Improving Collaborative Filtering's Rating Prediction Quality by Exploiting the Item Adoption Eagerness Information

Dionisis Margaris
margaris@di.uoa.gr
Department of Informatics and Telecommunications, University of Athens
Athens, Greece

Dimitris Spiliotopoulos
dspiliot@uop.gr
Department of Informatics and Telecommunications, University of the Peloponnese
Tripoli, Greece

Costas Vassilakis
costas@uop.gr
Department of Informatics and Telecommunications, University of the Peloponnese
Tripoli, Greece

ABSTRACT

Collaborative filtering generates recommendations tailored to the users' preferences by exploiting item ratings registered by users. Collaborative filtering algorithms firstly find people that have rated items in a similar fashion; these people are coined as "near neighbors" and their ratings on items are combined in the recommendation generation phase to predict ratings and generate recommendations. On the other hand, people exhibit different levels of eagerness to adopt new products: according to this characteristic, there is a set of users, termed as *"Early Adopters"*, who are prone to start using a product or technology as soon as it becomes available, in contrast to the majority of users, who prefer to start using items once they reach maturity; this important aspect of user behavior is not taken into account by existing algorithms. In this work, we propose an algorithm that considers the eagerness shown by users to adopt products, so as to leverage the accuracy of rating prediction. The proposed algorithm is evaluated using seven popular datasets.

CCS CONCEPTS

• **Information systems → Collaborative filtering**.

KEYWORDS

collaborative filtering, item adoption eagerness, rating prediction quality, evaluation, Pearson correlation coefficient, cosine similarity

ACM Reference Format:

Dionisis Margaris, Dimitris Spiliotopoulos, and Costas Vassilakis. 2019. Improving Collaborative Filtering's Rating Prediction Quality by Exploiting the Item Adoption Eagerness Information. In *IEEE/WIC/ACM International Conference on Web Intelligence (WI '19), October 14–17, 2019, Thessaloniki, Greece.* ACM, New York, NY, USA, 6 pages. https://doi.org/10.1145/3350546.3352544

1 INTRODUCTION

Collaborative filtering (CF) generates recommendations tailored to the users' preferences by exploiting item ratings registered by users. These ratings reflect the users' preferences and likings. For each user *u*, CF algorithms firstly find people that have rated items in a fashion similar to *u*; these people are coined as "*u*'s near neighbors" (NNs) and their ratings on items are combined in the recommendation generation phase to predict ratings that *u* would assign to items s/he has not reviewed yet [8], and ultimately generate recommendations for *u*. CF is based on the premise that people that have rated items similarly in the past are bound to continue doing so in the future as well [14, 20].

Many extensions of the basic CF algorithm have emerged, taking into account various features of the user profile [17, 46, 49, 60, 63], temporal behavior [15, 21, 25, 30, 31, 33, 35–37, 43] or inter-user relationships [6, 7, 22, 38]. Research has identified that people exhibit different levels of eagerness to adopt new products: according to this characteristic, there is a set of users, termed as *"Early Adopters"* (EA) [51], who are prone to adopting items eagerly, i.e. they start using a product or technology as soon as it becomes available, providing useful feedback, for other users and vendors. On the contrary, other users prefer to wait to use items, until they reach maturity. This behavioral aspect of users has been shown to be associated with different mentalities and psychological factors [26, 54], which in turn affect rating criteria. Nevertheless, this aspect is not taken into account by existing CF algorithms, while this is also true for individual elements of CF algorithms, most notably similarity measures [9, 45]

In this work, we introduce an algorithm that moderates the weight that each individual rating r on an item i is taken into account when formulating a rating prediction p on that specific item, by considering the item adoption eagerness information; this aspect is reflected on the particular item's adoption phases that the registered rating r and the prediction to be formulated p fall in. Effectively, the algorithm boosts the weight of r when both r and p fall into the early adoption phase, while it attenuates the weight of r when r and p fall into different adoption phases of i. We also evaluate the performance of the proposed algorithm, under different user similarity metrics and across seven datasets.

Furthermore, we have conducted and present an detailed comparative evaluation between (i) the proposed algorithm, (ii) the algorithm presented in [36] which is based on rating abstention intervals and (iii) the user variability-based algorithm presented in [37]. The algorithms in [36, 37] are state-of-the-art (both of them have been published in 2018) and utilize temporal information for increasing rating prediction accuracy. Furthermore, the algorithm in [37] does not necessitate additional information on users or items and maintains rating prediction accuracy levels, while the

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than the author(s) must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from permissions@acm.org.

WI '19, October 14–17, 2019, Thessaloniki, Greece

© 2019 Copyright held by the owner/author(s). Publication rights licensed to ACM.

ACM ISBN 978-1-4503-6934-3/19/10…$15.00
https://doi.org/10.1145/3350546.3352544

algorithm in [36] requires information about user social ties and exhibits coverage drops.

Finally, it is noteworthy that the algorithm proposed in this paper can be used in conjunction with other algorithms which aim to improve performance, reduce rating prediction errors and leverage recommendation quality in systems based on CF. Algorithms that can be used in conjunction with the proposed one include clustering-based techniques [16, 28], utilization of data sourced from social networks (SNs) [7, 38] or abrupt/gradual forgetting of old user ratings [30, 31].

The remaining of the paper is organized as follows: in Section 2 related work is overviewed, while the proposed algorithm is detailed in Section 3. Section 4 presents experiments conducted to tune and evaluate the proposed algorithm, as well as their results, while Section 5 concludes the paper and outlines future work.

