

# Étude de l'omniprésence des propriétés petit-monde et sans-échelle

Vincent Labatut

Université Galatasaray, Département d'informatique  
Université d'Avignon, LIA  
`vincent.labatut@univ-avignon.fr`

MARAMI

Jeudi 16 octobre 2014

# Plan

## 1 Propriétés topologiques

## 2 Méthode

- Données
- Analyse

## 3 Résultats

- Mesures topologiques
- Relations

## 4 Conclusion

# Propriété petit-monde

- Étude de Watts & Strogatz [WS98]
- Basée sur deux mesures topologiques :
  - Mesure *globale* : distance géodésique moyenne  $L$
  - Mesure *locale* : transitivité ou coefficient de clustering  $T$

# Propriété petit-monde

- Étude de Watts & Strogatz [WS98]
- Basée sur deux mesures topologiques :
  - Mesure *globale* : distance géodésique moyenne  $L$
  - Mesure *locale* : transitivité ou coefficient de clustering  $T$

Mesure	Graphes réguliers	Réseaux complexes	Graphes aléatoires
$L$	Élevée		Faible
$T$	Élevée		Faible

# Propriété petit-monde

- Étude de Watts & Strogatz [WS98]
- Basée sur deux mesures topologiques :
  - Mesure *globale* : distance géodésique moyenne  $L$
  - Mesure *locale* : transitivité ou coefficient de clustering  $T$

Mesure	Graphes réguliers	Réseaux complexes	Graphes aléatoires
$L$	Élevée	Faible	Faible
$T$	Élevée	Élevée	Faible

# Propriété petit-monde

- Étude de Watts & Strogatz [WS98]
- Basée sur deux mesures topologiques :
  - Mesure *globale* : distance géodésique moyenne  $L$
  - Mesure *locale* : transitivité ou coefficient de clustering  $T$

Mesure	Graphes réguliers	Réseaux complexes	Graphes aléatoires
$L$	Élevée	Faible	Faible
$T$	Élevée	Élevée	Faible

- Effet notable : facilitation de la propagation

# Propriété sans-échelle

- Étude de Barabási & Albert [BA99]
- Basée sur la distribution du degré  $k$

# Propriété sans-échelle

- Étude de Barabási & Albert [BA99]
- Basée sur la distribution du degré  $k$

	<b>Graphes réguliers</b>	<b>Réseaux complexes</b>	<b>Graphes aléatoires</b>
Distribution de $k$ Valeur caractéristique	Dégénérée Oui		Poisson Oui

# Propriété sans-échelle

- Étude de Barabási & Albert [BA99]
- Basée sur la distribution du degré  $k$

	<b>Graphes réguliers</b>	<b>Réseaux complexes</b>	<b>Graphes aléatoires</b>
Distribution de $k$	Dégénérée	Puissance	Poisson
Valeur caractéristique	Oui	Non	Oui

# Propriété sans-échelle

- Étude de Barabási & Albert [BA99]
- Basée sur la distribution du degré  $k$

	<b>Graphes réguliers</b>	<b>Réseaux complexes</b>	<b>Graphes aléatoires</b>
Distribution de $k$	Dégénérée	Puissance	Poisson
Valeur caractéristique	Oui	Non	Oui

- Loi de puissance :  $P(k) \sim k^{-\alpha}$
- Effet notable : queue épaisse  $\rightarrow$  hubs
  - Résistance aux attaques aléatoires
  - Seuil épidémiologique bas (durée des épidémies accrue)

# Remise en question

- Idée répandue :
  - Omniprésence des propriétés petit-monde et sans-échelle
  - Est-ce fondé ?

