

TWO-STEP RECOMMENDATIONS: CONTRAST ANALYSIS AND MATRIX FACTORIZATION TECHNIQUES

^{*}University of Lorraine, France

^{**}National Technical University of Ukraine “Kyiv Polytechnic Institute”, Kyiv, Ukraine

Анотація. У даній статті представлена двокрокова модель рекомендаційної системи, що використовує взаємодоповнюючим чином техніки контрастного аналізу та матричної факторизації. Також наданий короткий огляд варіацій методу матричної факторизації.

Ключові слова: системи рекомендацій, матрична факторизація, контрастний аналіз.

Аннотация. В данной статье представлена двухшаговая модель рекомендационной системы, которая использует взаимодополняющим образом техники контрастного анализа и матричной факторизации. Также приведен краткий обзор вариаций метода матричной факторизации.

Ключевые слова: системы рекомендаций, матричная факторизация, контрастный анализ.

Abstract. In this paper, we present a two-step recommendation model based on Contrast Analysis and Matrix Factorization techniques which mutually complement each other. We also provide a brief overview of different Matrix Factorization approaches.

Keywords: recommender systems, matrix factorization, contrast analysis.

1. Introduction

In modern numerical world, automatic recommendations are used in a wide area of applications starting from traditional recommendation of movies and music to considering search engines as a special type of recommender systems. The task of recommender engine is to predict how much a specific user will like a certain item and recommend the one, which has the highest predicted rating. It can be viewed as a task of filling in the unknown values of the rating matrix \mathbf{R} , the rows of which usually represent users and columns – items (Fig. 1). As a rule, rating matrix is very sparse. For example, GroupLens provided three datasets, which are widely used for testing different recommendation algorithms and contain 5,9%, 4,2% and 1,4% of known ratings respectively [1].

There exist many approaches for recommending, but they are usually classified into three categories [2]:

1. Content-based recommendations.
2. Collaborative recommendations.
3. Hybrid recommendations.

| | | Items | | | | |
|-------|-----------|-------|-------|-----|-----------|-------|
| | | i_1 | i_2 | ... | i_{n-1} | i_n |
| Users | u_1 | 4 | ? | ... | 4 | ? |
| | u_2 | ? | ? | ... | ? | 3 |
| | ... | ... | ... | ... | ... | ... |
| | u_{m-1} | ? | ? | ... | ? | 2 |
| | u_m | ? | 5 | ... | ? | ? |

Fig. 1. Rating matrix

Content-based recommender systems propose to user items with similar characteristics as those products, which were highly rated by a target user previously. Collaborative filtering approach will recommend those items, which were highly appreciated by users with similar interests. Collaborative filtering is more dynamic, as its recommendations can follow such events as, for example, fashion or change of interests in a certain user group. It also allows recommending items

with new characteristics. Still this group of methods can't propose anything, if there are not enough ratings in the system. Hybrid recommender systems use both collaborative and content-based methods. Relatively new and very promising approach in the field of recommender systems is Matrix Factorization, which belongs to the category of collaborative filtering [3].

In [4, 5] a novel Influence search algorithm, which is based on contrast analysis, was proposed for solving the task of searching levers of influence on the human decision-making process concerning such social problems as whether to bear a baby, whether to start studying etc. In [6] authors provided a general scheme of this algorithm and described its adoption to the recommender systems domain. Influence search algorithm essentially differs from traditional recommending approaches as in stead of solving the task of matrix filling it searches for the patterns of satisfied and dissatisfied users and provide recommendations how to improve their satisfaction.

2. Problem Formulation

In general, use of different nature methods can give essentially new and useful results. The purpose of this paper is to investigate possibility of joint usage of traditional recommendation approaches with a new Influence search algorithm.

3. Contrast Analysis for Recommender Systems

General scheme of the Influence search algorithm was provided in [6]. In this paper, we will consider its association rules based variant, which is presented in Fig. 2. Input of the proposed algorithm is presented by a set of records revealing information about user's interaction with the system. These records can also contain additional data, such as information about content of the items, user preferences, and demographic data.

