalue la densité d'un candidat en comparant sa densité avec celle de ses voisins dans le voisinage du temps de début. Il mesure à quel point le groupe est dense par rapport à d'autres groupes similaires dans le voisinage temporel.
- Score du voisinage de la durée (p_δ) : Il mesure la densité d'un candidat en comparant sa densité avec celle de ses voisins dans le voisinage de la durée. Il évalue la densité du groupe par rapport à d'autres groupes similaires avec des durées voisines.

Une fois les 3 scores calculés pour chaque candidat, l'article nous explique qu'il faut définir un seuil minimal par type score. Ce seuil est déterminé en traçant la courbe de distribution inverse dudit score et en observant le moment sur le graphique où il y a « courbure ». Une fois les seuils de chaque score fixé, il ne reste plus qu'à comparé les valeurs de scores de chaque candidat avec les valeurs « seuils » et les rejeter si les trois scores du candidat ne dépassent pas les seuils fixés.

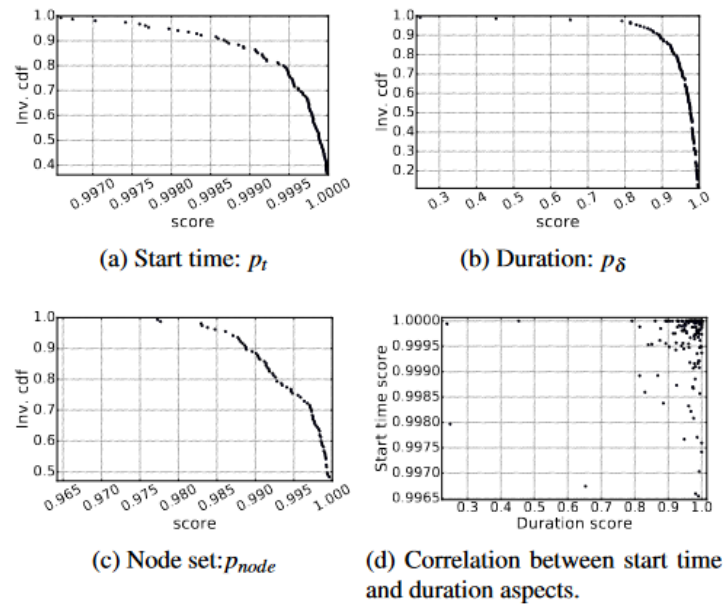


Fig. 6: Inverse cumulative distribution of scores for each aspect for the data set Socio Pattern.

Fig. 3 : Exemple extrait de l'article. Les seuils pour start time, duration et node aspect ont été fixés à : 0.998, 0.85 and 0.98

En résumé, les groupes pertinents sont déterminés en réalisant des comparaisons complexes de densité avec les voisins de chaque candidat (à partir du calcul des scores). L'objectif est d'identifier les groupes de liens qui sont à la fois denses et qui sont significatif dans le contexte global du flux de liens.

La méthode décrite ci-dessus est la méthode conseillée dans l'article pour sélectionner les candidats les plus pertinents. Malheureusement, nous ne sommes pas parvenus à calculer les scores p_t et p_δ , ce qui nous a contraint à déterminer les groupes pertinents seulement à partir des valeurs de densité des candidats. Cette valeur seule n'est bien évidemment pas suffisante pour déterminer les groupes pertinents mais représente tout de même une aide.

Ainsi, nous nous sommes dit que nous pouvions ne conserver que les 50% meilleurs candidats de notre partition. Autrement dit les 50% meilleurs en termes de valeurs de densité. Ensuite, à partir de là, nous pouvons filtrer ceux qui possèdent plus de 2 nœuds par exemple. Cela étant un choix arbitraire, nous pouvons penser intuitivement que plus un candidat possède de nœud et plus il est pertinent et dense. Enfin, le nombre de liens exacte composant un candidat peut également être prit en compte pour évaluer sa pertinence.

