

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

RJCIA 2014

Une nouvelle approche pour l'extraction efficace
des quadri-concepts fréquents

Mohamed Nader Jelassi

1^{er} juillet 2014

- 1 Introduction et Motivations
- 2 Le Problème d'extraction des quadri-concepts fréquents
- 3 L'Algorithme QuadriCons
- 4 Résultats et Discussion
- 5 Conclusion et Perspectives

- 1 Introduction et Motivations
- 2 Le Problème d'extraction des quadri-concepts fréquents
- 3 L'Algorithme QuadriCons
- 4 Résultats et Discussion
- 5 Conclusion et Perspectives

- 1 Introduction et Motivations
- 2 Le Problème d'extraction des quadri-concepts fréquents
- 3 L'Algorithme QuadriCons
- 4 Résultats et Discussion
- 5 Conclusion et Perspectives

- 1 Introduction et Motivations
- 2 Le Problème d'extraction des quadri-concepts fréquents
- 3 L'Algorithme QuadriCons
- 4 Résultats et Discussion
- 5 Conclusion et Perspectives

- 1 Introduction et Motivations
- 2 Le Problème d'extraction des quadri-concepts fréquents
- 3 L'Algorithme QuadriCons
- 4 Résultats et Discussion
- 5 Conclusion et Perspectives

Une nouvelle
approche pour
l'extraction
efficace des
quadri-
concepts
fréquents

Introduction et
Motivations

Le Problème
d'extraction des
quadri-concepts
fréquents

L'Algorithme
QuadriCons

Résultats et
Discussion

Conclusion et
Perspectives

1 Introduction et Motivations

2 Le Problème d'extraction des quadri-concepts fréquents

3 L'Algorithme QuadriCons

4 Résultats et Discussion

5 Conclusion et Perspectives

Folksonomie = Porte-Manteau

Une nouvelle
approche pour
l'extraction
efficace des
quadri-
concepts
fréquents

Introduction et
Motivations

Le Problème
d'extraction des
quadri-concepts
fréquents

L'Algorithme
QuadriCons

Résultats et
Discussion

Conclusion et
Perspectives

Utilisateurs

FOLK



Done by people

Organisation de Données

TAXONOMY



Classifications of items

3 acteurs principaux

Une nouvelle
approche pour
l'extraction
efficace des
quadri-
concepts
fréquents

Introduction et
Motivations

Le Problème
d'extraction des
quadri-concepts
fréquents

L'Algorithme
QuadriCons

Résultats et
Discussion

Conclusion et
Perspectives

Contexte triadique : Utilisateurs-Tags-Ressources



UTILISATEURS



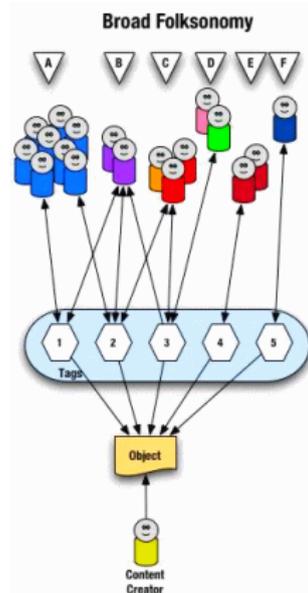
TAGS



RESSOURCES

Folksonomie

- L'internaute participe par le biais de mots clés : **les tags**
- Il partage et annote des ressources
- ⇒ L'ensemble des utilisateurs, tags et ressources forment une **folksonomie**.



Une nouvelle
approche pour
l'extraction
efficace des
quadri-
concepts
fréquents

Introduction et
Motivations

Le Problème
d'extraction des
quadri-concepts
fréquents

L'Algorithme
QuadriCons

Résultats et
Discussion

Conclusion et
Perspectives

- Surcharge d'information : Millions de données - Utilisateurs en nombre exponentiel
- Taille énorme des folksonomies
- Manque d'organisation des données

Problématiques

Une nouvelle
approche pour
l'extraction
efficace des
quadri-
concepts
fréquents

Introduction et
Motivations

Le Problème
d'extraction des
quadri-concepts
fréquents

L'Algorithme
QuadriCons

Résultats et
Discussion

Conclusion et
Perspectives



- Intelligence collective : Oui, mais à quel prix !
- Organiser ce flux important de données \implies Éviter l'anarchie !