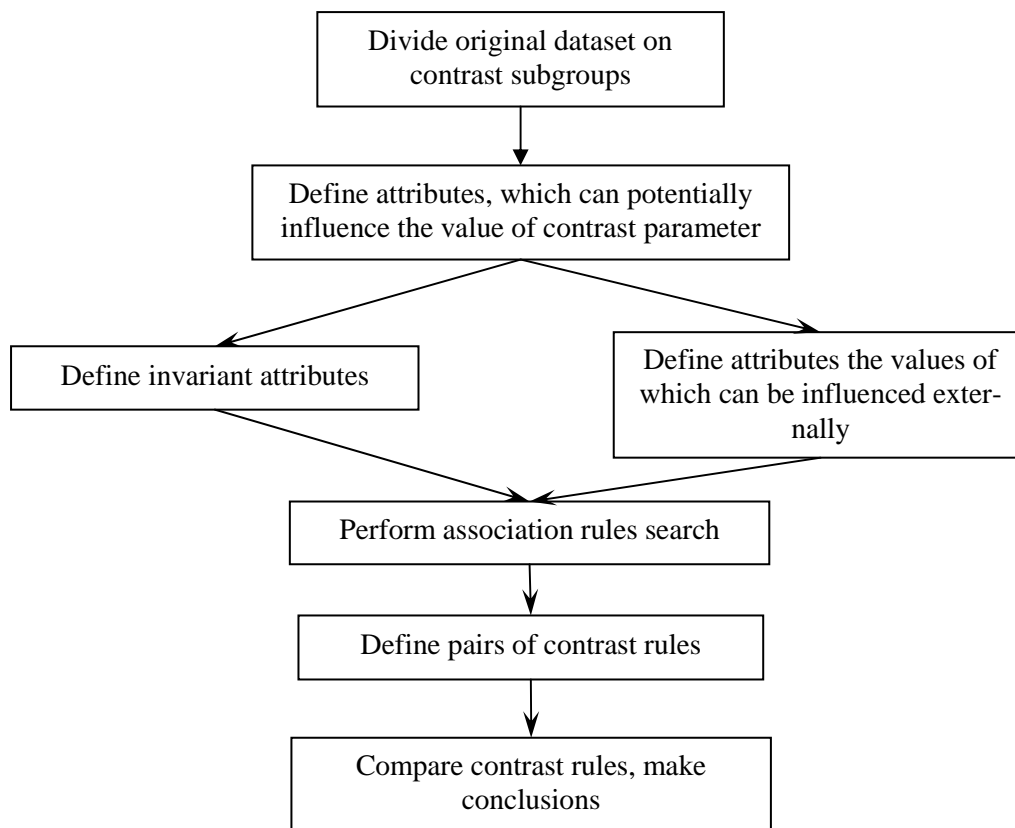


Fig. 2. Influence search algorithm

On the first step of the proposed algorithm, original data set is divided into two contrast groups basing on the value of contrast parameter. In the framework of recommender systems it is natural to use user satisfaction rate for contrasting, which can be estimated basing on the previous ratings.

The second step of the proposed approach is definition of attributes, which can potentially influence value of contrast parameter that is the level of satisfaction (influencing attributes). Satisfaction level of the user depends on the recommended items and user preferences; also, it can depend on the sequence of recommendations. So information about items (content) and users (demographic information, preferences) and sequence of recommended items must to be considered as influencing attributes.

Next, the set of chosen attributes is divided on two subsets: invariant or independent attributes and attributes the values of which can be influenced externally (that is by means of recommender system) or dependent attributes. It is obvious that among considered above attributes only proposed items and sequence of recommendations can be influenced by the system. That means that all other attributes (content, demographic information, preferences) belong to the invariant subset.

After that the search of contrast pairs of association rules is performed. A pair of association rules is considered to be contrast if two rules have different values of contrast parameter as conclusion of the rule and the bodies of the rules are constructed of premises with the same values of invariant attributes and different values of at least one influencing attribute. Contrast rules must have high confidence and not necessarily large support.

Let's examine the following example of contrast rules pair:

Rule 1. <sex=male>&<age=group1>&<preferences=actor1&actor2>&<recommended1 = comedy&actor1>&<recommended2=adventure&actor2> → <satisfied=YES> (*positive rule*)
support = 2%, confidence = 78%

Rule 2. <sex=male>&<age=group1>&<preferences=actor1&actor2>&<recommended1 = comedy&actor1>&<recommended2=comedy&actor2> → <satisfied=NO> (*negative rule*)
support=1,5%, confidence = 80%

Here sex, age, and preferences are invariant attributes and sequence of recommendations (recommended1 and recommended2) belongs to the group of depended attributes. Analyzing these 2 rules we can say that if we have a male user of age=group1, who likes actor1&actor2 and we have previously recommended him a comedy with actor1 now we need to recommend him an adventure with actor2 (but not a comedy with actor2), and with the probability of 78% (confidence of positive rule) user will like it. Obtained recommendations remain general and don't answer the question which film exactly we should recommend.

