Expert Based Prediction of User Preferences

M. Dima
Department of Informatics
National and Kapodistrian University of Athens, Greece

D. Vogiatzis
Institute of Informatics and Telecommunications,
National Centre for Scientific Research ”Demokritos”, Greece

American College of Greece, Deree, Greece

G. Paliouras
Institute of Informatics and Telecommunications
National Centre for Scientific Research ”Demokritos”, Greece

P. Stamatopoulos
Department of Informatics
National and Kapodistrian University of Athens, Greece

Abstract

Two influential strands in Recommender systems (RS) are the collaborative filtering and content based filtering that by taking into account user communities or interaction history suggest to the active user interesting items. However, the aforementioned approaches do not work well when confronted with new users with few interactions; or with the addition of new items. In such cases, the guidance of an expert could help the active user. In this paper we provide a definition of expert users *that can be reduced into two components the* expertise *and the* contribution*. The former is related to the content of items evaluated by an expert and the latter refers to the influence of the expert to the users of a RS. In particular, contribution is learnt with the aid of a perceptron. Experts users are defined for values of the features of the items. Furthermore, we have studied the* temporal *evolution of the experts, as new users, new items, or new item evaluations are added into the system. Moreover, we have compared the proposed expert based method with a stereotype based method, since for both methods a minimal interaction of the active user with the RS suffices. The data originated from the MovieLens set with enhancements from the IMDB.*

1 Introduction

The overabundance of information is well known problem of the information age, first attested many years ago. Given the current limits in the human cognitive abilities, a moderate amount of information can be digested at any given time. The field of information retrieval along with machine learning and user modeling aims to aid users into receiving exactly (or close to that) the information they need. Three basic approaches are followed, directories (such as dmoz [1]), search engines, (e.g. Google) and recommender systems (e.g. NetFlix [2]). A recommender system, based on certain assumptions, aims to suggest items (be they products or services), which to certain extent are of interest to a user. It is reasonable to assume that the past user's interests determine his future interests. Another approach, is that a resemblance of a user to an existing user group, with respect to certain common interests, suggests that the user inherits all the interests of that group. These two approaches reflect the content based and the collaborative based filtering approaches (see [1] for an overview of recommender systems).

As mentioned, in collaborative filtering we seek a community of similar users, usually with respect to past interactions with some items. If we could dispense for a mo-

[1] http://www.dmoz.org/
[2] http://www.netflix.com/

ment with similarity issues, and instead seek users that we can trust or expert users who can reliably distinguish between low quality and high quality items, then we could have some unique advantages when compared with simple collaborative systems. For instance, collaborative systems are based on the assumption that the user has already expressed interest in the products, or services that are offered. This assumption is not valid in the case of a novel user, or of a user, which has never evaluated an item of a specified category. For instance how could a user that has only evaluated adventure movies, obtain reliable information about drama movies in a movie recommender system? On the other hand an expert user could provide recommendations (at least in the categories of his expertise) irrespective of his similarity to other users. Second, the matrix that represents users and their ratings regarding the items of the RS is sparse, and thus it is not expected that there will be a large number of users sufficiently similar to the current user. Third, experts can be defined in ways that are focused on the features of items of the RS, and thus potentially more reliable; for instance if the item are movies then we could define experts on movie genres, on the actors etc. The concept of trust, or expert users is active field of research, and different approaches have been followed to address this issue (see [7] for an overview).

Paper contribution We have developed an expert based prediction method for the active user's preferences, In particular, experts have *expertise* and *contribution*. Expertise is defined in terms of the characteristics of the items the expert has interacted with, whereas contribution is based on the influence of the expert. Moreover, experts are defined for specific feature values of the application domain, and since our domain concerns movies, there are experts for specific actors, directors, genres, and movie tags or keywords. In addition, we studied the temporal evolution of experts, as new items, new item evaluation or new users are added. Finally, we compared our method with a stereotype based method on data from MovieLens dataset enhanced with data from IMDB.

The rest of the paper is organised as follows: In Section 2, we proceed with an exposure of expert or trust based methods in recommender systems. In Section 3, we expose our method; we define the experts and in particular the expertise and the influence, followed by the usage of experts for predicting a user's preferences. The experimental setting is described in Section 4, followed by results in Section 5. Finally conclusions are drawn in Section 6.

