



International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 2, Issue 8, August 2014

A Survey on Recommender Systems (RSS) and Its Applications

P. N. Vijaya Kumar¹, Dr. V. Raghunatha Reddy²

¹PhD Research Scholar, Department of Computer Science & Technology, Sri Krishnadevaraya University, Anantapuramu,
India.

²Assistant Professor, Department of Computer Science & Technology, Sri Krishnadevaraya University, Anantapuramu,
India.

ABSTRACT: The growth of the Internet has made it much more difficult to effectively extract useful information from all the available online information. The overwhelming amount of data necessitates mechanisms for efficient information filtering. Recommender systems have the effect of guiding users in a personalized way to interesting objects in a large space of possible options. Recommender Systems (RSs) are software tools and techniques providing suggestions for items to be of use to a user. In this paper we will look at three different recommender system approaches namely Collaborative filtering (CF), Content-based filtering, Hybrid recommender systems that can be used on different e-commerce websites. We briefly describe each type with pros and cons and will present some of the applications of Recommender Systems (RSs) in different domains.

KEYWORDS: Recommender systems, Collaborative filtering, Content-based filtering, Hybrid recommender systems, information filtering.

I. INTRODUCTION

Recommender systems or recommendation systems are a subclass of information filtering system that seek to predict the 'rating' or 'preference' that user would give to an item [1]. Recommender systems have become extremely common in recent years, and are applied in a variety of applications. The most popular ones are books, research articles, movies, music, news, search queries, social tags, and products on e-commerce websites in general. Recommender systems assist and augment this natural social process to help people sift through available books, articles, web pages, movies, music, restaurants, jokes, grocery products, and so forth to find the most interesting and valuable information for them. Recommender systems are widely used by online stores for they improve user convenience and store benefits. It turns browsers into buyers, cross-sell items that are suggested at the checkout page, increase users' loyalty by making the purchase only a few clicks away or awarding frequent customers with good deals and such. E-commerce recommendation algorithms often operate in a challenging environment, especially for large online shopping companies like eBay and Amazon. Usually, a recommender system providing fast and accurate recommendations will attract the interest of customers and bring benefits to companies. Usually Recommender systems produce a list of recommendations in one of three ways: Collaborative filtering (CF), Content-based filtering, and Hybrid recommender systems. RSs are primarily directed towards individuals who lack sufficient personal experience or competence to evaluate the potentially overwhelming number of alternative items that a Web site, for example, may offer [2].

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 2, Issue 8, August 2014

II. TASKS OF A RECOMMENDER SYSTEM

Herlocker has a list of eleven popular tasks a recommender system can assist to implement: [3]

- **Find some good items:** A featured list of ranked items found that fit the user's requirements.
- **Find all good items:** A list of all the items that satisfy all the criteria the user set from the item database.
- **Annotation in text:** A list of items that are recommended according to the current context and the long term user preference. A certain TV series on a certain channel can be recommended according to the user's long term viewing habits.
- **Recommend a sequence:** A list of the item been searched and some related items that are not necessarily fitting search criteria, but it may be interesting for the user.
- **Recommend a bundle:** A list of related items that can work together to serve a purpose better for the user. Typically when you buy a camera, you may consider buying a memory card, a pouch and complete the purpose of the camera.
- **Just browsing:** For users that browse without a prominent purpose, the recommender system's task is to assist the user to browse items within the scope that are interesting to the user at that specific browsing session.
- **Find credible recommender:** There are users that are skeptical of the recommendation yielded by the system. The recommender system's task is then to allow the user to test the system's behavior.
- **Improve the profile:** The system can take inputs from the user about his/her likes and dislikes, in general, explicit preference information.
- **Express self:** Some users care little about recommendations, but it is important for them to be able to express their opinions and beliefs of certain item. A comment section is where the system can take such inputs, and also the satisfaction it creates can leverage as a motivation for purchasing the item related.
- **Help others:** Certain users may be even more motivated to leave a full review or rating of the item due to their belief in this benefiting the community. And that can be hugely motivational for other potential buyers to set their minds.
- **Influencing others:** Certain users could be exclusively influential, trying to convince other users buying or not buying the product. Even malicious user can fall into this category.

