Soc. Netw. Anal. Min. (2015) 5:68
DOI 10.1007/s13278-015-0307-8

ORIGINAL ARTICLE

# Towards more targeted recommendations in folksonomies

**Mohamed Nader Jelassi[1,2,3] · Sadok Ben Yahia[1] · Engelbert Mephu Nguifo[2,3]**

Received: 6 April 2015 / Revised: 30 October 2015 / Accepted: 12 November 2015 / Published online: 30 November 2015


**Abstract** Recommender systems are now popular both commercially as well as within the research community, where many approaches have been suggested for providing recommendations. Folksonomies' users are sharing items (e.g., movies, books, and bookmarks) by annotating them with freely chosen tags. Within the Web 2.0 age, users become the core of the system since they are both the contributors and the creators of the information. In this respect, it is of paramount importance to match their needs for providing a more targeted recommendation. In this paper, we consider a new dimension in a *folksonomy* classically composed of three dimensions <users,tags,resources> and propose an approach to group users with close interests through quadratic concepts. Then, we use such structures in order to propose our personalized recommendation system of users, tags, and resources. We carried out extensive experiments on two real-life datasets, i.e., MovieLens and BookCrossing which highlight good results in terms of precision and recall as well as a promising social evaluation. Moreover, we study some of the key assessment metrics namely coverage, diversity, adaptivity, serendipity, and scalability. Finally, we conduct a user study as a valuable complement to our evaluation in order to get further insights.

**Keywords** Folksonomy · Quadri-concepts · Recommendation · Personalization · User study · Metrics

## 1 Introduction

A *folksonomy* is a practice of collaborative categorization done by internet users (Mika 2007). The idea is to allow users share and annotate resources via freely chosen keywords, i.e., tags which are a way of grouping content by category to make them easy to view by topic (Strohmaier et al. 2012). Roughly speaking, a *folksonomy* consists of three sets $\mathcal{U}$, $\mathcal{T}$, $\mathcal{R}$ as well as a ternary relation between them, where $\mathcal{T}$ is a set of tags which are arbitrary strings and $\mathcal{R}$ denotes resources, e.g., bookmarked websites, described movies, or personal shared videos, depending on the type of the considered *folksonomy* (Jäschke et al. 2008). While, the third set $\mathcal{U}$ consists of the set of users of the *folksonomy* which are generally described by the users' respective nicknames. Users are the main actors of such a system since they contribute to its content by adding resources and assigning tags to them; they are thus considered as the creators of the information and the key actors of the *folksonomy*. Obviously, *folksonomies* are expected to provide satisfactory answers to the needs of each user during the recommendation of tags or resources, which is unfortunately far from being the case. This shortage has led researchers to propose personalized recommender systems in order to suggest the most personalized and appropriate tags and resources to the users (Vallet et al. 2010). In this respect, the research area of personalization attempts to

✉ Mohamed Nader Jelassi
nader.jelassi@gmail.com

Sadok Ben Yahia
sadok.benyahia@fst.rnu.tn

Engelbert Mephu Nguifo
engelbert.mephu_nguifo@univ-bpclermont.fr

[1] Université de Tunis El-Manar, Faculté des Sciences de Tunis, LIPAH, 2092 Tunis, Tunisia

[2] Clermont Université, Université Blaise Pascal, LIMOS, BP 10448, 63000 Clermont-Ferrand, France

[3] CNRS, UMR 6158, LIMOS, Aubière, 63171 Clermont-Ferrand, France

provide solutions to help users share tags and resources among the huge amount of objects in *folksonomies* (Ricci et al. 2011). Hence, a recommender system typically provides the user with a list of recommended items they might prefer which eases the task of finding preferred items in the *folksonomy* (Ricci et al. 2011). For such a task, we consider introducing a fourth dimension in a *folksonomy* and we advocate an approach to group users with "close" interests into structures called *quadratic concepts* (Jelassi et al. 2013). That fourth dimension may cover different aspects: the profile (e.g., gender, age, profession), or the timestamp whenever we want to analyze the temporal dynamics of *folksonomies*. In this paper, we handle the fourth dimension indifferently for the methodological aspect, but, in order to compare our method versus those the pioneering ones of the literature, we will focus on the profile aspect. A quadratic concept sketches a shared conceptualization of the *folksonomy*. For example, a possible quadratic concept could be: *"Jack and Kate who are aged between 18 and 25 years old use tags 'action' and 'adventure,' among others, to annotate movies' 'diana Jones' and "Star Wars."'* A question then arises: why quadratic concepts ? On the one hand, if we can easily explore tags used by a single user on a single resource, such a task quickly becomes unmanageable for a set of taggings involving several users and several resources. On the other hand, recommended tags (or resources) do not appear to be very specific, i.e., tags which are "hackneyed" words or waves resources that do not match the specific needs of the user (Jäschke et al. 2007). Thanks to quadratic concepts, we can tackle both issues. Indeed, quadratic concepts are structures bringing together tags and resources in common to a maximal set of users. Thus, quadri-concepts are a reasonably sized representation of the folksonomy, which may contain thousands of quadruples in real-life datasets. Once extracted, such quadratic concepts stand for the core of our personalized recommendation system.

Moreover, *folksonomy* users might be interested in multi-mode recommendations, so algorithms that serve all modes are ideally indicated (Ricci et al. 2011). Doing so, we explore the following areas: the suggestion of tags, the recommendation of resources, and the proposition of users (friends) helping connect people with common interests. Furthermore, we are interested in a minority of new users who may not have recommendations. This issue is known as the *cold start problem* (Said et al. 2011). Finally, we carried out a thorough evaluation: first, we compute the precision and recall of our recommender system, then we analyze its different properties. In addition, we propose a user study in order to get feedback of users on our recommender system.

The remainder of the paper is organized as follows. In the next section, we motivate the personalization of recommendations. Then, we thoroughly study the related work in Sect. 3. We present the key notions that will be of use through the paper as well as the process of quadratic concept's extraction in Sect. 4. Then, in Sect. 5, we introduce our new personalized recommender system. In Sect. 6, we report the very promising results of our experimental validation. Finally, we conclude the paper with a summary and we sketch some avenues for future works in Sect. 7.

## 2 Motivations

The lack of organization of the shared resources and the (too) much freedom of the choice of tags led researchers to enhance the current recommendation. Many approaches tend to assist the user to choose the "good" resources, i.e., the most interesting among the overwhelming number of available ones to share as well as to use the "good" tags to annotate them. Hence, the recommendation has to address the personalization issue, i.e., additional information about users should be of use to provide a more personalized recommendation. This issue is of prominent importance since the users of a *folksonomy* have different needs and expectations that depend on their motivations (Landia and Anand 2009). Looking for meeting the needs of each user, different works were interested in the personalization of recommendations (Rikitianskii et al. 2014). Moreover, personalization attempts to provide solutions to help users tackling the information overload issue (Liang 2010). But why, in our case, is it so useful to know more about users? (Noll et al. 2007) In order to try to fulfill the expectations of each user of the *folksonomy*, it is useful to have more information about him/her (Das et al. 2012). Actually, recommender systems need a deeper understanding of users and their information seeking tasks to be able to provide better recommendations (Breuss and Tsagkias 2014). The goal is to produce recommendations that better target the needs of each individual user.

### 2.1 Contributions

In this paper, we propose a new personalized recommender algorithm FolkRec [1] that uses a ranking score to classify recommendations in order to improve the precision and recall of our system. Moreover, we have conducted a detailed analysis of FolkRec, through the following properties:

1. *Cold start problem* A brief analysis of the existing recommender systems shows that a minority of new

[1] Downloadable at this link http://www.isima.fr/~mephu/FILES/FolkRec/.

users may not have recommendations. This issue is known as the *cold start problem* (Said et al. 2011; Lika et al. 2014). We propose to include such feature to our new system in order to take into account new users who have not tagged anything yet.

2. *Coverage, diversity, adaptivity serendipity, and scalability* As a result of the first point, we aim to improve the user space coverage as well as the item space coverage, i.e., the portion of users and items that our system is able to cover. Then, we assess different properties of our recommender system as the diversity and serendipity, i.e., how diverse and surprising the recommendations are. Finally, we assess the *scalability* of the system in order to show how fast our recommendations are.
3. *User study* Some approaches rely on the interaction of users with the system and thus, off-line testing may be insufficient. Hence, in this paper, we conduct a user study by recruiting a set of test subjects and asking them to perform several tasks and then we record their behavior and report their feedback.

In the remainder, we scrutinize the main state-of-the-art approaches dealing with the recommendations in *folksonomies*.

## 3 Related work

Aiming to improve the recommendations in *folksonomies*, several works have been proposed (Agarwal and Chen 2010; Yang et al. 2011). We can split the main works into four categories:

1. *Works using similarity measures* In Diederich and Iofciu (2006), users' *personomies*, i.e., relative tags, are used to recommend users whom shared similar tags or similar resources. First, the authors looked at the most used tags w.r.t a given user, then, based on this tag-based profile, authors are able to recommend users (called *collaborators*) relying on a similarity measure between users. Besides focusing on a unique dimension (users), this approach suffers from an approximative measure that only relies on the tags used by the users. Doing so, no full information about users is obtained. In Landia and Anand (2009), the authors proposed an approach combining both resource similarity and user similarity to recommend personalized tags. Through identifying a set of users similar to the active user, the approach is able to suggest personalized tags. Thus, two users are considered as similar whenever they assigned the same tags to the same resources. However, in real-life, it is not so common to find such a situation in real-life *folksonomies* which tags used by different users on the same resources are the same.
2. *Works relying on popularity* Jaeschke et al. (2007) proposed tag recommendations in *folksonomies* based on the most used tags. The user dimension was also taken in consideration to recommend tags. However, the approach did not take further information about users, i.e., only their tagging history through their shared tags and resources. Lipczak, proposed in Lipczak (2008), a three-step tag recommendation system: starting from tags assigned to resources, the authors add tags proposed by a lexicon based on co-occurrences of tags within resource's posts. Then, the system filters those already used by the user to focus on its interest. Nevertheless, the recommendation did not seem personalized despite filtering users' tags since it seeks tags that co-occurred on other posts. The approach only comes back to remove tags previously used by a user from the list of recommended ones. De Meo et al. (2010) built and maintained a profile for each user. Hence, whenever a user submits a query to the *folksonomy* to retrieve a set of resources of his interest, the authors propose to find further *"authoritative"* tags to enrich his query and propose them to the user. Authoritative tags are those which have been most used by other users on the same resources. However, in such an approach, the user has to add tags to resources and/or select the most relevant tags to them; roughly speaking, the user has to provide an explicit feedback. In addition, the evaluation of tag authoritativeness involves an off-line data analysis.
3. *Hybrid Works using history tagging and similarity measures* In Hu et al. (2011), tag recommendations were based on user's tagging history as well as the personalized preference learned from social contacts. Indeed, according to the authors, the social contacts data could be of use to provide more personalized recommendations of tags for a user when annotating resources. Nevertheless, the moan that can be addressed stands in the requirement that a user must have social contacts to benefit from tag's recommendation. Hence, new users would not benefit from recommended tags. Basile et al. (2007) proposed a smart tag recommender which is able to learn from user's tagging history as well as the content of the resources to annotate. The system also has the ability to recommend a list of new meaningful tags used by other on the same resources.
4. *Works relying on user information* In Bellogìn et al. (2013), extended the common accuracy-oriented evaluation practice (i.e., recall and precision) with several metrics to measure additional recommendation quality dimensions (e.g., coverage, diversity, and novelty).

