Jul. 31
International Journal of advanced studies in Computer Science and Engineering
IJASCSE, Volume 4, Issue 7, 2015

Recommender System for volunteers in connection with NGO

Pooja V. Chavan
Department of Computer Engineering & Information Technology
Veermata Jijabai Technological Institute, Matunga
Mumbai, India

Prof. P.M. Chawan
Department of Computer Engineering & Information Technology
Veermata Jijabai Technological Institute, Matunga
Mumbai, India

***Abstract*— Recommender systems apply data processing techniques and prediction algorithms to predict users' interest on information, products and services among the tremendous amount of available things. The huge growth of information on the net still as vary as number of visitors to websites adds some key challenges to recommender systems. These are: producing correct recommendation, handling many recommendations expeditiously and managing huge growth of vary of participants at intervals the system. Therefore, new recommender system technologies are needed which is able to quickly manufacture top of the range recommendation even for huge information sets. Here our main goal is to 1) Make recommendation of NGO projects to volunteers 2) Make recommendation of volunteers to NGO.**

Keywords—Recommender system, collaborative filtering, content-based algorithm, profile centric matching.

1. INTRODUCTION

Recommender system or engine occurs for these several applications such as, e-shops, entertainment items (books, music, videos, Video on Demand, news, events, images etc.) or people (e.g. on dating sites) that are likely to be of interest to the user.

As ascertained by me, recommender system exists for several applications, but there is no recommender system on social facet. So, I am proposing recommender system for NGO's. A non-governmental organization (NGO) is that term normally used for a corporation that's is neither a district of a government nor a traditional for-profit business. Sometimes created by standard voters, NGOs could also be funded by governments, foundations, businesses, or non-public persons. Some avoid formal funding altogether and area unit run primarily by volunteers. NGOs area unit extremely various teams of organization engaged during a big selection of activities, and take completely different forms in in several components of planet. Some might have charitable standing, while others could also be registered for tax exemption supported recognition of social functions. Others could also be fronts for political, non-secular or alternative interest teams. The amount of NGOs within the United States is calculable at 1.5 million. Russia has 277,000 NGOs. Asian nations are calculable to possess had around 2 million NGOs in 2009, simply over one NGO per 600 Indians, and plenty of times the amount of primary facilities and first health centers in Asian nation. [1]

Here our goal is to 1) Make recommendation of NGO projects to volunteer of their interests. 2) Make recommendation of volunteers to NGO.

2. RECOMMENDER SYSTEM APPROACHES

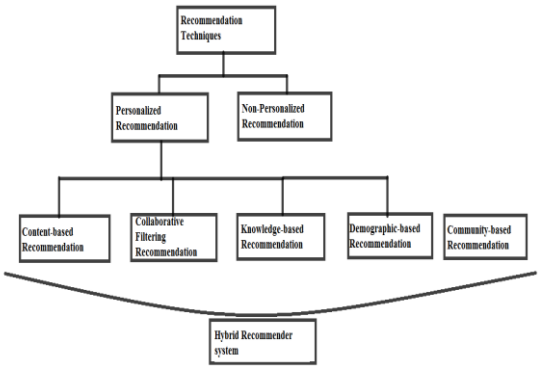

Fig. 2.1 Recommender system approaches

2.1 Personalized recommendation:

In personalized recommendation, recommendation is performed individually. One-to-one policy is followed. General workflow of personalized recommendation is, 1) System interacts with the user for information 2) Information is acquired from user either implicitly(via user transactional activity) or explicitly(via user inputs)3)Information pre-processing, that includes transformation of information of different types to the information required by recommendation algorithm. 4) Form recommendation after acquiring and pre-processing of information. 5) Show recommendation to users.
For example we can show personalized people recommendation on social networking sites. In

Facebook, friend recommendation feature known as, "People You will know", which recommends people to connect with based on "friend-of-friend" approach. Friend recommendation is different from e-commerce site product's recommendation.

