Knowledge-Based Systems xxx (2013) xxx–xxx

Contents lists available at SciVerse ScienceDirect

Knowledge-Based Systems

journal homepage: www.elsevier.com/locate/knosys

# Recommender systems survey

J. Bobadilla *, F. Ortega, A. Hernando, A. Gutiérrez

*Universidad Politécnica de Madrid, Ctra. De Valencia, Km. 7, 28031 Madrid, Spain*

ARTICLE INFO

*Article history:*
Received 7 October 2012
Received in revised form 4 March 2013
Accepted 19 March 2013
Available online xxxx

*Keywords:*
Recommender systems
Collaborative filtering
Similarity measures
Evaluation metrics
Prediction
Recommendation
Hybrid
Social
Internet of things
Cold-start

ABSTRACT

Recommender systems have developed in parallel with the web. They were initially based on demographic, content-based and collaborative filtering. Currently, these systems are incorporating social information. In the future, they will use implicit, local and personal information from the Internet of things. This article provides an overview of recommender systems as well as collaborative filtering methods and algorithms; it also explains their evolution, provides an original classification for these systems, identifies areas of future implementation and develops certain areas selected for past, present or future importance.



## 1. Introduction

Recommender Systems (RSs) collect information on the preferences of its users for a set of items (e.g., movies, songs, books, jokes, gadgets, applications, websites, travel destinations and e-learning material). The information can be acquired explicitly (typically by collecting users' ratings) or implicitly [134,60,164] (typically by monitoring users' behavior, such as songs heard, applications downloaded, web sites visited and books read). RS may use demographic features of users (like age, nationality, gender). Social information, like followers, followed, twits, and posts, is commonly used in Web 2.0. There is a growing tend towards the use of information from Internet of things (e.g., GPS locations, RFID, real-time health signals).

RS make use of different sources of information for providing users with predictions and recommendations of items. They try to balance factors like accuracy, novelty, dispersity and stability in the recommendations. Collaborative Filtering (CF) methods play an important role in the recommendation, although they are often used along with other filterning techniques like content-based, knowledge-based or social ones.

CF is based on the way in which humans have made decisions throughout history: besides on our own experiences, we also base our decisions on the experiences and knowledge that reach each of us from a relatively large group of acquaintances.

Recently, RS implementation in the Internet has increased, which has facilitated its use in diverse areas [171]. The most common research papers are focused on movie recommendation studies [53,230]; however, a great volume of literature for RS is centered on different topics, such as music [134,162,216], television [238,18], books [164,88], documents [206,184,183,185], e-learning [241,30], e-commerce [104,54], applications in markets [67] and web search [154], among others.

The kinds of filtering most used at the beginning of the RS (collaborative, content-based and demographic) were described in [177]. Breese et al. [43] evaluated the predictive accuracy of different algorithms for CF; later, the classical paper [94] describes the base for evaluating the Collaborative Filtering RS.

The evolution of RS has shown the importance of hybrid techniques of RS, which merge different techniques in order to get the advantages of each of them. A survey focused on the hybrid RS has been presented in [47]. However, it does not deal with the role of social-filtering, a technique which has become more popular in the recent years through social networks.

The neighborhood-based CF has been the recommendation method most popular at the beginning of the RS; Herlocker et al. [93] provides a set of guidelines for designing neighborhood-based prediction systems. Adomavicius and Tuzhilin [3] present an overview on the RS field standing out the most complex areas on which

* Corresponding author. Tel.: +34 913365133; fax: +34 913367527.
*E-mail address:* jesus.bobadilla@upm.es (J. Bobadilla).

0950-7051/$ - see front matter © 2013 Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.knosys.2013.03.012

researchers in RS should focus in the "next generation of RS": limited content analysis and overspecialization in content-based methods, cold-start and sparsity in CF methods, model-based techniques, nonintrusiveness, flexibility (real-time customization), etc.

While researchers have been developing RS, different survey papers have been published summarizing the most important issues in this field. In view of the impossibility of showing every detail of all these techniques in just a paper, this publication selects those issues the authors have felt most suitable to understand the evolution of RS.

While the existing surveys focus on the most relevant methods and algorithms of the RS field, our survey instead tries to enhance the evolution of the RS: from a first phase based on the traditional Web to the present second phase based on social Web, which is presently progressing to a third phase (Internet of things). With the purpose of being useful to the new readers of RS field, we have included in this survey some traditional topics: RS foundations, *k*-Nearest Neighbors algorithm, cold-start issues, similarity measures, and evaluation of RS. The rest of the paper deals with novel topics that existing surveys do not consider. Through this survey, advanced readers in RS will study in depth concepts, classifications and approaches related to social information (social filtering: followers, followed, trust, reputation, credibility, content-based filtering of social data; social tagging and taxonomies), recommending to groups of users and explaining recommendations. Readers interested in brand new and future applications will find this survey useful since it informs about the most recent works in location-aware RS trends and bio-inspired approaches. They will also discover some important issues, such as privacy, security, P2P information and Internet of things use (RFID data, health parameters, surveillance data, teleoperation, telepresence, etc.).

According to the idea that RS tend to make use of different sources of information (collaborative, social, demographic, content, knowledge-based, geographic, sensors, tags, implicit and explicit data acquisition, etc.), this survey emphasizes hybrid architectures, based on making recommendations through different known technologies (each one designed on behalf of a specific source of information).

Much of the quality of a survey can be measured by an appropriate choice of its references. This survey contains 249 references systematically obtained, which have been selected taking into account factors like the number of recent citations and the importance of the journal in which the paper has been published.

The remainder of this article is structured as follows: In Section 2, we explain concisely the methodology used to select the most significative papers on the RS field. Section 3 describes the RS foundations: methods, algorithms and models used for providing recommendations based from the information of the traditional web: ratings, demographic data and item data (CF, demographic filtering, content-based filtering and hybrid filtering). Section 4 describes measures for evaluating the quality of the RS predictions and recommendations. Section 5 shows the use of social information from Web 2.0 for making recomendations through concepts like trust, reputation and credibility. We will also describe techniques based on content-based for social information (e.g. tags and posts). Section 6 focusses on two important areas (although not very well studied yet): recommendation to group of users and explanation of recommendations. Section 7 focusses on recommender system trends, covering bio-inspired approaches and Web 3.0 information filtering such as location-aware RS. Section 8 explains related works and the original contributions of this survey.

The concluding section summarizes the RS history and focuses on the type of data used as well as the development of algorithms and evaluation measures. The conclusions section also indicates seven new areas that we consider likely to be the focus of RS research in the scientific community in the near future.

## 2. Methodology

An initial study was performed to determine the most representative topics and terms in the RS field. First, 300 RS papers were selected from journals, with a higher priority for current and for often-cited articles. Next, we extracted from these 300 papers the most significant terms. We gave the most emphasis to keywords, less emphasis to titles and, finally, the least emphasis to abstracts.

We have overlooked common words, like articles, prepositions and general-use words from the remaining pool, we selected 300 terms represented in the RS field. From a matrix of articles $\times$ words, wherein we stored the importance of each word from each article, we generated a tree of relationships between the words. Fig. 1 depicts the most significant section of the graph (due to space constraints, the entire tree is not shown, but it is provided as additional material in Fig. 1 *AdditionalData.png*). The short distances between words indicate the highest similarities; warm colors indicate a greater reliability for the relationships. The size of the nodes indicates the importance of the words as a function of the parameters $N^k$, $N^t$, $N^a$ (number of significative words in the keywords, title and abstract) and $N^k_w, N^t_w, N^a_w$ (number of times that the word $w$ appears in the keywords, title and abstract). The equation used to determine the importance of each word $w$ is as follows:

$$f_w = \frac{1}{3}\left(\frac{N^k_w}{N^k} + \frac{N^t_w}{N^t \log\frac{N^a}{N^t}} + \frac{N^a_w}{N^a\frac{N^a}{N^t}}\right)$$

*Example*: we will consider a paper where $N^k = 5$ keywords, $N^t = 11$ words in the title, and $N^a = 52$ words of abstract length. We will get the values of $f_{factorization}$ and $f_{matrix}$, where the word 'factorization' appears once as a keyword, once in the title and three times in the abstract; the word 'matrix' does not appear as a keyword, but it is contained once in the title and twice in the abstract. The importance of these words will be:

$$f_{factorization} = \frac{1}{3}\left(\frac{1}{5} + \frac{1}{11\log\frac{52}{11}} + \frac{3}{52\frac{52}{11}}\right) = 0.09$$

$$f_{matrix} = \frac{1}{3}\left(\frac{0}{5} + \frac{1}{11\log\frac{52}{11}} + \frac{2}{52\frac{52}{11}}\right) = 0.02$$

The information depicted in Fig. 1 is used to identify the most relevant aspects of RS. They are represented by the most significant words in the graph and the related terms. The articles referenced herein were chosen based on the following criteria: (a) the transcendence of the subject according to the importance of the words in Fig. 1; (b) its historical contribution (a significant fraction of the classic reference articles are included); (c) the number of times the article is cited; (d) articles published in journals with an impact factor were preferred over conferences and workshops; and (e) recent articles were preferred over articles published many years ago. Fig. 2 shows a temporal distribution for the referenced papers.

We use the clusters of words in Fig. 1 to structure the explications of the survey. For each concept explained: (1) we have obtained their keywords and all the words related to them according to Fig. 1; (2) we have identified, among the set of 300 papers, those which are more related to the set of words associated to the concept; (3) we have selected the subset of papers which deal with the concept, giving priority to those with high values in criteria like importance of the paper and the number of cites; and (4) we have tried to balance the number of times a paper is referenced in our survey, aiming to reference most of the 300 papers selected.

Please cite this article in press as: J. Bobadilla et al., Recommender systems survey, Knowl. Based Syst. (2013), http://dx.doi.org/10.1016/j.knosys.2013.03.012

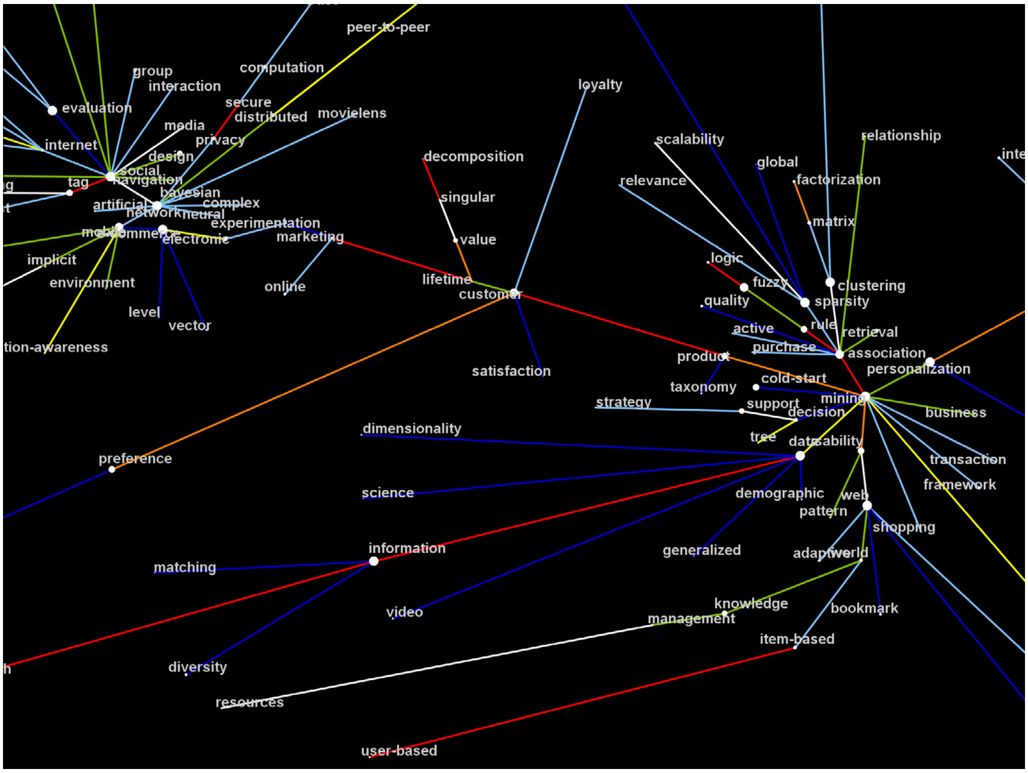


**Fig. 1.** Words represented in the recommender systems research field. Short distances indicate higher similarities, and a warm color indicates greater reliability. The size of the nodes is proportional to the importance of the words.

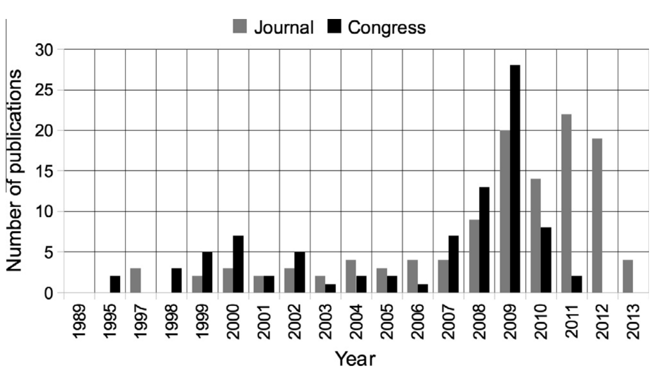


**Fig. 2.** Temporal distribution for the referenced papers.

## 3. Recommender systems foundations

This section presents the most relevant concepts on which the traditional RS are based. Here, we provide general descriptions on the classical taxonomies, algorithms, methods, filtering approaches, databases, etc. Besides, we show a graphic depicting the traditional models of recommendations and their relations. Next, we will describe the cold-start problem, which will illustrate the difficulty of making collaborative recommendation when the RS contains a small amount of data. Next, we will describe the *k*NN algorithm; the most used algorithm for implementing RS based on CF. Finally, we will describe different proposed similarity measures for comparing users or items. We will show graphics for measuring the quality of these similarity measures.

### 3.1. Fundamentals

The process for generating an RS recommendation is based on a combination of the following considerations:

- The type of data available in its database (e.g., ratings, user registration information, features and content for items that can be ranked, social relationships among users and location-aware information).
- The filtering algorithm used (e.g., demographic, content-based, collaborative, social-based, context-aware and hybrid).
- The model chosen (e.g., based on direct use of data: "memory-based," or a model generated using such data: "model-based").
- The employed techniques are also considered: probabilistic approaches, Bayesian networks, nearest neighbors algorithm; bio-inspired algorithms such as neural networks and genetic algorithms; fuzzy models, singular value decomposition techniques to reduce sparsity levels, etc.
- Sparsity level of the database and the desired scalability.
- Performance of the system (time and memory consuming).
- The objective sought is considered (e.g., predictions and top N recommendations) as well as
- The desired quality of the results (e.g., novelty, coverage and precision).

Research in RS requires using a representative set of public *databases* to facilitate investigations on the techniques, methods and algorithms developed by researchers in the field. Through these databases, the scientific community can replicate experiments to validate and improve their techniques. Table 1 lists the current public databases referenced most often in the literature. Last.Fm and Delicious incorporate implicit ratings and social information; their data were generated from the versions released in the HetRec, 2011 data sets, hosted by the GroupLens research Group.

The internal functions for RS are characterized by the *filtering algorithm*. The most widely used classification divides the filtering algorithms into [3,51,203]: (a) collaborative filtering, (b) demographic filtering, (c) content-based filtering and (d) hybrid filtering.

Please cite this article in press as: J. Bobadilla et al., Recommender systems survey, Knowl. Based Syst. (2013), http://dx.doi.org/10.1016/j.knosys.2013.03.012

**Table 1**
Most often used memory-based recommender systems public databases.

| | Without social information | | | | | | With social information (hosted by the GroupLens) | | |
|---|---|---|---|---|---|---|---|---|---|
| | MovieLens 1M | MovieLens 10M | Netflix | Jester | EachMovie | Book-crossing | ML | Last.Fm | Delicious |
| Ratings | 1 million | 10 million | 100 million | 4.1 million | 2.8 million | 1.1 million | 855,598 | 92,834 | 104,833 |
| Users | 6040 | 71,567 | 480,189 | 73,421 | 72,916 | 278,858 | 2113 | 1892 | 1867 |
| Items | 3592 | 10,681 | 17,770 | 100 | 1628 | 271,379 | 10,153 | 17,632 | 69,226 |
| Range | $\{1,\ldots,5\}$ | $\{1,\ldots,5\}$ | $\{1,\ldots,5\}$ | $-10$, 10 | [0,1] | $\{1,\ldots,10\}$ | $\{1,\ldots,5\}$ | Implicit | Implicit |
| Tags | N/A | N/A | N/A | N/A | N/A | N/A | 13222 | 11946 | 53388 |
| Tags assignment | N/A | N/A | N/A | N/A | N/A | N/A | 47957 | 186479 | 437593 |
| Friends relations | N/A | N/A | N/A | N/A | N/A | N/A | N/A | 25434 | 15328 |
| Items | Movies | Movies | Movies | Jokes | Movies | Books | Movies | Music | URL's |

*Content-based filtering* [131,11,158] makes recommendations based on user choices made in the past (e.g. in a web-based e-commerce RS, if the user purchased some fiction films in the past, the RS will probably recommend a recent fiction film that he has not yet purchased on this website). Content-based filtering also generates recommendations using the content from objects intended for recommendation; therefore, certain content can be analyzed, like text, images and sound. From this analysis, a similarity can be established between objects as the basis for recommending items similar to items that a user has bought, visited, heard, viewed and ranked positively.

*Demographic filtering* [177,126,185] is justified on the principle that individuals with certain common personal attributes (sex, age, country, etc.) will also have common preferences.

*Collaborative Filtering* [3,94,92,51,212] allows users to give ratings about a set of elements (e.g. videos, songs, films, etc. in a CF based website) in such a way that when enough information is stored on the system, we can make recommendations to each user based on information provided by those users we consider to have the most in common with them. CF is an interesting open research field [232,34,32]. As noted earlier, user ratings can also be

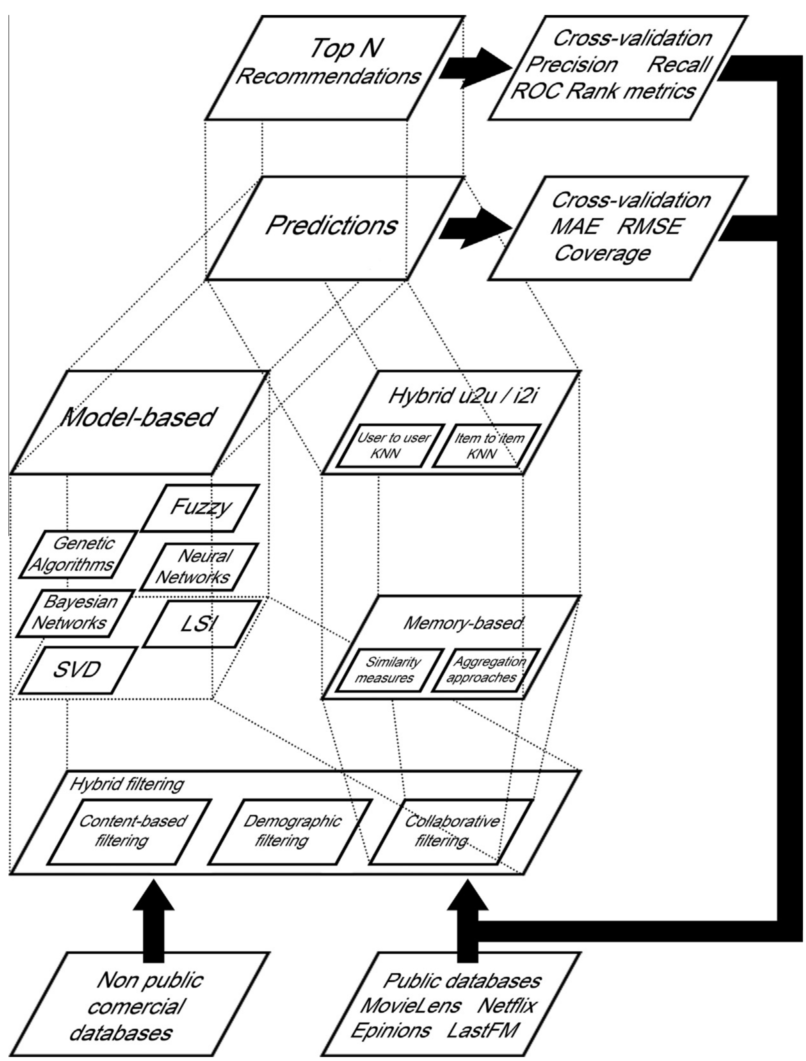


**Fig. 3.** Traditional models of recommendations and their relationships.

Please cite this article in press as: J. Bobadilla et al., Recommender systems survey, Knowl. Based Syst. (2013), http://dx.doi.org/10.1016/j.knosys.2013.03.012

implicitly acquired (e.g., number of times a song is heard, information consulted and access to a resource).