# Remise en question

- Idée répandue :
  - Omniprésence des propriétés petit-monde et sans-échelle
  - Est-ce fondé ?
- Taille des données
  - Watts & Strogatz : 3 réseaux (IMDb, électricité, *C.elegans*)
  - Barabási & Albert : 3 réseaux (IMDb, électricité, WWW)
  - Études ultérieures : au plus une trentaine [AB02, HG08]

# Remise en question

- Idée répandue :
  - Omniprésence des propriétés petit-monde et sans-échelle
  - Est-ce fondé ?
- Taille des données
  - Watts & Strogatz : 3 réseaux (IMDb, électricité, *C.elegans*)
  - Barabási & Albert : 3 réseaux (IMDb, électricité, WWW)
  - Études ultérieures : au plus une trentaine [AB02, HG08]
- Outils d'analyse
  - Petit-monde : propriété qualitative → à quantifier
  - Sans-échelle : outils inadaptés/mal utilisés [CSN09]

# Remise en question

- Idée répandue :
  - Omniprésence des propriétés petit-monde et sans-échelle
  - Est-ce fondé ?
- Taille des données
  - Watts & Strogatz : 3 réseaux (IMDb, électricité, *C.elegans*)
  - Barabási & Albert : 3 réseaux (IMDb, électricité, WWW)
  - Études ultérieures : au plus une trentaine [AB02, HG08]
- Outils d'analyse
  - Petit-monde : propriété qualitative → à quantifier
  - Sans-échelle : outils inadaptés/mal utilisés [CSN09]
- Approche proposée
  - 1 Constituer une grande base de données de réseaux complexes
  - 2 Identifier des méthodes d'analyse fiables
  - 3 Les appliquer de façon uniforme aux données

# Constitution de la base de données

- Problèmes

- Dispersion sur le Web : chercheurs, institutions, éditeurs, logiciels...
- Aspect légal de l'utilisation des données
- Hétérogénéité des formats

# Constitution de la base de données

- Problèmes
  - Dispersion sur le Web : chercheurs, institutions, éditeurs, logiciels...
  - Aspect légal de l'utilisation des données
  - Hétérogénéité des formats
- Collection
  - 611 réseaux
  - Sources : auteurs/labos importants, SNAP, Konect, Pajek...
  - 7 classes de systèmes :

# Constitution de la base de données

- Problèmes
  - Dispersion sur le Web : chercheurs, institutions, éditeurs, logiciels...
  - Aspect légal de l'utilisation des données
  - Hétérogénéité des formats
- Collection
  - 611 réseaux
  - Sources : auteurs/labos importants, SNAP, Konect, Pajek...
  - 7 classes de systèmes :

<b>Classe</b>	<b>Nombre</b>	<b>Classe</b>	<b>Nombre</b>
Bibliographiques	55	Programmation	23
Biologiques	179	Sociologiques	140
Trophiques	31	Web	63
Institutionnels	28	Autres	93

# Prétraitement des données

- Information nécessaire : structure
  - Suppression des orientations (385 réseaux)
  - Suppression des poids (224)
  - Projection des réseaux bipartis (92)
  - Filtrage des liens des réseaux multiplexes (50)
- Dénominateur commun
  - Suppression des nœuds isolés (396)
  - Suppression des boucles (126)
  - Suppression des liens multiples (153)
- Filtrage des réseaux trop grands
  - Ex. Twitter  $n = 50 \times 10^6$ ,  $m = 2 \times 10^9$
- Collection après nettoyage : 598 réseaux basiques

# Propriété petit-monde

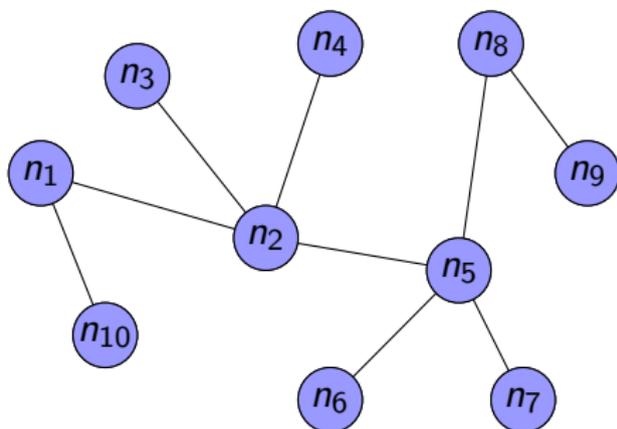
- Mesure  $\omega$  de Telesford *et al.* [TJH<sup>+</sup>11]
- Pour  $G$ , construction de 2 réseaux de référence :