According to authors, hybrid recommenders combining some strategies (e.g., content-based, collaborative filtering, and social) may provide more valuable recommendations in terms of performance metrics. Kim et al. (2011) proposed a recommendation procedure for online book communities. The proposed recommendation procedure consists of two steps. First, it finds *neighbors* using users preferences for books and their feature information (i.e., profile), and then it generates personalized recommendations. The second step removes *irrelevant* books from the recommendation list using the keyword preferences of each user. Qumsiyeh et al. (2012) proposed a personalized recommendation system that relies on several users information as ratings and reviews of different multimedia items. The authors aimed at matching the interests of each user to make recommendations.

Although these recent works rely on personal information about users to propose personalized recommendations, they address huge data which may alter recommendation quality in some contexts. Moreover, most of the other related works are limited to the information <user, tag, resource>, so, in our works, we extend this triplet by the information included in the fourth dimension. Hence, we propose to combine additional information about users with their shared tags and resources into structures called quadri-concepts in order to improve recommendations. Within such concepts, we can focus not only on which tags have been used, but rather on which tags have been used in combination. Moreover, quadri-concepts allow to focus on the most used tags and resources which help reduce significantly the huge input data before the recommendation step. However, the previous recommender system (i.e., PersoRec Jelassi et al. 2013), which relies on such concepts to provide recommendations, did not use a ranking score to classify recommendations. Moreover, PersoRec did not take into account new users who did not share anything yet. Finally, such algorithm did not filter out tags and resources already shared by users when providing recommendations. Hence, we propose our new personalized recommender system in order to overcome these problems by (i) considering a ranking score to classify recommendations in order to increase the precision and recall of our system; (ii) considering new users by providing them recommendations based on their profiles; (iii) filter out tags and resources already shared by users in order to have new and diverse recommendations.

In the following, we define a *v-folksonomy* as well as a quadri-concept.

## 4 Key notions

In this section, we introduce some key notions that will be of use throughout this paper. Indeed, we present the mathematical settings of core of our recommendation system, i.e., the quadratic concepts. But, we first introduce an extension of the notion of a *folksonomy* (Jäschke et al. 2008) by adding a fourth dimension.

**Definition 1** (V-Folksonomy; Jelassi et al. 2013) A *v-folksonomy* is a set of tuples $\mathcal{F}_v = (\mathcal{U}, \mathcal{T}, \mathcal{R}, \mathcal{V}, Y)$ where $\mathcal{U}$, $\mathcal{T}$, $\mathcal{R}$ and $\mathcal{V}$ are finite sets which elements are called, respectively, *users*, *tags*, *resources*, and *variables*. $Y \subseteq \mathcal{U} \times \mathcal{T} \times \mathcal{R} \times \mathcal{V}$ represents a quaternary relation where each $y \subseteq Y$ can be represented by a quadruple: $y = \{(u, t, r, v) \mid u \in \mathcal{U}, t \in \mathcal{T}, r \in \mathcal{R}, v \in \mathcal{V}\}$, i.e., the user $u$ has annotated the resource $r$ using the tag $t$ through the variable $v$. We consider that two users are *close* if they share at least one variable in common.

The introduced fourth dimension may serve to cover different aspects: the profile, (e.g., gender, age, profession, and location), or the timestamp whenever we want to analyze the temporal dynamics of *folksonomies*. Thus, the information included in the fourth dimension is completely correlated to the triple (user, tag, resource). For example, in a quadruple ($u$, $t$, $r$, *timestamp*), the *timestamp* information is correlated to the tagging operation made by $u$ with the tag $t$ and the resource $r$. While, in a quadruple ($u$, $t$, $r$, *profile*), the *profile* information is correlated to the user $u$ as well as to the tag $t$ and the resource $r$ shared by $u$. Indeed, a user $u$ may share a book about programming languages with the tag *programming* via one profile information $p_1$ (*student* for example) while (s)he may share a journal paper with the tag *paper* via another profile information $p_2$ (*researcher* for example). In this paper, we handle the fourth dimension indifferently from a methodological aspect point of view, but, in order to compare our method with the works of the literature, we will later focus on the profile aspect.

*Example 1* Table 1 depicts a toy example of a *v-folksonomy* $\mathcal{F}_v$ with $\mathcal{U} = \{u_1, u_2, u_3, u_4\}$, $\mathcal{T} = \{t_1, t_2, t_3, t_4\}$, $\mathcal{R} = \{r_1, r_2, r_3\}$ and $\mathcal{V} = \{v_1, v_2\}$. Each cross, within the quaternary relation, indicates a tagging operation by a user from $\mathcal{U}$ with a variable from $\mathcal{V}$, a tag from $\mathcal{T}$ and a resource from $\mathcal{R}$. For example, the user $u_1$, which has between 25 and 35 years old and is a student, has tagged the articles $r_1$, $r_2$ and $r_3$ through the tags *thesis*, *web* and *to_recommend*.

We now introduce a quadratic concept, which is the four-dimensional version of a frequent itemset in a dyadic context.

**Table 1** A toy example of a *v-folksonomy* with the following values: *science* ($t_1$), *thesis* ($t_2$), *web* ($t_3$), *to_recommend* ($t_4$), 25–35 years old ($v_1$), *student* ($v_2$) and $r_1$, $r_2$ and $r_3$ a set of three science articles

| $\mathcal{F}_v$ | $\mathcal{R}$ | $r_1$ | | | | $r_2$ | | | | $r_3$ | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| $\mathcal{V}$ | $\mathcal{U}/\mathcal{T}$ | $t_1$ | $t_2$ | $t_3$ | $t_4$ | $t_1$ | $t_2$ | $t_3$ | $t_4$ | $t_1$ | $t_2$ | $t_3$ | $t_4$ |
| | $u_1$ | | × | × | × | | × | × | × | | × | × | × |
| $v_1$ | $u_2$ | | × | × | × | × | × | × | × | × | × | × | × |
| | $u_3$ | | × | × | × | × | × | × | × | × | × | × | × |
| | $u_4$ | | × | × | | × | | | × | × | | | × |
| | $u_1$ | | × | × | × | | × | × | × | | × | × | × |
| $v_2$ | $u_2$ | | × | × | × | × | | | × | × | × | × | × |
| | $u_3$ | | | | | | | | | | | | |
| | $u_4$ | | | | | | | | | | | | |

**Definition 2** (QUADRATIC CONCEPT; Jelassi et al. 2013) A *quadratic concept* (or quadri-concept) of a v-folksonomy $\mathcal{F}_v = (\mathcal{U}, \mathcal{T}, \mathcal{R}, \mathcal{V}, \Upsilon)$ is a quadruple $(U, T, R, V)$ where $U \subseteq \mathcal{U}$, $T \subseteq \mathcal{T}$, $R \subseteq \mathcal{R}$ and $V \subseteq \mathcal{V}$ and $U \times T \times R \times V \subseteq \Upsilon$ such that the quadruple $(U, T, R, V)$ is maximal w.r.t to set inclusion, i.e., it is not possible to increment one of these sets without shrinking one of the three other sets. For a quadri-concept $QC = (U, T, R, V)$, the $U$, $R$, $T$, and $V$ parts are, respectively, called *Extent*, *Intent*, *Modus*, and *Variable* (Jäschke et al. 2008). Without support restrictions, we expect getting out a huge number of quadri-concepts from a given v-folksonomy. In order to keep the most interesting ones, we can set minimum thresholds on each dimension of the v-folksonomy, i.e., $minsupp_u$, $minsupp_t$, $minsupp_r$ and $minsupp_v$. Doing so, it results in quadri-concepts which are called *frequent*. The set of all frequent quadri-concepts extracted from a given v-folksonomy is denoted in the sequel $\mathcal{QC}$.

*Example 2* With respect to the v-folksonomy, depicted by Table 1, the quadruple $(\{u_1, u_2\}, \{t_2, t_3, t_4\}, \{r_1, r_2, r_3\}, v_1)$ is not a quadri-concept since its extent, i.e., the users part is not maximal. Indeed, the user $u_3$ as well as $u_1$ and $u_2$ (whom all have the variable $v_1$ in common) also shared the resources $r_1$, $r_2$ and $r_3$ through the tags $t_2$, $t_3$, $t_4$. Hence, if we maximize the extent, we get the quadruple $(\{u_1, u_2, u_3\}, \{t_2, t_3, t_4\}, \{r_1, r_2, r_3\}, v_1)$ which is actually a quadri-concept, highlighting that the users $u_1$, $u_2$ and $u_3$, aged between 25 and 35 years old, have tagged the science articles $r_1$, $r_2$ and $r_3$ via the tags $t_2$, $t_3$, $t_4$ (thesis, web and to_recommend).

In order to mine all (frequent) quadri-concepts from a given *v-folksonomy*, we may use one of the reported algorithms in the literature: QUADRICONS (Jelassi et al. 2012), an extension of the TRICONS algorithm (Trabelsi et al. 2012), or DATAPEELER (Cerf et al. 2009) which are both able to extract all frequent quadri-concepts fulfilling user-defined minimum thresholds on each dimension of the *v-folksonomy*. Both algorithms take as input a *v-folksonomy* as well as four minimum thresholds and output the set $\mathcal{QC}$ of frequent quadri-concepts fulfilling these thresholds.