Solution

Recours à la fouille de données (Data Mining) pour la recherche d'information dans les *folksonomies* : **Folksonomy Mining**

Problématiques

Une nouvelle
approche pour
l'extraction
efficace des
quadri-
concepts
fréquents

Introduction et
Motivations

Le Problème
d'extraction des
quadri-concepts
fréquents

L'Algorithme
QuadriCons

Résultats et
Discussion

Conclusion et
Perspectives



- Intelligence collective : Oui, mais à quel prix !
- Organiser ce flux important de données \implies Éviter l'anarchie !

Solution

Recours à la fouille de données (Data Mining) pour la recherche d'information dans les *folksonomies* : **Folksonomy Mining**

Solution : Représentations concises de la folksonomie

Une nouvelle
approche pour
l'extraction
efficace des
quadri-
concepts
fréquents

Introduction et
Motivations

Le Problème
d'extraction des
quadri-concepts
fréquents

L'Algorithme
QuadriCons

Résultats et
Discussion

Conclusion et
Perspectives

Concepts Triadiques

- Ensembles maximaux (Utilisateurs, Tags, Ressources)
- Représentation réduite sans perte d'information
- Méthodes existantes
 - Trias (Stumme et al., 2006)
 - CubeMiner (Jin et al., 2006)
 - Data Peeler (Cerf et al., 2009)
 - TriCons (Jelassi et al., 2011)

Data Peeler (Cerf *et al.*, 2009)

Une nouvelle
approche pour
l'extraction
efficace des
quadri-
concepts
fréquents

Introduction et
Motivations

Le Problème
d'extraction des
quadri-concepts
fréquents

L'Algorithme
QuadriCons

Résultats et
Discussion

Conclusion et
Perspectives

Modus Operandi (pour $n=4$)

Un noeud : un couple de quadri-sets

Pour chaque noeud :

- 1 Le partitionner suivant un élément à énumérer
- 2 est-ce un quadri-concept ?
 - Si **oui** : l'extraire + continuer la(les) décomposition(s)
 - Si **non** : continuer la(les) décomposition(s)

...jusqu'à "ensembles vides" sur tous les noeuds

Remarque

- L'exécution est indépendante de l'ordre des éléments à énumérer

Avantage et Inconvénients

Une nouvelle
approche pour
l'extraction
efficace des
quadri-
concepts
fréquents

Introduction et
Motivations

Le Problème
d'extraction des
quadri-concepts
fréquents

L'Algorithme
QuadriCons

Résultats et
Discussion

Conclusion et
Perspectives

Avantage

- Parcours rapide de l'arbre binaire (la structure utilisée)

Inconvénients

- Approche en profondeur d'abord
- Adapté surtout aux petites bases (problèmes de mémoire)

Conclusions

- Risque d'avoir les mêmes problèmes si on instancie le problème à $n=4$
- Nécessité de développer une méthode propre aux contextes quadratiques, adaptée à notre solution

Proposition : Folksonomy Mining + Une 4e dimension

Une nouvelle
approche pour
l'extraction
efficace des
quadri-
concepts
fréquents

Introduction et
Motivations

Le Problème
d'extraction des
quadri-concepts
fréquents

L'Algorithme
QuadriCons

Résultats et
Discussion

Conclusion et
Perspectives

Pourquoi?

- Utile dans plusieurs applications : systèmes de recommandation, étude temporelle des folksonomies, etc.

A quoi correspond la nouvelle dimension ?

- Le Temps
- Le Profil des utilisateurs

Un algorithme dédié

- QuadriCons

Une nouvelle
approche pour
l'extraction
efficace des
quadri-
concepts
fréquents

Introduction et
Motivations

Le Problème
d'extraction des
quadri-concepts
fréquents

L'Algorithme
QuadriCons

Résultats et
Discussion

Conclusion et
Perspectives

- 1 Introduction et Motivations
- 2 Le Problème d'extraction des quadri-concepts fréquents
- 3 L'Algorithme QuadriCons
- 4 Résultats et Discussion
- 5 Conclusion et Perspectives

- Une p-folksonomie est un ensemble de tuples $\mathcal{F}_p = (\mathcal{U}, \mathcal{T}, \mathcal{R}, \mathcal{P}, Y)$ où $\mathcal{U}, \mathcal{T}, \mathcal{R}$ et \mathcal{P} sont des ensembles finis dont les éléments sont appelés **utilisateurs**, **tags**, **ressources** et **profils**. $Y \subseteq \mathcal{U} \times \mathcal{T} \times \mathcal{R} \times \mathcal{P}$ représente une relation quadratique où chaque élément $y \subseteq Y$ peut être représenté par un quadruplet : $y = \{(u, t, r, p) \mid u \in \mathcal{U}, t \in \mathcal{T}, r \in \mathcal{R}, p \in \mathcal{P}\}$ ce qui veut dire que l'utilisateur u avec le profil p a annoté la ressource r via le tag t .
- Un quadri-set est une relation quadratique (A, B, C, D) avec $A \subseteq \mathcal{U}, B \subseteq \mathcal{T}, C \subseteq \mathcal{R}$ et $D \subseteq \mathcal{P}$
- Nous avons modélisé la 4e dimension par le **profil** afin d'extraire des résultats à partir d'un jeu de données du monde réel (MovieLens)