4. Matrix Factorization for Recommender Systems

Matrix Factorization (MF) gained its wide popularity after 2009 when a team BellKor's Pragmatic Chaos used it to win Netflix Prize competition [7]. The objective of this approach is to present rating matrix \mathbf{R} (where r_{ij} is a rating given by user i to the recommended item j) as a product of two matrices of a small rank (1).

$$\mathbf{R} \approx \mathbf{U}^T \mathbf{V}$$

$$\dim \mathbf{R} = m \times n \quad \& \quad \begin{cases} \dim \mathbf{U}^T = m \times k \\ \dim \mathbf{V} = k \times n \end{cases} \quad \& \quad k \ll \min\{m, n\}, \quad (1)$$

where m is total number of users, n – total number of items, k – number of features.

Values of the matrices \mathbf{U} and \mathbf{V} are usually calculated by a Gradient descend based or Alternating least squares [8, 9] methods using only known values of matrix \mathbf{R} . They minimize objective function (2)

$$\min \left(\left\| \mathbf{R} - \mathbf{U}^T \mathbf{V} \right\|^2 \right), \quad (2)$$

Where $\|\cdot\|$ is usually a Frobenius norm.

MF approach belongs to the class of latent factor models and aims to represent interaction between users and items with a small number of latent factors (features). It is obvious that if $k \ll \min\{m, n\}$ the complexity of the model is reduced significantly. In addition, it is easy to calculate unknown rating \hat{r}_{ij} using formulae (3).

$$\hat{r}_{ij} = \mathbf{u}_i^T \mathbf{v}_j, \quad (3)$$

where \mathbf{u}_i and \mathbf{v}_j are column-vectors of matrices \mathbf{U} and \mathbf{V} respectively.

There exist a number of variations of MF techniques, which can be classified on 3 groups (Fig. 3). The first group of methods builds model of the system using only one rating matrix. Such techniques as Regularized MF (RMF), Non-Negative MF (NNMF) and Matrix Tri-Factorization belong to this group. Second group analyses simultaneously two or more matrices (Collective MF or CMF) and the third one provides different generalizations of the basic MF approach (Kernel MF and Tensor Factorization).

