

A Hybrid Genre-based Recommender System for Movies using Genetic Algorithm and kNN Approach

Dr. Amit Verma

*Department Of Computer Science
Chandigarh University
Gharuan, India*

Harpreet Kaur Virk

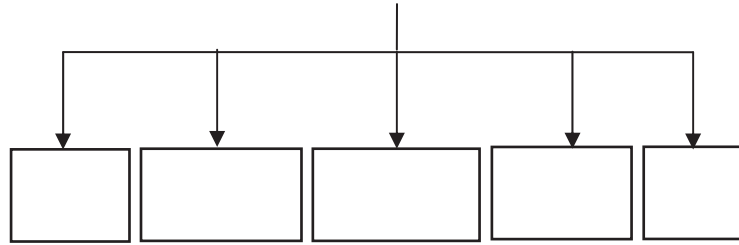
*Department of Computer Science and engineering
Chandigarh University, Mohali, Punjab, India*

Abstract- The need of personalization and filtering systems is growing permanently due to the immense information available online. Recommendation systems form or establish a specific type of information filtering approach that attempt to present items according to the interest of a user. In content-based filtering, the systems examine items previously preferred by the actual user, whereas in collaborative filtering, recommendations are based on the information of similar users or items. In content-based the recommendations are provided based on the profiles of users that are created at the beginning. For recommendation, the engine compares the items that were rated by the user with that are not rated by the user to calculate similarities. Items that are similar to rated items will be recommended to the user. In collaborative filtering the focus is on relationship between users and items. Similarity of items is calculated by similarity of ratings of those items by the users who have rated both items. We also analyzed that recommendations are also influenced by the factors such as age, gender and some other user profile information. In our work both content and collaborative techniques and some demographic information are combined into a hybrid approach, where additional content features are used to improve the accuracy of collaborative filtering. Also we are using the genetic algorithm and k-NN algorithm to provide recommendations to the user. To evaluate precision, recall and F1-measure performance parameters are used.

I. INTRODUCTION

Commercially recommender technology came in emergence in the late 1990s. Perhaps Amazon.com is the most popular application using this technology in which they recommend items for the users based on their purchase and browsing history. While adoption of recommender technology by Amazon frequently based on CF technique, has been incorporated in various e-commerce and online structures. A considerable inspiration to exploit this is to boost sales degree by way of the customers may pay for an element if it is recommended to them but might not look for it otherwise.

Various corporations have been built in order to present recommendation technology and services to online retailers, such as Strands and NetPerceptions. Beyond collaborative filtering, the toolbox of recommendation techniques has also incorporate CBF techniques grounded on information retrieval, Bayesian implication, and case-based interpretation approaches[10, 24]. These approaches or methodologies take into account the genuine features or characteristics of the things to be suggested rather than user rating patterns. As several recommender approaches have developed, hybrid recommender systems[12] have also emerged in which multiple algorithms are combined. However, collaborative filtering has remained a valuable technique, both unaided and crossbred with CBF approaches. However, collaborative filtering has remained a valuable technique, both unaided and hybridized with content-based approaches. To provide an overview of the different types of recommender systems, we want to cite a taxonomy provided by[12].



In content-based filtering method, the system investigate a set of portrayals of items that are earlier rated by a customer, and create a profile of user preferences grounded on the artifacts of the objects graded by that user[5]. The profile is an organized demonstration of user preferences which is assumed to suggest new interested items. Basically, the recommendation procedure comprise of comparing the features of the customer profile contrary to the features of a contented item. The outcome is a pertinence finding that denotes the level of user curiosity in that item. A profile can be used to filter the search outcomes by determining the fact that the user is fascinated by a particular item or not and this is possible only if the profile of the user precisely considers user likings. Items are recommended by this type of method grounded on demographic profile of the user. This sort of method suggests items grounded on the demographic outline of the user. Simple and efficient personalization solutions based on demographics are adopted by several websites. To cluster users the demographic information can be used such as country, age, gender, etc. This information is then compared with the existing users to find the most related cluster and finally, items in the most identical cluster are recommended to the user.

