



Recommendation approach of educational resources in a frequent subnet

Gasmi Saraa, Bouhadada Tahara, Benmachiche Abdelmadjid•

a Département d'informatique, Université Badji Mokhtar - Annaba, Algérie.

Université Chadli Bendjedid -El Tarf

Informations	Résumé
Mots clés : Recommender system Content based filtering Collaborative filtering Hibryd filtering Data mining Frequent subnet *Correspondance : Gasmisara23@gmail.com Requ le, Révisé le, Accepté le	Recommendation systems or recommender systems (RS) are subset of of information used to anticipate the "assessment" or "preference" that a user would transmit to an item. The main idea of recommender system is to build relationship between the articles, users and make the decision to select the most appropriate articles to a specific user. There are three main ways that recommender systems produce a list of recommendations for a user content based, Collaborative, and hybrid filtering. Data mining is one of the important analysis techniques used in RS to predict user interest in information, products and services among the vast amount of available items. The data mining techniques that are most commonly used in RS • classification, clustering and association rule discovery. The objective of our work is to propose a model to customize educational resources based on connections between users in frequent networks. This model makes it possible to calculate the similarity between a user and his friends of the same frequent subnet.

1. Introduction

The multitude of resources, relationships and interactions present in social media can lead users to suffer from information overload. The development of the Social Web has created a need for new techniques to help users find what they are looking for these techniques are called Recommendation Systems (RS). The recommendation system recommends and suggests elements to the user taking into account their purpose, their behavior and their preferences [1]. There are three main ways that recommender systems produce a list of recommendations for a user content based, Collaborative, and hybrid filtering[2]. In contentbased filtering, the model uses the specifications of an item to recommend additional items with similar properties. [3]. Collaborative filtering operates by collecting user comments as reviews of items of a given domain and exploiting the evaluation behavior similarities between users to determine how to recommend an item. CF uses the opinions of users similar to the active user. [4]. Majority of current research involves employing a hybrid approach, which combines methods from both types of recommender systems.

There are many websites that use the RS to recommend data or items to customers like Netflix., Rotten Tomatoes, Movielense, IMDB, Amazon, eBay, Google News and social networking sites like Facebook, twitter, google*, LinkedIn, Instagram, etc. to recommend friends or followers.

Data mining is one of the important analysis techniques used in RS to predict user interest in information, products and services among the vast amount of available items. The data mining techniques that are most commonly used in RS classification, clustering and association rule discovery [5]. Data mining is defined as the process of discovering patterns in data.

The process must be automatic or (more usually), semiautomatic Data mining is the analysis of data and the use of software techniques for finding patterns and regularities in sets of data. Data mining provides a number of algorithms to obtain profiles of users based on historical data, which are used to predict the preferences of new users that are : Decision tree ,Kmeans ,SVM, Bayesian Networks.

The objective of our work is to propose an approach to customize educational resources based on connections between users in frequent networks. We rely on two principles in our work , the first principle states that similar individuals tend to refer to or connect to the same resources. This principle is used in conventional recommendation systems. These are primarily based on user ratings similar to a given user in order to predict their preferences. The second principle indicates that users who are socially connected are likely to share the same interests. So users of a system can be easily influenced by their friends and thus be interested in their activities.

This paper is organized as : section 2 presents recommender systems techniques widely used today Contentbased recommendations, Collaborative recommendations, and Hybrid Approaches,we also discuss advantages and disadvantage of recommender systems .Section 3 presents data mining methods used in recommender, Section 4 presents our model of recommendation system and finally, Section 5 concludes our discussion.

Recommender systems techniques

The recommendation system CRS) obtains information about the user using different methods and sources to predict the user's needs and recommend items based on the results of that analysis process [51].

