

UNIVERSITÉ DE TUNIS EL MANAR (LIPAH)
UNIVERSITÉ BLAISE PASCAL CLERMONT FERRAND II (LIMOS)

Extraction de quadri-concepts à partir de d-folksonomies : Application à la détection de tendances

Présenté par : Mohamed Nader Jelassi

16 mai 2012

Plan

- 1 Précédemment...en Mastère !
- 2 Le contexte quadratique
- 3 Nouvelles Notions et algorithme dédié
- 4 Quadri-concepts - Règles quadratiques
- 5 Application : Détection de tendances
- 6 Conclusion et Perspectives

Plan

- 1 Précédemment...en Mastère !
- 2 Le contexte quadratique
- 3 Nouvelles Notions et algorithme dédié
- 4 Quadri-concepts - Règles quadratiques
- 5 Application : Détection de tendances
- 6 Conclusion et Perspectives

Plan

- 1 Précédemment...en Mastère !
- 2 Le contexte quadratique
- 3 Nouvelles Notions et algorithme dédié
- 4 Quadri-concepts - Règles quadratiques
- 5 Application : Détection de tendances
- 6 Conclusion et Perspectives

Plan

- 1 Précédemment...en Mastère !
- 2 Le contexte quadratique
- 3 Nouvelles Notions et algorithme dédié
- 4 Quadri-concepts - Règles quadratiques
- 5 Application : Détection de tendances
- 6 Conclusion et Perspectives

Plan

- 1 Précédemment...en Mastère !
- 2 Le contexte quadratique
- 3 Nouvelles Notions et algorithme dédié
- 4 Quadri-concepts - Règles quadratiques
- 5 Application : Détection de tendances
- 6 Conclusion et Perspectives

Plan

- 1 Précédemment...en Mastère !
- 2 Le contexte quadratique
- 3 Nouvelles Notions et algorithme dédié
- 4 Quadri-concepts - Règles quadratiques
- 5 Application : Détection de tendances
- 6 Conclusion et Perspectives

Plan

- 1 Précédemment...en Mastère !
- 2 Le contexte quadratique
- 3 Nouvelles Notions et algorithme dédié
- 4 Quadri-concepts - Règles quadratiques
- 5 Application : Détection de tendances
- 6 Conclusion et Perspectives

Petite présentation

Qui suis-je ?

- Mohamed Nader Jelassi
- Faculté des Sciences à Tunis, Université Blaise Pascal à Clermont

Parcours

- 2009 Maitrise en Informatique (Major de promo)
- 2011 Mastère en Informatique
- 2012 - Thèse de Doctorat

Publications (Conférences et Revues)

- CORIA 2011** Auto-complétion de requêtes par une base générique de règles triadiques
- PAKDD 2012** Scalable mining of triadic concepts from folksonomies
- RNTI 2012** (Version longue du papier de CORIA)

Data Mining : Contexte diadique

Exemple

Objets/Attributs	α	β	γ
1	×	×	×
2	×	×	×
3		×	×

- Objets = $\{1,2,3\}$, Attributs = $\{\alpha,\beta,\gamma\}$.
- Fermés = **Concepts formels** ; Règles d'association
- APriori, Close, ...

Contexte triadique

Introduit par Ganter et Wille en 1995 par ajout d'une nouvelle dimension : **"la condition"**

Exemple

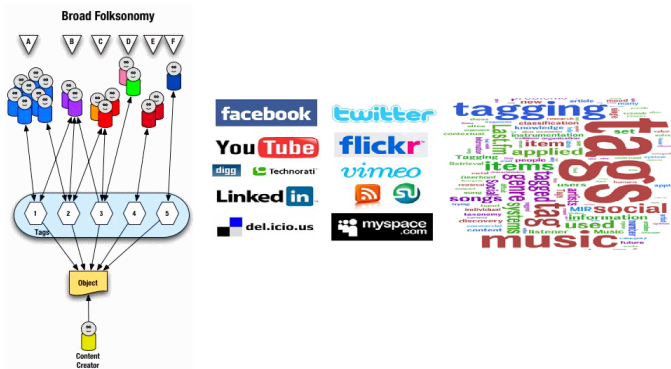
Objets/Attributs	α			β			γ		
Conditions	<i>a</i>	<i>b</i>	<i>c</i>	<i>a</i>	<i>b</i>	<i>c</i>	<i>a</i>	<i>b</i>	<i>c</i>
1	×	×	×	×	×		×	×	
2	×	×	×		×	×	×	×	
3	×	×	×		×	×	×	×	×

- Objets = $\{1,2,3\}$, Attributs = $\{\alpha,\beta,\gamma\}$, **Conditions** = $\{a,b,c\}$.
- Fermés = **Concepts triadiques**
- Règles d'association ?
- TriCons, Trias, ...