The most widely used algorithm for collaborative filtering is the *k Nearest Neighbors (kNN)* [3,203,32]. In the *user to user* version, *k*NN executes the following three tasks to generate recommendations for an active user: (1) determine *k users neighbors (neighborhood)* for the active user *a*; (2) implement an *aggregation approach* with the ratings for the neighborhood in items not rated by *a*; and (3) extract the predictions from in step 2 then select the *top N recommendations*.

*Hybrid filtering* [47,185]. Commonly uses a combination of CF with demographic filtering [224] or CF with content-based filtering [18,60] to exploit merits of each one of these techniques. Hybrid filtering is usually based on bioinspired or probabilistic methods such as genetic algorithms [76,99], fuzzy genetic [7], neural networks [133,62,192], Bayesian networks [50], clustering [209] and latent features [199].

A widely accepted taxonomy divides recommendation methods into memory-based and model-based method categories:

*Memory-based methods* [3,51,123,214]. Memory-based methods can be defined as methods that (a) act only on the matrix of user ratings for items and (b) use any rating generated before the referral process (i.e., its results are always updated). Memory-based methods usually use similarity metrics to obtain the distance between two users, or two items, based on each of their ratios.

*Model-based methods* [3,212]. Use RS information to create a model that generates the recommendations. Herein, we consider a method model-based if new information from any user outdates the model. Among the most widely used models we have *Bayesian classifiers* [59], *neural networks* [107], *fuzzy systems* [234], *genetic algorithms* [76,99], *latent features* [251] and *matrix factorization* [142], among others.

To reduce the problems from high levels of sparsity in RS databases, certain studies have used *dimensionality reduction* techniques [202]. The reduction methods are based on *Matrix Factorization* [124,142,143]. Matrix factorization is especially adequate for processing large RS databases and providing scalable approaches [215]. The model-based technique *Latent Semantic Index* (LSI) and the reduction method *Singular Value Decomposition* (SVD) are typically combined [224,244,48]. SVD methods provide good prediction results but are computationally very expensive; they can only be deployed in static off-line settings where the known preference information does not change with time.

RS can use *clustering* techniques to improve the prediction quality and reduce the cold-start problem when applied to hybrid filtering. It is typical to form clusters of items in hybrid RS [209,237]. A different common approach uses clustering both for items and users (*bi-clustering*) [252,85]. RS comprising social information have been clustered to improve the following areas: *tagging* [208], explicit *social links* [179] and explicit *trust information* [181,70].

The graph in Fig. 3 shows the most significant traditional methods, techniques and algorithms for the recommendation process as well as their relationships and groupings. Different sections of this paper provide more detail on the most important aspects involved in the recommendation process.

As may be seen in Fig. 3, we can use some of the traditional filtering methods (content-based, demographic and collaborative) applied to databases. Model-based technologies (genetic algorithms, neural networks, etc.) make use of this kind of information. Typical memory-based approaches are: item to item; user to user; and hybrids of the two previous. The main purpose of both memory-based and model-based approaches is to get the most accurate predictions in the tastes of users. The accuracy of these predictions may be evaluated through the classical information retrieval measures, like MAE, precision, and recall. Researchers make use of these measures in order to improve the RS methods and technologies.

### 3.2. Cold-start

The cold-start problem [203,3] occurs when it is not possible to make reliable recommendations due to an initial lack of ratings. We can distinguish three kinds of cold-start problems: *new community*, *new item* and *new user*. The last kind is the most important in RS that are already in operation.

The new community problem [204,129] refers to the difficulty, when starting up a RS, in obtaining, a sufficient amount of data (ratings) for making reliable recommendations. Two common ways are used for tackling this problem: to encourage users to make ratings through different means; to take CF-based recommendations when there are enough users and ratings.

The new item problem [174,172] arises because the new items entered in RS do not usually have initial ratings, and therefore, they are not likely to be recommended. In turn, an item that is not recommended goes unnoticed by a large part of the community of users, and as they are unaware of it they do not rate it; this way, we can enter a vicious circle in which a set of items of the RS are left out of the ratings/recommendations process. The new item problem has less of an impact on RS in which the items can be discovered via other means (e.g. movies) than in RS where this is not the case (i.e. e-commerce, blogs, photos, videos, etc.). A common solution to this problem is to have a set of motivated users who are responsible for rating each new item in the system.

The new user problem [190,197] represents one of the great difficulties faced by the RS in operation. Since new users in the RS have not yet provided any rating in the RS, they cannot receive any personalized recommendations based on memory-based CF; when the users enter their firsts ratings they expect the RS to offer them personalized recommendations, but the number of ratings introduced in the RS is usually not yet sufficient to be able to make reliable CF-based recommendations, and, therefore, new users may feel that the RS does not offer the service they expected and they may stop using it.

The common strategy to tackle the new user problem consists of turning to additional information to the set of ratings in order to be able to make recommendations based on the data available for each user. The cold-start problem is often faced using hybrid approaches (usually CF-content based RS, CF-demographic based RS, CF-social based RS) [118,140]. Leung et al. [135] propose a novel content-based hybrid approach that makes use of cross-level association rules to integrate content information about domains items. Kim et al. [118] use collaborative tagging employed as an approach in order to grasp and filter users' preferences for items and they explore the advantages of the collaborative tagging for data sparseness and cold-start users (they collected the dataset by crawling the collaborative tagging *delicious* site). Weng et al. [228] combine the implicit relations between users' items preferences and the additional taxonomic preferences to make better quality recommendations as well as alleviate the cold-start problem. Loh et al. [140] represent user's profiles with information extracted from their scientific publications. Martinez et al. [148] present a hybrid RS which combines a CF algorithm with a knowledge-based one. Chen and He [56] propose a number of common terms/ term frequency (NCT/TF) CF algorithm based on demographic vector. Saranya and Atsuhiro [199] propose a hybrid RS that utilizes latent features extracted from items represented by a multi-attributed record using a probabilistic model. Park et al. [173] propose a new approach: they use filterbots, and surrogate users that rate items based only on user or item attributes.

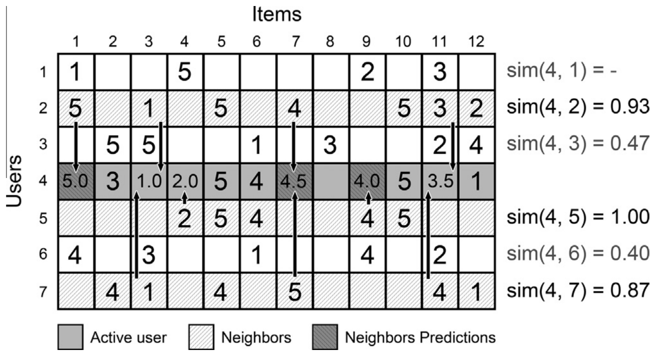


**Fig. 4.** User to user *k*NN algorithm example, *k* = 3. Similarity measure: 1 – (mean squared differences). Aggregation approach: average.

### 3.3. The k nearest neighbors recommendation algorithm

The *k Nearest Neighbors* (*k*NN) recommendation algorithm is the reference algorithm for the collaborative filtering recommendation process. Its primary virtues are simplicity and reasonably accurate results; its major pitfalls are low scalability and vulnerability to sparsity in the RS databases. This section provides a general explanation of this algorithm function.

CF based on the *k*NN algorithm is conceptually simple, with a straightforward implementation; it also generally produces good-quality predictions and recommendations. However, due to the high level of *sparsity* [142,29] in RS databases, similarity measures often encounter processing problems (typically from insufficient mutual ratings for a comparison of users and items) and *cold start* situations (users and items with low number of rankings) [204,98,36,135].

Another major problem for the *k*NN algorithm is its low *scalability* [142]. As the databases (such as Netflix) increase in size (hundreds of thousands of users, tens of thousands of items, and hundreds of millions of rankings), the process for generating a neighborhood for an active user becomes too slow; The similarity measure must be processed as often as new users are registered in the database. The item to item version of the *k*NN algorithm significantly reduces the scalability problem [200]. To this end, neighbors are calculated for each item; their top *n* similarity values are stored, and for a period of time, predictions and recommendations are generated using the stored information. Although the stored information does not include the ratings from previous processing/storage, outdated information for items is less sensitive than for the users.

A recurrent theme in CF research is generating *metrics* to calculate with accuracy and precision the existing similarity for the users (or items). Traditionally, a series of statistical metrics have been used [3,51], such as the *Pearson correlation*, *cosine*, *constraint Pearson correlation* and *mean squared differences*. Recently, metrics have been designed to fit the constraints and peculiarities of RS [31,35]. The *relevance* (*significance*) concept was introduced to afford more importance to more relevant users and items [34,227]. Additionally, a group of metrics was specifically designed to adequately function in cold-start situations [6,36].

The *k*NN algorithm is based on similarity measures. Next subsection provides further details on the current RS similarity measures. The similarity approaches typically compute the similarity between two users *x* and *y* (user to user) based on both users' item ratings. The item to item *k*NN version computes the similarity between two items *i* and *j*.

A formal approach of the *k*NN algorithm may be found in [32]. In this section, we will provide an illustrative example of this algorithm. The method for making recommendations is based on the following three steps:

(a) Using the selected similarity measure, we produce the set of *k* neighbors for the active user *a*. The *k* neighbors for *a* are the nearest *k* (similar) users to *u*.
(b) Once the set of *k* users (neighbors) similar to active *a* has been calculated, in order to obtain the prediction of item *i* on user *a*, one of the following aggregation approaches is often used: the average, the weighted sum and the adjusted weighted aggregation (deviation-from-mean).
(c) To obtain the top-*n* recommendations, we choose the *n* items, which provide most satisfaction to the active user according to our predictions.

Fig. 4 shows a case study using the user to user *k*NN algorithm mechanism.

In the *item to item* version [200,77] of the *k*NN algorithm, the following three tasks are executed: (1) determine *q* items neighbors for each item in the database; (2) for each item *i* not ranked by the active user *a*, calculate its prediction based on the ratings of *a* from the *q* neighbors of *i*; and (3) select the top *n* recommendations for the active user (typically the *n* major predictions from *a*). Step (1) can be executed periodically, which facilitates an accelerated recommendation with regard to the user to user version.

The item to item and user to user versions of the *k*NN algorithm can be combined [188] to take advantage of the positive aspects from each approach. These approaches are typically fused by processing the similarity between objects.

### 3.4. Similarity measures

A *metric* or a Similarity Measure (SM) determines the similarity between pairs of users (user to user CF) or the similarity between pairs of items (item to item CF). For this purpose, we compare the ratings of all the items rated by two users (user to user) or the ratings of all users who have rated two items (item to item).

The *k*NN algorithm is based essentially on the use of traditional *similarity metrics* of statistical origin. These metrics require, as the only source of information, the set of votes made by the users on the items (memory-based CF). Among the most commonly used traditional metrics we have: Pearson correlation (CORR), cosine (COS), adjusted cosine (ACOS), constrained correlation (CCORR), Mean Squared Differences (MSD) and Euclidean (EUC) [51,3].

We will describe and compare a representative group of SM used in the *k*NN algorithm. The SM discussed include the following variations: (a) cold-start and general cases, (b) based or not based on models, and (c) using trust information or only ratings. Table 2 shows a classification of the memory-based CF SM which will be tested in this section.

A new metric (JMSD) has recently been published, which besides using the numerical information from the ratings (via mean squared differences) also uses the non-numerical information provided by the arrangement of these (via Jaccard) [31]. Ortega et al. [169] use Pareto dominance to perform a pre-filtering process eliminating less representative users from the *k*-neighbur selection process while retaining the most promising ones.

A specialization of the memory-based CF SM, which appeared recently [35], uses the information contained in the votes of all users, instead of restricting it to the ratings of the two users compared (user to user) or the two items compared (item to item). We will call this SM SING (singularities).

The possibility exists to create a model (model-based CF) from the full set of users' ratings in order to later determine the similarity between pairs of users or pairs of items based on the model created. The potential advantages of this focus are an increase in the accuracy obtained, in the performance (time consuming) achieved or in both. The drawback is that the model must be regularly updated in order to consider the most recently entered set of ratings.

**Table 2**
Tested collaborative filtering similarity measures.

| | Not based on models | | Model-based |
|---|---|---|---|
| | No trust extraction | Trust extraction | |
| Traditional (only the ratings of both users or both items) | | | |
| Not tailored to cold-start users | JMSD, CORR, CCORR, COS, ACOS, MSD, EUC | | GEN |
| Tailored to cold-start users | PIP | UERROR | NCS |
| Extended to all the ratings | SING | TRUST | |

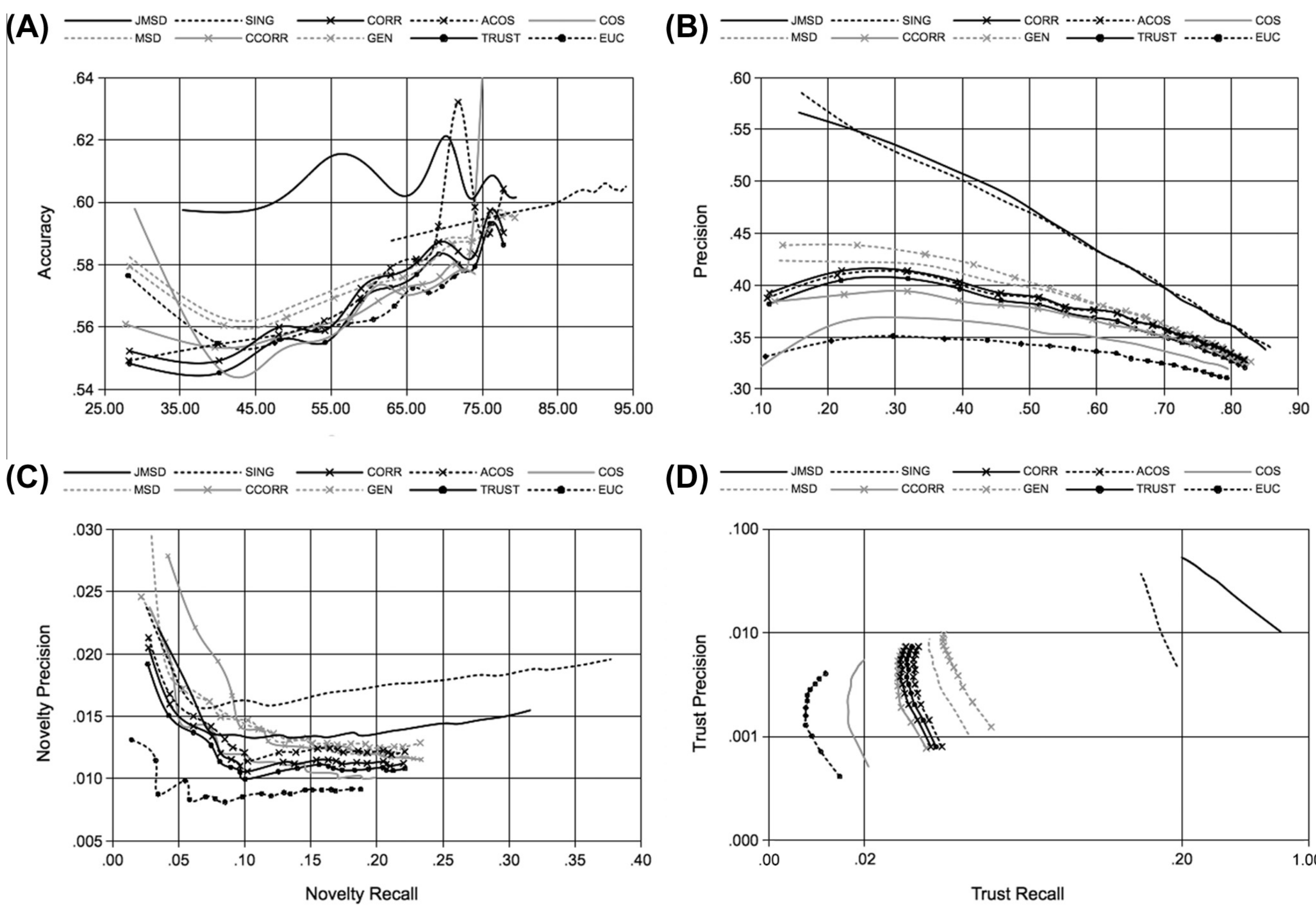


**Fig. 5.** Evaluation measures results obtained from current similarities measures; MovieLens database. (A) Prediction results, (B) recommendation results, (C) novelty results, and (D) trust results.

Bobadilla et al. [33] provides a metric based on a model generated using genetic algorithms. We will call this SM GEN (genetic-based).

As a result of the increase in web 2.0 websites on the Internet, a set of metrics has appeared which use the new social information available (friends, followers, followeds, etc.). Most of these SM are grouped in papers related to trust, reputation and credibility [71,239,138], although this situation is also produced in other fields [30]. These metrics could not be considered strictly memory-based CF, as they use additional information which not all RS have. In this sense, each SM proposed is tailored to a specific RS or at most to a very small set of RS which share the same structure in their social information.

There are SM [112,127] which aim to extract information related to trust and reputation by only using the users' set of ratings (memory-based CF). The advantage is that their use can be generalized to all CF RS; the drawback is that the social information extracted is really poor. We will call TRUST the SM proponed in Jeong et al. [112]. Currently, two new interesting SM get more coverage [38] and accuracy [61].

Fig. 5 shows the results from several evaluation measures generated by applying the SM discussed in this section. The results show that the RS-tailored SM are superior compared with the traditional SM from statistics. Processing for the memory-based information and results from Fig. 5 follow the framework schematic published previously [32].

