

Structure et dynamique des graphes de terrain bipartis : liens internes et prédiction de liens

Oussama Allali

Clémence Magnien et Matthieu Latapy
<http://complexnetworks.fr>

LIP6 CNRS-UPMC



Contexte

Analyse : liens
et pairs
internes

Dynamique :
prédiction des
liens

Conclusion et
perspectives

- 1 Contexte
- 2 Analyse : liens et pairs internes
- 3 Dynamique : prédiction des liens
- 4 Conclusion et perspectives

Contexte

Analyse : liens
et pairs
internes

Dynamique :
prédiction des
liens

Conclusion et
perspectives

- 1 Contexte
- 2 Analyse : liens et pairs internes
- 3 Dynamique : prédiction des liens
- 4 Conclusion et perspectives

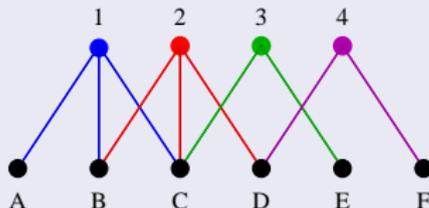
Contexte

Analyse : liens
et pairs
internes

Dynamique :
prédiction des
liens

Conclusion et
perspectives

Graphe biparti



- Acteurs-films.
- Auteurs-publications.
- Acheteur-produit.
- ...

Projection

Liens internes
et prédiction
de liens

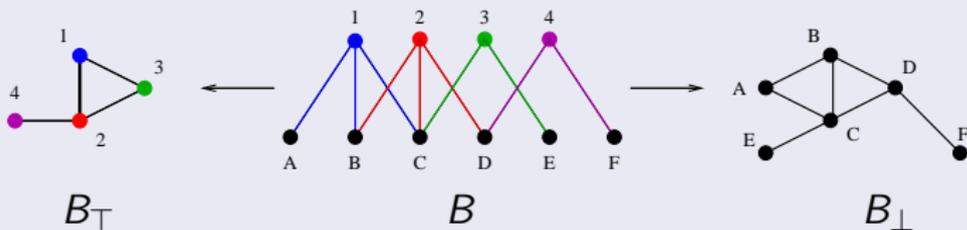
Contexte

Analyse : liens
et pairs
internes

Dynamique :
prédiction des
liens

Conclusion et
perspectives

Graphe biparti \rightarrow Graphe classique



- Perte d'information.

Projection

Liens internes
et prédiction
de liens

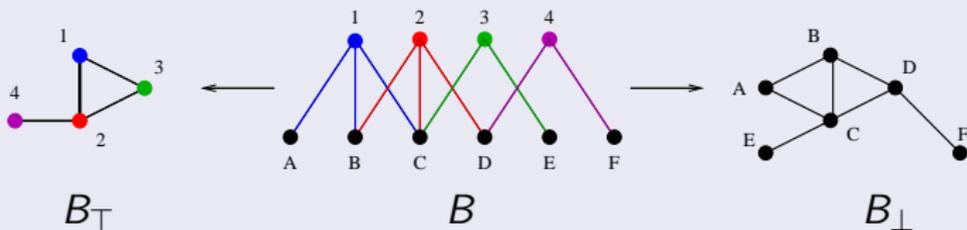
Contexte

Analyse : liens
et pairs
internes

Dynamique :
prédiction des
liens

Conclusion et
perspectives

Graphe biparti \rightarrow Graphe classique



- Perte d'information.

Projection valuée

Liens internes
et prédiction
de liens

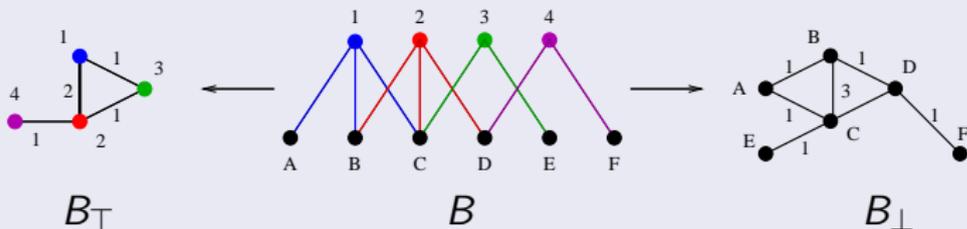
Contexte

Analyse : liens
et pairs
internes

Dynamique :
prédiction des
liens

Conclusion et
perspectives

Nombre de voisins en commun dans le biparti



Autres projections valuées

- Fraction de voisins en commun dans le biparti (*Jaccard*).
- Somme des votes des voisins en commun dans le biparti.

