

L'analyse de graphes : un aperçu avec *Cytoscape*

Guillaume Brysbaert - Unité de Glycobiologie Structurale et Fonctionnelle
guillaume.brysbaert@univ-lille.fr

G. Brysbaert & S. Plutniak
Atelier Cytoscape-R
« Analyse de réseaux et complexité »
25 Septembre 2018, Cargèse

Contenu du tutoriel

- [1 - Cytoscape, au secours! Avant toute chose...](#)
- [2 - Chargements et construction du graphe](#)
- [3 - Inspection : obtenir des informations générales](#)
- [4 - Visualisation](#)
 - [4.1 - Modifier le layout](#)
 - [4.2 - Créer un style](#)
 - [4.3 - Impression dans un fichier](#)
- [5 - Édition : modifier un graphe](#)
 - [5.1 - Éditer les sommets et arêtes](#)
 - [5.2 - Éditer les attributs](#)
- [6 - Indicage et extraction : manipuler les parties d'un graphe](#)
 - [6.1 - Sélection](#)
 - [6.2 - Extraction de sous-graphes](#)
- [7 - Métriques : mesurer les propriétés d'un graphe](#)
 - [7.1 - Mesures](#)
 - [7.2 - Visualisations](#)
- [8 - Classification des sommets d'un graphe](#)
- [9 - Appeler des commandes Cytoscape depuis R](#)

1 - Cytoscape, au secours! Avant toute chose...

Voici quelques liens vers le manuel et des tutoriels Cytoscape. Ce logiciel permet d'ajouter des outils de visualisation, d'analyse ou de construction de réseaux via des app accessibles via l'app manager intégré dans Cytoscape. Un site dédié aux apps (<http://apps.cytoscape.org/>) permet de rechercher par mot clé et donne une description de chacune d'elle qui inclut généralement un tutoriel.

- Manuel Cytoscape : <http://manual.cytoscape.org/en/stable/>
- Tutoriels Cytoscape : <https://github.com/cytoscape/cytoscape-tutorials/wiki>
- Présentations : <https://cytoscape.github.io/cytoscape-tutorials/contents/index.html#/>
- Tutoriel associé à chaque app : <http://apps.cytoscape.org/>
- Contact help desk : cytoscape-helpdesk@googlegroups.com
<https://groups.google.com/forum/?fromgroups#!forum/cytoscape-helpdesk>

N.B.: D'une manière générale, il est important de nommer les réseaux de manière à pouvoir comprendre facilement à quoi ils correspondent et de sauvegarder son projet ("File/Save as...").

2 - Chargements et construction du graphe

Commençons par charger Cytoscape et les données.

Pour cet exercice nous reprendrons le cas, classique, du réseau publié dans l'article de W. Zachazy en 1977¹. Il représente les relations d'amitié entre les 34 membres d'un club universitaire de karaté, répartis en deux groupes («factions»).

Les informations sont représentées dans deux tableaux : l'un contenant des informations sur les acteurs, l'autre contenant la description des liens. Chargeons ces deux tableaux.

importer les arêtes (appelés "edges" dans Cytoscape)

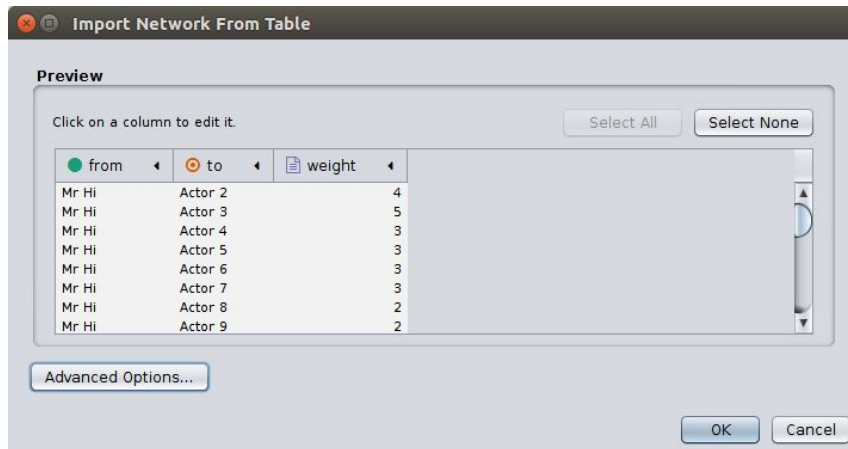
- Dans le menu "File", sélectionnez : "Import/Network/File"
- Choisissez le fichier "[karate_edges.csv](#)"

Une fenêtre "Import Network From Table" apparaît.

- Cliquez sur la colonne "from" et sélectionnez "Source node"
- Cliquez sur la colonne "to" et sélectionnez "Target node"
- Sélectionnez ";" dans les options avancées comme séparateur de colonne

La colonne "weight" sera considérée comme un attribut d'arête.

¹ W. Zachazy, 1977. «An information flow model for conflict and fission in small groups», Journal of Anthropological Research, 33, 452-473.



