

Socnetv

Kamel.derouiche@gmail.com



Le libre et vous !
**15èmes Rencontres Mondiales
du Logiciel Libre**
Du 5 au 11 juillet 2014



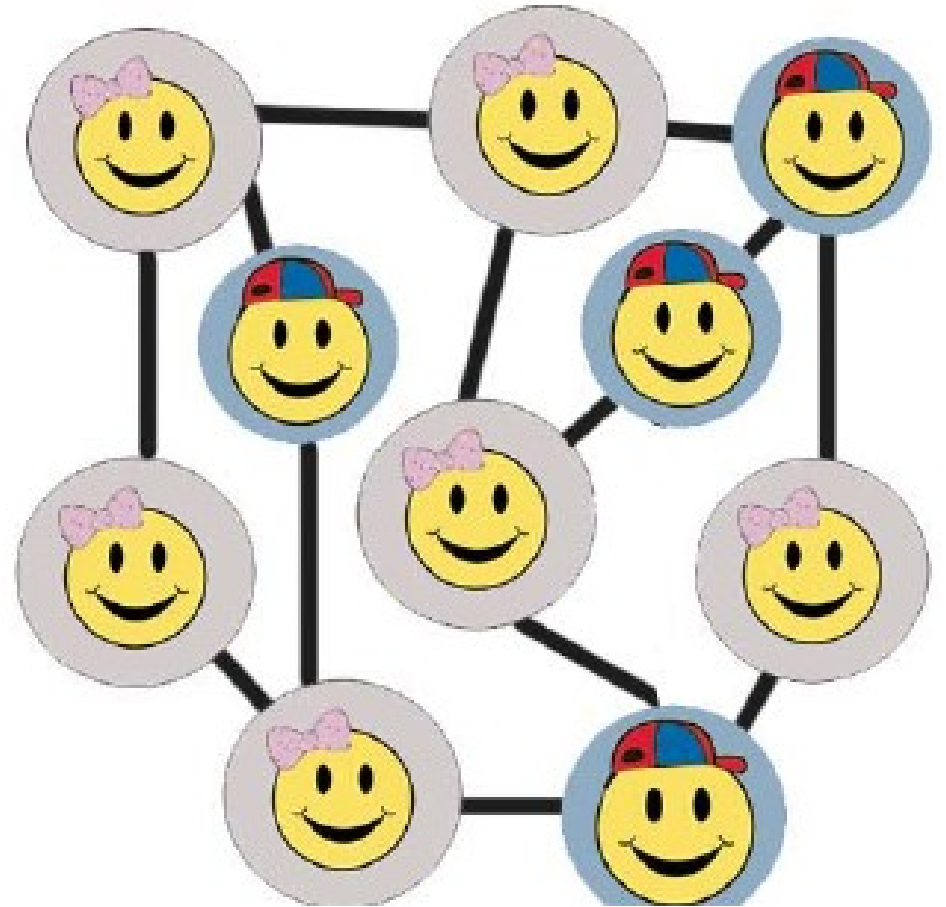
A propos

- Utilisateur de logiciel libre en Algerie depuis 2002!
- Développeur-mainteneur dans le projet NetBSD (Framework pkgsrc)
- Architect R&D chez Algerian IT Security Group
- Membre de l'Association Algérienne pour le développement de l'enseignement, des mathématiques et technologies de l'information

