Effects of the Placement of Diverse Items in Recommendation Lists

Mouzhi Ge[1], Dietmar Jannach[2], Fatih Gedikli[2] and Martin Hepp[1]
[1]Universität der Bundeswehr Munich, Werner-Heisenberg-Weg 39, Neubiberg, Germany
[2]Technische Universität Dortmund, Joseph-von-Fraunhofer 23, Dortmund, Germany

Keywords: Recommender System, Evaluation, Diversity, Item Ranking, User Satisfaction.

Abstract: Over the last fifteen years, a large amount of research in recommender systems was devoted to the development of algorithms that focus on improving the accuracy of recommendations. More recently, it has been proposed that accuracy is not the only factor that contributes to the quality of recommender systems. Among others, the diversity of recommendation lists has been considered as one of the additionally relevant factors. Therefore a number of algorithms were proposed to generate recommendations lists containing a diverse set of items. However, limited research has been done regarding how to position those diverse items in the list. In this paper we therefore investigate how to organize the diverse items to achieve a higher perceived quality. The results of an experimental study show that the perceived diversity of a recommendation list depends on the placement of the diverse items. Placing the diverse items dispersedly or together at the bottom of the list can increase the perceived diversity. In addition, we found that in the movie domain including diverse items in the recommendation list does not hurt user satisfaction, which means that recommender system providers have some flexibility to add some extra items to the lists, for example to increase the serendipity of the recommendations.

1 INTRODUCTION

Recommender systems are developed to help users find relevant products that may interest them. The goal of recommender systems is to support the human user with information processing task and to provide personalized recommendations for users. Over the last decade, recommender systems have been widely applied in e-commerce, for example, book recommendation on Amazon or movie recommendation on Netflix (Jannach et al. 2010). Moreover, some case studies have stated that the use of recommender systems can both increase user satisfaction and produce added value to the business (Dias et al., 2008); (Jannach and Hegelich, 2009); (Zanker et al., 2006).

As there is a growing popularity of using recommender systems in e-commerce, a variety of recommender algorithms have been proposed over the last fifteen years. Most of these algorithms focus on improving recommendation accuracy. Accordingly, the performance of recommender systems was evaluated by accuracy metrics such as Mean Absolute Error (MAE) or Precision and Recall. However, some researchers have proposed that being accurate alone is not enough (McNee et al., 2006). Additional and complementary metrics, including diversity, novelty and serendipity could be used to evaluate the quality of recommender systems (Castells et al., 2011); (Herlocker et al., 2004). Among the proposed metrics, diversity has been widely discussed and considered to be a factor that is equally important as accuracy (Smyth and McClave 2001); (Fleder and Hosanagar, 2007).

The concept of diversity in recommender system research can be generally divided into *inherent* diversity and *perceived* diversity. Inherent diversity considers diversity from an objective view and is often measured by the dissimilarity among the recommended items (Zhang and Hurley, 2008); (Ziegler et al., 2005). The set of recommended items can either refer to a single list of recommendations for a single user or the set of overall recommendations from the whole system. Thus the concept of inherent diversity comprises intra-list diversity as defined by (Ziegler et al., 2005) and aggregate diversity as proposed by (Adomavicius and Kwon, 2011a). While intra-list diversity means the diversity inside a particular recommendation list, aggregate diversity refers to the diversity among the recommendations across all users.

Perceived diversity, in contrast, defines diversity

Ge M., Jannach D., Gedikli F. and Hepp M..
Effects of the Placement of Diverse Items in Recommendation Lists.
DOI: 10.5220/0003974802010208
In *Proceedings of the 14th International Conference on Enterprise Information Systems* (ICEIS-2012), pages 201-208
ISBN: 978-989-8565-11-2

from a subjective perspective and can only be determined through a user evaluation. The advantage of focusing on perceived diversity is that we can directly capture the users' opinions. Lathia et al. (2010) found that perceived diversity is positively related to user satisfaction in the long term when using a recommender system. Regarding the importance of perceived diversity, this paper will analyze how end users perceive the diversity-increasing items in recommendation lists. Our experimental study will use movie recommendations as an example. The diversity will be varied by adding movies from different genres.