- Validez en cliquant sur “OK”

Le nom du réseau “karate_edges.csv” apparaît dans le menu de “Control Panel” à gauche, les données dans le “Table Panel” en bas et une représentation par défaut du graphe dans la “Network View”. Les attributs des arêtes contenus dans le fichier apparaissent dans la “Edge Table”.

importer les attributs des sommets (appelés “nodes” dans Cytoscape)

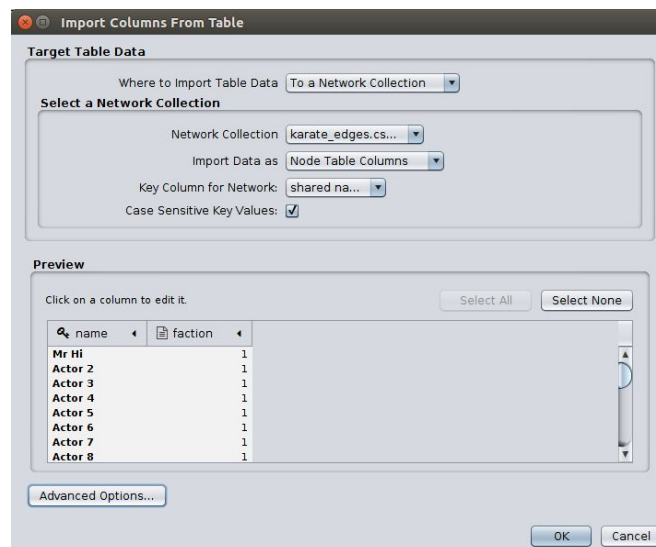
- Dans le menu “File”, sélectionnez : “Import/Table/File”
- Choisissez le fichier “karate_vertices.csv”. Le fichier “karate_vertices.csv” contient des attributs associés à chaque noeud.

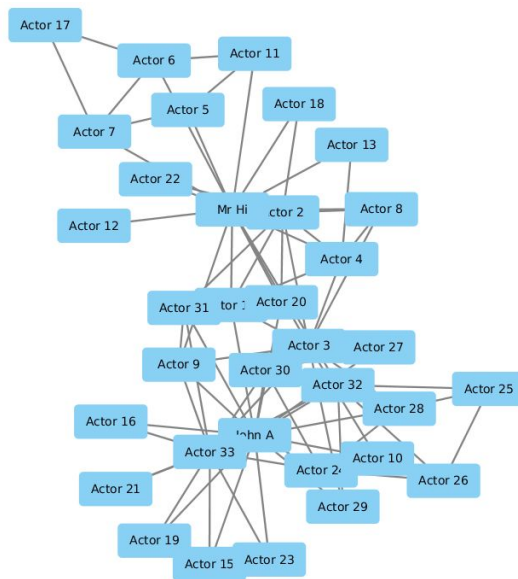
Une fenêtre “Import Column From Table” apparaît.

- Sélectionnez “;” dans les options avancées comme séparateur de colonne

Ne modifiez rien : l’icône de clef à côté de “name” indique que c’est cette colonne qui servira de référence pour associer chacun de ses éléments à ceux de la colonne “shared name” de la “Node Table” du réseau qu’on vient de générer. Les éléments des autres colonnes contenues dans “karate_vertices.csv” seront ajoutés comme attributs des noeuds dans la “Node Table”.

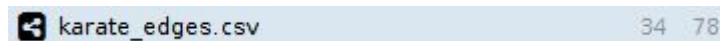
- Validez en appuyant sur “OK”





3 - Inspection : obtenir des informations générales

Des informations sur le graphe sont disponibles dans le “Control Panel”, notamment le nombre de sommets et le nombre d’arêtes.



Ici nous avons 34 sommets et 78 arêtes.

Les attributs sont listés dans les tables du “Table Panel”, “Node Table” pour les sommets et “Edge Table” pour les arêtes. Ils peuvent être triés par type de colonne par un clic gauche sur le nom de colonne.

4 - Visualisation

4.1 - Modifier le layout

Plusieurs algorithmes permettent d’afficher le réseau de différentes manières. Un algorithme est appliqué par défaut (en général “Prefuse Force Directed Layout”) et le menu “Layout” en propose d’autres.

- Créez une copie du réseau actuel : Ctrl+A, puis Ctrl+N (ou via les menus : “Select/Nodes/Select All Nodes” puis “File/New/Network/From selected nodes, all edges”)
- Renommez le nouveau réseau en “karate_edges_layout”

