Comparative Study on Approaches of Recommendation Systems

Chapter · April 2020

DOI: 10.1007/978-981-15-0947-6_72

5 authors, including:

Khalid al Fararni
Sidi Mohamed Ben Abdellah University

Badraddine Aghoutane
Faculté des Sciences de Meknès

Jamal Riffi
FSDM Sidi Mohamed Ben Abdellah University

Sabri Abdelouahed
Sidi Mohamed Ben Abdellah University

Comparative Study on Approaches of Recommendation Systems

Khalid Al Fararni, Badraddine Aghoutane, Jamal Riffi, Abdelouahed Sabri and Ali Yahyaouy

Abstract In the current context, characterized by an information overload, also known as infobesity, it has become essential to design mechanisms that allow users to access what interests them as quickly as possible. Hence, recommender systems have emerged. This article then consists of examining and comparing the different existing recommendation approaches: those content-based filtering, those collaborative filtering, and finally the demographic and social approaches, while indicating, for each approach, the fields of application, some interesting examples, as well as its advantages and limitations. Then, we will indicate the hybridization techniques available to overcome these limitations.

Keywords Recommender systems · Collaborative filtering · Content-Based filtering · Hybrid recommender system

1 Introduction

Frequently, we are faced with making choices. How to dress? Which movie to watch? Which item to buy? Where to travel? The list of possibilities available to us is

K. Al Fararni (✉) · J. Riffi · A. Sabri · A. Yahyaouy
LIIAN Laboratory, Department of Informatics, Faculty of Sciences Dhar El Mahraz, Sidi Mohamed Ben Abdellah University, 1796 Fez-Atlas, Fez, Morocco
e-mail: khalid.alfararni@usmba.ac.ma

J. Riffi
e-mail: jamal.riffi@usmba.ac.ma

A. Sabri
e-mail: abdelouahed.sabri@gmail.com

A. Yahyaouy
e-mail: ali.yahyaouy@usmba.ac.ma

B. Aghoutane
Team of Processing and Transformation of Information, Polydisciplinary Faculty of Errachidia, Moulay Ismaïl University, 11201 Zitoune, Meknes, Morocco
e-mail: b.aghoutane@fpe.umi.ac.ma

V. Bhateja et al. (eds.), *Embedded Systems and Artificial Intelligence*,
Advances in Intelligent Systems and Computing 1076,
https://doi.org/10.1007/978-981-15-0947-6_72

generally very large, the evaluation of these possibilities to find what suits us the most is a difficult task and can consume a lot of our time. To address this problem of information overload and choices, recommender systems have emerged in the last decade of the twentieth century.

Recommender systems have been studied in various fields such as the web, e-commerce, and many others.

This paper presents at first a state of the art on the various approaches of recommendation, not only based on preferences of the users, but also on the basis of their behaviors, their demographic profiles, and the judgments of other users (presented on the social networks). A summary table that compares these approaches is provided in Sect. 3. We conclude in Sect. 4, on the perspectives of our work.

2 Recommender Systems

Recommender systems are a specific form of *Filtering Information* to produce to an «active user» (website visitor, a potential buyer of a product, etc.) a personalized list of suggestions (services, products, or links to click, etc.) related to its concerns and expectations. These suggestions are based on what the user has purchased or previously viewed but also on the activity history of other buyers with a similar profile.

Thus, this is how we differentiate *Information Retrieval Systems* for which the user's request to be oriented to the appropriate choices is explicit and recommender systems where the participation of user is non-voluntary, induced in particular by trace engines [1].

Generally, a recommender system makes it possible to compare the profile of a user with certain reference characteristics, and seeks to predict the «opinion» that this user would give to a given item. These characteristics can come from:

- The object itself, we speak about a *Content-based approach*.
- Social environment, we speak about *Collaborative Filtering Approach* or *Social Approach*.
- Demographic data, we speak about a *Demographic Approach*.

2.1 Content-Based Approach

In this approach, also called cognitive filtering, each item is defined by a set of attributes and their value. The choice of the items to recommend is based on a comparison of the topics covered in these items with the topics relevant to the user.

In a content-based recommendation system, the attributes (keywords) that are used to describe the items, and the user profile is designed to indicate the type of item that this user likes. In other words, for the proposal of recommendations to the users, this technique is based on the analysis of the similarities of contents between the items previously consulted by the users (or examine in the present) and those which have not been yet consulted [2].

