International Journal of Advanced Research in Computer Engineering & Technology (IJARCET)
Volume 7, Issue 9, September 2018, ISSN: 2278 – 1323

Application of Recommender Systems in the Design of Complex Microsystem Devices

Mykhaylo Lobur[1], Mykhaylo Shvarts[2], Yuriy Stekh[3]

1. CAD Department, Lviv Polytechnic National University, UKRAINE, Lviv, S. Bandery street 12, Loburm@gmail.com

2. CAD Department, Lviv Polytechnic National University, UKRAINE, Lviv, S. Bandery street 12,, mykhaylo.shvarts@gmail.com,

3. CAD Department, Lviv Polytechnic National University, UKRAINE, Lviv, S. Bandery street 12,,yuriy.v.stekh@lpnu.ua,

Abstract **- In this article methods and algorithms for predicting recommendations for complex device designer communities are considered. Methods use the qualification profile of the of complex device designers. The search for complex device designer communities contains two stages. At the first stage, the community of designers are searched on the basis of the qualification vectors of the designer profiles. At the second stage, the community of designers are searched on the basis of the rating designer profiles A method for improving the accuracy of calculating the similarity coefficients in the method of collaborative filtering is considered.**

Index Terms **- complex device designer communities, recomendation system, clustering, collaborative filtering**

I. INTRODUCTION

The design of complex microsystem devices is automated at all stages. To obtain reliable, functional and relatively cheap, the design phase is very important. There are quite a lot of software tools that allow you to model complex devices. The most famous of them are: SUGAR, MEM Research, 3-D ANISOTROPIC ETCH SIMULATION ON-LINE , CorningIntelliSense Corporation, Nodal Design of Actuators and Sensors, Coventor Inc, MEMSCAP, Tanner Research [8]. Specific characteristics and differences between the design, production and application of microsystems in comparison with traditional (macro) implementations stem from their size. The design of microsystems is of paramount importance due to the overlap of price responsibility for the subsequent stages, due to the high cost of the prototype and the lack of repair capability. Design covers the steps from analysis of options to functional optimization and final production documentation (Fig.1) [9].

Microsystems consist of separate components, such as sensors and actuators, which are integrated and packaged together with control and computing electronics. Not all steps can be automated the same way. In particular, the conceptual design and development of the principles of operation is based on the creative ability of the developer and, therefore, can not be standardized. It can only be supported to a small extent by the design environment. The main types of microsystem devices are: optical, thermal, liquid, bio- devices [10]. Microsystem devices design requires knowledge and experience in such areas as mechanics, thermodynamics, field theory, optics, hydrodynamics [11].

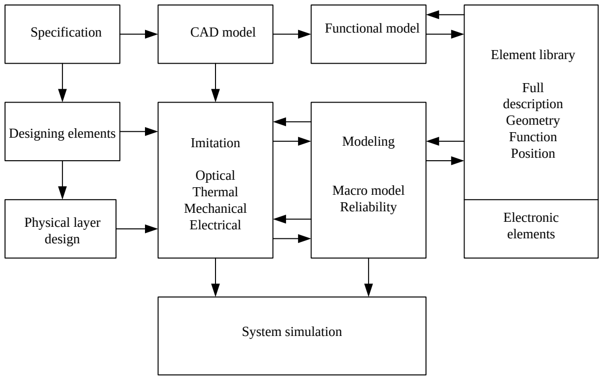

Fig. 1 Structural diagram of complex microsystem devices design

An urgent task in modern design of microsystem devices is the development of automated systems that would allow the exchange of experience and knowledge to the designers of microsystem devices. These systems include recommender systems [1]–[3]. Recommender systems are widely used in e-commerce, information retrieval systems, multimedia portals [4], [5], [7]. The task of recommender systems is to help users choose services and items that would meet their requirements. Recommendation systems can be used in the design of microsystem devices for the exchange of experience and knowledge between designers. The article proposes a new method for forecasting recommendations when designing microsystem devices. The method takes into account the qualification characteristics of the designers and the design experience of microsystem devices. The method allows designers to recommend the choice of the principles of operation, structure, materials and parameters when designing new microsystem devices. Recommendations can be provided to individual designers, as well as to the whole group of designers.

