

Structure de graphes des RL-

Comparaison formelle du RLF et du RLA

Sandrine Ollinger

13 juin 2013

Quels graphes ?

- Ensembles finis de lexies munis de relations de fonctions lexicales.

État du RLF

ordre	18 057
taille	25 932
composantes faiblement connexes	5 147
boucles	25
degré de connectivité	1,44
arcs symétriques	45,93%

Quels graphes ?

- Ensembles finis de lexies munis de relations de fonctions lexicales.

État du RLF

ordre	18 057
taille	25 932
composantes faiblement connexes	5 147
boucles	25
degré de connectivité	1,44
arcs symétriques	45,93%

Multigraphe orienté non
connexe.

Quels graphes ?

- Ensembles finis de lexies munis de relations de fonctions lexicales.

État du RLF

ordre	18 057
taille	25 932
composantes faiblement connexes	5 147
boucles	25
degré de connectivité	1,44
arcs symétriques	45,93%

Multigraphe orienté non
connexe.

État du RLA

206 976

Quels graphes ?

- Ensembles finis de lexies munis de relations de fonctions lexicales.

État du RLF

ordre	18 057
taille	25 932
composantes faiblement connexes	5 147
boucles	25
degré de connectivité	1,44
arcs symétriques	45,93%

Multigraphe orienté non
connexe.

État du RLA

206 976
847 799

Quels graphes ?

- Ensembles finis de lexies munis de relations de fonctions lexicales.

État du RLF

ordre	18 057
taille	25 932
composantes faiblement connexes	5 147
boucles	25
degré de connectivité	1,44
arcs symétriques	45,93%

Multigraphe orienté non
connexe.

État du RLA

206 976
847 799
60 183

Quels graphes ?

- Ensembles finis de lexies munis de relations de fonctions lexicales.

État du RLF

ordre	18 057
taille	25 932
composantes faiblement connexes	5 147
boucles	25
degré de connectivité	1,44
arcs symétriques	45,93%

Multigraphe orienté non
connexe.

État du RLA

206 976
847 799
60 183
0

Quels graphes ?

- Ensembles finis de lexies munis de relations de fonctions lexicales.

État du RLF

ordre	18 057
taille	25 932
composantes faiblement connexes	5 147
boucles	25
degré de connectivité	1,44
arcs symétriques	45,93%

Multigraphe orienté non
connexe.

État du RLA

206 976
847 799
60 183
0
4,10

Quels graphes ?

- Ensembles finis de lexies munis de relations de fonctions lexicales.

État du RLF

ordre	18 057
taille	25 932
composantes faiblement connexes	5 147
boucles	25
degré de connectivité	1,44
arcs symétriques	45,93%

Multigraphe orienté non
connexe.

État du RLA

206 976
847 799
60 183
0
4,10
99,77%

Quels graphes ?

- Ensembles finis de lexies munis de relations de fonctions lexicales.

État du RLF

ordre	18 057
taille	25 932
composantes faiblement connexes	5 147
boucles	25
degré de connectivité	1,44
arcs symétriques	45,93%

Multigraphe orienté non
connexe.

État du RLA

206 976
847 799
60 183
0
4,10
99,77%

Graphe orienté non connexe.

Quels outils ?

Yed

Editeur de graphes yWorks

(http://www.yworks.com/en/products_yed_about.html)

pedigree.py

Script Python développé par Emmanuel Navarro (IRIT, Toulouse)

Graphes « petits mondes » ?

Ni graphes aléatoires, ni graphes réguliers

- Relativement peu d'arcs/nombres de sommets.
- Coefficient d'agglomération élevé.
- Distribution des degrés suivant une loi de puissance.
- Longueur moyenne des plus courts chemins faible.

Graphes « petits mondes » ?

Ni graphes aléatoires, ni graphes réguliers

- Relativement peu d'arcs/nombres de sommets.
- Coefficient d'agglomération élevé.
- Distribution des degrés suivant une loi de puissance.
- Longueur moyenne des plus courts chemins faible.