Contexte triadique et folksonomie

Contexte triadique : Intérêt croissant depuis 2006

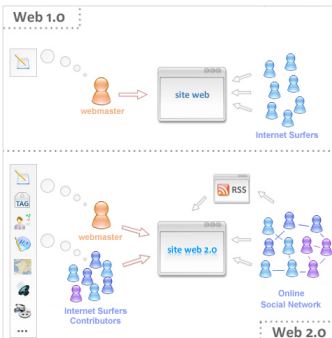
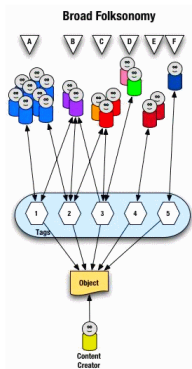
⇒ dû à la popularité des folksonomies qui imitent cette structure(3).
Objets = **Utilisateurs**, Attributs = **Tags**, Conditions = **Ressources**



Contexte triadique et folksonomie

Contexte triadique : Intérêt croissant

⇒ dû à la popularité des folksonomies qui imitent cette structure(3).
 Objets = **Utilisateurs**, Attributs = **Tags**, Conditions = **Ressources**



Contributions

Extraction efficace de tri-concepts

- Définition d'un opérateur de fermeture, **tri-générateur**, ...
- Algorithme TriCons (plus performant que ses concurrents : temps et mémoire)
- Publication à *PAKDD* 2012

Dérivation de règles d'association

- Définition d'une règle triadique, support et confiance conditionnés, base générique ...
- Ce type de règles **n'existait pas** dans la littérature
- Publication à *CORIA* 2011 (prix du meilleur papier)

Contributions

Extraction efficace de tri-concepts

- Définition d'un opérateur de fermeture, **tri-générateur**, ...
- Algorithme TriCons (plus performant que ses concurrents : temps et mémoire)
- Publication à *PAKDD* 2012

Dérivation de règles d'association

- Définition d'une règle triadique, support et confiance conditionnés, base générique ...
- Ce type de règles **n'existait pas** dans la littérature
- Publication à *CORIA* 2011 (prix du meilleur papier)

Contributions

Extraction efficace de tri-concepts

- Définition d'un opérateur de fermeture, **tri-générateur**, ...
- Algorithme TriCons (plus performant que ses concurrents : temps et mémoire)
- Publication à *PAKDD* 2012

Dérivation de règles d'association

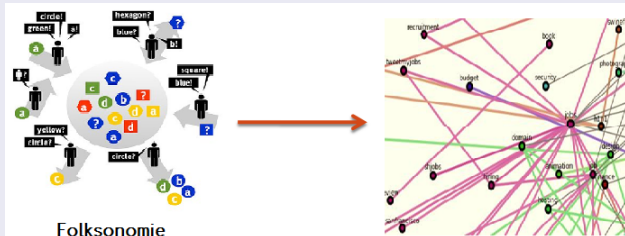
- Définition d'une règle triadique, support et confiance conditionnés, base générique ...
- Ce type de règles **n'existait pas** dans la littérature
- Publication à *CORIA* 2011 (prix du meilleur papier)

Résumé de l'approche



Folksonomie

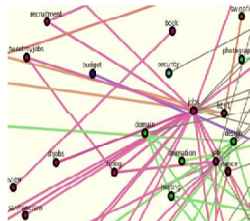
Résumé de l'approche



Résumé de l'approche

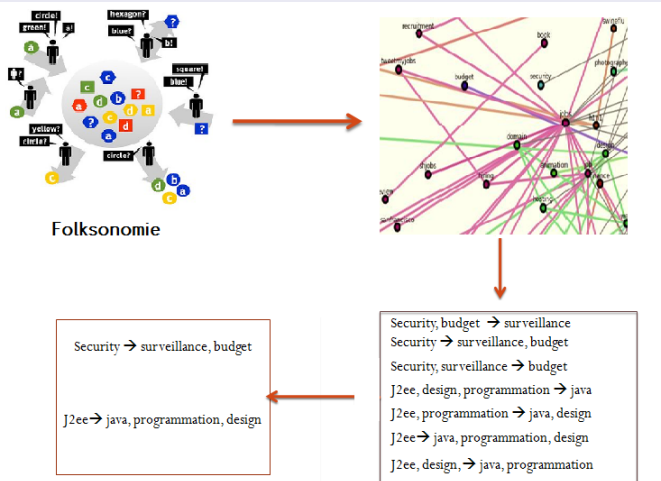


Folksonomie



Security, budget → surveillance
 Security → surveillance, budget
 Security, surveillance → budget
 J2ee, design, programmation → java
 J2ee, programmation → java, design
 J2ee → java, programmation, design
 J2ee, design, → java, programmation

Résumé de l'approche



Plan

- 1 Précédemment...en Mastère !
- 2 Le contexte quadratique**
- 3 Nouvelles Notions et algorithme dédié
- 4 Quadri-concepts - Règles quadratiques
- 5 Application : Détection de tendances
- 6 Conclusion et Perspectives

Contexte Quadratique

Objets	Attributs	α			β			γ		
Variables	Conditions	a	b	c	a	b	c	a	b	c
	1	×	×	×	×	×		×	×	
d	2	×	×	×		×	×	×	×	
	3	×	×	×		×	×	×	×	×
	1	×	×	×	×	×		×	×	
e	2	×	×	×		×	×	×	×	
	3	×	×	×		×	×	×	×	×

- Objets = $\{1,2,3\}$, Attributs = $\{\alpha,\beta,\gamma\}$, Conditions = $\{a,b,c\}$, Variables = $\{d, e\}$.
- Fermés = **Concepts quadratiques**
- Si (Objets, Attributs, Conditions) = (Utilisateurs, Tags, Ressources) = Folksonomie, Que peuvent bien représenter les **variables** ?