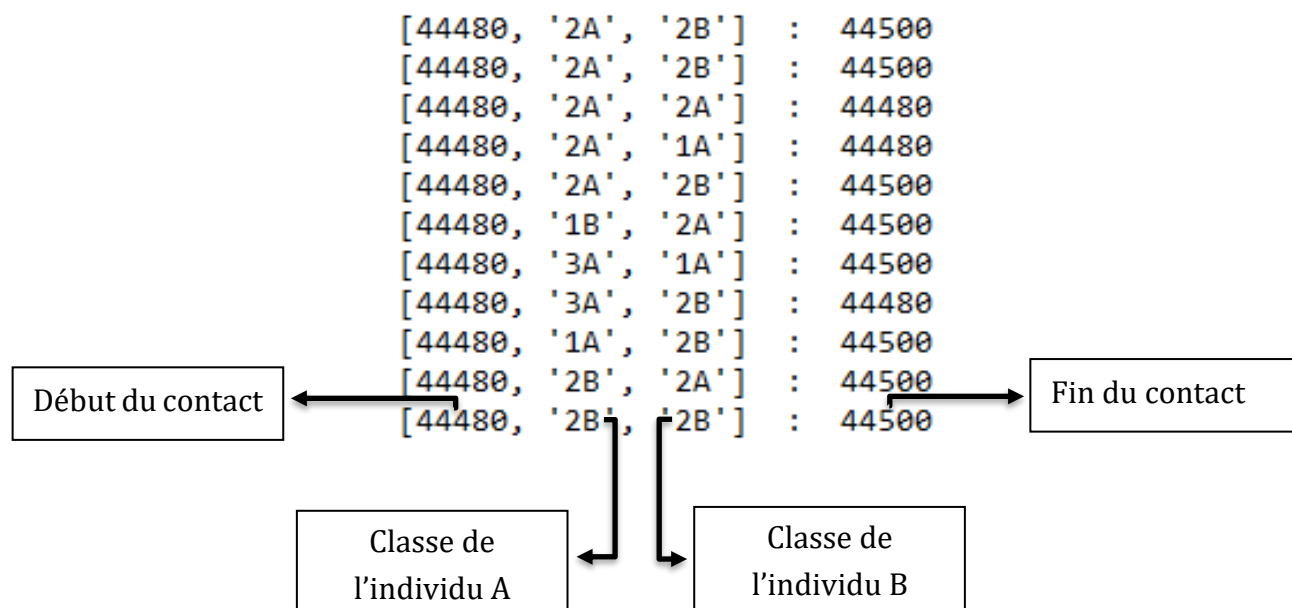
En ne sélectionnant que les candidats ayant une densité parmi les 50% les plus élevés, il n'en reste plus que 1972.

ANALYSE D'UN GROUPE PERTINENT

Parmi les 50% des candidats ayant la densité la plus élevée, on remarque qu'un de nos candidat (le n° 11713 au moment de l'exécution de notre code) est celui qui possède le plus grand nombre de nœud. En effet, on n'en dénombre pas moins de 11 dans le sous graphe correspondant à ce groupe. En outre, il concentre au total 36 liens, ce qui est un indicateur de densité. En regardant de plus près les heures exacte de début et de fin de chaque contact (donc ceux de chaque flux de liens du groupe) on remarque ceci :

Contact n° 0	Heure de début :	13:21:20	Heure de fin :	13:21:40
Contact n° 1	Heure de début :	13:21:20	Heure de fin :	13:21:40
Contact n° 2	Heure de début :	13:21:20	Heure de fin :	13:21:20
Contact n° 3	Heure de début :	13:21:20	Heure de fin :	13:21:20
Contact n° 4	Heure de début :	13:21:20	Heure de fin :	13:21:40
Contact n° 5	Heure de début :	13:21:20	Heure de fin :	13:21:40
Contact n° 6	Heure de début :	13:21:20	Heure de fin :	13:21:40
Contact n° 7	Heure de début :	13:21:20	Heure de fin :	13:21:20
Contact n° 8	Heure de début :	13:21:20	Heure de fin :	13:21:40
Contact n° 9	Heure de début :	13:21:20	Heure de fin :	13:21:40
Contact n° 10	Heure de début :	13:21:20	Heure de fin :	13:21:40