### 4.1 Illustrative example

Consider the *v-folksonomy* depicted by Table 1, with the following minimum thresholds: $minsupp_u = 2$, $minsupp_t = 2$, $minsupp_r = 1$ and $minsupp_v = 1$. Roughly speaking, these thresholds indicate that in each quadri-concept, two users (at least) with a same variable (at least) have shared a resource (at least) with the same two tags (at least). Applying one of the dedicated algorithms for the extraction of quadri-concepts on this *v-folksonomy* results in the mining of five frequent quadri-concepts:

1. $(\{u_1, u_2, u_3\}, \{t_2, t_3, t_4\}, \{r_1, r_2, r_3\}, v_1)$;
2. $(\{u_2, u_3, u_4\}, \{t_1, t_4\}, \{r_2, r_3\}, v_1)$;
3. $(\{u_2, u_3\}, \{t_1, t_2, t_3, t_4\}, \{r_2, r_3\}, v_1)$;
4. $(\{u_1, u_2\}, \{t_2, t_3, t_4\}, \{r_1, r_3\}, \{v_1, v_2\})$;
5. $(\{u_1, u_2, u_3, u_4\}, \{t_2, t_3\}, \{r_1\}, v_1)$

For example, the first quadri-concept highlights that the users $u_1$, $u_2$ and $u_3$, having between 25 and 35 years old, shared the resources $r_1$, $r_2$ and $r_3$ through the tags $t_2$, $t_3$ and $t_4$ (i.e., *thesis*, *web* and *to_recommend*). Whereas, the third quadri-concept shows that the users $u_2$ and $u_3$ (a subset of the previous set of users) shared the resources $r_2$ and $r_3$ (a subset of the previous set of resources) via the tags $t_1$, $t_2$, $t_3$ and $t_4$ (a superset of the previous set of tags). Such examples highlight the property of quadri-concepts, i.e., "none of its sets can be extended without shrinking one of the other three sets."

In what follows, we use the frequent quadri-concepts to introduce our new personalized recommender system of users, tags as well as resources.

## 5 FolkRec: a new personalized recommender system

In this section, we propose our new personalized recommender algorithm FOLKREC. The pseudocode of FOLKREC is sketched by Algorithm 1. FOLKREC takes $\mathcal{QC}$ (a set of frequent quadri-concepts) as input as well as a user $u$ with a variable $v$ and (optionally) a resource $r$ (to annotate). The set $\mathcal{QC}$ is first extracted (as a pre-processing step) by one of the algorithms of the literature dedicated to such a task from a given *v-folksonomy* using minimum thresholds (Jelassi et al. 2012; Cerf et al. 2013). Then, FOLKREC outputs three sets: a set of proposed users, a set of suggested tags and a set of recommended resources. Among the improvements achieved by FOLKREC, we can notice that our recommender system is able to give recommendations to a new user in order to solve the cold start problem. Moreover, FolkRec ensures that each user will not receive as recommendations tags or/and resources that (s)he has already shared.

**Algorithme 1 :** FolkRec

```
   Input : a set QC, a user u with a variable v and a resource r.
   Output : The sets PU, ST and RR.
 1 begin
 2   foreach quadri-concept qc ∈ QC do
 3     if u is an old user then
 4       u.Tags={t ∈ T / ∃ r ∈ R, (u,t,r,v) is a quadri-concept};
 5       u.Resources={r ∈ R / ∃ t ∈ T, (u,t,r,v) is a quadri-concept};
 6       if v ∈ qc.Variable then
 7         if u ∉ qc.Extent then
 8           /*User proposition*/
 9           PU = PU ∪ qc.extent;
10           /*Tag Suggestion*/
11           if r ∈ qc.Intent then
12             ST = ST ∪ qc.modus \ u.Tags;
13           end
14           /*Resource Recommendation*/
15           RR = RR ∪ qc.Intent \ u.Resources;
16         end
17       end
18     end
19     if u is an new user then
20       if v ∈ qc.Variable then
21         /*User proposition*/
22         PU = PU ∪ qc.extent;
23         /*Tag Suggestion*/
24         if r ∈ qc.Intent then
25           ST = ST ∪ qc.modus;
26         end
27         /*Resource Recommendation*/
28         RR = RR ∪ qc.Intent;
29       end
30     end
31   end
32   return (PU,ST,RR);
33 end
```

FolkRec operates as follows: depending of the statut of $u$ (old or new), it seeks for quadri-concepts whose users have the same variable in common than $u$ (Lines 6–16). If $u$ is an old user, then, we compute its tags and resources already shared (Lines 4–5). Then, if $u$ already belongs to the quadri-concept $qc$, then $qc$ is skipped (Line 7) to filter out tags and resources already shared by $u$. Such a strategy is inspired by that of (Lipczak 2008). Then, depending of the type of task, FolkRec operates as follows: for the *User Proposition* task (Line 9), we add to the set $\mathcal{PU}$, of proposed users, the extent of the quadri-concept $qc$. The user proposition helps connect users with common interests and thus promotes the sharing of content. For the *Tag Suggestion* task (Lines 11 and 12), the aim is to suggest personalized tags to a target user which shares a resource in the *v-folksonomy*. Such a suggestion would be of benefit for threefolds purposes: (i) increasing a resource annotation, (ii) telling (or reminding) a user what a resource is about, and (iii) consolidating the vocabulary across the users (Ricci et al. 2011). Thus, we could suggest for $u$ tags that were assigned by close users to the same resource. For this task, we need an additional information, i.e., the resource to be annotated ($r$). Then, we display suggested tags that were affected to the same resource $r$ into the set $\mathcal{ST}$. At Line 14 of FolkRec, the goal of the *Resource Recommendation* task is to propose a personalized list of resources to a target user that is aimed to be in compliance with its interests. Hence, the set $\mathcal{RR}$ contains resources which are recommended to $u$. However, if $u$ is a new user (lines 19-30), there is no need to any filter since $u$ has not shared tags and resources yet. Hence, the only information that we need is contained in the variable $v$ related to the user. Thus, FolkRec recommends to $u$ tags and resources shared by users having the same variable $v$ in common.

*Theoretical complexity issues* As in the triadic case (Jäschke et al. 2008), the number of (frequent) quadri-concepts may exponentially grow in the worst case. Hence, the theoretical complexity of the pre-processing phase of frequent quadri-concepts extraction is around $\mathcal{O}(2^n)$ with $n = |\mathcal{T}| + |\mathcal{R}| + |\mathcal{V}|$. Nevertheless, and as it will be shown in the section dedicated to experimental results, from a practical point of view, the actual performances are far from being exponential and the step of quadri-concepts extraction flags out the desired scalability feature. Therefore, we focus on empirical evaluations on large-scale real-world datasets. Furthermore, it is important to note that, even if the pre-processing step of quadri-concepts extraction could be time consuming, it happens *off-line* and is executed only *once* (i.e., when the system starts). Indeed, FolkRec runs already extracted frequent quadri-concepts. Thus, our recommender system does not bear the cost of quadri-concepts extraction on each recommendation. Let $m \ll n$ be the number of extracted frequent quadri-concepts. Hence, the theoretical complexity of our FolkRec algorithm is around $\mathcal{O}(m)$ since FolkRec runs the set of frequents quadri-concepts only once.

## 6 Results and discussion

In this section, we evaluate our approach on two real-life datasets. Then, we present some examples of frequent quadri-concepts as well as several personalized recommendations generated by FolkRec. Moreover, we evaluate our recommendations by computing their precision and recall as well as some metrics that are of common use within the recommendation community. Finally, we conduct a user study in order to get further insights and conclusions.

### 6.1 Datasets

The two real-life datasets used for our evaluation are described as follows:

- The MovieLens filmography dataset:[2] it is a recommender system and virtual community website that allows users to share movies using tags. The MovieLens dataset, used for our experiments, is freely downloadable[3] and contains 95580 tags applied to 10681 movies by 71567 users (e.g., ⟨*Alex*, *X-Files*, *sciencefiction*⟩). The MovieLens dataset, published by the GroupLens research group at University of Minnesota, is one of the most referenced and evaluated repositories within the Recommender Systems community.
- The BookCrossing library dataset:[4] it is a free online book club which was founded to encourage the practice, aiming to "make the whole world a library". Contrariwise to MovieLens, the BookCrossing dataset does not rely on tags to annotate the resources, i.e., the books, but rather on *ratings*. In such a library website, users are asked to annotate books by choosing an integer from 1 to 10, i.e., the higher the value is, the higher the appreciation is. The used dataset is freely downloadable[5] and contains 278858 users providing 1149780 ratings about 271379 books (e.g., <*Spender*, *DaVinciCode*, *9*>).

Tables 2 and 3 depict some examples of quadruples from, respectively, the MovieLens and the BookCrossing datasets. For purposes of comparison with the pioneering approaches of the literature, we choose, in what follows, the *user profile* in order to model the variable $v$. Hence, we consider now that two users are *close* if they share at least one profile information in common (e.g., a same age, a same profession, etc.). To this end, additional available information about users constitute their profiles (the fourth dimension of the *v-folksonomy*) which can be for MovieLens: the gender (male or female), the profession (21 in total, which can be administrator, artist, doctor, scientist, etc.). For BookCrossing, users are provided with their location (United States, Taiwan, France, etc.). Finally, both datasets provide information about the age of the users which are partitioned into five categories: (i) 7–18 years; (ii) 19–24 years; (iii) 25–35 years; (iv) 36–45 years and (v) 46–73 years. Note that, contrariwise to shared tags and resources which constantly evolve through time, the user's profile rarely changes. Indeed, users can change their profession or location but such change is not very frequent into the *v-folksonomy*.

[2] http://movielens.umn.edu/.

[3] http://www.grouplens.org/node/73.

[4] http://www.bookcrossing.com/.

[5] http://www.grouplens.org/node/74.

**Table 2** A toy screenshot of the MovieLens dataset

| User | Tag | Resource | Profile |
|---|---|---|---|
| *Mulder* | *action* | *X-Files* | *student* |
| *Mulder* | *sciencefiction* | *X-Files* | *25 years old* |
| *Scully* | *adventure* | *Jurassic Park* | *professor* |
| *Scully* | *bestmovie* | *Jurassic Park* | *female* |
| *Skinner* | *thriller* | *Carrie* | *Canada* |
| ⋮ | ⋮ | ⋮ | ⋮ |

**Table 3** A toy screenshot of the BookCrossing dataset

| User | Rating | Resource | Profile |
|---|---|---|---|
| *Jack* | *9* | *Da Vinci Code* | *doctor* |
| *Kate* | *8* | *I got you under my skin* | *artist* |
| *Kate* | *8* | *I got you under my skin* | *38 years old* |
| *Locke* | *10* | *Mouth of Madness* | *52 years old* |
| *Reyes* | *10* | *Mouth of Madness* | *Spain* |
| ⋮ | ⋮ | ⋮ | ⋮ |

## 6.2 Examples of extracted frequent quadri-concepts

In what follows, we present some interesting results of extracted frequent quadri-concepts from both datasets, i.e., results from the pre-processing step of FolkRec. For such an extraction, we ran one of the dedicated algorithms of the literature, i.e., QuadriCons (Jelassi et al. 2012) or Data Peeler (Cerf et al. 2013) on a machine with a processor Intel Core *i*5 and a 8 Go memory. Several tests, carried out on the operating system Linux Ubuntu 12.04 (64 bits), allow to generate frequent quadri-concepts ⟨users, tags, resources, profiles⟩.