Knowledge-based recommendation methods provide suggestions based on particular domain knowledge with reference to user preferences. Case-based [20, 21] systems are assumed as remarkable in the field of knowledge-based recommender systems. These systems estimate that exactly how much the user requirements match with the list of suggestions. Another type of knowledge-based systems is constraint-based systems. Knowledge-based systems have a tendency to work better systems than others on condition that they are furnished with learning mechanism otherwise they may be beaten by further trivial methods that can utilize the logs of human computer interaction. All kind of recommendation approaches possess some strengths and flaws. The flaws of pure recommender systems can be overcome through combination of different approaches[22].

This paper consists of five sections. Section II describes related work. Section III describes the methodology to be followed to recommend movies to the user. Section IV is the experimental analysis. Section V wraps up this paper with its conclusion and future scope.

II. RELATED WORK

Belkin, et al.[3] did a subjective comparison of different information filtering and information retrieval models. Authors present an advantages and disadvantages of different information filtering and information retrieval models.

Badrul Sarwar et al.[11] analyzed various item-based recommendation algorithms. An experimental result shows that item-based approaches perform better than the user-based approaches and also provide superior quality than the user-based algorithms.

Qing Li et al.[13] to resolve the cold start problem clustering techniques have been applied on item-based collaborative filtering approach. Content information is also integrated into collaborative filtering. To analyze the proposed technique various experiments have been conducted on MovieLens dataset. The results show the improvement in recommendations provided by using item-based collaborative filtering and also resolve the cold start problem.

Linden, et al. [14] presents an industry report on the recommendation system used in Amazon.com e-commerce site. Report presents Item-to- Item Collaborative filtering for the recommendation of product on their site. Amazon.com practices suggestions as a targeted promotion device in several email campaigns and on maximum of its Web sites'

pages, together with the great traffic Amazon.com homepage. Snapping on the “Your Recommendations” link points users to a region where they can filter their recommendations by product line and subject region, rate the suggested products, rate their preceding procurements, and understand why items are suggested.

Miyahara, K., et al.[15] discussed a collaborative filtering approach based on the modest Bayesian classifier. The modest Bayesian classifier is one of the supreme effective supervised machine-learning algorithms. It performs well in numerous classification tasks instead of its modesty. Two variations of the recommendation issue for the modest Bayesian classifier are defined. In this method, relationship among the users from negative and positive ratings independently is calculated and assessed these algorithms by means of a database of movie recommendations and joke recommendations.

Bamshad Mobasher et al.[27] analyzed the various issues to build secure recommendation systems and bring in various new attack models. Simulation-based assessment to show the most successful attack model against commonly used recommendation approaches was also performed. Analysis shows that hybrid algorithms might provide a high degree of robustness. Whereas, item-based and user-based algorithms are extremely susceptible to particular attack models considering equally the whole impact on the capability of the system to generate correct predictions, along with the amount of knowledge about the system essential by the invader to post a realistic attack. Finally, to detect attack profiles they develop a classification-based framework.

L. Baltrunas et al.[30] introduced user micro-profiling which is a new context-aware recommendation technique. It is a contextual pre-filtering technique that uses implicit user feedback. In micro-profiling, profile of user is divided into number of sub-profiles and each of which represents the user in a specific context. Recommendations are based on these micro-profiles rather than a single user model. Though, the user needs depends on the accurate division of the contextual variable. If we are using a continuous contextual variable and implicit feedback the detection of significant partition of the user profile and its assessment is a non-trivial subject. Considering this point an online assessment process is proposed for CARS.

Sang Hyun Choi et al.[32] proposed HYRED, a hybrid recommendation algorithm which combines collaborative filtering by means of modified Pearson’s binary correlation coefficients with content-based filtering by means of generalized distance to boundary based rating. It utilizes the nearest and farthest neighbors of a target user to yield useful information from large dataset avoiding sparsity and scalability problem. This increases the performance of the system and reduces the effort to perform computations.

Thai-Nghe, et al.[33] proposed a new approach which uses recommendation techniques for learning data mining, especially for forecasting student performance. To confirm this method, recommender system techniques with traditional regression procedures are compared such as logistic/linear regression by using learning information for intellectual tutoring systems. Experimental results demonstrate that the suggested approach can improve prediction results.

L.O. Colombo-Mendoza et al.[59] proposed a recommendation system in movie show times field. The presented system is named as RecomMetz which is a context aware mobile RS. The experimental results show that it performs efficiently in both a cold-start scenario and a no cold-start-scenario.