One factor that may affect the perceived diversity but has not been analyzed in research so far, is the placement of diversity-increasing items in the recommendation list. Suppose we have several diverse items that we can include in a recommendation list. We can place these items *dispersedly* within the list, for example, by randomly positioning the diverse items at different places in the list. On the other hand, the diverse items can be placed together in one *block* in the list. A block means that one section of the recommendation list contains only diverse items. Users may perceive the recommendation list with a block of diverse items to be more diversified than a list with dispersedly placed diverse items since it can be easier for users to discover the block of diverse items. Furthermore, the position of the diverse items may also affect a recommender system's overall perceived quality. For example, if the diverse items are placed together on the top of the list, users may get the impression that the recommender system's predicting ability is poor and therefore they may lose the trust in the system and stop using it in the future (Lathia et al., 2010).

To the best of our knowledge, how to place diverse items in a recommendation list has not been explored so far in recommender system research. Considering the possible effects of differently positioning the diverse items, we believe that the question of how to arrange the diverse items is an important research topic in recommender systems.

In order to tackle this problem, the aim of this paper is to investigate how to place the diverse items in a recommendation list and analyze the effects of different item placements on the perceived diversity, on serendipity, and on user satisfaction. As a final goal, we want to develop a set of guidelines of how to arrange diverse items so as to improve recommender system's overall perceived quality.

The remainder of this paper is organized as follows. In Section 2 we propose a set of hypotheses about the placement of diverse recommendations and their potential effects. In order to validate the hypotheses, in Section 3, we design an experiment to study the effects of the different placements of the diverse items. Next, we carry out a data analysis and summarize our results in Section 4. We conclude this paper by discussing our findings and providing indications how to better arrange the items in a recommendation list.

2 HYPOTHESIS DEVELOPMENT

Sakai (2011) pointed out that balancing relevance and diversity has been considered as a challenge in document retrieval (Clarke et al., 2011). This trade-off has been also noticed in the recommender system community. Adomavicius and Kwon (2011a) stated that increasing diversity in a recommender system can result in decreasing its accuracy and vice versa. Thus a number of recommender algorithms focus on combining diversity and accuracy (Smyth and McClave, 2001); (Ziegler et al., 2005) or increasing diversity with a minimal loss of accuracy (Adomavicius and Kwon, 2011a); (Zhou et al., 2010); (Zhang and Hurley, 2008).

The concept of diversity used in the papers above refers to inherent diversity, which is often measured by the dissimilarity between all pairs of recommended items. Therefore, inherent diversity does not depend on the order of the items and changing the order of diverse items in a recommendation list will not affect inherent diversity. Ziegler et al. (2005) therefore argued that rearranging the positions of the items in a recommendation list would not affect inherent diversity. However, as we discussed in the introduction, it may affect the perceived diversity. Specifically, it might be easier for users to discover diverse items when they are arranged in a block than dispersedly placed. We therefore propose the following hypothesis.

H1: A Recommendation List Containing a Block of Diverse Items is perceived to be more Diverse than one with Dispersedly Placed Diverse Items.

Changing the order of diverse items may also affect the serendipity of a recommendation list. Serendipity is considered to be an important factor to attract users to use recommender systems (Ge et al., 2010). McNee et al. (2006) propose to define it as the experience by the user who received an unexpected and fortuitous recommendation. Thus serendipity can be measured by the extent to which the recommendations are both attractive and surprising to the user (Herlocker et al., 2004). Moreover, Ge et al. (2010) found two essential aspects of serendipity: unexpectedness and usefulness. While unexpected recommendations refer to those recommendations that are significantly distant from the user's expectations, usefulness means the highest level of utility to the user. Diverse items are considered to play an important role in generating unexpected recommendations (Adamopoulos and Tuzhilin, 2011).

Intuitively, we assume that users are to some extent surprised when they see diverse recommendations. For example, users may be surprised when seeing a romantic movie within a list of action movie recommendations. Thus, if several diverse items are dispersedly placed in the recommendation list, users can regularly find unexpected items and may experience more "surprise times" than in the case that the diverse items are placed together in a block. We therefore establish hypothesis H2 as follows.