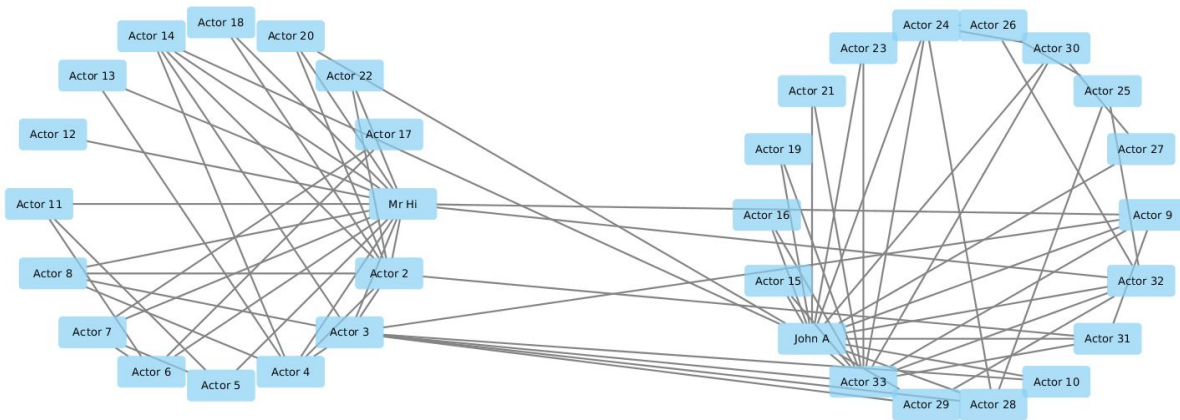
N.B.: attention, ce réseau est dans la même collection que le précédent, par conséquent, tous les attributs ajoutés, supprimés ou modifiés du premier réseau le seront aussi pour le deuxième réseau et inversement. Pour créer un réseau indépendant, il faut qu’il soit dans une autre collection, dans ce cas, il est possible de créer un clone dans “File/New/Network/Clone Current Network”.

- Dans le nouveau réseau, testez les différents layouts du menu “Layout” notamment “Circular Layout” et “yFiles Organic Layout”

Certains layout peuvent être appliqués sur les noeuds sélectionnés, c’est le cas par exemple du “Grid Layout”.

Si l'on souhaite par exemple grouper visuellement les sommets en fonction de leurs factions, il suffit de choisir la colonne "Faction" dans le "Group Attribute Layout".

- Créez une copie du réseau (Ctrl+A puis Ctrl+N)
- Nommez-la en "karate_edges_layout"
- Appliquez le "Group Attribute Layout" sur "Faction"




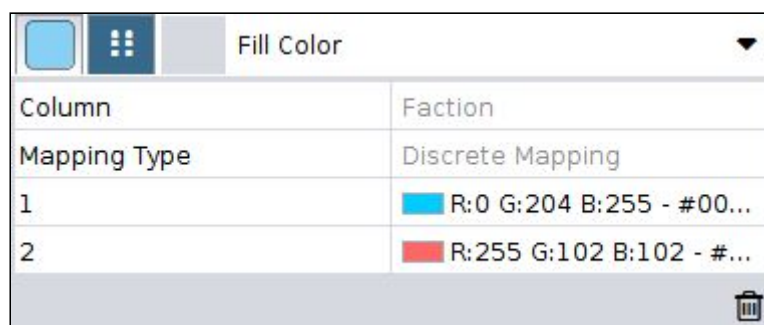
4.2 - Créer un style

La visualisation du réseau générée par défaut peut être modifiée à différents niveaux via l'onglet "Style" du "Control Panel".

- Cliquez sur l'onglet "style" du control panel
- Cliquez sur "default" et sélectionnez d'autres styles pour voir différents exemples de représentations de réseaux

Nous allons créer notre propre style "karate" :

- Cliquez à côté du style "default" sur 
- Choisissez "Copy style"
- Nommez le nouveau style "karate"
- Dans l'onglet "Fill Color", cliquez sur la petite flèche à droite de la propriété "Size".
- Choisissez la colonne "Faction" et le mapping type "Discrete Mapping". Choisissez deux couleurs en fonction de la Faction.



Mr Hi et John A sont deux personnes particulières dans le réseau, nous allons augmenter leur teinte pour mieux les visualiser :

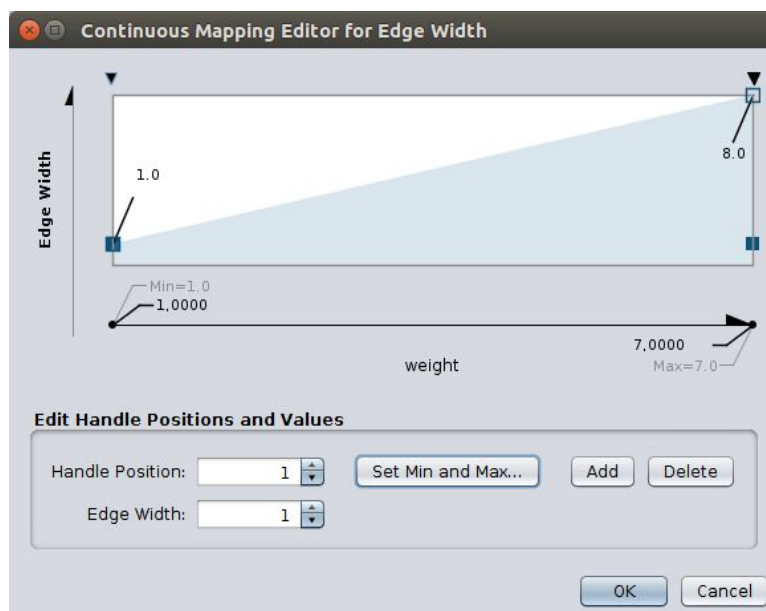
- Sélectionnez M Hi
- Cliquez sur la 3ème case de “Fill Color”, juste à côté du nom de la propriété. C’est la case de Bypass. Elle permet de spécifier une propriété différente de celle définie précédemment.
- Choisissez une couleur plus foncée
- Répétez l’opération avec John A

D’autres propriétés sur les noeuds comme la forme, la taille des labels, peuvent être modifiées en fonction (ou non) des attributs des noeuds. Nous allons modifier la transparence pour pouvoir mieux visualiser les noeuds en arrière-plan :

- Dans la propriété “Transparency”, cliquez sur la valeur par défaut “255” dans la case tout à gauche et remplacez là par “180”.