Contexte Quadratique

Une nouvelle
approche pour
l'extraction
efficace des
quadri-
concepts
fréquents

Introduction et
Motivations

Le Problème
d'extraction des
quadri-concepts
fréquents

L'Algorithme
QuadriCons

Résultats et
Discussion

Conclusion et
Perspectives

\mathcal{F}_p	\mathcal{R}	r_1				r_2				r_3			
\mathcal{P}	\mathcal{U}/\mathcal{T}	t_1	t_2	t_3	t_4	t_1	t_2	t_3	t_4	t_1	t_2	t_3	t_4
	u_1		×	×	×		×	×	×		×	×	×
p_1	u_2		×	×	×	×	×	×	×	×	×	×	×
	u_3		×	×	×	×	×	×	×	×	×	×	×
	u_4					×			×	×			×
	u_1		×	×	×		×	×	×		×	×	×
p_2	u_2		×	×	×	×			×	×	×	×	×
	u_3		×	×	×	×	×	×	×	×	×		
	u_4					×			×	×			×

Concept Quadratique (ou quadri-concept)

Une nouvelle
approche pour
l'extraction
efficace des
quadri-
concepts
fréquents

Introduction et
Motivations

Le Problème
d'extraction des
quadri-concepts
fréquents

L'Algorithme
QuadriCons

Résultats et
Discussion

Conclusion et
Perspectives

Un concept quadratique (ou quadri-concept) d'une p-folksonomie $\mathcal{F}_p = (\mathcal{U}, \mathcal{T}, \mathcal{R}, \mathcal{P}, Y)$ est un quadruplet (U, T, R, P) avec $U \subseteq \mathcal{U}$, $T \subseteq \mathcal{T}$, $R \subseteq \mathcal{R}$ et $P \subseteq \mathcal{P}$ avec $U \times T \times R \times P \subseteq Y$ tel que le quadruplet (U, T, R, P) est maximal, *i.e.*, aucun de ces ensembles ne peut être augmenté sans diminuer un des trois autres ensembles.
 $U = \text{Extent}$, $T = \text{Modus}$, $R = \text{Intent}$, $P = \text{Variable}$

- $(\{u_1\}, \{t_2, t_3, t_4\}, \{r_1, r_2, r_3\}, \{p_1\})$ n'est pas un quadri-concept.
- $(\{u_1, u_2, u_3\}, \{t_2, t_3, t_4\}, \{r_1, r_2, r_3\}, \{p_1\})$ est un quadri-concept.

Inspiré de (Lehmann et Wille, 1995)

Soient (A_1, B_1, C_1, E_1) et (A_2, B_2, C_2, E_2) deux quadri-sets avec A_i maximal pour $A_i \times B_i \times C_i \times E_i \subseteq Y$, pour $i = 1, 2$. Si $B_1 \subseteq B_2$, $C_1 \subseteq C_2$ et $E_1 \subseteq E_2$ alors $A_2 \subseteq A_1$. Il en est de même pour les trois autres dimensions.

- $(\{u_1, u_2\}, \{t_2, t_3\}, \{r_1, r_2, r_3\}, \{p_1\})$
 - $(\{u_1, u_2, u_3\}, \{t_2, t_3, t_4\}, \{r_1, r_2\}, \{p_1\})$
- Il s'agit de deux quadri-concepts.

Le problème d'extraction des quadri-concepts (fréquents)