H2: A Recommendation List with Dispersedly Placed Diverse Items is perceived to be more Unexpected than the One Containing a Block of Diverse Items. Our review above indicates that previous research has realized the potential value of diversity and serendipity in recommender systems. Adomavicius and Kwon (2011b) argue that increasing diversity can lead to an increase in sales diversity and user satisfaction. Also, as Ge et al. (2010) discussed, surprising and serendipitous recommendations can increase the user's interest in using a recommender system, and in turn lead to higher user satisfaction. Therefore maintaining a certain level of diversity and serendipity in a recommendation list can improve user satisfaction. According to the discussion when developing hypothesis H1 and H2, diverse items that are arranged in a block presumably will result in higher diversity, whereas diverse items that are dispersedly arranged will presumably increase serendipity. Increasing either diversity or serendipity can lead to a higher level of user satisfaction. We therefore propose a *null hypothesis,* H3, as follows.

H3: A Recommendation List Containing a Block of Diverse Items can Result in the Same User Satisfaction with a Recommendation List with Dispersedly Placed Diverse Items. Overall, our three hypotheses are proposed based on a literature review and our intuitive conjectures. In order to test the proposed hypothesis, we designed an experiment to empirically analyze the effects of different placements of diverse items, which we describe in the next section. Furthermore, as we are also interested in studying whether the presence of diverse items is beneficial for recommender systems in general, we will include a treatment without diverse items in the experiment.

3 EXPERIMENTAL DESIGN

In this section, we will present the experimental design and measurement technique used in our study. In addition to studying how to arrange diverse items in a recommendation list, we also study whether and to which extent diverse items influence the user-perceived quality of a recommender system. In our experiment, we employ a *within subjects* design, in which each subject can evaluate and compare all the treatments used in this user study.

Our experiment is implemented as an online Web site. There are three phases in the experiment. The first phase is to instruct the participants about the different phases of the experiment and how they can complete the experiment. The second phase is that subjects interact with a recommender system, rate items and are presented with movie recommendations. In the recommendation phase, we used four movie genres: action movies, romantic movies, comedy movies and animation movies. For each movie genre, we have developed two Web pages. In the first Web page, subjects are provided with a list of 20 well-known movies of one specific genre. Figure 1 shows an example snapshot in which a list of 20 action movies is presented to the subjects. The subjects will be asked to check the movies they have watched and also liked. After the subjects finished ticking their preferred movies, they can click on the "Get Recommendations" button to obtain recommendations. Then, on the second Web page, shown as Figure 2, a list of twelve recommended movies is presented to the subject. Three options are offered: "I would like to watch this movie", "I have watched this movie and liked it" and "I have watched this movie but I do not like it". Subjects can tick one of the options to report their opinions towards the recommendations. It is however not mandatory for subjects to tick an option for each recommendation. In order to support the subjects in the decision process, the plot of each recommended movie is also given by the system (refer to Figure 2). The movie plot and three options were used to let users carefully consider the recommendations.

It is important to know that in our experiment we do not use any recommender algorithm to compute

Figure 1: Screen 1 - Acquiring user preferences for action movies.