Contexte Quadratique

Objets	Attributs	α			β			γ		
Variables	Conditions	a	b	c	a	b	c	a	b	c
	1	×	×	×	×	×		×	×	
d	2	×	×	×		×	×	×	×	
	3	×	×	×		×	×	×	×	×
	1	×	×	×	×	×		×	×	
e	2	×	×	×		×	×	×	×	
	3	×	×	×		×	×	×	×	×

- Objets = $\{1,2,3\}$, Attributs = $\{\alpha,\beta,\gamma\}$, Conditions = $\{a,b,c\}$, Variables = $\{d, e\}$.
- Fermés = **Concepts quadratiques**
- Si (Objets, Attributs, Conditions) = (Utilisateurs, Tags, Ressources) = Folksonomie, Que peuvent bien représenter les **variables** ?

Motivation

Le temps dans les *folksonomies*

- Cette nouvelle variable/dimension, c'est le **temps** !
- Taggings \Rightarrow Temps (Dépendance naturelle)
- Associer chaque tagging à sa période (moment du tagging)

Motivation

- Les utilisateurs écoutent un artiste, regardent un film, vont sur un site selon la popularité de ces derniers...et des tags/notes affectés par les autres utilisateurs.
- Les sites (ou artistes, films, ...) les plus populaires changent d'une période à une autre !

Motivation

Le temps dans les *folksonomies*

- Cette nouvelle variable/dimension, c'est le **temps** !
- Taggings \Rightarrow Temps (Dépendance naturelle)
- Associer chaque tagging à sa période (moment du tagging)

Motivation

- Les utilisateurs écoutent un artiste, regardent un film, vont sur un site selon la popularité de ces derniers...et des tags/notes affectés par les autres utilisateurs.
- Les sites (ou artistes, films, ...) les plus populaires changent d'une période à une autre !

Exemple concret

Exemple : pour la ressource *www.tunisie.com*

$$\left\{ \begin{array}{l} 2006, \quad u_1, u_2, \text{ Tags : } \textit{Tourism, Travel}; \\ 2011, \quad u_1, u_2, \text{ Tags : } \textit{Revolution, Elections}. \end{array} \right.$$

Avant : Tri-Concept

- ① $\{(u_1, u_2), (\textit{Tourism, Travel, Revolution, Elections}), (\textit{tunisie.com})\}$
(sans information sur le temps ! Quelle sémantique?)

Maintenant : Quadri-Concept

- ① $\{(u_1, u_2), (\textit{Tourism, Travel}), (\textit{www.tunisie.com}), (\text{Mai 06})\}$
- ② $\{(u_1, u_2), (\textit{Revolution, Elections}), (\textit{www.tunisie.com}), (\text{Oct 11})\}$
(chaque quadri-concept est lié à une période précise de tagging.)

Exemple concret

Exemple : pour la ressource *www.tunisie.com*

$$\left\{ \begin{array}{l} 2006, \quad u_1, u_2, \text{ Tags : } \textit{Tourism, Travel}; \\ 2011, \quad u_1, u_2, \text{ Tags : } \textit{Revolution, Elections}. \end{array} \right.$$

Avant : Tri-Concept

- ① $\{(u_1, u_2), (\textit{Tourism, Travel, Revolution, Elections}), (\textit{tunisie.com})\}$
(sans information sur le temps ! Quelle sémantique?)

Maintenant : Quadri-Concept

- ① $\{(u_1, u_2), (\textit{Tourism, Travel}), (\textit{www.tunisie.com}), (\text{Mai 06})\}$
- ② $\{(u_1, u_2), (\textit{Revolution, Elections}), (\textit{www.tunisie.com}), (\text{Oct 11})\}$
(chaque quadri-concept est lié à une période précise de tagging.)