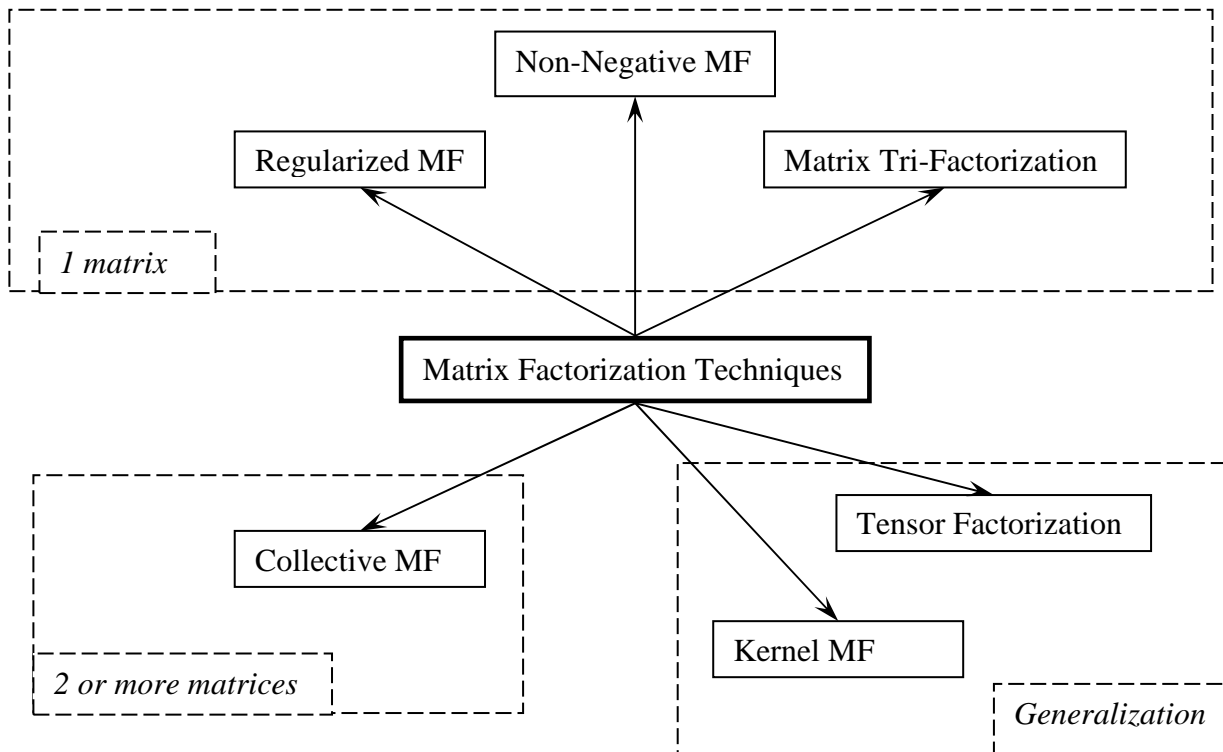


Fig. 3. Classification of Matrix Factorization Techniques

While using Regularized Matrix Factorization [10] an additional constant λ is added in order to avoid over fitting of the model, thus objective function is represented by equation (4).

$$\min \left(\left\| \mathbf{R} - \mathbf{U}^T \mathbf{V} \right\|^2 + \lambda \left(\sum_i \|\mathbf{u}_i\|^2 + \sum_j \|\mathbf{v}_j\|^2 \right) \right). \quad (4)$$

This approach is one of basic ones and it is widely used apart or in combination with other MF techniques [11–13].

In Non-Negative Matrix Factorization values of both \mathbf{U} and \mathbf{V} must satisfy the condition of positivity $u_{ij}, v_{ij} \geq 0$. This approach decomposes an object into a sum of its parts (allowing interpretation of the results). For example, for the task of image analysis it is possible to present a face as a sum of eyes, nose, lips and so on [14]. In the frame of recommender systems, basic parts can be considered as behavioral patterns [15] or groups of users [16].

Matrix Tri-Factorization [17] presents original rating matrix as a product of three matrices (5). In this model matrix \mathbf{U} represents interaction of m users and k_m user-related features, matrix \mathbf{V} – interaction of n items and k_n item-related features, and new matrix \mathbf{S} ($\dim(\mathbf{S}) = k_m \times k_n$) reveals interdependence between user-related and item-related features. Thus using Matrix Tri-Factorization it is possible to define different feature spaces for users and items. If holds $k_m, k_n \leq \min\{m, n\}$ complexity of the system is also reduced, as well as in the basic MF approach.

$$\mathbf{R} \approx \mathbf{U}^T \mathbf{S} \mathbf{V}. \quad (5)$$

Collective Matrix Factorization was proposed by Singh and Gordon in 2008 [18]. It performs simultaneous factorization of two or more matrices with condition of interdependence of feature spaces (6).

$$\begin{cases} \mathbf{X}^1 \approx (\mathbf{U}^1)^T \mathbf{V}^1 \\ \mathbf{X}^2 \approx (\mathbf{U}^2)^T \mathbf{V}^2 \end{cases}$$

$$\mathbf{U}^2 = f_u(\mathbf{U}^1, \mathbf{V}^1) \quad \text{and / or} \quad \mathbf{V}^2 = f_v(\mathbf{U}^1, \mathbf{V}^1). \quad (6)$$

CMF approaches are of particular interest, because depending on the nature of matrices \mathbf{X}^1 and \mathbf{X}^2 and dependences f_u and f_v it is possible to use additional information while building the model of the system. For example, in [19] authors took two rating matrices from different domains $\mathbf{X}^1 = \mathbf{R}^{src}$ (source domain) and $\mathbf{X}^2 = \mathbf{R}^{tgt}$ (target domain), thus using ratings from the source domain in order to build the model in the target one. In [18] matrix \mathbf{X}^1 was represented by a rating matrix and matrix \mathbf{X}^2 revealed information about item's content (characteristics). Performing collective factorization of these two matrices authors incorporated content information into collaborative-based recommendation method, hereby implementing a hybrid technique.