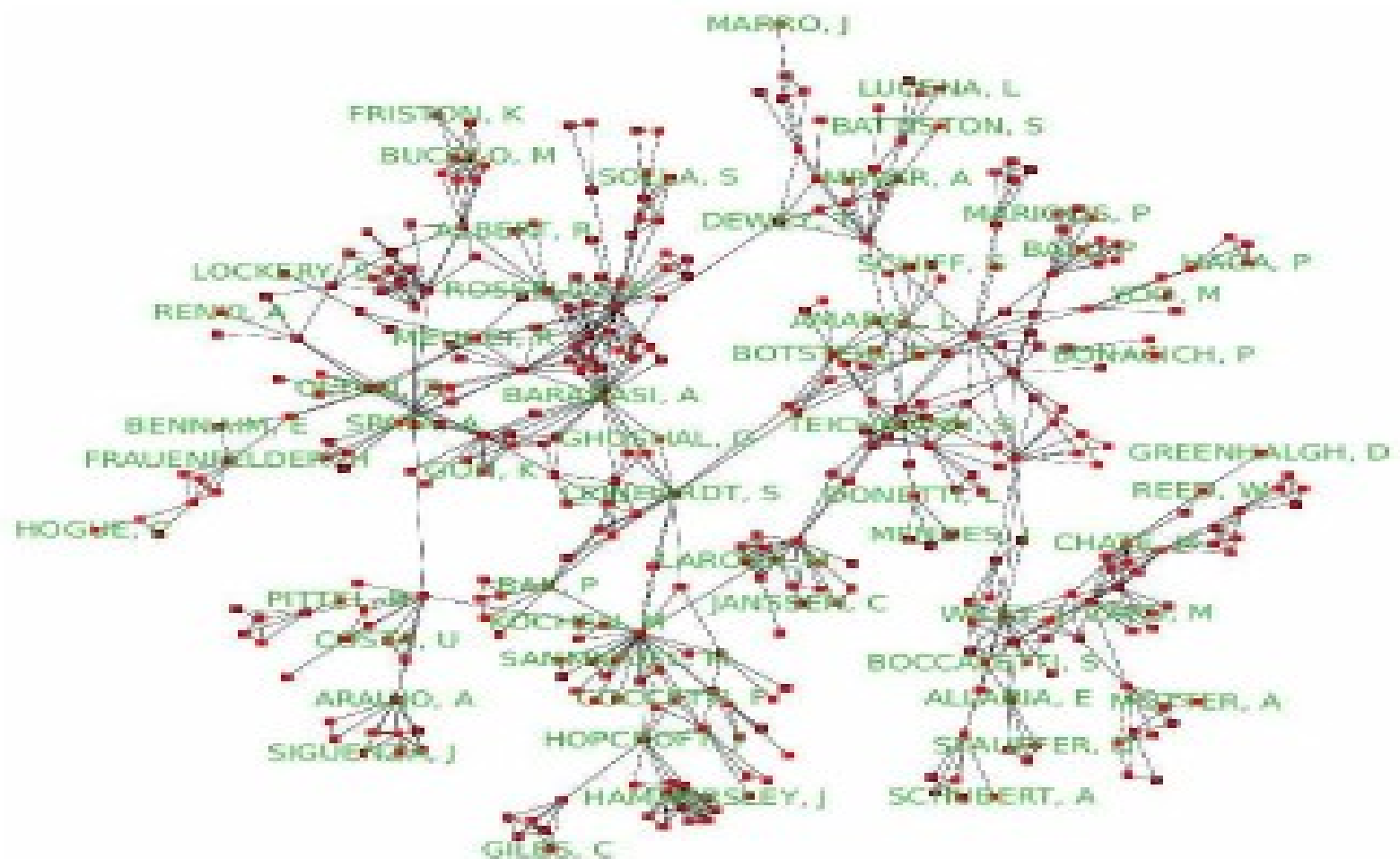


Qu'est-ce qu'un réseau social?

- Un réseau est un graphe (sommets et aretes)
- Un réseau social est un ensemble d'**acteurs** (individus, groups ou organisations) relies par des interactions sociales ils peuvent etre de différents natures (relation d'affaire, de travail, familiales, etc...)