In the following, we define the following empirical thresholds values of supports: $minsupp_u = 2$, $minsupp_t = 2$, $minsupp_r = 2$ and $minsupp_v = 2$. Obviously, it seems more interesting to set each minimum threshold to 2 (at least) in order to have frequent quadri-concepts with an added value illustrating shared tags and resources by a group (of two users at minimum) with same profile, i.e., two common information at least. Hence, Tables 4 and 5 depict some examples of frequent quadri-concepts among the respectively 10627 and 18756 frequent quadri-concepts fulfilling these aforementioned thresholds on each dimension, on respectively, the MovieLens and BookCrossing datasets. It takes, respectively, 3.79 and 15.92 s to QuadriCons to output the sets of frequents quadri-concepts from the MovieLens and BookCrossing datasets versus thousands of seconds for Data Peeler. In Table 4, for example, the first quadri-concept shows that the users *Saloua*, *Wafa*, and *Yasmine*, three retired women aged between 46 and 73

**Table 4** Examples of frequent quadri-concepts extracted from the MovieLens dataset

| | |
|---|---|
| *U* | {Saloua, Wafa, Yasmine} |
| *T* | {classic, dialog, oscar} |
| *R* | {Star Wars, M.A.S.H, Rear Window} |
| *P* | {Female, 46–73 years, retired} |
| *U* | {Mulder, Scully, Frohike} |
| *T* | {bestmovie, cult} |
| *R* | {Usual Suspects, Silence of Lambs} |
| *P* | {25–35 years, healthcare} |
| *U* | {Ross, Anlucia, Franela} |
| *T* | {classic, oldmovie, quotes} |
| *R* | {Rear Window, OZ, Gone with the Wind} |
| *P* | {Male, 36–45 years, Writer} |

**Table 5** Examples of frequent quadri-concepts extracted from the BookCrossing dataset

| | |
|---|---|
| *U* | {Regina, Rumple} |
| *T* | {8} |
| *R* | {The Chicago Manual of Style The Man Who Counts} |
| *P* | {36–45 years, Toronto} |
| *U* | Caphook, Madhat |
| *T* | 10 |
| *R* | Harry Potter and the Prisoner of Azkaban Tough Enough (Pokemon Chapter Book, 27) |
| *P* | 7–18 years, Chicago |
| *U* | {Allison15, Buffay} |
| *T* | {9} |
| *R* | {The Eye of the World (The Wheel of Time, Book 1) Dragons of Autumn Twilight (Vol 1)} |
| *P* | {18–25 years, Lisbon} |
| *U* | {Maryc, Kingg90} |
| *T* | {8} |
| *R* | {Le Cycle d'Ender, tome 4 : Les Enfants de l'esprit |
| | Le Cycle d'Ender, tome 3 : Xénocide Les amants maudits (10/18)} |
| *P* | {25–35 years, Paris} |
| *U* | {Brson,Manblack} |
| *T* | {10} |
| *R* | {Full Time: the Secret Life of Tony Cascarino Special Needs Education: Children... The Rasputin File} |
| *P* | {18–25 years, Hong Kong} |

years old, shared the movies *Star Wars*, *M.A.S.H* and *Rear Window* via the tags *classic*, *dialog* and *oscar*. While, in the third quadri-concept, three writers aged between 36 and 45 years, i.e., *Ross*, *Anlucia* and *Franela* opted for the tags *quotes*, *classic*, and *oldmovie* when they shared the movies *Braveheart*, *Magician of OZ* and *Gone with the Wind*. In Table 5, for example, the first quadri-concept shows that

**Table 6** Examples of quadri-concepts following the profile *profession*

| | |
|---|---|
| *U* | {Krycek, Deepthr, Mistx12} |
| *T* | {Author, based_on_a_book} |
| *R* | {The Fugitive, Dead poets society} |
| *P* | {Librarian} |
| *U* | {Fox16, Dana, Cgbspender} |
| *T* | {Adventure, polar} |
| *R* | {Pulp Fiction, The Godfather, The Fugitive} |
| *P* | {Lawyer} |
| *U* | {Mylafi, Jenifer, Nabilawi} |
| *T* | {thriller, action} |
| *R* | {Braveheart, Magician of OZ} |
| *P* | {Retired} |
| *U* | {Fran, Chandy, Joeytr, cacforever} |
| *T* | {Award, oldmovie, classic} |
| *R* | {Star Wars, Blade Runner, Monty Python} |
| *P* | {Engineer} |
| *U* | {Chedly50, Slioua7, Nina16} |
| *T* | {Classic, dialog, oscar} |
| *R* | {Seven, Appolo 13, Raiders of Lost Ark} |
| *P* | {Student} |

two users from Toronto, aged between 36 and 45 years, share the books *The Chicago Manual of Style* and *The Man Who Counts* with a rating equal to 8. Another quadri-concept depicts that two users, aged between 25 and 35 years and living in Paris, are interested in the franchise *Le Cycle d'Ender* as well as the franchise *Les Amants Maudits*. Furthermore, Table 6 depicts some examples of frequent quadri-concepts when the profile information concerns only the user's profession. It highlights a difference of vocabulary and interests between users of different professions. For example, we have the users *Ched50*, *Slioua7*, and *Nina16*, who are students, sharing the movies *Appolo 13* and *Raiders of Lost Ark* via the tags *adventure* and *action*. It is also worth of mention that students shared massively action movies within tags like *adventure* and *action*, while lawyers would rather share movies with a crime story like *The Fugitive* through tags like *detective* and *crime_story*.

### 6.3 Personalized recommendation

Quadri-concepts provide a new way for grouping under a same concept users with same profiles sharing tags and resources in common. There are several areas in which quadri-concepts can be useful for. Thanks to our FolkRec algorithm, we are able to illustrate three of them. In what follows, we give some examples of real cases for each application.

- *Tag suggestion* Consider a new user called *Davis* of the MovieLens dataset which looks for sharing the movie *The Fugitive*, then he will get a choice of tags that depends on his profile. For example, if *Davis* was a librarian, we would suggest him the tags *author* and *based_on_a_book*, while if he was a lawyer, he would be able to annotate the movie with the tags *adventure* and *polar* (cf., Table 6).
- *Resource recommendation* Users of different countries in the world are interested in different books, according to the culture and traditions of their lives. Consider two new users *Reyes* (28 years, from Paris) and *Zlatan* (20 years, from Chicago): our system will recommend the books *Le Cycle d'Ender* and *Les amants maudits* for *Reyes*, which are books that are popular in France. While, books like *Harry Potter* and *Pokemon*, very in vogue in Chicago among young people, will be recommend to *Zlatan*, which could correspond better to his interests. Since these two users are new in the *v-folksonomy*, we have no available information about their taggings. Hence, our recommendations for both users are only based on their profile (i.e., their age and location in this case).
- *Friend proposition* For example, consider the user *Krycek* who is a librarian (cf., Table 6); users that we could recommend him as friends are *deepthr* and *mistx12* since they have a same profile information, i.e., the profession, and same interests, i.e., shared tags and resources.

*Example 3* Consider the frequent quadri-concepts illustrated by Table 4 and 5 and let $u_1$ = Jacob (37 years, Writer) and $u_2$ = Ilana (Female, 63 years) be two new users of the dataset MovieLens and $u_3$ = Sheldon (Male, 26 years, Paris) be a new user of the dataset BookCrossing. First, suppose that Jacob and Ilana both want to share the movie Rear Window. Thanks to our personalized recommender system, we are able to provide two kinds of suggestion following the profile of each user. Hence, Jacob will have the following suggested tags: classic, quotes and oldmovie while tags like classic, dialog, and oscar will be suggested to Ilana in order to annotate that movie (cf., Table 4). Moreover, we may propose as friends the users Ross, Anlucia, and Franela to Jacob since they have close profiles and interests. While, it seems more suitable to propose the users Saloua, Wafa and Yasmine to Ilana since they all have about the same age and interests. Second, concerning Sheldon, who is a new user in the BookCrossing dataset, FolkRec searches for frequent quadri-concepts matching his profile, i.e., a 26-year-old male from Paris. Hence, we are able to recommend him the books Le Cycle d'Ender, tome 4: Les Enfants de l'esprit and Le Cycle d'Ender, tome 3: Xénocide, Les amants maudits (10/18) since both books were shared by users with same profile as Sheldon, i.e., Parisian men within the same category of age.

### 6.4 Evaluation of the recommendation: precision, recall, and F1-score

#### 6.4.1 Training set/test set

In our experiments, we used the fivefold cross-validation (Weiss and Kulikowski 1991) to evaluate the effectiveness of our approach. Both MovieLens and BookCrossing datasets were split into two-sub datasets: the first sub dataset, containing random 80 % of users, was used as *training set* while the second one, containing the remaining users (i.e., random 20 % of users), was retained as the validation data for tests (i.e., the *test set*). For each test user, a random 20 % of its tagging is considered as the test/answer set and 80 % as its training set. We repeat such experiments five times changing at each time the 20 % representing the test set in order to cover 100 % of the whole set. For each user of the test set, our recommender algorithm generates a list of items (tags, users or resources) based on the user's training set. If an item in the recommendation list was also in the user's test set, then the item is considered as pertinent. In our experiments, we also varied the number of recommendations provided to the user. This is known as top-*k* recommendations. With such requests, the user can specify the *k* recommendations considered as the most relevant that the system shall return to him. The first *k* answers are those which are ranked with the best scores (see below, Equation 1). This especially helps avoid overwhelming the user with a large number of answers by returning only the number of the most relevant answers that (s)he wishes (Fig. 1, 2, 3).

