Grands Réseaux d'Interaction
TP n° 6 : Centralités

Règles Générales

- Les règles générales en TP restent valides, voir feuille du TP 1
- Ce TP est à rendre sur Moodle pour le **10 novembre**

I) Centralités

Comme vous l'avez vu en cours, il existe de nombreuses mesures de centralité dans les graphes. Elles servent à identifier les sommets importants (*centraux*, d'où le nom).
Nous nous intéressons ici aux deux centralités les plus communes : les centralités de proximité (*closeness*) et d'intermédiarité (*betweenness*). Nous notons $G = (V, E)$ un graphe non orienté.

Votre travail lors de cette séance de TP sera d'**implémenter ces deux mesures de centralité pour des graphes non orientés**.

a) Centralité de proximité

Cette mesure est basée sur l'idée que, plus un sommet est central, plus il est proche des autres sommets (et moins il est éloigné d'eux). L'éloignement d'un sommet v dans G étant défini comme la somme des distances de tous les autres sommets de G à v, Bavelas (1950) a donc proposé de définir cette centralité, que nous notons C_p, comme étant l'inverse de l'éloignement, soit :

$$C_p(v) = \frac{1}{\sum_{u \in V \setminus v} d(u, v)}$$

où $d(u, v)$ représente la distance *géodésique* entre les sommets u et v dans G, c'est-à-dire la longueur d'un des plus courts chemins les séparant.

On utilise en général une forme normalisée de C_p, où l'on divise la formule précédente par $n - 1$, avec $n = |V|$. Comme le -1 est négligeable par rapport à n dans les graphes considérés, on utilise la formule simplifiée suivante :

$$C_p(v) = \frac{n}{\sum_{u \in V \setminus v} d(u, v)}$$

b) Centralité d'intermédiarité

Cette mesure, proposée par Anthonisse mais publiée pour la première fois par Freeman (1977), repose sur l'idée qu'un sommet figurant un grand nombre de fois dans un plus court chemin entre deux paires de sommets dans le graphe, a une haute centralité.
Pour cela, on calcule tous les plus courts chemins entre chaque couple de sommets $(s, t) \in V^2$, $s \neq t$. Le nombre total de plus courts chemins pour un couple (s, t) est noté $\sigma_{s,t}$.

La centralité d'intermédiarité, que nous notons C_i, est donc définie pour une sommet $v \in V$ comme :

$$C_i(v) = \sum_{s \neq t \neq v \in V} \frac{\sigma_{s,t}(v)}{\sigma_{s,t}}$$

où $\sigma_{s,t}(v)$ est le nombre de plus courts chemins de s à t contenant v.

De même que pour la centralité de proximité, la variante normalisée est souvent utilisée. La normalisation se fait par le nombre de paires de sommets possibles dans $G\backslash v$, soit $(n-1)(n-2)/2$ pour un graphe non orienté, aboutissant à la formule :

$$C_i(v) = \frac{2}{(n-1)(n-2)} \sum_{s \neq t \neq v \in V} \frac{\sigma_{s,t}(v)}{\sigma_{s,t}}$$

II) Implémentation

Il vous est demandé lors de cette séance de fournir une implémentation pour chacune des centralités. Votre exécutable doit pouvoir produire, à partir d'un fichier `.dot` donné en entrée, une sortie sur le terminal associant à chaque sommet sa centralité normalisée. Le graphe est non orienté et supposé connexe.
Exemple pour la centralité de proximité :

```
> ./proximite graphe.dot
1       0.1
2       0.1
3       0.5
4       0.1
5       0.7
(...)
```

a) Proximité

C'est à vous de définir une implémentation. Elle doit permettre de traiter un graphe de l'ordre de la centaine de milliers de sommets, peu dense, en quelques minutes.
Conseil de départ : vous pouvez reprendre une implémentation de l'algorithme de Dijkstra ou Bellman-Ford modifiée afin de calculer les plus courts chemins entre toutes les paires de sommet d'un graphe. De telles implémentations sont disponibles sur Internet dans la plupart des langages courants.

b) Intermédiarité

Nous vous proposons ici d'utiliser l'algorithme de Brandes (2001), l'un des algorithmes les plus rapides pour calculer la centralité d'intermédiarité d'un sommet.
Le pseudo-code de cet algorithme se trouve en page annexe.

Références

Bavelas, A. (1950). Communication Patterns in Task-Oriented Groups. *The Journal of the Acoustical Society of America*, *22*, 725–730.

Brandes, U. (2001). A faster algorithm for betweenness centrality. *The Journal of Mathematical Sociology*, *25*, 163–177.

Freeman, L. (1977). A Set of Measures of Centrality Based on Betweenness. *Sociometry*, *40*, 35–41.

Algorithm 1 Centralité d'intermédiarité (betweenness centrality)

Require: $G = (V, E)$ connexe, non orienté
Ensure: $C_i(v), \forall v \in V$

```
for all v ∈ V do
   C_i(v) ← 0
end for

for s ∈ V do
   // initialisations
   S ← pile vide
   Q ← file vide
   σ(s) ← 1
   d(s) ← 0
   for all u ∈ V do
      P(u) ← liste vide
      σ(u) ← 0
      d(u) ← −1
   end for

   // instructions
   ajouter dans file : s → Q
   while Q n'est pas vide do
      retirer de file : w ← Q
      mettre sur la pile : w → S
      for each γ ∈ Γ(w) do
         // premiere fois qu'on voit γ ?
         if d(γ) < 0 then
            ajouter dans file : γ → Q
            d(γ) ← d(w) + 1
         end if
         // plus court chemin vers γ passant par w ?
         if d(γ) = d(w) + 1 then
            σ(γ) ← σ(γ) + σ(w)
            ajouter à liste : w → P(γ)
         end if
      end for
   end while
   for all v ∈ V do
      δ(v) ← 0
   end for
   while S n'est pas vide do
      dépiler : w ← S
      for v ∈ P(w) do
         δ(v) ← δ(v) + σ(v)/σ(w) (1 + δ(w))
      end for
      if w ≠ s then
         C_i(w) ← C_i(w) + δ(w)
      end if
   end while
end for
```