Graphes « petits mondes » connus

- Mathématiciens ayant cosigné un article.
- Graphe physique du site web de l'Université Notre Dame (Indiana).
- Graphe des acteurs ayant joué dans un même film.
- Réseau d'interactions entre protéines.

Relativement peu d'arcs ?

Densité du RLF et du RLA :

	nœuds (n)	arcs		
RLF (28/05/13)	18 057	25 932		
RLA (28/05/13)	206 976	847 799		

Relativement peu d'arcs ?

Densité du RLF et du RLA :

	nœuds (n)	arcs	$n \log(n)$	
RLF (28/05/13)	18 057	25 932	76 862	
RLA (28/05/13)	206 976	847 799		

- 1/3 de ce qu'on observe dans les graphes « petits mondes » pour le RLF.

Relativement peu d'arcs ?

Densité du RLF et du RLA :

	nœuds (n)	arcs	$n \log(n)$	
RLF (28/05/13)	18 057	25 932	76 862	
RLA (28/05/13)	206 976	847 799	1 100 267	

- 1/3 de ce qu'on observe dans les graphes « petits mondes » pour le RLF.
- De l'ordre de ce qu'on observe dans les graphes « petits mondes » pour le RLA.

Relativement peu d'arcs ?

Densité du RLF et du RLA :

	nœuds (n)	arcs	$n \log(n)$	n^2
RLF (28/05/13)	18 057	25 932	76 862	326×10^6
RLA (28/05/13)	206 976	847 799	1 100 267	42×10^9

- 1/3 de ce qu'on observe dans les graphes « petits mondes » pour le RLF.
- De l'ordre de ce qu'on observe dans les graphes « petits mondes » pour le RLA.

Relativement peu d'arcs ?

Densité du RLF et du RLA :

	nœuds (n)	arcs	$n \log(n)$	n^2
RLF (28/05/13)	18 057	25 932	76 862	326×10^6
RLA (28/05/13)	206 976	847 799	1 100 267	42×10^9

- 1/3 de ce qu'on observe dans les graphes « petits mondes » pour le RLF.
- De l'ordre de ce qu'on observe dans les graphes « petits mondes » pour le RLA.
- Relativement peu d'arcs pour les 2 réseaux.

Coefficient d'agglomération

RLF

0,115

RLA

0,1042

Coefficient d'agglomération

RLF

0,115

RLA

0,1042

Premières impressions

Coefficient d'agglomération

RLF

0,115

RLA

0,1042

Premières impressions

- Faible coefficient d'agglomération dans les 2 réseaux.

Coefficient d'agglomération

RLF

0,115

RLA

0,1042

Premières impressions

- Faible coefficient d'agglomération dans les 2 réseaux.
- Des coefficients sensiblement identiques.

Coefficient d'agglomération

RLF

0,115

RLA

0,1042

Premières impressions

- Faible coefficient d'agglomération dans les 2 réseaux.
- Des coefficients sensiblement identiques.

Coefficients d'agglomération faibles dans les RL-

Pas de possibilité de supprimer une lexie sans modifier significativement les propriétés du réseau.

Coefficients d'agglomération faibles ?

RLF

0,115

RLF

aléatoire

0,0003

Coefficients d'agglomération faibles ?

RLF

0,115

RLF

aléatoire

0,0003

RLA

0,1042

RLA

aléatoire

0,000001

Coefficients d'agglomération faibles ?

RLF

0,115

RLF

aléatoire

0,0003

RLA

0,1042

RLA

aléatoire

0,000001

- Les coefficients d'agglomération sont bien plus élevés que ce qu'ils seraient pour des graphes aléatoires de mêmes ordres (n) et de mêmes tailles (m) ($2m/n^2$).

Coefficients d'agglomération faibles ?