Challenges

But

- ① Regrouper sous forme de quadri-concepts les $\langle \text{user}, \text{tag}, \text{ressource}, \text{date} \rangle$ en commun \Rightarrow Extraction de motifs séquentiels multi-dimensionnels ?
- ② Etudier l'évolution des tendances dans les *folksonomies*
- ③ Connaître qui (ou qu'est-ce qui) fait le *buzz* à une période donnée
- ④ Déterminer les utilisateurs influenceurs, tags ou ressources populaires \Rightarrow Gain ou perte de popularité à travers le temps

CHALLENGE ACCEPTED



Plan

- 1 Précédemment...en Mastère !
- 2 Le contexte quadratique
- 3 Nouvelles Notions et algorithme dédié**
- 4 Quadri-concepts - Règles quadratiques
- 5 Application : Détection de tendances
- 6 Conclusion et Perspectives

Folksonomie et d-folksonomie

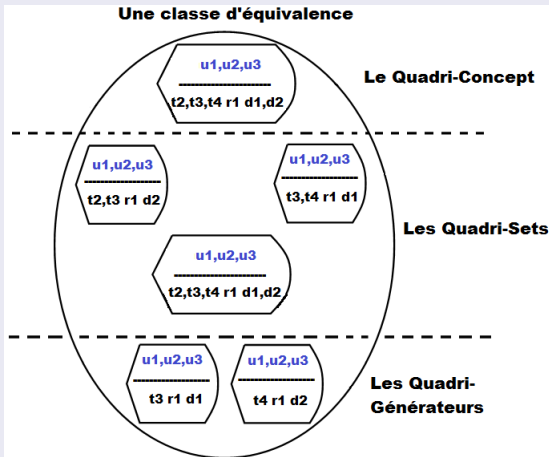
Folksonomie et d-folksonomie

$U/R-T$	r_1					r_2					r_3				
	t_1	t_2	t_3	t_4	t_5	t_1	t_2	t_3	t_4	t_5	t_1	t_2	t_3	t_4	t_5
u_1		x	x	x			x	x	x			x	x	x	
u_2		x	x	x			x	x	x			x	x	x	
u_3		x	x	x			x	x	x			x	x	x	
u_4									x						
u_5		x	x	x	x			x	x	x			x	x	x
u_6				x	x					x					
u_7	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x

\mathcal{F}	\mathcal{R}	$U-T$	r_1					r_2					r_3				
			t_1	t_2	t_3	t_4	t_5	t_1	t_2	t_3	t_4	t_5	t_1	t_2	t_3	t_4	t_5
		u_1		x	x	x			x	x	x			x	x	x	
		u_2		x	x	x			x	x	x			x	x	x	
		u_3		x	x	x			x	x	x			x	x	x	
		u_4									x						
		u_5		x					x	x	x			x	x	x	
		u_6				x	x	x		x	x	x					x
		u_7	x			x	x	x	x	x	x	x	x	x	x	x	x
		u_1		x	x	x			x	x	x			x	x	x	x
		u_2		x	x	x			x		x			x	x	x	
		u_3		x	x	x			x	x	x			x	x		
		u_4							x		x			x		x	
		u_5							x	x	x			x	x	x	
		u_6	x							x		x			x		x
		u_7			x				x	x	x			x	x	x	x

Classe d'équivalence / Quadri-générateur / Quadri-concept

Classes d'équivalences



Quadri-Concepts / opérateur de fermeture

Tri-Concepts / Quadri-Concepts

Tri-Concept $(U, T, R) \neq$ Quadri-Concept $(U, T, R, \textcolor{red}{D})$

Nouvel opérateur de fermeture

$S = (A, B, C, E)$: un quadri-set. Nous avons :

$h(S) = h(A, B, C, E) = (U, T, R, D) \mid U = \{u_i \in \mathcal{U} \mid (u_i, t_i, r_i, e_i) \in \mathcal{Y} \forall t_i \in B, \forall r_i \in C, \forall e_i \in E\}$

$\wedge T = \{t_i \in \mathcal{T} \mid (u_i, t_i, r_i, e_i) \in \mathcal{Y} \forall u_i \in U, \forall r_i \in C, \forall e_i \in E\}$

$\wedge R = \{r_i \in \mathcal{R} \mid (u_i, t_i, r_i, e_i) \in \mathcal{Y} \forall u_i \in U, \forall t_i \in T, \forall e_i \in E\}$

$\wedge \textcolor{red}{D} = \{\textcolor{red}{d}_i \in \mathcal{D} \mid (\textcolor{red}{d}_i, t_i, r_i, e_i) \in \mathcal{Y} \forall u_i \in U, \forall t_i \in T, \forall r_i \in R\}$

Plus généralement...

Définition d'un opérateur de fermeture pour un contexte à n dimensions ($n > 3$)

Quadri-Concepts / opérateur de fermeture

Tri-Concepts / Quadri-Concepts

Tri-Concept $(U, T, R) \neq$ Quadri-Concept (U, T, R, D)

Nouvel opérateur de fermeture

$S = (A, B, C, E)$: un quadri-set. Nous avons :

$h(S) = h(A, B, C, E) = (U, T, R, D) \mid U = \{u_i \in \mathcal{U} \mid (u_i, t_i, r_i, e_i) \in \mathcal{Y} \forall t_i \in B, \forall r_i \in C, \forall e_i \in E\}$

$\wedge T = \{t_i \in \mathcal{T} \mid (u_i, t_i, r_i, e_i) \in \mathcal{Y} \forall u_i \in U, \forall r_i \in C, \forall e_i \in E\}$

$\wedge R = \{r_i \in \mathcal{R} \mid (u_i, t_i, r_i, e_i) \in \mathcal{Y} \forall u_i \in U, \forall t_i \in T, \forall e_i \in E\}$

$\wedge D = \{d_i \in \mathcal{D} \mid (d_i, t_i, r_i, e_i) \in \mathcal{Y} \forall u_i \in U, \forall t_i \in T, \forall r_i \in R\}$

Plus généralement...