According to Venkatesan et al [6] «Recommender system can be considered as a decision supportive framework that facilitates strategies and conventions to the service users for obtaining optimal recommendations towards availing the resources or services on the basis of the personalization or preferences of the user of concerned »

Fig.1 illustrates the general concept of the recommendation system while matching the user profile with that of other users, and recommends the difference between these items

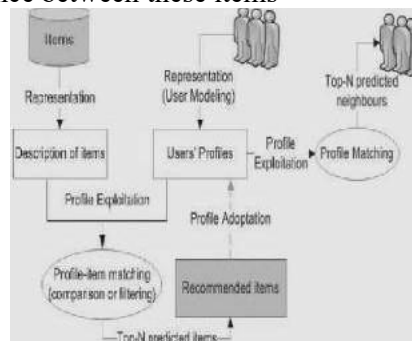


Fig.1. General Concept of Recommender Systems [7].

The RS are generally classified into three filtering methods to provide personalized recommendations to users; the most popular ones are collaborative filtering (CF), contentbased filtering and hybrid filtering.

2.1 Collaborative filtering

Collaborative filtering operates by collecting and analysing a large amount of information about user behaviors, activities, or preferences, and exploiting the evaluation similarities between users to determine how to recommend an item [81].

2.2 Content based filtering

Content-based filtering (CBF) uses only the preferences of the active user. It tries to recommend items by correlating content with the user profile. The user's profile is compared to the characteristics of the content of the article for which the user has already made the evaluation [91].

2.3 Hybrid filtering

Recent research has demonstrated that a hybrid approach, which combines methods from both types of recommender systems could be more effective in some cases. These methods can also be used to avoid cold start and sparsity problems [5].

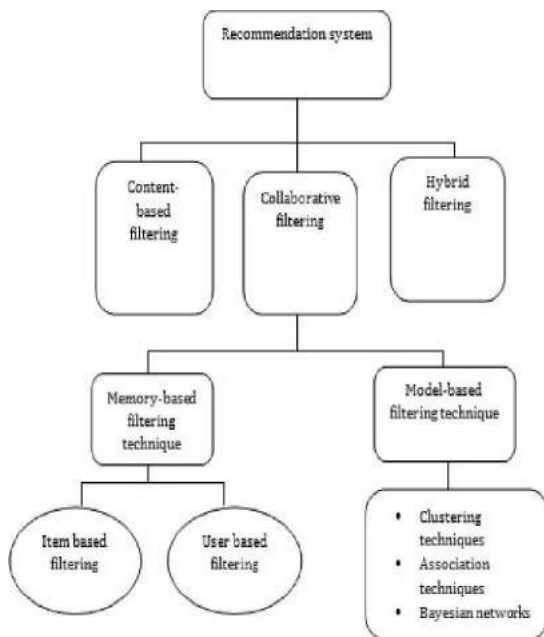


Fig.4. Recommendation technique

In this section we will discuss advantages and disadvantage of recommender systems techniques shown in table 1

TABLE 1. List advantages and disadvantages of recommender systems

dependency Of variables through various independent variables.

Techniques	Advantages	Disadvantages
ContentBased Filtering	Frustrate any cold start and sparsity problem.	Una ble to notice users interest.
Collaborative Filtering	The quality of articles can be easily evaluated. Recommend items to users outside their references.	Problem how to recommend items (Cold start roblem
Hybrid Approaches	Overcomes the difficulties of content based filtering and collaborative filtering.	Suffer from the cold start and sparsity problem.

Data mining methods used with a recommender system

The vast growth of information on the Internet and the number of visitors on websites add major challenges for recommendation systems. Recommendation systems uses Knowledge Discovery in Database techniques and prediction algorithms to predict user interest in information , products and services among the vast amount of available items. Data mining is also known as Knowledge Discovery in Database refers to finding or "mining" knowledge from large amounts of data. Data mining techniques are used to operate on large volumes of data to discover hidden patterns and relationships helpful in decision making-[10].

This study presents various methods used in data mining and recommender systems

- Outlier detection: the extreme values that

detach from the rest of the data are called outliers.