2.2 Non-personalized recommendation:

The recommendations are independent of the user, so each user gets the recommendation as other user. Non-personalized recommender systems recommend items based on what other users have said about that item, independent of user, so, it does not require storing of user sessions and therefore recommendations are same for all. For instance, if I visit any site like amazon.com, since site does not store user session information, it will recommend items which other users have viewed. This is one of the non-personalized recommendation examples.

2.3 Collaborative filtering recommendation:

Collaborative filtering methods are based on collecting and analyzing an oversized quantity of knowledge on users' activities or preferences and predicting user's interest in similarity with other users. Collaborative Filtering are based on the assumption that people who agree in past will agree in future too and that they will like the similar kinds if items they like in the past. This is the reason why refers to collaborative filtering as "people-to-individual correlation." Within the newer, limited sense, cooperative (collaborative) filtering could be a technique of constructing predictions (filtering) about the interests of a user by collecting preferences or taste information from many users (collaborating). Consider that there are 2 persons that are like- minded, means there is a probability that person A's opinion match with person B's opinion on items. Collaborative filtering algorithms often require (1) users' active participation, (2) users' interests to the system, and (3) similar taste people matching. [3]

Problems with user-based collaborative filtering:

User cold-start problem:

Not enough known about new user to decide who is similar (and perhaps no other users yet).

Sparsity:

When recommending from an outsized item set, users can have rated just some of the items (makes it laborious to seek out similar users).

Scalability:

With millions of ratings, computations become slow.

Item Cold-Start problem:

Cannot predict ratings for new item till some similar users have rated it [No problem for content-based].

Item to Item collaborative filtering:

In item to item collaborative filtering prediction are generated based on similarities between items. The prediction for an item is based on the user's ratings for similar items.

User is likely to have the same opinion for similar items. Similarity between items is decided by looking at how other users have rated them. Two approaches can be used to compute the similarity between users; explicitly and implicitly.

Advantage (compared to user-based CF):

– Prevents User Cold-Start problem

– Improves scalability (similarity between items is more stable than between users).

Problems with user-based collaborative filtering:

Item Cold-Start problem:

Cannot pr

edict items similarity until we have ratings for this item.

2.4 Content-Based filtering:

Content-based recommendation systems include matching up users profile with item profile based on attributes (metadata) defined. First attributes of user and items are found, then user profile and item profile is constructed. And then similarity is calculated.

In content based filtering first text mining of document take place. In text mining important information/keywords are extracted from bulk information. Then, content based filtering is performed, i.e. filtering based on user given input to the system. Then, similarity calculation is done. And after that relevant results are displayed.[3]

The process of recommendation is performed in step by step manner, every of that is handled by a separate component:

- Content analyser
- Profile learner
- Filtering Component

2.5 Demographic based Recommendation:

Demographic based recommendation system recommend items based on user's demographic profile. Demographic profile of user includes location. Demography is all about statistical study of human populations. Demographic recommendation system may suffer cold start problem. Using Demographic information alone can achieve only limited accuracy. [4]

2.6 Knowledge-based recommender system:

Knowledge based recommender system is based on collecting explicit and huge amount of knowledge about items. This system is useful where collaborative filtering and content based filtering cannot be used. It considers attributes like user ratings, user preferences, user comments, user feedback etc. while recommending item. Knowledge based recommender system can be of types conversational, search based and navigational based recommendations. Since, it first collects all the knowledge about item, it does not suffer cold start problem.[4]

2.7 Community-based recommender system:
Community based recommender system recommends item based on or taking into consideration taste of other similar users. Recommendation can be based on ratings provided by other similar user. People more rely on groups or subgroups of similar user. Trust factor matters a lot in this system for critical decisions. It generally does not suffer cold start problem, because even if the user is new but similar users may exist.[4]

2.8 Hybrid recommender system:
Hybrid recommendation system combines two or more different techniques for recommendation. It generally combines collaborative filtering with other techniques. Due to combination drawbacks of one can be overcome by other and performance and accuracy of recommendation can be improved. [4]

3. DATA MINING RECOMMENDATION METHODS

Data mining methods most commonly used in RS are: Classification, Clustering and Association rule mining.