Un réseau de collaboration scientifique



Source: Nathalie Villa

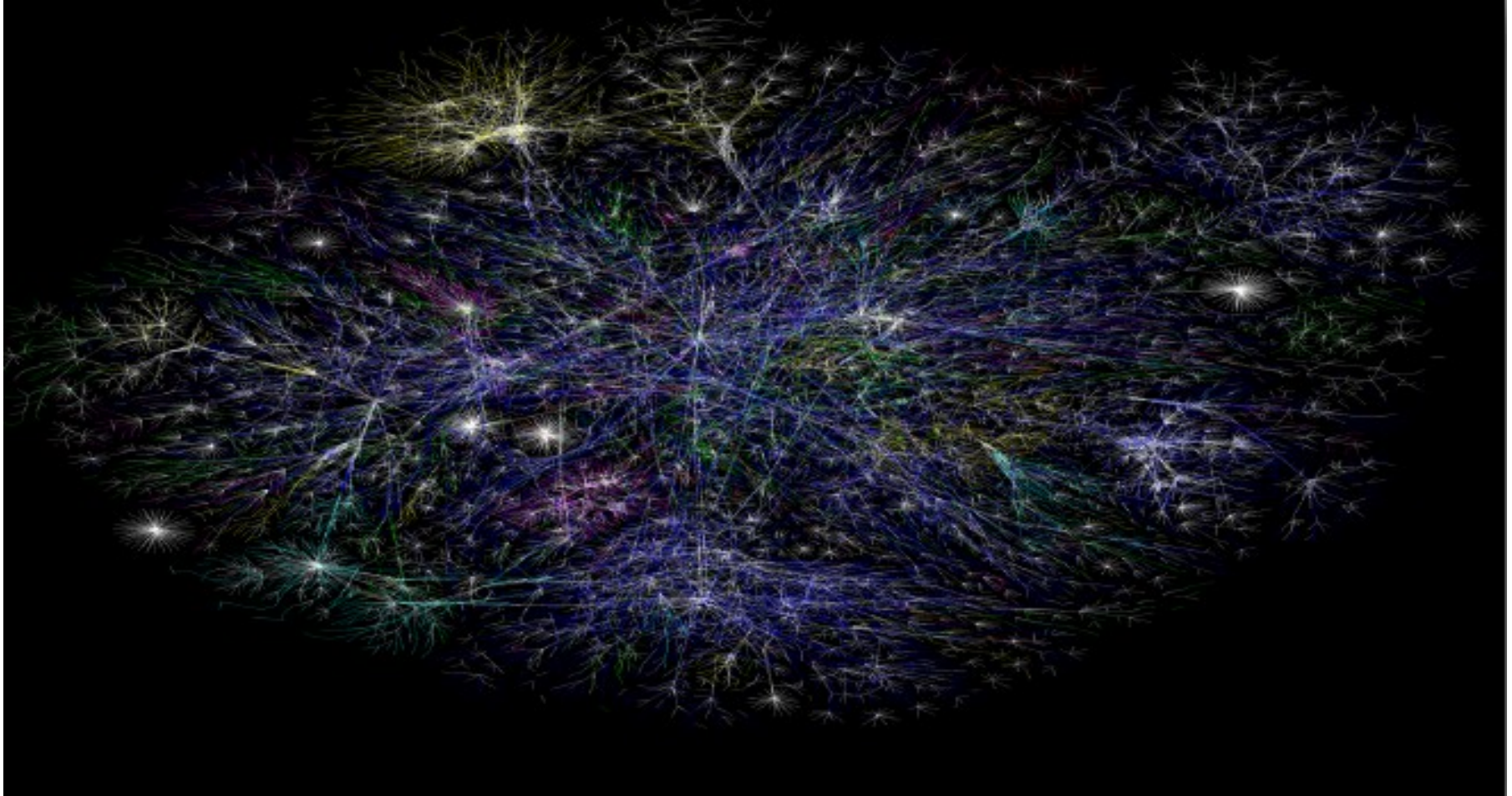


RMLL
MONTPELLIER 2014

Le libre et vous !
15èmes Rencontres Mondiales
du Logiciel Libre
Du 5 au 11 juillet 2014



Réseau Internet, réseau des amis de facebook, réseau de protéine, pair-a-pair (grands réseaux)



RMLL
MONTPELLIER 2014

Le libre et vous !
15èmes Rencontres Mondiales
du Logiciel Libre
Du 5 au 11 juillet 2014



Pourquoi analyser les réseaux sociaux?

- Pour comprendre la propagation de maladies infectieuses, rumeurs, formations d'opinions, phénomènes coopératifs
- On cherche à d'identifier des sous-groupes homogènes d'individus (comme des communautés) et à comprendre la manière dont ils sont structurés entre eux liens forts (amitiés, sentimentales) et faibles (relation de travail),
- Pour la recherche d'information, chercher à grouper les notes des étudiants par niveau pour faciliter l'identification des meilleurs



Les outils pour analyser des réseaux sociaux: bibliothèques

- networkx
- tweepy
- Networkit: Interactive tool suite for high-performance network analysis
- graphviz
- Graph-tools
- R (modules: statnet, sna, igraph): permet le calcul d'indicateurs réseaux sur un grand nombre de réseaux



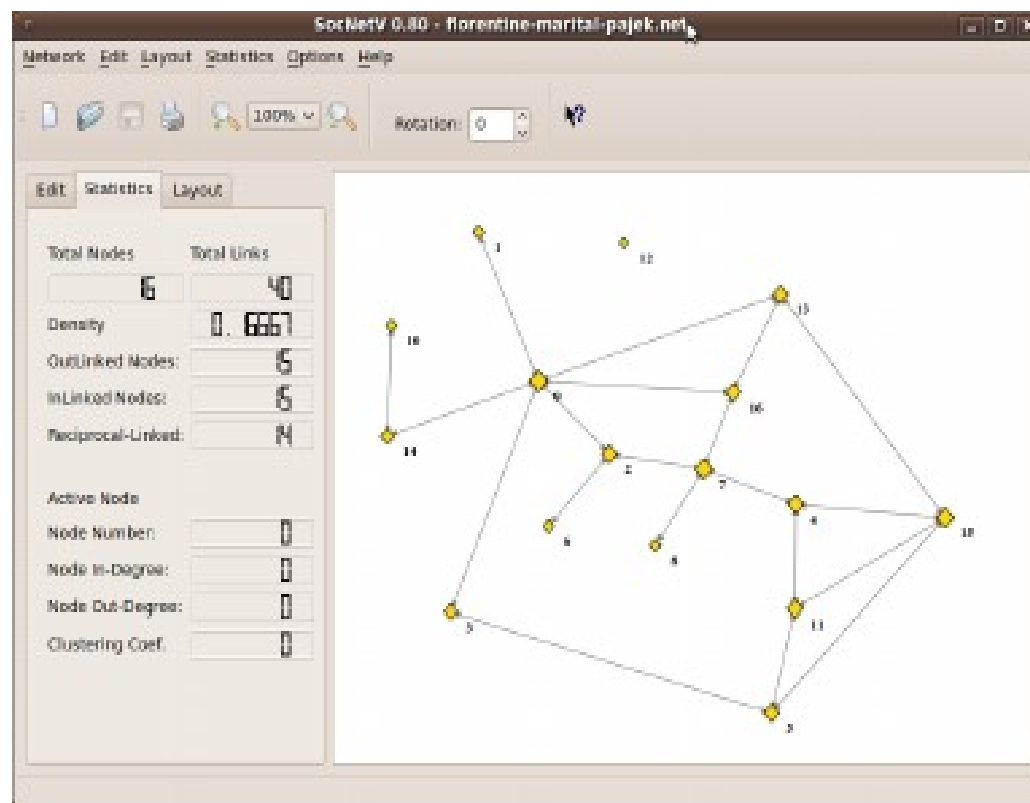
Les outils pour analyser des réseaux sociaux: Logiciel