From the literature survey we conclude that the sometime recommendations are not according to the need and preferences of the users. Collaborative filtering approach suffers from several shortcomings. Scalability factor is attempted to improve the performance of the system in this paper. Recommendations are also influenced by the factors such as age, gender and some other user profile information. Considering this, age and gender information of the user along with collaborative filtering to provide recommendations is used. Further the genetic algorithm along with genre correlation to provide recommendations is used in this paper.

III. PROPOSED WORK

Recommendation process for our work mainly consists following steps. Firstly, find neighbors of the active or target user. Second, compute the similarities among the neighbors and the active users to calculate prediction. After the prediction is calculated, the items will be recommended by the systems but the item should attain a definite threshold of ranking to the active user. And finally predict fast recommendations to the user. The Android Studio is

positive refers to predicted classes of items that are not actually recommended and True negative is the difference of all three classes from total items. Table 2 shows the structure of confusion matrix.

Table 1: Confusion Matrix

	Relevant	Irrelevant
Recommended	True Positive (tp)	False Positive (fp)
Not Recommended	False Negative (fn)	True Negative (tn)

Precision or true positive accuracy is calculated as:

$$\text{Precision} = \text{tpa} = \frac{tp}{tp + fp} \times 100$$

Recall or true positive rate is calculated as:

$$\text{Recall} = \text{tpr} = \frac{tp}{tp + fn} \times 100$$

The F1-measure tries to combine precision and recall into a single score by calculating different types of means of both metrics. The F1-measure or F1-score is calculated as the standard harmonic mean of precision and recall:

$$F1 = \frac{2 \times \text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}}$$

The experimental analysis has been done using the above described parameters. The results obtained from the implementation of proposed system have been provided in the form of graphs. Figure 5.1 shows that the system will perform better as precision is high when the number of recommendations is less. Figure 5.2 shows the results obtained using recall. Figure 5.3 shows the results obtained using F1-measure which is harmonic mean of precision and recall.