There are so far research papers dealing with the cold-start problem through the users' ratings information. Ahn [6] presents a heuristic SM named PIP, that outperforms the traditional statistical SM

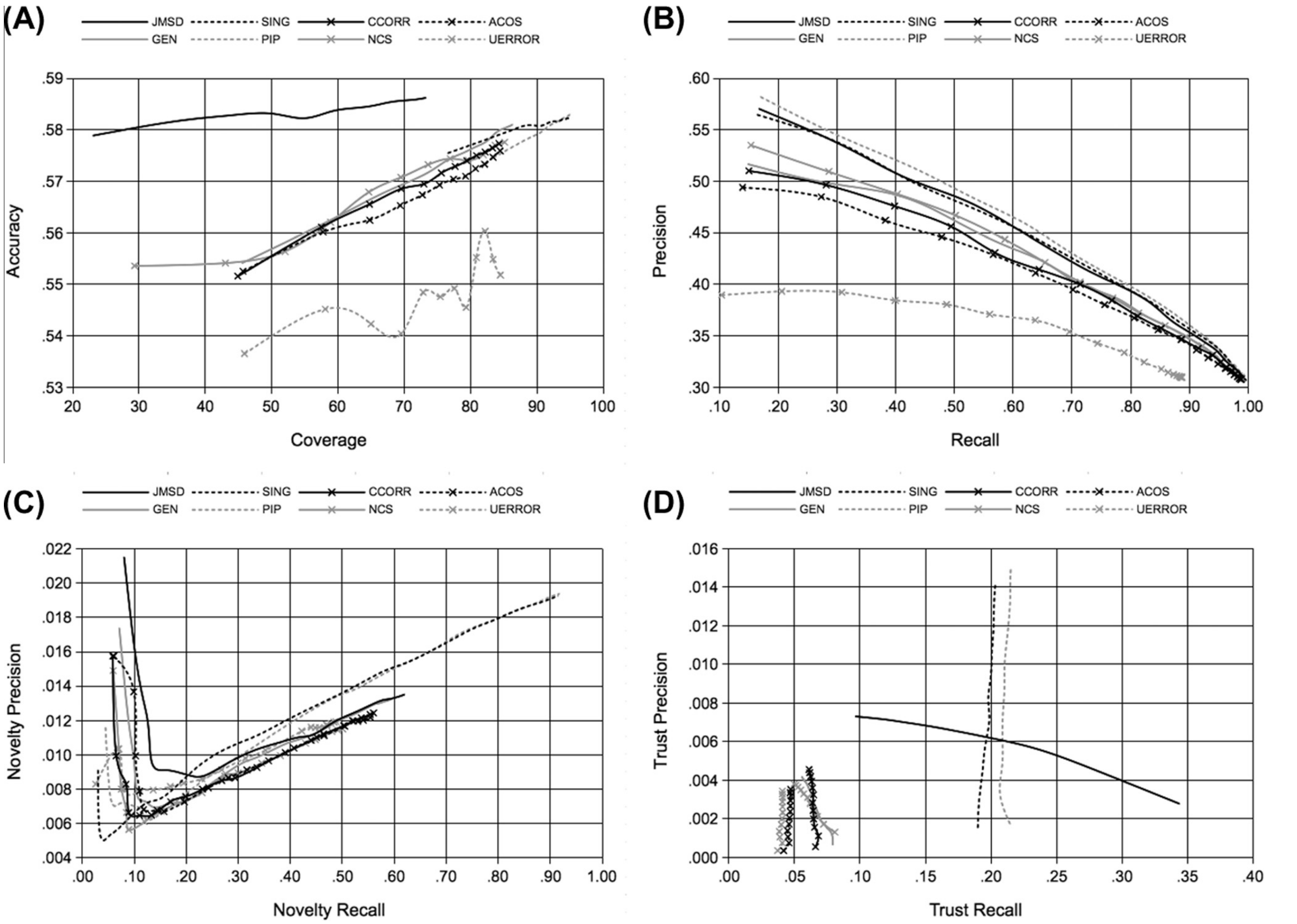


**Fig. 6.** Evaluation results obtained from current cold-start similarities measures. (A) Prediction results, (B) recommendation results, (C) novelty results, and (D) trust results.

(Pearson correlation, cosine, etc.). Heung-Nam et al. [98] proposes a method (UERROR) that predicts first actual ratings and subsequently identifies prediction errors for each user. Taking into account this error information, some specific "error-reflected" models, are designed. Bobadilla et al. [36] presents a metric based on neural learning (model-based CF) and adapted for new user cold-start situations, called NCS.

Fig. 6 shows results from several evaluation measures generated by applying the cold-start SM presented in this section; These results show that the RS-tailored SM are superior compared with the traditional SM from statistics. Since the database Movielens does not take into account cold-start users, we have removed ratings of this database in order to achieve cold-start users. Indeed, we have removed randomly between 5 and 20 ratings of those users who have rated between 20 and 30 items. In this way, we will regard those users who now result to rate between 2 and 20 items as cold-start users.

## 4. Evaluation of recommender systems results

Since RS research began, evaluation of predictions and recommendations has become important [94,201]. Research in the RS field requires *quality measures* and *evaluation metrics* [90] to know the quality of the techniques, methods, and algorithms for predictions and recommendations. *Evaluation metrics* [94,95] and *evaluation frameworks* [92,32] facilitate comparisons of several solutions for the same problem and selection from different promising lines of research that generate better results.

Because of evaluation measures, RS recommendations have gradually been tested and improved [48]. A representative set of existing evaluation measures has standard formulations, and a group of open RS public databases has been generated. These two advances have facilitated quality comparisons for new proposed recommendation methods and previously published methods; thus, RS methods and algorithms research has progressed continuously.

The most commonly used quality measures are the following [90,95]: (1) prediction evaluations, (2) evaluations for recommendation as sets, and (3) evaluations for recommendations as ranked lists. Fig. 5 shows results from applying several evaluation measures to a set of representative similarity measures.

Evaluation metrics [12] can be classified as [94,95] (a) prediction metrics: such as the *accuracy* ones: *Mean Absolute Error* (*MAE*), *Root of Mean Square Error* (*RMSE*), *Normalized Mean Average Error* (*NMAE*); and the *coverage* (b) set recommendation metrics: such as *Precision*, *Recall* and *Receiver Operating Characteristic* (*ROC*) [204] (c) rank recommendation metrics: such as the *half-life* [43] and the *discounted cumulative gain* [17] and (d) diversity metrics: such as the *diversity* and the *novelty* of the recommended items [105]. The validation process is performed by employing the most common cross validation techniques (*random sub-sampling* and *k-fold cross validation*) [21]; for cold-start situations, due to the limited number of users (or items) votes involved, the usual method chosen to carry out the experiments is *leave-one-out cross validation* [36].

Hernández and Gaudioso [95] propose an evaluation process based on the distinction between interactive and non-interactive subsystems. General publications and reviews also exist which include the most commonly accepted evaluation measures: *mean absolute error*, *coverage*, *precision*, *recall* and derivatives of these: *mean squared error*, *normalized mean absolute error*, *ROC* and *fallout*; Goldberg et al. [87] focuses on the aspects not related to the evaluation, Breese et al. [43] compare the predictive accuracy of various methods in a set of representative problem domains.

The majority of articles discuss attempted improvements to the accuracy of RS results (*RMSE*, *MAE*, etc.). It is also common to attempt an improvement in recommendations (precision, recall, ROC, etc.). However, additional objectives should be considered for generating greater user satisfaction [253], such as *topic diversification* and coverage *serendipity*.

Currently, the field has a growing interest in generating algorithms with diverse and innovative recommendations, even at the expense of accuracy and precision. To evaluate these aspects, various metrics have been proposed to measure recommendation *novelty* and *diversity* [105,220].

The frameworks aid in defining and standardizing the methods and algorithms employed by RS as well as the mechanisms to evaluate the quality of the results. Among the most significant papers that propose CF frameworks are Herlocker et al. [92] which evaluates the following: similarity weight, significance weighting, variance weighting, selecting neighborhood and rating normalization; Hernández and Gaudioso [95] proposes a framework in which any RS is formed by two different subsystems, one of them to guide the user and the other to provide useful/interesting items. Koutrika et al. [125] is a framework which introduces levels of abstraction in CF process, making the modifications in the RS more flexible. Antunes et al. [12] presents an evaluation framework assuming that evaluation is an evolving process during the system lifecicle.

The majority of RS evaluation frameworks proposed until now present two deficiencies: the first of these is the lack of formalization. Although the evaluation metrics are well defined, there are a variety of details in the implementation of the methods which, in the event they are not specified, can lead to the generation of different results in similar experiments. The second deficiency is the absence of standardization of the evaluation measures in aspects such as novelty and trust of the recommendations.

Bobadilla et al. [32] provides a complete series of mathematical formalizations based on sets theory. Authors provide a set of evaluation measures, which include the quality analysis of the following aspects: predictions, recommendations, novelty and trust.

Presented next is a representative selection of the RS evaluation quality measures most often used in the bibliography.

### 4.1. Quality of the predictions: mean absolute error, accuracy and coverage

In order to measure the accuracy of the results of an RS, it is usual to use the calculation of some of the most common prediction error metrics, amongst which the Mean Absolute Error (MAE) and its related metrics: mean squared error, root mean squared error, and normalized mean absolute error stand out.

We define $U$ as the set of the RS users, $I$ as the set of the RS items, $r_{u,i}$ the rating of user $u$ on item $i$, $\bullet$ the lack of rating ($r_{u,i} = \bullet$ means user $u$ has not rated item $i$), $p_{u,i}$ the prediction of item $i$ on user $u$.

Let $O_u = \{i \in I | p_{u,i} \neq \bullet \wedge r_{u,i} \neq \bullet\}$, set of items rated by user $u$ having prediction values. We define the MAE and RMSE of the system as the average of the user's MAE. We remark that the absolute difference between prediction and real value, $|p_{u,i} - r_{u,i}|$, informs about the error in the prediction.

$$\mathrm{MAE} = \frac{1}{\#U}\sum_{u\in U}\left(\frac{1}{\#O_u}\sum_{i\in O_u}|p_{u,i} - r_{u,i}|\right) \tag{1}$$

$$\mathrm{RMSE} = \frac{1}{\#U}\sum_{u\in U}\sqrt{\frac{1}{\#O_u}\sum_{i\in O_u}(p_{u,i} - r_{u,i})^2} \tag{2}$$

The coverage could be defined as the capacity of predicting from a metric applied to a specific RS. In short, it calculates the percentage of situations in which at least one $k$-neighbor of each active user can rate an item that has not been rated yet by that active user. We defined $K_{u,i}$ as the set of neighbors of $u$ which have rated the item $i$. We define the coverage of the system as the average of the user's coverage:

Let

$$C_u = \{i \in I | r_{u,i} = \bullet \wedge K_{u,i} \neq \varnothing\}, \quad D_u = \{i \in I | r_{u,i} = \bullet\}$$

$$coverage = \frac{1}{\#U}\sum_{u\in U}\left(100 \times \frac{\#C_u}{\#D_u}\right) \tag{3}$$

### 4.2. Quality of the set of recommendations: precision, recall and F1

The confidence of users for a certain RS does not depend directly on the accuracy for the set of possible predictions. A user gains confidence on the RS when this user agrees with a reduced set of recommendations made by the RS.

In this section, we define the following three most widely used recommendation quality measures: (1) precision, which indicates the proportion of relevant recommended items from the total number of recommended items, (2) recall, which indicates the proportion of relevant recommended items from the number of relevant items, and (3) F1, which is a combination of precision and recall.

Let $X_u$ as the set of recommendations to user $u$, and $Z_u$ as the set of $n$ recommendations to user $u$. We will represent the evaluation precision, recall and $F1$ measures for recommendations obtained by making $n$ test recommendations to the user $u$, taking a $\theta$ relevancy threshold. Assuming that all users accept $n$ test recommendations:

$$precision = \frac{1}{\#U}\sum_{u\in U}\frac{\#\{i \in Z_u | r_{u,i} \geqslant \theta\}}{n} \tag{4}$$

$$recall = \frac{1}{\#U}\sum_{u\in U}\frac{\#\{i \in Z_u | r_{u,i} \geqslant \theta\}}{\#\{i \in Z_u | r_{u,i} \geqslant \theta\} + \#\{i \in Z_u^c | r_{u,i} \geqslant \theta\}} \tag{5}$$

$$F1 = \frac{2 \times precision \times recall}{precision + recall} \tag{6}$$

### 4.3. Quality of the list of recommendations: rank measures

When the number $n$ of recommended items is not small, users give greater importance to the first items on the list of recommendations. The mistakes incurred in these items are more serious errors than those in the last items on the list. The ranking measures consider this situation. Among the ranking measures most often used are the following standard information retrieval measures: (a) *half-life* (7) [43], which assumes an exponential decrease in the interest of users as they move away from the recommendations at the top and (b) *discounted cumulative gain* (8) [17], wherein *decay* is logarithmic.

$$\mathrm{HL} = \frac{1}{\#U}\sum_{u\in U}\sum_{i=1}^{N}\frac{\max(r_{u,p_i} - d, 0)}{2^{(i-1)/(\alpha-1)}} \tag{7}$$

$$\mathrm{DCG}^k = \frac{1}{\#U}\sum_{u\in U}\left(r_{u,p_1} + \sum_{i=2}^{k}\frac{r_{u,p_i}}{\log_2(i)}\right) \tag{8}$$

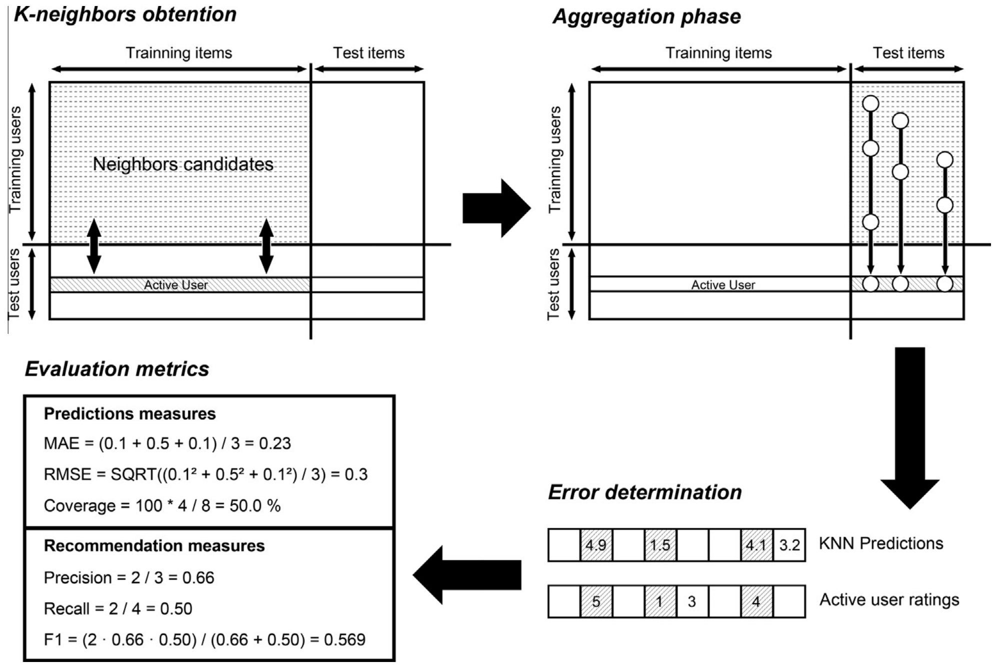


**Fig. 7.** Recommender systems evaluation process.

$p_1, \ldots, p_n$ represents the recommendation list, $r_{u,pi}$ represents the true rating of the user $u$ for the item $p_i$, $k$ is the rank of the evaluated item, $d$ is the default rating, $\alpha$ is the number of the item on the list such that there is a 50% chance the user will review that item.

### 4.4. Novelty and diversity

The novelty evaluation measure indicates the degree of difference between the items recommended to and known by the user. The diversity quality measure indicates the degree of differentiation among recommended items.

Currently, novelty and diversity measures do not have a standard; therefore, different authors propose different metrics [163,220]. Certain authors have [105] used the following:

$$diversity_{Z_u} = \frac{1}{\#Z_u(\#Z_u - 1)} \sum_{i \in Z_u} \sum_{j \in Z_u, j \neq i} [1 - sim(i,j)] \tag{9}$$

$$novelty_i = \frac{1}{\#Z_u - 1} \sum_{j \in Z_u} [1 - sim(i,j)], \quad i \in Z_u \tag{10}$$

Here, $sim(i,j)$ indicates item to item memory-based CF similarity measures. $Z_u$ indicates the set of $n$ recommendations to user $u$.

### 4.5. Stability

The stability in the predictions and recommendations influences on the users' trust towards the RS. A RS is stable if the predicitions it provides do not change strongly over a short period of time. Adomavicius and Zhang [4] propose a quality measure of stability, MAS (Mean Absolute Shift). This measure is defined through a set of known ratings $R_1$ and a set of predictions of all unknown ratings, $P_1$. For an interval of time, users of the RS will have rated a subset $S$ of these unknown ratings and the RS can now make new predictions, $P_2$. MAS is defined as follows:

$$stability = \mathrm{MAS} = \frac{1}{|P_2|} \sum_{(u,i) \in P_2} |P_2(u,i) - P_1(u,i)| \tag{11}$$

### 4.6. Reliability

The reliability of a prediction or a recommendation informs about how seriously we may consider this prediction. When RS recommends an item to a user with prediction 4.5 in a scale $\{1, \ldots, 5\}$, this user hopes to be satisfied by this item. However, this value of prediction (4.5 over 5) does not reflect with which certain degree the RS has concluded that the user will like this item (with value 4.5 over 5). Indeed, this prediction of 4.5 is much more reliable if it has obtained by means of 200 similar users than if it has obtained by only two similar users.

In Hernando et al. [96], a realibility measure is proposed according the usual notion that the more reliable a prediction, the less liable to be wrong. Although this reliability measure is not a quality measure used for comparing different techniques of RS through cross validation, this can be regarded as a quality measure associated to a prediction and a recommendation. In this way, the RS provides a pair of values (prediction value, reliability value), through which users may balance its preference: for example users would probably prefer the option (4,0.9) to the option (4.5,0.1). Consequently, the reliability measure proposed in Hernando et al. [96] provides a new understandable factor, which users may consider for taking its decisions. Nevertheless, the use of this reliability measure is just constrained to those RS based on the $k$NN algorithm.

The definition of reliability on the prediction, $p_{u,i}$, is based on two numeric factors: $s_{u,i}$ and $v_{u,i}$. $s_{u,i}$ measures the similarity of the neighbors used for making the prediction $p_{u,i}$; $v_{u,i}$ measures the degree of disagreement between these neighbors rating the item $i$. Finally, the reliablity measure is defined as follows:

$$f_S(s_{u,i}) = 1 - \frac{\bar{s}}{\bar{s} + s_{u,i}}, \quad s_{u,i} = \sum_{v \in K_{u,i}} sim(u,v) \tag{12}$$

where

$$f_S(s_{u,i}) = 1 - \frac{\bar{s}}{\bar{s} + s_{u,i}}, \quad s_{u,i} = \sum_{v \in K_{u,i}} sim(u,v) \tag{13}$$

Please cite this article in press as: J. Bobadilla et al., Recommender systems survey, Knowl. Based Syst. (2013), http://dx.doi.org/10.1016/j.knosys.2013.03.012

$$f_v(v_{u,i}) = \left(\frac{\max - \min - v_{u,i}}{\max - \min}\right)^{\frac{\ln 0.5}{\ln \frac{\max - \min - \bar{v}}{\max - \min}}}, \quad v_{u,i} = \frac{\sum_{v \in K_{u,i}} sim(u,v)(r_{v,i} - \bar{r}_v - p_{u,i} + \bar{r}_u)^2}{\sum_{v \in K_{u,i}} sim(u,v)} \tag{14}$$

where $\bar{s}$ and $\bar{v}$ are respectively the median of the values of $s_{u,i}$ and $v_{u,i}$ in the specific RS. $K_{u,i}$ is the set of neighbors of $u$ which have rated the item $i$. $\{\min, \ldots, \max\}$ is the discrete range of rating values.

Fig. 7 shows the general mechanism for *cross validation* used to generate quality results form the evaluation measures. The database is divided in training and test areas for both users and items. In the first phase (top on the left side), $k$-neighbors are calculated for the active user (while the active user is selected from the set of test users, the $k$-neighbors are selected from the set of training users). In the aggregation phase (top on the right side), predictions are calculated for the active user (from the set of test items). Finally, evaluation metrics are used to compare the predictions and recommendations obtained with the real ratings of the user; the more accurate the predictions and recommendations, better quality of the proposed recommendation algorithm.

## 5. Social information

As the web 2.0 has developed, RS have increasingly incorporated social information (e.g., trusted and untrusted users, followed and followers, friends lists, posts, blogs, and tags). This new contextual information [145,216] improves the RS. Social information improves the sparsity problem inherent in memory-based RS because social information reinforces traditional memory-based information (users ratings): users connected by a network of trust exhibit significantly higher similarity on items and meta-data that non-connected users [132].