Une nouvelle
approche pour
l'extraction
efficace des
quadri-
concepts
fréquents

Introduction et
Motivations

Le Problème
d'extraction des
quadri-concepts
fréquents

L'Algorithme
QuadriCons

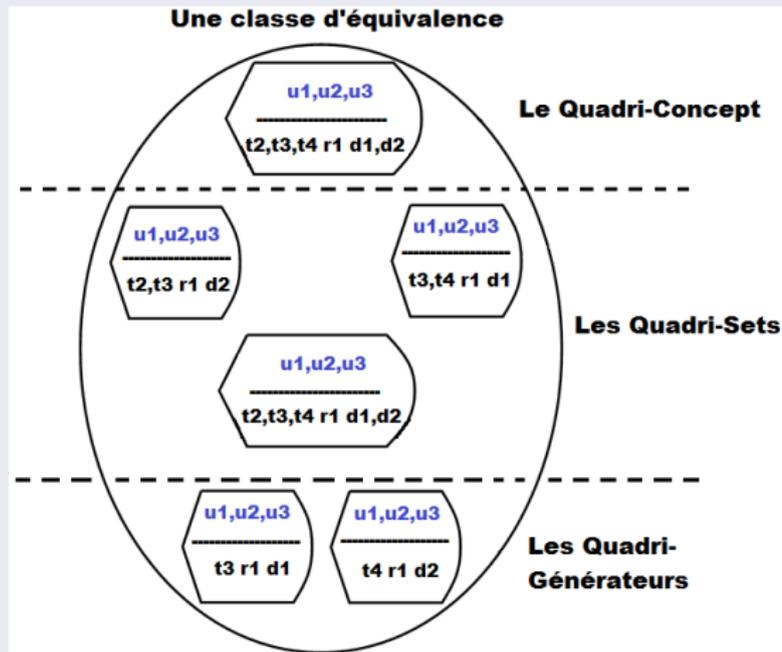
Résultats et
Discussion

Conclusion et
Perspectives

Soit $\mathbb{F}_p = (\mathcal{U}, \mathcal{T}, \mathcal{R}, \mathcal{P}, Y)$ une p -folksonomie et minsupp_u , minsupp_t , minsupp_r et minsupp_p des seuils minimaux de support. Le problème d'extraction de tous les quadri-concepts fréquents consiste à déterminer tous les quadri-concepts (U, T, R, P) de \mathbb{F}_p avec $|U| \geq \text{minsupp}_u$, $|T| \geq \text{minsupp}_t$, $|R| \geq \text{minsupp}_r$ et $|P| \geq \text{minsupp}_p$.

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

Classe d'équivalence



Une nouvelle
approche pour
l'extraction
efficace des
quadri-
concepts
fréquents

Introduction et
Motivations

Le Problème
d'extraction des
quadri-concepts
fréquents

L'Algorithme
QuadriCons

Résultats et
Discussion

Conclusion et
Perspectives

Quadri-générateur

Une nouvelle
approche pour
l'extraction
efficace des
quadri-
concepts
fréquents

Introduction et
Motivations

Le Problème
d'extraction des
quadri-concepts
fréquents

L'Algorithme
QuadriCons

Résultats et
Discussion

Conclusion et
Perspectives

Quadri-générateur

Soit $g = (A, B, C, E)$ un quadri-set de \mathbb{F}_p tel que $A \subseteq \mathcal{U}$, $B \subseteq \mathcal{T}$, $C \subseteq \mathcal{R}$ et $E \subseteq \mathcal{P}$ et $qc \in \mathcal{QC}$. Le quadruplet g est un quadri-générateur minimal (ou quadri-générateur) de qc ssi $h(g) = qc$ et $\nexists g_1 = (A_1, B_1, C_1, E_1)$ tel que :

- 1 $A = A_1$,
- 2 $(B_1 \subseteq B \wedge C_1 \subseteq C \wedge E_1 \subset E) \vee (B_1 \subseteq B \wedge C_1 \subset C \wedge E_1 \subseteq E)$, et
- 3 $h(g) = h(g_1) = qc$.

Opérateur de fermeture

Une nouvelle
approche pour
l'extraction
efficace des
quadri-
concepts
fréquents

Introduction et
Motivations

Le Problème
d'extraction des
quadri-concepts
fréquents

L'Algorithme
QuadriCons

Résultats et
Discussion

Conclusion et
Perspectives

Nouvel opérateur de fermeture

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

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

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

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

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

Une nouvelle
approche pour
l'extraction
efficace des
quadri-
concepts
fréquents

Introduction et
Motivations

Le Problème
d'extraction des
quadri-concepts
fréquents

L'Algorithme
QuadriCons

Résultats et
Discussion

Conclusion et
Perspectives

- 1 Introduction et Motivations
- 2 Le Problème d'extraction des quadri-concepts fréquents
- 3 L'Algorithme QuadriCons
- 4 Résultats et Discussion
- 5 Conclusion et Perspectives

L'algorithme QuadriCons

Une nouvelle
approche pour
l'extraction
efficace des
quadri-
concepts
fréquents

Introduction et
Motivations

Le Problème
d'extraction des
quadri-concepts
fréquents

L'Algorithme
QuadriCons

Résultats et
Discussion

Conclusion et
Perspectives

Un algorithme en 4 étapes

- 1 Extraction des quadri-générateurs
- 2 Calcul du modus des quadri-concepts
- 3 Calcul de l'intent des quadri-concepts
- 4 Calcul de la variable des quadri-concepts