# Propriété petit-monde

- Mesure  $\omega$  de Telesford *et al.* [TJH<sup>+</sup>11]
- Pour  $G$ , construction de 2 réseaux de référence :
  - $G_A \simeq$  graphe aléatoire

# Propriété petit-monde

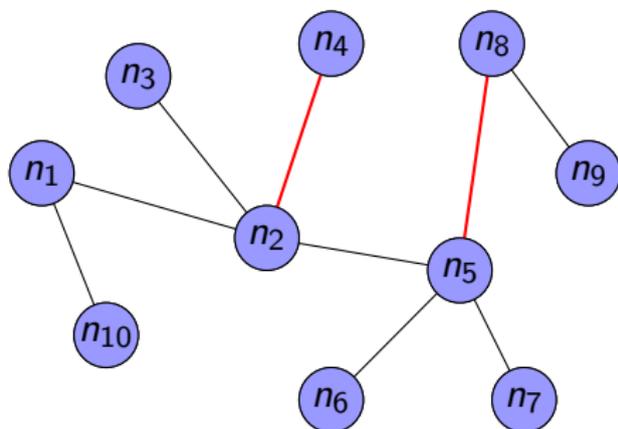
- Mesure  $\omega$  de Telesford *et al.* [TJH<sup>+</sup>11]
- Pour  $G$ , construction de 2 réseaux de référence :
  - $G_A \simeq$  graphe aléatoire



$$\begin{pmatrix} - & 1 & - & - & - & - & - & - & - & 1 \\ - & - & 1 & 1 & 1 & - & - & - & - & - \\ & - & - & - & - & - & - & - & - & - \\ & & - & - & - & - & - & - & - & - \\ & & & - & 1 & 1 & 1 & - & - & - \\ & & & & - & - & - & - & - & - \\ & & & & & - & - & - & - & - \\ & & & & & & - & 1 & - & - \\ & & & & & & & - & - & - \\ & & & & & & & & - & - \end{pmatrix}$$

# Propriété petit-monde

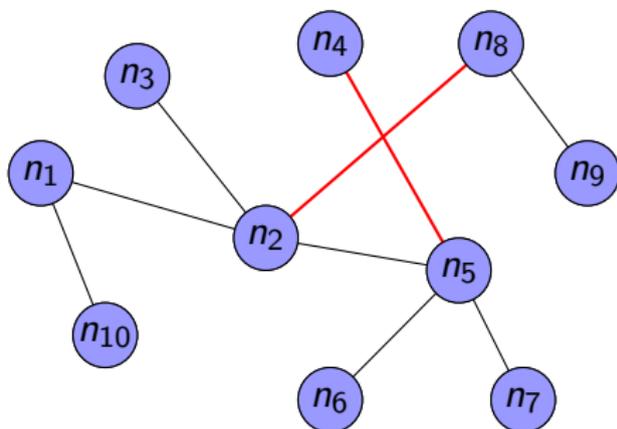
- Mesure  $\omega$  de Telesford *et al.* [TJH<sup>+</sup>11]
- Pour  $G$ , construction de 2 réseaux de référence :
  - $G_A \simeq$  graphe aléatoire



$$\begin{pmatrix} - & 1 & - & - & - & - & - & - & - & 1 \\ - & - & 1 & \mathbf{1} & \mathbf{1} & - & - & - & - & - \\ & - & - & - & - & - & - & - & - & - \\ & & - & - & - & - & - & - & - & - \\ & & & - & - & \mathbf{1} & \mathbf{1} & \mathbf{1} & - & - \\ & & & & - & - & - & - & - & - \\ & & & & & - & - & - & - & - \\ & & & & & & - & - & - & - \\ & & & & & & & - & \mathbf{1} & - \\ & & & & & & & & - & - \\ & & & & & & & & & - \end{pmatrix}$$