In particular, various items of candidates are compared with items previously rated by the user and the most relevant items are recommended [3, 4].

This approach has its roots in Information Retrieval (IR) and Artificial Intelligence. In *Information Retrieval*, it is considered that users wanting recommendations are engaged in an information retrieval process. The user expresses a specific need by giving a request (usually a list of keywords). In Information Filtering (IF) Systems, the need is represented by the user's profile.

In *Artificial Intelligence*, the recommendation task can be expressed as a learning problem that exploits past user knowledge. In a simple way, user profiles are in the form of keyword vectors and reflect the long-term interests of the user. Often, it is best for the system to learn the user's profile rather than forcing the user to provide it. This usually involves the application of Machine Learning (ML) techniques. Their purpose is to learn how to categorize new information based on the information previously seen, and which has been implicitly or explicitly labeled as interesting or not by the user. With these labels, Machine Learning's methods can generate a predictive model which, given a new item, will help decide how much the item can interest the user.

There are two types of recommendation based on the content:

- *Keywords-based recommendation*: Consider the characteristics of items that the user has rated in the past to correlate them with their profile. In fact, user and item profiles are represented as weighted keyword vectors. To recommend interesting new items, these SRs attempt to match item attributes with the user's preferences and interests. For a new item, the system performs a cosine similarity calculation between the profile vector and the item vector to predict the user's score on the item.

It should be noted that these keywords are usually either extracted based on automatic indexing or manually assigned.

- *Recommendation based on semantics*: Semantics has been introduced by several methods in the recommendation process. These methods are discussed taking into account several criteria:
  - The type of source of knowledge involved (lexicon, ontology, etc.);
  - The techniques adopted for the annotation or the representation of items;
  - The type of content included in the user profile;
  - The matching strategy between items and profile.

2.2 Collaborative Filtering

Draws on the ratings given by a set of users on a set of articles, to generate the predictions. These appreciations, translated into numerical values, can be notes, accounts of purchases made, numbers of visits, etc. The key idea is that a note of the user *u* for a new item *i* is likely to be similar to that given by another user *v*, if *u* and *v* have similarly noted other items. Similarly, *u* is likely to rate two items *i* and *j* in the same way, if other users gave similar ratings to these two items. There are two main approaches to collaborative filtering:

- *User-Based Approach* [1]: compare users with each other and find those with common tastes, the notes of a user is then predicted according to his neighborhood.
- *Item-Based Approach* [5]: is to bring together items liked by the same people and to predict user ratings based on articles closest to those they have already noted.

2.3 Demographic Approach

This is a simple recommendation that proposes items relative to the demographic profile of the user [6]. It involves categorizing users into several classes or groups in relation to demographic information such as gender, age, occupation, location, language, country, etc. The principle of this approach is that two users who have evolved in a similar environment share common tastes than two users who have evolved in different environments and therefore do not share the same codes. Many sites use this solution to offer a "personalized" content offer. For example, users are redirected to a particular website based on their language or country. This approach has been very popular in the marketing literature but has received little attention in the field of recommendation algorithms.

2.4 Social Approach

This type of system offers recommendations based on the user's relationships with his friends in social networks, and sometimes this recommendation also depends on the value of user trust in each of his friends. Facebook (2.32 billion users), YouTube (1.9 billion users) are famous examples.

So, the question that arises is how can social models be exploited in information filtering systems?

The basic idea would be to replace the traditional community formation on the basis of votes with that induced by social media (friends and friends of friends). [7] compared the classic collaborative recommendations with those made by friends on three movie recommendation systems (Amazon.com, MovieCritic and Reel.com) and three others for a book recommendation (Amazon.com, RatingZone, and Sleeper). The results showed that users preferred those made by their friends. This can be explained by the fact that the friends are more qualified to advise them since they are supposed to know more about the preferences of the users.

2.5 *Hybrid Approach*

The approaches mentioned previously are not exclusive and different hybrid methods, combining these different types of filtering, have been developed [3]. The use of hybrid approaches improves the relevance of filtering system results by overcoming some of the limitations of the types of filtering presented in Table 1 as Overspecialization problem in content-based filtering; obtaining judgments which is an expensive task for users; etc.