II. FORMULATION OF THE PROBLEM

Formal statement of the task of forecasting personalized recommendations is as follows. Let **U** be a set of designers, |**U**| = m, **I** is a set of objects (principles of operation, structure, materials and parameters when designing microsystem devices), |**I**| = n, **D** is the set of actions giving personalized recommendations for designers. It is necessary to carry out the forecast of personalized recommendations $\mathbf{U} \times \mathbf{I} \rightarrow \mathbf{D}$.

Let $\mathbf{R}$ be the set of microsystem device designers ratings (items ratings), the assessments given by previous designers, $\mathbf{U}_p$ - the set of user profiles for microsystem device designers, $\mathbf{I}_p$ - the set of the profile of the items (principles of operation, structure, materials and parameters when designing microsystem devices), $\mathbf{U_{p_i}} \in \mathbf{U_p}$ - the vector of the profile of the i-th user $\mathbf{U_{p_i}} = ((i_1, r(u_{p_i}, j_1), (i_2, r(u_{p_i}, j_2), \ldots, (i_n, r(u_{p_i}, j_n))$, $r(u_{p_i}, j_k) = r_{ij}$ - the estimation of the i-th user for the j-th item, $\mathbf{I_{p_j}} \in \mathbf{I_p}$ - the profile vector of the j-th item $\mathbf{I_{p_j}} = ((u_1, r(u_1, i_j), (u_2, r(u_2, i_j), \ldots, (u_m, r(u_m, i_j))$, $r(u_i, i_{p_j}) = r_{ij}$ -estimation of the j-th item for the i-th user, $r(u_{p_i}, j_k) = r(u_i, i_{p_j})$.

A personalized recommendation forecast is required $\hat{r}_{ij} = \text{Predict}(u, i, \ldots) \approx r_{ij}$, where $\hat{r}_{ij}$ - the predicted value of the estimate, r_{ij} - a user-rated score. The task of the system is to achieve the minimum difference between $\hat{r}_{ij}$ and r_{ij}

$$\left|\hat{r}_{ij} - r_{ij}\right| \Rightarrow \min \qquad (1)$$

The task of the recommendation system can be described by the following map

$$RS_A : \mathbf{U_p} \times \mathbf{I_p} \rightarrow \hat{\mathbf{R}} \qquad (2)$$

where $\hat{\mathbf{R}}$ - is the set of predicted values of object valuations for which a set of recommended items can be determined.

The objective function of forecasting recommendations for the recommendation system is represented by the following equation

$$F_{RS} = \min\left(\sum_{i=1}^{m} \sum_{j=1}^{n} \left|\hat{r}_{ij} - r_{ij}\right|\right) \qquad (3)$$

The formulation of the problem of forecasting recommendations for communities of complex microsystem device designers is as follows. We split the set of designers $\mathbf{U}$ in a partition of m groups $\mathbf{G} = \{\mathbf{G_1}, \mathbf{G_2}, \ldots, \mathbf{G_m}\}$, so that for each group $\mathbf{G_i} \subseteq \mathbf{U}\ (i \in \{1, 2, \ldots, m\})$ every user $\mathbf{U_j} \in \mathbf{G_i}$ receives the same recommendations.

In methods for predicting recommendations for user communities, a user- item matrix is used. This matrix has a high degree of sparsity (6-7% of non-zero elements). Preliminary implemented prediction of non-zero elements by the weighted sum method and matrix filling is carried out. In the next step, user communities are selected based on the clustering of the user profile vector of the user-item matrix. For this, the k-means method and its modifications are used. The result of clustering is the set of user-item matrices for each user communities. For each user communities, the average value of the rating of each item in the group is calculated. Users are offered top - N items that have the largest average ratings in each user community. The disadvantage of such methods is that they take into account only the current values of the item rating by users. These methods give a significant error in the absence of estimates or with a small number of estimates.