Définition d'un opérateur de fermeture pour un contexte à n dimensions ($n > 3$)

Quadri-Concepts / opérateur de fermeture

Tri-Concepts / Quadri-Concepts

Tri-Concept $(U, T, R) \neq$ Quadri-Concept (U, T, R, D)

Nouvel opérateur de fermeture

$S = (A, B, C, E)$: un quadri-set. Nous avons :

$h(S) = h(A, B, C, E) = (U, T, R, D) \mid U = \{u_i \in \mathcal{U} \mid (u_i, t_i, r_i, e_i) \in \mathcal{Y} \forall t_i \in B, \forall r_i \in C, \forall e_i \in E\}$

$\wedge T = \{t_i \in \mathcal{T} \mid (u_i, t_i, r_i, e_i) \in \mathcal{Y} \forall u_i \in U, \forall r_i \in C, \forall e_i \in E\}$

$\wedge R = \{r_i \in \mathcal{R} \mid (u_i, t_i, r_i, e_i) \in \mathcal{Y} \forall u_i \in U, \forall t_i \in T, \forall e_i \in E\}$

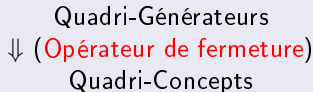
$\wedge D = \{d_i \in \mathcal{D} \mid (d_i, t_i, r_i, e_i) \in \mathcal{Y} \forall u_i \in U, \forall t_i \in T, \forall r_i \in R\}$

Plus généralement...

Définition d'un opérateur de fermeture pour un contexte à n dimensions ($n > 3$)

L'algorithme QuadriCons

Principe simple



Point Fort / Originalité

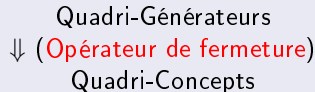
- 1 Les Quadri-Générateurs réduisent l'espace de recherche \Rightarrow Facilitent l'extraction des quadri-concepts ! (par augmentations successives)
- 2 Le recours à un opérateur de fermeture spécialement dédié au contexte quadratique
- 3 Non-redondance de calcul des quadri-concepts

Concurrents

- Data Peeler, Voutsadakis : extraction *brute* sans intérêt !

L'algorithme QuadriCons

Principe simple



Point Fort / Originalité

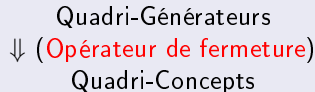
- 1 Les Quadri-Générateurs réduisent l'espace de recherche \Rightarrow Facilitent l'extraction des quadri-concepts ! (par augmentations successives)
- 2 Le recours à un opérateur de fermeture spécialement dédié au contexte quadratique
- 3 Non-redondance de calcul des quadri-concepts

Concurrents

- Data Peeler, Voutsadakis : extraction *brute* sans intérêt !

L'algorithme QuadriCons

Principe simple



Point Fort / Originalité

- 1 Les Quadri-Générateurs réduisent l'espace de recherche \Rightarrow Facilitent l'extraction des quadri-concepts ! (par augmentations successives)
- 2 Le recours à un opérateur de fermeture spécialement dédié au contexte quadratique
- 3 Non-redondance de calcul des quadri-concepts

Concurrents

- Data Peeler, Voutsadakis : extraction *brute* sans intérêt !

Plan

- 1 Précédemment...en Mastère !
- 2 Le contexte quadratique
- 3 Nouvelles Notions et algorithme dédié
- 4 Quadri-concepts - Règles quadratiques**
- 5 Application : Détection de tendances
- 6 Conclusion et Perspectives

Quadri-Concept : Exemple

Exemple de Quadri-Concept

\mathbb{F}_d	\mathcal{R}	r_1					r_2					r_3				
\mathcal{D}	\mathcal{U}/\mathcal{T}	t_1	t_2	t_3	t_4	t_5	t_1	t_2	t_3	t_4	t_5	t_1	t_2	t_3	t_4	t_5
	u_1		x	x	x			x	x	x			x	x	x	
	u_2		x	x	x			x	x	x			x	x	x	
	u_3			x	x			x	x	x	x		x	x	x	
d_1	u_4							x		x			x		x	
	u_5							x	x	x	x		x	x	x	
	u_6			x	x	x		x		x	x				x	
	u_7	x		x	x	x		x	x	x	x		x	x	x	x
	u_1		x	x	x			x	x	x			x	x	x	x
	u_2		x	x	x			x		x			x	x	x	x
	u_3		x	x	x			x	x	x	x		x	x		
d_2	u_4							x		x			x		x	
	u_5							x	x	x	x		x	x	x	
	u_6	x								x	x			x		x
	u_7			x				x	x	x			x	x	x	x

