

Installation et prise en main des outils d'ARS

iGraph, Gephi et NetworkX

Ahmed Mohameden
amed.mohameden@gmail.com

Université Cheikh Anta Diop
Faculté des Sciences et Techniques
Département de Mathématiques et Informatique
Section Informatique

26 mars 2018

A propos

Enseignant

Ahmed Mohameden

- Ingénieur en Systèmes d'Information Repartis ;
- Doctorant en Analyse des Réseaux Sociaux.

Pages du cours

- Logiciels et outils ;
- Supports de cours, scripts et jeux de données.

Évaluation

Ce module est constitué de 6 travaux pratiques. Chaque séance comprend un TP à rendre.

Thèmes abordés

- Installation et prise en main des outils d'ARS ;
- Manipulation et visualisation des graphes avec R/iGraph ;
- Manipulation et visualisation des graphes avec Gephi ;
- Mesures locales et globales ;
- Algorithmes de prédiction de liens ;
- Détection de communautés statiques ;
- Analyse des réseaux sociaux dynamiques.

Plan

Introduction

Pourquoi utiliser l'ARS ?

Démarche d'ARS

Outils d'ARS

- Quel outil choisir ?

- R, RStudio et iGraph

- Gephi

- NetworkX

- Comparaison des outils d'ARS

TP à rendre

Introduction

L'Analyse des Réseaux Sociaux (ARS) est une approche permettant de modéliser, de visualiser et d'analyser les relations sociales entre les éléments du réseau.



FIGURE: Illustration d'un réseau social.

Exemples de réseaux sociaux

Des structures sociales couramment étudiées grâce à l'analyse des réseaux sociaux comprennent :

- Réseau d'amitié ;
- Réseau de proximité ;
- Réseau d'appels téléphoniques ;
- Réseau de personnes touchées par une maladie.

Pourquoi utiliser l'Analyse des Réseaux Sociaux (ARS) ?

L'ARS est utilisée pour répondre à des questions telles que :

- Qui sont les influenceurs ? Les intermédiaires ? Les points isolés ?
- Où sont les communautés ? Et qui en fait partie ?
- Qui est au cœur du réseau ? Et qui en est à la périphérie ?
- Les amis de mes amis tendent-ils à devenir mes amis ?

Toutes ces questions relèvent des réponses permettant de mieux comprendre le comportement des éléments du réseau et pourraient, donc, aider à prendre des décisions stratégiques.

Démarche d'ARS

La démarche d'analyse d'un réseau social est composée de 3 phases :

- 1 Collecte d'information
- 2 Saisie de données
- 3 Visualisation et interprétation

Quel outil choisir ? (1/3)

Le choix d'un outil d'analyse des réseaux sociaux dépend de plusieurs critères :

- Type de licence : la licence de l'outil est-elle propriétaire ? Libre ? Ou gratuite ?
- Portabilité : l'outil est-il multi-plateforme ?
- Récence : l'outil est-il mis à jour fréquemment ?
- Performance : à quel degré l'outil permet de supporter la manipulation des graphes de grande taille ?

Quel outil choisir ? (2/3)

- Fichiers d'entrée/de sortie : l'outil permet-il d'importer et d'exporter des graphes attribués ? Multiplexes ? Temporels ?
- Manipulation de graphes : ajout/suppression de nœuds/liens, génération de graphes, extraction des voisinages, etc. ;
- Métriques : quelles sont les mesures locales et globales que l'outil implémente ?

Quel outil choisir ? (3/3)

- Détection de communautés : les algorithmes fournis détectent quels types de communautés (disjointes, chevauchantes, hiérarchiques, etc.) ?
- Prédiction de liens : voisins communs, coefficient de Jaccard, attachement préférentiel, etc. ;
- Layouts : l'outil offre-t-il une visualisation interactive ? Les layouts permettent-ils de regrouper des parties du réseau en cluster ? etc. ;
- Réseaux dynamiques : l'outil permet-il d'analyser et de visualiser les réseaux qui évoluent au fil du temps ?

R, RStudio et iGraph

R

R est un langage de programmation et aussi un logiciel libre multi-plateforme dédié au calcul et à l'analyse statistique.

RStudio

RStudio est un environnement de développement des applications R. Nous utilisons RStudio lors de la manipulation des réseaux dynamiques.

iGraph

iGraph est un package d'analyse des réseaux sociaux statiques, disponible en 3 langages : C, Python et R.

Téléchargement

R est disponible sur le site du CRAN (Comprehensive R Archive Network) : <http://cran.r-project.org>

RStudio est disponible en plusieurs éditions. Dans ce cours, nous nous intéressons à la version RStudio Desktop, disponible sur :

<https://www.rstudio.com/products/rstudio/download/>

Installation

Après avoir téléchargé les programmes d'installation de R et RStudio, exécutez-les et suivez les étapes d'installation. Il est préférable de commencer par installer R avant d'installer RStudio.