# Propriété petit-monde

- Mesure  $\omega$  de Telesford *et al.* [TJH<sup>+</sup>11]
- Pour  $G$ , construction de 2 réseaux de référence :
  - $G_A \simeq$  graphe aléatoire



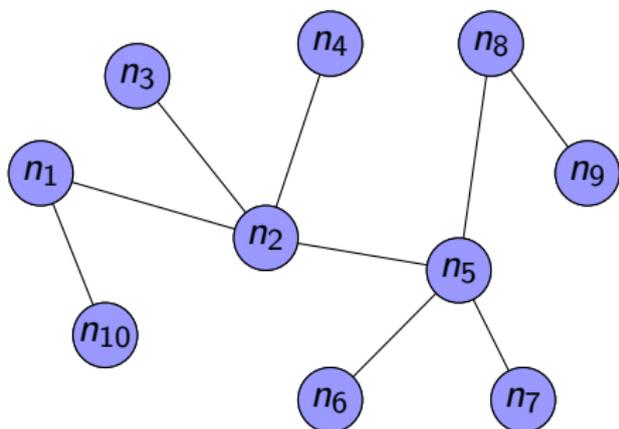
$$\begin{pmatrix} - & 1 & - & - & - & - & - & - & - & 1 \\ - & 1 & 0 & 1 & - & - & - & 1 & - & - \\ & - & - & - & - & - & - & - & - & - \\ & & - & 1 & - & - & - & - & - & - \\ & & & - & 1 & 1 & 0 & - & - & - \\ & & & & - & - & - & - & - & - \\ & & & & & - & - & - & - & - \\ & & & & & & - & - & 1 & - \\ & & & & & & & - & - & - \\ & & & & & & & & - & - \end{pmatrix}$$

# Propriété petit-monde

- Mesure  $\omega$  de Telesford *et al.* [TJH<sup>+</sup>11]
- Pour  $G$ , construction de 2 réseaux de référence :
  - $G_A \simeq$  graphe aléatoire
  - $G_T \simeq$  treillis

# Propriété petit-monde

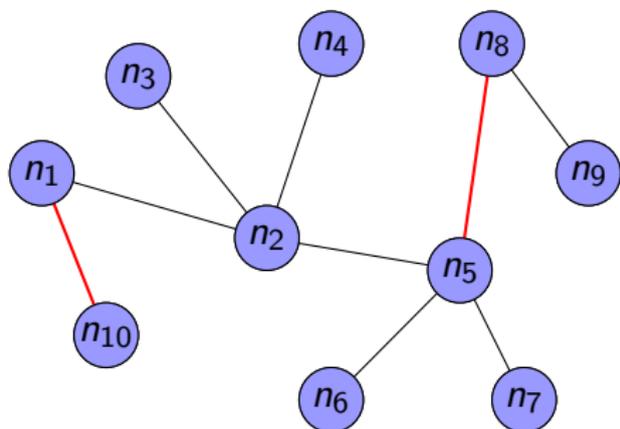
- Mesure  $\omega$  de Telesford *et al.* [TJH<sup>+</sup>11]
- Pour  $G$ , construction de 2 réseaux de référence :
  - $G_A \simeq$  graphe aléatoire
  - $G_T \simeq$  treillis



$$\begin{pmatrix} - & 1 & - & - & - & - & - & - & - & 1 \\ - & - & 1 & 1 & 1 & - & - & - & - & - \\ & - & - & - & - & - & - & - & - & - \\ & & - & - & - & - & - & - & - & - \\ & & & - & 1 & 1 & 1 & - & - & - \\ & & & & - & - & - & - & - & - \\ & & & & & - & - & - & - & - \\ & & & & & & - & 1 & - & - \\ & & & & & & & - & - & - \\ & & & & & & & & - & - \end{pmatrix}$$