2 RELATED WORK

Considerable research efforts have targeted the issue of CF-based systems accuracy. In this context, algorithms utilizing numerous characteristics of the ratings database and/or information from linked databases have been proposed [12, 20, 32, 40].

Rating timestamps is a feature utilized in numerous algorithms aiming to improve rating prediction accuracy. The algorithms presented in [30, 31] examine different methods for tacking the fact that old-aged ratings may not be aligned with the current user preferences. Towards this direction, old-aged ratings are either removed from the database (a method termed as *abrupt forgetting*), or their importance is attenuated (gradual forgetting). The abrupt forgetting methods have been shown to achieve higher benefits in terms of rating prediction accuracy, while additionally reducing the size of the ratings database. On the other hand, abrupt forgetting increases the sparsity of the ratings database, leading thus to some decrements in the rating prediction coverage, i.e. the capability to generate personalized rating predictions for users.

Knowledge-based recommender systems (KB-RSs) utilize higher-level knowledge regarding CF entities (i.e. items and users) to determine the items which meet the requirements of individual users and generate thus successful recommendations. Margaris et al. [39] present a KB-RS targeted to leisure time recommendations in the context of social media. This algorithm identifies influencing relationships among the users of the social network, while it additionally exploits (i) qualitative attributes associated with venues (e.g. atmosphere, price and service levels), (ii) the actual distance between the locations where venues are located and (iii) the individual users' profile and venue selection patterns.

With the proliferation of SN, many algorithms have emerged targeting the formulation of recommendations in the context of SN. Margaris et al. [38] consider the aspect of information diffusion in SNs in the context of recommendation generation, asserting that users' receptiveness to recommendations is not uniform across different item categories. Consequently, identifying and utilizing item category-specific sets of influencers for each user can lead to the formulation of more reliable recommendations, as compared to a model that employs a single set of influencers for each user. Trust propagation mechanisms within SN are considered in [41], which embeds this aspect in a matrix factorization-based RS.

Recently, the variability of user ratings has been recognized as a feature that can be exploited to improve rating prediction accuracy [37]. Additionally, the work in [36] exploits temporal information from the user rating database to identify periods that users have not submitted new ratings, which are termed as rating abstention intervals; the presence of rating abstention intervals is shown to be positively associated with a shift of interest, and therefore can provide the basis for amplifying or attenuating the weight assigned to user ratings in the recommendation generation process. The algorithm presented in [36] also calculates and utilizes influence levels between users, by considering the interaction that have taken place among users in social networks.

However, none of the aforementioned works considers the aspect of item adoption eagerness in the rating prediction computation. The present paper fills this gap by presenting an algorithm that leverages the similarity score of users whose ratings both belong to a particular item's EA phase, or both do not, and evaluates its performance using different user similarity metrics and datasets.

3 THE PROPOSED ALGORITHM

The basic CF formula for calculating a rating prediction $p_{U,i}$ for the rating of user U on item i is depicted in formula (1) [8]:

$$p_{U,i} = \overline{r_U} + \frac{\sum_{V \in NN_u} sim(U,V) * (r_{V,i} - \overline{r_V})}{\sum_{V \in NN_u} sim(U,V)} \tag{1}$$

where $\overline{r_U}$ and $\overline{r_V}$ are the mean values or ratings submitted by users U and V, *sim(U,V)* is a quantification of the similarity between users U and V, while NN_u denotes U's NNs.

The algorithm proposed in this paper adapts the prediction computation formula, by introducing terms that correspond to items' adoption eagerness (IAE). More specifically, formula (1) is modified as follows:

$$p_{U,i} = \overline{r_U} + \frac{\sum_{V \in NN_u} sim(U,V) * IAE\_factor(U,V,i) * (r_{V,i} - \overline{r_V})}{\sum_{V \in NN_u} sim(U,V) * IAE\_factor(U,V,i)} \tag{2}$$

where the $IAE\_factor(U,V,i)$ is a factor moderating the importance of rating $r_{V,i}$ in the context of the computation of prediction $p_{U,i}$, taking into account whether users U and V have adopted item i eagerly ("early adopters") or not ("late adopters"). The rationale behind the usage of $IAE\_factor(U,V,i)$ is that users exhibiting different degrees of item adoption eagerness assess and rate items with different mentalities and criteria, hence ratings entered within an item's early adoption phase convey an eager adopter's view and will thus be more useful for other eager adopters, but of less value for late adopters. A similar remark holds for the ratings of late adopters.

More specifically, the computation of the $IAE\_factor(U,V,i)$ quantity takes into account whether the (factual) rating of user V for item i and the (predicted) rating of user U on item i both belong to the early adoption phase of item i (denoted as EA_i), or not (i.e. they both belong to the late adoption phase or they belong to different phases). Formula (3) illustrates the computation method for the $IAE\_factor(U,V,i)$ quantity.

$$IAE\_factor(U,V,i) = \begin{cases} EA & \text{if } r_{U,i} \in EA_i \land r_{V,i} \in EA_i \\ LA & \text{if } r_{U,i} \notin EA_i \land r_{V,i} \notin EA_i \\ DIFF & \text{otherwise} \end{cases} \quad (3)$$

In formula (3) the *EA* is a constant that is used when both users' ratings on item *i*, belong to the item's *EA* lifetime phase; and similarly *LA* is a constant employed when both users' ratings on item *i*, belong to its Late Adoption lifetime phase. The *DIFF* constant is employed when the two ratings belong to different lifetime phases of item *i* (Early and Late).