Point Fort / Originalité

- 1 Réduction de l'espace de recherche avec les quadri-générateurs
⇒ Facilité d'extraction des quadri-concepts ! (par augmentations successives)
- 2 Recours à un opérateur de fermeture spécialement dédié au contexte quadratique
- 3 Non-redondance de calcul des quadri-concepts

L'algorithme QuadriCons : Pseudo-code

Une nouvelle
approche pour
l'extraction
efficace des
quadri-
concepts
fréquents

Introduction et
Motivations

Le Problème
d'extraction des
quadri-concepts
fréquents

L'Algorithme
QuadriCons

Résultats et
Discussion

Conclusion et
Perspectives

ALGORITHM 1 : QUADRICONS

Data :

- 1) $\mathbb{F}_d(\mathcal{U}, \mathcal{T}, \mathcal{R}, \mathcal{D}, \mathcal{Y})$: A d -folksonomy.
- 2) $minsupp_u, minsupp_l, minsupp_r, minsupp_d$:
User-defined thresholds.

Results : QC : {Frequent quadri-concepts}.

```
1 Begin
2   /*Step 1 : The extraction of quadri-generators*/
3   FINDMINIMALGENERATORS( $\mathbb{F}_d, MG,$ 
4      $minsupp_u$ );
5   /*Step 2 : The computation of the modus part*/
6   Foreach quadri-gen  $g \in MG$  do
7     CLOSURECOMPUTE( $MG, minsupp_u,$ 
8        $minsupp_l, minsupp_r, g, QS, 1$ );
9   End
10  PRUNEFREQUENTSETS( $QS, minsupp_l$ );
11  /*Step 3 : The computation of the intent part*/
12  Foreach quadri-set  $s \in QS$  do
13    CLOSURECOMPUTE( $QS, minsupp_u,$ 
14       $minsupp_l, minsupp_r, s, QS, 2$ );
15  End
16  PRUNEFREQUENTSETS( $QS, minsupp_r$ );
17  /*Step 4 : The computation of the variable part*/
18  Foreach quadri-set  $s \in QS$  do
19    CLOSURECOMPUTE( $QS, minsupp_u,$ 
20       $minsupp_l, minsupp_r, s, QC, 3$ );
21  End
22  PRUNEFREQUENTSETS( $QC, minsupp_d$ );
23  End
24  return  $QC$  ;
```

Etape 1

Etape 2

Etape 3

Etape 4

L'algorithme QuadriCons : Propriétés

Une nouvelle
approche pour
l'extraction
efficace des
quadri-
concepts
fréquents

Introduction et
Motivations

Le Problème
d'extraction des
quadri-concepts
fréquents

L'Algorithme
QuadriCons

Résultats et
Discussion

Conclusion et
Perspectives

Correction

QuadriCons est correct : extraction de TOUS les quadri-concepts fréquents.

Terminaison

QuadriCons termine.

Complexité (Pire des cas)

QuadriCons est de complexité 4^n avec $n = \min(|\mathcal{U}|, |\mathcal{T}|, |\mathcal{R}|, |\mathcal{V}|)$.

L'algorithme QuadriCons : Exemple illustratif (1)

Une nouvelle
approche pour
l'extraction
efficace des
quadri-
concepts
fréquents

Introduction et
Motivations

Le Problème
d'extraction des
quadri-concepts
fréquents

L'Algorithme
QuadriCons

Résultats et
Discussion

Conclusion et
Perspectives

Etape 1

$\{u_1 u_2 u_4\}$

$t_1 r_1 d_1\}$

$\{u_1 u_2 u_3 u_4\}$

$t_2 r_1 d_1\}$

$\{u_1 u_3 u_4\}$

$t_3 r_1 d_1\}$

 $\{u_1 u_2 u_4\}$

$t_1 r_2 d_1\}$

$\{u_1 u_2 u_3 u_4\}$

$t_2 r_2 d_1\}$

$\{u_1 u_3 u_4\}$

$t_3 r_2 d_1\}$

 $\{u_1 u_2\}$

$t_1 r_1 d_1\}$

$\{u_1 u_2\}$

$t_2 r_1 d_2\}$

$\{u_2 u_4\}$

$t_3 r_1 d_2\}$

 $\{\emptyset\}$

$t_1 r_2 d_2\}$

$\{u_1 u_2\}$

$t_2 r_2 d_2\}$

$\{u_3 u_4\}$

$t_3 r_2 d_2\}$

**Générateurs
Minimaux**

\mathbb{F}_d	\mathcal{R}	r_1			r_2		
\mathcal{D}	$U11$	t_1	t_2	t_3	t_1	t_2	t_3
	u_1	×	×	×	×	×	×
d_1	u_2	×	×		×	×	
	u_3		×	×		×	×
	u_4	×	×	×	×	×	×
	u_1	×	×			×	
d_2	u_2	×	×			×	
	u_3			×			×
	u_4			×			×