#### 6.4.2 Recommendations ranking

For a given dataset, the top-*k* recommendations consist of a list of items ranked by a decreasing score value. In the following, we focus only on the resource recommendation since that such task is considered as the most important in the process of recomme ndation. Hence, for generating a resource recommendation for a given user, we compute the ranking as described above, and then restrict the result set to the top-*k* first results (with the higher scores). The score measure (denoted *rec_score*) corresponding to a profile *v* is defined as follows:

$$rec\_score(r_i, v) = \frac{|u_i|}{|UU|} / \exists t_i \exists r_i \exists v_i \in v, (u_i, t_i, r_i, v_i) \in \mathcal{F}_v \quad (1)$$

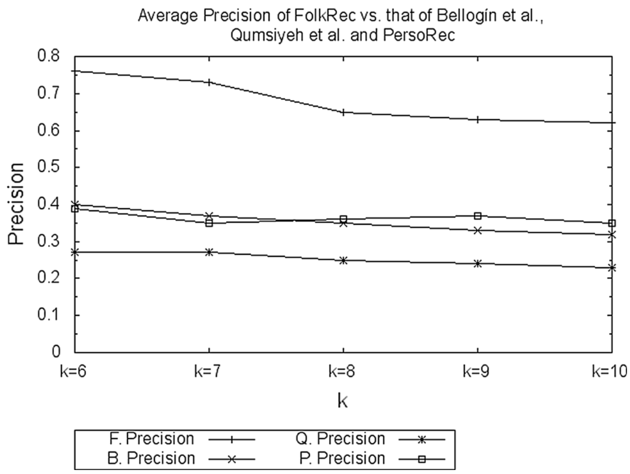


**Fig. 1** Average precision of FolkRec versus its competitors for the recommendation of resources on MovieLens. (F) FolkRec (B) Bellogìn et al. (Q) Qumsiyeh et al. (P) PersoRec (cf., Table 8)

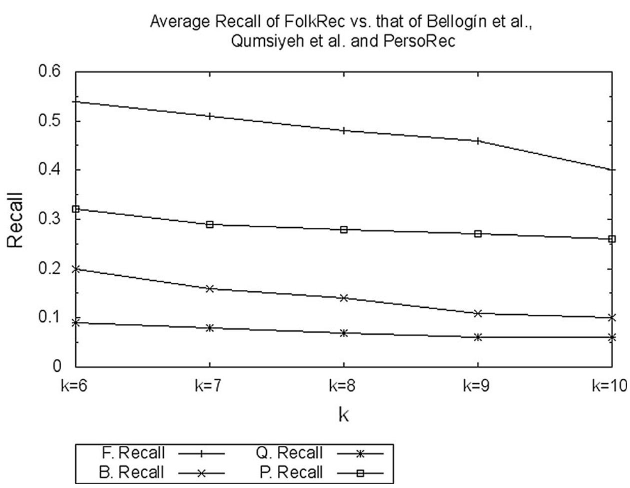


**Fig. 2** Average recall of FolkRec versus its competitors for the recommendation of resources on MovieLens. (F) FolkRec (B) Bellogìn et al. (Q) Qumsiyeh et al. (P) PersoRec (cf., Table 10)

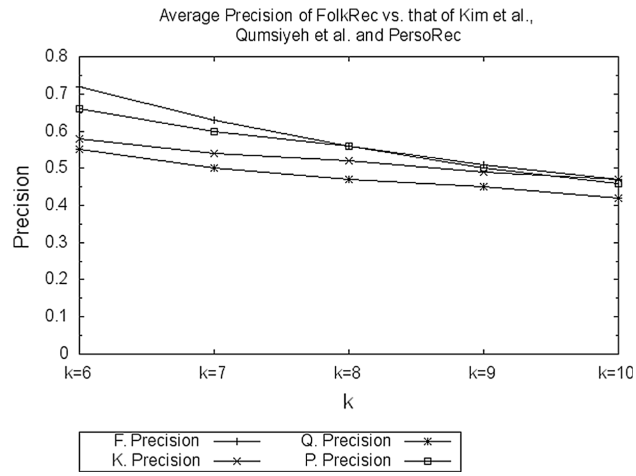


**Fig. 3** Average precision of FolkRec versus its competitors for the recommendation of resources on BookCrossing. (F) FolkRec (K) Kim et al. (Q) Qumsiyeh et al.(P) PersoRec (cf., Table 9

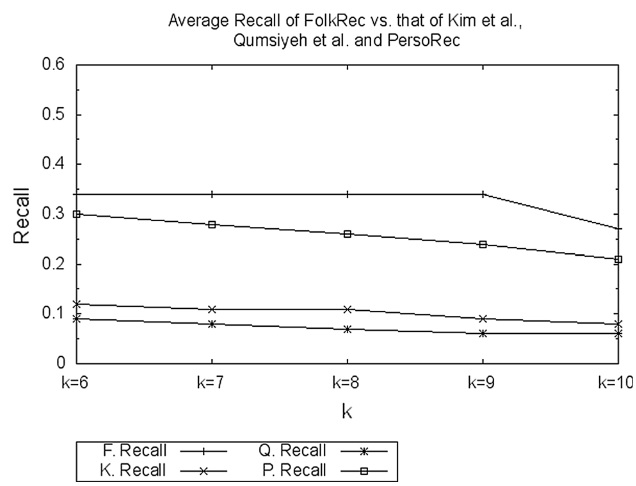


**Fig. 4** Average recall of FolkRec versus its competitors for the recommendation of resources on BookCrossing. (F) FolkRec (K) Kim et al. (Q) Qumsiyeh et al.(P) PersoRec (cf., Table 11)

Hence, the *rec_score* of a resource $r_i$ corresponding to a profile $v$ is the number of unique users having the same profile $v$ (or at least one profile information $v_i \in v$) who have shared such resource divided by the total number of unique users in the set of frequent quadri-concepts (denoted $UU$). For example, if a resource $r_1$ was shared by 7 different users among a set of 67 unique users, its score will be equal to 0.104 while another resource $r_2$ shared by 16 different users among the same set will have a score equal to 0.238. Taking this fact within the evaluation of recommender systems, we could apply classic information retrieval metrics to evaluate those engines: Precision and Recall (Baeza-Yates and Ribeiro-Neto 1999). These metrics are of common use within information retrieval community scenario and applied to several domains. Figures 2 and 4 summarize the different results obtained by FolkRec versus its competitors in both datasets.

### 6.4.3 Precision

Assessing the efficiency of a recommendation algorithm is far from being a trivial task. First, since different algorithms can be better or worse according to the dataset on which they are applied (Herlocker et al. 2004). Second, the goals of a recommendation system can be different and varied. For example, a recommendation system can be implemented in order to accurately estimate the score that a

**Table 7** The surveyed approaches at a glance

| Approach | Datasets | Validation Protocol | Pertinence | Score | Top-k |
|---|---|---|---|---|---|
| Qumsiyeh et al. | MovieLens, BookCrossing Yahoo!music, Netflix | Fivefold cross-validation | The rating is pertinent if it is actually given by the user to the resource | A measure score based on past ratings | No |
| Kim et al. | BookCrossing (a small representation) | Not defined | The book recommendation is pertinent if the book is selected by the user | A similarity measure between communities of users | Yes |
| Bellogìn et al. | MovieLens, Delicious, LastFM | Fivefold cross-validation | The recommendation is pertinent if it belongs to the test set of the target user | A measure of preference and a measure of items the most used by friends | Yes |
| PersoRec | MovieLens and BookCrossing | Holdout | The recommendation is pertinent if it belongs to the test set of the target user | Not used | Yes |
| FolkRec | MovieLens and BookCrossing | Fivefold cross-validation | The recommendation is pertinent if it belongs to the test set of the target user | A measure score based on user's profiles (See Eq. 1) | Yes |

user would give to an element, whereas other ones would pay attention to avoid proposing incorrect recommendations. One can therefore legitimately wonder to what extent these different methods of recommendation are actually effective. Nevertheless, to determine the efficiency of a system, the most common indicators in the literature are the *precision*, the *recall* and the *F1-score*. Such measures represent the quality of the recommendation, i.e., to what extent the proposed suggestions are in line with the interests of the user.

The precision computes the probability that a recommended element is relevant. Thus, the best measure of the efficiency of a recommendation algorithm and the relevance of the recommendations is to evaluate the precision of the prediction performed by the system by comparing the predictions with the choices that would have given the user in the real case. In what follows, we focus on one of the three applications of our recommendation system, i.e., the recommendation of resources and we assess the precision of our approach versus the *pioneering* approaches of the literature that have close goals to our approach, i.e., those of Bellogin et al. (2013), Qumsiyeh et al. (2012), Kim et al. (2011) and the recommender system PersoRec (Jelassi et al. 2013). First, Table 7 shows some comparisons between the different approaches and FolkRec. It highlights for each approach, the datasets used for the experiments, the validation protocol, how the authors define the pertinence of the recommendations, the ranking score used for tests and finally if the approach rely on top-k results. Tables 8 and 9 show the different values of precision obtained by our recommendation algorithm versus its competitors for different values of $k$ going from 6 to 10 for both datasets. Generally, our recommendations proposed to the users of MovieLens and BookCrossing reflect their expectations. Indeed, the recommendations are relevant in an average of 67 and 57 % on, respectively, MovieLens and

**Table 8** Average precision of FolkRec versus its competitors for the recommendation of resources on MovieLens

| $k$ | FolkRec | Bellogìn et al. | Qumsiyeh et al. | PersoRec |
|---|---|---|---|---|
| 6 | **0.76** | 0.40 | 0.27 | 0.39 |
| 7 | **0.73** | 0.37 | 0.27 | 0.35 |
| 8 | **0.65** | 0.35 | 0.25 | 0.36 |
| 9 | **0.63** | 0.33 | 0.24 | 0.37 |
| 10 | **0.62** | 0.32 | 0.23 | 0.35 |

**Table 9** Average precision of FolkRec versus its competitors for the recommendation of resources on BookCrossing

| $k$ | FolkRec | Kim et al. | Qumsiyeh *et al.* | PersoRec |
|---|---|---|---|---|
| 6 | **0.72** | 0.58 | 0.55 | 0.66 |
| 7 | **0.63** | 0.54 | 0.50 | 0.60 |
| 8 | **0.56** | 0.52 | 0.47 | **0.56** |
| 9 | **0.51** | 0.49 | 0.45 | 0.50 |
| 10 | **0.47** | **0.47** | 0.42 | 0.46 |

BookCrossing, which outperforms, for most of the values, the precision of its competitors. Hence, for MovieLens, our precision score is, respectively, 91, 168, and 86 % greater than that of Bellogin et al., Qumsiyeh et al. and PersoRec. While, for BookCrossing, our precision is respectively 11, 21 and 4 % higher than those of Kim et al., Qumsiyeh *et al.* and PersoRec. Results also show that the best performances of FolkRec were obtained with a value of $k = 6$. This is explained by the fact that the first five recommendations meet the expectations of users. In addition, as far as the number of recommendations increases, this inevitably leads to a decrease in precision as the user selects fewer resources than those recommended. We explain the difference between our precision scores and those of competitors by the fact that the use of quadri-concepts improves the recommendations by suggesting the closest tags and resources to users' needs. Indeed, while related works focus on most used items (books, movies, tags), quadri-concepts offer, to our users, tags and resources that have been shared in common by a set of users with close profiles. In fact, it turns out that users tend to share tags and resources already shared by users within the same profile. Doing so, the precision of FolkRec increases. While, we also improve the precision values of PersoRec for the MovieLens dataset by enhancing the recommendations thanks to our ranking score and some features in the algorithm (consideration of new users, filtering already shared resources, etc.).