Le premier contact débute à 13h21 et 20s et le dernier se termine à 13h21 et 40s ; ce qui est un intervalle de temps très court. On peut faire l'hypothèse que cela est survenue pendant un intercoures. Cependant, quand on regarde plus attentivement les classes auxquelles font parties les élèves du groupe pertinent on observe ceci :



Il y a un total de 5 classes différentes impliquées dans ce groupe (les classes 2A, 2B, 1A, 1B, et 3A). Leurs contacts étant de très courte durée et aux alentours de 13h20, on peut supposer qu'il s'agissait probablement de la fin de la pause midi, au moment où les élèves regagnent leurs classes et se croisent forcément dans les couloirs.

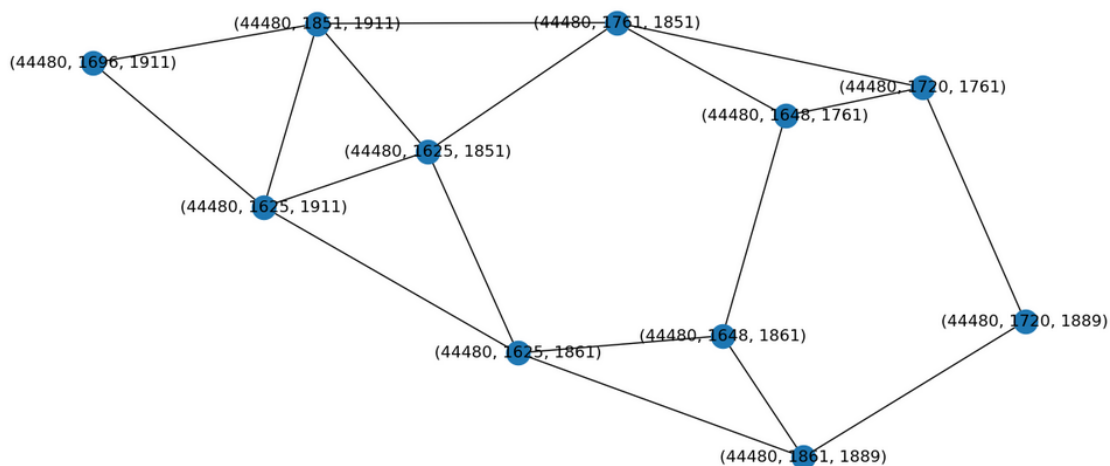


Fig. 4 : Sous-graphe non orienté du groupe pertinent en question

Il est important de rappeler que notre analyse n'a pas porté sur le groupe qui serait « le plus » pertinent mais plutôt sur un groupe que nous avons jugé suffisamment intéressant au vu du nombre de nœud qui le compose, du nombre de lien et de sa densité jugée suffisante.

Au moment d'une pandémie, dans un tel cas de figure, une mesure qui pourrait être prise serait de limiter au maximum les contacts entre les élèves de classes différentes et faire en sorte qu'ils ne se croisent pas où le moins possible, afin d'étouffer tout départ d'épidémie.