2 Trust in recommender systems

One way to define trust is in terms of item ratings and their differences from average ratings. For instance, given the rating R_{ui} of user u to item i, the average rating $\bar{R}_u$ of user u over all items, the average rating $\bar{R}_i$ of item i over all users, and the average rating of all users over all items $\bar{R}$ and a constant β, then trustworthiness of user u over all items he has evaluated is defined as follows [2]:

$$trust = \beta \left| \frac{\sum_{i \in S_u} (R_{ui} - \bar{R}_u)(\bar{R}_i - \bar{R})}{\sqrt{\sum_{i \in S_u} (R_{ui} - \bar{R}_u)^2 \sum_{i \in S_u} (\bar{R}_i - \bar{R})^2}} \right| \quad (1)$$

Trust can also be defined in terms of the accuracy of user's u prediction (or P_{uv}) of user's v rating $R_v(i)$ of item i. In this case, users' satisfaction is the central issue. Thus the trust of user's u for user v regarding item i is high if $|P_{uv}(i) - R_v(i)| < \epsilon$. This definition can be extended to express the trustworthiness of user u regarding item i in the whole user community or even the general trustworthiness of user u regarding all items and all users [6].

There are also efforts to combine collaborative filtering with trust based recommendations. For example, two types of communities can be defined: the one of similar users and that of expert users, and they are both used in suggesting items of interest to the active user (see [3] for the DISCORS system).

Moreover, dispensing with formal definitions of trust, and allowing users to rate each other based on personal notions of reliability and trustworthiness is common in reputation systems and social networks [5], [4].

3 Expert Based Prediction

The definitions are drawn from the domain of movies, but they bear general relevance. First, we assume the existence of a database of items (i.e. movies), that have been evaluated by users; the matrix of user evaluations and movies is sparse. Second, each item can decomposed into a number of features such as genre, actors, directors and keywords. Each feature assumes a number of feature values (for example the genre category assumes a subset of the values: drama, comedy etc.). In addition, there is demographic information regarding the users (age, gender, profession), and a time stamp for each user evaluation.

Experts will be discovered for each feature value (e.g. for a specific actor, or for a specific director); because of the large number of feature values and the subsequent computational complexity some sort of feature value selection must be implemented so as to keep the *important values*. In the context of the current work, we have suggested three plausible ways of feature selection:

1. Preferred features' values (Pr): feature values that receive high evaluations.

2. Popular features' values (Pop): feature values that are frequently evaluated.

Table 1. Method for expert based prediction

Preliminary Steps
Experts will be discovered for *important* feature values 1. Preferred feature values (Pr) 2. Popular feature values (Pop) 3. Extreme feature values (E)

Experts definition
Experts are the users that maximise: *expertise*$\times$*contribution* expertise is defined in terms of the item features contribution is defined in terms of the influence of that expert to users

Expert based Prediction
Prediction of user's u preference for an item depends on the user's u average rating of all items and on the voting of each expert relevant to a feature value of the current item according to equation 6.

3. Extreme features' values (E): feature values that receive very high, or very low evaluations.

Let us assume there is an item m whose rating is to be predicted for user u, the role of experts in predicting the user's preference is as follows:

- detect the *important* feature values for the item m
- discover the experts for those values
- experts will denote their preferences for values for which they are experts
- the preference of each expert is represented by the product: *expertise* $\times$ *contribution*, which is analysed next

3.1 Experts definition

Experts are characterised by two values, the *expertise* and the *contribution*. The expertise is somewhat objective being derived from the features of the items of the application domain, whereas contribution reflects the influence of the expert on users, and is somewhat subjective.

3.1.1 Expertise

The expertise being one of the components that characterise an expert, as is calculated for *selected users* and for *selected (important) feature values*. There are definitions of expertise that could be domain independent as mentioned in the literature review. However, he have chosen to define expertise in ways that incorporate domain knowledge. In the current movie application domain, we have selected 4 feature categories: genres, actors, directors, and tags. Experts will discovered for specific actors, directors, movie genres or tags. The users interacting heavily will be selected for expertise calculations. The definitions of expertise are as follows:

genres, keywords The definition incorporates number of movies evaluated in the past that fall into a specific genre or contain a specific keyword, as well as the popularity of that feature value with respect to other feature values. It was assumed that the rarer a movie is, the more expert user u is considered: Thus for feature values $g \in \{\text{genres, keywords}\}$ the expertise of user u is:

$$expertise_{ug} = \frac{r_g}{m_g^2} \sum_i \frac{1}{r_i} \tag{2}$$

where r_g is the number of raters of movies with feature value g, m_g is the number of movies with feature value g, i is an index to the movies seen by user u, and r_i represents the raters of movie i.

actors, directors The definition incorporates the number of movies evaluated including length of acting or directing period, their role in the movies, and the variance of the movie genres of the actor or the director. For actor a the following definition of expertise of user u holds:

$$expertise_{ua} = 2\frac{movies\_acted(S_u)}{movies\_acted} + \frac{acting\_period(S_u)}{acting\_period} + \frac{mean\_credit\_postion(S_u)}{mean\_credit\_position} + \frac{genres\_acted(S_u)}{genres\_acted} \tag{3}$$

where S_u denotes the movies seen by user u.