According to [51], the early adoption phase for an item corresponds to the initial 16% of the item lifespan in the market. In some cases, the lifetime of the product is available through the official vendor pages (e.g. [10]). In the absence of official information, the lifetime of the product is approximated as follows:

(1) the beginning of the product lifespan is set to the timestamp of the earliest rating on the product within the ratings database,
(2) if the category of the product has a nominal lifespan (e.g. [26] reports that the lifetime of mobile phones is 3 years while that of cars is 10 years), then the end the product lifespan is computed by adding the nominal product category lifespan to the beginning of the product lifespan. Otherwise, the end of the product lifespan is set equal to the timestamp of the most recent rating on the product within the ratings database.

Taking the above into account, a rating $r_{U,i}$ belongs to the early adoption phase of item *i* if the timestamp of $r_{U,i}$ belongs to the first 16% of the lifespan of item *i* in the market; otherwise $r_{U,i}$ belongs to the product's late adoption phase.

In the next section, we explore the optimal setting for parameters *EA*, *LA* and *DIFF*, while we also evaluate the proposed algorithm's performance.

4 ALGORITHM TUNING AND PERFORMANCE EVALUATION

In this section, we present the experiments conducted to:

(1) Calculate the optimal values for parameters *EA*, *LA* and *DIFF*, which are used in the *IAE_factor* function of the presented algorithm.
(2) Compute the prediction error reduction, introduced by the presented algorithm, due to the consideration of the item adoption eagerness information in the CF rating prediction computation process.

Due to space limitations, detailed information on the experiments and their results is listed in [29].

In order to compute the optimal values for parameters *EA*, *LA* and *DIFF*, we experimentally searched the parameter value assignment solution space, by iteratively selecting parameter value assignments and assessing the impact that each particular parameter value assignment had on the accuracy of rating prediction. Rating prediction accuracy was quantified using two popular error metrics, namely the Mean Absolute Error (MAE), and the Root Mean Squared Error (RMSE). We opted to use two metrics, because each one of them highlights different aspects of the quality of the results: more specifically, the MAE metric treats all error magnitudes uniformly since it averages the absolute values of errors; on the other hand the RMSE metric squares error magnitudes before summing them up, therefore larger errors are emphasized. The error between an individual prediction and the corresponding actual rating was computed using the standard "hide one" technique [13, 14, 27, 36, 37]: the rating was hidden, and its value was subsequently predicted by combining non-hidden ratings. In our first experiment, only the last rating of each user was hidden and then its value was predicted; we also executed a second experiment, where -for each user- a random rating was hidden and then its value was predicted. The rating prediction quality metrics obtained from these two experiments were in close agreement (the absolute magnitude of their differences had an upper bound of 1.8% in all cases), and thus we confine ourselves to presenting only the results of the first experiment. All reported experiments were run on seven datasets, five which have been sourced from Amazon [2, 42], while the remaining two are sourced from MovieLens [18, 44]. The datasets sourced from Amazon are relatively sparse, while the ones sourced from MovieLens are relatively dense; a dataset's density is computed as $density(DS) = \frac{\#ratings}{\#items * \#users}$ [52].

In the next two subsections, we outline and discuss the results obtained from the conducted experiments.

4.1 Tuning of algorithm parameters

The first experiment aimed to determine the optimal values for the parameters EA, LA and DIFF. Since only the ratios EA/LA and $LA/DIFF$ (and not the actual parameter values) affect the algorithm performance, we fix *LA* to 1 and vary the values of *EA* and *DIFF*. We explored both the Pearson and cosine similarity measures (PCC and CF, respectively) and under both metrics, the setting attaining the highest reduction in the MAE and the RMSE is when the *DIFF* parameter is set to 0.5 and the *EA* parameter is set to 2.0. More specifically, this setting achieves an average MAE reduction of 3.7% and an average RMSE reduction of 3.18%, when the PCC metric is used. The respective reductions concerning the CS metric are 3.55% and 3.02%. Figure 1 depicts MAE reduction under different *EA* and *DIFF* parameter value combinations, using the PCC measure.

4.2 Performance evaluation

In this subsection, we present the results produced by the presented algorithm and contrast them with the ones produced by the algorithm proposed in [37], i.e. the CF variability algorithm. We selected the CF variability algorithm for the comparison because (i) it has been published in 2018, and thus is a state-of-the-art algorithm, (ii) it targets the improvement of prediction accuracy in the context of CF, (iii) it does not require any additional information about users or items (e.g. item taxonomical information or social ties between users) and (iv) it maintains prediction coverage levels. We note here that no other algorithm addresses the particular aspect of user behavior considered by the proposed algorithm; thus, in the absence of such an algorithm, the comparison is made with an algorithm that exploits similar features of ratings (i.e. temporal features).

Taking into account the results of subsection 4.1, we set parameters *EA* and *DIFF* to 2.0 and 0.5, respectively.