The hybridization takes place in two phases:

1. Independently perform item filtering through collaborative methods or content (or other) to generate candidate recommendations (local scores);
2. Combine these sets of recommendations through hybridization methods such as weighting, mixing, cascading, switching, etc. [2], to produce the final recommendations for users (final scores).

2.5.1 Hybridization Techniques

The difference between hybridization techniques lies in the strategy chosen to combine local scores. Robin Burke summarized hybridization techniques in seven techniques [2], including:

- Weighted: The scores calculated by the different recommendation techniques are combined numerically and weighted to generate a single recommendation.
  The advantage of this method is that the combination is simple to perform and allows to adjust the hybridization depending on performance. The following linear equation can be used:

$$\mathrm{Score}_{\mathrm{finale}} = \alpha \times \mathrm{Score}_1 + (1-\alpha) \times \mathrm{Score}_2;\ 0 < \alpha < 1 \tag{1}$$

  This strategy is flexible, it is necessary to adjust the parameter α to quantify the contribution of each approach.
- Switching: This technique chooses one of several recommendation approaches, based on several criteria. The system must then define the switching criteria or the cases where the use of another technique is recommended. We can consider that the system tries to predict the note by the first approach and uses the second one in case of failure.

$$\mathrm{Score}_{\mathrm{finale}} = \begin{cases} \mathrm{Score}_1, \ \mathrm{si\ Score}_1 \neq \mathrm{null} \\ \mathrm{Score}_{\mathbf{2}}, \ \mathrm{si\ Score}_1 = \mathrm{null} \end{cases} \tag{2}$$

- Mixed: Recommendations from different techniques are all presented simultaneously to users. The problem that can arise with the use of this technique is the difficulty of calculating the scores to order a recommendation list, when all the techniques recommend the same items but with different notes. To overcome this problem, we can choose for example a list that gathers the five best items according to each technique, aggregate them in a single list, and present the list to the user, or each approach calculates scores of all items; each item has the highest score, the items are then listed according to their scores, and the best ones are recommended.

$$\text{Score}_{\text{finale}} = \max(\text{Score}_1, \text{Score}_2) \tag{3}$$

- Features combination: In a hybrid based on the combination of features, data from collaborative techniques is treated as a feature, and a content-based approach is used on that data.
- Cascade: The cascade involves a step-by-step process. In this case, a recommendation technique is applied first, producing a set of potential candidates. Then, a second technique refines the results obtained in the first step. This method has the advantage that if the first technique generates few recommendations, or if these recommendations are ordered to allow a quick selection, the second technique will no longer be used.
- Features augmentation: is similar to the cascade, but in this case, the results obtained (classification) of the first technique are used by the second as an added feature.
  Although both techniques, in cascade and by augmentation of features, chain two recommenders, with the first one has an influence on the second, they are fundamentally very different. In a hybrid by augmentation, the characteristics used by the second recommender include the output of the first, such as scores for example. In a hybrid by cascade, the second approach does not use the output of the first approach in the production of its ratings, but the results of the two recommenders are combined as a priority.
- Meta-level: In a meta-level hybrid, a first technique is used, but differently than the previous method (feature augmentation), not to produce new features, but to produce a model. And in the second stage, it is the whole model that will serve as input for the second technique [6].

3 Summary Table

In this table, we have reviewed and compared several recommendation approaches, indicating each approach the fields of application and some interesting examples, as well as its advantages and limitations.