- rocs, langage de script (Scripting Engine API) plus manipulation de graph.
- Gephi: écrit en Java, recherche et industrie
- Pajek: logiciel libre disponible sous windows, adapté pour les très grands réseaux (plus de 1000 individus), nombreuses options de visualisation graphique (3D.)
- Socnetv (Social Networks Visualizer)



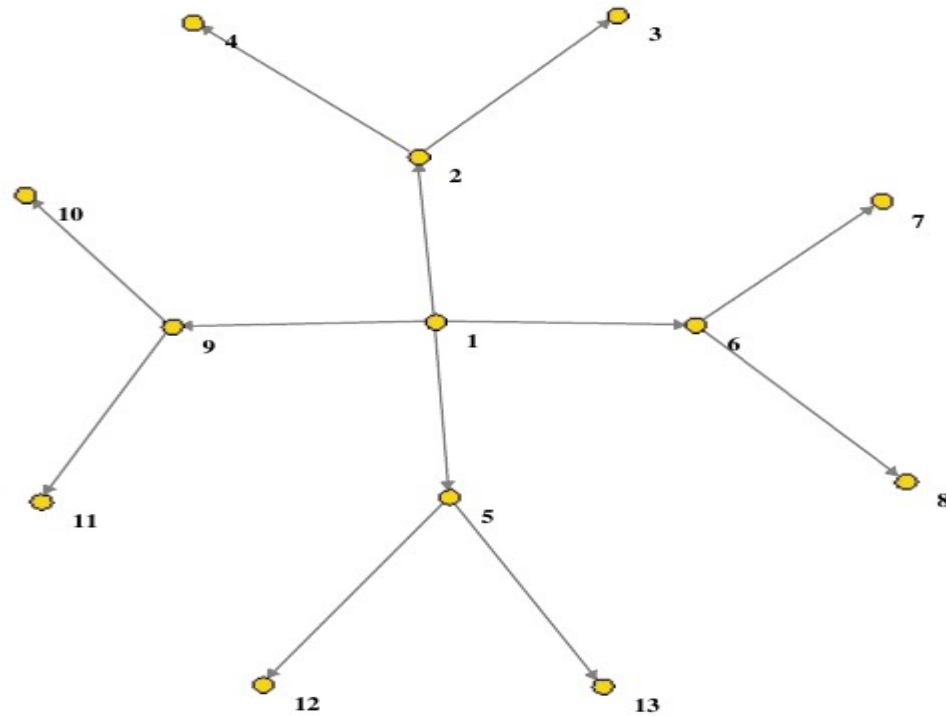
Socnetv (Social Networks Visualizer)

- Auteur: Dimitris Kalamaras
- Ecrit en C++/Qt
- License (GPLv3)
- Education
- Plus adaptés aux réseaux de petite et moyenne taille.
- Editeur (impression, import, export, sauvegarde)



Degré

- $\delta = \#(\text{nb_relations_observées}) / \#(\text{nb_sommets}) (\#(\text{nb_sommets}) - 1)$



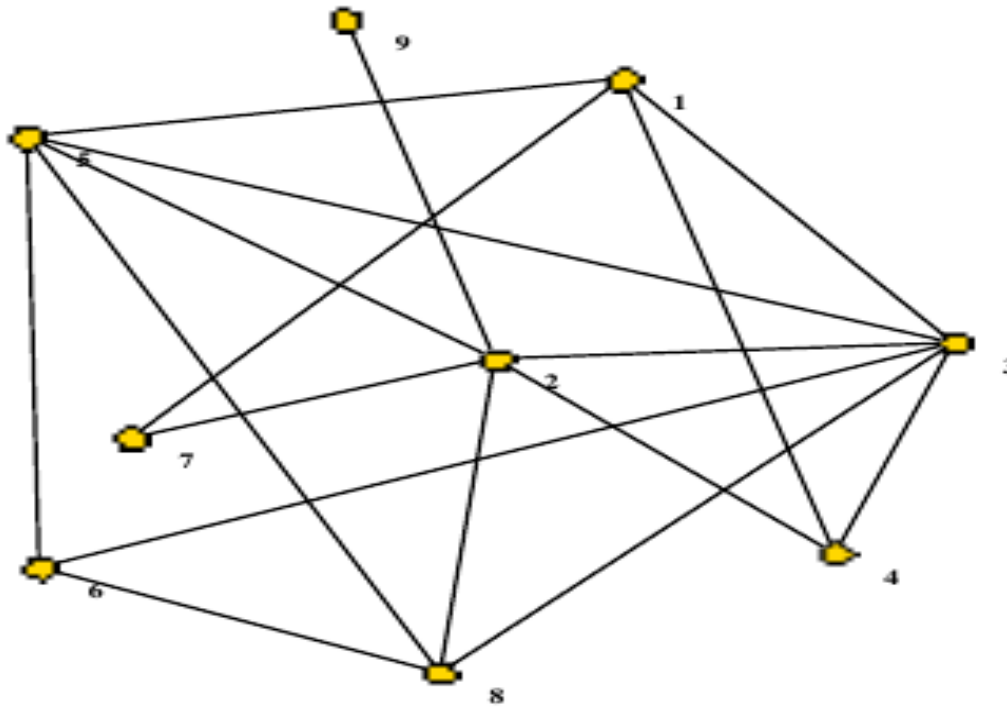
RMLL
MONTPELLIER 2014

Le libre et vous !
**15èmes Rencontres Mondiales
du Logiciel Libre**
Du 5 au 11 juillet 2014