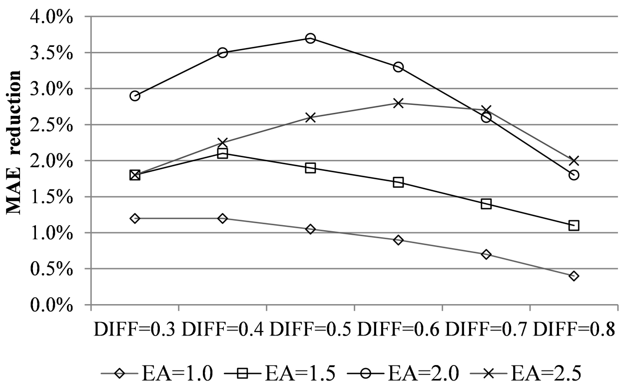

Figure 1: MAE reduction under different EA and DIFF parameter value combinations, using the PCC measure

Figure 2 depicts the reduction in the MAE attained by the proposed algorithm, along with the respective attainment of the CF variability algorithm, proposed in [37]. In both cases, MAE reduction percentages are calculated against the performance of the plain CF algorithm.

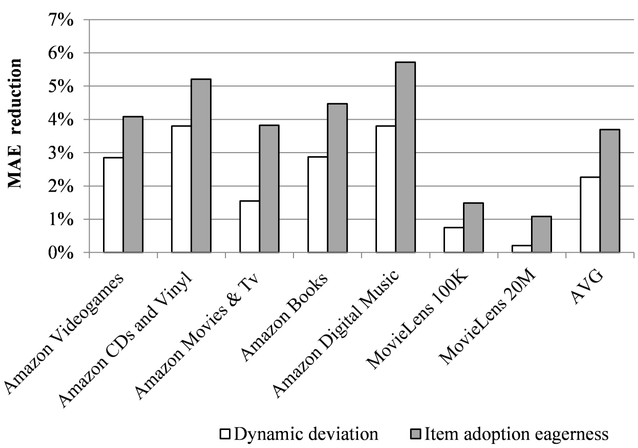

Figure 2: MAE reduction achieved by the proposed algorithm, in comparison to the CF variability algorithm [37]

In Fig. 2 we can notice that the proposed algorithm outperforms CF variability algorithm, in all tested datasets. More specifically, the proposed algorithm reduces the MAE by 3.7%, which is 63.4% higher than the reduction attained by the CF variability algorithm (2.26%). The widest performance margin is observed for the MovieLens 20M dataset (414%), while the narrowest one for the Amazon CDs and Vinyl dataset (37%). As far as the RMSE is concerned, the reduction achieved by the proposed algorithm is equal to 3.18%, while the reduction attained by the CF variability algorithm is 1.48%, therefore the performance edge is 114.6%. The experiment and its results is described in more detail in [29].

Finally, we give a performance comparison between the proposed algorithm and the rating abstention-based algorithm presented in [36]; the algorithm in [36] is a state-of-the-art algorithm utilizing temporal, within-user history information to reduce prediction errors, while it has been demonstrated to achieve higher error reductions than other state-of-the art algorithms. As noted above, the MAE reduction attained by the proposed algorithm averages to 3.7% over all tested datasets, surpassing the corresponding improvements of the rating abstention-based algorithm reported in [36], which average to 2.99%. While the absolute difference is limited to 0.7% and the relative difference is 23.7%, we emphasize that the rating abstention-based algorithm [36]:

- necessitates the existence of information about user social ties, which is not always available.
- exhibits a drop in coverage, which is substantial in the context of sparse datasets.

On the contrary, the algorithm presented in this paper fully maintains coverage levels and does not necessitate any additional information. It is also noteworthy that the rating abstention-based algorithm presented in [36] has been demonstrated to outperform other state-of-the-art algorithms, e.g. the pruning-based algorithms in [31, 34] and the temporal dynamics-based algorithm reported in [23].

5 CONCLUSION AND FUTURE WORK

In this paper, we proposed an algorithm that incorporates, in the rating prediction computation process, the aspect of the users' eagerness to adopt new items and technologies, aiming to increase prediction accuracy. We also reported on a set of experiments conducted to validate the performance of the proposed algorithm: in these experiments we used two user similarity metrics and seven datasets, both sparse and dense. These experiments showed that the consideration of the item adoption eagerness aspect entails considerable prediction accuracy gains.

We have also compared the proposed algorithm against (i) the user rating variability algorithm [37] and (ii) the rating abstention-based algorithm presented in [36]. The proposed algorithm outperformed both these algorithms.

The proposed algorithm can be straightforwardly incorporated in a CF-based RS, since (1) it does not require any extra information on users or items, (2) it introduces only minimal processing overhead for rating prediction formulation, which has been quantified to be less than 3%, indicating the feasibility of this approach, (3) it needs minimal additional storage space, for computing and storing only each item's EA phase, (4) it can be directly implemented as a modification of existing CF-based systems and (5) it can be used in conjunction with other algorithms that aim to leverage rating prediction accuracy, performance and/or coverage.

In our future work we plan to study additional methods to improve rating prediction quality in CF. Furthermore, we will study the algorithm's performance under more similarity metrics, such as the Euclidean distance and the Spearman coefficient [19]. We will also work on a suitable integration of the proposed method into matrix factorization techniques [24]. Finally, this work may be used with other algorithms to improve accuracy in language analysis [1, 3, 55, 58, 59, 61, 62], semantics [11, 48, 50, 56] and analysis of social media content [4, 5, 47, 53, 57].