3.1 Classification:
Classification is a data analysis process in which application is classified into different models based on classes or features and labels. There are many example exist for classification. For eg. Consider bank application, classifying customers on the basis of loan accounts, whether giving loan to customer is safe or not. Again in general classifying customers on the basis of what types of computers they buy. In this way, Different classification models are built. Steps that take place in classification are data cleaning, data relevance, and data transformation and reduction.[4]

3.1.1 Nearest Neighbors:
Nearest neighbor classifier classifies data on the basis of statistics and not parameters. In this object is classified and assigned to one of the class from nearest neighbor. Neighbors are taken from object property. In K-NN, K is user-defined constant. K is decided on the basis of data. Larger values of k reduce effect of noise in classification but makes boundaries between classes less distinct. Finding similar taste users is like finding nearest neighbors for users. Advantages are statistical and non-parametric, simple, powerful and require no training time. Drawbacks are memory intensive and classification is slow. [4]

3.2.2 Decision Trees:
Decision tree is graph based classification model in which decisions are made at each level of tree. Primary decision is supposed to be taken at root node (topmost node). Each inner node is tested for attribute value or input variable. In decision tree structure, leaf nodes represent classification. Inner nodes represent predictive attributes and braches represent decisions to be taken that lead to final classification. The popular decision trees algorithms include ID3, C4.5 which is an extension of ID3 algorithm and CART. [4]

3.2.3 Ruled-based Classifiers:
Rule based classifier classifies data based on rules. We define a rule and we say rule covers particular instance if attributes of that instance satisfy rule. We say rules are mutually exclusive if every record is covered by at most one of the rules defined. We say rules are exhaustive if every record is covered by at least one of the rues defined. It is very difficult to build complete recommendation model using rule based classifier but it can help to improve accuracy by adding some domain knowledge or business rules. [4]

3.2.4 Bayesian Classifiers:
Bayesian classification is a probabilistic approach for classification based on Bayes theorem. It is statistical classifier. If given record belongs to particular class or not is decided by Bayesian classifier. There are two types of probabilities 1. Posterior probability 2. Prior probability.

3.2.5 Artificial Neural Networks:
An Artificial Neural Network (ANN) is an assembly of inter-connected nodes and weighted links that is inspired in the architecture of the biological brain. Nodes in an Artificial neural network (ANN) are called neurons as an analogy with biological neurons. These simple functional units are composed into networks that have the ability to learn a classification problem after they are trained with sufficient data.
ANN's can be used in a similar way as Bayesian Networks to construct model based
RS's.[4]

3.2 Clustering:
Clustering, also referred to as unsupervised learning, consists of assigning items to groups so that the items in the same groups are more similar than items in different groups: the goal is to discover natural (or meaningful) groups that exist in the data. The goal of a clustering algorithm is to minimize intra-cluster distances while maximizing inter-cluster distances. There are two main categories of clustering algorithms: hierarchical and partitioned. Partitioned clustering algorithms divide data items into non-overlapping clusters such that each data item is in exactly one cluster. Hierarchical clustering successively cluster items within found clusters, producing a set of nested cluster organized as a hierarchical tree.
K-Means clustering is a partitioning method. The function partitions the data set of N items into k disjoint subsets Sj that contain Nj items so that they are as close to each other as possible according a given distance measure. Each cluster in the partition is defined by its Nj members and by its centroid cj. There are some shortcomings of K-

means clustering algorithm, so, alternatives for k-means is, Density-based clustering, Message-passing clustering, hierarchical clustering.[4]

3.3 Association rule mining:

Association Rule Mining focuses on finding rules that will predict the occurrence of an item based on the occurrences of other items in a transaction. The fact that two items are found to be related means co-occurrence but not causality.[4]

4. PROPOSED SYSTEM ARCHITECTURE

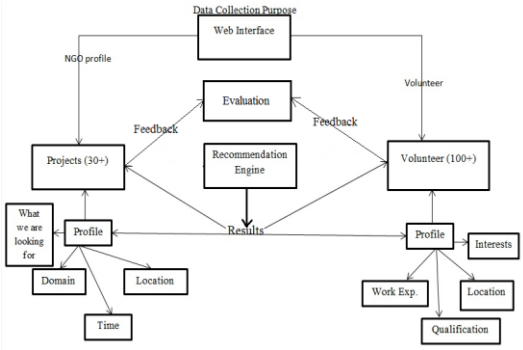

Fig. 4.1 Proposed system architecture

Steps to be followed in implementation:

Step 1: Studying the current suggested techniques.