Une fois R installé, ouvrez-le et tapez la commande suivante pour installer le package iGraph :

```
install.packages("igraph")
```

Premières pas

- Charger le package iGraph ;
- Exécuter le script [demo_igraph.R](#) pour visualiser quelques démonstrations de iGraph.

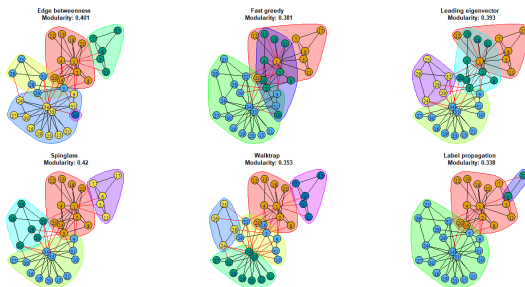


FIGURE: Illustration de quelques algorithmes de détection de communautés implémentés en iGraph.

Gephi

Gephi est un logiciel libre d'analyse et de visualisation de réseaux statiques et dynamiques. Il est développé en Java et basé sur la plateforme NetBeans.

Le lien ci-dessous comprend une courte présentation de la version 0.7 de Gephi.



Téléchargement

Pour télécharger la dernière version de Gephi, cliquez [ici](#).
Plusieurs ressources intéressantes sont disponibles sur le site, telles que :

- [Guide d'utilisation](#) ;
- [Jeux de données](#) ;
- [Forums d'assistance](#) ;
- Et encore plus sur <https://gephi.org/users/>.

Installation

Après avoir téléchargé le programme d'installation de Gephi, exécutez-le et suivez les étapes d'installation.

Prérequis :

- ① JRE : la présence d'une JRE est obligatoire puisque Gephi est développé en Java ;
 - Pour voir la version de Java installée sur votre système, tapez `java -version` dans l'invite de commandes ;
- ② OpenGL : Gephi utilise un moteur OpenGL 3D pour accélérer la visualisation de graphes. Il serait recommandé de mettre à jour le pilote de la carte graphique ;
 - Vous pouvez utiliser l'outil [OpenGL Extensions Viewer](#) pour afficher la version installée d'OpenGL.

Premières pas

- Visualiser le réseau des Misérables ;
- Décrire brièvement les 3 onglets principaux de Gephi.

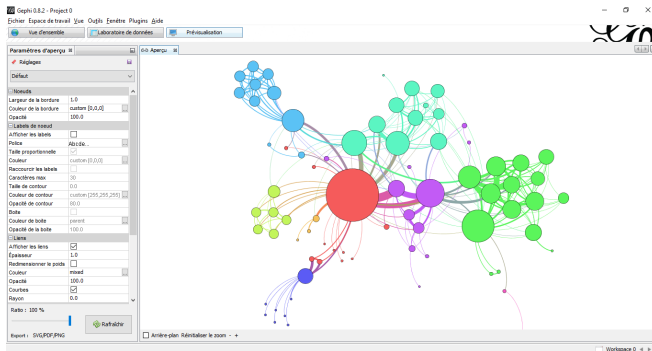


FIGURE: Interface d'accueil de Gephi.

NetworkX

NetworkX est un package Python permettant la création, la manipulation et l'étude de la structure des réseaux réseaux.

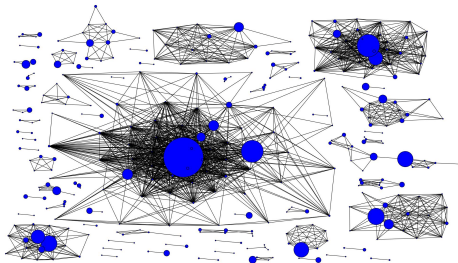


FIGURE: Exemple d'utilisation de NetworkX montrant toutes les communautés d'un graphe.

Téléchargement

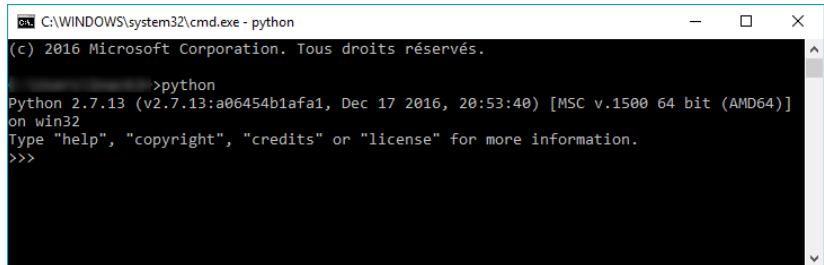
Pour pouvoir installer NetworkX, vous devez d'abord installer Python sur votre système d'exploitation. La version que je vous recommande est 2.7.x, disponible sur le lien ci-dessous.

<https://www.python.org/downloads/>

Configuration

Pour que vous puissiez exécuter des commandes Python à travers l'invite de commandes, vous devez ajouter le chemin d'installation de Python à la variable d'environnement PATH.