Social information is used by researchers with three primary objectives: (a) to improve the quality of predictions and recommendations [53,13], (b) propose or generate new RS [139,210], and (c) elucidate the most significant relationships between social information and collaborative processes [100,178].

*Trust* and *reputation* is an important area of research in RS [166]; this area is closely related to the social information currently included in RS [114]. The most common approachs to generating trust and reputation measurements are the following: (a) user trust: to calculate the credibility of users through explicit information of the rest of users [239,138] or to calculate the credibility of users through implicit information obtained in a social network [59,150] and (b) item trust: to calculate the reputation of items through a feedback of users [114] or to calculate the reputation of items studying how users work with these items [58,122].

In the social RS field, users can introduce *labels* associated with items. The set of triples $\langle$user, item, *tag*$\rangle$ form information spaces referred to as *folksonomies*. Fundamentally, folksonomies are used in the following two ways: (1) to create *tag recommendation* systems (RS based only on tags) [147] and (2) to enrich the recommendation processes using tags [81].

Content-based filtering has recently become more important due to the surge in social networks. RS show a clear trend to allow users to introduce content [13,178], such as comments, critiques, ratings, opinions and labels as well as to establish social relationship links (e.g., followed, followers, like user and dislike user). This additional information increases the accuracy of predictions and recommendations, which has generated a variety of research articles: Kim et al. [117], Zheng and Li [248] and Carrer-Neto et al. [53].

The rest of this section deal is dealt with the concepts and research in the two lines considered previously: Filtering of social information and content filtering.

### 5.1. Social Filtering

Social information can be gathered explicitly or implicitly through identification of a *community network* or *affinity network* [196] using the individual information that users generate (e.g., communications and web logs) [178]. Even using only the ratings from the users, it is possible to improve the RS results creating an implicit social networking [180]. Both implicit and explicit information sources can be combined to generate recommendations [144].

The explicit social information can be used via a trust-based CF in order to improve the quality of recommendations. Trust information can be generated or used through different approaches, such as trust propagation mechanisms [42], a 'follow the leader' approach [8,186], personality-based similarity measures [101], trust networks [239,221], distrust analysis [223,20], and dynamic trust based on the ant colonies metaphor [20].

Most of the research work that uses social information applied to RS aims to obtain improvements in the recommendations made by referring to the extra information provided by the social information used. Among the most relevant current work which uses this approach we have: Woerndl and Groh [231] use social networks to enhance collaborative filtering; Their evaluation shows that the social recommender outperforms traditional collaborative filtering algorithms in the used scenario. Arazy et al. [13] improve accuracy by using data from online social networks and electronic communication tools. Xin et al. [233] propose an approach for improving RS through exploiting the learners note taking activity. They maintain that notes' features can be exploited by collaborative learning systems in order to enrich and extend the user profile and improve personalized learning. The Bonhard and Sasse [41] research has shown that the relationship between advice-seeker and recommender is extremely important, so ways of indicating social closeness and taste overlap are required. They thus suggest that drawing on similarity and familiarity between the user and the persons who have rated the items can aid judgment and decision making. Fengkun and Hong [75] developed a way to increase recommendation effectiveness by incorporating social network information into CF. They collected data about users' preference ratings and their social network relationships from a social networking web site; then, they evaluated CF performance with diverse neighbor groups combining groups of friends and nearest neighbors. Carmagnola et al. [52] state that joining in a network with other people exposes individuals to social dynamics which can influence their attitudes, behaviors and preferences: They present *SoNARS*, an algorithm for recommending content in social RS. SoNARS targets users as members of social networks, suggesting items that reflect the trend of the network itself, based on its structure and on the influence relationships among users. In Ramaswamy et al. [189] the design of the social network based RS incorporates three features that complement each other to derive highly targeted ads. First, they analyze information such as customer's address books to estimate the level of social affinity among various users. This social affinity information is used to identify the recommendations to be sent to an individual user.

Another group of research work uses social information to create or enable RS. That is, the aim is not to improve the results of a particular RS in operation, the aim is to propose or make possible RS which still do not exist, or if they do exist they are not based on social information: The Siersdorfer and Sergei [210] objective is to construct social recommender systems that predict the utility of items, users, or groups based on the multi-dimensional social environment of a given user; they do a mining of the rich set of structures and social relationships that provides the folksonomies. In the Li and Chen [137] study they propose a blog recommendation mechanism that combines trust model, social relation and

**Table 3**
State of the art on trust and reputation.

| | User trust | Item trust |
|---|---|---|
| Explicit trust systems | The 'credibility' of users is calculated through explicit information of the rest of users. [71,239,240]. Services P2P usually implement this technique [138] | The 'reputation' of items is calculated by means of a feedback of users who are asked about their opinions [114]. E-commerce services often use this technique |
| Implicit trust systems | The 'credibility' of users is calculated through implicit information obtained in a social network [59,150,200] | The 'reputation' of items is calculated studying how users work with these items (for example, the number of times a song is played) [58,122] |
| Memory based trust | The 'credibility' measure is calculated taking into account the users' ratings [112,127,145] | |

semantic analysis and illustrates how it can be applied to a prestigious online blogging system. In the Jason [111] research project, they have applied a system to discover the social networks between mobile users by collecting a dataset from about two millions of users. They argue that social network is applicable to generate context-based recommendation services. Jyun and Chui [115] paper uses trading relationships to calculate level of recommendation for trusted online auction sellers. They demonstrate that network structures formed by transactional histories can be used to expose such underlying opportunistic collusive seller behaviors. In Dell'amico and Capra [69] users' trustworthiness has been measured according to one of the following two criteria: taste similarity (i.e., "I trust those who agree with me"), or social ties (i.e., "I trust my friends, and the people that my friends trust"). They argue that, in order to be trusted, users must be both well intentioned and competent. Based on this observation, they propose a novel approach that they call social filtering.

A third group of work provides the foundation of the research to discover the most significant relationships between social information and collaborative processes, without creating, proposing or improving any particular RS. This research moves at a higher level of abstraction, with the aim of establishing bases and general principles. Bonhard [40] paper explains that qualitative research conducted to date has shown that the relationship between recommender and recommendee has a significant impact on decision-making. Hossain and Fazio [100] present a study exploring the connection between social networks and collaborative process. They focus on exploring academics' network position and its effect on their collaborative networks. By defining network position in this way, they develop a social network that uses the academics as nodes within the network instead of each published paper. The Esslimani et al. [72] paper presents a new CF approach based on a behavioral network that uses navigational patterns to model relationships between users and exploits social networks techniques. Golbeck and Kuter [86] present an experimental study of several types of trust inference algorithms to answer the following questions on trust and change: How far does a single change propagate through the network? How large is the impact of that change? How does this relate to the type of inference algorithm? The experimental results provide insights into which algorithms are most suitable for certain applications.

Research in the field of trust and reputation could provide a suitable starting point to create social interaction among users of the RS, however, the most relevant work on the subject is limited to the use of trust relationships to improve the quality of the recommendation services. O'donovan [165] book chapter examines the diversity of sources from which trust information can be harnessed within social web applications and discusses a high level classification of those sources. It is shown that harnessing an increased amount of information upon which to make trust decisions greatly enhances the user experience with the social web application. Massa and Avesani [151] explain that RS making use of trust information are the most effective in term of accuracy while preserving a good coverage. This is especially evident on users who provided few ratings. Yuan et al. [239] choose the trust aware RS as an example to demonstrate the advantages by making use of the verified small-world nature of the trust network. Li and Kao [138] present a RS based on the trust of social networks; Through the trust computing, the quality and the veracity of peer production services can be appropriately assessed. The experimental results show that the proposed RS can significantly enhance the quality of peer production services.

Table 3 classifies the current approaches to address user credibility and item reputation in social-based RS.

In the CF field, the trust of users is used to make predictions, weighting trust values. That is to say, the more trust a user has, the more important its ratings are for making predictions [58,112,239]. In Ma et al. [145], they propose a probabilistic factor analysis framework, combining ratings and trusted friends; this framework can be applied to pure user-item rating matrix.

### 5.2. Content-based filtering

Content-based filtering (CBF) tries to recommend items to the active user similar to those rated positively in the past. It is based on the concept that items with similar attributes will be rated similarly [16,177,203]. For example, if a user likes a web page with the words "car", "engine" and "gasoline", the CBF will recommend pages related to the automotive world.

CBF is becoming especially important as RS incorporate information on items from users working in web 2.0 environments, such as tags, posts, opinions and multimedia material.

Two challenging problems for content-based filtering are *limited content analysis* and *overspecialization* [3]. The first problem arises from the difficulty in extracting reliable automated information from various content (e.g., images, video, audio and text), which can greatly reduce the quality of recommendations. The second problem (overspecialization) refers to the phenomenon in which users only receive recommendations for items that are very similar to items they liked or preferred; therefore, the users are not receiving recommendations for items that they might like but are unknown (e.g., when a user only receives recommendations about fiction films). Recommendations can be evaluated for *novelty* [32,105].

For CBF to operate, attributes of the items you wish to recommend must be extracted [176]. Typically, a set of attributes is manually defined for each item depending on its domain. In certain instances, such as when it is desired to recommend textual information, classic information retrieval techniques must be used to automatically define such attributes (e.g., *term frequency, inverse document frequency* and*normalization* to page length).

Fig. 8 shows the CBF mechanism, which includes the following steps: (1) extract the attributes of items for recommendation, (2) compare the attributes of items with the preferences of the active user, and (3) recommend items with characteristics that fit the user's interests.

When the attributes of the items and the user profiles are known, the key purpose for CBF [158] is to determine whether a user will like a specific item. This task is resolved traditionally by using heuristic methods [198,15,79] or *classification algorithms*, such us: *rule induction* [65,119], *nearest neighbors* methods [236,27], *Rocchio's algorithm* [131,16], *linear classifiers* [113], and *probabilistic methods* [175,160,84].

The pure CBF has several shortcomings [16,176,212]:

(a) In certain domains (e.g., music, blogs, and videos), it is a complicated task to generate the attributes for items.
(b) CBF suffers from an overspecialization problem because by nature it tends to recommend the same types of items.
(c) It is more difficult to acquire feedback from users because with CBF, users do not typically rate the items (as in CF), and, therefore, it is not possible to determine whether the recommendation is correct.

Because of these shortcomings, it is rare to find a pure CBF implementation. It is more common to use the hybrid CBF/CF Burke 2002. CF solves CBF's problems because it can function in any domain; it is less affected by overspecialization; and it acquires feedback from users. CBF adds the following qualities to CF: improvement to the quality of the predictions, because they are calculated with more information, and reduced impact from the cold-start and sparsity problems.

CBF and CF can be combined in different ways [3]. Fig. 9 shows the different alternatives.

Fig. 9a shows the methods that calculate CBF and CF recommendations separately and subsequently combine them. Claypool et al. [64] propose to use a weighted average for combining CBF and CF predictions depending on the type of prediction. In another study, Pazzani [177] proposes combining the CBF and CF recommendation lists by assigning the items scores according to their position on the lists. Additionally, Billsus and Pazzani [26] and Tran and Cohen [218] propose to select the CBF or CF prediction in accordance with the quality.

Fig. 9b depicts the methods that incorporate CBF characteristics into the CF approach. Balabanovic and Shoham [16] maintain user profiles based on content analysis and directly compare the profiles to determine similar users for CF recommendations. Good et al. [89] construct specialized filterbots using CBF techniques, which later act as neighbors in the CF stage. Melville et al. [157] propose to add predictions from the CBF into the ratting matrix employed by the CF. Li [136] modifies the ratting matrix, which is input for the CF, by combining it with another matrix generated from clustering the items according to their attributes. In Hu and Pu [101], authors incorporate personality characteristics in the CF similarity measure to minimize the new-user problem.

Fig. 9c illustrates the methods to construct a unified model with both CBF and CF characteristics. Basu et al. [19] propose using CBF and CF characteristics in a single rule-based classifier. Popescul et al. [182] and Schein et al. [204] propose using probability models to combine CBF and CF recommendations. In another studies [66,10,50], the authors employ Bayesian networks to combine CBF and CF characteristics and generate more accurate recommendations. Burke [45] and Middleton et al. [159] propose using knowledge-based techniques to solve the cold-start problem.

Fig. 9d shows the methods that incorporate CF characteristics into a CBF approach. In Soboroff and Nicholas [211], the authors use LSI to create the user profiles used in CBF recommendations beginning with the CF ratting matrix. Mooney and Roy [160] use CF system predictions as input for CBF.

The current trend in CBF is to add social information to the items attributes, such as tags, comments, opinion, and social network sharing. *Social tagging* systems are the most popular because they allow users to annotate online resources with arbitrary labels, which produces rich information spaces (*folksonomies*). These new components have opened novel lines of RS research that can be divided into two categories: (1) *tag recommendation* systems and (2) use of tags in the recommendation process:

(1) RS tags attempt to provide personalized item recommendations to users through the most representative tags. In Jächke et al. [110], the authors compare different mechanisms for tags recommendations. Marinho and Schmidt-Thieme [147] improve tags recommendations by applying classic recommendation methods. Additionally, Landia and Anand [130] propose a method that combines clustering-based CBF with CF to suggest new tags to users.
(2) The methods using tags in the recommendation process increase the capacity of traditional RS. Tso-Sutter et al. [219] propose a generic method that allows tags to be incorporated to standard CF algorithms. Bogers and Van Den Bosh [39] examine how to incorporate the tags and other metadata into a hybrid CBF/CF algorithm by replacing the traditional user-based and item-based similarity measures by tag overlap. Gemmell et al. [83] propose a weighted hybrid recommender, wherein they combine the graph-based tag recommendations with user-based CF and item-based CF. Gedikli and Jannach [81] propose to use tags as a means to express which features of an item users particularly like or dislike. In Gemmell et al. [82], the authors offer a hybrid RS, wherein they predict the user preferences for items by only consulting the user's tagging history.

## 6. Additional recommender systems objectives

Commercial RS compete in the market by offering the best content and quality in recommendations as well as greatest variety of services. Recommendations to user groups [108] facilitate joint recommendations to user groups (e.g., a group of four friends who wish to choose a movie). For CF, four design approaches offer an opportunity for action: (1) acting into the similarity measures stage [168], (2) acquiring neighbors [37], (3) acquiring predictions [63], and (4) generating recommendations [17]. Research results [168] indicate that the quality of the recommendations does not vary greatly between the different approaches, but the execution time is dramatically reduced as we advance when it is used (when the design of a similarity measure for groups is the most efficient solution).

For the RS generated recommendations to be valuable for users, they must be explained well in a simple, compelling and accurate manner. The *recommendation explanation* field has been investigated with new developments in RS [91] until now [170]. Traditionally, the explanation type is divided into the following categories: (a) *human style* (user to user approach), (b) *item style* (item to item approach), (c) *feature style*(items features), and (d) *hybrid*. It also employs the use of *conversational techniques* [155] and incorporates *geo-social information* [235].

### 6.1. Recommending to groups of users

RS that consider groups of users [108] are starting to expand and to be used in different areas: tourism [14], music [55], TV [238], web [176].

Given the specific characteristics of the recommendation to groups, it is appropriate to establish a consensus for different

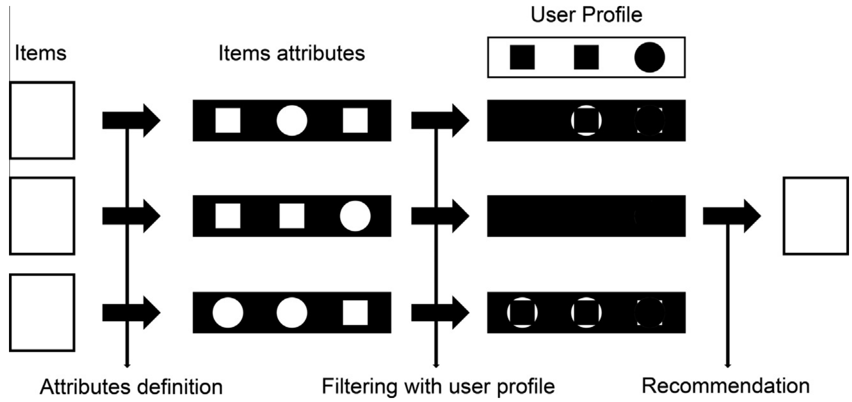


**Fig. 8.** Content-based filtering mechanism.

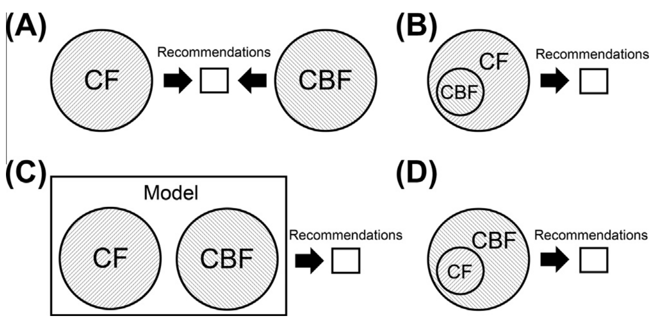


**Fig. 9.** Different alternatives for combining CF and CBF.

group semantics that formalize the agreements and disagreements among users [195].

With the aim of presenting the work carried out to date in a structured way, we provide a classification of the recommendation to groups in CF RS. Fig. 10 graphically illustrates the four basic levels on which we can act in order to unify the group's users' data with the objective of obtaining the data of the group of users: similarity metric, establishing the neighborhood, prediction phase, determination of recommended items.

In Fig. 10, the individual members of a group are represented on the left, in grey; each graticule represents the matrix of ratings by the users (horizontal) on the items (vertical). The graph shows the four representative cases of tackling the solution to recommendation by groups (one case for each matrix on the left of the figure). The circles show key information: they indicate the CF process phase where the unification is performed: "$n$ users → 1 group".

In the first case, at the top of the graph, the data unification is performed in the prediction phase of the CF process: $n$ individual predictions of $n$ users of the group are combined in one prediction of the group (predictions aggregation). This approach has been used by Berkovsky and Freyne [22], García et al. [78] and Christensen and Schiaffino [63].

The second case acts on the sets of neighbors of the group's users, by unifying them in one neighborhood for the whole group. This approach has been studied by Bobadilla et al. [37], proposing the intersection of a large number ($k$) of neighbors of each user of the group.

In the third case, the recommendations obtained for each individual user of the group are merged into one recommendation for the group. Baltrunas et al. [17] use rank aggregation of individual lists of recommendations.

The fourth case [168] uses a similarity metric that acts directly on the set of ratings of the group of users. This solution is the only one that directly provides a set of neighbors for the group of users.

A study exists [9] which, prior to any of the previous cases, proposes, as a front-end, the incorporation of a process of estimation of missing information when dealing with incomplete fuzzy linguistic preference relations.

### 6.2. Explaining recommendations

An important research subject in the RS field focuses on providing *explanations* that justify the recommendations the user has received. This is an important aspect of an RS because it aids in maintaining a higher degree of user confidence in the results generated by the system.

The type of explanations used thus far can be classified as follows [170].

*Human style explanations* (user to user approach). For example, we recommend movie $i$ because it was liked by the users who rated movies $j$, $k$, $m$, ... very positively ($j$, $k$, $m$, ... are movies rated well by the active user).

*Item style explanations* (item to item approach). For example, we recommend the vacation destination $i$ because you liked the vacation destinations $g$, $c$, $r$, ... ($g$, $c$, $r$, ... are vacation destinations similar to $i$ and rated well by the active user).

*Feature style explanations* (it is recommended based on items' features). For example, we recommend movie $i$ because it was directed by director $d$, it features actors $a$, $b$, and it belongs to genre $g$ ($d$, $a$, $b$, $g$ are features the active user is interested in).