De la même manière, des propriétés des arêtes peuvent être modifiées.

- Cliquez sur “Edge” dans l’onglet en bas du “Control panel”
- Dans la propriétés “width”, adaptez la largeur à la colonne “weight” avec un “Continuous mapping”. La fenêtre suivante apparaît :

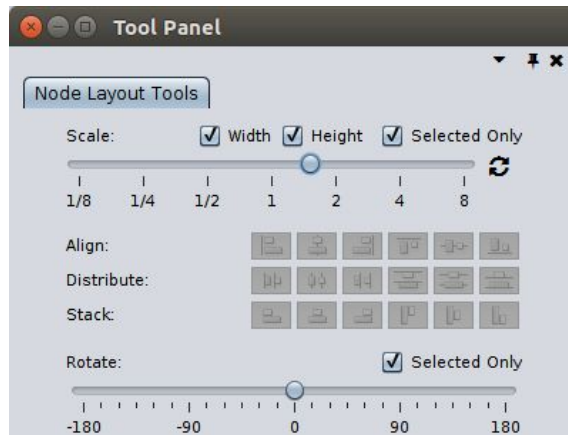


- Cliquez sur la flèche vers le bas situé en haut à gauche du graphe puis notez “1” dans “Edge Width”
- Cliquez sur la flèche vers le bas situé en haut à droite du graphe puis notez “7” dans “Edge Width”

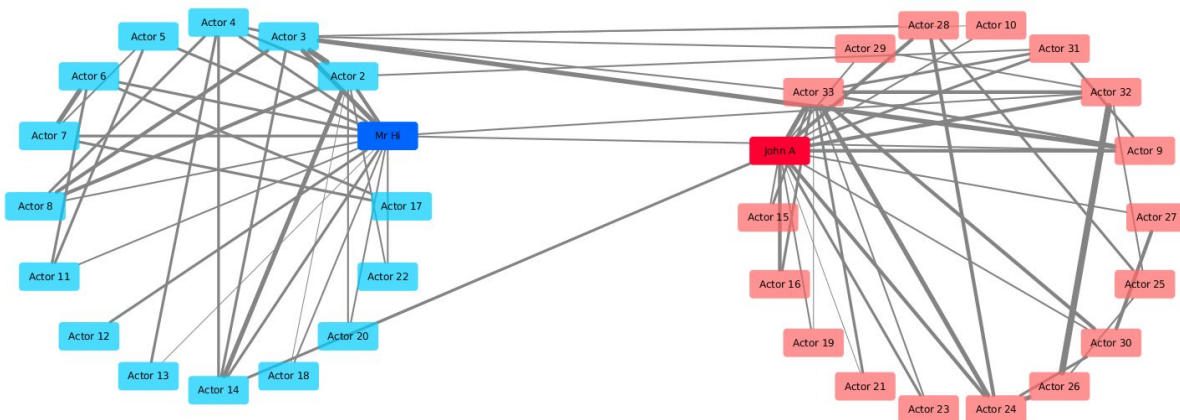
La largeur des arêtes varie maintenant entre 1 et 7 selon la valeur du poids du lien d’amitié entre les karatekas.

Dans certains cas, le réseau est tellement compact qu’il est nécessaire d’écarter les sommets.

- Dans le menu “View”, choisissez “Show Tool Panel” et positionnez le curseur d’échelle entre 1 et 2



- Ajustez le réseau à la taille de la fenêtre en cliquant sur l'icône 



N.B.: Si certains éléments du réseau n'apparaissent pas, vous devriez les voir en zoomant. Cette restriction permet de limiter les temps de calcul pour l'affichage d'une vue de réseau, ce qui peut être utile pour des réseaux de taille importante. Vous pouvez modifier ces paramètres dans "Edit/Preferences/Properties". Les propriétés de rendering concernent principalement l'affichage des labels et des bordures.

4.3 - Impression dans un fichier

Un ou plusieurs réseaux peuvent être exportés sous différents formats via le menu "File/Export/Network" mais il est souvent utile de pouvoir exporter un réseau en image :

- Dans le menu "File", choisissez "Export as image". Attention, c'est exactement l'image de réseau affichée à l'écran qui sera exportée.
- Dans les options d'export, choisissez le format et la destination, augmentez le zoom permet d'obtenir une image de plus grande définition.
- Validez

5 - Édition : modifier un graphe

5.1 - Éditer les sommets et arêtes

La suppression de sommets et/ou d'arêtes sélectionnés s'effectue dans le menu "Edit/Remove selected Nodes and edges".

- sélectionnez 5 karatékas et supprimez les. Vous pouvez ensuite annuler la suppression avec "Ctrl+Z" ou "Edit/Undo".