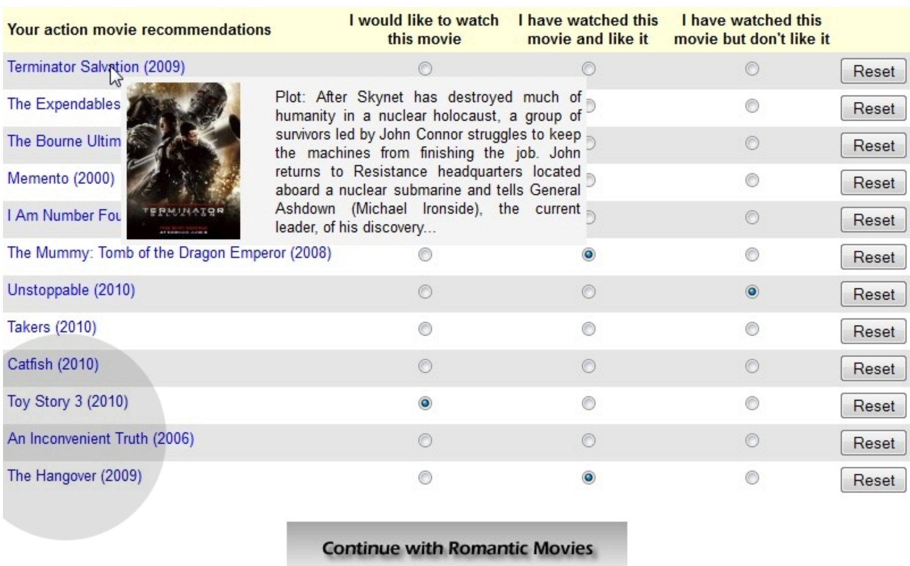

Figure 2: Screen 2 - Displaying recommended action movie recommendations to users.

the recommendations. Instead, we manually create a static list of recommended movies for each genre and present it to users. Therefore each subject will obtain exactly the same set of recommendations. We can thus eliminate possible effects from recommender algorithms. In order to give the user an impression that there is a recommender system running in the background, we not only ask the users about their preferences but also show a message that the recommendations are being computed for two seconds after the subject clicks the "Get Recommendations" button.

In our experiment, we determine diverse movie recommendations based on differences with respect to the movie genre. For example, among the recommended action movies, an animation movie, Toy Story 3, is considered as a diverse item. In the experiment, each recommendation list contains twelve items. Four of them are diverse items. For example, in Figure 2, there is a list of twelve recommendations. The four diverse recommendations are placed at the bottom of the list. We use a round grey shadow to highlight the four diverse items in Figure 2. Note that this shadow was of course not visible during the study.

We designed the different placements of diverse items as follows. In the list of action movie recommendations, the four diverse items are organized together in one block at the end of the list. For romantic movie recommendations, the four diverse items are arranged in the middle block of the list. Among the comedy movie recommendations, the four diverse items are respectively placed at positions 3, 6, 9, and 12. We suppose that diverse items are dispersedly placed in this list. In addition, we use the recommended animation movies as our control group, which contains no diverse items.

After the subjects have gone through every recommendation list, in the last phase they are again presented all the four manually designed recommendation lists. Subjects are then asked to evaluate each list on a five point Likert scale. The evaluation is based on the following questions, which are designed to test our proposed hypotheses.

- Do you think this recommendation list is diversified?

  (1: not at all, 5: very diversified)

- Does this recommendation list surprise you?

  (1: not at all, 5: very surprised)

- Are you satisfied with the movie recommendations?

  (1: not satisfied, 5: very satisfied)

In the end of the evaluation, the system also displays a textbox where the subjects can leave a feedback regarding the recommendations. After finishing the evaluation, the subject needs to click the "Submit" button to complete the experiment. The whole experiment procedure is supervised in case the subjects need an explanation about system functions or the meanings of some terms. During the experiment there is no interaction between the subjects.

4 DATA ANALYSIS

A total of 52 subjects were involved in the experiment. All the subjects were researchers or students from the computer science department at the Technical University of Dortmund. 35% of the subjects were female and 65% were male. The average age of subjects was 29. For each subject, it took on average about 15 minutes to finish the whole experiment.

As our experiment used a Likert scale, the data collected from the experiment were ordinal data. We therefore choose a non-parametric test to analyze our collected data. Since the same subjects have participated in all the experimental treatments, the Friedman Test is used to test, whether or not there is any difference among the experimental treatments. Once a significant difference is found, the Wilcoxon Signed-Rank Test would be performed to find where the differences actually occur. SPSS 19.0 was used for data analysis and all the tests were done at a 95% confidence level. We report the analysis results in the following.

As a first step, we performed a Friedman test on perceived diversity. In the test, there are four buckets of data, which are named "Dispersedly", "Bottom", "Middle" and "Without". "Dispersedly", "Bottom" and "Middle" denote recommendation lists where the diverse items are placed dispersedly, at the bottom, or in the middle respectively. "Without" stands for our control group that contains no diverse items. This naming scheme is also applied in all the following tests. The results of the Friedman test are shown in Table 1.

Table 1: Friedman test for perceived diversity.