Projection valuée

Liens internes
et prédiction
de liens

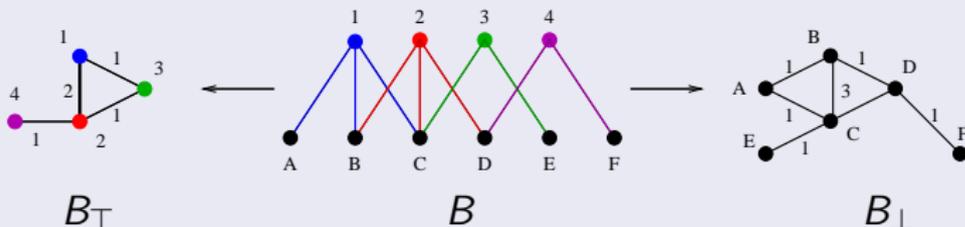
Contexte

Analyse : liens
et pairs
internes

Dynamique :
prédiction des
liens

Conclusion et
perspectives

Nombre de voisins en commun dans le biparti



Autres projections valuées

- Fraction de voisins en commun dans le biparti (*Jaccard*).
- Somme des votes des voisins en commun dans le biparti.

Contexte

Analyse : liens
et pairs
internes

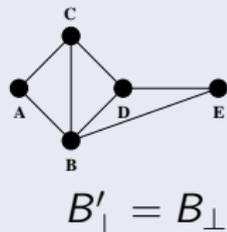
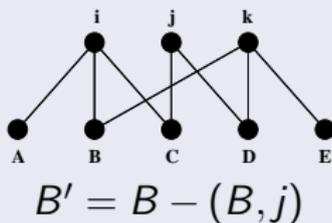
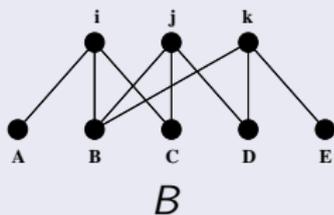
Dynamique :
prédiction des
liens

Conclusion et
perspectives

- Analyse des graphes de terrain bipartis :
liens internes.

- 1 Contexte
- 2 Analyse : liens et pairs internes**
- 3 Dynamique : prédiction des liens
- 4 Conclusion et perspectives

Exemple d'un lien \perp -interne.



- Retirer le lien ne change pas la projection.

Pairs internes

Liens internes
et prédiction
de liens

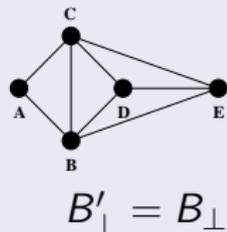
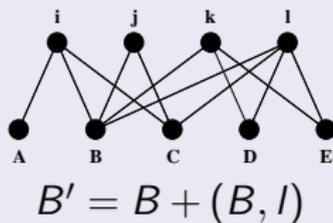
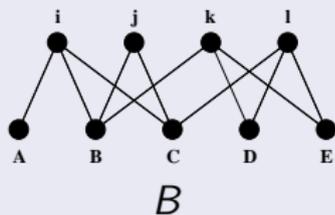
Contexte

Analyse : liens
et pairs
internes

Dynamique :
prédiction des
liens

Conclusion et
perspectives

Exemple d'une pair \perp -interne.



- Ajouter le lien ne change pas la projection.

- *Flickr-comts*
- *Flickr-favorites*
- *P2P-files*
- *PRL-papers*

	$f_{E_I}(\perp)$	$\frac{\mathcal{P}_I(\perp)}{\mathcal{P}_I^*(\perp)}$	$\frac{E_I(\perp)}{E_I^*(\perp)}$	$f_{E_I}(\top)$	$\frac{\mathcal{P}_I(\top)}{\mathcal{P}_I^*(\top)}$	$\frac{E_I(\top)}{E_I^*(\top)}$
<i>Flickr-comts</i>	0.398	0.258	4.22	0.002	0.151	22.0
<i>Flickr-favorites</i>	0.172	0.574	2.02	0.002	0.704	12.4
<i>P2P-files</i>	0.337	0.082	8.53	0.136	0.092	1430
<i>PRL-papers</i>	0.718	0.033	7.17	0.487	0.001	11.2

- f_{E_I} : fraction de liens internes.
- \mathcal{P}_I : nombre de paires internes.
- E_I : nombre de liens internes.
- \mathcal{P}_I^* et E_I^* : valeurs pour des graphe aléatoire même distributions des degrés.