Etape 1

Localiser le quadri-générateur

Quadri-Concept : Exemple

Exemple de Quadri-Concept

\mathbb{F}_d	\mathcal{R}	r_1					r_2					r_3				
\mathcal{D}	\mathcal{U}/\mathcal{T}	t_1	t_2	t_3	t_4	t_5	t_1	t_2	t_3	t_4	t_5	t_1	t_2	t_3	t_4	t_5
	u_1		x	x	x			x	x	x			x	x	x	
	u_2		x	x	x			x	x	x			x	x	x	
	u_3			x	x			x	x	x	x		x	x	x	
d_1	u_4							x		x			x		x	
	u_5							x	x	x	x		x	x	x	
	u_6			x	x	x		x		x	x				x	
	u_7	x		x	x	x		x	x	x	x	x	x	x	x	x
	u_1		x	x	x			x	x	x			x	x	x	x
	u_2		x	x	x			x		x			x	x	x	x
	u_3		x	x	x			x	x	x	x		x	x		
d_2	u_4							x		x			x		x	
	u_5							x	x	x	x		x	x	x	
	u_6	x								x	x			x		x
	u_7			x				x	x	x			x	x	x	x

Etape 2

Maximiser l'ensemble de tags

Quadri-Concept : Exemple

Exemple de Quadri-Concept

\mathbb{F}_d	\mathcal{R}	r_1					r_2					r_3				
\mathcal{D}	\mathcal{U}/\mathcal{T}	t_1	t_2	t_3	t_4	t_5	t_1	t_2	t_3	t_4	t_5	t_1	t_2	t_3	t_4	t_5
	u_1		x	x	x		x	x	x			x	x	x	x	
	u_2		x	x	x		x	x	x			x	x	x	x	
	u_3			x	x		x	x	x	x		x	x	x	x	
d_1	u_4						x		x			x			x	
	u_5						x	x	x	x		x	x	x		
	u_6			x	x	x	x		x	x					x	
	u_7	x		x	x	x	x	x	x	x	x	x	x	x	x	x
	u_1		x	x	x		x	x	x			x	x	x	x	
	u_2		x	x	x		x		x			x	x	x	x	
	u_3		x	x	x		x	x	x	x		x	x			
d_2	u_4						x		x			x			x	
	u_5						x	x	x	x		x	x	x		
	u_6	x							x	x			x		x	
	u_7			x			x	x	x			x	x	x	x	x

Etape 3

Maximiser l'ensemble de ressources

Quadri-Concept : Exemple

Exemple de Quadri-Concept

$\mathbb{F}_d \mathcal{R}$	r_1					r_2					r_3					
\mathcal{D}	u_1	t_1	t_2	t_3	t_4	t_5	t_1	t_2	t_3	t_4	t_5	t_1	t_2	t_3	t_4	t_5
u_1		x	x	x				x	x	x			x	x	x	
		x	x	x				x	x	x			x	x	x	
			x	x				x	x	x	x		x	x	x	
d_1							x		x			x				
								x	x	x	x		x	x	x	
			x	x	x			x	x	x						
u_7	x		x	x	x	x	x	x	x	x	x	x	x	x	x	x
			x	x	x								x	x	x	x
			x	x	x								x	x	x	
u_2		x	x	x				x	x	x			x	x	x	
		x	x	x				x				x	x	x		
		x	x	x				x	x	x	x		x			
d_2							x		x				x			
								x	x	x	x		x	x	x	
		x							x	x				x		x
u_7			x					x	x	x			x	x	x	x

Etape 4

Maximiser l'ensemble de dates

Quadri-Concept obtenu

$\{(u_1, u_2), (t_1, t_2, t_3), (r_1, r_3), (d_1, d_2)\}$

Quadri-Concept : Exemple

Exemple de Quadri-Concept

$\mathbb{F}_d \mathcal{R}$	r_1					r_2					r_3				
$\mathcal{D} \quad \mathcal{U}/\mathcal{T}$	t_1	t_2	t_3	t_4	t_5	t_1	t_2	t_3	t_4	t_5	t_1	t_2	t_3	t_4	t_5
u_1		×	×	×			×	×	×			×	×	×	
u_2		×	×	×		×		×	×		×	×	×	×	
u_3			×	×		×	×	×	×		×	×	×	×	
$d_1 \quad u_4$						×			×		×			×	
u_5							×	×	×	×		×	×	×	
u_6			×	×	×		×		×	×					×
u_7	×		×	×	×	×	×	×	×	×	×	×	×	×	×
u_1		×	×	×			×	×	×			×	×	×	×
u_2		×	×	×		×		×	×		×	×	×	×	
u_3		×	×	×		×	×	×	×		×	×			
$d_2 \quad u_4$						×			×		×			×	
u_5							×	×	×	×		×	×	×	
u_6	×								×	×			×		×
u_7			×			×	×	×			×	×	×	×	×

Quadri-Concept ??