*Hybrid methods*. This category primarily includes the following: human/item, human/feature, feature/item, and human/feature/item.

Additionally, in geo-social RS (Foursquare, Google latitude, etc.), location information exists that must be used in the recommendation explanation mechanism [235]. *Geo-social* RS typically adopt a hybrid human/item explanation method based on social, location and memory-based information.

A reference publication that is a helpful introduction to the RS explanations research field has been published previously [91]. They explore the utility of explanations in CF RS, and they stated three key research questions: (1) What models and techniques are effective in supporting explanations? (2) Can explanation facilities increase the acceptance of CF RS? (3) Can explanation facilities increase the filtering performance of the CF RS users? To answer to the first question, they propose using rating histograms, indications of past performance, comparisons to similar rated items, and use of domain specific content features. The results from the experiments conducted with RS users support an affirmative response to the second question. The third question is unanswered because users perform filtering based on many different channels of input.

A dynamic approach that favors the mechanisms for RS explanations includes using conversational techniques, such as the CCBR (conversational case-base reasoning), explained into McSherry [155]. As CCBR they use an incremental nearest neighbor process based on the Pareto case dominance approach. In a different study [153], a dynamic approach is also adopted, but it employs a different perspective. Instead of attempting to justify a particular recommendation they focus on how explanations can help users to understand the recommendation opportunities that remain if the current recommendation should not meet their requirements. They generate compound critiques as explanations: Users have the opportunity to accept or critique recommendations. If they critique a recommendation, the critique acts as a filter over the remaining recommendations.

In a separate study [24], authors differentiate between the concepts *promotion* (increasing of the acceptance of the recommended item) and *satisfaction* (user satisfaction with the recommended item). They also produced better results by using the *keyword style explanation* (based on content data) compared with the *neighbor style explanation* (human style explanation). Authors propose a new classification of the recommendation justifications: Keyword Style Explanation (for content-based RS), Neighbor Style Explanation (for collaborative filtering RS) and Influence Style Explanation (tells the user how their interactions with the RS influences the recommendation). Tintarev and Masthoff [217] describe the advantages of making justifications in recommendations: transparency, scutability, trustworthiness, effectiveness, persuasiveness, efficiency and satisfaction.

Billus and Pazzani [25] propose a recommendation system on news, which provides keyword style justifications of the recommendations through the weights used for obtaining these recommendations. Wang et al. [226] describe a system of justifications based on the features of users' preference. Tintarnev and Masthoff [217] design a recommedation system on films whose recommendations are justified through the features. Vig et al. [222] propose a mechanism for justifying recommendations called tagsplanations, which is based on community tags. Trangsplanations have two key components: tag relevance, the degree to which a tag describes an item; and tag preference, the user's sentiment toward a tag.

Fahri [73] provides a framework for organizing justifications, used to categorize explanations; they propose the categorization of the discourse: explicative, theoretical, pragmatic, ethical, moral, legal, aesthetic, and personal. Although this theoretical framework has not been used into the research literature, it can be used to design new types of explanations. Hernando et al. [97] present a novel explanation technique based on the visualization of trees of items; these trees provide valuable information about the reliability of recommendations and the importance of the ratings the user has made.

The most relevant investigations that produce justifications in recommender systems include a study [187] wherein the authors design a new organization interface where results are grouped according to their tradeoff properties. They have developed a trust model for recommender agents based on the Pareto algorithm (excluding dominated categories). Symeonidis et al. [213] first construct a feature profile for the users to reveal their favorite features, later they group users into biclusters to exploit partial matching between de preferences of groups of users over groups of items. Additionally they propose a metric to measure the quality of justifications: the explain coverage ratio. In Symeonidis et al. [214] they use a prototype "MoviExplain" to put into the test the research showed into Symeonidis et al. [213]. In Hu et al. [102] they use implicit feedback to derive an estimate of the user preference (like or dislike an item) and user confidence for each user-item pair.

## 7. Recommender systems trends

From the evolution of existing RS and research papers in the field, there is a clear tendency to collect and integrate more and different types of data. This trend is parallel to the evolution of the web, which we can define through the following three primary stages: (1) at the genesis of the web, RS used only the explicit ratings from users as well as their demographic information and content-based information included by the RS owners. (2) For the web 2.0, in addition to the above information, RS collect and use social

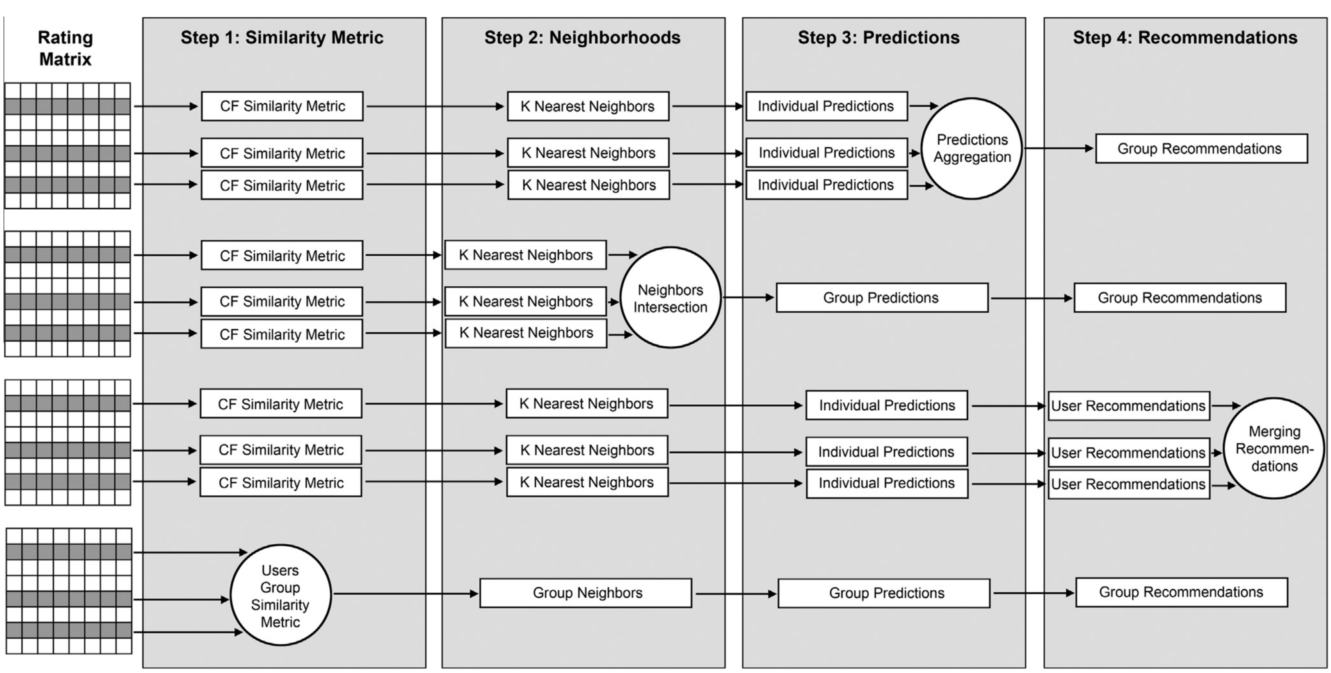


**Fig. 10.** Classification of the recommendations to groups in CF RS. The figure represents the four representative cases for approaching the solution to group recommendations.

information, such as friends, followers, followed, both trusted and untrusted. Simultaneously, users aid in the collaborative inclusion of such information: blogs, tags, comments, photos and videos. (3) For the web 3.0 and the Internet of things, context-aware information from a variety of devices and sensors will be incorporated with the above information. Currently, geographic information is included, and the expected trend is gradual incorporation of information, such as *radio frequency identification* (RFID) data, *surveillance*data, on-line *health parameters* and food and shopping habits, as well as teleoperation and telepresence.

*Context-aware recommender systems* [5,1], focus on additional contextual information, such as time, location, and wireless sensor networks [80]. The contextual information can be obtained explicitly, implicitly, using data mining or with a mixture of these methods (hybrid). Currently, mobile applications increasingly use geographic information; this information enables *geographic* RS that can be considered as *location-aware* RS. For geographic RS [167,152], recommendations are typically generated by considering the geographical position of the user that receives the recommendation.

This section provides an introduction of concepts, which are gaining popularity in the RS research field: Internet of things, privacy preservation, shilling attacks, new frameworks, etc. In this introduction, we provide a novel classification for analyzing these RS concepts. Next, we will deal with the research on the location-aware RS, which may be regarded as the first steps for future RS based on Web 3.0. Finally, we will describe the most significative results on a promising research field: the RS based on bio-inspired models.

### 7.1. Introduction

There is a clear trend towards collection of implicit information instead of a traditional explicit evaluation of items by ratings. Last.Fm is a good example of this situation; the user ratings are inferred by the number of times they have heard each song. The same can be applied in a number of everyday situations, such as for access to web addresses, use of various public transport systems, food purchased, access to sports facilities and access to learning resources.

Incorporation of implicit information on the daily habits of users allows RS to use a variety of data; these data will be used in future CF processes, which are increasingly useful and accurate. Privacy and security considerations will be increasingly important with the widespread trend in using, with consent, devices and sensors for the *Internet of things.*

*Privacy* is an important issue for RS [23] because the systems contain information on large numbers of registered users. For *privacy preservation* in RS, a certain level of uncertainty must be introduced into the predictions [156], primarily through tradeoffs between accuracy and privacy [146]. Furthermore, privacy can be preserved when different RS companies share information (combining their data) [116,242]. Privacy becomes more important as RS increasingly incorporate social information.

Because RS are often used in *electronic commerce*, unscrupulous producers may find profitable to shill RS by lying to the systems in order to have their products recommended more often than those of their competitors. RS can experience *shilling attacks* [128,57], which generate many positive ratings for a product, while products from competitors receive negative ratings. RS are still highly vulnerable to such attacks [191].

Knowledge-based filtering is emerging as an important field of RS. Knowledge RS [46] ‘‘use knowledge about users and products to pursue a knowledge-based approach to generating recommendations, reasoning about what products meet the user's requeriments’’. Recommendations are based on inferences about users needs and preferences. User models are based on knowledge structures such as querys (preferred features por products) [109], cases (case-based reasoning) [44], constraints (constraint-based reasoning) [74], ontologies [159], matching metrics and knowledge vectors [194], and social knowledge [53].

Workflow is a current knowledge field where the user model is based on ‘‘users-roles-tasks reference information that describes which member plays which roles or fulfills which tasks’’ [245,246]. Peer-to-peer (P2P) networks are other current knowledge field, where user information is based on the distributed information existing from each peer and the set of peers who may need her [247].

Gradual incorporation of different types of information (e.g., explicit ratings, social relations, user contents, locations, use trends, knowledge-based information) has forced RS to use hybrid approaches. Once the memory-based, social and location-aware methods and algorithms are consolidated, the evolution of RS demonstrates a clear trend toward combining existing collaborative methods.

The latest research in the CF field has generated only modest improvements for predictions and recommendations from a single type of information (e.g., when the only information used is user ratings, information from social relations, or item content). The results improve further when several algorithms are combined with their respective data types. A growing number of publications address hybrid approaches that use current databases to simultaneously incorporate memory-based, social and content-based information.

To unify the above concepts, Fig. 11 provides an original taxonomy for RS. The taxonomy is classified depending on the nature of the data rather than according to the methods and algorithms used. The core of the taxonomy focuses on data classification by three factors: (1) the target of the data: user or item; (2) mode of acquisition: explicit (i.e., ratings to items made by users) or implicit (e.g., number of times a user has heard a song); and (3) information level: memory, content or social context.

Fig. 11 shows the recommender methods and algorithms (labeled as ‘‘collaborative filtering algorithms’’). Depending on the information type in each RS database, it adopts a hybrid filtering approach. Each hybrid approach will use an appropriate subset of algorithms to consider processing of existing information in a coordinated manner. Future developments will include different recommendation frameworks that address the most common situations. These frameworks allow RS to incorporate the CF kernel with the most appropriate recommendations methods based on the available information in a simple and straightforward manner.

At higher levels (prediction and recommendation), Fig. 11 incorporates current evaluation quality measures, such as those for diversity and novelty. The importance of such measures, and measures developed in the future will grow as users demand novel, stable and less predictable recommendations.

### 7.2. Location-aware recommender systems

Due to the increasing use of mobile devices, *location-aware* systems are becoming more widespread. These systems show a tendency towards their consolidation as web 3.0 services and this naturally leads to location-aware CF and location-aware RS, which can be called geographic CF and geographic RS.

We introduce a classification for geographic CF RS and focus on the most relevant section of the classification obtained. Table 4 establishes the different possibilities of tackling a geographic RS according to the nature of the ratings made (‘‘rating stage’’) and the recommendation process followed (‘‘recommendation stage’’). ‘‘User’’ indicates that the rating and/or recommendation are made without having or using the user's *Geographic Information* (GI).

Similarly, "Item" indicates that the rating and/or recommendation are made without having or using the item's GI. In the cases labeled as "User$^g$" and "Item$^g$" the GI is used.

The cases identified are:

- RS: Traditional RS, in which ratings and recommendations are made without using geographical information.
- RS + G: Traditional RS, which also contributes the item's geographical position. These RS cannot be regarded as geographic RS, as the GI does not play a part in the recommendation process.
- GRS: This group of Geographic RS is most likely to become popular in the near future. In these, ratings are made in a traditional way, whilst recommendations are made by considering the geographical position of the user to whom the recommendation is to be made. A representative example is that of a RS for restaurants; the users rate a restaurant using very diverse concepts, which do not include the distance at the time of voting between the user and the restaurant. However, users of a Geographic RS expects a restaurant to be recommended to them not only because of good ratings from similar users (*k*-neighbors), but also according to the distance between their current position and that of the restaurant. Other possible examples are RS for cinemas, pubs, supermarkets, cultural activities in a city, language learning centers, gyms and sports clubs, etc.
- GRS$^+$: In this case, users establish ratings on items by weighting the distance between them and the items rated. In this type of geographic RS two possibilities can be established:
  1. Hybrid CF/Demographic filtering: Each item accepts a maximum of one vote per user, to which the geographical position from which it has been issued is associated.
  2. Geographic RS where each item accepts more than one rating for each user, depending on the geographical position from which each rating is made.
  3. The hybrid RS in case 1 respond to regional or national geographical approaches, in which recommendations can be established according to weighting between the similarity of the votes (CF) and their origin. This type of GRS may be regarded as an extended case of hybrid CF/demographic filtering, in which the GI is given for each vote instead of for each user.

From a theoretical point of view, Type 2 GRS$^+$ are the most complete; however, from a practical point of view, they involve a semantic difficulty in the item rating process, which makes their use very difficult. Rating items in this GRS$^+$ involves that each user can rate items according to the relative distances between the user and the items. In this way, a user can rate a restaurant from their home differently to how they would rate it from their workplace; and when the distances are very different, the ratings are also likely to be so. The mental process would be something like this: I am 1 km from the restaurant and I rate very positively travelling 1 km to go to that restaurant which I think is good; but after some time, the same user, who is at work, 24 km away from the restaurant, could cast a vote indicating they do not consider it to be positive to travel 24 km to go to the restaurant even if they think it is good.

In summary, GRS$^+$ have the advantage that they accept a wider variety of ratings and that these also contain the relative importance that each user gives to the items according to the distance required to access them. The disadvantage is that it is difficult to involve users in a particularly complex and demanding ratings process.

This subsection focuses on the GRS-type geographic CF RS. At present, there are few publications regarding GI-based RS; This is due, to a great extent, to the lack of public databases that include ratings and geographic positions capable of being combined in an RS. Some of the publications that focus more closely on the field are as follows:

Martinez et al. [149] and Biuk-Aghai et al. [28] are examples of the RS + G group. In Schlieder [205], they propose a novel approach for modeling the collaborative semantics of geographic folksonomies. This approach is based on multi-object tagging, that is, the analysis of tags that users assign to composite objects. This paper is based on the concept of groups of people who share a common geospatial feature data dictionary (including definitions of feature relationships) and a common metadata schema.

Wan-Shiou et al. [225] can be considered as a hybrid content based/geographic RS. The core of the system is a hybrid content based/geographic recommendation mechanism that analyzes a customer's history and position so that vendor information can be ranked according to the match with the preferences of a customer.

Matyas and Schlieder [152] show a collaborative system that we could situate between a RS and a GRS. In this case, the users' ratings are taken based on the photos they have downloaded from a Web 2.0 and the photos they have uploaded to the same Web (the photos have a GPS address associated to them). After this, a search of *k*-neighborhoods based on this data is carried out. The recommendation process does not take into account the user's position.

It is possible to collect travel GPS traces from users and use the database to generate recommendations [249]. The travel GPS traces can be reinforced with social information based on friends [250]. Both papers can be classified as GRS$^+$.

### 7.3. Bio-inspired approaches

Much of the proposed model-based RS are based on bio-inspired approaches, which primarily use *Genetic Algorithms* (GAs) and *Neural Networks* (NNs). Models have also been proposed based on *Artificial Immune Networks* (AINs).

GA are heuristic approaches based on evolutionary principles such as natural selection and survival of the fitest. GA have mainly been used in two aspects of RS: clustering [120,243] and hybrid user models [76,99,7]. A common technique to improve the features of RS consists of initially carrying out a clustering on all of the users, in such a way that a group of classes of similar users is obtained, after this, the desired CF techniques can be applied to each of the clusters, obtaining similar results but in much shorter calculation times; It is usual to use common genetic clustering algorithms such as GA-based *K*-means [121].

The RS hybrid user models commonly use a combination of CF with demographic filtering or CF with content based filtering, to exploit merits of each one of these techniques. In these cases, the chromosome structure can easily contain the demographic characteristics and/or those related to content-based filtering.

In order to tackle location-based advertisement, Dao et al. [68] propose a model-based CF using GA. They combine both user's preferences and interaction context. Bobadilla et al. [33] use GA to create a similarity metric, weighting a set of very simple similarity measures. Hwang et al. [106] employ a GA to learn personal preferences of customers.

NN is a model based on the observed behavior of biological neurons. This model, intended to simulate the way the brain processes information, enables the computer to "learn" to a certain degree. A NN typically consists of a number of interconnected nodes. Each handles a designated sphere of knowledge, and has several inputs from the network. Based on the inputs it gets, a node can "learn" about the relationships between sets of data, pattern, and, based upon operational feedback, are molded into the pattern required to generate the required results.