For director d the following definition of expertise of user u holds:

$$expertise_{ud} = 2\frac{movies\_directed(S_u)}{movies\_directed_d} + \frac{directing\_period(S_u)}{directing\_period_d} + \frac{genres\_directed(S_u)}{genres\_directed_d} \tag{4}$$

3.1.2 Contribution

As mentioned, the contribution of an expert is the influence that he exercises over users of the system. We assume the existence of a linear function from the domain $\mathcal{E}$ of expertise to the domain of item predictions (or evaluations) $\mathcal{P}$.

Thus, $f : \mathcal{E} \times \mathcal{C} \rightarrow \mathcal{P}$, where $\mathcal{C}$ is the contribution. For each item (movie in our case) of the database, the best experts for the selected feature values will be discovered. They will contribute their expertise, which will be multiplied by the relevant contributions, then summed up to produce a prediction of a user's output. Contribution, is not constant but it is learnt with the aid of a perceptron, where the inputs represent expertise, and the weights contribution.

Let us elaborate on the above succinct description of contribution discovery. The perceptron's inputs x_i are the weighted preferences of the experts regarding feature values. Thus, $x_i = expertise_i \times (R_{e_i} - \bar{R}_{e_i})$, where $expertise_i$ represents the expertise value for the selected feature value, R_{e_i} ($\bar{R}_{e_i}$) is the rating (average rating) of the expert for the feature value in question. All the x_i for the current movie form a training vector.

Assuming the existence of experts for the selected feature values of item m, and i an index to the experts space, then the perceptron's output is as follows,

$$o = \frac{\sum_i x_i w_i}{\sum_i |expertise_i \times w_i|} \quad (5)$$

where w_i is the contribution of expert i.

The target output is the normalised rating of the user u for which the prediction of item m is aimed at: $t_{um} = R_{um} - \bar{R}$, where $\bar{R_u}$ is the average rating of the user.

The difference between the target output and the perceptron's output constitutes the error $error_{um} = t_{um} - o_{um}$, which is used to train the perceptron with the stochastic gradient descent algorithm: $w_i = w_i + \eta * (t_{um} - o_{um}) * x_i$.

3.2 Predicting active user's preferences

Prediction of active user's u preference for item m, denoted as P_{um} will be performed with experts. First, the experts for each feature value of m are discovered; then, each expert will vote according to his expertise and contribution. The steps are summarised in Table 1. In addition, the average movies' rating R_u of the active user u is taken into account. Thus,

$$P_{um} = \bar{R}_u + \frac{\sum_{i=1}^{n} expertise_i (R_{e_i} - \bar{R}_{e_i}) contr_i}{\sum_{i=1}^{n} |expertise_i \times contr_i|} \quad (6)$$

where $expertise_i$ and $contr_i$ represent the expertise and the contribution of expert i; R_{e_i} ($\bar{R}_{e_i}$) is the rating (average rating) of expert; i is an index over the selected experts. The second term of the above equation is the perceptron's output.

4 Experimental Setting

The purpose of the experiments is to test the prediction accuracy of methods that apply primarily to novel users, with few interactions. The methods under study are: the expert based method that we have proposed, and another one that associates user demographics with user preferences, named stereotype based method.

Another point of the study aimed at the temporal evolution of experts. In particular, we studied the evolution of the product *expertise* $\times$ *contribution* in time, which we have named *experience*. Moreover, we have studied the addition of users that become experts in the passage of time. We have selected 500 extreme feature values, i.e. according to our definition, feature values that receive very high or very low ratings. The total number of experts has been set to 2000. As explained before, experts are defined for specific actors, directors, movie genres or movie keywords, and one expert could exhibit expertise in many feature values.

The data sets that we have employed is the MovieLens [3] and the IMDB [4]. The MovieLens data set is comprised of 100,000 evaluations for 1682 movies by 943 users. Each rating is accompanied by a time stamp, and the whole data set covers a period of 7 months (20/09/1997-23/04/1998). There is also, demographic information for each user (gender, age, profession and postal code); and each movie is characterised by its title, release date, and it belongs to one or more genres. We used the IMDB data to enhance MovieLens with extra information about movie genres, tags, actors, directors, and other things. In the current experiments movie genres, keywords, actors and directors as well as user's demographics were used. The data have been split into 27 time periods of equal length.