$\{(u_1, u_2), (t_1, t_2, t_3), (r_1, r_2, r_3), (d_1, d_2)\}$

Quadri-Concept : Exemple

Exemple de Quadri-Concept

$\mathbb{F}_d \mathcal{R}$	r_1					r_2					r_3				
$\mathcal{D} \quad \mathcal{U}/\mathcal{T}$	t_1	t_2	t_3	t_4	t_5	t_1	t_2	t_3	t_4	t_5	t_1	t_2	t_3	t_4	t_5
u_1		×	×	×			×	×	×			×	×	×	
u_2		×	×	×			×	×	×			×	×	×	
u_3			×	×			×	×	×			×	×	×	
$d_1 \quad u_4$							×		×			×		×	
u_5							×	×	×	×		×	×	×	
u_6			×	×	×		×		×	×					×
u_7	×		×	×	×	×	×	×	×	×	×	×	×	×	×
u_1		×	×	×			×	×	×			×	×	×	×
u_2		×	×	×			×		×			×	×	×	
u_3		×	×	×			×	×	×			×		×	
$d_2 \quad u_4$							×		×			×		×	
u_5							×	×	×	×		×	×	×	
u_6	×								×	×			×		×
u_7			×				×	×	×			×	×	×	×

Quadri-Concept ??

$\{(u_1, u_2, u_3), (t_1, t_2, t_3), (r_1, r_3), (d_1, d_2)\}$

Règles et mesures

2 types de règles

- Règles intra-classe : entre le quadri-générateur et le quadri-concept (d'une même classe)
- Règles inter-classe : entre un quadri-générateur et un quadri-concept (de deux classes différentes)

Quelles mesures ?

- Quelles mesures de qualité ?
- Extension de la définition de support et confiance conditionnés ?

Base Générique ?

- Définition d'une base générique de règles quadratiques ?

Quelles applications ?

Timeline

- 1 (Date 1) $tag_1 \Rightarrow tag_2$
- 2 (Date 2) $tag_1 \Rightarrow tag_5$

Application 1 : Analyse de tendances des tags (resp., utilisateurs, ressources)

Utilisateurs

- Quels utilisateurs sont les plus populaires/influents à une période donnée ?

Tags

- Quels tags sont les plus fréquemment associés à cette période ?

Ressources

- Quels ressources (sites web, films, artistes, ...) font le "buzz" en ce moment ?

Quelles applications ?

Timeline

- 1 (Date 1) $tag_1 \Rightarrow tag_2$
- 2 (Date 2) $tag_1 \Rightarrow tag_5$

Application 1 : Analyse de tendances des tags (resp., utilisateurs, ressources)

Utilisateurs

- Quels utilisateurs sont les plus populaires/influents à une période donnée ?

Tags

- Quels tags sont les plus fréquemment associés à cette période ?

Ressources

- Quels ressources (sites web, films, artistes, ...) font le "buzz" en ce moment ?

Quelles applications ?

Timeline

- 1 (Date 1) $tag_1 \Rightarrow tag_2$
- 2 (Date 2) $tag_1 \Rightarrow tag_5$

Application 1 : Analyse de tendances des tags (resp., utilisateurs, ressources)

Utilisateurs

- Quels utilisateurs sont les plus populaires/influents à une période donnée ?

Tags

- Quels tags sont les plus fréquemment associés à cette période ?

Ressources

- Quels ressources (sites web, films, artistes, ...) font le "buzz" en ce moment ?

Quelles applications ?

Timeline

- 1 (Date 1) $tag_1 \Rightarrow tag_2$
- 2 (Date 2) $tag_1 \Rightarrow tag_5$

Application 1 : Analyse de tendances des tags (resp., utilisateurs, ressources)

Utilisateurs

- Quels utilisateurs sont les plus populaires/influents à une période donnée ?

Tags

- Quels tags sont les plus fréquemment associés à cette période ?

Ressources

- Quels ressources (sites web, films, artistes, ...) font le "buzz" en ce moment ?

Motivations

Timeline

- ① (Date 1) $tag_1 \Rightarrow tag_2$
- ② (Date 2) $tag_1 \Rightarrow tag_5$

Application 2 : Filtrage par date (Recommandation de tags)

- L'utilisateur pourra trier par date les tags qui lui sont recommandés (choisir forcément la plus récemment affectée ?)

Motivations

Timeline

- ① (Date 1) $tag_1 \Rightarrow tag_2$
- ② (Date 2) $tag_1 \Rightarrow tag_5$

Application 3 : Détection des spammeurs

- (Mulder, Date 1) $tag_1 \Rightarrow tag_2$
- (Mulder, Date 2) $tag_1 \Rightarrow tag_2$
- \vdots
- (Mulder, Date n) $tag_1 \Rightarrow tag_2$

\Rightarrow L'utilisateur *Mulder* est susceptible d'être un spammeur !