### 6.4.4 Recall

The recall is the number of the recommendations that are pertinent reported to the total number of existing pertinent recommendations in the dataset. Hence, the recall measure highlights the portion of pertinent recommendations that have been returned to the user from the whole set of pertinent recommendations. The recall would then assess the proportion of all relevant results included in the top-*k* results.

Tables 10 and 11 show the different recall values obtained by FolkRec versus its competitors for different values of *k* going from 6 to 10 on both MovieLens and BookCrossing datasets. Results highlight that our algorithm sharply outperforms state-of-the-art approaches. Indeed, FolkRec has an average recall value of 47 % on MovieLens versus 14 % for that of Bellogin et al., 7 % for that of Qumsiyeh et al. and 28 % for that of the recommender system PersoRec. Whereas, for BookCrossing, our recall is, respectively, 219, 352, and 26 % greater than the respective recall values of Kim *et al.*, Qumsiyeh et al. approaches and PersoRec. Such difference shows that among pertinent items, FolkRec is better able to recommend a significant portion to users. Thanks to frequent quadri-concepts which are representative structures of the *folksonomy* which target better users' needs, FolkRec recommends items shared by users which are likely to be shared by other users with close profile. Moreover, the strategy advocated by FolkRec which consists in filtering tags and resources already shared by users as well as considering new users tend to increase its recall versus the other approaches.

### 6.4.5 F1-Score

The *F1-Score* considers both precision and recall measures of the test to compute the score. We could interpret it as a weighted average of the precision and recall, where the best *F1-score* has its value at 1 and worst score at the value 0. Within the recommendation domain, it is considered as a single value obtained combining both the precision and recall measures and indicates an overall utility of the recommendation list. In this respect, Tables 12 and 13 show the different values of the *F1-Score* obtained by FolkRec versus its competitors for different values of *k* going from 6 to 10 on the MovieLens and BookCrossing datasets. It is worth of mention that FolkRec outperforms its competitors on both datasets since our recall and precision values are higher than those of other considered approaches.

**Table 10** Average recall of FolkRec versus its competitors for the recommendation of resources on MovieLens

| *k* | FolkRec | Bellogìn et al. | Qumsiyeh *et al.* | PersoRec |
|---|---|---|---|---|
| 6 | **0.54** | 0.20 | 0.09 | 0.32 |
| 7 | **0.51** | 0.16 | 0.08 | 0.29 |
| 8 | **0.48** | 0.14 | 0.07 | 0.28 |
| 9 | **0.46** | 0.11 | 0.06 | 0.27 |
| 10 | **0.40** | 0.10 | 0.06 | 0.26 |

**Table 11** Average recall of FolkRec versus its competitors for the recommendation of resources on BookCrossing

| *k* | FolkRec | Kim et al. | Qumsiyeh *et al.* | PersoRec |
|---|---|---|---|---|
| 6 | **0.34** | 0.12 | 0.09 | 0.30 |
| 7 | **0.34** | 0.11 | 0.08 | 0.28 |
| 8 | **0.34** | 0.11 | 0.07 | 0.26 |
| 9 | **0.34** | 0.09 | 0.06 | 0.24 |
| 10 | **0.27** | 0.08 | 0.06 | 0.21 |

**Table 12** Average F1-Score of FolkRec versus its competitors for the recommendation of resources on MovieLens

| *k* | FolkRec | Bellogìn et al. | Qumsiyeh *et al.* | PersoRec |
|---|---|---|---|---|
| 6 | **0.57** | 0.20 | 0.09 | 0.35 |
| 7 | **0.56** | 0.16 | 0.08 | 0.31 |
| 8 | **0.54** | 0.14 | 0.07 | 0.31 |
| 9 | **0.56** | 0.11 | 0.06 | 0.31 |
| 10 | **0.52** | 0.10 | 0.06 | 0.29 |

**Table 13** Average F1-Score of FolkRec versus its competitors for the recommendation of resources on BookCrossing

| *k* | FolkRec | Kim et al. | Qumsiyeh *et al.* | PersoRec |
|---|---|---|---|---|
| 6 | **0.48** | 0.20 | 0.15 | 0.41 |
| 7 | **0.45** | 0.18 | 0.14 | 0.38 |
| 8 | **0.40** | 0.18 | 0.13 | 0.35 |
| 9 | **0.41** | 0.15 | 0.12 | 0.32 |
| 10 | **0.39** | 0.13 | 0.12 | 0.28 |

### 6.5 Social evaluation of the recommendation

In what follows, we focus on the study of the social evaluation of our recommendations. Indeed, assessing the efficiency of a recommendation algorithm is far from being trivial and as mentioned before, the precision and recall do not globally evaluate the pertinence of an algorithm since datasets and objectives may vary from an algorithm to another one. Hence, in what follows, we innovate by looking at what happens *after* the recommendation step, i.e., if the target user really enjoyed the recommendation and if the users (friends) proposed to him/her are getting the same social behavior. To do so, we study three different cases of recommendation in the BookCrossing dataset and one case in the MovieLens dataset. Thus, for the first dataset, we choose three different users with different ages and countries: *Skinner* (38 years, New York, USA), *Herge* (26 years, Seixal, Portugal) and *Benjamin* (15 years, Texas, USA). At first, our algorithm recommends to *Skinner* three books from the franchise *Harry Potter* as well as four new friends: *Snowh* (43 years, *Illinois*, *USA*), *Char_dav* (54 years, *California*, *USA*), *Emma* (40 years, *Oregon*, *USA*) and *Henry36* (36 years, *Teheran*, *Iran*). It later turned out that these new friends also share all books of the franchise *Harry Potter*. In addition, *Skinner*, as well as, his recommended friends, have rated the recommended books with a rating equal to 9 which highlights that they really appreciate the recommended books. Furthermore, we recommend to *Herge* three different books (*Da Vinci Code*, *Wild Animus* and *The Joy Luck Club*) as well as four users aged between 25 and 35 years: three of them from the *USA* (respectively from *Kansas*, *Wisconsin* and *Virginia*) and one from *Canada* (*Ottawa*). However, although he shares all these books, *Herge* did not rate them in order to report his appreciation. Interestingly enough, among its "new friends," one of them was really interested in the same books. Finally, our algorithm generates for *Benjamin* two books (*Harry Potter and the Prisonnier of Azkaban* and *Harry Potter and Cup of Fire*) as well as a new friend, i.e., *Baefire* (12 years, *Illinois*, *USA*). It turns out that, after the recommendation, both users share these books with the top high rating (10) which indicates that they really appreciate the recommended books. For the MovieLens dataset, our test user is *Bruce* (47 years, Male, Educator). FolkRec recommends to him four movies: *Star Wars*, *The Return of the Jedi*, *God Father 1 and 2* as well as two friends: *Marina* (49 years, Female, Educator) and *Joey* (49 years, Male, Educator). First, we can see that *Bruce* enjoys his movie recommendations as he rates the four movies with the top note of 5. Second, it turns out that his recommended friends have also shared the same movies with an average rating equal to 4 which shows that Bruce and his new friends have actually common interests in the same movies.

In the following, we proposed several metrics capturing different item recommendation properties, namely coverage, diversity, scalability, serendipity, adaptivity, and novelty. The latter is defined as the capability of a recommender to suggest a user with relevant but non-popular items.

### 6.6 Properties of the FolkRec algorithm

As seen in the previous subsections, a good item recommendation is an item that belongs to the test set of the target user. However, recommending the right items (users, tags and resources) is a crucial task, but might be insufficient to deploy a good recommendation system. Sometimes, users may be interested by more than a good recommendation: discovering new items, the diversity of items, and many other properties of the recommender system (Said and Bellogín 2014). Hence, we must identify the set of properties that may influence the success of a recommender system (Ricci et al. 2011).

#### 6.6.1 User space coverage

As the precision of a recommendation system, especially in social networks, is in a snugness connection with the growth of data size, some algorithms may provide recommendations with high quality, but only for a small portion of the users of data size. The user space coverage can then be defined as the proportion of users for which the system can recommend items (Ricci et al. 2011). Algorithms which are able to give recommendations to a majority of users are particulary desired. That said, FolkRec is able to provide recommendation to all users of the *folksonomy* regardless of the fulfillment of a given condition, e.g., that a user has tagged less than a defined number of items or a user must have a certain number of friends. With FolkRec, once a user is added to the *folksonomy*, its personal information is sufficient to provide him/her a recommendation of tags, resources and users. Also, we compute the percentage of profiles covered by FolkRec that results in the following results: 100 % of genders (male and female), 100 % of age categories, 100 % of professions, and 88 % of cities.[6] While, the approach of Bellogìn et al. reaches an average coverage of 84.6 %.

#### 6.6.2 Cold start problem

A recurrent issue in recommendation process is the so-called *cold start problem* (Ricci et al. 2011; Said et al.

[6] From the 13625 cities represented in BookCrossing, we evaluate the coverage of FolkRec above the most represented ones, i.e., cities present in more than 500 quadruples in the *v-folksonomy*.

2011), i.e., the performance of the system when handling new users. Until now, we made use of the history of tagging and profile information of users to recommend tags and resources. However, new users just start to use the system and we do not know their preferences yet. Indeed, a new user is a user which has no tagging history. Many recommender systems tend to solve the cold start problem by asking new users to rate a set of resources at the start (Ricci et al. 2011). However, it is not easy to decide which resources to get a user to rate/tag. Hence, contrariwise to the majority of approaches of the literature, FolkRec does not ask a new user to share a minimum number of resources before being considered by the recommender system. It amounts to the fact that FolkRec looks first for a user's personal information to provide recommendations before looking to its tagging history. After creating such stereotypes as a starting point (e.g., according to the age, gender or country of the users), FolkRec follows the new users and suggests them further recommendations based on their shared items (through an incremental learning). FolkRec is then able to take into account new users without restarting the process of the quadri-concept's extraction. However, the limit of our approach is that there is a problem of incrementality of new items (tags and resources) which are not included into the set of quadri-concepts already extracted.