REFERENCES

[1] Jan Alexandersson, Maria Aretoulaki, Nick Campbell, Michael Gardner, Andrey Girenko, Dietrich Klakow, Dimitris Koryzis, Volha Petukhova, Marcus Specht, Dimitris Spiliotopoulos, Alexander Stricker, and Niels Taatgen. 2014. Metalogue: A Multiperspective Multimodal Dialogue System with Metacognitive Abilities for Highly Adaptive and Flexible Dialogue Management. In *2014 International Conference on Intelligent Environments*. IEEE, 365–368. https://doi.org/10.1109/IE.2014.67

[2] Amazon. 2019. Amazon product data. http://jmcauley.ucsd.edu/data/amazon/links.html

[3] Ion Androutsopoulos, Dimitris Spiliotopoulos, Konstantinos Stamatakis, Aggeliki Dimitromanolaki, Vangelis Karkaletsis, and Constantine D. Spyropoulos. 2002. Symbolic Authoring for Multilingual Natural Language Generation. In *Methods and Applications of Artificial Intelligence. Proceedings of the 2nd Hellenic Conference in Artificial Intelligence*, Ioannis P. Vlahavas and Constantine D. Spyropoulos (Eds.). Springer Berlin Heidelberg, Berlin, Heidelberg, 131–142.

[4] Despoina Antonakaki, Dimitris Spiliotopoulos, Christos V. Samaras, Sotiris Ioannidis, and Paraskevi Fragopoulou. 2016. Investigating the complete corpus of referendum and elections tweets. In *2016 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining (ASONAM)*. IEEE, 100–105. https://doi.org/10.1109/ASONAM.2016.7752220

[5] Despoina Antonakaki, Dimitris Spiliotopoulos, Christos V. Samaras, Polyvios Pratikakis, Sotiris Ioannidis, and Paraskevi Fragopoulou. 2017. Social media analysis during political turbulence. *PLOS ONE* 12, 10 (10 2017), 1–23. https://doi.org/10.1371/journal.pone.0186836

[6] Ofer Arazy, Nanda Kumar, and Bracha Shapira. 2009. Improving Social Recommender Systems. *IT Professional* 11, 4 (July 2009), 38–44. https://doi.org/10.1109/MITP.2009.76

[7] Eytan Bakshy, Dean Eckles, Rong Yan, and Itamar Rosenn. 2012. Social Influence in Social Advertising: Evidence from Field Experiments. In *Proceedings of the 13th ACM Conference on Electronic Commerce (EC '12)*. ACM, New York, NY, USA, 146–161. https://doi.org/10.1145/2229012.2229027

[8] Marko Balabanović and Yoav Shoham. 1997. Fab: Content-based, Collaborative Recommendation. *Commun. ACM* 40, 3 (March 1997), 66–72. https://doi.org/10.1145/245108.245124

[9] Laurent Candillier, Frank Meyer, and Françoise Fessant. 2008. Designing Specific Weighted Similarity Measures to Improve Collaborative Filtering Systems. In *Proceedings of the 8th Industrial Conference on Advances in Data Mining: Medical Applications, E-Commerce, Marketing, and Theoretical Aspects (ICDM '08)*. Springer-Verlag, Berlin, Heidelberg, 242–255. https://doi.org/10.1007/978-3-540-70720-2_19

[10] CISCO. 2013. *End-of-Life and End-of-Sale Notices for Cisco ACE 4700 Series Application Control Engine Appliances*. CISCO. https://www.cisco.com/c/en/us/products/application-networking-services/ace-4700-series-application-control-engine-appliances/eos-eol-notice-listing.html

[11] Elena Demidova, Nicola Barbieri, Stefan Dietze, Adam Funk, Helge Holzmann, Diana Maynard, Nikolaos Papailiou, Wim Peters, Thomas Risse, and Dimitris Spiliotopoulos. 2014. Analysing and Enriching Focused Semantic Web Archives for Parliament Applications. *Future Internet* 6, 3 (2014), 433–456. https://doi.org/10.3390/fi6030433

[12] Ricardo Dias and Manuel J. Fonseca. 2013. Improving Music Recommendation in Session-Based Collaborative Filtering by Using Temporal Context. In *2013 IEEE 25th International Conference on Tools with Artificial Intelligence*. IEEE, 783–788. https://doi.org/10.1109/ICTAI.2013.120

[13] Margaris Dionisis, Vassilakis Costas, and Georgiadis Panagiotis. 2013. An Integrated Framework for QoS-based Adaptation and Exception Resolution in WS-BPEL Scenarios. In *Proceedings of the 28th Annual ACM Symposium on Applied Computing (SAC '13)*. ACM, New York, NY, USA, 1900–1906. https://doi.org/10.1145/2480362.2480714

[14] Michael D. Ekstrand, John T. Riedl, and Joseph A. Konstan. 2011. Collaborative Filtering Recommender Systems. *Found. Trends Hum.-Comput. Interact.* 4, 2 (Feb. 2011), 81–173. https://doi.org/10.1561/1100000009

[15] João Gama, Indrė Žliobaitė, Albert Bifet, Mykola Pechenizkiy, and Abdelhamid Bouchachia. 2014. A Survey on Concept Drift Adaptation. *ACM Comput. Surv.* 46, 4, Article 44 (March 2014), 37 pages. https://doi.org/10.1145/2523813

[16] Songjie Gong. 2010. A collaborative filtering recommendation algorithm based on user clustering and item clustering. *Journal of Software* 5, 7 (2010), 745–752.

