

Target Tracking in the recommender space - Toward a new recommender system based on Kalman filtering

Samuel Nowakowski, Anne Boyer, Cédric Bernier

▶ **To cite this version:**

Samuel Nowakowski, Anne Boyer, Cédric Bernier. Target Tracking in the recommender space - Toward a new recommender system based on Kalman filtering. INTERNET 2011, Jun 2011, Luxembourg, Luxembourg. pp.non connus. hal-00591812

HAL Id: hal-00591812
https://hal.science/hal-00591812

Submitted on 10 May 2011

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Target Tracking in the recommender space

Toward a new recommender system based on Kalman filtering

Samuel Nowakowski – Anne Boyer
LORIA-KIWI
Campus scientifique – BP239
F-54506 Vandœuvre-lès-Nancy Cedex
Samuel.nowakowski@loria.fr
Anne.boyer@loria.fr

Cédric Bernier
Alcatel-Lucent Bell Labs France – LORIA-KIWI
Route de Villarceaux
F-91600 Nozay
Cedric.bernier@loria.fr

***Abstract*— We assume that users and their consumptions of television programs are vectors in the multidimensional space of the categories of the resources. Knowing this space, we propose an algorithm based on a Kalman filter to track the user's profile and to foresee the best prediction of their future position in the recommendation space. From this prediction, we build a recommendation of contents.**

Keywords-recommender system; user profile; group profile; Kalman filter; target tracking

I. INTRODUCTION

In Web-based services of dynamic content, recommender systems face the difficulty of identifying new pertinent items and providing pertinent and personalized recommendations for users.

Personalized recommendation has become a mandatory feature of Web sites to improve customer satisfaction and customer retention. Recommendation involves a process of gathering information about site visitors, managing the content assets, analyzing current and past user interactive behavior, and, based on the analysis, delivering the right content to each visitor.

Recommendation methods can be distinguished into two main approaches: content based filtering [9] and collaborative filtering [10]. Collaborative filtering (CF) is one of the most successful and widely used technology to design recommender systems. CF analyzes users ratings to recognize similarities between users on the basis of their past ratings, and then generates new recommendations based on like-minded users' preferences. This approach suffers from several drawbacks, such as cold start, latency, sparsity [11], even if it gives interesting results.

The main idea of this paper is to propose an alternative way for recommender systems. Our work is based on the following assumption: we consider Users and Web resources as a dynamic system described in a state space. This dynamic system can be modeled by techniques coming from control system methods. The obtained state space is defined by state variables that are related to the users. We consider that the states of the users (by states, we understand « what are the resources they want to see in the next step ») are measured by the grades given to one resource by the users.

In this paper, we are going to present the effectiveness of Kalman filtering based approach for recommendation. We will detail the backgrounds of this approach i.e. state space description and Kalman filter. Then, we expose the applied methodology. Our conclusion will give some guidelines for future works.

II. PRINCIPLES

Kalman filter is an optimal state estimator of a linear system. It can estimate the state of the system using a priori knowledge of the evolution of the state and the measurements.

A. *Target tracking in the cyberspace*

Hypothesis: the user is a target which is moving along a a priori unknown trajectory in the multidimensional space of the categories. Figure 1 shows the principle of our approach.

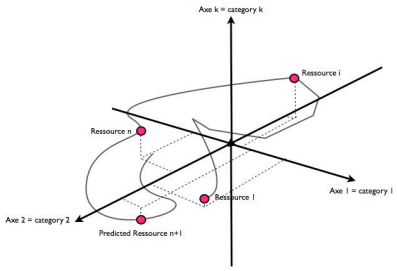


Figure 1. Trajectory in the recommender space

B. *Kalman filter: equations*

How can we know about a target moving in the recommender space ?

Using its position, speed and acceleration, we choose as the state vector the following form:

$$X_k = \begin{bmatrix} x \\ \dot{x} \\ \ddot{x} \end{bmatrix}_k \quad (1)$$

where :

x is the vector containing the position vector

$\dot{x}$ is the vector containing the speed vector,

$\ddot{x}$ is the vector containing the acceleration vector.