Step 2: Collection of required data and preparing a dataset. Either by manually or through web interface.

Step 3: Selecting the Input and output variables with the use of previous studies. Determining the membership functions for the variables. Determining the relevant attributes to split the dataset. Specifying rules to make clear the relations between Inputs and outputs. Developing the algorithm. Running recommendation algorithm considering some aspects of data attributes.

Step 4: Verify recommendation results in real time.

Step 5: Evaluating recommendation algorithm on the basis of feedback given**.**

Objectives:

1. Recommending NGO projects to volunteer based on inputs given by volunteers plus keeping track of user's activity.

2. Recommending volunteers to NGO based on inputs given by NGO.

3. Recommending NGO projects based on other user's activity.

4. Get information of all NGO's with filtration as per user's input at just one click.

5. Get information of all Volunteers with filtration as per user's input at just one click.

General approach towards the proposed system:

a) Find the attributes of user and items (projects).

b) Construct interest profile of the user.

c) Use this profile to rank-order the unseen items by similarity to the profile, thereby approximately possible interest in those items.

Algorithmic approach for proposed system:

a) Collate all users' assigned metadata into a single user
profile (By aggregating all of the metadata assigned by a user, his/her interest can be completely captured)

b) Similarly, construct item profiles that collate all of the
metadata assigned to those items by all users in the training set.

c) Match active user profiles against the item profiles
on similarity to produce ranking of all items.

d) Remove items already in the active user's profile &
get the final recommendation test.

5. CHALLENGES AND ISSUES

5.1 Cold-start:

Whenever new user creates his/her profile, their profile is almost empty, so, it becomes difficult to recommend any item to such a user, whose taste is unknown to the system. This problem is called cold-start problem. This problem can be solved by doing survey while creating a profile, to get at least some information about user's taste. Same problem occurs with newly added items, which is not rated by anyone yet. By using some hybrid approaches, this problem can be solved.

5.2 Trust:

Evaluation of a customer is done through the history of the user profile. Some user profiles contain more history and some are poor in history. One way to solve this problem is to distribute the priority to the users.

5.3 Scalability:

As the number of users and items increases, the requirement for resources by the system for processing of information and forming recommendation also increases. Most of the resources are consumed for finding out users with similar behavior and taste. This problem can be solved by using different types of filters and improving physical physical state of the system.

5.4 Sparsity:

In any application where more number of users and items are involved, sometimes it happens that some users rate less number of items.

6. CONCLUSION AND FUTURE WORK

In this paper many collaborative and content-based recommendation approaches are presented and briefly described. Many recommendation methods are also described.

Proposed architecture is also presented. Proposed system works based on content-based matching algorithm. In future this can be extended to collaborative filtering, demographic recommendation, community- based recommendation, and hybrid recommender system. Recommender systems are powerful systems that give an added-value to business and corporation. They are a relatively recent technology and they will only keep improving in the future.

7. REFERENCES

[1] Recommender Systems in E-Commerce, J. Ben Schafer, Joseph Konstan, John Riedl, GroupLens Research Project Department of Computer Science and Engineering University of Minne sota Minneapolis.

[2] http://www.slideshare.net/VikrantArya/recommendation-system-33379953

[3] Algorithms and Methods in Recommender Systems, Daniar Asanov Berlin Institute of Technology Berlin, Germany

[4] Recommender Systems Handbook, Francesco Ricci, Lior Rokach, Bracha Shapira, Paul B. Kantor

[5] Relevant Words Extraction Method for Recommendation System, International Journal of Emerging Technology and Advanced Engineering, (ISSN 2250-2459, ISO 9001:2008 Certified Journal, Volume 3, Issue 1, January 2013