Table 1 Summary table on recommendation approaches

	Representation and recommendation techniques		Examples of realization and their fields of application	Advantages	Limitations
Content-based approach	**For user/item profiling**: the vector model of Salton (or the logarithmic model) with a basic weighting such as TF-IDF (term frequency-inverse document frequency), decision trees, ontologies, etc. **For the measure of similarity and the prediction of the votes**: cosine similarity, Sørensen-Dice coefficient, Jaccard index, etc.	Keywords-based filtering	**In the web**: Letizia [8], Syskill and Webert [9], Amalthea [10], Personal WebWatcher [11] **News**: NewT [12], NewsDude [4], YourNews [13] **Books**: Citeseer [14], LIBRA [15] **Music**: Pandora [16] **Movies**17: Movies2GO [], INTIMATE [18]	• *User independence*: no need for a large community of users to make recommendations. A list of recommendations can be generated even if there is only one user • *No cold start problem for new items*: valid for the recommendation of new items or unpopular items • *Transparency*: an explanation of how the recommender system works can be provided by explicitly listing the characteristics or descriptions that led to the appearance of an item in a list of recommendations • *The quality increases with time*: the more the user will use the system, the more the relevance of the items that will be proposed to him will be fine	• *Traditional keyword-based profiles are not able to capture the semantics of users' interests*: suffers from polysemy and synonymy problems • *Knowledge of the field is often necessary*: for example, for the recommendation of films, the system needs to know the actors, the directors, etc. • *Overspecialization problem*: if a user is only interested in articles of sports, he will never be offered a political article • *Cold start*[a] (*new user*) • Problem with recommending multimedia documents (images, videos, etc.) in the absence of metadata
		Semantics-based	**Web**: SiteIF [19], SEWeP [20] **News**: Informer [21], News@hand [22] **Academic research papers**: Quickstep and Foxtrot [23] **Other fields**: ITR [24], Informed Recommender [25]		

(continued)

[a]Cold start refers to the lack of information on a new user or a new item that has just been added to the recommendation system.

Table 1 (continued)

	Representation and recommendation techniques		Examples of realization and their fields of application		Advantages	Limitations
Collaborative approach	**Item/user profiling**: rating matrix, models of latent factors, etc. **For the measure of similarity and the prediction of the votes**: Pearson correlation coefficient, cosine similarity, Spearman correlation coefficient, Bayesian clustering, maximum entropy, support-vector machine, singular value decomposition…	Memory-based filtering	User-based	**Movies**: GroupLens Movie (MovieLens) [26] Bellcore Video [27], Firefly [28] **Music**: Ringo [28]	• *Independence from the content*: does not require any knowledge of the content of the item or its semantics • *Works for any type of items*: whose content is either unavailable or difficult to analyze • *No overspecialization problem*: the recommendations are not based on the thematic dimension of the profiles	• *Request enough user to get satisfactory results* • *Sparse matrix* (*sparsity*): difficult to find users who voted the same items • *Cold start* (*new user/item*) • *Scalability*: often, the platforms on which collaborative filters are used have millions of users, products, and content. So it requires a lot of computing power to be able to offer suggestions to users • *Recommends only the most popular items and cannot recommend items not voted*
			Item-based	**TV**: Tivo [29] **E-mails**: Tapestry [30] **E-commerce**: Amazon [31] **Music**: LastFM [32], Mures [6]		
Demogr. approach	**For the measure of similarity and the prediction of the votes**: classification, selection…		**Restaurants**: [32] **E-commerce**: LifeStyle Finder [33], Alambic [34]		• *Simple to implement* • *Requires no history of estimates*	• *Privacy issues* • *The lack of diversity* • *Cold start* (*New item*)

(continued)

Table 1 (continued)

	Representation and recommendation techniques	Examples of realization and their fields of application	Advantages	Limitations
Social approach	**For item/user profiling**: ratings matrix, matrix of trust scores between users, etc. **For the prediction of the votes**: Web of Trust [35], TidalTrust [36]…	**Movies**: FilmTrust [36] **E-commerce**: MoleTrust [35]	• *Adaptability*: as the evaluations database (number of friends) increases, the recommendation becomes more precise	• *Cold start* (*a new item*) • *The accuracy decreases as the research moves away from the source user*
Hybrid approach	**For the prediction of the votes**: weighting, mixing, switching, features combination, cascade, features increase, and meta-level	**Web**: Fab [37] **News**: P-Tango [38], DailyLearner [39] **Movies**: Netflix, IMDb, CinemaScreen [40], More [41] **Restaurants**: EntreeC [2]	• *Adapts better with some pure filtering problems,* such as cold start, overspecialization, etc.	• *Requires additional settings*: related to the combination of different methods • *Unjustified proposals*: cannot explain why an article was recommended to a user

4 Conclusions

The field of recommendation systems, especially those using a hybrid approach, is still very recent.