# Propriété petit-monde

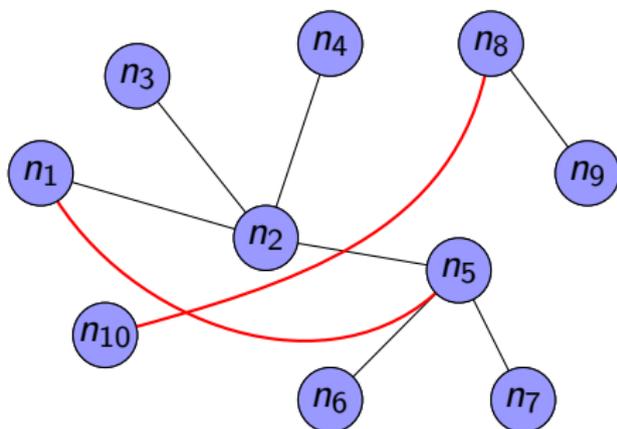
- Mesure  $\omega$  de Telesford *et al.* [TJH<sup>+</sup>11]
- Pour  $G$ , construction de 2 réseaux de référence :
  - $G_A \simeq$  graphe aléatoire
  - $G_T \simeq$  treillis



$$\begin{pmatrix} - & 1 & - & - & - & - & - & - & - & 1 \\ - & - & 1 & 1 & 1 & - & - & - & - & - \\ & - & - & - & - & - & - & - & - & - \\ & & - & - & - & - & - & - & - & - \\ & & & - & 1 & 1 & 1 & - & - & - \\ & & & & - & - & - & - & - & - \\ & & & & & - & - & - & - & - \\ & & & & & & - & - & - & - \\ & & & & & & & - & 1 & - \\ & & & & & & & & - & - \\ & & & & & & & & & - \end{pmatrix}$$

# Propriété petit-monde

- Mesure  $\omega$  de Telesford *et al.* [TJH<sup>+</sup>11]
- Pour  $G$ , construction de 2 réseaux de référence :
  - $G_A \simeq$  graphe aléatoire
  - $G_T \simeq$  treillis



$$\begin{pmatrix} - & 1 & - & - & 1 & - & - & - & - & 0 \\ - & - & 1 & 1 & 1 & - & - & - & - & - \\ - & - & - & - & - & - & - & - & - & - \\ - & - & - & - & - & - & - & - & - & - \\ - & - & - & - & - & - & - & - & - & - \\ - & - & - & - & - & 1 & 1 & 0 & - & - \\ - & - & - & - & - & - & - & - & - & - \\ - & - & - & - & - & - & - & - & - & - \\ - & - & - & - & - & - & - & - & 1 & 1 \\ - & - & - & - & - & - & - & - & - & - \end{pmatrix}$$

# Propriété petit-monde

- Mesure  $\omega$  de Telesford *et al.* [TJH<sup>+</sup>11]
- Pour  $G$ , construction de 2 réseaux de référence :
  - $G_A \simeq$  graphe aléatoire
  - $G_T \simeq$  treillis