L'algorithme QuadriCons : Exemple illustratif (2)

Une nouvelle
approche pour
l'extraction
efficace des
quadri-
concepts
fréquents

Introduction et
Motivations

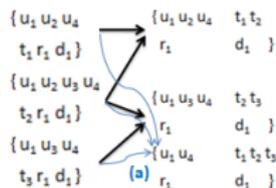
Le Problème
d'extraction des
quadri-concepts
fréquents

L'Algorithme
QuadriCons

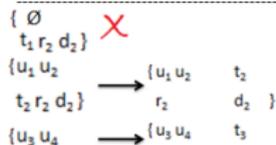
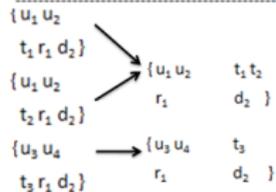
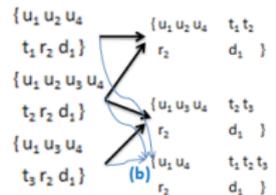
Résultats et
Discussion

Conclusion et
Perspectives

Etape1



Etape 2



L'algorithme QuadriCons : Exemple illustratif (3)

Une nouvelle
approche pour
l'extraction
efficace des
quadri-
concepts
fréquents

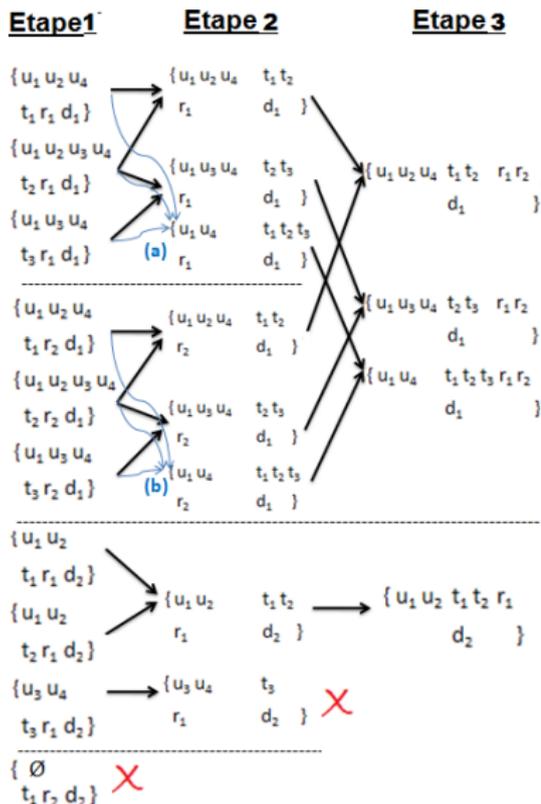
Introduction et
Motivations

Le Problème
d'extraction des
quadri-concepts
fréquents

L'Algorithme
QuadriCons

Résultats et
Discussion

Conclusion et
Perspectives



L'algorithme QuadriCons : Exemple illustratif (4)