5.2 - Éditer les attributs

Nous voulons ajouter un «age» aux 34 karatekas (donc, formellement, un attribut aux sommets du graphe) :

- Cliquez sur l'icône "+" en haut de la "Node table", puis "New Single Column/Integer"
- Nommez la colonne "age"
- Double-cliquez dans la première cellule de la colonne "age" pour assigner un âge
- Cliquez-droit sur cette cellule et sélectionnez "Apply to entire column"

N.B.: De manière alternative, il serait possible de charger l'attribut age depuis un fichier externe qui contiendrait l'âge de chaque karatéka, de la même manière que dans la partie 2 ("File/Import/Table")


Pour supprimer l'attribut "age" :

- Cliquez droit sur le nom de la colonne
- Delete column

6 - Indicage et extraction : manipuler les parties d'un graphe

6.1 - Sélection

Plusieurs méthodes permettent de sélectionner des noeuds et d'identifier leurs attributs. Nous allons d'abord inspecter la table de sommets.

- Affichez l'identifiant interne de chaque sommet en cliquant sur l'icône  de la Node Table et en ajoutant "SUID".


Puis identifiez :


- Le premier sommet: triez le SUID par ordre croissant pour l'identifier
- Visualisez les 10 premiers sommets

Tout noeud ou lien peut être sélectionné par un simple clic dans le graphe ou une recherche dans la barre de recherche en haut de la fenêtre. Un ctrl+clic (ou shift+clic) sur plusieurs sommets ou arêtes dans la vue les sélectionnera tous.

- Sélectionnez 3 karatekas dans la vue en une fois

La barre de recherche permet de rechercher par mot-clef, en spécifiant ou non un nom de colonne (nom de la colonne+" : " +recherche).

- Cherchez John A dans la barre de recherche en tapant “John A” (les “” ici sont uniquement nécessaire à cause du caractère d’espace)
- Cherchez les liens d’amitié de poids 5 en tapant *weight: 5* (le résultat apparaît en bas à droite de la vue du réseau )

N.B.: suite à une recherche dans de grands graphes, il est parfois difficile de visualiser où se trouve le ou les noeuds sélectionnés. L’icône  permet de cadrer la vue sur les noeuds sélectionnés. La molette de la souris ou les icônes de zoom permettent ensuite d’ajuster la vue.

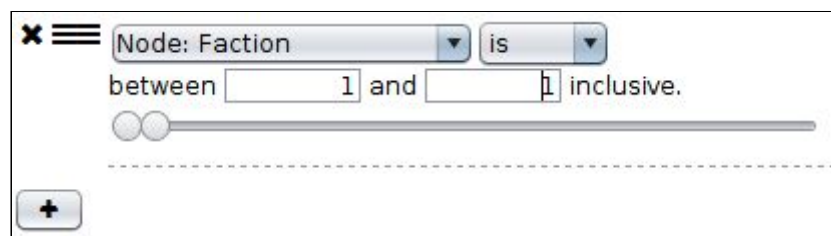
Le menu “Select” permet de sélectionner des sommets et arêtes. En particulier, il permet de sélectionner des sommets et/ou arêtes connectées à un ou plusieurs sommets :

- Sélectionnez le sommet “John A”
- Dans “Select/Nodes/First neighbor of selected nodes”, choisissez “Undirected”

Le sommet “John A” et ses voisins sont sélectionnés. Cliquez dans l’arrière plan du graphe pour désélectionner tous les noeuds et sélectionnez à nouveau “John A”. L’équivalent pour sélectionner les arêtes adjacentes est accessible dans “Select/Edges/Select Adjacent Edges”.

L’onglet “Select” du control panel permet d’affiner une sélection en s’appuyant sur les attributs des sommets ou des arêtes.

- Cliquez sur le “+”, puis “Column filter” et choisissez “Node: Faction”
- Notez “1” dans les deux cases pour ne sélectionner que les karateka de faction “1”



Par défaut, les sommets ou arêtes sont sélectionnées automatiquement.

Il est possible d’ajouter d’autres critères de filtrage en appuyant à nouveau sur “+”. On peut alors choisir l’intersection (AND) ou l’union des résultats des différents filtres (OR).

- Ajoutez un critère de filtrage “Degree filter” et sélectionnez seulement les karatekas de faction 1 qui ont au moins 4 amis.

6.2 - Extraction de sous-graphes

L’extraction s’effectue après avoir sélectionné les sommets et/ou arêtes désirées. Ensuite, il suffit de créer un nouveau réseau à partir de la sélection :

- Une fois les karatékas de faction 1 qui ont au moins 4 amis sélectionnés, sélectionnez “File/New/Network/From selected nodes, all edges” ou tapez Ctrl+N
- Renommez le nouveau réseau en “karate_edges_faction1_degreeSupeq4”
- Appliquez un “yFiles Organic Layout” et écartez un peu les noeuds grâce au Tool Panel si nécessaire

Un sous-réseau est alors créé dans la même collection, les attributs sont donc les mêmes et sont liés à ceux du réseau d'origine.