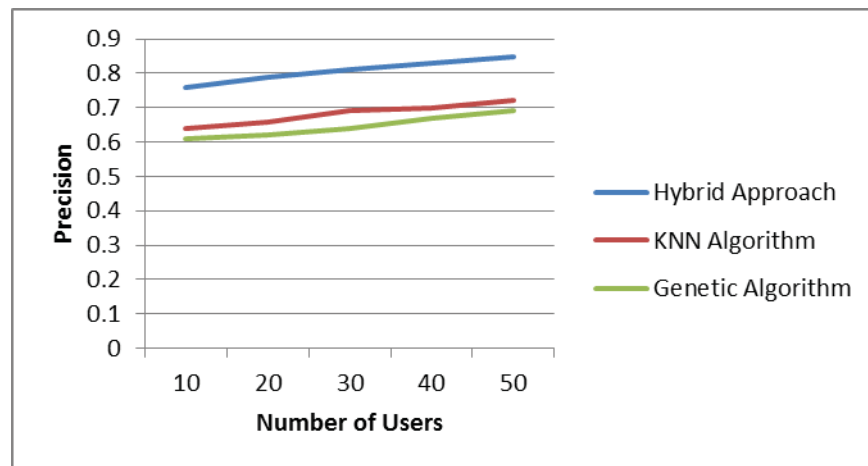


Figure 2. Precision vs Number of Users

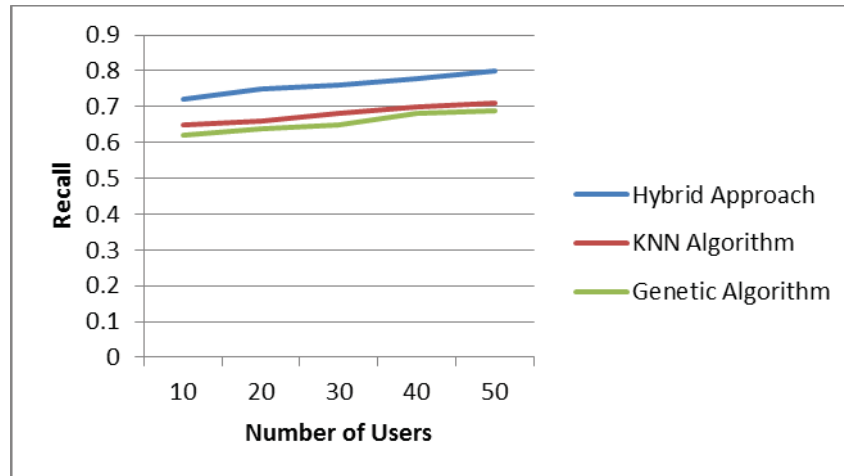


Figure 3. Recall vs Number of Users

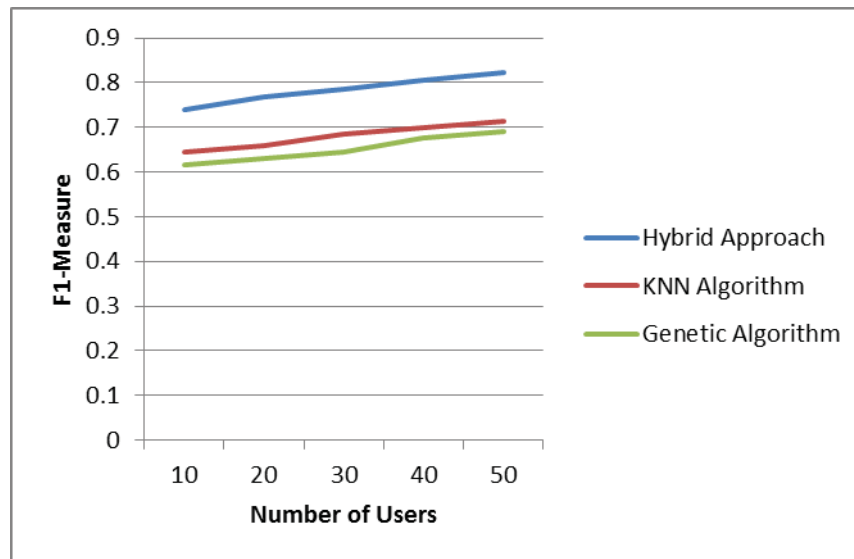


Figure 4. F1-measure vs Number of Users

V. CONCLUSION

To find out related content is very difficult task in current scenario where there are huge amount of data is stored in the databases. Recommender systems are solution to this problem and attracting researchers to explore this area in past few years. This thesis tries to solve the problem for recommending movies. In this thesis, an application is proposed which recommends relevant data to the user according to their preferences and history in movies. Wrong recommendations assigned to the user tend to decrease the efficiency of the system but this problem is reduced by using content and collaborative approach and optimal recommendations are provided to the user. . In our work, both content and collaborative techniques and some demographic information are combined into a hybrid approach, where additional content features are used to improve the accuracy of collaborative filtering. Also the genetic algorithm is used to provide recommendations to the user.

To build tags into the interaction model of applications, turn the actions and intentions of users into inferred or implied tags, then re-surface that information as a basis for explicit meta-tagging action later. Further research can be performed in the field of non-technical aspects like privacy resulting from user based tagging.

VI. ACKNOWLEDGEMENTS

We would like to thank Dr. (Prof.) Amit Verma the head of department at Chandigarh University and Assistant Prof. Maninder Singh at Chandigarh University and Prof. Amritpal Singh (Thapar University) for their ideas that helped guide our research.