Une nouvelle
approche pour
l'extraction
efficace des
quadri-
concepts
fréquents

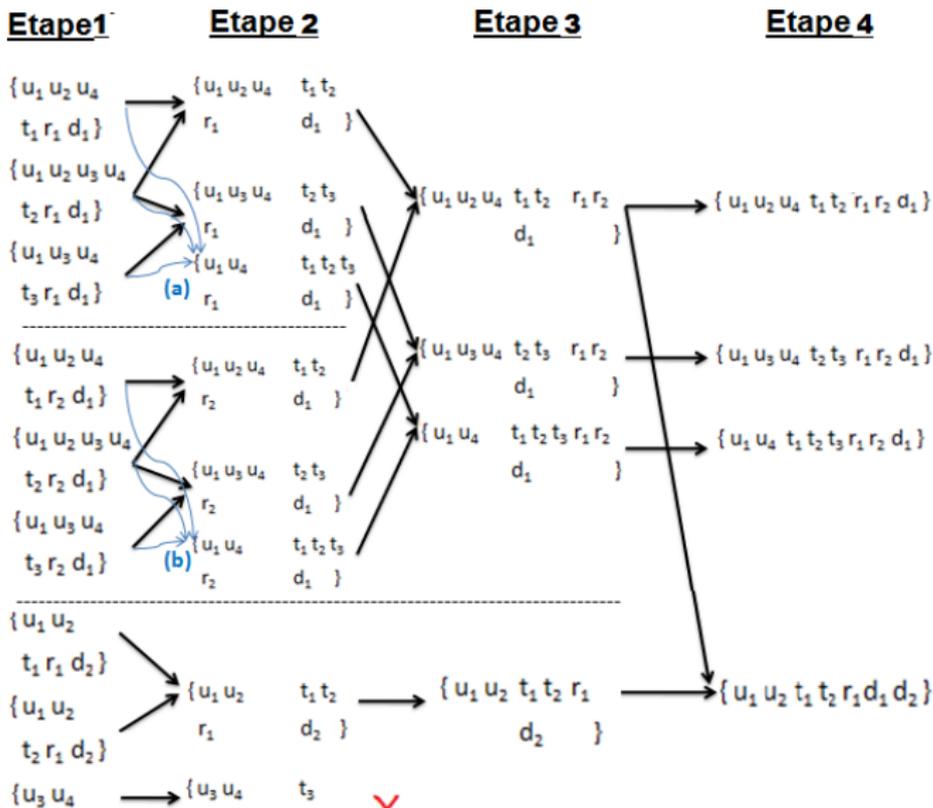
Introduction et
Motivations

Le Problème
d'extraction des
quadri-concepts
fréquents

L'Algorithme
QuadriCons

Résultats et
Discussion

Conclusion et
Perspectives



- 1 Introduction et Motivations
- 2 Le Problème d'extraction des quadri-concepts fréquents
- 3 L'Algorithme QuadriCons
- 4 Résultats et Discussion
- 5 Conclusion et Perspectives

Jeux de données

- 1 Le jeu de données filmographique MovieLens : 95580 tags appliqués à 10681 films par 71567 utilisateurs
- 2 Profils : Genre, Age, Profession, Localisation

Exemples de quadri-concepts

Une nouvelle
approche pour
l'extraction
efficace des
quadri-
concepts
fréquents

Introduction et
Motivations

Le Problème
d'extraction des
quadri-concepts
fréquents

L'Algorithme
QuadriCons

Résultats et
Discussion

Conclusion et
Perspectives

Utilisateurs	Tags	Ressources	Profil
{saloua, yasmine, wafa}	{classic, dialog, oscar}	{Star Wars, Magician of OZ, Rear Window}	{Femme, 46-73 ans, retraité}
{mulder, scully, krycek}	{bestmovie, cult}	{Usual Suspects, Silence of the Lambs, X-Files}	{Homme , 25-35 ans, médecin}
{ross, anlucia, franela}	{classic, oldmovie, quotes}	{Rear Window, Magician of OZ, Gone with the Wind}	{Homme, 36-45 ans, Ecrivain}

Temps d'exécution et mémoire consommée

Y	Quadri Cons (sec)	Mémoire Consommée (kb)	Data Peeler (sec)	Mémoire Consommée (kb)
$minsupp_u = 3, minsupp_t = 2,$ $minsupp_r = 1, minsupp_d = 1$				
25000	0. 86	542	43.10	209843
50000	2. 05	1361	110.72	378907
70000	3. 08	1760	198.33	509541
95580	4. 61	2087	288.00	654761
$minsupp_u = 2, minsupp_t = 2,$ $minsupp_r = 2, minsupp_d = 1$				
25000	0. 36	198	39.98	399672
50000	0. 97	431	107.71	508943
70000	1 .96	567	227.65	667006
95580	3. 79	1182	472.87	842551

Temps d'exécution et mémoire consommée

$ Y $	Quadri Cons (sec)	Mémoire Consommée (kb)	Data Peeler (sec)	Mémoire Consommée (kb)
$minsupp_u = 2, minsupp_t = 1,$ $minsupp_r = 1, minsupp_d = 1$				
25000	97.56	10982	1022.12	1272988
50000	188.61	14671	1987.06	1561992
70000	263.63	19548	2876.02	1751258
95580	528.58	38762	5965.94	2098452

- Espace de recherche beaucoup plus réduit
- Non stockage de la base en mémoire
- Ensemble réduit de candidats grâce aux quadri-générateurs

Domaine d'application - Recommandation : Exemple d'illustration (1)

Une nouvelle approche pour l'extraction efficace des quadri-concepts fréquents

Introduction et Motivations

Le Problème d'extraction des quadri-concepts fréquents

L'Algorithme QuadriCons

Résultats et Discussion

Conclusion et Perspectives

- danascully : un nouvel utilisateur qui veut tagger le film « The Fugitive » Quels tags lui seront proposés ?