4. Exemples d'applications concrète des résultats

Comme introduit plus tôt, dans un contexte de propagation d'une maladie, l'intérêt d'identifier les groupes pertinents est multiple :

- Pouvoir compléter nos données initiales avec les caractéristiques des personnes appartenant à chaque groupe pertinents, telles que l'âge, le sexe, la profession, etc. Cela peut aider à identifier des tendances ou des facteurs de risque spécifiques dans certains groupes, et de déterminer les facteurs qui rendent certains groupes plus vulnérables à l'infection.
- Explorer la manière dont le virus s'est propagé à l'intérieur et entre les groupes. Identifiez les membres qui ont été infectés et examinez comment la transmission s'est produite (par exemple, contact direct, contact indirect, ...). Nous pouvons également analyser la vitesse de propagation et les schémas de transmission entre les groupes. Cela permet de mieux comprendre comment la maladie se propage et ainsi pouvoir prendre des mesures efficaces.
- Evaluer l'efficacité des mesures mises en place dans chaque groupe, telles que le port de masques, la distanciation sociale, le lavage des mains, etc. Le but est d'identifier quelles mesures sont les plus efficaces pour endiguer la propagation de la maladie.
- Cibler les mesures de prévention et de contrôle des infections de manière plus efficace. Autrement dit, les autorités sanitaires peuvent concentrer leurs efforts sur les personnes appartenant à ces groupes et ainsi limiter l'étendue des mesures restrictives à une partie limitée de la population.

Il est important de noter que tout cela doit se faire dans le respect de la vie privée et des données confidentielles de chaque individu. Cela implique de traiter les données de manière sécurisée, de les anonymiser si nécessaire et de limiter leur accès aux personnes autorisées uniquement.

5. Conclusion

En conclusion, notre étude repose sur la méthode décrite par Noé Gaumont, Clémence Magnien et Matthieu Latapy dans leur article intitulé « Finding remarkably dense sequences of contacts in lien streams ». Cette méthode vise à détecter les groupes denses dans un flux de liens, où les groupes denses sont définis comme des ensembles de nœuds ayant un grand nombre de connexions entre eux.

Nous avons suivi les trois étapes décrite dans l'article. Tout d'abord, nous avons projeté sur un graphe non orienté le flux de liens, où chaque lien est représenté par un nœud et où deux nœuds de ce graphe sont reliés si leur lien se chevauchent à la fois structurellement mais aussi temporellement. Ensuite, nous avons appliqué un algorithme de détection de communauté sur cette projection pour obtenir une partition des liens. Enfin, la partition obtenue est filtrée pour ne conserver que les groupes denses, ceux qui ont une densité plus élevée que leurs voisins et qui sont suffisamment grands pour être considérés comme des groupes significatifs dans le flux de liens.

Dans le cadre de ce projet, un jeu de données simplifié d'une école primaire de Lyon a été utilisé pour illustrer la méthode. Le jeu de données contient un réseau temporel de contacts entre les enfants et les enseignants de cette école.

Le rapport du projet présente les différentes étapes de la méthode et les résultats obtenus sur le jeu de données de l'école primaire de Lyon. Les groupes denses détectés dans ce contexte peuvent fournir des informations pertinentes sur les interactions entre les enfants et les enseignants, ce qui peut être utile pour analyser la propagation de maladies ou pour comprendre les dynamiques sociales dans un environnement scolaire.

6. Bibliographie

- S Contreras et al. "Low case numbers enable long-term stable pandemic control without lockdowns". en. In: Science Advances 7.41 (Oct. 2021). doi: 10.1126/sciadv.abg2243.
- N Gaumont, C Magnien, and M Latapy. "Finding remarkably dense sequences of contacts in link streams". In: Social Network Analysis and Mining 6.1 (Sept. 2016). doi: 10.1007/s13278-016-0396-z.
- J Persson, JF Parie, and S Feuerriegel. "Monitoring the COVID-19 epidemic with nationwide telecommunication data". In: Proceedings of the National Academy of Sciences 118.26 (June 2021). doi: 10.1073/pnas.2100664118.
- Article "Finding remarkably dense sequences of contacts in link streams":
<https://www-complexnetworks.lip6.fr/~magnien/Publis/44density/article.pdf>
- Méthode de Louvain : <https://scikit-network.readthedocs.io/en/latest/tutorials/clustering/louvain.html>
- Téléchargement du dataset : <http://www.sociopatterns.org/datasets/primary-school-temporal-network-data/>