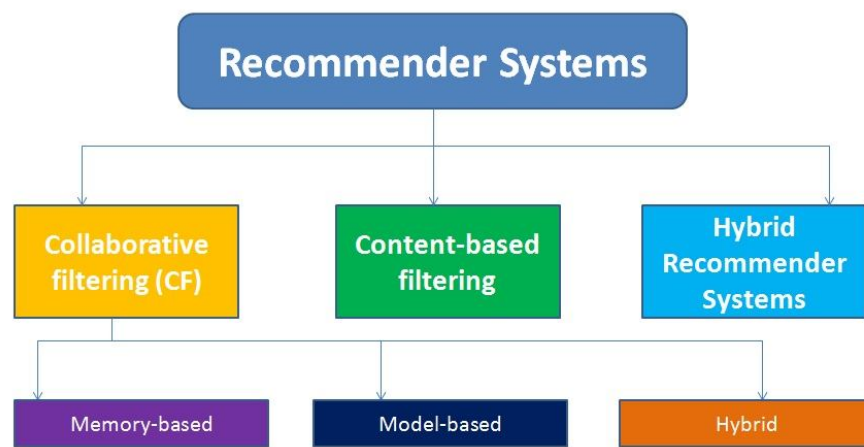


Fig. 1. Different types of Recommender systems

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 2, Issue 8, August 2014

III. COLLABORATIVE FILTERING

Collaborative filtering is the process of filtering information or patterns using techniques which involves collaboration among various agents, viewpoints, data sources, etc.[4] Applications of collaborative filtering involves very large data sets. Collaborative filtering methods have been applied to many different kinds of data including: financial data, such as financial service institutions that integrate many financial sources; monitoring and sensing data, such as in mineral exploration, environmental sensing over large areas or multiple sensors; or in electronic commerce and web applications where the focus is on user data, etc. collaborative filtering can be used for making automatic predictions about the interests of a user by collecting preferences or taste information from many users by means of collaboration.

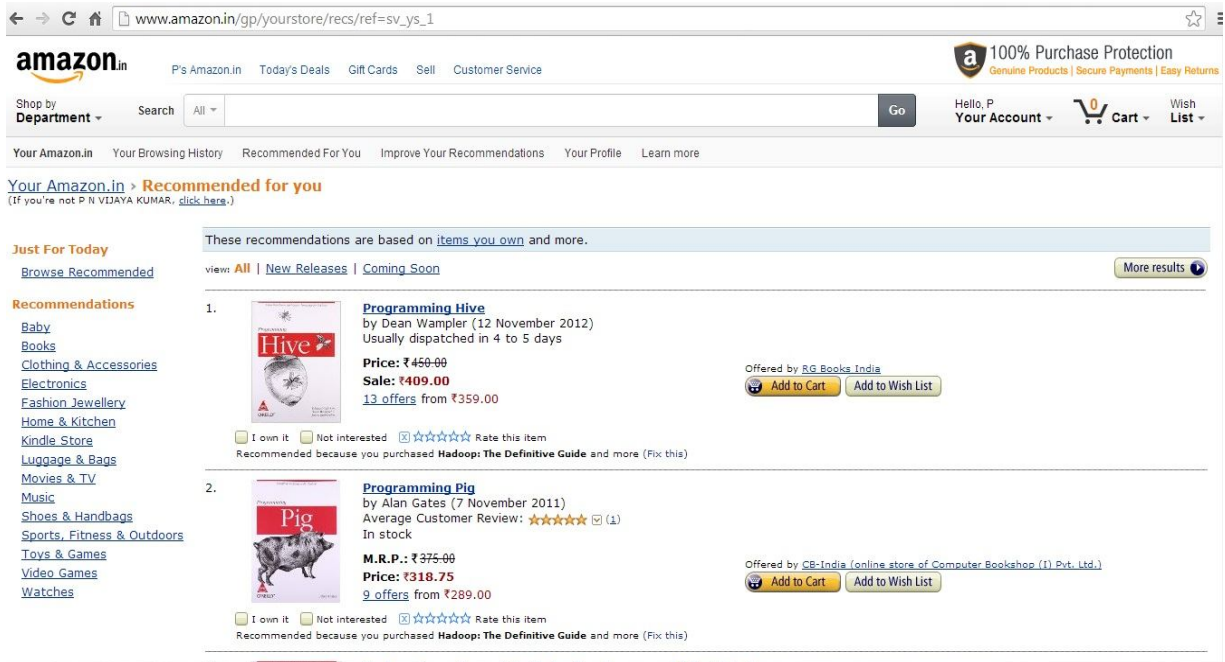


Fig. 2. Amazon recommends products to customers by customizing Collaborative filtering (CF) systems.