Mean Ranks

Bottom	3.13
Dispersedly	2.64
Without	2.56
Middle	1.68

Test Statistics[a]

N	52
Chi-Square	30.890
df	3
Asymp. Sig.	.000

In Table 1 we can see that there was a significant difference in perceived diversity depending on the placement of diverse items ($\chi^2(3) = 30.890$, $p < 0.05$). This means that different placements of the diverse items significantly affected the perceived diversity of the recommendation list. Thus we arranged the mean ranks in descending order and further performed the Wilcoxon Signed-Rank test to find which group caused the significant difference. The result of the Wilcoxon test for perceived diversity is shown in Table 2.

Table 2: Wilcoxon Signed-Rank test for perceived diversity.

	Dispersedly & Bottom	Middle & Bottom	Without & Bottom
Z Asymp. Sig.	-1.950[a] .051	-4.295[a] .000	-2.856[a] .004
	Middle & Dispersedly	**Without & Dispersedly**	**Without & Middle**
Z Asymp. Sig.	-3.980[a] .000	-.557[a] .577	-3.541[b] .000

[a] *Based on negative ranks*
[b] *Based on positive ranks*

In order to interpret our Wilcoxon test result, a Bonferroni correction was accordingly applied and thus all the effects are reported at a $p < 0.008$ level of significance.

The result show that it appears that placing the diverse items dispersedly in the recommendation lists is perceived to be more diverse than in the case where the diverse items are placed in the middle ($Z = -3.980$, $p < 0.008$). H1 is therefore rejected and placing the diverse items, for example, in the middle of the recommendation list, does not lead to a higher level of perceived diversity. However, there was no significant difference between placing diverse items dispersedly and at the bottom ($Z = -1.950$, $p = 0.051$). We therefore found that a recommendation list with dispersedly placed diverse items can achieve equal or higher perceived diversity than the one containing a block of diverse items.

Regarding the issue of whether or not including diverse items will increase the perceived diversity, our analysis showed that including diverse items in a recommendation list can both increase and sometimes even decrease the perceived diversity. It depends on how to arrange the diverse items. If the diverse items are placed together in the bottom of a list, the perceived diversity is significantly higher than the list without diverse items ($Z = -2.856$, $p = 0.004$). However, when we place the diverse items in the middle of the recommendation list, the list's perceived diversity is even significantly lower than the one without diverse items ($Z = -3.541$, $p < 0.008$). One possible explanation is that users may stop reading the recommendation list when they meet diverse items in the middle. However they may have inspected the whole list without any diverse items and thus found it to be more diverse than the one with diverse items placed in the middle.

In order to examine H2, we performed a Friedman test on the perceived surprise level. The result of the analysis is shown in Table 3.

Table 3: Friedman test for the surprise level of the recommendation list.

Mean Ranks

Dispersedly	2.83
Bottom	2.58
Without	2.53
Middle	2.06

Test Statistics[a]

N	52
Chi-Square	8.817
df	3
Asymp. Sig.	.032

Table 3 shows that there was a significant difference among the four experimental treatments ($\chi^2(3) = 8.817$, $p = 0.032$), indicating that different placements of diverse items perform differently in surprising the users. Therefore we further used the Wilcoxon Signed-Rank test to find the details regarding this significant difference. The result of this Wilcoxon test is shown in Table 4.

Similar to the analysis for perceived diversity, the Wilcoxon Test was conducted with a Bonferroni correction, resulting in a significance level at $p < 0.008$. The analysis shows that placing the diverse items in a recommendation list dispersedly can lead to a higher surprise level than the in the case where the diverse items are placed in the middle of the list ($Z = -2.755$, $p = 0.006$). There was no significant difference in surprising users when the diverse items are placed dispersedly or at the bottom ($Z = -0.426$, $p = 0.670$). Therefore H2 is partially supported. Interestingly, we found that including diverse items does not significantly increase the surprise level of the recommendation list. This indicates that including diverse items in a recommendation list to the extent we did in our experiment will not increase the surprise level independent of the placement of these items.

Table 4: Wilcoxon Signed-Rank test for user satisfaction.

	Dispersedly & Bottom	Middle & Bottom	Without & Bottom
Z Asymp. Sig.	-.426[a] .670	-2.240[b] .025	-.906[b] .365
	Middle & Dispersedly	**Without& Dispersedly**	**Without & Middle**
Z Asymp. Sig.	-2.755[b] .006	-1.271[b] .204	-2.462[a] .014

[a] *Based on negative ranks*
[b] *Based on positive ranks*

Finally, we carried out a Friedman test on user satisfaction. The analysis result can be found in Table 5.