### 6.6.3 Serendipity

Serendipity is a measure of how surprising the recommendations are (Ricci et al. 2011). A recommender system tries to surprise users by recommending resources that a user may not know. We use a distance metric $d$ (c.f., Eq. 2) that measures the distance between a recommended item and a set of items already tagged by a user. Doing so, we experiment this distance on the BookCrossing dataset in order to assess to what extent the recommendations are surprising the users, i.e., by recommending books that the reader is less familiar with. The distance metric $d$ is defined as follows (Ricci et al. 2011):

$$d(b, B) = \frac{1 + C_B - C_B.w(b)}{1 + C_B} \quad (2)$$

where $b$ is the recommended book, $B$ the set of books previously read by the target user, $C_B$ the maximal number of books from a single author in $B$ and $C_B.w(b)$ the number of books by the author of $b$ in $B$. Note that the value of $d$ stands within the unit interval. Then, to measure serendipity, we combine this metric with the pertinence of items (i.e., the precision). Indeed, serendipity is reached when the item is both surprising and pertinent. We tested the metric $d$ on the same users from the social evaluation (cf., Sect. 6.10). The average score of the distance metric is then equal to 0.44, and the final metric (combination between the metric $d$ and the precision) is equal to 0.53 which highlights that *folksonomy*'s users are quite surprised by the recommended books. If the first and the third users are not specially surprised, since we recommended to them books of their favorite author *J.K.Rowling*, we quite succeed to surprise the second user. Indeed, we recommend to him books of *R. Shapero*, *D. Brown* and *Tan* while he used to read books from *A. Gide* and *M. Duras* (3 books read of each).

### 6.6.4 Diversity

Recommending a set of similar items is not as useful for users which prefer diversity, i.e., different recommendations that are *distant*. For example, a user may prefer five recommended books from five different authors to a recommendation including five books from a single author. To measure the diversity of our recommendations on the BookCrossing dataset, we use the same distance metric $d$ used for the serendipity criteria (cf., Eq. 2) as follows: we compute the distance between each recommended book and the remainder of the list (of recommended books) and then we average the result to obtain a diversity score. We get an average diversity score equal to 0.56 for all users (of social evaluation experimentation) with a top-score of 1 for *herge*. Such a score is explained by the fact that our recommended books for this user are diverse (from three different authors) which may please to users interested in having diverse recommendations. For comparison purposes, the diversity score achieved by the approach of Bellogin et al. is about 0.35 while the other approaches do not calculate such score.

### 6.6.5 Adaptivity

Testing the adaptivity of the system comes back to compute the rate by which the system adapts to a user's personal preferences, or to changes in a user profile (Ricci et al. 2011). For example, when a user shares new resources, (s)he is expecting to have new recommendations related to its new preferences. Such a system behavior encourages users to share more resources since they can see that the recommender system is able to change according to their profiles. In the following, in order to evaluate the adaptivity of our system, we measure the difference between the recommendation lists before and after a new information (about a user) is added. Let us reconsider some users (from Sect. 6.10) and see the system's reaction to some changes in their profiles. At first, we introduce a change in the profile of *Skinner* (38 years, New York, USA) from BookCrossing with the following scenario: the user moves to (Ottawa, Canada). FolkRec takes into account such profile information change and proposes to Skinner a new recommendation that is more suitable, i.e., we recommend to him the books *The Alchimist*, *The Red Tent* and *Love in the time of cholera*, three books that have been shared by canadian users of Ottawa offering to Skinner a chance to know more about books in vogue in his new town. Moreover, we recommend him three new users from Ottawa: *tchang* (27 years), *archie* (49 years) and *grump* (40 years). Furthermore, we consider the user *Bruce* (47 years, Male, Educator) from MovieLens and we assume that he changes his job, i.e., working in healthcare instead of being educator. Thus, FolkRec will generate a new recommendation for Bruce with regard to his new statuts: *When Harry Met Sally*, *Silence of Lambs* and *A Streetcar Named Desire* as recommended movies and *Michael* (50 years, Male) and *Janice* (43 years, Female) as recommended friends. On the one hand, we recommend him users within the same profession and category of age that might henceforth interest him. On the other hand, we recommend him movies shared by users within the same profile.

In this subsection, we focus on changes in the profile (change in profession, location, etc.) but we may also focus on changes in tagging history of users. It seems obvious that a user who shares new resources, i.e., becomes interested by a new category of movies or books (e.g., action movies instead of comedy ones for example) will have new recommendations that may match its new preferences.

#### 6.6.6 Scalability

One standard approach to evaluate the scalability of a system is to evaluate the complexity of the dedicated algorithm in terms of time and/or memory requirements. For this purpose, we compute the average response time (in milliseconds) of recommendations provided by FolkRec on both datasets for the tasks of resource recommendation (denoted Task 1) and user proposition (denoted Task 2).[7] Table 14 shows the average response time of FolkRec [8] for the MovieLens dataset which contains about 100,000 quadruples ⟨user,tag,resource,profile⟩. We assume that each extracted quadri-concept, from this dataset, contains at least one tag, one resource and one profile information and we vary the minimum threshold of users, i.e., the minimum number of users by quadri-concept. For example, when $minsupp\_u = 6$, we get 13461 quadri-concepts such as each concept contains 6 users at least. In addition, we compute the number of (unique) users on the set of frequent quadri-concepts. Table 14 highlights the good performances of FolkRec for all values of *minsupp_u*. While the number of quadri-concepts grows rapidly (from 221 to 13461), the average time of recommendations generated by FolkRec is around 2 ms for the recommendation of resources and around 8 ms for the user proposition task. The number of total recommendations is equal to the number of unique users (one recommendation for each user) which is up to 865 for the smallest value of *minsupp_u*.

**Table 14** Average response time of FolkRec recommendations above the MovieLens (up) and BookCrossing (down) datasets

| *min_u* | $\lvert \mathcal{QC} \rvert$ | *# Unique users* | Task 1 (ms) | Task 2 (ms) |
|---|---|---|---|---|
| MovieLens | | | | |
| 20 | 221 | 526 | 0.1 | 2.6 |
| 14 | 795 | 634 | 0.4 | 4.9 |
| 10 | 2430 | 718 | 1.7 | 9.3 |
| 6 | 13461 | 865 | 12.7 | 23.3 |
| BookCrossing | | | | |
| 30 | 553 | 6789 | 0.9 | 149.8 |
| 20 | 1486 | 9092 | 4.9 | 296.9 |
| 14 | 3720 | 11257 | 23.6 | 494.4 |
| 10 | 10100 | 13457 | 114.7 | 586.8 |

Table 14 also highlights the performance of the FolkRec algorithm in terms of response time over the BookCrossing dataset, which contains about 762,000 quadruples ⟨user,-tag,resource,profile⟩. Each quadri-concept of the set $\mathcal{QC}$ contains at least one tag, one resource and one profile information, and we vary the minimum number of users, i.e., the *minsupp_u* threshold from 30 to 10. Firstly, we can see that the highest number of quadri-concepts is equal to 10100 which represents only 1.76 % of the *v-folksonomy*; it shows the usefulness of quadri-concepts which are a small representation of the dataset. Contrariwise to MovieLens, the number of unique users, i.e., the total number of recommendations drastically increases while the number of frequent quadri-concepts slightly grows. However, if FolkRec flags out good performances for the task of resource recommendation where the average response time is around 28 ms, the task of user proposition become harder since that each quadri-concept contains at least 10 users. Nevertheless, the average response time of the second task remains reasonable which is around 384 ms.

#### 6.6.7 Overview of the surveyed approaches

To highlight the added value of our approach over its predecessors, we underline in Table 15 the different criteria of recommendation systems (as discussed above) shall fulfill (Ricci et al. 2011). The question mark ("?") denotes that such information is missing in the original paper and

[7] We omit the tag suggestion task since that BookCrossing rather considers ratings than tags.

[8] Unfortunately, the codes of our competitors are not available. Moreover, The runtime of the competitors were not specified in the original papers.

**Table 15** The surveyed approaches at a glance

| | M-Mode | Cov. | C-Start | Adap. | Div. | Ser. | Sca. |
|---|---|---|---|---|---|---|---|
| Diederich and Iofciu (2006) | No | No | No | No | Yes | ? | No |
| Basile et al. (2007) | No | No | No | No | ? | ? | ? |
| Jäschke et al. (2007) | No | No | No | No | Yes | ? | ? |
| Landia and Anand (2009) | No | No | No | No | No | ? | ? |
| De Meo et al. (2010) | No | No | No | No | ? | ? | ? |
| Hu et al. (2011) | No | No | No | No | Yes | ? | ? |
| Kim et al. (2011) | No | No | No | No | Yes | ? | ? |
| Qumsiyeh and Ng (2012) | No | No | No | No | Yes | ? | ? |
| Bellogín et al. (2013) | No | Yes | No | No | Yes | ? | ? |
| PersoRec Jelassi et al. (2013) | Yes | No | No | No | No | No | Yes |
| FolkRec | Yes | Yes | Yes | Yes | Yes | Yes | Yes |

thus hard to check from an external point of view. We can notice, for example, that none of the presented approaches offer a multi-mode recommendation (of users, tags and resources at the same time). As we could see in the evaluation section, our approach differs from its competitors by taking into account new users (*(user space) coverage* and *cold start* criteria) and providing them recommendations without asking them to have a tagging history. Besides, most of the introduced approaches fulfill the diversity of recommendations while the serendipity and the scalability criteria are hard to check as the authors do not provide sufficient information about their approaches.

### 6.7 User study

In this subsection, we report the results of the user study carried out to get information about users' feedback on our personalized recommender system. We choose six test subjects with different profile information as follows: (*Nidhal*, Male, 30 years, Academic assistant, Tunisia), (*Imen*, Female, 26 years, Student, Tunisia), (*Roxane*, Female, 27 years, Educator, France), (*Raymond*, Male, 58 years, Retired, Belgium), (*Wassim*, Male, 24 years, Engineer, Canada) and (*Quentin*, Male, 28 years, Optician, France). Test subjects are asked to perform a set of tasks using the system and answering questions afterwards about their experience. Such question's answers could provide insights for data that is not directly observable, such as assessing whether the subject enjoyed the user interface, or whether the user appreciated recommendations. The study is divided into four tasks: (i) *Quality of the recommendation* We recommend to each user a list of recommended resources (movies and books) and we ask them to rate each item. The following scale was used for the rating of resources: 5:*Very good*, 4:*Good*, 3:*Fair*, 2:*Poor*, 1:*Very Poor*; (ii) *Resource Recommendation* For each test subject, we provide an expanded list of resources including five pertinent ones[9] and ask the user to select a set of items that (s)he considers as interesting. Then, we compute the portion of pertinent selected resources. Ideally, the resources selected by the user correspond to the resources recommended by FolkRec, i.e., the pertinent ones; (iii) *Tag Suggestion* When a user is about sharing a book or a movie, we suggest him (her) an expanded list of tags including three pertinent ones and ask the user to select three tags that (s)he considers appropriate for such resource. Then, we compare the selected tags with the pertinent ones in order to see whether users actually chose the tags recommended by FolkRec; and (iv) *User Proposition* For each test subject, we propose a list of users with their respective profiles and ask them to select the user(s) that they would add as a friend. Then, we check which user they choose among the list. Finally, test subjects can justify their different choices by writing a free-form text.