Please cite this article in press as: J. Bobadilla et al., Recommender systems survey, Knowl. Based Syst. (2013), http://dx.doi.org/10.1016/j.knosys.2013.03.012

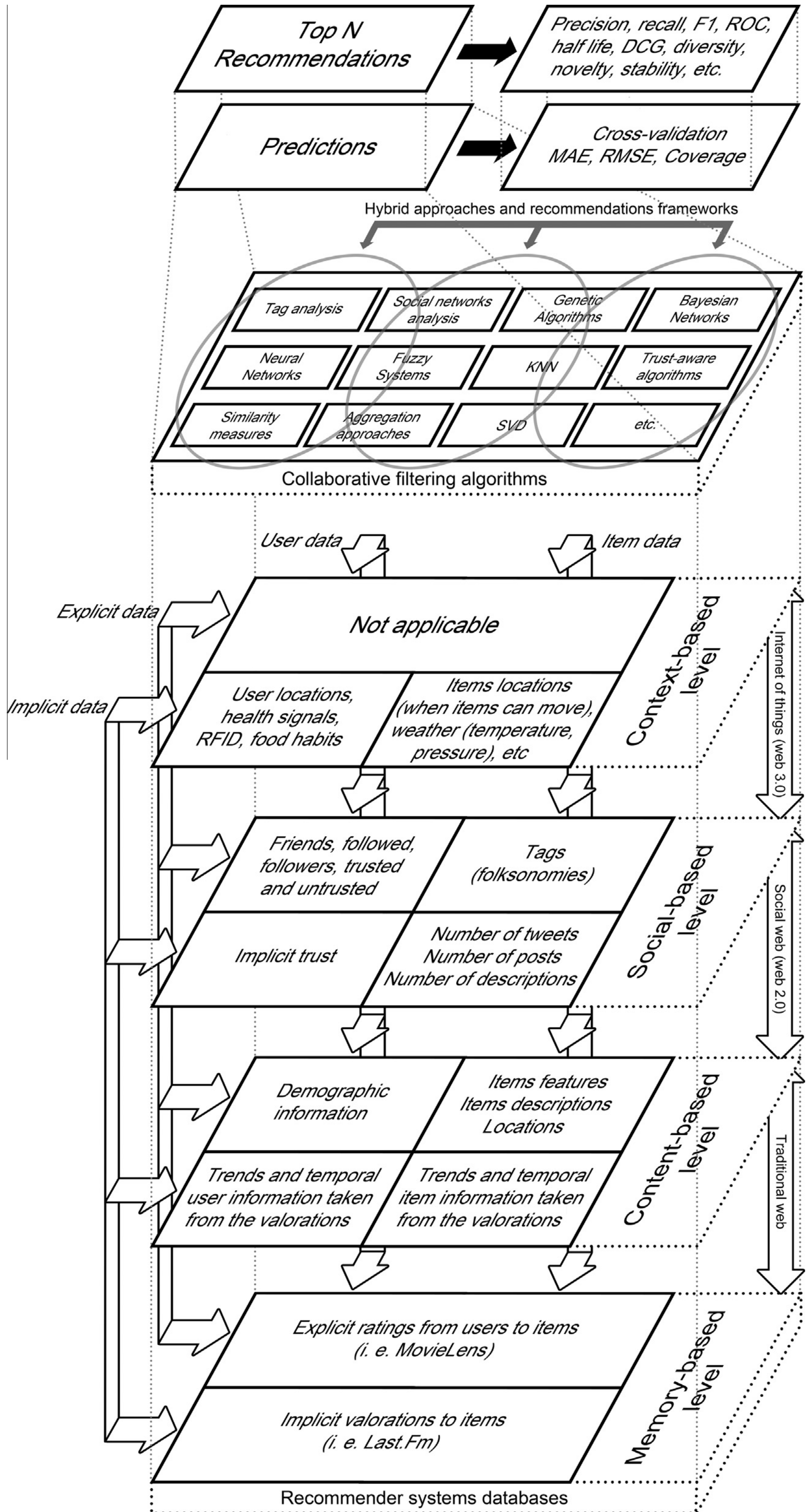


**Fig. 11.** Recommender systems taxonomy.

The RS most relevant research available in which NN usually focuses is hybrid RS, in which NN are used for learn users profiles; NN have also been used in the clustering processes of some RS.

The hybrid approaches enable NN to act on the additional information to the ratings. In Ren et al. [192] they propose a hybrid recommender approach that employs Widrow-Hoff [229] algorithm to learn each user's profile from the contents of rated items. This improves the granularity of the user profiling. In Christakou and Stafylopatis [62] they use a combination of content-based and CF in order to construct a system that provides more precise recommendations concerning movies. In Lee and Woo [133] first, all users are segmented by demographic characteristics and users in

**Table 4**
Geographic collaborative filtering recommender systems classification.

| | Rating stage | | Recommendation stage | | User GI |
|---|---|---|---|---|---|
| | Item | Item$^g$ | Item | Item$^g$ | |
| User | RS/GRS | – | RS | RS + G | Not |
| User$^g$ | – | GRS$^+$ | – | GRS/GRS$^+$ | Yes |
| Item GI | Not | Yes | Not | Yes | |

each segment are clustered according to the preference of items using the Self-Organizing Map (SOM) NN. Kohonon's SOMs are a type of unsupervised learning; their goal is to discover some underlying structure of the data.

Two alternative NN uses are presented in Huang et al. [103] and Roh et al. [193]. In the first case paper, authors use a training back-propagation NN for generating association rules that are mined from a transactional database; in the second paper, authors propose a model that combines a CF algorithm with two machine learning processes: SOM and Case Based Reasoning (CBR) by changing an unsupervised clustering problem into a supervised user preference reasoning problem.

Neuro-fuzzy inference has been used in Sevarac et al. [207] to create pedagogical rules in e-learning. A new cold-start similarity measure has been perfected in Bobadilla et al. [36] using optimization based on neural learning.

Artificial immune systems are distributed and adaptive systems using the models and principles derived form the human immune system. They model the defence system which can protect our body against infections. In order to tackle the RS sparsity problem and to make algorithms more scalable, Acilar and Arslan [2] present a new CF model based on the AIN Algorithm (aiNet). AIN were previously proposed to general recommendations [49] and to recommend web sites [161].

## 8. Related works and original contributions of the paper

As CF has become more complex, different survey papers have been published in this area. Schafer et al. [203] introduces the core concepts of CF: the theory and practice, the rating systems and their acquisition, evaluation, interaction interfaces and privacy issues. Candillier et al. [51] review the main CF filtering methods and compare their results.

Su and Khoshgoftaar [212] presents a survey of CF techniques. Authors introduce the theory on CF and concisely deal with the main challenges: sparsity, scalability, synonymy, gray sheep, shilling attacks, privacy, etc. They also expose an overview table of CF techniques.

Park et al. [171] review 210 papers on RS and classifies them by the year and journal of the publication, their application fields, and their data mining techniques. Additionaly, they categorized the papers into eight application fields (films, music, etc.).

A review in RS algorithms is presented in [141]. This paper focuses on explaining carefully how the most used algorithms in RS work. The paper presents also the basic concepts of CF and their evaluation metrics, dimensionality reduction techniques, diffusion-based methods, social filtering and meta approaches.

Our survey tries to include the most novel issues that have not been dealt carefully in the previous papers. Next, we will stand out the most outstanding features of this survey:

- Uses a methodology for selecting the most suitable papers in the RS, standing out the latest and most cited papers in the area of RS.
- Provides an updated overview table of the most used RS public databases, including tags and friend relations information.
- Studies the cold-start problem inherent to all the RS.
- Presents a novel overview table informing both the classical similarity measure and those which have recently been proposed. It includes both the tailored metrics for cold-start users and the general-purpose metrics. Besides, we show the quality measures obtained when evaluating such metrics.
- Includes the recent quality measurements, beyond accuracy, to evaluate RS: novelty, diversity and stability. Additionaly, we include a reliability measure associated to predictions and recommendations.
- Provides a comprehensive survey on social filtering, presenting a novel overview table on trust, reputation and credibility.
- Introduces the content-based filtering from a modern perspective standing out its application for dealing with social information, such as social tagging.
- Presents a summary of the most relevant contributions in the RS for group of users. We will show a novel classification for the existing methods.
- Deals with a fast growing RS field: the location-aware RS, based on geographic information. This section is estructured with the help of a novel geographic RS classification table.
- Summarizes the most relevant contributions on the use of bio-inspired approaches.
- Describes the RS trends to implicitly collect data (specially those derived from the use of Internet of things).
- Provides an RS taxonomy for classifying the RS through three factors: source of data (traditional web, social web 2.0, Internet of things/web 3.0); target of data (users, items); method for extracting data (explicit, implicit).

## 9. Conclusions

Recommender systems are proving to be a useful tool for addressing a portion of the information overload phenomenon from the Internet. Its evolution has accompanied the evolution of the web. The first generation of recommender systems used traditional websites to collect information from the following three sources: (a) content-based data from purchased or used products, (b) demographic data collected in users' records, and (c) memory-based data collected from users' item preferences. The second generation of recommender systems, extensively use the web 2.0 by gathering social information (e.g., friends, followers, followed, trusted users, untrusted users). The third generation of recommender systems will use the web 3.0 through information provided by the integrated devices on the Internet. The use of location information already incorporated in many recommender systems will be followed by data from devices and sensors, which will be widely used (e.g., real-time health signals, RFID, food habits, online local weather parameters such as temperature and pressure).

The firsts recommender systems were focused on improving recommendation accuracy through filtering. Most memory-based methods and algorithms were developed and optimized in this context (e.g., *k*NN metrics, aggregation approaches, singular value decomposition, diffusion-based methods, etc.). At this stage, hybrid approaches (primarily collaborative–demographic and collaborative–content filtering) improved the quality of the recommendations. In the second stage, algorithms that included social information with previous hybrid approaches were adapted and developed (e.g., trust-aware algorithms, social adaptive approaches, social networks analysis, etc.). Currently, the hybrid ensemble algorithms incorporate location information into existing recommendation algorithms.

Evaluation of the predictions and recommendations has evolved since the origins of recommender systems, which weighted prediction errors (accuracy) heavily. They also recognized the convenience of evaluating the quality of the top $n$ recommendations as a set; evaluation of the top $n$ recommendations as a ranked list was then incorporated. Currently, there is a tendency to assess new evaluation measures, such as diversity and novelty.

Future research will concentrate on advancing the existing methods and algorithms to improve the quality of recommender systems predictions and recommendations. Simultaneously, new lines of research will be developed for fields and aims, such as on: (1) proper combination of existing recommendation methods that use different types of available information, (2) to get the maximum use of the individual potential of various sensors and devices on the Internet of things, (3) acquisition and integration of trends related to the habits, consumption and tastes of individual users in the recommendation process, (4) data mining from RS databases for non-recommendation uses (e.g., market research, general trends, visualization of differential characteristics of demographic groups), (5) enabling security and privacy for recommender systems processes, (6) new evaluation measures and developing a standard for non-standardized evaluation measures, and (7) designing flexible frameworks for automated analysis of heterogeneous data.

## References

[1] S. Abbar, M. Bouzeghoub, S. Lopez, Context-aware recommender systems: a service oriented approach, in: Proceedings of the 3rd International Workshop on Personalized Access, Profile Management and Context Awareness in Databases, 2009.
[2] A.M. Acilar, A. Arslan, A collaborative filtering method based on artificial immune network, Expert Systems with Applications 36 (4) (2008) 8324–8332.
[3] G. Adomavicius, A. Tuzhilin, Toward the next generation of recommender systems: a survey of the state-of-the-art and possible extensions, IEEE Transactions on Knowledge and Data Engineering 17 (6) (2005) 734–749.
[4] G. Adomavicius, J. Zhang, On the stability of recommendations algorithms, in: ACM Conference on Recommender Systems, 2010, pp. 47–54.
[5] G. Adomavicius, A. Tuzhilin, Context-Aware recommender Systems, in: F. Ricci, et al. (Ed.), Recommender Systems Handbook, 2011, pp. 217–253.
[6] H.J. Ahn, A new similarity measure for collaborative filtering to alleviate the new user cold-starting problem, Information Sciences 178 (2008) 37–51.
[7] M.Y.H. Al-Shamri, K.K. Bharadwaj, Fuzzy-genetic approach to recommender systems based on a novel hybrid user model, Expert Systems with Applications 35 (3) (2008) 1386–1399.
[8] J. Al-Sharawneh, M.A. Williams, Credibility-aware Web-based social network recommender: follow the leader, in: Proceedings of the 2010 ACM Conference on Recommender Systems, 2010, pp. 1–8.
[9] S. Alonso, F.J. Cabrerizo, F. Chiclana, F. Herrera, E. Herrera-Viedma, Group decision making with incomplete fuzzy linguistic preference relations, International Journal of Intelligent Systems 24 (2009) 201–222.
[10] A. Ansari, S. Essegaier, R. Kohli, Internet recommendation systems, Journal of Marketing Research 37 (3) (2000) 363–375.
[11] N. Antonopoulus, J. Salter, Cinema screen recommender agent: combining collaborative and content-based filtering, IEEE Intelligent Systems (2006) 35–41.
[12] P. Antunes, V. Herskovic, S.F. Ochoa, J.A. Pino, Structuring dimensions for collaborative systems evaluation, ACM Computing Surveys 44 (2) (2012). Article 8.
[13] O. Arazy, N. Kumar, B. Shapira, Improving Social Recommender Systems, Journal IT Professional 11 (4) (2009) 31–37.
[14] L. Ardissono, A. Goy, G. Petrone, M. Segnan, P. Torasso, INTRIGUE: Personalized recommendation of tourist attractions for desktop and handset devices, Applied Artificial Intelligence 17 (8-9) (2003) 687–714.
[15] R. Baeza-Yates, B. Ribeiro-Neto, Modern Information Retrieval, Addison-Wesley, 1999.
[16] M. Balabanovic, Y. Shoham, Content-based, collaborative recommendation, Communications of the ACM 40 (3) (1997) 66–72.
[17] L. Baltrunas, T. Makcinskas, F. Ricci, Group recommendation with rank aggregation and collaborative filtering, in: Proceedings of the 2010 ACM Conference on Recommender Systems, 2010, pp. 119–126.
[18] A.B. Barragáns-Martınez, E. Costa-Montenegro, J.C. Burguillo, M. Rey-López, F.A. Mikic-Fonte, A. Peleteiro, A hybrid content-based and item-based collaborative filtering approach to recommend TV programs enhanced with singular value decomposition, Information Sciences 180 (22) (2010) 4290–4311.
[19] C. Basu, H. Hirsh, W. Cohen, Recommendation as classification: using social and content-based information in recommendation, in: Proceedings of the Fifteenth National Conference on Artificial Intelligence, 1998, pp. 714–720.
[20] P. Bedi, R. Sharma, Trust based recommender system using ant colony for trust computation, Expert Systems with Applications 39 (1) (2012) 1183–1190.
[21] Y. Bengio, Y. Grandvalet, No umbiased estimator of the variance of k-fold cross-validation, Journal of Machine Learning Research 5 (2004) 1089–1105.
[22] S. Berkovsky, J. Freyne, Group-based recipe recommendations: analysis of data aggregation strategies, in: Proceedings of the 2010 ACM Conference on Recommender Systems, 2010, pp. 111–118.
[23] A. Bilge, H. Polat, An improved privacy-preserving DWT-based collaborative filtering scheme, Experts Systems with Applications 39 (3) (2012) 3654–3841.
[24] M. Bilgic, R. Mooney, Explanation for recommender systems: satisfaction vs. promotion, in: Next Stage of Recommender Systems Research Workshop (IUI conference), 2005, pp. 13–18.
[25] D. Billsus, M. Pazzani, A personal news agent that talks, learns and explains, in: Proc. Auton. Agents Conf., 1999, pp. 268–275.
[26] D. Billsus, M. Pazzani, User modeling for adaptive news access, User Modeling and User-Adapted Interaction 10 (2–3) (2000) 147–180.
[27] D. Billsus, M. Pazzani, J. Chen, A learning agent for wireless news access, in: Proceedings of the International Conference on Intelligent User Interfaces, 2002, pp. 33–36.
[28] R.P. Biuk-Aghai, S. Fong, S. Yain-Whar, Design of a recommender system for mobile tourism multimedia selection, in: 2nd International Conference on Internet Multimedia Services Architecture and Applications (IMSAA), 2008, pp. 1–6.
[29] J. Bobadilla, F. Serradilla, The effect of sparsity on collaborative filtering metrics, in: Australian Database Conference, 2009, pp. 9–17.
[30] J. Bobadilla, F. Serradilla, A. Hernando, Collaborative filtering adapted to recommender systems of e-learning, Knowledge Based Systems 22 (2009) 261–265.
[31] J. Bobadilla, F. Serradilla, J. Bernal, A new collaborative filtering metric that improves the behavior of recommender systems, Knowledge Based Systems 23 (2010) 520–528.
[32] J. Bobadilla, A. Hernando, F. Ortega, J. Bernal, A framework for collaborative filtering recommender systems, Expert Systems with Applications 38 (12) (2011) 14609–14623.
[33] J. Bobadilla, F. Ortega, A. Hernando, J. Alcalá, Improving collaborative filtering recommender systems results and performance using genetic algorithms, Knowledge Based Systems 24 (8) (2011) 1310–1316.
[34] J. Bobadilla, A. Hernando, F. Ortega, A. Gutiérrez, Collaborative filtering based on significances, Information Sciences 185 (1) (2012) 1–17.
[35] J. Bobadilla, F. Ortega, A. Hernando, A collaborative filtering similarity measure based on singularities, Information Processing and Management 48 (2) (2012) 204–217.
[36] J. Bobadilla, F. Ortega, A. Hernando, J. Bernal, A collaborative filtering approach to mitigate the new user cold start problem, Knowledge Based Systems 26 (2012) 225–238.
[37] J. Bobadilla, F. Ortega, A. Hernando, J. Bernal, Generalization of recommender systems: collaborative filtering extended to groups of users and restricted to groups of items, Expert Systems with Applications 39 (2012) 172–186.
[38] J. Bobadilla, F. Ortega, A. Hernando, A. Arroyo, A balanced memory-based collaborative filtering similarity measure, International Journal of Intelligent Systems 27 (10) (2013) 939–946.
[39] T. Bogers, A. Van Den Bosch, Collaborative and content-based filtering for item recommendation on social bookmarking websites, in: Proceedings of the 2009 ACM Conference on Recommender Systems, 2009, pp. 9–16.
[40] P. Bonhard, Who do trust? Combining recommender systems and social networking for better advice, in: International Conference on Intelligent User Interfaces, 2005.
[41] P. Bonhard, M.A. Sasse, 'Knowing me, knowing you'—Using profiles and social networking to improve recommender Systems, BT Technology Journal 24 (3) (2006) 84–98.
[42] P. Borzymek, M. Sydow, A. Wierbicki, Enriching trust prediction model in social network with user rating similarity, in: Proceedings of the 2009 International Conference on Computational Aspects of Social Network, 2009, pp. 40–47.
[43] J.S. Breese, D. Heckerman, C. Kadie, Empirical analysis of predictive algorithms for collaborative filtering, in: 14th Conference on Uncertainty in Artificial Intelligence, 1998, pp. 43–52.
[44] D. Bridge, M.H. Goker, L. McGinty, B. Smyth, Case-based recommender systems, The Knowledge Engineering Review 20 (3) (2005) 315–320.
[45] R. Burke, Encyclopedia of library and information systems, in: A. Kent (Ed.), vol. 69(Suppl. 32), Marcel Dekker, 2000 (Chapter: Knowledge-Based Recommender Systems).
[46] R. Burke, A case-based reasoning approach to collaborative filtering, in: EWCBR 2000, 2000, pp. 370–379.
[47] R. Burke, Hybrid recommender systems: survey and experiments, User Modeling and User-Adapted Interaction 12 (4) (2002) 331–370.
[48] F. Cacheda, V. Carneiro, D. Fernández, V. Formoso, Comparison of collaborative filtering algorithms: limitations of current techniques and proposals for scalable, high-performance recommender Systems, ACM Transactions on the Web 5 (1) (2011). Article 2.
[49] S. Caizer, U. Aickelin, A recommender system based on idiotypic artificial immune networks, Journal of Mathematics, Models and Algorithms 4 (2) (2005) 181–198.