Collaborative filtering approach assumes that if a person X has the same opinion as a person Y on an issue, X is more likely to have Y's opinion on a different issue 'a' than to have the opinion on 'a' of a person chosen randomly. For example, a collaborative filtering recommendation system for laptop tastes could make predictions about which laptop brand a user should like given a partial list of that user's tastes (likes or dislikes).

Typical workflow of a collaborative filtering system is as under:

- A user expresses his or her preferences by rating items (e.g. articles, videos or books) of the system. These ratings can be viewed as an approximate representation of the user's interest in the corresponding domain.
- The system matches these users' ratings against other users and finds the people with most similar tastes.
- With similar users, the system recommends items that the similar users have rated highly but not yet being rated by this user.



International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 2, Issue 8, August 2014

A. Types of collaborative filtering

a) **Memory-based:** This mechanism uses user rating data to compute similarity between users or items. This is used for making recommendations. This was the earlier mechanism and is used in many commercial systems. It is easy to implement and is effective. Typical examples of this mechanism are neighborhood based CF and item-based/user-based top-N recommendations [5]. The Representative techniques are Neighbor-based CF (item-based/user-based CF algorithms with Pearson/vector cosine correlation) and Item-based/user-based top-N recommendations.

Advantages:

- easy implementation
- new data can be added easily and incrementally
- need not consider the content of the items being recommended
- scale well with co-rated items

Shortcomings:

- are dependent on human ratings
- performance decrease when data are sparse
- cannot recommend for new users and items
- have limited scalability for large datasets

b) **Model-based:** Models are developed using data mining, machine learning algorithms to find patterns based on training data. These are used to make predictions for real data. There are many model-based CF algorithms. This approach has a more holistic goal to uncover latent factors that explain observed ratings.[6] These include Bayesian networks, clustering models, latent semantic models such as singular value decomposition, probabilistic latent semantic analysis, Multiple Multiplicative Factor, Latent Dirichlet allocation and markov decision process based models [7].

Advantages:

- better address the sparsity, scalability and other problems
- improve prediction performance
- give an intuitive rationale for recommendations

Shortcomings:

- expensive model-building
- have trade-off between prediction performance and scalability
- lose useful information for dimensionality reduction techniques

c) **Hybrid recommenders:** A number of applications combine the memory-based and the model-based CF algorithms. These overcome the limitations of native CF approaches. It improves the prediction performance. Importantly, it overcomes the CF problems such as sparsity and loss of information. Usually most of the commercial recommender systems are hybrid, for example, Google news recommender system [7].

Advantages:

- overcome limitations of CF and content-based or other recommenders
- improve prediction performance
- overcome CF problems such as sparsity and gray sheep

Shortcomings:

- have increased complexity and expense for implementation,
- need external information that usually not available

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 2, Issue 8, August 2014

IV. CONTENT-BASED FILTERING

Content-based filtering methods are based on a description of the item and a profile of the user's preference [8]. These algorithms try to recommend items that are similar to those that a user liked in the past or is examining in the present. In particular, various candidate items are compared with items previously rated by the user and the best-matching items are recommended. This approach has its roots in information retrieval and information filtering research. Basically, these methods use an item profile characterizing the item within the system. The system creates a content-based profile of users based on a weighted vector of item features. The weights denote the importance of each feature to the user and can be computed from individually rated content vectors using a variety of techniques. Simple approaches use the average values of the rated item vector while other sophisticated methods use machine learning techniques such as Bayesian Classifiers, cluster analysis, decision trees, and artificial neural networks in order to estimate the probability that the user is going to like the item.