Plan

- 1 Précédemment...en Mastère !
- 2 Le contexte quadratique
- 3 Nouvelles Notions et algorithme dédié
- 4 Quadri-concepts - Règles quadratiques
- 5 Application : Détection de tendances**
- 6 Conclusion et Perspectives

Application : Détection de tendances

Datasets

Plusieurs datasets intègrent la temporalité

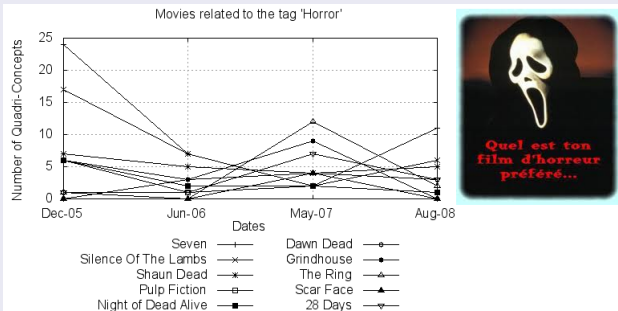
MovieLens utilisateurs, tags, films

LastFM utilisateurs, tags, artistes

Etat de l'art

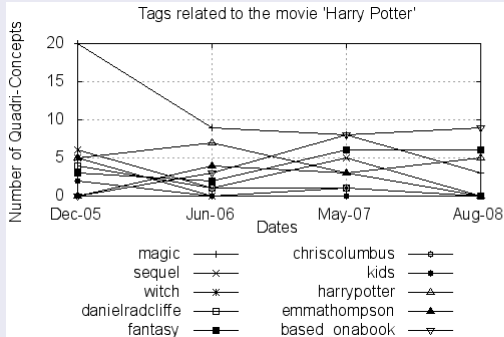
- Hotho *et al.*, Amitay *et al.*, ...
- **Problème majeur : perte d'information**
- \Rightarrow Faut Tenir compte des 4 dimensions : garder la dépendance utilisateurs-tags-ressources-dates !
(Etude sans perte d'information grâce aux Quadri-Concepts !)
- Plusieurs mesures utilisées...mais peu crédibles ! (ex : *interestingness*)

Quelques résultats...



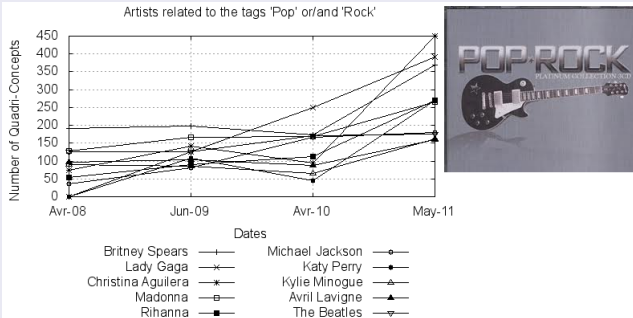
- Recommandation de ressources "*similaires*" (suivant un tag donné)
- Quels évènements influencent le gain (ou perte) de popularité des films ?

Quelques résultats...



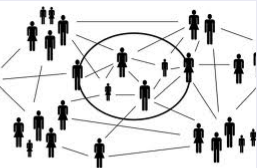
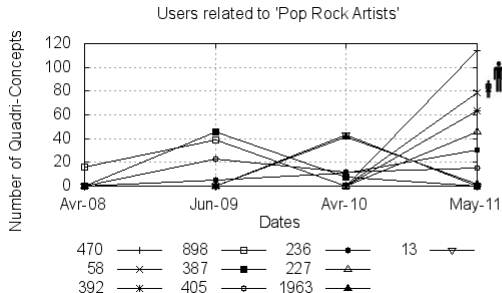
- Aide à la recherche d'information (query=harry potter)
- Recommandation des "trendy" tags les plus appropriés

Quelques résultats...



- Prédire les événements qui participent à l'augmentation de popularité d'un artiste (award, décès, ...)
- Expliquer le buzz produit par les artistes (nouvel album, mariage, vidéo censurée, ...)

Quelques résultats...



- Recommandation d'utilisateurs ayant les mêmes intérêts
- Détecter les utilisateurs les plus influents (cibler pour promouvoir des artistes)

Plan

- 1 Précédemment...en Mastère !
- 2 Le contexte quadratique
- 3 Nouvelles Notions et algorithme dédié
- 4 Quadri-concepts - Règles quadratiques
- 5 Application : Détection de tendances
- 6 Conclusion et Perspectives**

Conclusion et Perspectives

- Nouvelles notions pour un Contexte quadratique :
Quadri-générateur, Quadri-concept, ...
 - Opérateur de fermeture
 - Algorithme dédié à l'extraction des quadri-concepts
 - Etude de tendances/popularité dans les réseaux sociaux à travers le temps
-
- Dériver les règles intra-classe et inter-classe : chaque quadri-concept représentant une classe d'équivalence
 - Adaptation de mesures de qualité du cas triadique
 - Ordonner les quadri-concepts par inclusions/extensions
 - Trouver d'autres domaines d'application pour cette approche

That's all folks !