III. THE METHOD OF SEQUENTIAL CLUSTERING FOR PREDICTING RECOMMENDATIONS

In this method it is proposed to isolate communities of microsystem device designers using a sequential two-stage clustering (Fig. 2).

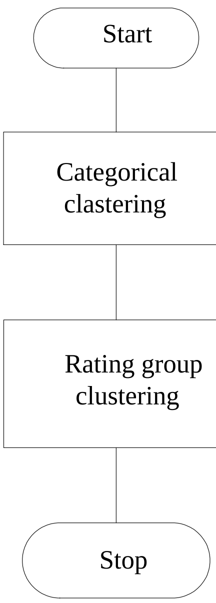

Fig.2 Block diagram of sequential clustering algorithm

Stage 1. Community of designers are searching on the basis of the qualification characteristics of designers.

Each designer is characterized by a profile vector, which contains the following categorical parameters: area of interest in designing microsystem devices; experience in designing microsystem devices in years; the availability and number of publications in the field of interest in the design microsystem devices; the availability and number of patents in the field of interest in the design microsystem devices; education (technical, physical or mathematical), positions occupied in the design of microsystem devices in years. The k-mod categorization clustering method [6] is used for searching community of designers.

Stage 2. Clustering by the similarity of the numerical vectors of designer profiles is carried out in each of the groups identified by the qualification characteristics of designers.

For solving the problem of the second stage is used modified k-means method [19]. This method is seeking the necessary optimal number of clusters and breaking objects into clusters with clustering accuracy control (Fig. 3).

Each received cluster is considered as a separate group of users, for which a forecast of recommendations is made.

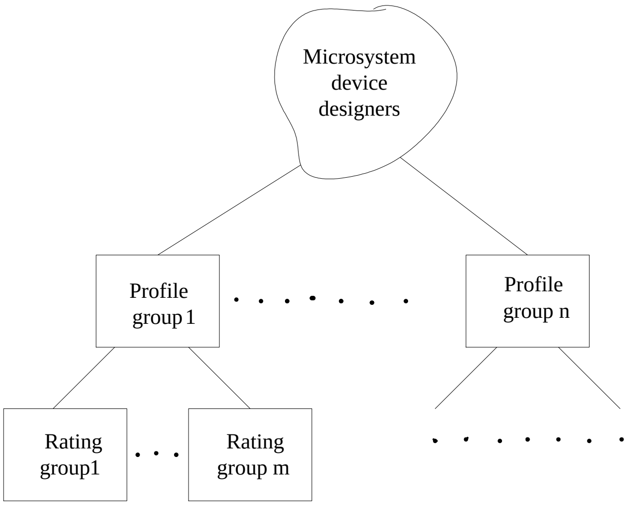

Fig.3 Structure of the method of sequential clustering

The process of forecasting recommendations for a new user has two steps:

1. Based on the profile characteristics of the new user, it refers to a specific profile group (profile cluster).
2. If the rating vector of the profile is empty, it predicts objects based on the averaged profile vectors of all group members.
3. If the rating vector of the profile contains estimates, it refers to a certain rating group and recommends items that take into account the interests of all members of the group.

Additive and multiplicative utility is used to predict objects in a group The proposed method solves the problem of sparseness of the user-item matrix, the problem of inaccuracy in calculating the similarity coefficients between the user profile vectors, the scalability problem, the problem of the new user.

IV. USE OF DESIGNER QUALIFICATION CHARACTERISTICS TO IMPROVE THE ACCURACY OF FORECASTING RECOMMENDATIONS

The basic method of forecasting recommendations in modern recommendation systems is the method of collaborative filtering [16], [17]. A generalized scheme of the method of collaborative filtration is presented in Fig. 4.