The estimation of this state vector will give the necessary knowledge of the trajectory in the recommender space. We use the following state space model:

$$\begin{cases} X_{k+1} = AX_k + w_k \\ Z_k = HX_k + v_k \end{cases} \quad (2)$$

where matrix A includes the relationship between the position, speed and acceleration.

$$A = \begin{bmatrix} \alpha & T & \frac{1}{2}T^2 \\ 0 & \alpha & T \\ 0 & 0 & \alpha \end{bmatrix} \quad (3)$$

Matrix H will have the following structure as shown in the figure 2.

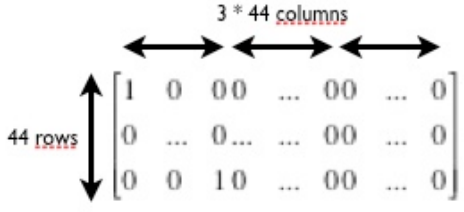


Figure 2. Structure of Matrix H

The Kalman filter equations are then given [6]:

Prediction

$$\begin{cases} \hat{X}_{k+1/k} = \hat{X}_{k/k-1} + K_k\left(Z_k - H\hat{X}_{k/k-1}\right) \\ \qquad = \left(A - K_k C\right)\hat{X}_{k/k-1} + K_k Z_k \end{cases} \quad (4)$$

Kalman gain:

$$K_k = AP_{k/k-1}H^T\left(HP_{k/k-1}H^T + R\right)^{-1} \quad (5)$$

The evolution of the uncertainty on the estimation is then given by:

$$P_{k+1/k} = AP_{k/k-\&}A^T - AP_{k/k-1}H^T\left(HP_{k/k-1}H^T + R\right)^{-1}HP_{k/k-1}A^T \quad (6)$$

where the initial conditions are given by:

$$\hat{X}_{0/-1} = X_0\ ,\ P_{0/-1} = P_0$$

and the state prediction by:

$$\hat{X}_{k+1/k}$$

III. APPLICATION

A. Description of the experiment

This experiment is based on TV consumption. The dataset is the TV consumption of 6423 English households over a period of 6 months (from 1st September 2008 to 1st March 2009) (Broadcaster Audience Research Board, [7]), [8]. This dataset contains information about the user, the household and about television program. Each TV program is labelled by one or several genres. In the experiment, a user profile build for each person. The user profile is the set of genres associated to the value of interest of the user for each genre. This user profile is elaborated in function of the quality of a user's TV consumption: if a TV program is watched entirely, the genre associated to this TV program increases in the user profile. Several logical rules are applied to estimate the interest of a user for a TV program.

The methodology of the experimentation is the following:

- Each user profile is computed at different instants (35) from the TV viewing data.
- The Kalman filter is applied iteratively to estimate the following positions of the user profile in the space of the genre.

The entire consumption is described by 44 types which will define the 44 dimensions space where users are "moving".

B. Numerical results

The obtained results can be exposed as follows:

- Kalman filter predicts the interest of a specific user for one gender knowing his past.

Using this prediction, we can propose a new recommendation strategy:

- If the Quantity of Interest (QoI) of the user is predicted to be in one specific region of the space, we can recommend something inside this specific region:
- For example, if the specific region is defined by dimensions Documentary and Drama, we can recommend contents related to these two dimensions
- If the predicted quantity of interest (QoI) changes to another dimension of the space, we can automatically recommend content from this new region of the space.

C. Results

The results can be analyzed as follows: Kalman filter predicts the specific interest for a category of contents of one user.

Figures 3 and 4 shows Estimation / Prediction computed by Kalman filtering. Doted-lines show the evolution of the real values. Continuous lines show the obtained predictions.

In the figures given below, we can see the estimation/prediction given by the Kalman filter. The green lines show the prediction obtained at each time using the knowledge we have of the degree of interest of each user.