[17] Alejandro González, Javier Torres-Niño, Enrique Jiménez-Domingo, Juan Gómez Berbis, and Giner Alor-Hernández. 2012. AKNOBAS: A Knowledge-based Segmentation Recommender System based on Intelligent Data Mining Techniques. *Computer Science and Information Systems* 9 (06 2012). https://doi.org/10.2298/CSIS110722008R

[18] Franklin Maxwell Harper and Joseph A. Konstan. 2015. The MovieLens Datasets: History and Context. *ACM Trans. Interact. Intell. Syst.* 5, 4, Article 19 (Dec. 2015), 19 pages. https://doi.org/10.1145/2827872

[19] Jonathan L. Herlocker, Joseph A. Konstan, Loren G. Terveen, and John T. Riedl. 2004. Evaluating Collaborative Filtering Recommender Systems. *ACM Trans. Inf. Syst.* 22, 1 (Jan. 2004), 5–53. https://doi.org/10.1145/963770.963772

[20] Kai Yu, Anton Schwaighofer, Volker Tresp, Xiaowei Xu, and Hans-Peter Kriegel. 2004. Probabilistic memory-based collaborative filtering. *IEEE Transactions on Knowledge and Data Engineering* 16, 1 (Jan 2004), 56–69. https://doi.org/10.1109/TKDE.2004.1264822

[21] Noam Koenigstein, Gideon Dror, and Yehuda Koren. 2011. Yahoo! Music Recommendations: Modeling Music Ratings with Temporal Dynamics and Item Taxonomy. In *Proceedings of the Fifth ACM Conference on Recommender Systems (RecSys '11)*. ACM, New York, NY, USA, 165–172. https://doi.org/10.1145/2043932.2043964

[22] Ioannis Konstas, Vassilios Stathopoulos, and Joemon M. Jose. 2009. On Social Networks and Collaborative Recommendation. In *Proceedings of the 32Nd International ACM SIGIR Conference on Research and Development in Information Retrieval (SIGIR '09)*. ACM, New York, NY, USA, 195–202. https://doi.org/10.1145/1571941.1571977

[23] Yehuda Koren. 2009. Collaborative Filtering with Temporal Dynamics. In *Proceedings of the 15th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD '09)*. ACM, New York, NY, USA, 447–456. https://doi.org/10.1145/1557019.1557072

[24] Yehuda Koren, Robert Bell, and Chris Volinsky. 2009. Matrix Factorization Techniques for Recommender Systems. *Computer* 42, 8 (Aug. 2009), 30–37. https://doi.org/10.1109/MC.2009.263

[25] Lei Li, Li Zheng, Fan Yang, and Tao Li. 2014. Modeling and broadening temporal user interest in personalized news recommendation. *Expert Systems with Applications* 41, 7 (2014), 3168 – 3177. https://doi.org/10.1016/j.eswa.2013.11.020

[26] John W. Loy. 1969. Social Psychological Characteristics of Innovators. *American Sociological Review* 34, 1 (1969), 73–82. http://www.jstor.org/stable/2092788

[27] Dionisis Margaris, Panagiotis Georgiadis, and Costas Vassilakis. 2013. Adapting WS-BPEL scenario execution using collaborative filtering techniques. In *IEEE 7th International Conference on Research Challenges in Information Science (RCIS)*. IEEE, 1–11. https://doi.org/10.1109/RCIS.2013.6577691

[28] Dionisis Margaris, Panagiotis Georgiadis, and Costas Vassilakis. 2015. A collaborative filtering algorithm with clustering for personalized web service selection in business processes. In *2015 IEEE 9th International Conference on Research Challenges in Information Science (RCIS)*. IEEE, 169–180. https://doi.org/10.1109/RCIS.2015.7128877

[29] Dionisis Margaris, Dimitris Spiliotopoulos, and Costas Vassilakis. 2019. *Experimental results for considering Item Adoption Eagerness Information in Collaborative Filtering's Rating Prediction.* Technical Report. Software and Database Systems lab, University of the Peloponnese. TR-19003,https://soda.dit.uop.gr/?q=TR-19003

[30] Dionisis Margaris and Costas Vassilakis. 2016. Pruning and aging for user histories in collaborative filtering. In *2016 IEEE Symposium Series on Computational Intelligence (SSCI)*. IEEE, 1–8. https://doi.org/10.1109/SSCI.2016.7849920

[31] Dionisis Margaris and Costas Vassilakis. 2017. Enhancing User Rating Database Consistency Through Pruning. In *Transactions on Large-Scale Data- and Knowledge-Centered Systems XXXIV: Special Issue on Consistency and Inconsistency in Data-Centric Applications*, Abdelkader Hameurlain, Josef Küng, Roland Wagner, and Hendrik Decker (Eds.). Springer Berlin Heidelberg, Berlin, Heidelberg, 33–64. https://doi.org/10.1007/978-3-662-55947-5_3

[32] Dionisis Margaris and Costas Vassilakis. 2017. Exploiting Internet of Things Information to Enhance Venues' Recommendation Accuracy. *Serv. Oriented Comput. Appl.* 11, 4 (Dec. 2017), 393–409. https://doi.org/10.1007/s11761-017-0216-y

[33] Dionisis Margaris and Costas Vassilakis. 2017. Improving Collaborative Filtering's Rating Prediction Quality by Considering Shifts in Rating Practices. In *2017 IEEE 19th Conference on Business Informatics (CBI)*, Vol. 01. IEEE, 158–166. https://doi.org/10.1109/CBI.2017.24