RLF

0,115

RLF

aléatoire

0,0003

RLA

0,1042

RLA

aléatoire

0,000001

- Les coefficients d'agglomération sont bien plus élevés que ce qu'ils seraient pour des graphes aléatoires de mêmes ordres (n) et de mêmes tailles (m) ($2m/n^2$).
- Le coefficient d'agglomération du RLA est plus élevé que celui du RLA.

Coefficients d'agglomération faibles ?

RLF

0,115

RLF

aléatoire

0,0003

RLA

0,1042

RLA

aléatoire

0,000001

- Les coefficients d'agglomération sont bien plus élevés que ce qu'ils seraient pour des graphes aléatoires de mêmes ordres (n) et de mêmes tailles (m) ($2m/n^2$).
- Le coefficient d'agglomération du RLA est plus élevé que celui du RLA.

Coefficient d'agglomération élevé dans les RL-

Possibilité de supprimer une lexie sans modifier significativement les propriétés du réseau.

Degrés sortants

Degré
sortant
moyen RLF
3,4148

Degré
sortant
moyen RLA
7,2124

Degrés sortants

Degré
sortant
moyen RLF

3,4148

Degré
sortant
moyen RLA

7,2124

Degré moyen sortant dans les RL-

Le nombre moyen de liens de FL présents dans la description d'une lexie.

Distribution des degrés sortants

Distribution des degrés sortants

Associe à chaque entier k le nombre de nœuds du graphe ayant un degré sortant k .

Distribution des degrés sortants

Distribution des degrés sortants

Associe à chaque entier k le nombre de nœuds du graphe ayant un degré sortant k .

Distribution des degrés sortants suivant une loi de puissance dans les RL-

Distribution des degrés sortants

Distribution des degrés sortants

Associe à chaque entier k le nombre de nœuds du graphe ayant un degré sortant k .

Distribution des degrés sortants suivant une loi de puissance dans les RL-

- Un nombre important de liens de FL ont pour source un très petit ensemble de lexies. Les lexies ainsi liées à cet ensemble ont elles-mêmes davantage de liens de FL que d'autres, etc.

Distribution des degrés sortants

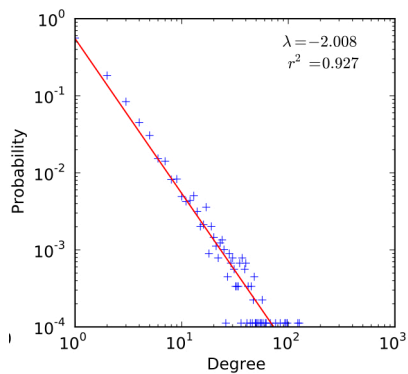
Distribution des degrés sortants

Associe à chaque entier k le nombre de nœuds du graphe ayant un degré sortant k .

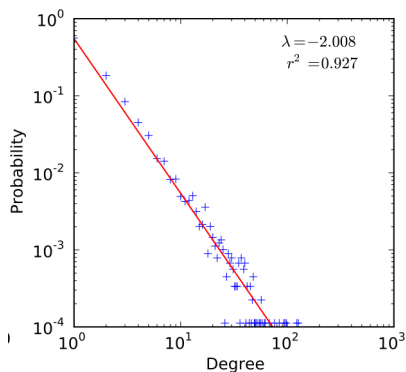
Distribution des degrés sortants suivant une loi de puissance dans les RL-

- Un nombre important de liens de FL ont pour source un très petit ensemble de lexies. Les lexies ainsi liées à cet ensemble ont elles-mêmes davantage de liens de FL que d'autres, etc.
- L'organisation des lexies est hiérarchique.