Distance

- node_to_node
- Différence entre graphe orienté et non-orienté
- Matrice des distances, génère le fichier **distance-matrix.dat**.
On déduit facilement le diamètre du réseau = max de la matrice **distance-matrix.dat**.



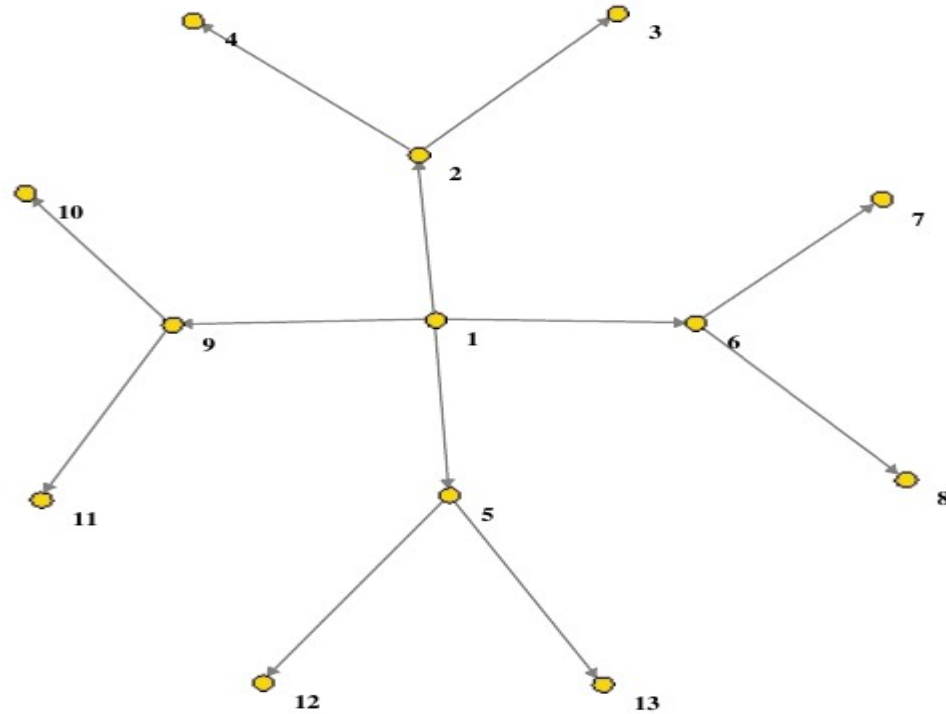
RMLL
MONTPELLIER 2014

Le libre et vous !
15èmes Rencontres Mondiales
du Logiciel Libre
Du 5 au 11 juillet 2014



Densité

- $\delta = \frac{\text{\#(nb_relations_observées)}}{\text{\#(nb_sommets)} (\text{\#(nb_sommets)} - 1)}$
- Interet: à quel point tout le monde est lié (information global)



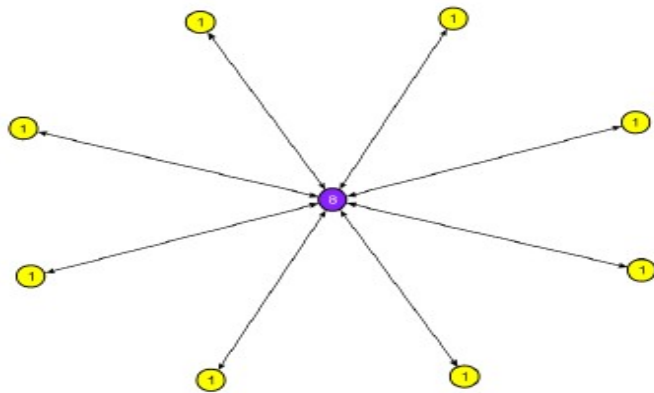
RMLL
MONTPELLIER 2014

Le libre et vous !
**15èmes Rencontres Mondiales
du Logiciel Libre**
Du 5 au 11 juillet 2014