REFERENCES

- [1] E. Rich, "User modeling via stereotypes," In: *Cognitive Science*, vol. 3, no. 4, pp. 329–354, October 1979.
- [2] D. Goldberg, D. Nichols, B. M. Oki, and D. Terry, "Using collaborative filtering to weave an information tapestry," In: *Communications of the ACM*, vol. 35, no. 12, pp. 61–70, 1992.
- [3] Belkin, Nicholas J., and W. Bruce Croft, "Information filtering and information retrieval: two sides of the same coin?" In: *Communications of the ACM* 35, no. 12, 29-38, 1992.
- [4] P. Resnick, N. Iacovou, M. Suchak, P. Bergstrom, and J. Riedl, "GroupLens: an open architecture for collaborative filtering of net news." In: *ACM CSCW '94*, pp. 175–186, ACM, 1994.
- [5] W. Hill, L. Stead, M. Rosenstein, and G. Furnas, "Recommending and evaluating choices in a virtual community of use." In: *ACM CHI '95*, pp. 194–201, ACM Press/Addison-Wesley Publishing Co., 1995.
- [6] U. Shardanand and P. Maes, "Social information filtering: Algorithms for automating "word of mouth"," In: *ACM CHI '95*, pp. 210–217, ACM Press/Addison-Wesley Publishing Co., 1995.
- [7] P. M. West, D. Ariely, S. Bellman, E. Bradlow, J. Huber, E. Johnson, B. Kahn, J. Little, and D. Schkade, "Agents to the Rescue?," In: *Marketing Letters*, vol. 10, no. 3, pp. 285–300, 1999.
- [8] Claypool, M., Gokhale, A., Miranda, T., Murnikov, P., Netes, D. and Sartin, M. "Combining Content-Based and Collaborative Filters in an Online Newspaper," In: *Proceedings of ACM SIGIR Workshop on Recommender Systems*, 1999.
- [9] A. Ansari, S. Essegaiier, and R. Kohli, "Internet recommendation systems," In: *Journal of Marketing Research*, vol. 37, no. 3, pp. 363–375, August 2000.
- [10] J. B. Schafer, J. A. Konstan, and J. Riedl, "E-Commerce recommendation applications," In: *Data Mining and Knowledge Discovery*, vol. 5, no. 1, pp. 115–153, January 2001.
- [11] Badrul Sarwar, George Karypis, Joseph Konstan, and John Riedl, "Item-Based Collaborative Filtering Recommendation Algorithms." In: *ACM*, 2001.
- [12] R. Burke, "Hybrid recommender systems: Survey and experiments," In: *User Modeling and User-Adapted Interaction*, vol. 12, no. 4, pp. 331–370, November 2002.
- [13] Qing Li, Byeong Man Kim, "Clustering Approach for Hybrid Recommender System," In: *Proceedings of the IEEE/WIC International Conference on Web Intelligence*, 2003.
- [14] Linden, G., Smith, B., & York, J., "Amazon. Com recommendations: Item-to-item collaborative filtering", In: *Internet Computing, IEEE*, 7(1), 76-80, 2003.
- [15] Miyahara, K., & Pazzani, M. J., "Collaborative filtering with the simple Bayesian classifier" In: *PRICAI 2000 Topics in Artificial Intelligence. Springer Berlin Heidelberg*, pp.679-689, 2000.
- [16] Wen-Yang Lin, Wen-Yuan Lee And Tzung-Pei Hong, "Adapting Crossover and Mutation Rates in Genetic Algorithms." In: *Journal of Information Science And Engineering* 19, pp. 889-903, 2003.