[34] Dionisis Margaris and Costas Vassilakis. 2017. Improving collaborative filtering's rating prediction quality in dense datasets, by pruning old ratings. In *2017 IEEE Symposium on Computers and Communications (ISCC)*. 1168–1174. https://doi.org/10.1109/ISCC.2017.8024683

[35] Dionisis Margaris and Costas Vassilakis. 2018. Enhancing Rating Prediction Quality Through Improving the Accuracy of Detection of Shifts in Rating Practices. In *Transactions on Large-Scale Data- and Knowledge-Centered Systems XXXVII*, Abdelkader Hameurlain and Roland Wagner (Eds.). Springer Berlin Heidelberg, Berlin, Heidelberg, 151–191. https://doi.org/10.1007/978-3-662-57932-9_5

[36] Dionisis Margaris and Costas Vassilakis. 2018. Exploiting Rating Abstention Intervals for Addressing Concept Drift in Social Network Recommender Systems. *Informatics* 5, 2 (2018), 20. https://doi.org/10.3390/informatics5020021

[37] Dionisis Margaris and Costas Vassilakis. 2018. Improving Collaborative Filtering's Rating Prediction Accuracy by Considering Users' Rating Variability. In *2018 IEEE 16th Intl Conf on Dependable, Autonomic and Secure Computing, 16th Intl Conf on Pervasive Intelligence and Computing, 4th Intl Conf on Big Data Intelligence and Computing and Cyber Science and Technology Congress (DASC/PiCom/DataCom/CyberSciTech)*. IEEE, 1022–1027. https://doi.org/10.1109/DASC/PiCom/DataCom/CyberSciTec.2018.00145

[38] Dionisis Margaris, Costas Vassilakis, and Panagiotis Georgiadis. 2016. Recommendation information diffusion in social networks considering user influence and semantics. *Social Network Analysis and Mining* 6, 1 (25 Nov 2016), 108.

https://doi.org/10.1007/s13278-016-0416-z

[39] Dionisis Margaris, Costas Vassilakis, and Panagiotis Georgiadis. 2017. Knowledge-Based Leisure Time Recommendations in Social Networks. In *Current Trends on Knowledge-Based Systems*, Giner Alor-Hernández and Rafael Valencia-García (Eds.). Springer International Publishing, Cham, 23–48. https://doi.org/10.1007/978-3-319-51905-0_2

[40] Dionisis Margaris, Costas Vassilakis, and Panagiotis Georgiadis. 2018. Query personalization using social network information and collaborative filtering techniques. *Future Generation Computer Systems* 78 (2018), 440 – 450. https://doi.org/10.1016/j.future.2017.03.015

[41] Ibrahim Mashal, Tein-Yaw Chung, and Osama Alsaryrah. 2015. Toward service recommendation in Internet of Things. In *2015 Seventh International Conference on Ubiquitous and Future Networks*. IEEE, 328–331. https://doi.org/10.1109/ICUFN.2015.7182559

[42] Julian McAuley, Christopher Targett, Qinfeng Shi, and Anton van den Hengel. 2015. Image-Based Recommendations on Styles and Substitutes. In *Proceedings of the 38th International ACM SIGIR Conference on Research and Development in Information Retrieval (SIGIR '15)*. ACM, New York, NY, USA, 43–52. https://doi.org/10.1145/2766462.2767755

[43] Leonardo, L Minku, Alan, P. White, and Xin Yao. 2010. The Impact of Diversity on Online Ensemble Learning in the Presence of Concept Drift. *IEEE Transactions on Knowledge and Data Engineering* 22, 5 (May 2010), 730–742. https://doi.org/10.1109/TKDE.2009.156

[44] MovieLens. 2019. MovieLens datasets. http://grouplens.org/datasets/movielens/

[45] Bidyut Kr. Patra, Raimo Launonen, Ville Ollikainen, and Sukumar Nandi. 2015. A new similarity measure using Bhattacharyya coefficient for collaborative filtering in sparse data. *Knowledge-Based Systems* 82 (2015), 163 – 177. https://doi.org/10.1016/j.knosys.2015.03.001

[46] David M. Pennock, Eric Horvitz, Steve Lawrence, and C. Lee Giles. 2000. Collaborative Filtering by Personality Diagnosis: A Hybrid Memory and Model-Based Approach. In *Proceedings of the 16th Conference on Uncertainty in Artificial Intelligence (UAI '00)*. Morgan Kaufmann Publishers Inc., San Francisco, CA, USA, 473–480. http://dl.acm.org/citation.cfm?id=647234.720062

[47] Georgios Petasis, Dimitrios Spiliotopoulos, Nikos Tsirakis, and Panayiotis Tsantilas. 2014. Sentiment Analysis for Reputation Management: Mining the Greek Web. In *Artificial Intelligence: Methods and Applications*, Aristidis Likas, Konstantinos Blekas, and Dimitris Kalles (Eds.). Springer International Publishing, Cham, 327–340.

[48] Alexandros Pino, Georgios Kouroupetroglou, Hernisa Kacorri, Anna Sarantidou, and Dimitris Spiliotopoulos. 2010. An Open Source / Freeware Assistive Technology Software Inventory. In *Proceedings of the 12th International Conference on Computers Helping People with Special Needs: Part I (ICCHP'10)*. Springer-Verlag, Berlin, Heidelberg, 178–185. http://dl.acm.org/citation.cfm?id=1886667.1886700

[49] Lara Quijano-Sánchez, Juan A Recio-García, and Belén Díaz-Agudo. 2011. Group recommendation methods for social network environments. In *3rd Workshop on Recommender Systems and the Social Web within the 5th ACM International Conference on Recommender Systems (RecSys' 11)*. 24–31.