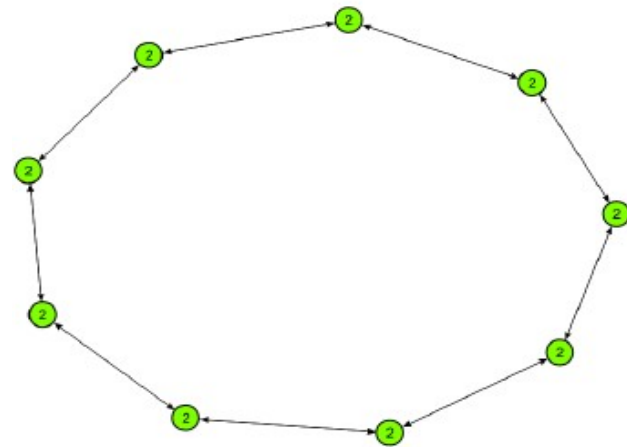


Centralité

Acteur central est l'acteur le plus actif de point de vue communication



Centralisation = 100%

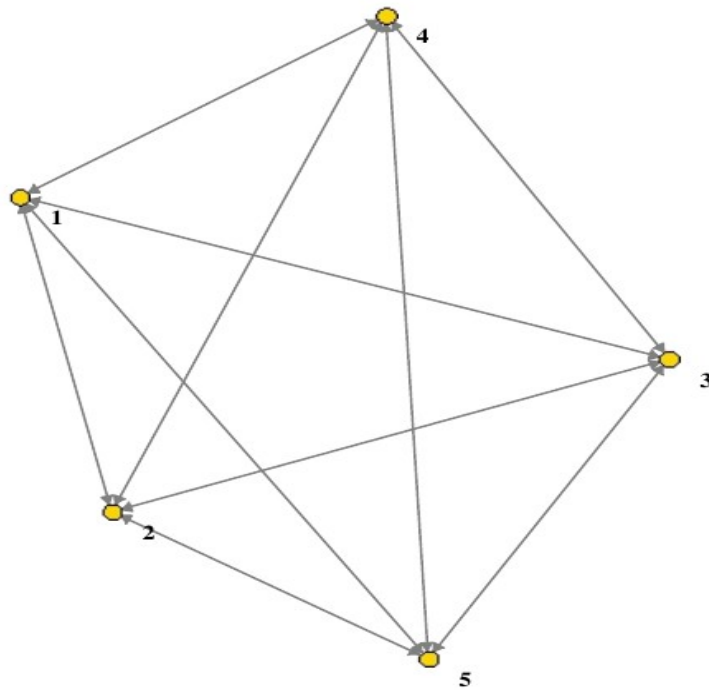


Centralisation = 0



Degré centralité d'un acteur dans graph non-orienté

- contenant n nœuds, d(i) étant le degré du nœud i : $C_D(i) = \text{deg}(i)/n-1$



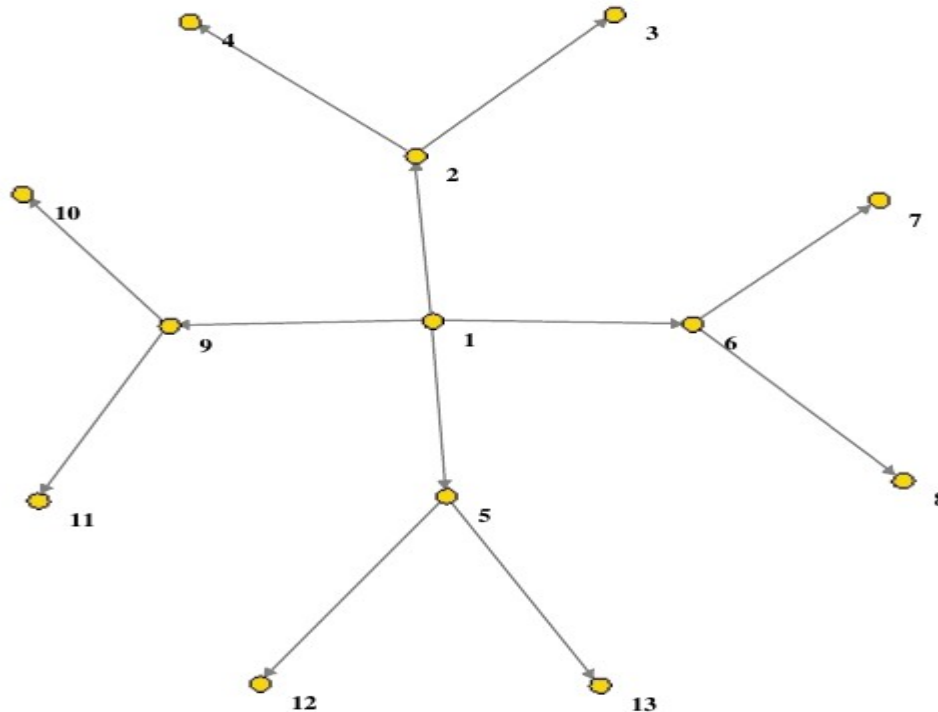
RMLL
MONTPELLIER 2014

Le libre et vous !
15èmes Rencontres Mondiales
du Logiciel Libre
Du 5 au 11 juillet 2014



Degré centralité d'un acteur dans graph orienté

- contenant n nœuds, etant le degré (liens sortants) du nœud: $C_D_0(i) = \text{deg}_0(i)/n-1$