Distribution des degrés sortants du RLF

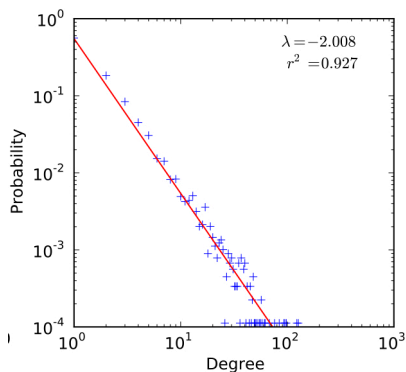


Distribution des degrés sortants du RLF



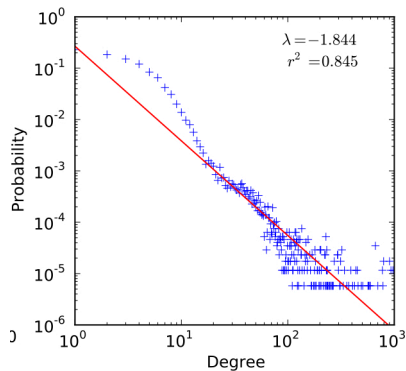
- La probabilité qu'un nœud du RLF ait k voisins est proportionnelle à $k^{-2.008}$.

Distribution des degrés sortants du RLF

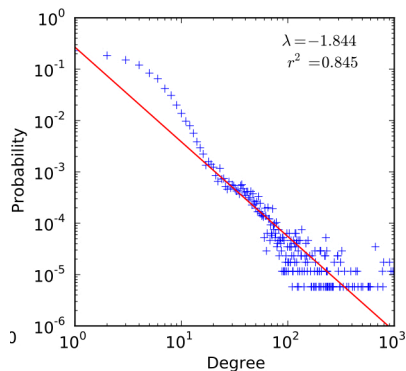


- La probabilité qu'un nœud du RLF ait k voisins est proportionnelle à $k^{-2.008}$.
- La distribution des degrés sortants du RLF est fortement corrélée à la loi de puissance correspondante.

Distribution des degrés sortants du RLA

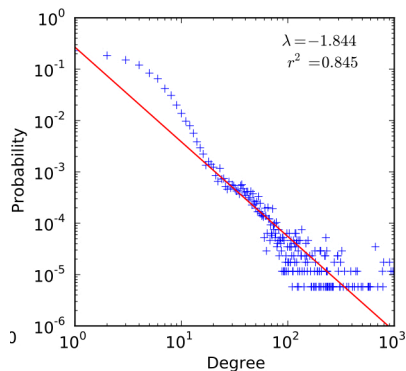


Distribution des degrés sortants du RLA



- La probabilité qu'un nœud du RLA ait k voisins est proportionnelle à $k^{-1.844}$.

Distribution des degrés sortants du RLA



- La probabilité qu'un nœud du RLA ait k voisins est proportionnelle à $k^{-1.844}$.
- Le RLA est moins bien corrélé à une loi de puissance que le RLF.

Faible moyenne des plus courts chemins ?

Faible moyenne des plus courts chemins dans les RL-

Facilité de passer d'une lexie à n'importe qu'elle autre en suivant les liens de FL.

Faible moyenne des plus courts chemins ?

Faible moyenne des plus courts chemins dans les RL-

Facilité de passer d'une lexie à n'importe quelle autre en suivant les liens de FL.

Difficulté

Mesure problématique pour les graphes de plus d'une composante.

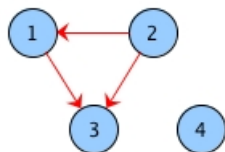
Faible moyenne des plus courts chemins ?

Faible moyenne des plus courts chemins dans les RL-

Facilité de passer d'une lexie à n'importe quelle autre en suivant les liens de FL.

Difficulté

Mesure problématique pour les graphes de plus d'une composante.



Longueurs des plus courts chemins

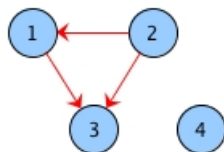
Faible moyenne des plus courts chemins ?

Faible moyenne des plus courts chemins dans les RL-

Facilité de passer d'une lexie à n'importe quelle autre en suivant les liens de FL.

Difficulté

Mesure problématique pour les graphes de plus d'une composante.



Longueurs des plus courts chemins

- (1,2) : 1
- (1,3) : 1
- (2,3) : 1

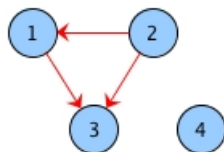
Faible moyenne des plus courts chemins ?