# Propriété petit-monde

- Mesure  $\omega$  de Telesford *et al.* [TJH<sup>+</sup>11]
- Pour  $G$ , construction de 2 réseaux de référence :
  - $G_A \simeq$  graphe aléatoire
  - $G_T \simeq$  treillis
- Comparaison de ( $L$  avec  $L_A$ ) et ( $T$  avec  $T_T$ )

$$\omega(G) = \frac{L_A}{L} - \frac{T}{T_T}$$

# Propriété petit-monde

- Mesure  $\omega$  de Telesford *et al.* [TJH<sup>+</sup>11]
- Pour  $G$ , construction de 2 réseaux de référence :
  - $G_A \simeq$  graphe aléatoire
  - $G_T \simeq$  treillis
- Comparaison de ( $L$  avec  $L_A$ ) et ( $T$  avec  $T_T$ )

$$\omega(G) = \frac{L_A}{L} - \frac{T}{T_T}$$

- Interprétation :
  - Réseau  $G$  proche d'un treillis :  $L > L_A$  et  $T \simeq T_T \rightarrow \omega < 0$
  - Réseau  $G$  proche d'un graphe aléatoire :  $L \simeq L_A$  et  $T < T_T \rightarrow \omega > 0$
  - Réseau  $G$  petit-monde :  $L \simeq L_A$  et  $T \simeq T_T \rightarrow \omega \simeq 0$
  - $\omega \in [-0,5; +0,5] \rightarrow$  réseau petit monde [TJH<sup>+</sup>11]

# Propriété sans-échelle

- Enjeux :
  - Présence de hubs (propriétés fonctionnelles)
  - Validation du modèle générateur

# Propriété sans-échelle

- Enjeux :
  - Présence de hubs (propriétés fonctionnelles)
  - Validation du modèle générateur
- Risques de confusion :
  - Autre distributions à queue (moins) lourde
  - Loi de puissance partielle

# Propriété sans-échelle

- Enjeux :
  - Présence de hubs (propriétés fonctionnelles)
  - Validation du modèle générateur
- Risques de confusion :
  - Autre distributions à queue (moins) lourde
  - Loi de puissance partielle
- Méthode de Clauset *et al.* [CSN09]

# Propriété sans-échelle

- Enjeux :
  - Présence de hubs (propriétés fonctionnelles)
  - Validation du modèle générateur
- Risques de confusion :
  - Autre distributions à queue (moins) lourde
  - Loi de puissance partielle
- Méthode de Clauset *et al.* [CSN09]
  - ① Estimation des valeurs de  $\alpha$  et  $x_{min}$

# Propriété sans-échelle

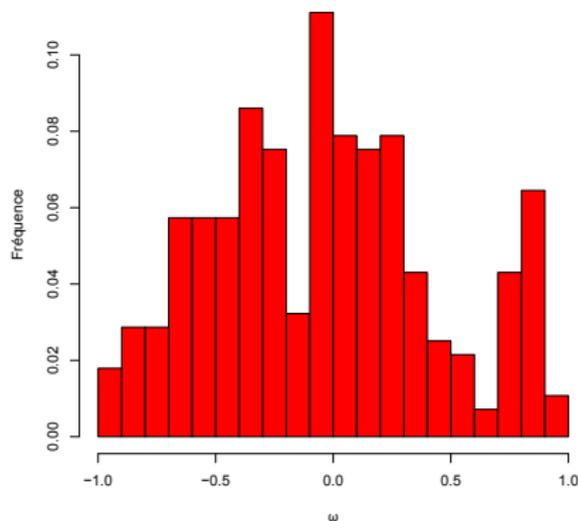
- Enjeux :
  - Présence de hubs (propriétés fonctionnelles)
  - Validation du modèle générateur
- Risques de confusion :
  - Autre distributions à queue (moins) lourde
  - Loi de puissance partielle
- Méthode de Clauset *et al.* [CSN09]
  - 1 Estimation des valeurs de  $\alpha$  et  $x_{min}$
  - 2 Test d'ajustement (KS) : plausibilité que les données suivent cette loi de puissance