- Revenez sur le réseau initial
- Sélectionnez les liens de poids compris entre 3 et 5
- Réalisez l'extraction du sous-graphe des sommets connectés par ces liens dans "File/New/Network/From selected nodes, selected edges" (ou Ctrl+Maj+N)
- Renommez ce réseau en "karate_edges_weight_supeq3_infeq5"
- Réalisez enfin l'extraction du voisinage de degré 1 de John A à partir du réseau initial

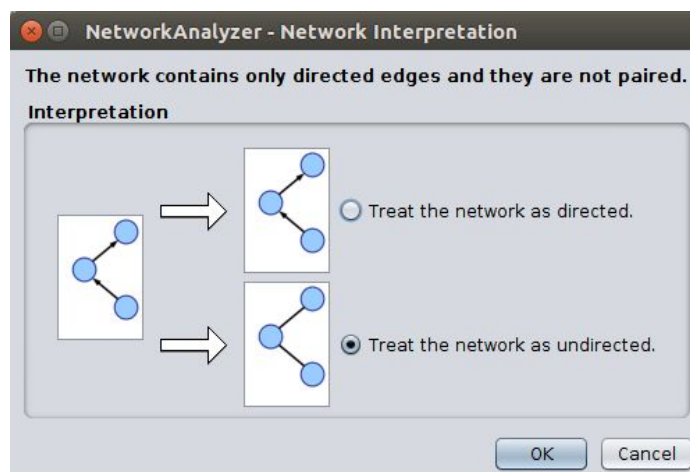
7 - Métriques : mesurer les propriétés d'un graphe

7.1 - Mesures

Une fonction dans Cytoscape permet de calculer des mesures globales et locales en une étape.

- Créer un clone du réseau "karate_edges" en le sélectionnant dans le "Control Panel", puis "File/New/Network/Clone Current Network"
- Renommez la collection et le réseau généré en "karate_edges_analysis"
- Dans le menu "Tools", sélectionnez "NetworkAnalyzer/Network Analysis/Analyze network..."

Vous devriez voir apparaître la fenêtre suivante :



- Sélectionnez "Treat the network as undirected" et validez

Plusieurs colonnes sont ajoutées à la table du réseau, elles contiennent les valeurs calculées pour chaque noeud (Degree, BetweennessCentrality, ClosenessCentrality, ClusteringCoefficient...)

Des résultats d'analyses apparaissent dans le "Result panel", ils présentent des résultats de calculs sur le réseau (telles que son diamètre, sa transitivité (clustering coefficient), son average shortest path length... et des graphiques.

- Affichez le panneau dans le menu "View/Show Results Panel"

- Visualisez les mesures sur le réseau
- Affichez la distribution des degrés en cliquant sur “Node degree distribution”.

N.B.: pour les réseaux dirigés, il faudrait sélectionner “Treat the network as directed”, ce qui donnerait deux colonnes pour “Degree” dans la “Node Table” et non une seule, soit “InDegree” et “OutDegree”. Dans la version 3.6.1 de Cytoscape, le diamètre n’est pas calculé pour les réseaux dirigés. Il devrait l’être dans une version future.

Si l’on souhaite exporter les résultats des calculs sous forme tabulaire, pour exporter toute une table :

- Dans le menu “File”, Cliquez sur “Export/Table”
- Choisissez le tableau à exporter et la destination
- Validez

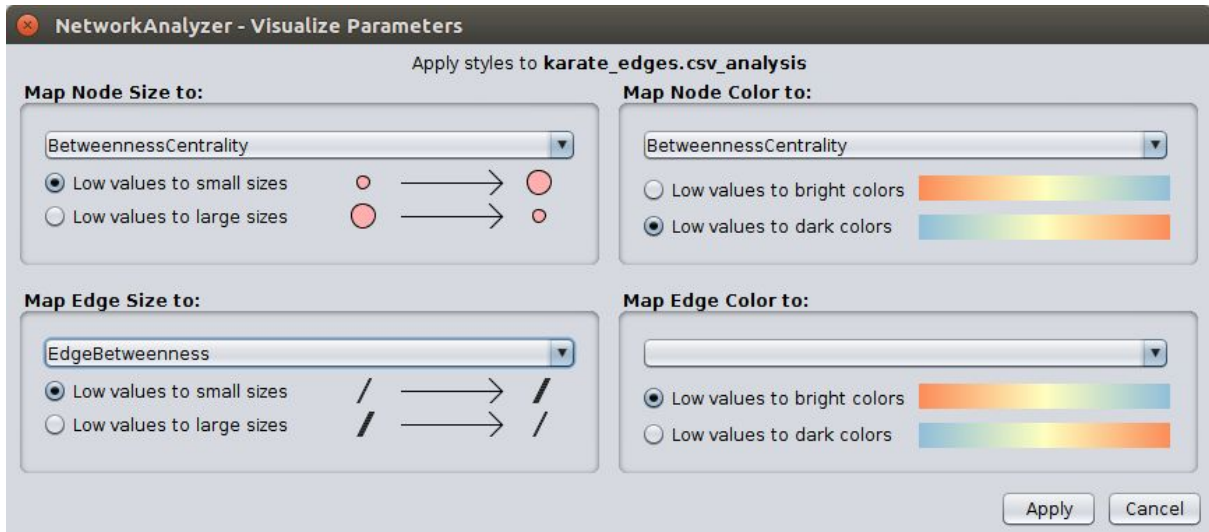
Pour exporter une partie de table : il est possible de sélectionner des données dans le “Table Panel” par un cliquer/glisser, de copier (Ctrl+C) et coller dans un tableur de type Excel (Ctrl+V). Pour exporter les degrés de chaque noeuds :