- [17] P Han, B Xie, F Yang, R Shen, "A scalable P2P recommender system based on distributed collaborative filtering." In: *Expert systems with applications* 27, pp. 203-210, 2004.
- [18] Jonathan L. Herlocker, Joseph A. Konstan, Loren G. Terveen, And John T. Riedl, "Evaluating Collaborative Filtering Recommender Systems," In: *ACM Transactions on Information Systems*, Vol. 22, No. 1, pp. 5–53, January 2004.
- [19] Mathes, A., "Folksonomies – Cooperative Classification and Communication Through Shared Metadata", In: *Adam Mathes.com*, USA, 2004.
- [20] Bridge, D., G'oker, M., McGinty, L., Smyth, B., "Case-based recommender systems," In: *Knowledge Engineering review* 20(3), pp. 315–320, 2006.
- [21] Ricci, F., Cavada, D., Mirzadeh, N., Venturini, A., "Case-based travel recommendations," In: *D.R. Fesenmaier, K.Woeber, H.Werthner (eds.) Destination Recommendation Systems: Behavioural Foundations and Applications*, pp. 67–93. CABL, 2006.
- [22] Janusz Sobecki, "Implementations of Web-based Recommender Systems Using Hybrid Methods", In: *International Journal of Computer Science & Applications*, Vol. 3. Issue 3, pp. 52 – 64, 2006.
- [23] Bhatia, Kapil. , "Collaborative Tagging for Software Reuse. Computer Science & Engineering Department," In: *Thapar Institute of Engineering & Technology*, Deemed University, 2006.
- [24] B. Smyth, "Case-based recommendation," In: *The Adaptive Web*, vol. 4321 of *Lecture Notes in Computer Science*, (P. Brusilovsky, A. Kobsa, and W. Nejdl, eds.), pp. 342–376, Springer, 2007.
- [25] Bamshad Mobasher, "Recommender Systems," In: *Kunstliche Intelligenz, Special Issue on Web Mining*, No. 3, pp. 41-43, 2007.
- [26] Burke, R., "Hybrid web recommender systems. In: *The Adaptive Web*," In: Springer Berlin / Heidelberg, pp. 377–408, 2007.
- [27] Bamshad Mobasher, Robin Burke, Runa Bhaumik, and Chad Williams, "Towards Trustworthy Recommender Systems: An Analysis of Attack Models and Algorithm Robustness," In: *ACM Transactions on Internet Technology*, Vol. 7, No. 2, May 2007.
- [28] V Formoso, F Casheda, V Carneiro, "Algorithms for Efficient Collaborative Filtering," In: *Efficiency Issues in Information Retrieval Workshop*, p. 17. 2008.
- [29] Manning, Christopher D., Prabhakar Raghavan, and Hinrich Schütze, "Introduction to information retrieval," In: *Vol. 1. Cambridge: Cambridge university press*, 2008.
- [30] Linas Baltrunas, Xavier Amatriain, "Towards Time-Dependant Recommendation based on Implicit Feedback." In: *ACM CARS, New York, USA*, 2009.
- [31] Gipp, Bela, Jöran Beel, and Christian Hentschel., "Scienstein: A research paper recommender system", In: *International Conference on Emerging Trends in Computing*, pp. 309-315, 2009.
- [32] Sang Hyun Choi, Young-Seon Jeong, and Myong K. Jeong, "A Hybrid Recommendation Method with Reduced Data for Large-Scale Application," In: *IEEE Transactions on Systems, Man, And Cybernetics—Part C: Applications And Reviews*, Vol. 40, No. 5, September 2010.