Expert based prediction

The predictions are calculated according to equation 6, that is the method we have introduced. In particular to calculate the contribution, a perceptron is trained on 75% of the data available for a specific time period, the rest is used for testing. The trained perceptron is then used for the next time period, where it is re-trained with the available data.

Stereotype based prediction

The users of the movie lens data set can be grouped according to their demographics (i.e. age, gender and occupation). Then the preferences of each demographic group can be discovered, according to a formula that we describe next. The association of a demographic group with feature value preferences creates a stereotype.

First, 30 demograpic groups were formed as follows: 5 groups combining gender and age,

```
gender=f and age<18
gender=f and age>=18 and age<25
```

[3] www.grouplens.org/node/73
[4] http://www.imdb.com/interfaces

```
...
gender=f and age>=56
```

similarly, 5 demographic groups were defined for the male gender. Finally, 20 demographic groups were defined to include occupation only.

The prediction of the ratings of the new item m for user u is based on the average rating of user u, and on the feature value preferences of his group associates,

$$P_{um} = \bar{R}_u + \frac{\sum_{f \in IPF_m} (\sum_{d \in D_u} \frac{pref_{df}}{|d \in D_u|})}{|f \in IPF_m|} \qquad (7)$$

where $pref_{df}$ represents the preference of demographic group or stereotype d regarding feature value f, U is the set of all system users, IPF_m is the set of important feature values of item m, and UPF_u is the set of important feature values of item m with which user u has interacted. D_u are the demographic groups user u belongs to

$$pref_{df} = \sum_{v \in U \bigcap d,\ f \in UPF_v} \frac{\bar{r}_{vf} - \bar{R}_v}{|v \in U \bigcap d,\ f \in UPF_v|} \qquad (8)$$

5 Results and Discussion

In figure 1 we depict the prediction accuracy P_{um} (of active user's u preference of item m) for the stereotype and the expert based methods. Comparison is based on Root Mean Square Error (RMSE) and it is caclulated for the 27 time periods. As it can be seen the expert based method, performs better than the stereotype method. The differences in prediction of the two methods is 0.40.

In figure 2 we depict the additon of new experts for the 27 time periods, for directors, keywords, actors and genres that concern the movies. As it can be seen, the addition of new experts is decreased. Moreover, in figure 3 the experience, or the product of expertise $\times$ contribution is depicted for directors, keywords, actors and genres. For the keywords, and there is a substantial increase in the experience, it is less for genres, and slight for actors and directors.

We should note that in the comparison we have included methods that propose interesting items to the active user, largely irrespective of his interactions, for it was the point of the paper to focus on the cold start problem. For the given stereotypes the expert based method is better. But it could be further improved by assuming the existence of a non-linear relation between the expertise and the prediction of the expert; whereas so far a linear one was assumed, which facilitates the computation of contribution 5. Admittedly, the definition of stereotypes could be improved to better reflect more realistic groupings of users. This could be achieved by machine learning methods.

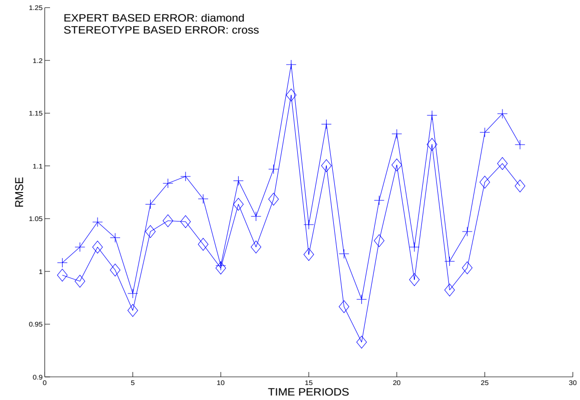

Figure 1. RMSE for each period for the four methods

6 Conclusions

We have proposed an expert based method for predicting user's preferences with respect to items, movies in particular. Each expert specialises in a movie feature value (e.g. an actor, a director, a movie genre etc.) and is deemed as possesing extensive knowledge. For practical reasons relating to computational speed and space, the experts are not defined for every feature value, but for selected or important ones; i.e. for the preferred, popopular and extreme feature values, as defined. In addition the experts' prowess or experience is reduced into the expertise and contribution components. The first one is calculated based on domain specific knowledge. The second one, depends on the influence the expert exercises on the existing users. The contribution variable is calculated with the aid of perceptron, where the input is the expertise, the weights represent the contribution, and the output represents the prediction of movies' ratings to users. The main reason for the introduction of the expert based method is in situations where there are few interactions for the active user, or when a new item has been introduced. These are cases of the cold start problem.