Figure 5 shows the results of the cosine distance which has been computed between the true values and the prediction by the filter.

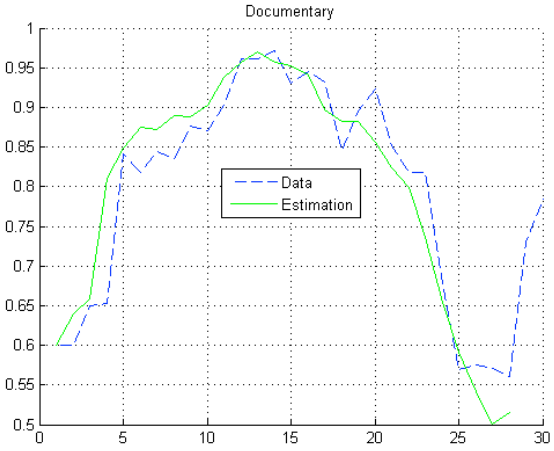


Figure 3. Prediction for Drama

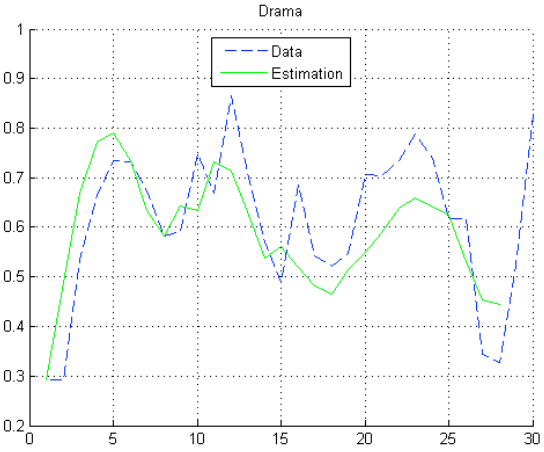


Figure 4. Prediction for Documentary

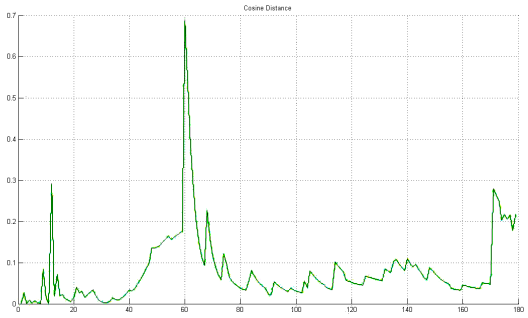


Figure 5. Cosine distance

IV. RECOMMENDATION STRATEGY

In this approach, we can build a recommendation by analyzing the estimation provided by Kalman filter.

The profile is built from the consumption of TV programs. Each TV program is defined by concepts such as entertainment, science fiction, talk show, etc. The analysis of the way different TV programs are watched allows deducing the interest of a user for each concept. Hence, the user profile is calculated from the TV consumption and it is represented by a vector of valuated concepts.

The user profile is considered as a point in the space of the concepts (in our case, 44 dimensions). This point moves at each different time in the space and so forms a trajectory. With the Kalman filter, we estimate the next position of the user profile. The estimation shows the evolution within each dimension, hence for each concept.

For our new recommending strategy (see Figure 6), we observe the difference between the estimated concept and the calculated concept. If the calculated concept is superior to the estimated concept (noted negative difference), then the user's interest for this concept is decreasing. On the contrary, if the estimated concept is superior to the calculated concept (noted positive difference), then the user's interest for this concept is increasing. We concentrate on the concepts showing up a big difference: the concepts with an important positive difference influence the recommendation towards these concepts, whereas the concepts with an important negative difference discourage the recommendation towards these concepts.