[50] L.M. Campos, J.M. Fernández-Luna, J.F. Huete, M.A. Rueda-Morales, Combining content-based and collaborative recommendations: a hybrid approach based on Bayesian Networks, International Journal of Approximate Reasoning 51 (7) (2010) 785–799.
[51] L. Candillier, F. Meyer, M. Boullé, Comparing state-of-the-art collaborative filtering systems, Lecture Notes in Computer Sciece 4571 (2007) 548–562.
[52] F. Carmagnola, F. Vernero, P. Grillo, SoNARS: a social networks-based algorithm for social recommender systems, in: Proceedings of the 17th International Conference on User Modeling, Adaptation, and Personalization: Formerly UM and AH, 2009, pp. 223–234.
[53] W. Carrer-Neto, M.L. Hernández-Alcaraz, R. Valencia-García, F. Garcıa-Sánchez, Social knowledge-based recommender system, Application to the movies domain. Expert Systems with Applications 39 (12) (2012) 10990–11000.
[54] J.J. Castro-Sanchez, R. Miguel, D. Vallejo, L.M. López-López, A highly adaptive recommender system based on fuzzy logic for B2C e-commerce portals, Expert Systems with Applications 38 (3) (2011) 2441–2454.
[55] D. Chao, J. Balthrop, S. Forrest, Adaptive radio: achieving consensus using negative preferences, in: International ACM SIGGROUP Conference on Supporting Group Work, 2005, pp. 120–123.
[56] T. Chen, L. He, Collaborative filtering based on demographic attribute vector, in: Proceedings of the International Conference on Future Computer and Communication, 2009, pp. 225–229.
[57] P.A. Chirita, W. Nejdl, C. Zamfir, Preventing shilling attacks in online recommender systems, in: Workshop on Web Information and Data Management, 2005, pp. 67–74.
[58] J. Cho, K. Kwon, Y. Park, Q-rater: a collaborative reputation system based on source credibility theory, Expert Systems with Applications 36 (2009) 3751–3760.
[59] S.B. Cho, J.H. Hong, M.H. Park, Location-based recommendation system using Bayesian user's preference model in mobile devices, Lecture Notes in Computer Science 4611 (2007) 1130–1139.
[60] K. Choi, D. Yoo, G. Kim, Y. Suh, A hybrid online-product recommendation system: combining implicit rating-based collaborative filtering and sequential pattern analysis. Electronic Commerce Research and Applications, in press, doi: 10.1016/j.elerap.2012.02.004.
[61] K. Choi, Y. Suh, A new similarity fuction for selecting neighbors for each target item in collaborative filtering, Knowledge Based Systems 37 (2013) 146–153.
[62] C. Christakou, A. Stafylopatis, A hybrid movie recommender system based on neural networks, in: International Conference on Intelligent Systems Design and Applications, 2005, pp. 500–505.
[63] I.A. Christensen, S. Schiaffino, Entertainment recommender systems for group of users, Expert Systems with Applications 38 (2011) 14127–14135.
[64] M. Claypool, A. Gokhale, T. Miranda, P. Murnikov, D. Netes, M. Sartin, Combining content-based and collaborative filters in an online newspaper, in: Proceedings of ACM SIGIR Workshop on Recommender Systems, 1999, pp. 40–48.
[65] W. Cohen, Fast effective rule induction, in: Proceedings of the Twelfth International Conference on Machine Learning, 1995, pp. 115–123.
[66] M. Condliff, D. Lewis, D. Madigan, C. Posse, Bayesian mixed-effects models for recommender systems, in: ACM SIGIR '99 Workshop on Recommender Systems: Algorithms and Evaluation, 1999, pp. 23–30.
[67] E. Costa-Montenegro, A.B. Barragáns-Martínez, M. Rey-López, Which App? A recommender system of applications in markets: implementation of the service for monitoring users' interaction, Expert Systems with Applications 39 (10) (2012) 9367–9375.
[68] T.H. Dao, S.R. Jeong, H. Ahn, A novel recommendation model of location-based advertising: context-aware collaborative filtering using GA approach, Expert Systems with Applications 39 (3) (2012) 3731–3739.
[69] M. Dell'amico, L. Capra, SOFIA: social filtering for robust recommendation, IFIP Advances in Information and Communication Technology 263 (2008) 135–150.
[70] T. Dubois, J. Golbeck, J. Kleint, A. Srinivasan, Improving recommendation accuracy by clustering social networks with trust, in: Proceedings of the 2009 ACM Conference on Recommender Systems, 2009, pp. 1–8.
[71] M. Ekström, H. Björnsson, C. Nass, A reputation mechanism for business-to-business electronic commerce that accounts for rater credibility, Journal of Organizational Computing and Electronic Commerce 15 (1) (2005) 1–18.
[72] I. Esslimani, A. Brun, A. Boyer, From social networks to behavioral networks in recommender systems, in: Proceedings of the 2009 International Conference on Advances in Social Network Analysis and Mining, 2009, pp. 143–148.
[73] Y. Fahri, A Framework for Organizing Justifications for Strategic use in Adaptive Iteraction Contexts, ECIS, 2008. Article 250.
[74] A. Felfernig, R. Burke, Constraint-based recommender systems: technologies and research issues, in: 10th International Conference on Electronic Commerce, 2008 (Article No. 3).
[75] L. Fengkun, J.L. Hong, Use of social network information to enhance collaborative filtering performance, Expert Systems with Applications 37 (7) (2010) 4772–4778.
[76] L.Q. Gao, C. Li, Hybrid personalizad recommended model based on genetic algorithm, in: International Conference on Wireless Communication, Networks and Mobile Computing, 2008, pp. 9215–9218.
[77] M. Gao, Z. Wu, F. Jiang, Userrank for item-based collaborative filtering recommendation, Information Processing Letters 111 (9) (2011) 440–446.
[78] I. García, L. Sebastia, E. Onaindia, On the design of individual and group recommender systems for tourism, Expert Systems with Applications 38 (2011) 7683–7692.
[79] R. García, X. Amatriain, Weighted content based methods for recommending connections in online social networks, in: Proceedings of the 2010 ACM conference on Recommender Systems, 2010, pp. 68–71.
[80] D. Gavalas, M. Kenteris, A web-based pervasive recommendation system for mobile tourist guides, Personal and Ubiquitous Computing 15 (7) (2011) 759–770.
[81] F. Gedikli, D. Jannach, Rating items by rating tags, in: Proceedings of the 2010 ACM Conference on Recommender Systems, 2010, pp. 25–32.
[82] J. Gemmell, T. Schimoler, B. Mobasher, R. Burke, Resource recommendation for social tagging: a multi-channel hybrid approach, in: Proceedings of the 2010 ACM Conference on Recommender Systems, 2010, pp. 60–67.
[83] J. Gemmell, T. Schimoler, M. Ramezani, L. Christiansen, B. Mobasher, Improving FolkRank with item-based collaborative filtering, in: Proceedings of the 2009 ACM conference on Recommender Systems, 2009, pp. 17–24.
[84] M. Gemmis, P. Lops, G. Semeraro, P. Basile, Integrating tags in a semantic content-based recommender, in: Proceedings of the 2008 ACM conference on Recommender Systems, 2008, pp. 163–170.
[85] T. George, S. Meregu, A scalable collaborative filtering Framework base don co-clustering, in: IEEE International Conference on Data Mining (ICDM), 2005, pp. 625–628.
[86] J. Golbeck, U. Kuter, The ripple effect: change in trust and its impact over a social network, in: Computing with Social Trust, Human–Computer Interaction Series, Part II, 2009, pp. 169–181 (Chapter 7).
[87] K. Goldberg, T. Roeder, D. Gupta, C. Perkins, Eigentaste: a constant time collaborative filtering algorithm, Information Retrieval 4 (2) (2001) 133–151.
[88] R. González-Crespo, O. Sanjuán-Martınez, J. Manuel-Cueva, B. Cristina-Pelayo, J.E. Labra-Gayo, P. Ordoñez, Recommendation system based on user interaction data applied to intelligent electronic books, Computers in Human Behavior 27 (4) (2011) 1445–1449.
[89] N. Good, J.B. Schafer, J.A. Konstan, A. Borchers, B. Sarwar, J.L. Herlocker, J. Riedl, in: Proceedings of the Sixteenth National Conference on Artificial Intelligence and the Eleventh Innovative Applications of Artificial Intelligence Conference Innovative Applications of Artificial Intelligence, 1999, pp. 439–446.
[90] A. Gunawardana, G. Shani, A survey of accuracy evaluation metrics of recommender tasks, Journal of Machine Learning Reearch 10 (2009) 2935–2962.
[91] J.L. Herlocker, J.A. Konstan, J. Riedl, Explaining collaborative filtering recommendations, in: ACM Conference on Computer Supported Cooperative Work (CSCW), 2000, pp. 241–250.
[92] J.L. Herlocker, J.A. Konstan, A.L. Borchers, J.T. Riedl, An algorithmic framework for performing collaborative filtering, in: Proceedings of the 22nd Annual International ACM SIGIR Conference on Research and Development in Information Retrieval, 1999, pp. 230–237.
[93] J.L. Herlocker, J.A. Konstan, J.T. Riedl, An empirical analysis of design choices in neighborhood-based collaborative filtering algorithms, Information Retrieval 5 (2002) 287–310.
[94] J.L. Herlocker, J.A. Konstan, J.T. Riedl, L.G. Terveen, Evaluating collaborative filtering recommender systems, ACM Transactions on Information Systems 22 (1) (2004) 5–53.
[95] F. Hernández, E. Gaudioso, Evaluation of recommender systems: a new approach, Expert Systems with Applications 35 (2008) 790–804.
[96] A. Hernando, J. Bobadilla, F. Ortega, J. Tejedor, Incorporating reliability measurements into the predictions of a recommender systems. Information Sciences, in press, doi: 10.1016/j.ins.2013.03.018.
[97] A. Hernando, J. Bobadilla, F. Ortega, A. Gutiérrez, Trees for explaining recommendations made through collaborative filtering, Information Sciences 218 (2013) 1–16.
[98] K. Heung-Nam, E.S. Abdulmotaleb, J. Geun-Sik, Collaborative error-reflected models for cold-start recommender systems, Decision Support Systems 51 (3) (2011) 519–531.
[99] Y. Ho, S. Fong, Z. Yan, A hybrid ga-based collaborative filtering model for online recommenders, in: International Conference on e-Business, 2007, pp. 200–203.
[100] L. Hossain, D. Fazio, The social networks of collaborative process, The Journal of High Technology Management Research 20 (2) (2009) 119–130.
[101] H.R. Hu, P. Pu, Using personality information in collaborative filtering for new users, in: Proceedings of the 2010 ACM Conference on Recommender Systems, 2010, pp. 17–24.
[102] Y. Hu, Y. Koren, C.H. Volinsky, Collaborative filtering for implicit feedback datasets, in: IEEE International Conference on Data Mining (ICDM), 2008, pp. 263–272.
[103] Y.P. Huang, W.P. Chuang, Y.H. KE, F.E. Sandnes, Using back-propagation to learn association rules for service personalization, Expert Systems with Applications 35 (2008) 245–253.
[104] Z. Huang, D. Zeng, H. Chen, A comparison of collaborative filtering recommendation algorithms for e-commerce, IEEE Intelligent Systems 22 (5) (2007) 68–78.
[105] N. Hurley, M. Zhang, Novelty and diversity in top-N recommendations-analysis and evaluation, ACM Transactions on Internet Technology 10 (4) (2011) 1–29.

[106] CH.S. Hwang, Y.CH. Su, K.CH. Tseng, Using genetic algorithms for personalized recommendation, Lecture Notes in Computer Science 6422 (2010) 104–112.
[107] H. Ingoo, J.O. Kyong, H.R. Tae, The collaborative filtering recommendation based on SOM cluster-indexing CBR, Expert Systems with Applications 25 (2003) 413–423.
[108] A. Jameson, B. Smyth, Recommendation to groups, in: P. Brusilovsky, A. Kobsa, W. Nejdl (Eds.), The Adaptive Web, 2007, pp. 596–627 (Chapter 20).
[109] D. Jannach, Fast computation of query relaxations for knowledge-based recommenders, AI Communications 22 (4) (2009) 235–248.
[110] R. Jäschke, L. Marinho, A. Hotho, L. Schmidt-Thieme, G. Stumme, Tag Recommendations in Folksonomies, in: Proceedings of the 11th European Conference on Principles and Practice of Knowledge Discovery in Databases, 2007, pp. 506–514.
[111] J.J. Jason, Contextualized mobile recommendation service based on interactive social network discovered from mobile users, Expert Systems with Applications 36 (9) (2009) 11950–11956.
[112] B. Jeong, J. Lee, H. Cho, User credit based collaborative filtering, Expert Systems with Applications 36 (2009) 7309–7312.
[113] T. Joachims, Text categorization with support vector machines: learning with many relevant features, in: European Conference on Machine Learning, 1998, pp. 137–142.
[114] A. Jøsang, R. Ismail, C. Boyd, A survey of trust and reputation systems for online service provision, Decision Support Systems 43 (2) (2007) 618–644.
[115] CH.W. Jyun, CH.CH. Chui, Recommending trusted online auction sellers using social network analysis, Expert Systems with Applications 34 (3) (2008) 1666–1679.
[116] C. Kaleli, H. Polat, Privacy-preserving SOM-based recommendations on horizontally distributed data, Knowledge Based Systems 33 (2012) 124–135.
[117] H.N. Kim, A. Alkhaldi, A.E. Saddik, G.S. Jo, Collaborative user modeling with user-generated tags for social recommender Systems, Expert Systems with Applications 38 (7) (2011) 8488–8496.
[118] H.N. Kim, A.T. Ji, I. Ha, G.S. Jo, Collaborative filtering based on collaborative tagging for enhancing the quality of recommendations, Electronic Commerce Research and Applications 9 (1) (2010) 73–83.
[119] J. Kim, B. Lee, M. Shaw, H. Chang, W. Nelson, Application of decision-tree induction techniques to personalized advertisements on internet storefronts, International Journal of Electronic Commerce 5 (3) (2001) 45–62.
[120] K. Kim, H. Ahn, Using a clustering genetic algorithm to support customer segmentation for personalizad recommender systems, in: Proceedings of the 13th International Conference on AI, Simulation, and Planning in High Autonomy Systems, 2004, pp. 409-415.
[121] K. Kim, H. Ahn, A recommender system using GA K-means clustering in an online Shopping market, Expert Systems with Applications 34 (2) (2008) 1200–1209.
[122] S. Kitisin, C. Neuman, Reputation-based trust-aware recommender system, in: Securecomm and Workshops, 2009, pp. 1–7.
[123] F. Kong, X. Sun, S. Ye, A comparison of several algorithms for collaborative filtering in startup stage, IEEE Transactions on Networks, Sensing and Control (2005) 25–28.
[124] Y. Koren, R. Bell, CH. Volinsky, Matrix factorization techniques dor recommender systems, IEEE Computer 42 (8) (2009) 42–49.
[125] G. Koutrika, B. Bercovitz, H. Garcia, FlexRecs: expressing and combining flexible recommendations, in: Proceedings of the 35th SIGMOD International Conference on Management of Data, 2009, pp. 745–757.
[126] B. Krulwich, Lifestyle finder: intelligent user profiling using large-scale demographic data, Artificial Intelligence Magazine 18 (2) (1997) 37–45.
[127] K. Kwon, J. Cho, Y. Park, Multidimensional credibility model for neighbor selection in collaborative recommendation, Expert Systems with Applications 36 (2009) 7114–7122.
[128] S.K. Lam, J. Riedl, Shilling recommender systems for fun and profit, in: International Conference on World Wide Web, 2004, pp. 393–402.
[129] X.N. Lam, T. Vu, T.D. Le, A.D. Duong, Addressing cold-start problem in recommendation systems, in: Conference On Ubiquitous Information Management And Communication, 2008, pp. 208–211.
[130] N. Landia, S.S. Anand, Personalised tag recommendation, in: Proceedings of the 2009 ACM Conference on Recommender Systems, 2009, pp. 83–36.
[131] K. Lang, NewsWeeder: learning to filter netnews, in: Proceedings 12th International Conference on Machine Learning, 1995, pp. 331–339.
[132] D.H. Lee, P. Brusilovsky, Does trust influence information similarity? in: Proceedings of the 2009 ACM Conference on Recommender Systems, 2009, pp. 71–74.
[133] M. Lee, Y. Woo, A hybrid recommender system combining collaborative filtering with neural network, Lecture Notes on Computer Sciences 2347 (2002) 531–534.
[134] S.K. Lee, Y.H. Cho, S.H. Kim, Collaborative filtering with ordinal scale-based implicit ratings for mobile music recommendations, Information Sciences 180 (11) (2010) 2142–2155.
[135] C.W. Leung, S.C. Chan, F.L. Chung, An empirical study of a cross-level association rule mining approach to cold-start recommendations, Knowledge Based Systems 21 (7) (2008) 515–529.
[136] Q. Li, Clustering approach for hybrid recommender system, in: Proceedings of the 2003 IEEE/WIC International Conference on Web Intelligence, 2003, pp. 33–38.
[137] Y.M. Li, CH.W. Chen, A synthetical approach for blog recommendation: combining trust, social relation, and semantic análisis, Expert Systems with Applications 36 (3) (2009) 6536–6547.
[138] Y.M. Li, CH.P. Kao, TREPPS: a trust-based recommender system for peer production services, Expert Systems with Applications 36 (2) (2009) 3263–3277.
[139] Y.M. Li, T.F. Liao, CH.Y. Lai, A social recommender mechanism for improving knowledge sharing in online forums, Information Processing and Management, in press, doi: 10.106/j.ipm.2011.10.004.
[140] S. Loh, F. Lorenzi, R. Granada, D. Lichtnow, L.K. Wives, J.P. Oliveira, Identifying similar users by their scientific publications to reduce cold start in recommender systems, in: Proceedings of the 5th International Conference on Web Information Systems and Technologies (WEBIST2009), 2009, pp. 593–600.
[141] L. Lü, M. Medo, Ch.H. Yeung, Y.Ch. Zhang, Z.K. Zhang, T. Zhou, Recommender systems, Physics Reports 519 (2012) 1–49.
[142] X. Luo, Y. Xia, Q. Zhu, Incremental collaborative filtering recommender based on regularizad matrix factorization, Knowledge-Based Systems 27 (2012) 271–280.
[143] X. Luo, Y. Xia, Q. Zhu, Applying the learning rate adaptation to the matrix factorization based collaborative filtering, Knowledge Based Systems 37 (2013) 154–164.
[144] H. Ma, I. King, M.R. Lyu, Learning to recommend with explicit and implicit social relations, ACM Transactions on Intelligent Systems and Technology 2 (3) (2011). Article 29.
[145] H. Ma, T. CH. Zhou, M.R. Lyu, I. King, Improving recommender systems by incorporating social contextual information, ACM Transactions on Information Systems 29 (2) (2011). Article 9.
[146] A. Machanavajjhala, A. Korolova, A.D. Sharma, Personalized social recommendations: accurate or private, in: Proceedings of the VLDB Endowment, vol. 4, issue 7, 2011, pp. 440–450.
[147] L.B. Marinho, L. Schmidt-Thieme, Collaborative tag recommendations, in: Proceedings of the 31st Annual Conference of the German Classification Society, 2008, pp. 533–540.
[148] L. Martinez, L.G. Perez, M.J. Barranco, Incomplete preference relations to smooth out the cold-start in collaborative recommender systems, in: Proceedings of the 28th North American Fuzzy Information Processing Society Annual Conference (NAFIPS2009), 2009a, pp. 1–6.
[149] L. Martinez, R.M. Rodriguez, M. Espinilla, REJA: a georeferenced hybrid recommender system for restaurants, in: IEEE/WIC/ACM International Joint Conference on Web Intelligence and Intelligent Agent Technonolgy (WI-IAT 3), 2009b, pp. 187–190.
[150] P. Massa, P. Avesani, Trust-aware collaborative filtering for recommender systems, Lecture Notes in Computer Science 3290 (2004) 492–508.
[151] P. Massa, P. Avesani, Trust-aware recommender Systems, in: Proceedings of the 2007 ACM conference on Recommender Systems, 2007, pp. 17–24.
[152] C. Matyas, C. Schlieder, A spatial user similarity measure for geographic recommender systems, in: Proceedings of the 3rd International Conference on GeoSpatial Semantics, 2009, pp. 122–139.
[153] K. Mccarthy, J. Reilly, L. Mcginty, B. Smyth, Thinking positively-explanatory feedback for conversational recommender systems, in: European Conference on Case-based reasoning (ECCBR), 2004, pp. 115–124.
[154] K. Mcnally, M.P. O'mahony, M. Coyle, P. Briggs, B. Smyth, A case study of collaboration and reputation in social web search, ACM Transactions on Intelligent Systems and Technology 3 (1) (2011). Article 4.
[155] D. Mcsherry, Explanation in recommender systems, Artificial Intelligence Review 24 (2) (2005) 179–197.
[156] F. Mcsherry, I. Mironov, Differentially Private recommender systems: building privacy into the netflix prize contenders, in: Proceedings of the 15th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD), 2009, pp. 627–636.
[157] P. Melville, R.J. Mooney, R. Nagarajan, Content-boosted collaborative filtering for improved recommendations, in: Proceeding Eighteenth National Conference on Artificial Intelligence, 2002, pp. 187–192.
[158] R. Meteren, M. Someren, Using content-based filtering for recommendation, in: Proceedings of ECML 2000 Workshop: Maching Learning in Information Age, 2000, pp. 47–56.
[159] S.E. Middleton, N.R. Shadbolt, D.C. De Roure, Ontological user profiling in recommender systems, ACM Transactions on Information Systems (TOIS) 22 (1) (2004) 54–88.
[160] R.J. Mooney, L. Roy, Content-based book recommending using learning for text categorization, in: Proceedings of the Fifth ACM Conference on Digital Libraries, 2000, pp. 195–204.
[161] T. Morrison, U. Aickelin, An artificial immune system as a recommender for Web sites, in: International Conference on Artificial Immune Systems, 2002, pp. 161–169.
[162] A. Nanolopoulus, D. Rafailidis, P. Symeonidis, Y. Manolopoulus, Music Box: personalized music recommendation based on cubic analysis of social tags, IEEE Transactions on Audio, Speech and Language Processing 18 (2) (2010) 407–412.
[163] K. Nehring, C. Puppe, A theory of diversity, Econometrica 70 (3) (2002) 1155–1198.
[164] E.R. Núñez-Valdéz, J.M. Cueva-Lovelle, O. Sanjuán-Martínez, V. García-Dıaz, P. Ordoñez, C.E. Montenegro-Marín, Implicit feedback techniques on recommender systems applied to electronic books, Computers in Human Behavior 28 (4) (2012) 1186–1193.