- Cliquez-glisser la colonne “Degree” à côté de la colonne “shared name” (cliquez sur le titre de colonne)
- Cliquez sur le premier élément de la colonne “Shared name” et glissez jusqu’au dernier élément de la colonne “Degree”
- Copiez-les dans un tableur (Ctrl+C/Ctrl+V)

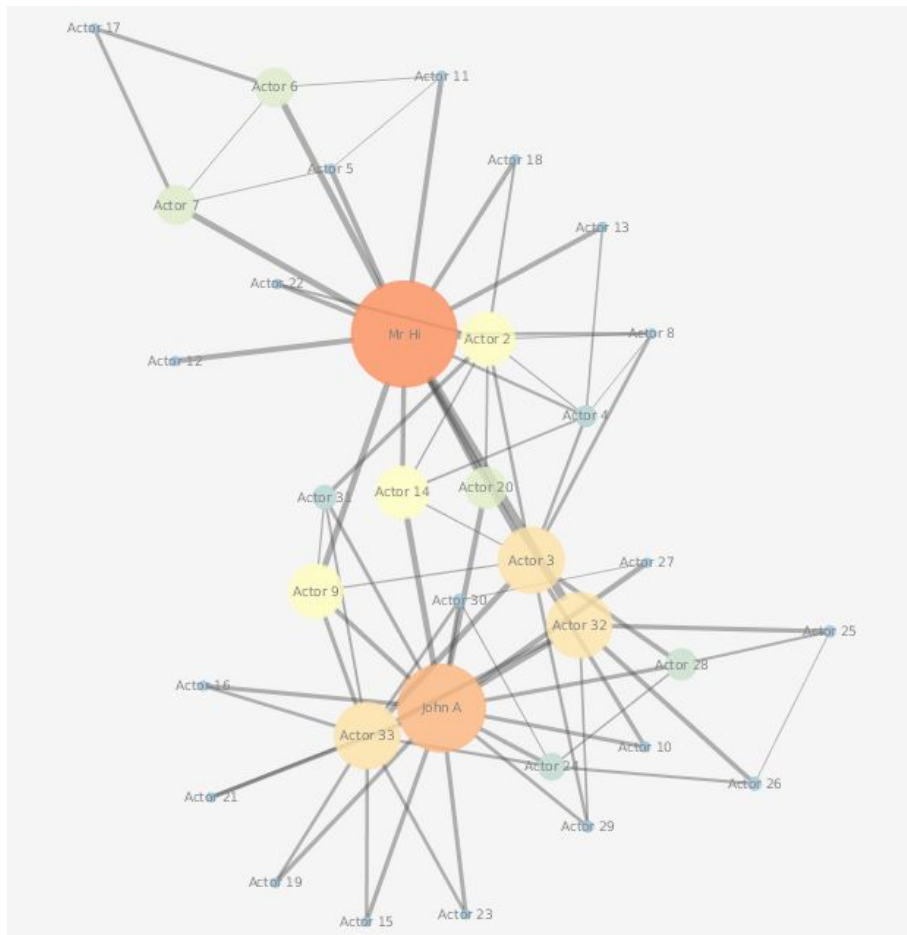
7.2 - Visualisations

Dans le “Results Panel”, un bouton “Visualize Parameters” présente une interface simplifiée pour appliquer un style au réseau qui présente les valeurs de closeness, de betweenness, degree, clustering coefficient...

- Cliquez sur “Visualize Parameters”
- Choisissez “BetweennessCentrality” dans “Map Node Size to” et “Map Node Color to”.
- Cliquez sur “Low values to dark colors” pour que les sommets qui ont les plus hautes valeurs de betweenness soient rouge et non bleu.
- Choisissez “EdgeBetweenness” pour “Map Edge Size to”
- Validez



Un nouveau style est alors créé, renommé-le en “Betweenness” dans l’onglet “Style” du “Control Panel”.



- Procédez de même avec la “ClosenessCentrality” en sélectionnant “Weight” dans “Map Edge Size To” et renommez le style en “Closeness”

Vous pouvez ainsi changer le style de “Betweenness” à “Closeness” et visualiser les sommets qui ont les centralités les plus importantes. Une alternative est d’afficher ClosenessCentrality et BetweennessCentrality sur le même graphe en faisant varier taille de noeud pour l’un et couleur de noeud pour l’autre.