In this article, we examined and compared several recommendation approaches, while indicating, for each approach the fields of application and some interesting examples. Then, we recalled that recommendation approaches generally suffer from a set of problems that reduce performance for a subset of users. We rely on hybrid SRs to overcome these problems. We then indicated the available hybridization techniques. However, it should be noted that these techniques do not take into account a lot of useful information to explain why an article was recommended to a user. This requires taking into account different sources of information on the user: his demographic data (Demographic filtering), his social network (Community filtering), etc. and their correlations with the content of the items that are proposed to him.

Thus, current recommender systems still require improvement and thus becoming a rich research area.

References

1. Resnick, P., Varian, H.R.: Recommender systems. Commun. ACM **40**, 56–58 (1997)
2. Burke, R.: Hybrid recommender systems: survey and experiments. User Model. User-Adap. Inter. **12**(4), 331–370 (2002)
3. Pazzani, M.J.: A framework for collaborative, content-based, and demographic filtering. Artif. Intell. Rev. **13**(5–6), 393–408 (1999)
4. Billsus, D., Pazzani, M.: A hybrid user model for news story classification. In: Seventh International Conference on User Modeling, Banff, Canada (1999)
5. Sarwar, B., Karypis, G., Konstan, J., Reidl, J.: Item-based collaborative filtering recommendation algorithms. In Proceedings of the 10th International Conference on World Wide Web (WWW'01), pp. 285–295. New York, NY, USA. ACM (2001)
6. Arnautu, O.R.: Mures: Un système de recommandation de musique. Master's thesis, La Faculté des arts et des sciences Université de Montréal (2012)
7. Sinha, R., Swearingen, K.: Comparing recommendations made by online systems and friends. In: DELOS-NSF Workshop on Personalization and Recommender Systems in Digital Libraries (2001)
8. Liberman, H.: Letizia: an agent that assists web browsing. In: International Joint Conference on Artificial Intelligence (IJCAI-95), pp. 924–929. Morgan Kaufmann publishers Inc., Montreal, Canada (1995)
9. Pazzani, M., Muramatsu, J., Billsus, D.: Syskill and Webert: identifying interesting web Sites. In: Thirteenth National Conference on Artificial Intelligence and the Eighth Innovative Applications of Artificial Intelligence Conference, pp. 54–61. AAAI Press/MIT Press, Menlo Park (1996)
10. Moukas, A.: Amalthaea information discovery and filtering using a multi-agent evolving ecosystem. Appl. Artif. Intell., 437–457 (1997)
11. Mladenic, D.: Machine learning used by personal webwatcher. In: ACAI-99 Workshop on Machine Learning and Intelligent Agents (1999)
12. Sheth, B., Maes, P.: Evolving agents for personalized information filtering. In: 9th Conference on Artificial Intelligence for Applications, pp. 345–352. IEEE Computer Society Press (1993)