[165] J. O'donovan, Capturing trust in social web applications, in: J. Golbeck (Ed.), Computing with Social Trust, 2009, pp. 213–257.
[166] J. O'donovan, B. Smyth, Trust in recommender systems, in: International Conference on Intelligent user Interfaces, 2005, pp. 167–174.
[167] K. Oku, R. Kotera, K. Sumiya, Geographical recommender system based on interaction between map operation and category selection, in: Workshop on Information Heterogeneity and Fusion in Recommender Systems, 2010, pp. 71–74.
[168] J. Ortega, J. Bobadilla, A. Hernando, A. Gutiérrez, Incorporating group recommendations to recommender systems: alternatives and performance, Information Processing and Management (2013), http://dx.doi.org/10.1016/j.ipm.2013.02.003.
[169] J. Ortega, J.L. Sánchez, J. Bobadilla, A. Gutiérrez, Improving collaborative filtering-based recommender systems results using Pareto dominance, Information Sciences (2013), http://dx.doi.org/10.1016/j.ins.2013.03.011.
[170] A. Papadimitriou, P. Symeonidid, Y. Manolopoulus, A generalized taxonomy of explanations styles for traditional and social recommender systems, Data Minning Knowledge Discovery 24 (3) (2012) 555–583.
[171] D.H. Park, H.K. Kim, I.Y. Choi, J.K. Kim, A literature review and classification of recommender Systems research, Expert Systems with Applications 39 (2012) 10059–10072.
[172] S.T. Park, W. Chu, Pairwise preference regression for cold-start recommendation, in: Proceedings of the 2009 ACM Conference on Recommender Systems, 2009, pp. 21–28.
[173] S.T. Park, D.M. Pennock, O. Madani, N. Good, D. Coste, Naïve filterbots for robust cold-start recommendations, in: Proceedings of Knowledge Discovery and Data Mining (KDD2006), 2006, pp. 699–705.
[174] Y.J. Park, A. Tuzhilin, The long tail of recommender systems and how to leverage it, in: Proceedings of the 2008 ACM Conference on Recommender Systems, 2008, pp. 11–18.
[175] M. Pazzani, D. Billsus, Learning and revising user profiles: the identification of interesting web sites, Machine Learning 27 (3) (1997) 313–331.
[176] M.J. Pazzani, D. Billsus, Content-based recommender systems, in: P. Brusilovsky, A. Kobsa, W. Nejdl (Eds.), The Adaptive Web, 2007, pp. 291–324 (Chapter 10).
[177] M. Pazzani, A framework for collaborative, content-based, and demographic filtering, Artificial Intelligence Review-Special Issue on Data Mining on the Internet 13 (5-6) (1999) 393–408.
[178] S. Perugini, M.A. Gonçalves, E.A. Fox, Recomender systems research: a connection-centric Surrey, Journal of Intelligen Information Systems 23 (2) (2004) 107–143.
[179] M.C. Pham, Y. Cao, R. Klamma, M. Jarke, A clustering approach for collaborative filtering recommendation using social network analysis, Journal of Universal Computer Science 17 (4) (2011) 583–604.
[180] G. Pitsilis, S.J. Knapskog, Socila trust as a solution to address sparsity-inherent problems of recommender Systems, in: Proceedings of the 2009 ACM Conference on Recommender Systems, 2009, pp. 33–40.
[181] G. Pitsilis, X. Zhang, W. Wang, Clustering recommenders in collaborative filtering using explicit trust information, Advances in Information and Communication Technology 358 (2011) 82–97.
[182] A. Popescul, L.H. Ungar, D.M. Pennock, S. Lawrence, Probabilistic models for unified collaborative and content-based recommendation in sparse-data environments, in: Proceeding UAI '01 Proceedings of the 17th Conference in Uncertainty in Artificial Intelligence, 2001, pp. 437–444.
[183] C. Porcel, E. Herrera-Viedma, Dealing with incomplete information in a fuzzy linguistic recommender system to disseminate information in university digital libraries, Knowledge-Based Systems 23 (1) (2010) 32–39.
[184] C. Porcel, J.M. Moreno, E. Herrera-Viedma, A multi-disciplinar recommender system to advice research resources in university digital libraries, Expert Systems with Applications 36 (10) (2009) 12520–12528.
[185] C. Porcel, A. Tejeda-Lorente, M.A. Martínez, E. Herrera-Viedma, A hybrid recommender system for the selective dissemination of research resources in a technology transfer office, Information Sciences 184 (1) (2012) 1–19.
[186] J. Preece, B. Shneiderman, The reader to leader framework: motivating technology-mediated social participation, AIS Transactions on Human–Computer Interaction 1 (1) (2009) 13–32.
[187] P. Pu, L. Chen, Trust-inspiring explanation interfaces for recommender systems, Knowledge Based Systems 20 (2007) 542–556.
[188] W. Qin, L. Xin, H. Liang, Unifying user-based and item-based algorithm to improve collaborative filtering accuracy, Energy Procedia 13 (2011) 8231–8239.
[189] L. Ramaswamy, P. Deepak, R. Polavarapu, K. Gunasekera, D. Garg, K. Visweswariah, S. Kalyanaraman, CAESAR: a context-aware, social recommender system for low-end mobile devices, in: International Conference on Mobile Data Management: Systems, Services and Middleware, 2009, pp. 338–347.
[190] A.M. Rashid, G. Karypis, J. Riedl, Learning preferences of new users in recommender systems: an information theoretic approach, in: ACM SIGKDD Explorations Newsletter, vol. 10, issue 2, 2008, pp. 90–100.
[191] S. Ray, A. Mahanti, Strategies for effective shilling attacks against recommender systems, Lecture Notes in Computer Science 5456 (2009) 111–125.
[192] L. Ren, L. HE, J. Gu, W. Xia, F. Wu, A hybrid recommender approach based on Widrow–Hoff learning, in: International Conference on Future Generation Communication and Networking, 2008, pp. 40–45.
[193] T.H. Roh, K.J. Oh, I. Han, The collaborative filtering recommendation based on SOM cluster-indexing CBR, Expert Systems with Applications 25 (2003) 413–423.
[194] J.A. Rodrigues, L.F. Cardoso, J. Moreira, G. Xexeo, Bringing knowledge into recommender systems, The Journal of Systems and Software, in press, http://dx.doi.org/10.1016/j.jss.2012.10.002.
[195] S.B. Roy, S. Amer-Yahia, A. Chala, G. Das, C. Yu, Space efficiency in group recommendation, The International Journal on Very Large Data Bases 19 (6) (2010) 877–900.
[196] G. Ruffo, R. Schifanella, A peer-to-peer recommender system base don spontaneous affinities, ACM Transactions on Internet Technology 9 (1) (2009) 1–34.
[197] P.B. Ryan, D. Bridge, Collaborative recommending using formal concept analysis, Knowledge Based Systems 19 (5) (2006) 309–315.
[198] G. Salton, Automatic Text Processing: The Transformation, Analysis, and Retrieval of Information by Computer, Addison-Wesley, Reading, MA, 1989.
[199] M. Saranya, T. Atsuhiro, Hybrid recommender systems using latent features, in: Proceedings of the International Conference on Advanced Information Networking and Applications Workshops, 2009, pp. 661–666.
[200] B. Sarwar, G. Karypis, J.A. Konstan, J. Riedl, Item-based collaborative filtering recommendation algorithms, in: 10th International Conference on World Wide Web, 2001, pp. 285–295.
[201] B. Sarwar, G. Karypis, J. Konstan, J. Riedl, Analysis of recommendation algorithms for e-commerce, in: ACM Conference on Electronic Commerce, 2000a, pp. 158–167.
[202] B. Sarwar, G. Karypis, J. Konstan, J. Riedl, Application of dimensionality reduction in recommender system – a case study, in: ACM WebKDD Workshop, 2000b, pp. 264–272.
[203] J.B. Schafer, D. Frankowski, J. Herlocker, S. Sen, Collaborative filltering recommender systems, in: P. Brusilovsky, A. Kobsa, W. Nejdl (Eds.), The Adaptive Web, 2007, pp. 291–324 (Chapter 9).
[204] A.I. Schein, A. Popescul, L.H. Ungar, D.M. Pennock, Methods and metrics for cold-start recommendations, in: Proceeding SIGIR '02 Proceedings of the 25th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval, 2002, pp. 253–260.
[205] C. Schlieder, Modeling collaborative semantics with a geographic recommender, in: Workshop on Semantic and Conceptual Issues in Geographic Information Systems, 2007, pp. 336–347.
[206] J. Serrano-Guerrero, E. Herrera-Viedma, J.A. Olivas, A. Cerezo, F.P. Romero, A google wave-based fuzzy recommender system to disseminate information in University Digital Libraries 2.0., Information Sciences 181 (9) (2011) 1503–1516.
[207] Z. Severac, V. Devedzic, J. Jovanovic, Adaptive neuro-fuzzy pedagogical recommender, Expert Systems with Applications 39 (10) (2012) 9797–9806.
[208] A. Shepitsen, J. Gemmell, B. Mobasher, R. Burke, Personalized recommendation in social tagging systems using hierarchical clustering, in: Proceedings of the 2008 ACM Conference on Recommender Systems, 2008, pp. 259–266.
[209] S.K. Shinde, U. Kulkami, Hybrid personalizad recommender system using centering–bunching based clustering algorithm, Expert Systems with Applications 39 (1) (2012) 1381–1387.
[210] S. Siersdorfer, S. Sergei, Social recommender systems for web 2.0 folksonomies, in: 20th ACM conference on Hypertext and hipermedia, 2009, pp. 261–269.
[211] I. Soboroff, C. Nicholas, Combining content and collaboration in text filtering, in: Proceedings of the IJCAI'99 Workshop on Machine Learning for Information Filtering, 1999, pp. 86–91.
[212] X. Su, T.M. Khoshgoftaar, A survey of collaborative filtering techniques, Advance in Artificial Intelligence 2009 (2009) 1–19.
[213] P. Symeonidis, A. Nanopoulus, Y. Manolopoulus, Providing justifications in recommender systems, IEEE Transactions on Systems, Man and Cybernet 38 (6) (2008) 1262–1272.
[214] P. Symeonidis, A. Nanopoulus, Y. Manolopoulus, MovieExplain: a recommender system with explanations, in: Proceedings of the 2009 ACM Conference on Recommender Systems, 2009, pp. 317–320.
[215] G. Takács, I. Pilászy, B. Németh, D. Tikk, Scalable collaborative filtering approaches for large recommender systems, Journal of Machine Learning Research 10 (2009) 623–656.
[216] S. Tan, J. Bu, CH. Chen, X. He, Using rich social media information for music recommendation via hypergraph model, ACM Transactions on Multimedia Computing, Communications, and Applications 7 (1) (2011). Article 7.
[217] N. Tintarev, J. Masthoff, A survey of explanations in recommender systems, in: IEEE 23rd International Conference on Data Engineering Workshop, 2007, 801–810.
[218] T. Tran, R. Cohen, Hybrid recommender systems for electronic commerce, in: Proceedings of the 17th National Conference on Artificial Intelligence, AAAI, 2000, pp. 78–84.
[219] K.H.L. Tso-Sutter, L.B. Marinho, L. Schmidt-Thieme, Tag-aware recommender systems by fusion of collaborative filtering algorithms, in: Proceedings of the 2008 ACM Symposium on Applied Computing, 2008, pp. 1995–1999.
[220] S. Vargas, P. Castells, Rank and relevance in novelty and diversity metrics for recommender systems, in: Proceedings of the 2011 ACM Conference on Recommender Systems, 2011, pp. 109–116.
[221] P. Victor, CH. Cornelis, M. De-Cock, Trust Networks for Recommender Systems, Antalis Press, 2011.

[222] J. Vig, S. Sen, J. Riedle, Tagsplanations: Explaining recommendations using tags, Proceedings of the 13th international conference on Intelligent user interfaces, 2009, pp. 47-56.
[223] P. Victor, CH. Cornelis, M. De-Cock, P.P. DA-SILVA, Gradual tust and distrust in recommender systems, Fuzzy Sets and Systems 160 (10) (2009) 1367–1382.
[224] M.G. Vozalis, K.G. Margaritis, Using SVD and demographic data for the enhancement of generalized collaborative filtering, Information Sciences 177 (2007) 3017–3037.
[225] Y. Wan-Shiou, CH. Hung-Chi, D. Jia-Ben, A location-aware recommender system for mobile shopping environments, Expert Systems with Applications 34 (1) (2008) 437–445.
[226] J. Wang, A. Vries, M. Reinders, Unifying user-based and item-based collaborative filtering approaches by similarity fusion, in: Proc. SIGIR Conf., 2006, pp. 501–508.
[227] J. Wang, A.P. Vries, M.J. Reinders, Unified relevance models for rating prediction in collaborative filtering, ACM Transactions in Information Systems 26 (3) (2008) 1–42.
[228] L.T. Weng, Y. Xu, Y. Li, R. Nayak, Exploiting item taxonomy for solving cold-start problem in recommendation making, in: Proceedings of the 20th IEEE International Conference on Tools with Artificial Intelligence (ICTAI2008), 2008, pp. 113–120.
[229] B. Widrow, M.E. Hoff, Adaptive switching circuits, in: Convention Record, IRE WESCON, 1960, pp. 96–104.
[230] P. Winoto, T.Y. Tang, The role of user mood in movie recommendations, Expert Systems with Applications 37 (8) (2010) 6086–6092.
[231] W. Woerndl, G. Groh, Utilizing physical and social context to improve recommender systems, in: IEEE/WIC/ACM International Conferences on Web Intelligence and Intelligent Agent Technology, 2007, pp. 123–128.
[232] B. Xie, P. Han, F. Yang, R.M. Shen, H.J. Zeng, Z. Chen, DCFLA: a distributed collaborative-filtering neighbor-locating algorithm, Information Sciences 177 (6) (2007) 1349–1363.
[233] W. Xin, Q. Jamaliding, T. Okamoto, Discovering social network to improve recommender system for group learning support, in: International Conference on Computational Intelligence and Software Engineering, 2009, pp. 1–4.
[234] R.R. Yager, Fuzzy logic methods in recommender systems, Fuzzy Sets and Systems 136 (2) (2003) 133–149.
[235] W.S. Yang, H.CH. Cheng, J.B. Dia, A location-aware recommender system for mobile shopping environments, Expert Systems With Applications 34 (1) (2008) 437–445.
[236] Y. Yang, An evaluation of statistical approaches to text categorization, Information Retrieval 1 (1) (1999) 67–88.
[237] Z. Yao, Q. Zhang, Item-based clustering collaborative filtering algorithm under high dimensional sparse data, in: International Joint Confeence on Computational Sciences and Optimization, 2009, pp. 787–790.
[238] Z. Yu, X. Zhou, Y. Hao, J. Gu, TV program recommendation for multiple viewers based on user profile merging, User Modeling and User-Adapted Interaction 16 (1) (2006) 63–82.
[239] W. Yuan, D. Guan, Y.K. Lee, S. Lee, S.J. Hur, Improved trust-aware recommender system using small-worldness of trust networks, Knowledge Based Systems 23 (3) (2010) 232–238.
[240] G. Zacharia, A. Moukas, P. Maes, Collaborative reputation mechanisms for electronic marketplaces, Decision Support Systems 29 (2000) 371–388.
[241] O. Zaiane, Building a recommender agent for e-learning systems, in: Proceedings of the International Conference on Computers Education (ICCE'02), vol. 1, 2002, pp. 55–59.
[242] J. Zhan, Privacy-preserving collaborative recommender systems, IEEE Transactions on Systems, Man and Cybernetics 40 (4) (2010) 472–476.
[243] F. Zhang, H.Y. Chang, A collaborative filtering algorithm employing genetic clustering to ameliorate the scalability issue, in: IEEE International Conference on e-Business Engineering, 2006, pp. 331–338.
[244] S. Zhang, W. Wang, J. Ford, F. Makedon, Using singular value decomposition approximation for collaborative filtering, in: IEEE International Conference on E-Commerce Technology, 2005, pp. 1–8.
[245] L. Zhen, G.Q. Huang, Z. Jiang, Recommender systems based on workflow, Decision Support Systems 48 (2009) 237–245.
[246] L. Zhen, G.Q. Huang, Z. Jiang, Collaborative filtering based on workflow space, Expert Systems with Applications 36 (2009) 7873–7881.
[247] L. Zhen, Z. Jiang, H. Song, Distributed recommender for peer-to-peer knowledge sharing, Information Sciences 210 (2010) 3546–3561.
[248] N. Zheng, Q. Li, A recommender system based on tag and time information for social tagging systems, Expert Systems with Applications 38 (4) (2011) 4575–4587.
[249] Y. Zheng, X. Xie, Learning travel recommendations from user-generated GPS traces, ACM Transactions on Intelligent Systems and Technology 2 (2011) 1. Article 2.
[250] Y. Zheng, L. Zhang, Z. Ma, X. Xie, W.Y. Ma, Recommending friends and locations based on individual location history, ACM Transactions on the Web 5 (2011) 1. Article 5.
[251] J. Zhong, X. Li, Unified collaborative filtering model based on combination of latent features, Expert Systems with Applications 37 (2010) 5666–5672.
[252] R.L. Zhu, S.J. Gong, Analyzing of collaborative filtering using clustering technology, international colloquium on computing, in: ISECS International Colloquium on Computing, Communication, Control, and Management, 2009, pp. 57–59.
[253] C.N. Ziegler, S.M. Mcnee, J.A. Konstan, G. Lausen, Improving recommendation lists through topic diversification, in: Proceedings of the 14th International Conference on World Wide Web, 2005, pp. 22–32.

Please cite this article in press as: J. Bobadilla et al., Recommender systems survey, Knowl. Based Syst. (2013), http://dx.doi.org/10.1016/j.knosys.2013.03.012