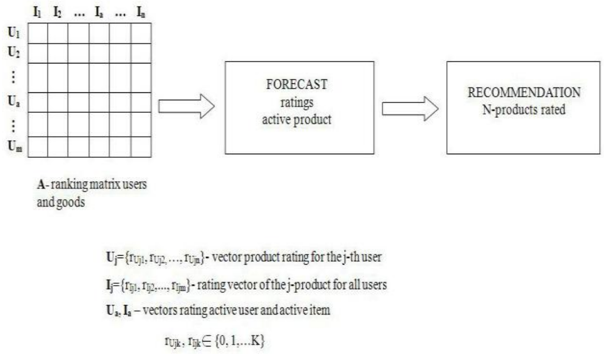

Fig.4 Generalized scheme of the method of collaborative filtering

A weighted sum method is used to predict the ratings of active user products:

$$r_{a,i} = \overline{r_a} + \frac{\sum\limits_{l \in T_a} (r_{l,i} - \overline{r_l}) w_{a,l}}{\sum\limits_{l \in T_a} \left| w_{a,l} \right|}, \qquad (4)$$

where $r_{a,i}$ - rating of the *i*-th item for the active user;

$\bar{r}_a, \bar{r}_l$ - average ratings of the respective user profile vectors;

$w_{a,l}$ - the coefficient of similarity between the vectors of the profiles *a* - th and *l* - th user;

T_a - a set of vector user profiles that have collectively rated products.

Typically, similarity coefficients are calculated using the cosine degree of similarity, the Pearson correlation coefficient, or the inverse Euclidean distance [18]. The kNN method is used to search for a set T_a of user profile vectors, which reduces the accuracy of the calculation. The prediction accuracy also reduces the high sparsity of the user-item matrix. User - item matrix contains 6-7% of non-zero elements [12]–[14]. Main of the problems with modern recommendation systems is the problem of a new user [15]. The new user has an empty profile vector when entering in the recommendation system (the profile vector contains only zero elements). A promising way to solve these problems is to use demographic information about users. Demographic information about the user can be obtained by registering the user in the system, from social networks, from the analysis of the content of other sites. In the developed method, instead of the demographic characteristics of users, qualification characteristics of the microsystem device designers are used. Qualification characteristic of designers are analyzed and qualification profiles of designers are formed:

$$\mathbf{U}_{\text{qual}} = \{experience, publication, patent, education, position occupied\} \quad (5)$$

The qualification profile vector can be effectively used to improve the accuracy of calculating the similarity coefficients in the formula (4). The total categorized vector of the qualification characteristic of microsystem device designers has 27 components. The component of the categorized qualification characteristic of microsystem device designers profile vector contains 1 in the position that corresponds to the category of qualification characteristics and 0 in the remaining positions. After the categorization, the components of the vector (5) contain binary bit values with 27 bit components. Despite the fact that the dimension of the vector increases to 27, the binary bit content of the vector allows to effectively determine the similarity between categorized qualification characteristic of microsystem device designer vectors. To do this, we use the Jaccard coefficient of similarity:

$$J = \frac{|\mathbf{A} \cap \mathbf{B}|}{|\mathbf{A} \cup \mathbf{B}|} \quad (6)$$

where **A** and **B** are vectors that can contain arbitrary real values or alphanumeric information.

For binary vectors, the formula for calculating the similarity coefficient has the following form

$$J = \frac{M_{11}}{M_{10} + M_{01} + M_{11}} \quad (7)$$

where M_{11} - the total number of elements, where the components of vectors A and B have a value of 1;

M_{10} - the total number of elements where the components of the vector A have a value of 1 and the components of the vector B have a value of 0;

M_{01} - the total number of elements where the components of the vector A have a value of 0 and the components of the vector B have a value of 1;

M_{00} - the total number of elements where the components of the vectors A and B have a value of 0.

Each component of vectors **A** and **B** must be in one of four categories:

$$M_{11} + M_{10} + M_{01} + M_{00} = n \quad (8)$$

where n - the dimension of vectors **A** and **B**.