For the first task, users rate the recommended movies with an average note equal to 3, 68 and 3, 1 for recommended books which highlights a quite good appreciation of the recommendations provided by FolkRec. Among users, *Roxane* had really enjoyed the recommended movies (twice "very good" and twice "good") *"I really enjoyed that the system recommends my favourite movies: Godfather and Star Wars."* while the same user also appreciate the recommended books *"I was unfamiliar with most of the recommended books and i really appreciate that."* which highlights the *novelty* aspect of our recommendations. The second task shows that 49.8 % of the movies selected by test subjects match the pertinent ones. For example, *Roxane* selected two movies in addition to the four recommended ones. The pertinence of selected books are also up to 49.8 %: for example, 66 % of the books selected by *Raymond* are pertinent; the subject comments that *"the books were very*

[9] Pertinent resources (*resp.* tags or users) are those (*resp.* tags or users) recommended by FolkRec.

*diverse, from classic to adventure ones."* While *Raymond* appreciate the *diversity* of our recommendations, the subject *Nidhal* underlines the *serendipity* of the propositions *"I was nicely surprised by recommended books, it was an unexpected result. Most of the books are unknown for me. It allowed me to rediscover my tastes."* Contrariwise to resources which may evolve from a user to another one within same profiles, the results of the third task (i.e., tag suggestion) highlight that users with same profile converges to a common vocabulary. Thus, 66.4 % of the tags selected by subjects matches the tags used by users with same profile. For example, when annotating the movie *Jurassic Park*, *Raymond* selects the tags *dinosaurs*, *steven_spielberg* and *oscar_winner* instead of *genetics*, *action* and *thriller*, which perfectly match the tags used by users within the same profile. Finally, the last task shows that half of the test subjects choose a user with a same profile. It also highlights that if some users (*Roxane*, *Raymond* and *Quentin*) look for friends with the same profile, other users are interested in users with a different profile. For example, the academic assistant *Nidhal* would add as a friend *Patrick* who is a librarian interested by movies based on a true story. As overall feedback, the users were generally pleased with the recommendations of resources, tags and users as mentioned by *Wassim*: *"I found the questionnaire very interesting. First because it's mainly about movies which is one of my hobbies. I enjoyed answering the questions and it made me remember some old movies that I like: Titanic, Braveheart, Blade Runner, etc."*

However, the limit of our user study is that it is restricted to the profile of the test subjects which is an incomplete information for recommendations. One of our test subjects, *Imen* notice: *"I don't appreciate the recommended books! And I don't specially like girlie movie, I prefer action movie like Seven or Die Hard."* Such issue can be resolved by an online evaluation through a real-time analysis of tags and resources shared by users. Such analysis may for example tell us that the user *Imen* is interested not only in romance but also in action movies. We conclude that the user study is able to provide users with good recommendations; however, an online tracking of its tags and resources will improve the results of our recommender system.

## 7 Conclusion and perspectives

In this paper, we considered a new dimension into the *folksonomy* in order to mine quadri-concepts. Such concepts are then used in order to offer a personalized choice of tags and resources to users. The evaluation of our recommender algorithm showed good results towards the different metrics and the general feedback of users. However, our recommender system still dependant, at the beginning, of the extraction of quadri-concepts. Moreover, dynamic updates in the *folksonomy* are not taking into account and involve the restart of the quadri-concept's extraction process. Among perspectives of our works, when considering the users' profile, we may explore other implicit and explicit information as reviews, comments or browsing history and click streams in order to have the evolutionary aspect of the users' profiles to better purchase their interests. Moreover, we may propose an online evaluation to have real-time information about tags and resources shared by users in order to improve the recommendations by incremental updates of the system (Valtchev et al. 2003). In addition, it would be interesting if we could vary the degree of closeness of users in order to evaluate the impact on experimental results. Finally, we could evaluate which user's information are the most important for recommendation. Indeed, it could be interesting for further experiment to only use the information the most discriminant and also to evaluate recommendations whenever the user does not give his personal information.

**Acknowledgments** Thanks to the PHC Utique project EXQUI 11G1417 for financial support to the first author, as well as project (French-Brazilian GDRI "Web of Sciences") for fruitful discussion on this topic. We also thank the anonymous referees for their valuable remarks and suggestions.

## References

Agarwal D, Chen B-C (2010) fLDA: matrix factorization through latent dirichlet allocation. In: Proceedings of the third ACM international conference on web search and data mining, WSDM '10, ACM, New York, pp 91–100

Baeza-Yates RA, Ribeiro-Neto B (1999) Modern information retrieval. Addison-Wesley Longman Publishing Co., Inc., Boston

Basile P, Gendarmi D, Lanubile F, Semeraro G (2007) Recommending smart tags in a social bookmarking system. Bridg Gap Between Sem Web Web 2:22–29

Bellogín A, Cantador I, Castells P (2013) A comparative study of heterogeneous item recommendations in social systems. Inf Sci 221:142–169

Breuss M, Tsagkias M (2014) Learning from user interactions for recommending content in social media. In: 36th European conference on information retrieval (ECIR'14)

Cerf L, Besson J, Nguyen K, Boulicaut J (2013) Closed and noise-tolerant patterns in n-ary relations. Data Min Knowl Discov 26(3):574–619

Cerf L, Besson J, Robardet C, Boulicaut J-F (2009) Closed patterns meet n-ary relations, ACM TKDD 3 3:1–3:36

Das M, Thirumuruganathan S, Amer-Yahia S, Das G, Yu C (2012) Who tags what? An analysis framework. Proc PVLDB 5(11):1567–1578

De Meo P, Quattrone G, Ursino D (2010) A query expansion and user profile enrichment approach to improve the performance of recommender systems operating on a folksonomy. User Model User-Adapt Interact 20(1):41–86

Diederich J, Iofciu T (2006) Finding communities of practice from user profiles based on folksonomies. In Proceedings of the 1st international workshop on TEL-CoPs, Crete, Greece, pp 288–297
Herlocker JL, Konstan JA, Terveen LG, Riedl JT (2004) Evaluating collaborative filtering recommender systems. ACM Trans Inf Syst 2004:5–53
Hu J, Wang B, Tao Z (2011) Personalized tag recommendation using social contacts. In: Proceedings. of workshop SRS'11, in conjunction with CSCW
Jelassi MN, Ben Yahia S, Mephu Nguifo E (2012) A scalable mining of frequent quadratic concepts in d-folksonomies, ArXiv e-printsarXiv:1212.0087
Jelassi MN, Ben Yahia S, Mephu Nguifo E (2013) A personalized recommender system based on users' information in folksonomies. In: Proceedings of the 22nd international conference on world wide web companion, WWW '13 Companion, pp 1215–1224
Jäschke R, Hotho A, Schmitz C, Ganter B, Stumme G (2008) Discovering shared conceptualizations in folksonomies. Web Sem 6:38–53
Jäschke R, Marinho L, Hotho AA, Lars S-T, Stum G (2007) Tag recommendations in folksonomies. In: Proceedings of the 11th ECML PKDD, Warsaw, Poland, pp 506–514
Kim HK, Oh HY, Gu JC, Kim JK (2011) Commenders: a recommendation procedure for online book communities. Electron Commer Rec Appl 10(5):501–509
Landia N, Anand S (2009) Personalised tag recommendation. Recommender Systems & the Social Web, New York
Liang H (2010) User profiling based on folksonomy information in web 2.0 for personalized recommender systems, Ph.D. thesis, Queensland University of Technology
Lika B, Kolomvatsos K, Hadjiefthymiades S (2014) Facing the cold start problem in recommender systems. Expert Syst Appl 41(4):2065–2073
Lipczak M (2008) Tag recommendation for folksonomies oriented towards individual users. In: Proceedings of the ECML/PKDD discovery challenge, Antwerp, Belgium, pp 84–95
Mika P (2007) Ontologies are us: a unified model of social networks and semantics. J Web Sem 5(1):5–15
Noll M, Michael G, Meinel C (2007) Web search personalization via social bookmarking and tagging. In: Proceedings of the 6th ISWC/ASWC, Busan, Korea, pp 367–380
Qumsiyeh R, Ng Y-K (2012) Predicting the ratings of multimedia items for making personalized recommendations. In: SIGIR'12, ACM, New York, pp 475–484
Ricci F, Rokach L, Shapira B, Kantor PB (eds) (2011) Recommender systems handbook. Springer, New York
Rikitianskii A, Harvey M, Crestani F (2014) A personalised recommendation system for context-aware suggestions. In: Advances in Information Retrieval. Lecture Notes in Computer Science, vol 8416. Springer, New York, pp 63–74
Said A, Bellogín A (2014) Comparative recommender system evaluation: benchmarking recommendation frameworks. In: Proceedings of the 8th ACM conference on recommender systems, RecSys, ACM, ACM, Foster City
Said A, Kille B, De Luca EW, Albayrak S (2011) Personalizing tags: a folksonomy-like approach for recommending movies. In: 2nd international HetRec, pp 53–56
Strohmaier M, KöRner C, Kern R (2012) Understanding why users tag: a survey of tagging motivation literature and results from an empirical study. Web Sem 17:1–11
Trabelsi C, Jelassi N, Ben Yahia S (2012) Scalable mining of frequent tri-concept. In: Proceedings of the 15th PAKDD, Kuala Lampur, Malaysia, pp 231–242
Vallet D, Cantador I, Jose JM (2010) Personalizing web search with folksonomy-based user and document profiles. In: Proceedings of the 32nd ECIR, Berlin, Heidelberg, pp 420–431
Valtchev P, Hacene MR, Missaoui R (2003) A generic scheme for the design of efficient on-line algorithms for lattices. In: ICCS, pp 282–295
Weiss SM, Kulikowski CA (1991) Computer systems that learn: classification and prediction methods from statistics, neural nets, machine learning, and expert systems. Morgan Kaufmann Publishers Inc., San Francisco
Yang S-H, Long B, Smola AJ, Zha H, Zheng Z (2011) Collaborative competitive filtering: Learning recommender using context of user choice. In: Proceedings of the 34th international ACM SIGIR conference on research and development in information retrieval, SIGIR '11, ACM, New York, pp 295–304