Faible moyenne des plus courts chemins dans les RL-

Facilité de passer d'une lexie à n'importe quelle autre en suivant les liens de FL.

Difficulté

Mesure problématique pour les graphes de plus d'une composante.



Longueurs des plus courts chemins

- (1,2) : 1
- (1,3) : 1
- (2,3) : 1
- (1,4), (2,4), (3,4) : ?

Conclusions

Graphes « petits mondes » ?

	RLF	RLA
Relativement peu d'arcs		
Coefficient d'agglomération		
Distribution des degrés sortants corrélée à une loi de puissance		
Longueur moyenne des plus courts chemins		

Graphes « petits mondes » ?

	RLF	RLA
Relativement peu d'arcs	oui	oui +
Coefficient d'agglomération		
Distribution des degrés sortants corrélée à une loi de puissance		
Longueur moyenne des plus courts chemins		

Les caractéristiques du RLF et du RLA les classent dans la catégorie des graphes « petits mondes ».

Graphes « petits mondes » ?

	RLF	RLA
Relativement peu d'arcs	oui	oui +
Coefficient d'agglomération	élevé	élevé +
Distribution des degrés sortants corrélée à une loi de puissance		
Longueur moyenne des plus courts chemins		

Les caractéristiques du RLF et du RLA les classent dans la catégorie des graphes « petits mondes ».

Graphes « petits mondes » ?

	RLF	RLA
Relativement peu d'arcs	oui	oui +
Coefficient d'agglomération	élevé	élevé +
Distribution des degrés sortants corrélée à une loi de puissance	oui +	oui
Longueur moyenne des plus courts chemins		

Les caractéristiques du RLF et du RLA les classent dans la catégorie des graphes « petits mondes ».





Conclusions

Graphes « petits mondes » ?

	RLF	RLA
Relativement peu d'arcs	oui	oui +
Coefficient d'agglomération	élevé	élevé +
Distribution des degrés sortants corrélée à une loi de puissance	oui +	oui
Longueur moyenne des plus courts chemins	?	?

Les caractéristiques du RLF et du RLA les classent dans la catégorie des graphes « petits mondes ».

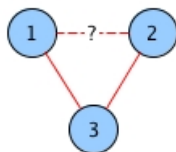
Références

-  Béla Bollobás and Oliver Riordan, *The diameter of a scale-free random graph*, *Combinatorica* **24** (2004), no. 1, 5–34.
-  Bruno Gaume, *Balades Aléatoires dans les petits mondes lexicaux*, *Information Interaction Intelligence* (2004).
-  M. E. J. Newman, *The structure and function of complex networks*, *SIAM REVIEW* **45** (2003), 167–256.
-  Duncan J. Watts and Steven H. Strogatz, *Collective dynamics of 'small-world' networks*, *Nature* **393** (1998), no. 6684, 440–442.

Coefficient d'agglomération

Coefficient d'agglomération

Probabilité que 2 nœuds voisins d'un même nœud soient eux-mêmes voisins.



Selon type de graphes

	Coefficient d'agglomération
aléatoires	faible
« petits mondes »	élevé
réguliers	élevé

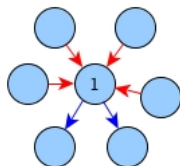
Degrés

Degrés d'un nœud

Degré global Nombre de liens reliés à un nœud.

Degré sortant Nombre de liens dont le nœud est la source.

Degré entrant Nombre de liens dont le nœud est la cible.



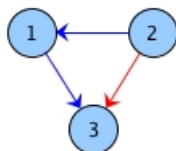
Degrés du nœud 1

- global : 6
- sortant : 2
- entrant : 4

Moyenne des plus courts chemins

Plus court chemin

Nombre minimal d'arcs à emprunter pour passer d'un noeud du graphe à un autre.



Selon type de graphes

	Moyenne des plus courts chemins
aléatoires	faible
« petits mondes »	faible
réguliers	élevée