Calculating the coefficients in expression (8) requires performing operations on the binary content of qualification profile vectors, which requires considerably less time costs than calculating similarity coefficients for the rating vectors of user profiles. To predict the recommendations in formula (4), two variants of calculation of the modified coefficient of similarity are used:

$$w'_{ij} = sim(i, j) + J_{ij} \times sim(i, j) \quad (9)$$

$$w'_{ij} = J_{ij} + J_{ij} \times sim(i, j) \quad (10)$$

where $sim(i, j)$ - the coefficient of similarity between rating vectors of user profiles;

J_{ij} - the similarity coefficient between the qualification profile vector.

The expression (9) takes into account the dominant significance of the similarity of users with the rating vectors of the profiles. The expression (10) takes into account the dominant meaning of similarity of users by qualification profile vectors. To calculate the similarity coefficients between the rating vectors of qualification profile, the method based on the inverse Euclidean distance between vectors is used. The proposed method allows you to solve the problem of a new user. For the new user, expression (10) will look like this.

$$w'_{ij} = J_{ij} \quad (11)$$

The similarity coefficients are used to predict recommendations in the formula (4).

The test data of the GroupGlens group was used to test the developed methods [19]. The research group GroupLens Research offers several sets of data about the ratings of films. The sets contain movie ratings, metadata about films (genre and year of release) and demographic information about users (age, postal code, sex and occupation). The MovieLens 1M kit contains 1 million ratings of 4000 movies, affixed to 6000 users. Data is divided into three tables: ratings, user information and information about the films.

The test results are shown in Fig.5. The abscissa indicates the size of the k-nearest area, within which predictions are made. The absolute error of prediction is plotted along the ordinate axis.

UB1 - cosine degree of similarity of rating vectors of profiles;

UB2 - modified degree of similarity, dominant significance of qualification characteristics;

UB3 - modified degree of similarity, dominant rating similarity.

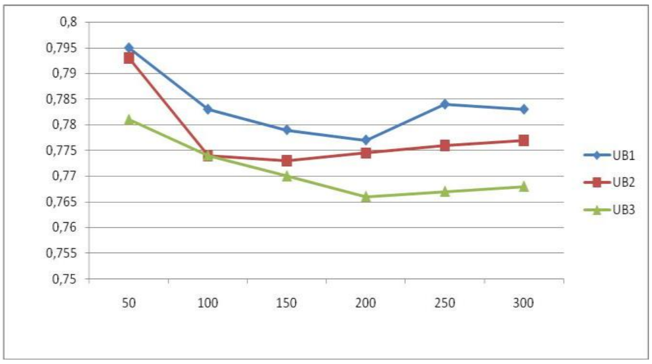

Fig.5 Test results on the set Movilens

III. Conclusion

This article presents new methods for using qualification information when predicting recommendations in recommendation systems for microsystem device designers communities. The method takes into account the qualification characteristics of the microsystem device designers and the quantitative characteristics of the utility of the application of components and technologies at the stages of conceptual design and the development of the principles of the work of microsystem devices. The method also allows to consider separately only the qualification characteristics or the utility characteristics of the application of microsystem device components and technologies. This approach is especially effective for microsystem device designers with extensive experience in designing certain classes of devices in the transition to the design of new classes for them. In general, the proposed methods allow to accumulate and transfer experience of microsystem device design both between designer communities, and between individual designers. The effectiveness of using the method was investigated on a test set of data. The experiments proved the high efficiency of the joint accounting of qualification characteristics of the microsystem device designers and the quantitative characteristics of the utility of the application of components and technologies in comparison with their separate use.

References

[1] K. McCarthy, M. Salamo, L. Coyole, L. McGinty, B. Smyth, P. Nixon, "Group RecommenderSystems: A Critiquing Based Approach, " In *IUI '06: Proceedings of the 11th international conferenceon Intelligent user interfaces*, 2006. – pp. 267-269.