Equation (3) can be written in a form of inner product (7) that means that interaction between users, features and items can be considered as a linear kernel.

$$\hat{r}_{ij} = \langle \mathbf{u}_i, \mathbf{v}_j \rangle. \quad (7)$$

If dependence is more complex (not linear) it is possible to use other types of kernels, for example polynomial or Gauss (8), and perform Kernel Matrix Factorization [11].

$$\begin{aligned}
K_l(\mathbf{u}, \mathbf{v}) &= \langle \mathbf{u}, \mathbf{v} \rangle && \text{linear} \\
K_p(\mathbf{u}, \mathbf{v}) &= (1 + \langle \mathbf{u}, \mathbf{v} \rangle)^d && \text{polynomial} \\
K_g(\mathbf{u}, \mathbf{v}) &= \exp\left(-\frac{\|\mathbf{u} - \mathbf{v}\|^2}{2\sigma^2}\right) && \text{Gauss.}
\end{aligned} \tag{8}$$

Provided each rating depends not only on user and item but also on other variables, MF turns into tensor factorization, which was proposed in [20]. In this article, authors analyzed ratings depending on context that is different conditions on which a certain item was proposed to a specific user. For example, whether food was recommended when a person was hungry or not, or in what season the system recommended a user to watch a comedy.

Additional conditions imposed by each discussed above method are not always mutually exclusive. So it is possible to use different approaches simultaneously depending on the solved task and nature of the data. For example, authors of [17] used non-negative matrix tri-factorization and in [13] collective regularized MF was performed.

5. Joint Use of Matrix Factorization and Contrast Analysis

Because usually rating matrices are very sparse over fitting remains a problem for MF approaches. In addition, if available information is not fully representative (doesn't correctly represent all interaction between users and items), built model will lack for accuracy. That is why it seems promising to divide original rating matrix on sub-matrices depending on the nature of the data and perform sub-matrix factorization. It is also possible to use different MF approaches for different sub-matrices according to their properties.

We can use Influence search algorithm for the task of identification of essentially different sub-matrices because it extracts groups of related users. So these two methods can be used as complements to each other. For example, considering pair of contrast rules discussed above, we can identify what film exactly we should recommend with the help of MF approach used for sub-matrix consisting of users $\langle \text{sex}=\text{male} \rangle \& \langle \text{age}=\text{group1} \rangle \& \langle \text{preferences}=\text{actor1} \& \text{actor2} \rangle$ and items with $\langle \text{content}=\text{adventure} \& \text{actor2} \rangle$. Thus we can consider a two-step recommender system, where first more general recommendation is generated by means of contrast analysis and after that, whenever possible, recommendation is personalized by means of Matrix Factorization.

6. Conclusion

In this paper, we discussed peculiarities of Influence search algorithm usage in the frame of recommender systems. We also presented a brief overview of Matrix Factorization approaches and outlined advantages of each of them. In the end we proposed a two-step recommender system model, which incorporates Contrast Analysis and Matrix Factorization and allows generate more general recommendation with possibility of their further personalization.

REFERENCES

1. MovieLens | GroupLens [Електронний ресурс]. – Режим доступу: <http://grouplens.org/datasets/movielens>.
2. Campos P.G. Towards a More Realistic Evaluation: Testing the Ability to Predict Future Tastes of Matrix Factorization-based Recommenders / P.G. Campos, F. Díez, M. Sánchez-Montañés // Proc. Conference on Recommender Systems (RecSys 2011), (Chicago, October 23 – 27, 2011). – NY: ACM, 2011. – P. 309 – 312.
3. Wu M. Collaborative Filtering via Ensembles of Matrix Factorizations / M. Wu // Proc. KDDCup and Workshop, (San Jose, August 12, 2007). – San Jose, 2007. – P. 43 – 47.