# Propriété sans-échelle

- Enjeux :
  - Présence de hubs (propriétés fonctionnelles)
  - Validation du modèle générateur
- Risques de confusion :
  - Autre distributions à queue (moins) lourde
  - Loi de puissance partielle
- Méthode de Clauset *et al.* [CSN09]
  - 1 Estimation des valeurs de  $\alpha$  et  $x_{min}$
  - 2 Test d'ajustement (KS) : plausibilité que les données suivent cette loi de puissance
  - 3 Test du rapport de vraisemblance : comparaison à des distributions alternatives
    - Lois exponentielle, exponentielle étirée, log-normale
    - Loi de puissance tronquée
    - Loi de Poisson

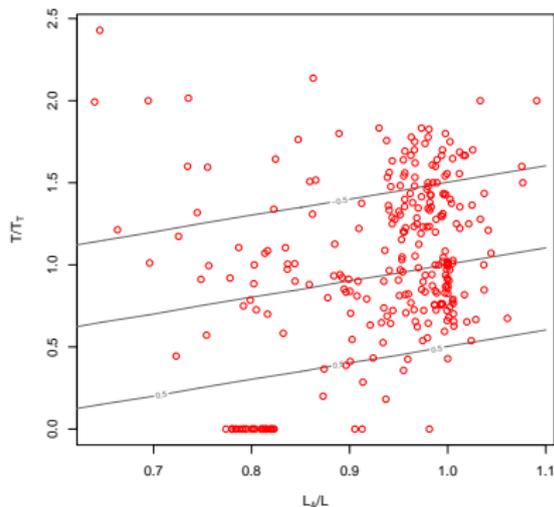
# Mesure $\omega$

- Coût calculatoire  $\rightarrow$  285 réseaux traités
- Mesure  $\omega$ 
  - $\omega \in [-0,5; +0,5]$  pour 65% des réseaux
  - Loi normale centrée sur 0



# Distance moyenne et transitivité

- Distance moyenne  $L$ 
  - $0,9 \leq L_A/L \leq 1,1$  pour 72% des réseaux
  - $L < L_A$  pour 80% des réseaux
- Transitivité  $T$ 
  - $0,5 \leq T/T_T \leq 1,5$  pour 67% des réseaux
  - Relativement symétrique par rapport à 1



# Distribution du degré

- 510 réseaux traités

Résultat	Réseaux
Improbable	
Modéré	
Probable	
Coupure	

# Distribution du degré

- 510 réseaux traités

Résultat	Réseaux
Improbable	81 (16%)
Modéré	6 (1%)
Probable	280 (55%)
Coupure	143 (28%)

# Distribution du degré

- 510 réseaux traités

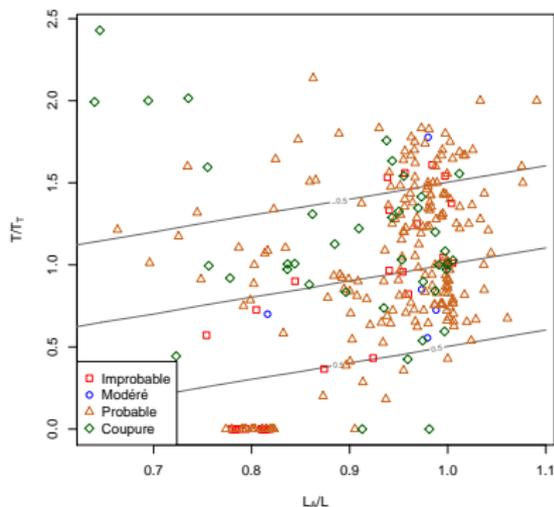
Résultat	Réseaux
Improbable	81 (16%)
Modéré	6 (1%)
Probable	280 (55%)
Coupure	143 (28%)