Table 5: Friedman test for user satisfaction.

Mean Ranks

Without	2.68
Middle	2.61
Dispersedly	2.50
Bottom	2.21

Test Statistics[a]

N	52
Chi-Square	3.359
df	3
Asymp. Sig.	. 340

Surprisingly, we found no significant differences among the four experimental treatments ($\chi^2(3) = 3.359$, $p = 0.340$). This indicates that placing the diverse items in a recommendation list dispersedly, at the bottom, in the middle or without diverse items results in the same level of user satisfaction. We therefore fail to reject the null hypothesis H3. That means we found there exists the possibility that all of our experimental treatments result in the same level of user satisfaction. Because there is no significant difference found in the Friedman test, there is no need to carry out the Wilcoxon Signed-Rank test for user satisfaction. As a practical consequence, we are able to add a certain number of diverse items in the recommendation list without hurting user satisfaction. This implies that in practice we can add some extra items to promote certain products or increase sales diversity.

5 DISCUSSION AND CONCLUDING REMARK

A number of algorithms have been proposed to increase diversity or generate diverse items in the recommendation list (Zhang and Hurley, 2008); (Ziegler et al., 2005). However, the issue of how to place the diverse items is still not in the focus of recommender system research. We propose in this work that the question of how to place diverse items is an important issue because differently placing the diverse items can affect perceived diversity and the level of serendipity. Based on our findings, if the goal of recommender systems is to increase the perceived diversity, we suggest positioning the diverse items dispersedly or together in the bottom of the list. It is also important to note that placing the diverse items in the middle of the recommendation list may even reduce the perceived diversity. Furthermore, as we can use the placement of the diverse items to control the perceived diversity, our result might be used to manipulate perceived diversity in future experiment such as in factorial design.

Additionally, we found that in the movie domain including a certain amount of diverse items in the recommendation list does not surprise the users too much. When investigating the role of serendipity in recommender systems, we therefore suggest that further studies should focus on the cross-domain product recommendations. Also, the possibility of improving serendipity might be increased when recommending products from different domains.

A number of studies are based on the assumption that increasing diversity will lead to higher user satisfaction. We therefore tried to analyze whether increasing diversity results in higher user satisfaction. However, we found that there was no significant difference between the groups that received diverse recommendations and the group whose list was more monotonous. One possible explanation is that in the movie domain users usually have a strong or relatively fixed movie preference. Therefore the diverse movies might have been of limited interest to the users. In other domains such as tourism, users might however be interested to see quite different travel destinations. Thus we argue that this can be seen as a domain specific problem and our conclusions are limited to the movie domain.

While we see our work as a further step toward a better understanding of the role of diversity and serendipity of recommendation lists, we are aware of some limitations of our work. First, there might be an effect related to the different movie genres in the experiment. Different movie genres might for example influence the user's evaluation of the system. In order to minimize the effect of different genres, we clearly instructed the subjects that in the experiment the four movie genres are four different scenarios. In a future study, we will further improve the design of the experiment and focus on a single movie genre so as to eliminate the effects of genres.

Second, user preference is an external factor that may influence the experiment. User satisfaction might not only depend on the position of diverse item, but also on their personal preference. We tried to avoid this influence by using only very popular and well-known movies in the experiment. Note that users have selected the movies they have watched and also liked in the experiment. Considering this data, we have excluded the subjects with strong movie preferences. In the future, we will further conduct an experiment with the subjects who have similar movie preferences.

In addition, our future work will further investigate user's personal valuation of diversity in the results, for example, the subject's degree of knowledge of a particular topic, the certainty in what he or she is looking for and the objective fitness criteria of objects for the searcher's purpose.

ACKNOWLEDGEMENTS

Parts of the work presented in this paper have been supported by the German Federal Ministry of Research (BMBF) by a grant under the KMU Innovativ program as part of the Intelligent Match project (FKZ 01IS10022B).