We have compared the expert based method for predicting user preferences, with a stereotype based method. The second one, is based on user groups that have common demographic characteristics. Also, we have studied the temporal evolution of experience as well as the addition of new experts as more transactions enter the database of the recommender system. We have observed, an increase in the experience, and a decrease in the addition of new experts. The increase in experience has been observed for keywords, and genres, which are semantically close.

In the future, we plan to integrate the expert based

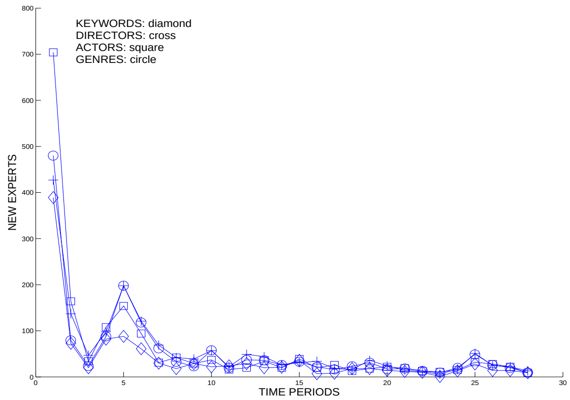

Figure 2. Addition of new experts over time periods

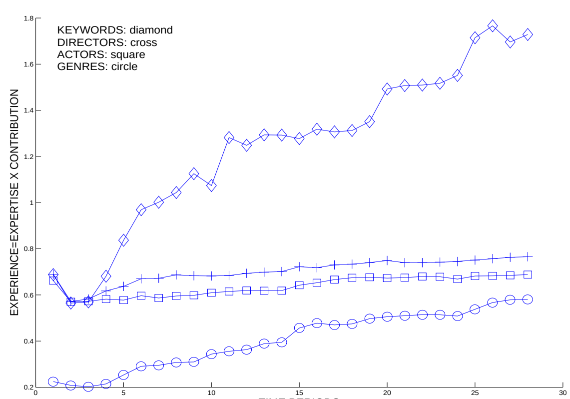

Figure 3. Evolution of experience over time

method with the stereotype method. A weighting scheme will be needed to balance the two methods. Second, the aformention approaches will probably be incorporated in switching scheme with collaborative based methods. The gist is to give emphasis to experts or stereotypes for new user with few ratings, and to collaborative filtering later on. Also, stereotypes that are more complex to the ones we have designed could be discovered with machine learning techniques.

Another assumption to investigate is the linear relationship between the expertise and the prediction as stipulated by the perceptron. Non-linear relations could be better models, although more difficult to interpret.

Acknowledgement

This work was carried out in the context of project CP-TP 214455-2 SERVice Oriented Intelligent Value Adding nEtwork for Clothing-SMEs embarking in Mass-Customisation (SERVIVE), which is partially funded by the European Union.

References

[1] G. Adomavicious and 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*, 17:734–749, 2005.

[2] J. Cho and K. Kwon. Source Credibility Model for Neighbor Selection in Collaboarative Web Content Recommendation. In Y. Zhang, G. Yu, E. Bertino, and G. Xu, editors, *Progress in WWW Research and Development (APWeb), LNCS*, pages 68–80. Springer-Verlag, 2008.

[3] J. Cho, K. Kwon, and Y. Park. Collaborative Filtering Using Dual Information Sources. *IEEE Intelligent Systems Computer*, 22(3):30–38, 2007.

[4] A. Jsang, R. Ismail, and C. Boyd. A Survey of Trust and Reputation Systems for Online Service Provision. *Decision Support Systems*, 43(2):618–644, 2007.

[5] L. Mui, A. Halberstadt, and M. Mohtashemi. Notions of Reputation in Multi-Agents Systems: A Review. In *Proceedings of the First International Joint Conference on Autonomous Agens and Multiagent Systems: Part 1 (AAMAS)*, pages 280–287, 2002.

[6] J. O'Donovan and B. Smyth. Trust in recommender systems. In *Proceedings of the 10th International Conference on Intelligent User Interfaces (IUI'05)*, pages 167–174, 2005.

[7] Y. Wand and H. Emurian. An overview of online trust: Concepts elements and implications. *Computers in human behaviour*, 21(1):105–125, 2005.