Corrélation entre le nombre de liens internes et le degré des nœuds

Liens internes
et prédiction
de liens

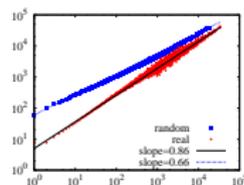
Contexte

Analyse : liens
et pairs
internes

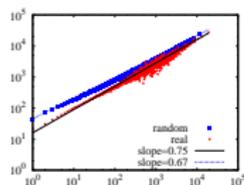
Dynamique :
prédiction des
liens

Conclusion et
perspectives

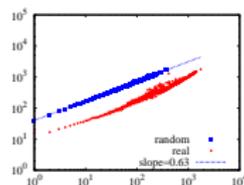
Flickr-comts



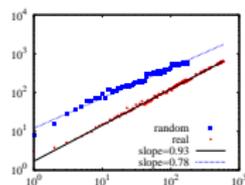
Flickr-favorites



P2P-files



PRL-papers



- Degré moyen en fonction du degré \perp -interne.

Suppression des liens internes

Liens internes
et prédiction
de liens

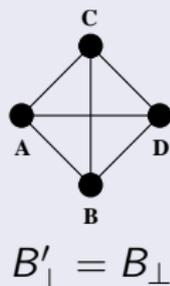
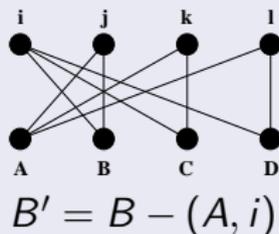
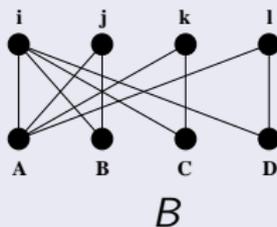
Contexte

Analyse : liens
et paires
internes

Dynamique :
prédiction des
liens

Conclusion et
perspectives

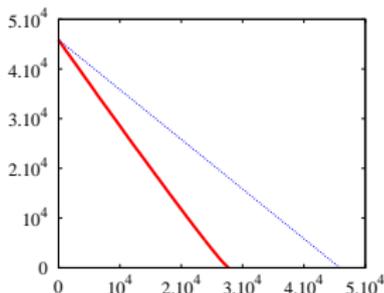
Exemple de suppression d'un lien \perp -interne.



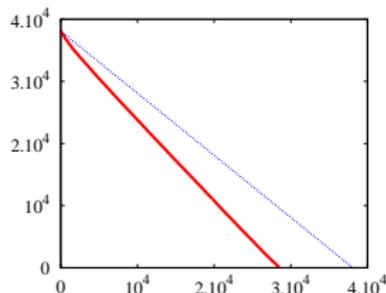
- $\{(A, i), (B, j), (C, k), (D, l)\}$ sont liens \perp -internes de B .
- La suppression du lien $(A, i) \implies \{(B, j), (C, k), (D, l)\}$ ne sont plus des liens internes.

P2P-files

⊥-internal links



⊤-internal links



- Nombre de liens internes restants en fonction du nombre de liens supprimer.
- La ligne rouge : processus de suppression aléatoire.
- La ligne bleue : borne supérieure théorique.

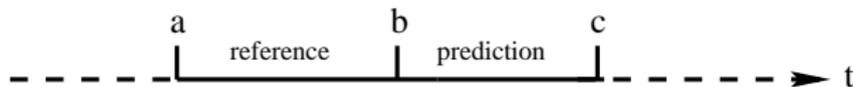
Contexte

Analyse : liens
et pairs
internes

Dynamique :
prédiction des
liens

Conclusion et
perspectives

- 1 Contexte
- 2 Analyse : liens et pairs internes
- 3 Dynamique : prédiction des liens
- 4 Conclusion et perspectives



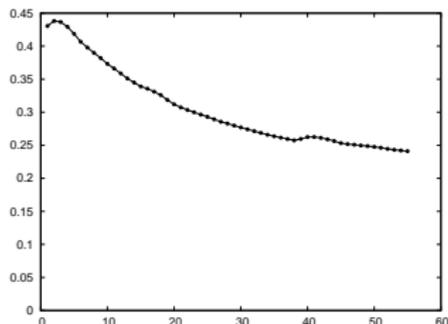
Méthodologie

- En entrée : B graphe de référence, période $[a, b[$.
- En sortie : P liens prédits comme apparaissant dans la période $[b, c[$.