Users	Tags	Ressources	Profil
u1, u2	harrison_ford , adventure	The Fugitive , Dead poets	bibliothécaire
u4, u5	crime_story , detective	Godfather , The Fugitive	avocat
u7, u8	classic , oldmovie	Braveheart , MagicianOZ	retraité
u9, u10	adventure , thriller	Star Wars , Blade Runner	ingénieur
u13, u15	adventure , action	Seven , Appolo , Lost Ark	étudiant

Recommandation : Exemple d'illustration (2)

Une nouvelle
approche pour
l'extraction
efficace des
quadri-
concepts
fréquents

Introduction et
Motivations

Le Problème
d'extraction des
quadri-concepts
fréquents

L'Algorithme
QuadriCons

Résultats et
Discussion

Conclusion et
Perspectives

- Deux nouveaux utilisateurs dans la base : **Jacob**(51 ans) et **Reyes** (22 ans) : quels films recommandés à ces deux utilisateurs ?

Utilisateurs	Tags	Ressources	Profil
u1, u2, u3	sciencefiction, cult	Star Wars, The Rock	7-18 ans
u4, u5, u6	adventure, polar	Toy Story , Seven , Braveheart	19-24 ans
u9, u10	thriller, action	Taxi Driver, Die Hard	25-35 ans
u13, u15, u16	oldmovie, classic	Casablanca, Usual Suspects	36-45 ans
u17, u19	classic, dialog, oscar	Magician of Oz , Casablanca	46-73 ans

Recommandation : Exemple d'illustration (3)

Une nouvelle
approche pour
l'extraction
efficace des
quadri-
concepts
fréquents

Introduction et
Motivations

Le Problème
d'extraction des
quadri-concepts
fréquents

L'Algorithme
QuadriCons

Résultats et
Discussion

Conclusion et
Perspectives

- Zoomons sur **Krycek**, un utilisateur de la base ; Alors quels utilisateurs (amis probables) proposés à **Krycek** ?

Users	Tags	Ressources	Profil
krycek, mulder, dana	author, story	Fugitive, Seven	bibliothécaire
u5, u6, u7	crimestory, detective	P.Fiction, Godfather	avocat
u8, u9	classic, oldmovie	Braveheart, OZ	retraité
u1, u2, u3, u4	adventure, thriller	Star Wars, Monty	ingénieur
u10, u11, u12	adventure, action	Seven, Lost	étudiant

Contribution des Quadri-Concepts par rapport à l'existant

Une nouvelle approche pour l'extraction efficace des quadri-concepts fréquents

Introduction et Motivations

Le Problème d'extraction des quadri-concepts fréquents

L'Algorithme QuadriCons

Résultats et Discussion

Conclusion et Perspectives

- Les Quadri-Concepts permettent de regrouper sous une même structure des informations sur les utilisateurs (profil) et leurs comportements (tags/ressources).
- Cibler les tags et les ressources les plus utilisés **EN COMMUN** par les utilisateurs
- Recommandation personnalisée d'amis aux intérêts **ET** profils similaires

- 1 Introduction et Motivations
- 2 Le Problème d'extraction des quadri-concepts fréquents
- 3 L'Algorithme QuadriCons
- 4 Résultats et Discussion
- 5 Conclusion et Perspectives

Conclusion

Une nouvelle
approche pour
l'extraction
efficace des
quadri-
concepts
fréquents

Introduction et
Motivations

Le Problème
d'extraction des
quadri-concepts
fréquents

L'Algorithme
QuadriCons

Résultats et
Discussion

Conclusion et
Perspectives

- Fouille des folksonomies : Approche quadratique (QuadriCons)
- Performances meilleures (temps et mémoire) que son concurrent sur la base MovieLens
- Domaine d'application : recommandation personnalisée

- Autres domaines d'application (e.g., recherche sociale)
- Exploitation d'autres informations que le profil (historique de navigation, commentaires, etc.)

**Une nouvelle
approche pour
l'extraction
efficace des
quadri-
concepts
fréquents**

Introduction et
Motivations

Le Problème
d'extraction des
quadri-concepts
fréquents

L'Algorithme
QuadriCons

Résultats et
Discussion

**Conclusion et
Perspectives**

Merci !
Vos Questions ?