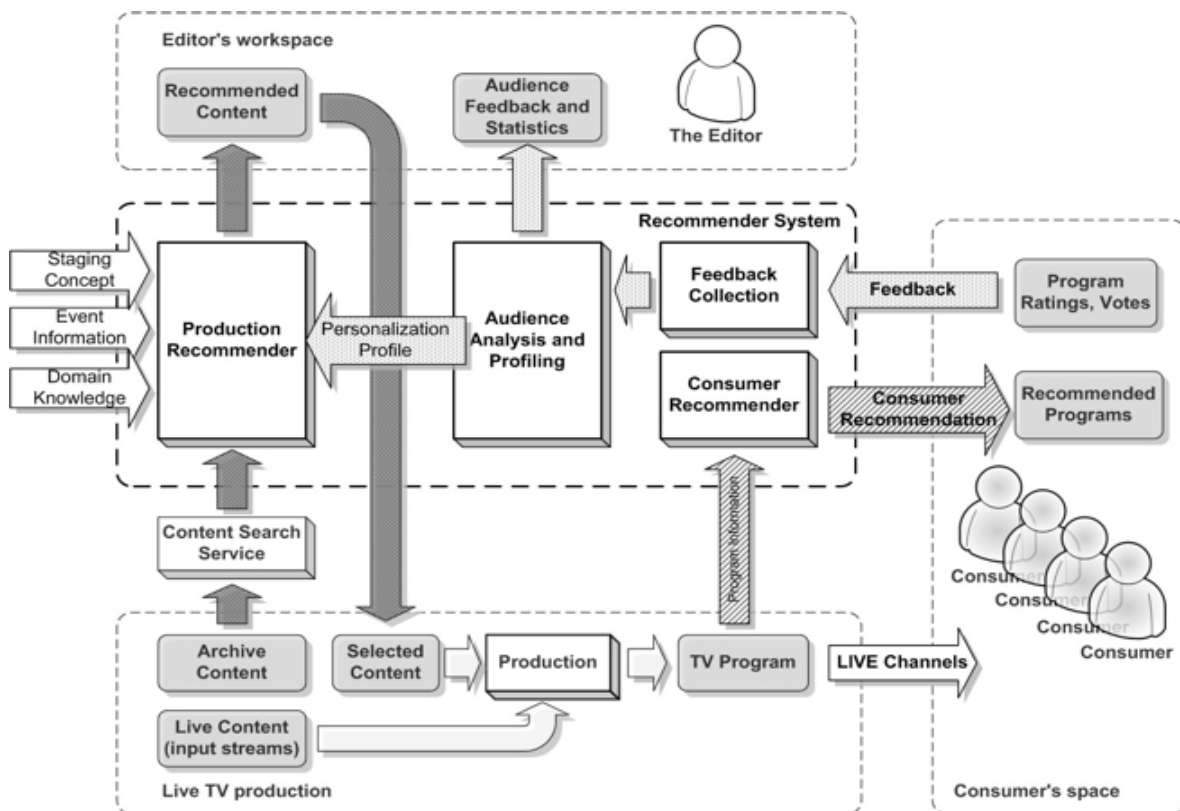


Fig. 3. Content recommender for the TV production, which includes Recommender GUI, and provides content search and personalization functionalities

Advantages: [9]

- **User Independence** - Content-based recommenders exploit solely ratings provided by the active user to build her own profile. Instead, collaborative filtering methods need ratings from other users in order to find the 'nearest neighbors' of the active user, i.e., users that have similar tastes since they rated the same items similarly. Then, only the items that are most liked by the neighbors of the active user will be recommended.



International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 2, Issue 8, August 2014

- **Transparency** - Explanations on how the recommender system works can be provided by explicitly listing content features or descriptions that caused an item to occur in the list of recommendations. Those features are indicators to consult in order to decide whether to trust a recommendation. Conversely, collaborative systems are black boxes since the only explanation for an item recommendation is that unknown users with similar tastes liked that item.
- **New Item** - Content-based recommenders are capable of recommending items not yet rated by any user. As a consequence, they do not suffer from the first-rater problem, which affects collaborative recommenders which rely solely on users' preferences to make recommendations. Therefore, until the new item is rated by a substantial number of users, the system would not be able to recommend it.