RMLL
MONTPELLIER 2014

Le libre et vous !
15èmes Rencontres Mondiales
du Logiciel Libre
Du 5 au 11 juillet 2014



Centralité de proximité dans un graph non-orienté

- Acteur central: est l'acteur qui communique facilement avec les autres
- $d(i, j)$ est la distance entre deux acteurs mesurée en nombre minimal de lien
- Dans le cas graphe orienté, la distance doit prendre en compte les sens des liens

$$C_c(i) = n-1/\#som(dist(i, j))$$



RMLL
MONTPELLIER 2014

Le libre et vous !
15èmes Rencontres Mondiales
du Logiciel Libre
Du 5 au 11 juillet 2014

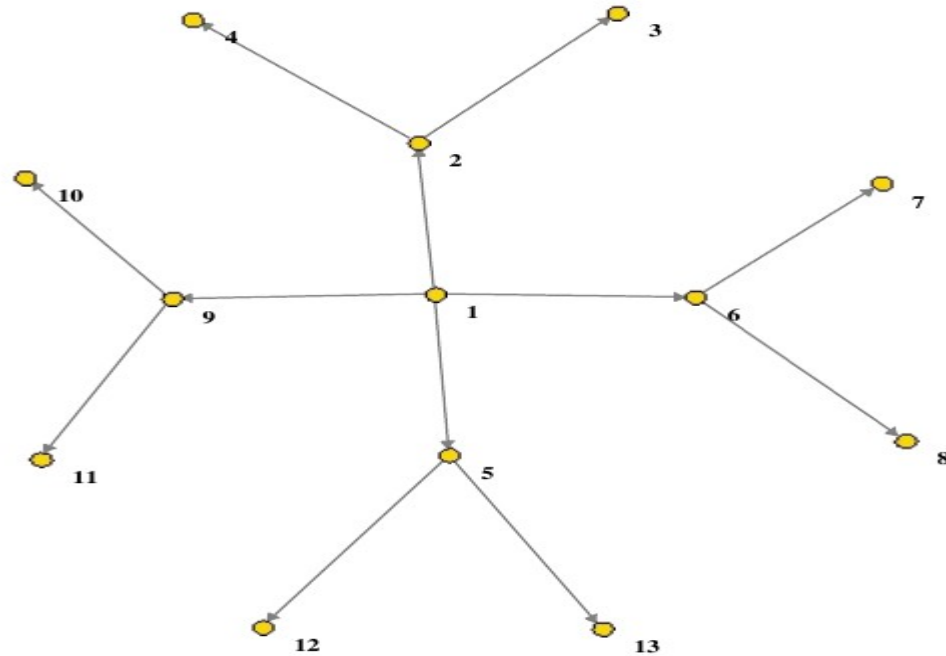


Coefficient de clustering

Identifier les communautés dans un réseau

- basée sur le postulat : l'ami de mon ami est mon ami se traduit par le nombre de triangles dans le réseau d'où la mesure du taux de clustering suivant :

$$C = 3 * (\#Nb_triangle_reseau) / (\#nb_triples_connectés_de_noeuds)$$



RMLL
MONTPELLIER 2014

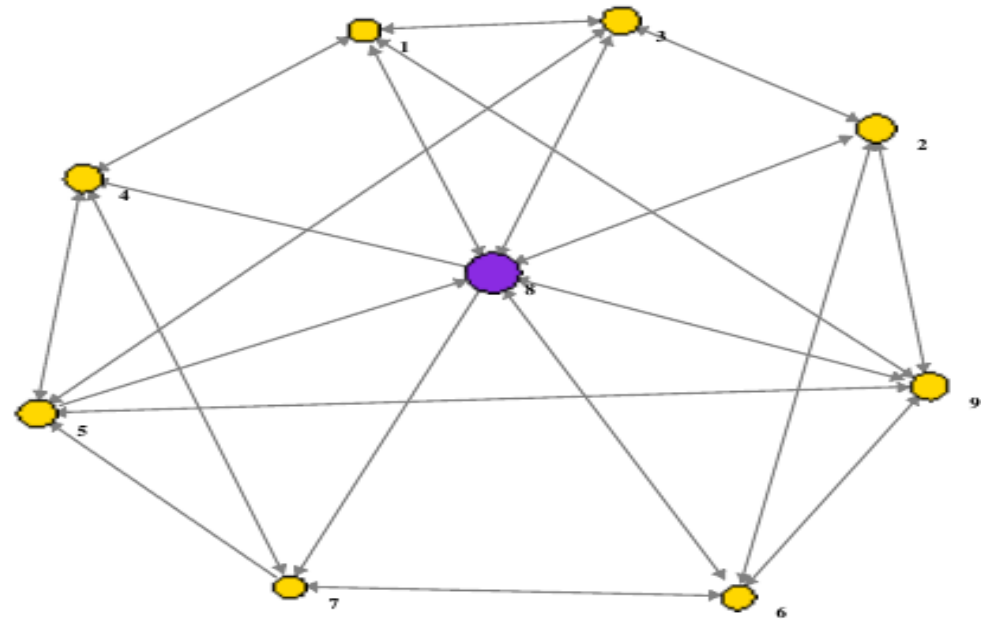
Le libre et vous !
15èmes Rencontres Mondiales
du Logiciel Libre
Du 5 au 11 juillet 2014