4. Chertov O. Fuzzy Clustering with Prototype Extraction for Census Data Analysis / O. Chertov, M. Aleksandrova // *Soft Computing: State of the Art Theory and Novel Applications*. – 2013. – Vol. 291. – P. 289 – 313.
5. Chertov O. Using Association Rules for Searching Levers of Influence in Census Data / O. Chertov, M. Aleksandrova // *Procedia – Social and Behavioral Sciences*; eds.: G. Giannakopoulos, D. Sakas, D. Vlachos, D. Kyriaki-Manessi. – Amsterdam: Elsevier. – 2013. – Vol. 73. – P. 475 – 478.
6. Chertov O. Data Mining Methods Usage for the Task of Automated Forming of Recommendations / O. Chertov, M. Aleksandrova // *Scientific journal “Bulletin of Volodymyr Dahl East Ukrainian National University”*. – 2013. – N 15 (204), Part 1. – P. 237 – 242.
7. NetflixPrize: Home [Електронний ресурс]. – Режим доступу: <http://www.netflixprize.com>.
8. Parallel matrix factorization for recommender systems / H.-F. Yu, C.-J. Hsieh, S. Si [et al.] // *Knowledge and Information Systems*. – 2013. – Vol. 37, N 3. – P. 629 – 651.
9. Pilászy I. Fast ALS-based Matrix Factorization for Explicit and Implicit Feedback Datasets / I. Pilászy, D. Zibriczky, D. Tikk // *Proc. Conference on Recommender Systems (RecSys 2010)*, (Barcelona, September 26 – 30, 2010). – NY: ACM, 2010. – P. 71 – 78.
10. Koren Y. Matrix factorization techniques for recommender systems / Y. Koren, R. Bell, C. Volinsky // *Computer*. – 2009. – Vol. 42, N 8. – P. 30 – 37.
11. Rendle S. Online-Updating Regularized Kernel Matrix Factorization Models for Large-Scale Recommender Systems / S. Rendle, L. Schmidt-Thieme // *Proc. Conference on Recommender Systems (RecSys 2008)*, (Lausanne, October 23 – 25, 2008). – NY: ACM, 2008. – P. 251 – 258.
12. Matrix Factorization and Neighbor Based Algorithms for the Netflix Prize Problem / G. Takacs, I. Pilászy, B. Nemeth [et al.] // *Proc. Conference on Recommender Systems (RecSys 2008)*, (Lausanne, October 23 – 25, 2008). – NY: ACM, 2008. – P. 267 – 274.
13. Yuan Q. Factorization vs. Regularization: Fusing Heterogeneous Social Relationships in Top-N Recommendation / Q. Yuan, L. Chen, S. Zhao // *Proc. Conference on Recommender Systems (RecSys 2011)*, (Chicago, October 23 – 27, 2011). – NY: ACM, 2011. – P. 245 – 252.
14. Wang Y. Nonnegative Matrix Factorization: A Comprehensive Review / Y. Wang, Y. Zhang // *IEEE Transactions on Knowledge and Data Engineering*. – 2013. – Vol. 25, N 6. – P. 1336 – 1353.
15. Learning from Incomplete Ratings Using Non-negative Matrix Factorization / S. Zhang, W. Wang, J. Ford [et al.] // *Proc. 6th SIAM Conference on Data Mining*, (Bethesda, April 20 – 22, 2006). – Bethesda, 2006. – P. 548 – 552.
16. Factorisation en matrices non-négatives pour le filtrage collaboratif / J.F. Pessiot, N. Usunier, M. Amini [et al.] // *Proc. Conférence en Recherche d'Informations et Applications (CORIA 2006)*, 3rd French Information Retrieval Conference, (Lyon, March 15 – 17, 2006). – Lyon, 2006. – P. 315 – 326.
17. Chen G. Collaborative filtering using orthogonal nonnegative matrix tri-factorization / G. Chen, F. Wang, C. Zhang // *Information Processing & Management*. – 2009. – Vol. 45, N 3. – P. 368 – 379.
18. Singh A.P. Relational Learning via Collective Matrix Factorization / A.P. Singh, G.J. Gordon // *Proc. 14th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD 08)*, (Las Vegas, August 24 – 28, 2008). – Las Vegas, 2008. – P. 650 – 658.
19. Huang Y. J. Constrained Collective Matrix Factorization / Y.J. Huang, E.W. Xiang, R. Pan // *Proc. Conference on Recommender Systems (RecSys 2012)*, (Dublin, September 9 – 13, 2012). – NY: ACM, 2012. – P. 237 – 240.
20. Multiverse Recommendation: N-dimensional Tensor Factorization for Context-aware Collaborative Filtering / A. Karatzoglou, X. Amatriain, L. Baltrunas [et al.] // *Proc. Conference on Recommender Systems (RecSys 2010)*, (Barcelona, September 26 – 30, 2010). – NY: ACM, 2010. – P. 79 – 86.

Стаття надійшла до редакції 20.01.2014