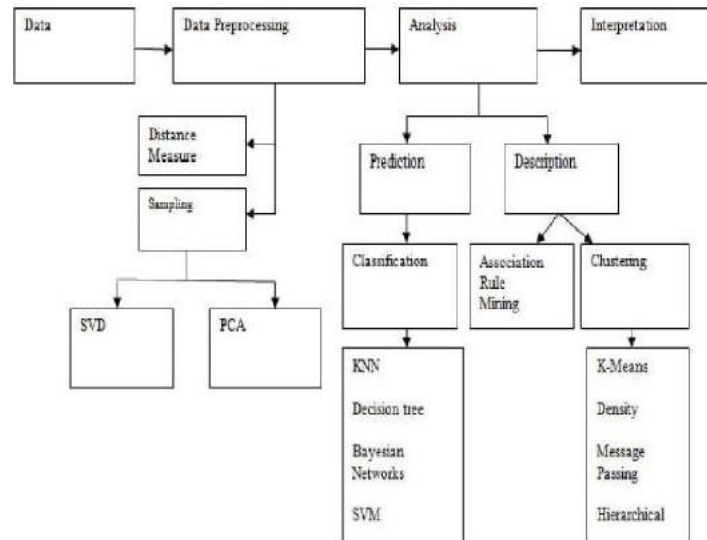
- Cluster analysis: Clusters is a collection of similar data objectsThe goal of this analysis is to segment unstructured data. To this end, algorithms are used to search for similarities in the structures of large data sets, in order to identify new clusters.

- Classification : While cluster analysis focuses on identifying new groups, classification involves the use of predefined classes. They are assigned using the

corresponding characteristics of the dataset.

- Association analysis : This technique aims to extract the relationships between attributes and items that can be formulated as inference rules.

- Regression analysis: This technique used to create models to find the Fig.5 .Data



mining methods with recommender systems.

The proposed approach

In the proposed model, the algorithm AGM(Apriori-based Graph Mining) [9] [10] aims to discover all the frequent subnets, in other words the patterns reporting support greater than a threshold minimum support. Indeed, the most widespread definition of a "pattern" in the context of networks is that of the "subgraph" [11]. The process of extracting frequent patterns takes place in two stages.

- A step of generating the candidate subgraphs, which is based on the frequent subgraphs of size n to generate the candidates of size $n + 1$.

- An evaluation phase which evaluates the frequency of each candidate using the isomorphism property in the graphs. This system builds a set of graph where the nodes represent the learners and the edges represent the relations of friendship between the users.

The system generates lists of recommended resources for a given user based on user characteristics, resources and the links between them. These lists are customized for each user in the subnet.

Each user is described with a set of information that characterizes them such as his interests, his habits, his actions, his preferences

expressed in the form of evaluations as well as his relations with other users. The recommendation process follows the following steps:

- 1) Choose a frequent subnetwork.

- 2) Select users related to the active user in the frequent subnet by a friendship link.

- 3) Calculate the degree of similarity between the active user and his friends.

- 4) Present to the learner a list of educational resources that have been liked by similar learners.

- 5) Repeat these points for each subnet,

Extraction process of

Social network

Frequent subnetwork

Frequent subnetwork

Social data Profile data

Recommender system

List of educational resources

Fig.6. The proposed approach of recommender system

a Algorithm Apriori

Function apriori(S_k , D , $\min_sup$)

D - Graph Database

S_k — frequent subgraphs of size ' k ' $\min_sup$

— minimum support value

L_{k+1} -loifs graphs whose size is ' $k+1$ '

S_{k+1}

For all g_i for all $g_j \in S_k$

L_{k+1} form as many as graphs possible by combining g_i and g_j : the size of the resultant graph is

For all

[If [(g_u is frequent in D) AND

($g_u \subseteq S_k$)] Then

$S_{k+1} \leftarrow S_{k+1} \cup g_u$;

```

End if]
If [Sk+1<> NUL
Then ipriori(Sk+1, D, min_sup);
End if] end function..

```

b Similarity

Collaborative filtering uses several measures in order to evaluate the similarities of assessments between users. Among these measures we can mention: the correlation coefficient of Pearson .

In the recommendation systems this coefficient is used to measure the dependence between two vectors of resource evaluations belonging to two users.