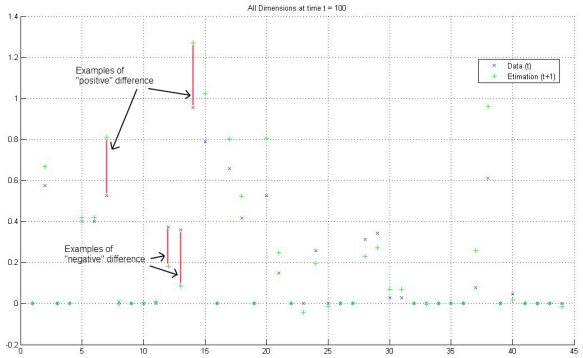


Figure 6. Analysis of the evolution of the prediction for recommendation

Conversely to existing methods which recommend precise contents for a given user, this method takes into account the user's state of mind. Our method performs on the macroscopic level. We find out the type of content the user appreciates and can determine some dimensions that can deliberately be closed out.

The recommendation is based on the two preceding arguments.

- the user's actual state of mind
- the subset of retained dimensions.

From these “positive” or “negative” dimensions and from the TV program, we have to define the recommendation for a set of TV programs for that day. Furthermore, according to what the user watched during the day, we can refine our recommendation. Indeed, in our example, if the user is interested in contents of types x, y and z and if he has already watched content of type x and y that day, the recommendation would essentially concentrate on content of type z.

Hence we need to make a last step which is to find a content which corresponds to the estimation of the dimensions' evolution.

V. CONCLUSION

In this paper, the main idea is to consider that the one who chooses films as a target which moves along a trajectory in the recommender space. The recommender space is seen as a 44 dimensions spaces based on the main concepts describing the films. The position of the target is measured by the index of interest expressed for each concept. Then the Kalman filter applied using a tracking model predicts the “positions” in the recommender space.

Then, knowing the past positions of the user in this space along the different axis of the 44 dimensions space, our Kalman filter based recommender system will suggest:

- if the user is interested in contents of types x, y and z and if he has already watched content of type x and y that day, the recommendation would essentially concentrate on content of type z
- knowing the position in the space, the best prediction for his next positions in the recommender space i.e. his best index of interest related to the favorite contents.

The strength of our approach is in its capability to make recommendations at a higher level which fit users habits i.e. given main directions to follow knowing the trajectory in the space and not to suggest specific resources.

Future works will be focused on tracking groups of users and on the definition of the topology of the recommendation space as a space including specific mathematical operators.

REFERENCES

[1] Anderson, B. and Moore, J. B., 1977. Optimal filtering. Prentice Hall – Information and System Sciences Series

[2] Gibson, W, 1988. Neuromancien. Collection J’ai Lu

[3] Söderström, T., 1994. Discrete-time stochastic systems : estimation and control. Prentice Hall International.

[4] Box, G.E.P. and Jenkins, G.M., 1970. Time series analysis : forecasting and control. Holden Day.

[5] Bernier C., Brun A., Aghasaryan A., Bouzid M., Picault J., Senot C. and Boyer A., 2010. Topology of communities for the collaborative recommendations to groups, SIIE 2010 conference, (Sousse, Tunisia, February 17 – 19, 2010).

[6] Gevers, M. and Vandendorpe, L. Processus stochastiques, estimation et prediction. DOI=http://www.tele.ucl.ac.be/EDU/INMA2731/].

[7] BARB: Broadcaster Audience Research Board, DOI=http://www.barb.co.uk/.

[8] Senot, C., Kostadinov D., Bouzid M., Picault J., Aghasaryan A. and Bernier C., 2010. Analysis of strategies for building group profiles. User Modeling, Adaptation and Personalization.

[9] M Pazzani, D. Billsus, 2007. Content-Based Recommendation Systems. The AdaptiveWeb, pages 325–341.

[10] D. Goldberg, D. Nichols, B. Oki, D. Terry, 1992. Using Collaborative Filtering to Weave an Information Tapestry. Communications of the ACM, pages 61–70

[11] M. Grcar, D. Mladenic, B. Fortuna, M. Groblenik, 2006. Data Sparsity Issues in the Collaborative Filtering Framework. Advances in Web Mining and Web Usage Analysis, pages 58–76

[12] Smith, J., 1998. The book, The publishing company. London, 2nd edition