Traduction matricielle

- Matrice d'adjacence, socnetv génère un texte fichier adjacency-matrix.dat

```
001100011
001001011
110010010
100010110
001100011
010000111
000111010
111101101
110011010
```

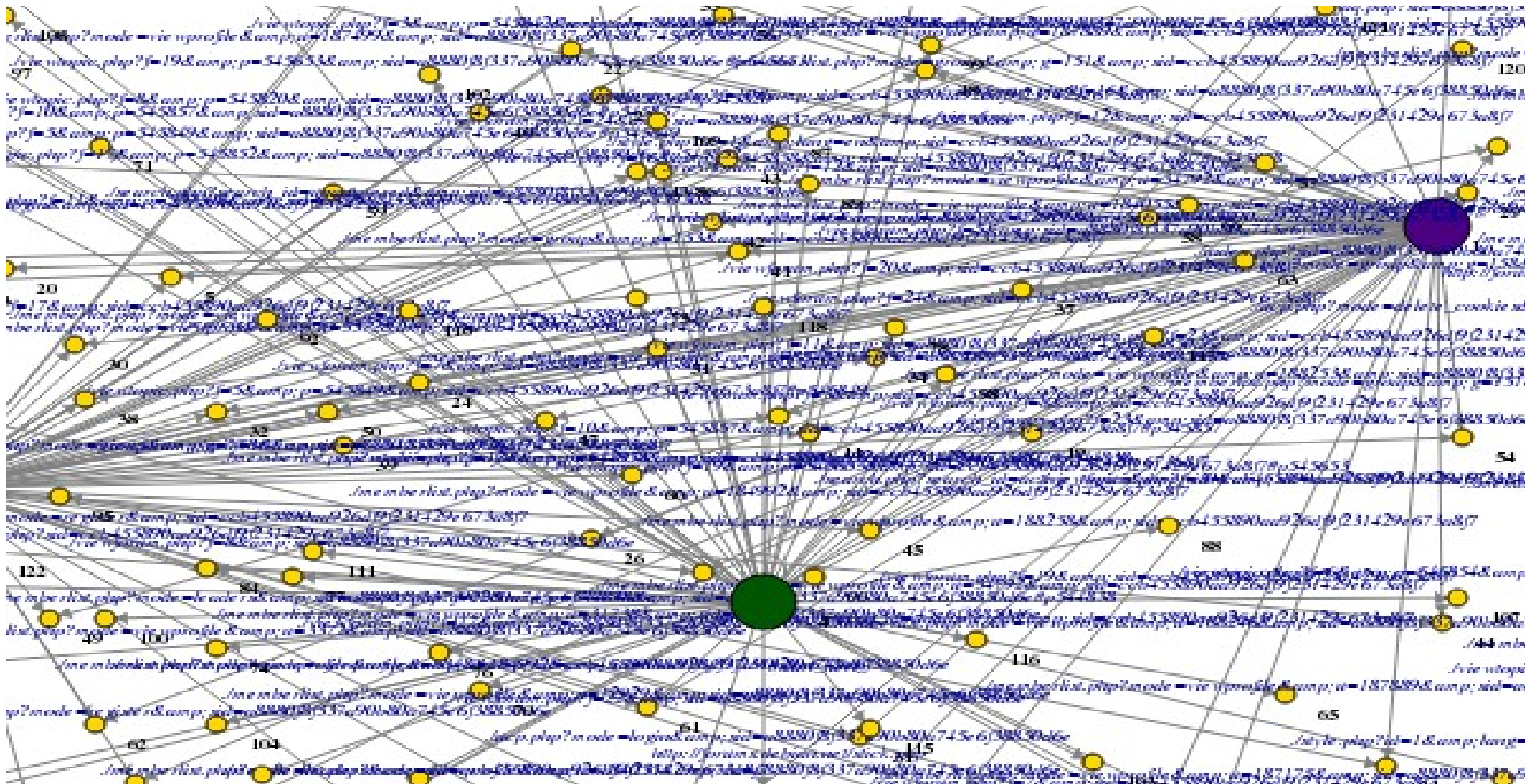


Web crawler

- Analyser des reseaux en temps réel qui évolue dynamiquement
- Pratique pour avoir un reseau sans recours à la fonctionnalité import
- Tester:
 - Avec des http et https
<http://www.debian.org> et <https://2014.rml.info>
 - Depasser les 1000 noeuds et voirs (surpirse)



Web crawler



Formats supportés

- GraphML (.graphml) (demo)
- GraphViz (.dot)
- Adjacency matrix (.net, .txt)
- Pajek-like (.net), (demo)
- UCINET's Data Language (.dl)



Questions ?



RMLL
MONTPELLIER 2014 

Le libre et vous !
**15èmes Rencontres Mondiales
du Logiciel Libre**
Du 5 au 11 juillet 2014