- [33]Thai-Nghe, Nguyen, Lucas Drumond, Artus Krohn-Grimberghe, and Lars Schmidt-Thieme, "Recommender system for predicting student performance", In: *Procedia Computer Science* 1, no. 2, pp. 2811-2819, 2010.
- [34]Pinata Winoto, Tiffany Y. Tang, "The role of user mood in movie recommendations." In: *Expert Systems with Applications* 37, pp: 6086–6092, 2010.
- [35]mTrip, Inc, www.mTrip-Intelligent Travel Guides, 2011.
- [36]Nan Zheng, Qiudan Li, "A recommender system based on tag and time information for social tagging system," In: *Expert Systems with Applications* 38, pp. 4575–4587, 2011.
- [37]Jesus Bobadilla, Fernando Ortega, Antonio Hernando, Javier Alcala, "Improving collaborative filtering recommender system results and performance using genetic algorithms," In: *Knowledge-Based Systems* 24, pp. 1310–1316, 2011.
- [38]Changbok, Hyokyung Chang, Yongho Kang, Yumi Bae, Hyosik Ahn, Euiin Choi, "Filtering Technique on Mobile Cloud Computing," In: *Energy Procedia* 16, pp. 1305 – 1311, 2011.
- [39]Bobadilla, J., Ortega, F., Hernando, A., & Alcalá, J., "Improving collaborative filtering recommender system results and performance using genetic algorithms", In: *Knowledge-based systems*, 24(8), 1310-1316, 2011.
- [40]Bahls, BradleyH, "Pedestrian Pal: A Route Recommendation System for the Android Mobile Phone" In: *Theses, Dissertations*, Professional Paper, Paper737, 2011.
- [41]K. Choi, D. Yoo, G. Kim, and Y. Suh, "A hybrid online-product recommendation system: Combining implicit rating-based collaborative filtering and sequential pattern analysis," In: *Electron. Commer. Res. Appl.*, vol. 11, no. 4, pp. 309–317, Jul. 2012.
- [42]J. Lee, M. Sun, and G. Lebanon, "A comparative study of collaborative filtering algorithms," In: *arXiv Prepr. arXiv1205.3193*, pp. 1–27, 2012.
- [43]Punam Bedi, Ravish Sharma, "Trust based recommender system using ant colony for trust computation," In: *Expert Systems with Applications* 39, pp. 1183–1190, 2012.
- [44]Tourist Eye, Inc, www.touristeye.com.TouristEye Web Application, 2012.
- [45]GuidePal, Inc, guidepal.com.GuidePal Home, 2012
- [46]Kate Starbird, Grace Muzny, Leysia Palen, "Learning from the Crowd: Collaborative Filtering Techniques for Identifying On-the-Ground Twitterers during Mass Disruptions," In: *Proceedings of the 9th International ISCRAM Conference – Vancouver, Canada*, April 2012.
- [47]Duda, Richard O., Peter E. Hart, and David G. Stork. Pattern classification. John Wiley & Sons, 2012.
- [48]Maria S. Pera, Yiu-Kai Ng, "A group recommender for movies based on content similarity and popularity," In: *Information Processing and Management* 49, pp. 673–687, 2013.
- [49]Beel, J., Langer, S., Genzmehr, M., Gipp, B. and Nurnberger, A., "A Comparative Analysis of Offline and Online Evaluations and Discussion of Research Paper Recommender System Evaluation," In: *Proceedings of the Workshop on Reproducibility and Replication in Recommender Systems Evaluation (RepSys) at the ACM Recommender System Conference (RecSys)*, 2013.
- [50]Mojtaba Salehi, Mohammad Pourzaferani, Seyed Amir Razavi, "Hybrid attribute-based recommender system for learning material using genetic algorithm and a multidimensional information model," In: *Egyptian Informatics Journal* 14, pp. 67–78, 2013.
- [51]Soojung Lee, "Collaborative Filtering based on Genre Preference." In: *International Journal of Advancements in Computing Technology (IJACT)*, Vol. 5, No 16, December 2013.
- [52]Woon-hae Jeong, Se-jun Kim, Doo-soon Park and Jin Kwak, "Performance Improvement of a Movie Recommendation System based on Personal Propensity and Secure Collaborative Filtering." In: *J Inf Process Syst*, Vol.9, No.1, March 2013.
- [53]Cui, L., & Shi, Y., "A Method based on one-class SVM for News Recommendation. *Procedia Computer Science*", 31, pp. 281-290, 2014.
- [54]Artem Umanets "GuideMe-A tourist Guide with a Recommender System and Social Interaction", In: *Conference on Electronics, Telecommunications and Computers-CETC*, ELSVIER, pp.407-414, 2014.
- [55]Peter Aksenov, Astrid Kemperman and Theo Arentze., "Toward personalised and dynamic cultural routing: a three-level approach", In: *12th International Conference on Design and Decision Support Systems in Architecture and Urban Planning*, DDS, ELSVIER, pp.257-269, 2014.
- [56] Linas Baltrunas, Francesco Ricci, "Experimental evaluation of context-dependent collaborative filtering using item splitting," In: *Springer, User Model User-Adap Inter*, pp. 7–34, 2014.
- [57]Ahmad Basheer Hassanat1, Mohammad Ali Abbadi2,Ghada Awad Altarawneh, Ahmad Ali Alhasanat, "Solving the Problem of the K Parameter in the KNN Classifier Using an Ensemble Learning Approach." In: *International Journal of Computer Science and Information Security*, Vol. 12, No. 8, August 2014.
- [58]Parivash Pirasteh, Dosam Hwang, Jason J. Jung, "Exploiting matrix factorization to asymmetric user similarities in recommendation systems." In: *Elsevier, Knowledge-Based Systems*, 2015.
- [59]Luis Omar Colombo-Mendoza, Rafael Valencia-Garcia, Alejandro Rodriguez-Gonzalez, Giner Alor-Hernandez, Jose Javier Samper-Zapater, "RecomMetz: A context-aware knowledge-based mobile recommender system for movie showtimes." In: *Expert Systems with Applications* 42 , pp: 1202–1222, 2015.
- [60]Youngki Park, Sungchan Park, Woosung Jung, Sang-goo Lee, "Reversed CF: A fast collaborative filtering algorithm using a k-nearest neighbor graph." In: *Expert Systems with Applications* 42, pp: 4022–4028, 2015.