The Pearson correlation coefficient represents the correlation between two users u_a and u_b . I_a and I_b represent all the resources noted by u_a and u_b . $v(u_a, i)$ denotes the average of notes of u_a and $v(u_a; i)$ indicates the note of u_a on item i . I_c denotes the co-noted items (noted in common) between the active user u_a and the user u_b [12].

user u_a and the user u_b [12]. $\text{corrP}(u_a, u_b)$

$$\frac{\sum_{i \in I_c} (v(u_a, i) - v(u_a)) (v(u_b, i) - v(u_b))}{\sqrt{(\sum_{i \in I_c} (v(u_a, i) - v(u_a))^2) (\sum_{i \in I_c} (v(u_b, i) - v(u_b))^2)}}$$

According to Anand and Mobasher [13] when the correlation is 1, it means that the users u_a and u_b [21 Lalita, S., Anju, 2013) . A Study of Recommendation System: Research Challenges. International Journal of Engineering Trends and Technology(IJETT), vol 4(5), pp,1989-1992 .

[31 Hussien,M., Helmy,M., Hasan,M.(2019).

Recommender Systems Challenges and Solutions Survey. International Conference on Innovative Trends in Computer Engineering(ITCE'2019).

[4] Ricci, F., Rokach, L., & Shapira, B . (2011). Introduction to recommender systems handbook. In Recommender systems handbook(pp. 1-35)..

[5] Priyanga , P.,Nadira A.. ganu Kamal. (2017). Methods of Mining the Data from Big Data and Social Networks Based on Recommender System. International Journal of Advanced Networking & Applications (IJANA), vol 8(5), pp.55-60.

[6] Venkatesana , T., Ramkumarb , T., and Saravananc, K. (2017). Mining Big Data: Towards a Machine Learning Framework Based on Collaborative Filtering.

[7] Shah ,K., Zafar ,A., and Irfan U . (2016) .Recommender Systems Issues, Challenges, and Research Opportunities. [n formation Science and Applications (ICISA), pp . 1179-1189-.

[81 Priyanga , P., Nadira ,P., Banu KI (2017) .

Methods of Mining the Data from Big Data and Social Networks Based on Recommender System. International Journal Of Advanced

are strongly correlated. Now, if the correlation is -1, it means that u_a and u_b have totally opposite evaluations. When this correlation is 0, no relationship exists between the two users [12].

5. Conclusion

Recommender system apply various techniques to predict user interest on information. Recommender systems have different types of filtering such as (Collaborative, content-based, and hybrid) to make an effective recommendation engine. This paper presented our approach of recommendation system , this system generates lists of recommended resources for a given user in frequents subnetwork , We also introduce various data mining techniques used in recommender system, Data mining is one of the important analysis techniques used in RS to predict user interest in information, products and services among the vast amount of available items. We also discuss advantages and disadvantages of recommender systems.

Référence

111 Pu, P., Chen, Hu. R .(2012). Evaluating recommender systems from the user's perspective : survey of the state of the art. User modeling and User-Adapted Interaction 22, 317355. Networking & Applications (IJANA), vol 8(5), pp.55-60 .

[9] Aarti ,S . (2014).Application of Data Mining - A Survey Paper. International Journal of Computer Science and Information technologies , Vol. 5 (2). [10] Agrawal,R.,Srikant,R. (1994). Fast Algorithms for Mining Association Rules in Large Databases. In Proceedings of the 20th International Conference on Very Large Data Bases, pages 487—499.

[12]| Inokuchi,A, Washio,T., Motoda,H.(2000). An Apriori-Based Algorithm for Mining Frequent Substructures from Graph Data. In 4th European Conference on Principles of Data Mining .

[13] ChengH., Yan,X., Han,J. (2010).) Mining Graph Patterns. Managing and Mining Graph Data, pages 365-392. . (Cité en pages 26 et 110).

[14] Esslimani, (2010). Vers une approche comportementale de recommandation : apport de l'analyse des usages dans le processus de personnalisation. These de doctorat en informatique. Université de Nancy 2, France.