- 44% ne possèdent pas la propriété sans-échelle

# Mesure $\omega$ vs. distribution du degré

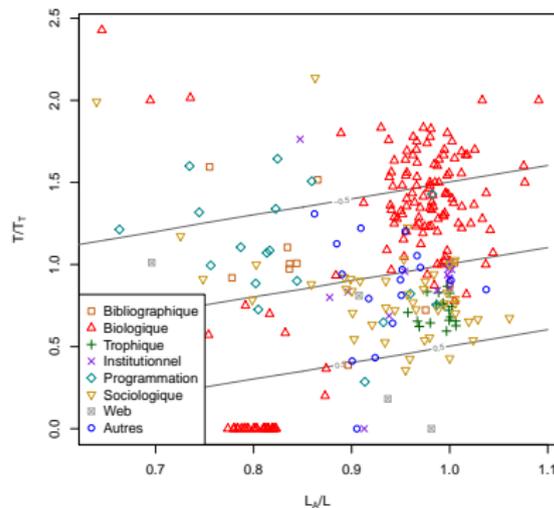
- 285 réseaux traités
- *Modéré* : très peu représenté
- *Improbable* : pas normal
- *Probable* et *coupure* : normaux, moyennes pas significativement différentes

Résultat	Réseaux
Improbable	37 (13%)
Modéré	6 (2%)
Probable	203 (71%)
Coupure	39 (13%)



# Classe de système vs. Mesure $\omega$

- Différences significatives :
  - Biologiques
  - Trophiques, Sociologiques



## Classe de système vs. distribution du degré

	<b>Improbable</b>	<b>Modéré</b>	<b>Probable</b>	<b>Coupure</b>
<b>Bibliographiques</b>	8	0	12	26
<b>Biologiques</b>	34	2	117	24
<b>Trophiques</b>	0	0	26	5
<b>Institutionnels</b>	7	1	11	9
<b>Programmation</b>	6	0	15	2
<b>Sociologiques</b>	9	3	57	38
<b>Web</b>	2	0	12	10
<b>Autres</b>	15	0	30	29

- Association mesurée :  $NMI \simeq 0.06$

# Conclusion

- Contributions

- Constitution du corpus
- Test à grande échelle des propriétés :
  - Petit-monde : largement répandue (mais pas omniprésente)
  - Sans-échelle : concerne seulement la moitié des réseaux considérés
  - Pas de relation forte entre les deux propriétés
  - Pas de relation forte avec le type de système

# Conclusion

## ● Contributions

- Constitution du corpus
- Test à grande échelle des propriétés :
  - Petit-monde : largement répandue (mais pas omniprésente)
  - Sans-échelle : concerne seulement la moitié des réseaux considérés
  - Pas de relation forte entre les deux propriétés
  - Pas de relation forte avec le type de système

## ● Perspectives

- Accroître la taille de la collection
- Considérer autres mesures topologiques (centralité, modularité...)
- Recherche systématique de relations entre mesures
- Identifier mesures caractéristiques d'une classe de systèmes
- Considérer classes de *relations* (et non pas *systèmes*)

# Références I

- [AB02] Reka Albert and Albert-Laszlo Barabási.  
Statistical mechanics of complex networks.  
*Reviews of Modern Physics*, 74(1) :47–96, 2002.
- [BA99] A.-L. Barabási and R Albert.  
Emergence of scaling in random networks.  
*Science*, 286(5439) :509, 1999.
- [CSN09] Aaron Clauset, Cosma Rohilla Shalizi, and M. E. J. Newman.  
Power-law distributions in empirical data.  
*SIAM Review*, 51(4) :661–703, 2009.

## Références II

- [HG08] Mark D. Humphries and Kevin Gurney.  
Network 'small-world-ness' : A quantitative method for determining canonical network equivalence.  
*PLoS ONE*, 3(4) :e0002051, 2008.
- [TJH<sup>+</sup>11] Qawi K. Telesford, Karen E. Joyce, Satoru Hayasaka, Jonathan H. Burdette, and Paul J. Laurienti.  
The ubiquity of small-world networks.  
*Brain Conn.*, 1 :367–375, 2011.
- [WS98] D. J. Watts and S. H. Strogatz.  
Collective dynamics of 'small-world' networks.  
*Nature*, 393(6684) :440–442, 1998.