[2] I. Garcia, L. Sebastia, E. Onaindia, C. Guzman, "A Group Recommender System for Tourist Activities," In *EC-Web 2009*: *Proceedings of E-Commerce and Web Technologies*, 2009. – pp. 26-37.

[3] S.K. Moon, T.W. Simpson, S.R.T. Kumara, "An agent-based recommender system for developing customized families of products" // *Journal of Intelligent Manufacturing*, Vol. 20, pp. 649-659, 2009.

[4] Y.-J. Chen, Y.-M. Chen, M.-S. Wu, "An expert recommendation system for product empirical knowledge consultation," In *International Conference on Computer Science and Information Technology ICCSIT2010*, 2010. – pp. 23-27.

[5] E.-A. Baatarjav, S. Phithakkitnukoon, R. Dantu, "Group Recommendation System for Facebook," In *OTM 2008: Proceedings of On the Move to Meaningful Internet Systems Workshop (2008),* Springer, LNCS 5333, 2009. – pp. 211-219.

[6] H.C. Romesburg,"Cluster Analysis for Researchers," Lulu Press, California, 2004.

[7] G.W. Flake, S. Lawrence, C.L. Giles, F. Coetzee, "Self-Organization and identification of Web Communities" // *IEEE Computer*, vol. 35, pp. 66-71, 2002.

[8] M. Kasper "Mikrosystementwurf. Entwurf und Simulation von Mikrosystemen", Springer-Verlag, Berlin Heidelberg, 2000

[9] J.A. Kubby "A Guide to Hands-on MEMS Design and Prototyping", Cambridge University Press, 2011

[10] J. A. Pelesko, D. H. Bernstein Modeling MEMS and NEMS, CRC press company, 2003

[11] N. Maluf, K. Williams, A. House, An Introduction to Microelectromechanical Systems Engineering, ARTECH HOUSE, INC., 2004

[12] J. Bobadilla, F. Ortega, A. Hernando, A. Gutiérrez, "Recommender systems survey" // *Knowledge-Based Systems*, vol. 46, pp.109-132, 2013.

[13] L. Lü, M. Medo, C.H.Yeung, Y.-C. Zhang, Z.-K. Zhang, T. Zhou, "Recommender systems" // *Physics Reports*, vol.519, pp.1-49, 2012.

[14] G. Adomavicius, 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*, vol.17, pp.734-749, 2005.

[15] J.B. Schafer, J. Konstan, J. Riedl, "E-commerce recommendation applications" // *Applications of Data Mining to Electronic Commerce*, pp. 115-153, 2001.

[16] P. Resnick, H.R. Varian, "Recommender systems" // Communications of the ACM, vol.40, pp.56-58, 1997.

[17] P. Resnick, N. Iacovou, M. Suchak, P. Bergstrom, J. Riedl, "GroupLens: an open architecture for collaborative filtering of netnews" In *Proceeding of the ACM Conference on Computer Supported Cooperative Work,* pp. 175–186, 1994.

[18] Y. Stekh, V. Artsibasov, "Some methods in software development recommendation systems" // *Visnyk NU "Lvivska politekhnika" : Kompiuterni systemy proektuvannia. Teoriia i praktyka*, № 777, pp. 74-78, 2013.

[19] Y. Stekh, V. Artsibasov, "Adaptive clustering algorithm for recommender systems" // *Visnyk NU "Lvivska politekhnika" : Kompiuterni systemy proektuvannia. Teoriia i praktyka*, № 747, pp. 75-78, 2013.

Dr. Stekh Yuriy, PhD, Assistant Professor of CAD department of Lviv Polytechnic National University. Research interests: databases, recommender systems, artificial intelligence methods, data mining.
Prof. Mykhaylo Lobur, DSc, PhD, Academitian of Academy of Science in applied radioelectronics, Head of CAD department of Lviv Polytechnic National University. Research interests: design of CAD tools for microelectronic devices design.
Mykhaylo Shvarts, Master of Computer Science. Research interests: recommender systems, artificial intelligence methods, data mining