Shortcomings:

- A key issue with content-based filtering is whether the system is able to learn user preferences from user's actions regarding one content source and use them across other content types.
- Content-based recommenders have no inherent method for finding something unexpected. The system suggests items whose scores are high when matched against the user profile; hence the user is going to be recommended items similar to those already rated. This drawback is also called serendipity problem to highlight the tendency of the content-based systems to produce recommendations with a limited degree of novelty.

V. HYBRID RECOMMENDER SYSTEMS

Hybrid recommender systems are based on the combination of Collaborative filtering and Content-based filtering. These overcome the limitations of native CF approaches. It improves the prediction performance. Importantly, it overcomes the CF problems such as sparsity and loss of information. Given two or more basic recommender systems techniques, several ways have been proposed for combining them to create a new hybrid system [10]. However, they have increased complexity and are expensive to implement [11]. Usually most of the commercial recommender systems are hybrid, for example, Google news recommender system [12].

Advantages:

- The hybrid approach tries to overcome the limitations of the other approaches and combine the advantages of the other approaches by using two or more approaches.

Disadvantages:

- As they combine two or more approaches to give better recommender system with increased complexity, they are expensive to implement.

VI. APPLICATIONS OF RECOMMENDER SYSTEMS

The most common recommender systems applications include:

- **Entertainment** - recommendations for movies, music, and IPTV.
- **Content** - personalized newspapers, recommendation for documents, recommendations of Web pages, e-learning applications, and e-mail filters.
- **E-commerce** - recommendations for consumers of products to buy such as books, cameras, PCs etc.
- **Services** - recommendations of travel services, recommendation of experts for consultation, recommendation of houses to rent, or matchmaking services.



International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 2, Issue 8, August 2014

VII. CONCLUSION AND FUTURE SCOPE

The three recommender systems have their advantages and disadvantages in performing their job. Most of the limitations in each one of the approaches can be complimented by the other. A good recommender system should be able to provide positive and relevant recommendations from time to time and also provide alternative recommendations to break the fatigue of the users seeing the same items in the recommendation list. Future recommendation systems should be dynamic, and the profiles should be able to be updated in real time. This and the synchronization of various profiles implies the need of huge amount of computational power, network bandwidth etc. Current algorithms and techniques all have relatively high memory computational complexity, and that leads to long system processing time and data latency. Therefore, new algorithms and techniques that can reduce memory computational complexity eventually eliminate synchronization problems will be the one of development orientations.

REFERENCES

1. Francesco Ricci and Lior Rokach and Bracha Shapira, Introduction to Recommender Systems Handbook, Springer, pages 1-35, 2011,
2. Resnick, P., Varian, H.R., Recommender systems. Communications of the ACM 40(3), Pages 56–58, 1997.
3. Herlocker, J.L.; Konstan, J.A.; Terveen, L.G.; Riedl, J.T., Evaluating collaborative filtering recommender systems. ACM Transaction on Information Systems 22(1), Pages 5–53, 2004,
4. Terveen, Loren; Hill, Will, "Beyond Recommender Systems: Helping People Help Each Other". Addison-Wesley. Page 6, 2001.
5. An integrated approach to TV Recommendations by Red Bee Media.
6. Factor in the Neighbors: Scalable and Accurate Collaborative Filtering.
7. Xiaoyuan Su, Taghi M. Khoshgoftaar, A survey of collaborative filtering techniques, Advances in Artificial Intelligence archive, 2009.
8. Peter, Brusilovsky, the Adaptive Web. ISBN 978-3-540-72078-2., Page. 325, 2007.
9. Pasquale Lops, Marco de Gemmis, Giovanni Semeraro, Content-based Recommender Systems: State of the Art and Trends, Recommender Systems Handbook, Springer, Pages 73-105, 2011.
10. Burke, R.: Hybrid web recommender systems. In: The AdaptiveWeb, Springer Berlin / Heidelberg, Pages. 377–408, 2007.
11. Kernel Mapping Recommender System Algorithms, www.sciencedirect.com/science/article/pii/S0020025512002587.
12. Google News Personalization: Scalable Online Collaborative Filtering.