REFERENCES

Adomavicius, G., Kwon, Y., 2011a, Improving aggregate recommendation diversity using ranking-based techniques. *IEEE Transactions on Knowledge and Data Engineering*, 99, 1-15.

Adomavicius, G., Kwon, Y., 2011b. Maximizing aggregate recommendation diversity: a graph-theoretic approach, In *Proceedings of Workshop on Novelty and Diversity in Recommender Systems*, Chicago, Illinois, USA, 3-10.

Adamopoulos, P., and Tuzhilin, A., 2011. On unexpectedness in recommender systems: or how to expect the unexpected, In *Proceedings of Workshop on Novelty and Diversity in Recommender Systems*, Chicago, Illinois, USA.

Castells, P., Vargas, S., Wang, J., 2011. Novelty and diversity metrics for recommender systems: choice, discovery and relevance. In *Proceedings of International Workshop on Diversity in Document Retrieval*, Dublin, Ireland, 29-37.

Clarke, C. L. A., Craswell, N., Soboroff, I. and Ashkan, A., 2011. A comparative analysis of cascade measures for novelty and diversity, In Proceedings of the fourth ACM international conference on web search and data mining, Hong Kong, China, 75-84.

Dias, M. B., Locher, D., Li, M., El-Deredy,W. and Lisboa, P. J., 2008. The value of personalised recommender systems to e-business: a case study. In Proceedings of the 2008 ACM Conference on Recommender Systems, Lausanne, Switzerland, 291–294.

Fleder, D., Hosanagar, K., 2007, Recommender systems and their impact on sales diversity. In *Proceedings of the 8th ACM Conference on Electronic Commerce,* San Diego, CA, USA, 192-199.

Ge, M., Delgado-Battenfeld, C., and Jannach, D., 2010. Beyond accuracy: evaluating recommender systems by coverage and serendipity. In *Proceedings of the fourth ACM Conference on Recommender Systems,* New York, 257-260.

Herlocker, L., Konstan, J., Terveen, L., Riedl, J., 2004. Evaluating collaborative filtering recommender systems, *ACM Transactions on Information Systems* 22,1: 5-53

Jannach, D., Hegelich K., 2009. A case study on the effectiveness of recommendations in the mobile Internet, *ACM Conference on Recommender Systems*, New York, 205-208.

Jannach, D., Zanker, M., Felfernig, A., Friedrich G., 2010. Recommender systems: an Introduction, Cambridge University Press.

Lathia, N., Hailes, S., Capra, L., Amatriain, X., 2010. Temporal diversity in recommender systems. In *Proceedings of the 33rd International ACM SIGIR Conference on Research and Development in Information Retrieval*, Geneva, Switzerland, 210-217.

McNee, S, Riedl, J., Konstan, J., 2006. Being accurate is not enough: how accuracy metrics have hurt recommender systems, In *Proceedings of the ACM SIGCHI Conference on Human Factors in Computing Systems*. Montréal, Canada, 1097-1101.

Smyth, B. and McClave, P., 2001. Similarity vs. Diversity. In *Proceedings of 4th International Conference on Case-Based Reasoning*, Vancouver, Canada, 348-361.

Sakai, T., 2011. Challenges in diversity evaluation, *In Proceedings of International Workshop on Diversity in Document Retrieval*. Dublin, Ireland, 1-7.

Zanker, M., Bricman, M., Gordea, S., Jannach, D., Jessenitschnig, M., 2006. Persuasive online selling in quality & taste domains, Proceedings EC-Web'06, Krakow, Poland, Springer LNCS 4082.

Zhou, T., Kuscsika, Z., Liua, J., Medoa, M., Wakelinga, J., Zhang. Y., 2010. Solving the apparent diversity-accuracy dilemma of recommender systems. *National Academy of Sciences of the USA*. 107, 10, 4511-4515.

Zhang, M., Hurley, N., 2008. Avoiding monotony: improving the diversity of recommendation lists. In *Proceedings of the 2nd ACM conference on recommender Systems*, Lausanne, Switzerland, 123-130.

Ziegler, C., McNee, S., Konstan, J., Lausen, G., 2005. Improving Recommendation Lists through Topic Diversification. In *Proceedings of the 14th World Wide Web Conference*. Chiba, Japan, 22-32.