13. Ahn, J., Brusilovsky, P., Grady, J., He, D., Syn, S.: Open user profiles for adaptive news systems: help or harm? In: 16th International Conference on World Wide Web, pp. 11–20. ACM (2007)
14. Bollacker, K., Giles, C.: CiteSeer: an autonomous web agent for automatic retrieval and identification of interesting publications
15. Mooney, R.J., Roy, L.: Content-based book recommending using learning for text categorization. In: Proceedings of the fifth ACM conference on digital libraries, pp. 195–204. ACM Press (1999)
16. Bu, J., Tan, S., Chen, C., Wang, C., Wu, H., Zhang, L., He, X.: Music recommendation by unified hypergraph: combining social media information and music content. In: Proceedings of the International Conference on Multimedia, ACM (2010)
17. Mukherjee, R., Jonsdottir, G., Sen S., Sarathi, P.: MOVIES2GO: an online voting based movie recommender system. In: Fifth International Conference on Autonomous Agents, pp. 114–115. ACM Press (n.d.)
18. Mak, H., Koprinska, I., Poon, J.: INTIMATE: a web-based movie recommender using text categorization. In: IEEE/WIC International Conference on Web Intelligence, pp. 602–605. IEEE Computer Society (2003)
19. Magnini, B., Strapparava, C.: Improving user modelling with content-based techniques. In: 8th International Conference of User Modeling, pp. 74–83 (2001)
20. Eirinaki, M., Vazirgiannis, M., Varlamis, I.: SEWeP: using site semantics and a taxonomy to enhance the web personalization process. In: 9th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, pp. 99–108 (2003)
21. Sorensen, H., O'Riordan, A., O'Riordan, C.: Profiling with the informer text filtering agent. J. Univers. Comput. Sci., 988–1006 (1997)
22. Cantador, I., Bellogin, A., Castells, P.: News@hand: a semantic web approach to recommending news. In: Adaptive Hypermedia and Adaptive Web-Based Systems, pp. 279–283 (2008)
23. Middleton, S., Shadbolt, N., De Roure, D.: Ontological user profiling in recommender systems. In: ACM Trans. Inf. Syst., 54–88 (2004)
24. Degemmis, M., Lops, P., Semeraro, G.: A content-collaborative recommender that exploits WordNet-based user profiles for neighborhood formation. User Model. User-Adap. Inter. J. Personalization Res. (UMUAI), 217–255 (2007)
25. Aciar, S., Zhang, D., Simoff, S., Debenham, J.: Informed recommender: basing recommendations on consumer product reviews. IEEE Intell. Syst. **22**, 39–47 (2007)
26. Konstan, J., Miller, B., Maltz, D., Herlocker, J., Gordon, L., Riedl, J.: GroupLens: applying collaborative filtering to usenet news. Commun. ACM, 77–87 (1997)
27. Hill, W., Stead, L., Rosenstein, M., Furnas, G.: Recommending and evaluating choices in a virtual community of use. In: Human Factors in Computing Systems (1995)
28. Maes, P., Shardanand, U.: Social information filtering: algorithms for automating "Word of Mouth". In: The SIGCHI Conference on Human Factors in Computing Systems. ACM Press, Addison-Wesley Publishing Co., Denver, Colorado, United States (1995)
29. Ali, K., van Stam, W.: TiVo: making show recommendations using a distributed collaborative filtering architecture. In: 10th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, pp. 394–401 (2004)
30. Goldberg, D., Nichols, D., Oki, B.M., Terry, D.: Using collaborative filtering to weave an information tapestry. Commun. Assoc. Comput. Mach. **35**(12), 61–70 (1992)
31. Linden, G., Smith, B., York, J.: Amazon.com recommendations: item-to-item collaborative filtering. IEEE Internet Comput. **7**, 76–80 (2003)
32. Jaschke, R., Marinho, L., Hotho, A., Schmidt-Thieme, L., Stumme, G.: Tag recommendations in folksonomies. In: Knowledge Discovery in Databases (PKDD 2007), pp. 506–514 (2007)
33. Krulwich, B.: LifeStyle finder: intelligent user profiling using large-scale demographic data. AI Magazine. **18**(2), 37–45 (1997)
34. Aimeur, E., Brassard, G., Fernandez, J.M., Onana, F.S.: Privacy-preserving demographic filtering. In: Proceedings of the ACM Symposium on Applied Computing, pp. 872–878 (2006)
35. Massa, P., Avesani, P.: Trust-aware collaborative filtering for recommender systems. In: On the Move to Meaningful Internet Systems 2004: CoopIS, DOA, and ODBASE. Springer, Berlin, Heidelberg, pp. 3–17 (2004)

36. Golbeck, J., Hendler, J.: FilmTrust: movie recommendations using trust in web-based social networks. Proceedings of the IEEE Consumer Communications and Networking Conference **96**, 2006 (2006)
37. Balabanovic, M., Shoham, Y.: Fab: content-based, collaborative recommendations. Commun. ACM **40**(3), 66–72 (1997)
38. Claypool, M., Gokhale, A., Miranda, T., Murnikov, P., Netes, D., Sartin, M.: Combining content-based and collaborative filters in an online newspaper. In: SIGIR'99 Workshop on Recommender Systems: Algorithms and Evaluation. Berkeley, CA (1999)
39. Billsus, D., Pazzani, M., Chen, J.: A learning agent for wireless news access. In: Proceedings of the 5th International Conference on Intelligent User Interfaces (IUI'00), pp. 33–36. ACM, New York, NY, USA (2000)
40. Salter, J., Antonoupoulos, N.: CinemaScreen recommender agent: combining collaborative and content-based filtering. IEEE Intell. Syst., 35–41 (2006)
41. Lekakos, G., Caravelas, P.: A hybrid approach for movie recommendation. Multimedia Tools Appl. **36**(1–2), 2006 (2006)