Validation

- E les liens apparus dans la période $]b, c]$
- Précision $\frac{|P \cap E|}{|P|}$.
- Rappel $\frac{|P \cap E|}{|E|}$.

Période de prédiction



- % pairs internes en fonction de la période de prédiction.
- période de référence : $[0, 1\text{jour}[$.
- période de prédiction : $[1, x[$, pour $x = 2, \dots, 55$ jours.

Méthode de prédiction

- Ensemble de liens dans le projeté tels que :
 $L = \{(u, v) \in E_{\perp}, \text{poids}(u, v) \geq \text{seuil}\}.$
- On prédit les paires internes de B qui ont au moins un lien induit par ces derniers dans $L : P.$

Prédiction des liens internes

Liens internes
et prédiction
de liens

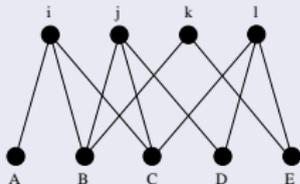
Contexte

Analyse : liens
et pairs
internes

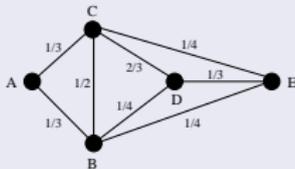
Dynamique :
prédiction des
liens

Conclusion et
perspectives

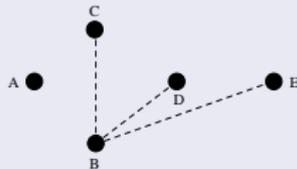
Exemple de prédiction d'une paire interne



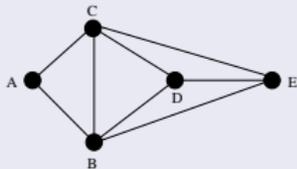
B



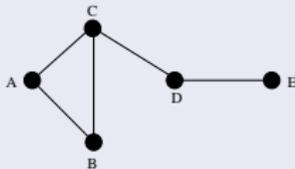
$B \perp, \text{Jaccard}$



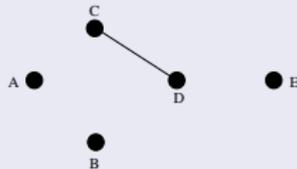
les liens induits
par (B, l)



seuil $\tau = \frac{1}{4}$

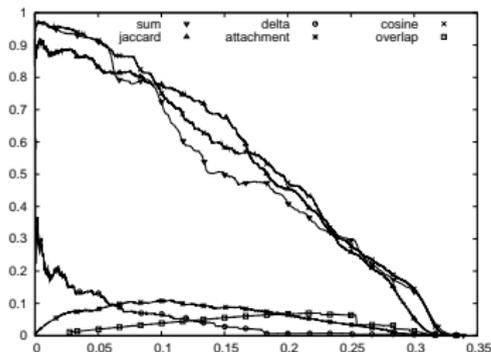


seuil $\tau = \frac{1}{3}$

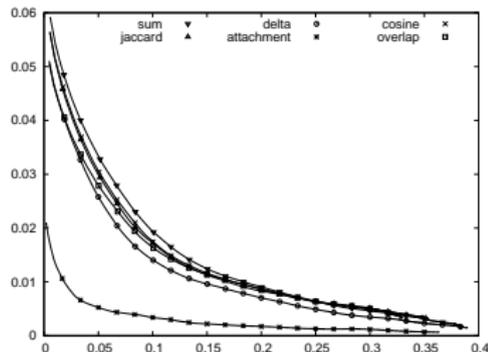


seuil $\tau = \frac{2}{3}$

prédiction de liens internes



filtrage collaboratif



- Précision en fonction du rappel.

Contexte

Analyse : liens
et pairs
internes

Dynamique :
prédiction des
liens

Conclusion et
perspectives

- 1 Contexte
- 2 Analyse : liens et pairs internes
- 3 Dynamique : prédiction des liens
- 4 Conclusion et perspectives

Conclusion

- Un nouveau type de liens : les liens internes.
- Approche par graphe évaluée pertinente.
- Comparer au filtrage collaboratif.

Perspectives.

- Autres jeux de données.
- Autres méthodes de pondération.
- Prédiction de paires non-internes.

Contexte

Analyse : liens
et pairs
internes

Dynamique :
prédiction des
liens

Conclusion et
perspectives

Merci