8 - Classification des sommets d’un graphe

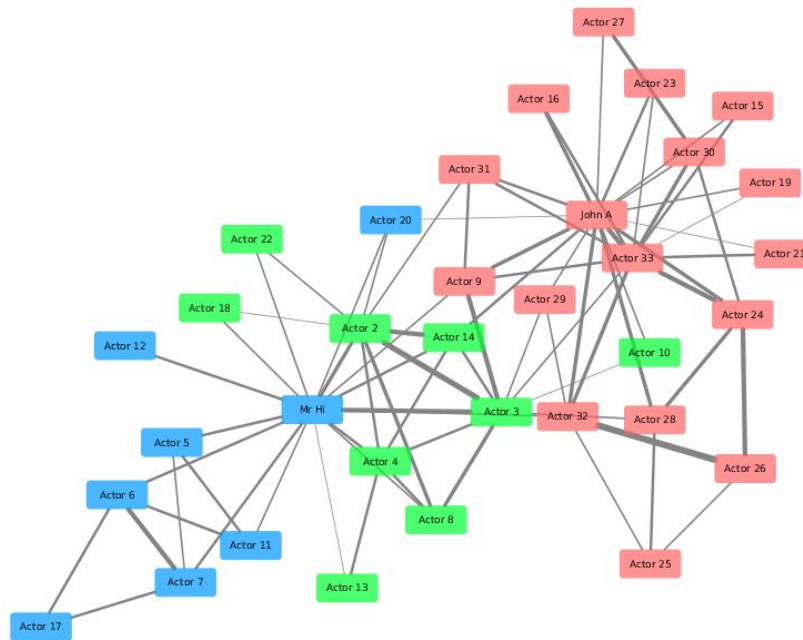
Plusieurs algorithmes de classification sont implémentés dans des app. ClusterMaker2 par exemple (<http://www.rbvi.ucsf.edu/cytoscape/clusterMaker2/>) propose des algorithmes de classification de sommets dans la section “Network Cluster Algorithms” de “Apps/ClusterMaker2”. Nous allons utiliser l’algorithme de clustering de communauté Girvan-Newman fast greedy.

- Clonez le réseau initial “karate_edges.csv” et renommez-le
- Sélectionnez “Apps/ClusterMaker/Community Cluster (GLay)”
- Validez les paramètres pré-sélectionnés

Trois clusters sont identifiés et à chaque sommet est associé son numéro de cluster en attribut “_glayCluster” dans la “Node Table”.

Nous allons créer un style qui affiche le graphe avec une couleur par classe.

- Copiez le style “karate” et nommez-le “karate_cluster”
- Modifiez la propriété “Fill color” pour afficher une couleur par noeud à l’aide d’un “Discrete Mapping” sur la colonne “_glayCluster”
- Supprimez les Bypass de Mr Hi et John A par un clic-droit sur la 3ème case + “Remove Bypass” après les avoir sélectionnés.
- Changez de Layout pour optimiser la visualisation (Group Attributes Layout - _glayCluster, Circular Layout, Prefuse Force Directed Layout...)



- testez d'autres algorithmes de classification (ex: Affinity Propagation Cluster) et visualisez les résultats en créant un style adapté.

9 - Appeler des commandes Cytoscape depuis R

Depuis la version 3.3.0, Cytoscape embarque un service REST qui permet de communiquer depuis un langage extérieur comme R ou Python. L'envoi de commande se fait donc par l'intermédiaire d'URLs spécifiques à chacune d'elle, avec éventuellement des paramètres. Les URLs de chaque commande sont décrites dans "Help/Automation/CyREST API" et "Help/CyREST Command API", et doivent être interrogées au format JSON. Certaines app intègrent cette fonctionnalité. L'intérêt est de pouvoir bénéficier des fonctionnalités des deux logiciels. Ici, nous allons simplement exporter les images de tous les réseaux générés.

- créez un fichier R
- copiez/collez le code :

```
#required packages
#install.packages("httr") #uncomment if necessary
require("httr")
#install.packages("jsonlite") #uncomment if necessary
require("jsonlite")

#parameters
height <- "2000"
options <- "jpg"
resolution <- "150"
units <- "pixels"
width <- "2000"
zoom <- "200"

urlList <- "http://localhost:1234/v1/commands/network/list"
```

```

urlSetNetwork <- "http://localhost:1234/v1/commands/network/set%20current"
urlExport <- "http://localhost:1234/v1/commands/view/export"

#function that sends a POST request in JSON format and returns the content of the result
requestPOST <- function(urlString, bodyList, encodeString)
{
  request <- POST(url=urlString, body=bodyList, encode=encodeString)
  resultFromJSON <- unname(fromJSON(rawToChar(request$content)))
  return(resultFromJSON[[1]])
}

#gets the list of networks SUIDs
networksList <- requestPOST(urlString=urlList, bodyList=NULL, encodeString="json")

#calculates centralities for each network
for(i in 1:length(networksList$networks))
{
  #sets the network
  networkID <- paste("SUID:",networksList$networks[i],sep="")
  requestPOST(urlString=urlSetNetwork, bodyList=list(network=networkID), encodeString="json")

  #exports the view
  requestPOST(urlString=urlExport, bodyList=list(Height=height, options=options, Resolution=resolution, Units=units,
Width=width, Zoom=zoom), encode="json")
}

```

- exécutez-le

Une image jpeg est créée pour chaque réseau dans le dossier courant. Le nom de chaque image est le même que celui du réseau correspondant.