

Objectif:

Dans de nombreux systèmes, qu'ils soient naturels ou sociaux, les entités ont souvent tendance à s'organiser en groupe. L'identification et la compréhension de ces groupes s'avèrent donc être un défi important pour comprendre les structures sous-jacentes et les interactions entre les entités. L'objectif de ce TP est d'étudier ces structures dynamiques en s'intéressant plus particulièrement aux cas des réseaux sociaux basés sur la proximité (*proximity-based social networks*) et aux patterns intéressants qui peuvent être identifiés.

Exercice 1: Prétraitement et description du jeu de données

Le jeu de données que nous utiliserons correspond aux contacts de proximité collectés sur 4 jours au sein d'un service hospitalier de Lyon [1].

"This dataset contains the temporal network of contacts between patients, patients and health-care workers (HCWs) and among HCWs in a hospital ward in Lyon, France, from Monday, December 6, 2010 at 1:00 pm to Friday, December 10, 2010 at 2:00 pm."



1. Télécharger le jeu de données à l'adresse http://www.erickstattner.com/?page_id=21
2. Jeu de données:
 1. Comment sont formatées les données ?
 2. Que représente l'attribut **statut** sur les nœuds?
 3. Que représente les attributs **start** et **endopen** sur les liens?
3. Utiliser la fonction "**importer**" de Gephi pour charger les contacts et les nœuds

Exercice 2: Analyse globale

Le jeu de données représente l'ensemble des contacts ayant eu lieu sur les 4 jours de l'étude.

1. Décrire et commenter le réseau (nombre de nœuds, de liens, degré moyen, densité, distribution du degré, etc.)
2. Combien de nœuds de type "*patient*", "*infirmier*", "*médecin*" et "*administratif*" possède ce réseau?
3. Proposer un affichage de ce réseau dans lequel une même catégorie de nœud est représentée par une couleur unique
Ex. les patients en rouge, les infirmiers en jaune, etc.
4. En utilisant la notion de liens externes, afficher uniquement les contacts ayant eu lieu entre
 1. Les patients et les infirmiers
 2. Les patients et les médecins
5. Il y a-t-il des acteurs centraux de ce réseau ?
6. Proposer un affichage de tous les liens impliquant uniquement les infirmiers et les médecins.
7. Commenter ce réseau.

Exercice 3: Identification de patterns

Dans cet exercice, l'objectif sera d'identifier les situations anormales, qui peuvent potentiellement être révélatrice de "conflits relationnels" au sein de la structure.

8. Afficher les contacts entre médecins uniquement. Commenter ce réseau.
9. Afficher les contacts entre infirmiers uniquement. Commenter ce réseau.
10. Donner la liste des 5 infirmiers ayant eu le moins de contacts avec leurs collègues infirmiers.
11. Il y a t-il un ou des leaders dans le groupe d'infirmiers?
12. Cette situation est-elle révélatrice de "*clans*" au sein du groupe d'infirmiers?

Pour vérifier cela, nous allons utiliser un algorithme d'extraction de communautés. Utiliser la notion de modularité pour proposer un affichage de ces possibles clans (i.e. les communautés).

13. Afficher les contacts ayant eu lieu entre les infirmiers et les médecins.
14. Donner l'identifiant des deux infirmiers qui ont eu le moins d'échanges avec des médecins durant la période d'étude.
15. Quel est l'identifiant du médecin qui a été le moins contacté par des infirmiers?
16. Il y a t-il des patients qui n'ont aucun contacts avec les autres patients? Montrer le à travers un affichage.
17. De façon analogue, analyser d'autres types de contacts (ex. Infirmier<->Patient, Administratif<->Medecin, Administratif<->Infirmier, etc.), et tentez de mettre en évidence deux ou trois situations qui peuvent vous paraître problématiques.

[1] P. Vanhems et al., Estimating Potential Infection Transmission Routes in Hospital Wards Using Wearable Proximity Sensors, PLoS ONE 8(9): e73970 (2013)