[50] Thomas Risse, Elena Demidova, Stefan Dietze, Wim Peters, Nikolaos Papailiou, Katerina Doka, Yannis Stavrakas, Vassilis Plachouras, Pierre Senellart, Florent Carpentier, Amin Mantrach, Bogdan Cautis, Patrick Siehndel, and Dimitris Spiliotopoulos. 2014. The ARCOMEM Architecture for Social- and Semantic-Driven Web Archiving. *Future Internet* 6, 4 (2014), 688–716. https://doi.org/10.3390/fi6040688

[51] Everett M. Rogers. 2003. *Diffusion of innovations* (5th ed.). Free Press, New York, NY [u.a.]. 576 pages.

[52] J. Ben Schafer, Dan Frankowski, Jon Herlocker, and Shilad Sen. 2007. Collaborative Filtering Recommender Systems. In *The Adaptive Web: Methods and Strategies of Web Personalization*, Peter Brusilovsky, Alfred Kobsa, and Wolfgang Nejdl (Eds.). Springer Berlin Heidelberg, Berlin, Heidelberg, 291–324. https://doi.org/10.1007/978-3-540-72079-9_9

[53] Guenther Schefbech, Dimitris Spiliotopoulos, and Thomas Risse. 2012. The Recent Challenge in Web Archiving: Archiving the Social Web. In *Proceedings of the International Council on Archives Congress*. 1–5.

[54] Jagdish N Sheth and Walter H Stellner. 1979. *Psychology of innovation resistance: The less developed concept (LDC) in diffusion research*. Vol. BEBR No. 622. College of Commerce and Business Administration, University of Illinois at Urbana-Champaign Stacks.

[55] Dimitris Spiliotopoulos, Panagiota Stavropoulou, and Georgios Kouroupetroglou. 2009. Acoustic Rendering of Data Tables Using Earcons and Prosody for Document Accessibility. In *Proceedings of the 13th International Conference on Human-Computer Interaction*, Constantine Stephanidis (Ed.). Springer Berlin Heidelberg, Berlin, Heidelberg, 587–596.

[56] Dimitris Spiliotopoulos, Pepi Stavropoulou, and Georgios Kouroupetroglou. 2009. Spoken Dialogue Interfaces: Integrating Usability. In *Proceedings of the 5th Symposium of the Austrian HCI and Usability Engineering Group*, Andreas Holzinger and Klaus Miesenberger (Eds.). Springer Berlin Heidelberg, Berlin, Heidelberg, 484–499.

[57] Dimitris Spiliotopoulos, Efstratios Tzoannos, Pepi Stavropoulou, Georgios Kouroupetroglou, and Alexandros Pino. 2012. Designing User Interfaces for Social Media Driven Digital Preservation and Information Retrieval. In *13th International Conference on Computers Helping People with Special Needs*. Springer Berlin Heidelberg, Berlin, Heidelberg, 581–584.

[58] Dimitris Spiliotopoulos, Gerasimos Xydas, and Georgios Kouroupetroglou. 2005. Diction Based Prosody Modeling in Table-to-Speech Synthesis. In *Proceedings of the 8th International Conference on Text, Speech and Dialogue*. Springer Berlin Heidelberg, Berlin, Heidelberg, 294–301.

[59] Dimitris Spiliotopoulos, Gerasimos Xydas, Georgios Kouroupetroglou, Vasilios Argyropoulos, and Kalliopi Ikospentaki. 2010. Auditory universal accessibility of data tables using naturally derived prosody specification. *Universal Access in the Information Society* 9, 2 (01 Jun 2010), 169–183. https://doi.org/10.1007/s10209-009-0165-0

[60] Kazunari Sugiyama, Kenji Hatano, and Masatoshi Yoshikawa. 2004. Adaptive Web Search Based on User Profile Constructed Without Any Effort from Users. In *Proceedings of the 13th International Conference on World Wide Web (WWW '04)*. ACM, New York, NY, USA, 675–684. https://doi.org/10.1145/988672.988764

[61] Gerasimos Xydas, Dimitris Spiliotopoulos, and Georgios Kouroupetroglou. 2003. Modelling Emphatic Events from Non-Speech Aware Documents in Speech Based User Interfaces. In *Proceedings International Confonference on Human-Computer Interaction (HCII-2003)*. Lawrence Erlbaum Associates Inc. Publishers, New Jersey, U.S.A., 806–810.

[62] Gerasimos Xydas, Dimitris Spiliotopoulos, and Georgios Kouroupetroglou. 2004. Modeling Prosodic Structures in Linguistically Enriched Environments. In *Proceedings of the 7th International Conference on Text, Speech and Dialogue*. Springer Berlin Heidelberg, Berlin, Heidelberg, 521–528.

[63] Lu Yang and Anilkumar Kothalil Gopalakrishnan. 2014. A collaborative filtering recommendation based on user profile and user behavior in online social networks. In *2014 International Computer Science and Engineering Conference (ICSEC)*. IEEE, 273–277. https://doi.org/10.1109/ICSEC.2014.6978207