Pour vérifier si Python est accessible à travers la console, tapez Python :



```
C:\WINDOWS\system32\cmd.exe - python
(c) 2016 Microsoft Corporation. Tous droits réservés.

>python
Python 2.7.13 (v2.7.13:a06454b1afa1, Dec 17 2016, 20:53:40) [MSC v.1500 64 bit (AMD64)]
on win32
Type "help", "copyright", "credits" or "license" for more information.
>>>
```

Installation (1/3)

Avant d'installer le package NetworkX, vous devez installer le package Setuptools.

Si vous avez Python 2 \geq 2.7.9 ou bien Python 3 \geq 3.4, vous aurez déjà Pip et Setuptools. Ainsi, vous devez juste les mettre à jour.

- Sur Linux ou OS X :

```
pip install -U pip setuptools
```

- Sur Windows :

```
python -m pip install -U pip setuptools
```

Par contre, si vous avez une autre version de Python, visitez la page d'installation et suivez les instructions :

<https://packaging.python.org/installing/>

Installation (2/3)

A ce stade, il vous reste de faire une légère configuration : ajouter le chemin des packages Python à la variable PATH pour pouvoir utiliser Pip à partir de l'invite de commandes.

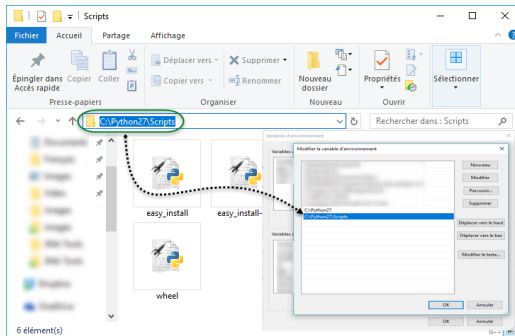
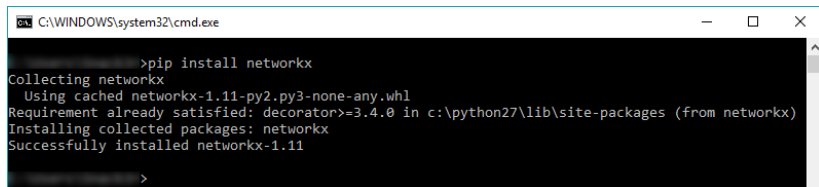


FIGURE: Configuration de l'outil Pip.

Installation (3/3)

Installer le package NetworkX en saisissant la commande suivante :

```
pip install networkx
```



```
C:\WINDOWS\system32\cmd.exe

>pip install networkx
Collecting networkx
  Using cached networkx-1.11-py2.py3-none-any.whl
Requirement already satisfied: decorator>=3.4.0 in c:\python27\lib\site-packages (from networkx)
Installing collected packages: networkx
Successfully installed networkx-1.11

>
```

FIGURE: Fenêtre d'installation du package NetworkX sur Windows 10.

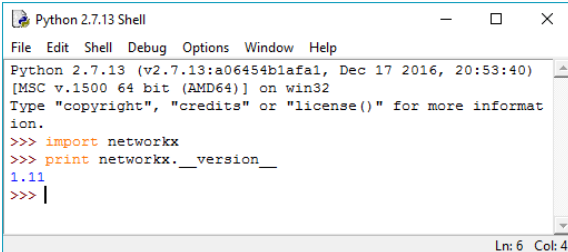
Premières pas

- Importer le package NetworkX

```
import networkx
```

- Afficher la version installée de NetworkX

```
print networkx.__version__
```



The screenshot shows a 'Python 2.7.13 Shell' window with a menu bar (File, Edit, Shell, Debug, Options, Window, Help). The terminal text is as follows:

```
Python 2.7.13 (v2.7.13:a06454b1afa1, Dec 17 2016, 20:53:40)
[MSC v.1500 64 bit (AMD64)] on win32
Type "copyright", "credits" or "license()" for more informat
ion.
>>> import networkx
>>> print networkx.__version__
1.11
>>> |
```

The status bar at the bottom right indicates 'Ln: 6 Col: 4'.

FIGURE: Fenêtre d'accueil de NetworkX.

Comparaison des outils d'ARS

Le tableau ci-dessous fournit une comparaison des outils précédemment présentés sur la base de 4 critères.

	Métriques & Manipulation de graphes	Visualisation de graphes	Détection de communautés	Prédiction de liens
iGraph	Parfait	Moyenne	Parfait	Faible
Gephi	Parfait	Parfait	Faible	Faible
NetworkX	Parfait	Faible	Faible	Parfait

TP à rendre

L'objectif de ce TP est de vous familiariser à la procédure d'installation des outils, du téléchargement jusqu'à l'utilisation. Pour ce faire, pour chaque outil, réalisez une vidéo illustrant la démarche d'installation. Vous devez obligatoirement commenter le processus d'installation tout au long de la vidéo.