Computer Science and Information Systems 16(1):179–203 https://doi.org/10.2298/CSIS171217012S

Towards Understandable Personalized Recommendations: Hybrid Explanations

Martin Svrcek, Michal Kompan, and Maria Bielikova

Slovak University of Technology in Bratislava
Faculty of Informatics and Information Technologies
Ilkovicova 2, 841 04 Bratislava, Slovakia
{name.surname}@stuba.sk

Abstract. Nowadays, personalized recommendations are widely used and popular. There are a lot of systems in various fields, which use recommendations for different purposes. One of the basic problems is the distrust of users of recommended systems. Users often consider the recommendations as an intrusion of their privacy. Therefore, it is important to make recommendations transparent and understandable to users. To address these problems, we propose a novel hybrid method of personalized explanation of recommendations. Our method is independent of recommendation technique and combines basic explanation styles to provide the appropriate type of personalized explanation to each user. We conducted several online experiments in the news domain. Obtained results clearly show that the proposed personalized hybrid explanation approach improves the users' attitude towards the recommender, moreover, we have observed the increase of recommendation precision.

Keywords: recommendations explanation, eye-tracking, collaborative filtering, personalized recommendation.

1. Introduction

Nowadays, information is an important part of our lives. The information we have available influences our decisions, opinions and ideas. Therefore, it is important that the available information is suitable and appropriate to our interests. Currently, there are lots of sources from which we can derive information, such as newspapers, books, television, friends or relatives. The availability of information is much higher than it was in the past, which is mainly due to the growth of the Internet[1].

The Web is integrated into everyday life and provides a large amount of information, which in most cases is freely available. People use the Internet to obtain information regardless of time or their physical location. However, the vast amount of data in the form of text, images, recordings or video brings many advantages but also some disadvantages (e.g., availability of information, amount of information).

In this context, there is a very important concept of information availability. Different search engines on the web allow us to get information quickly. Unlike books, we get specific information without a long search in the content or the registry.

The variety of information provided is another important aspect. The Internet provides a place where we can meet the different opinions, thoughts or different views on certain

[1] https://www.internetlivestats.com/total-number-of-websites/

issues. Internet is also a place of different qualities and styles of information such as professional, scientific or artistic as well.

The amount of information as another aspect has two main points of view. A lot of information on the Internet is on one side the positive feature, but on the other side it often results in information overload of users. Information overload relates to the problems of making certain decisions or understanding of certain issues due to the vast amount of information that is offered to the person concerned [36].

Information overload, unsuitability or poor quality of information causes a form of distrust among users. Often, the website displays information, products or services that are not interesting from the perspective of particular user or there are so much of these information that the user cannot find the important pieces. Problems like these can be solved by use of personalization.

Personalized Web minimizes the fundamental problem of classical websites and applications. This problem is that they provide the same information to all users without distinction of their knowledge or passions [7].

Personalized Web is based on the concept of adaptive hypermedia. This concept is related to creating, maintaining and using a user model as a base for adapting to the needs and preferences of particular users [7]. Just as there are several types of information that can be used for personalization such as search results, content or design, there exist a number of ways or opportunities of the personalization realization. One of these opportunities are recommendations as a form of an adaptation of the content to the needs of a particular user.

However, recommender systems in general often produce (recommend to the user) some type of information, product, article etc. without any explanation why they think that this particular item is suitable for particular user. Users usually do not know how these systems deduced presented recommendation and more importantly how recommender systems deduced information about the users, and how they used this information. We can speak about some kind of distrust of users in recommendations [41]. The problem of the trust in the human-computer interaction is a widely researched topic nowadays [38], as the need for such interaction is increasing day-by-day (often including new groups of users, e.g., elder people). Similarly, social media receive tremendous research attention in revealing, simulating and usage of the trust in such a complex environment [31,21].

One of the solutions for this type of problem is an explanation of recommendations, which has the power to clear the process of finding the recommended items and make the whole system more understandable. In this paper, we address the following research questions:

- RQ1: Can we increase the understandability and precision of recommendations with use of explanations?
- RQ2: Is there one explanation style that is preferred by users?

In order to find an answer to these questions, we have proposed a novel method of explaining recommendations. Several experiments were conducted with the effort to evaluate our method. Our contributions that we present in this paper are:

- Hybrid method of explanation of recommendations that combines different explanation styles.

– Improvement of precision for lower-positioned items by using proposed explanation method.

The paper is organized as follows. The brief overview of basic recommender approaches is in section 2. Section 3 is about explanations, reasons why we use them and also about different styles and attributes of explanations. In section 4 we analyze different approaches to recommendations and explanations related to them. Important part of this section is also detailed analysis of explanation approaches used in real web sites. Section 5 presents our novel method to explaining recommendation items. We evaluated our approach by user study which evaluation process consisting of experiment, its setup and results is described in Section 5. Summary and future works is in section 7.

2. Recommender Systems

Main purpose of recommender system is to provide (recommend) objects that would be helpful and suitable for the user [28]. Nowadays, such systems are relatively common and affect different domains as entertainment, content, e-commerce, etc.

In the context of the problem of information overload (which is related to various activities [12]), it is necessary to somehow reduce this overload. Our goal as users is to facilitate and simplify the process of selecting a particular item (ask friends, search for reviews). Recommender systems try to automate these activities (asking friends, etc.) and offer a recommendation, which should also be appropriate and interesting for the user [18].

There are many different types of recommended systems. Many of these systems are personalized to user needs. Personalized recommenders use personal preferences to generate recommendations. Source of such features and preferences is mainly implicit feedback (e.g., time spent, purchase, click) and explicit feedback (e.g., like, review).

There is a lot of systems that employ some of these methods to achieve better results [20]. Generally, the explicit feedback is considered more accurate then implicit feedback [3]. Thus, when choosing a method it is necessary to take into account the characteristics of the system and the opportunities that this system offers us according to specific domain.

Recommendations can be applied as part of systems in many different areas like news articles, movie databases, video portal or e-commerce. They are useful in each of these areas differently. Today, we can point to the different roles that recommender systems play and the different reasons why various service providers use them [28]: increase of items sold, diversity of sale, customer satisfaction or customer loyalty.

The most important part of the recommendation is the prediction of items that may be interesting to the user. Currently, there are several approaches that use different principles and methods to generate assumptions about what can help the user.

Collaborative filtering is considered the most popular and most used solution [8]. The basic idea of collaborative filtering recommendation is to recommend items that are interesting for users with similar interests [28]. An important part is therefore to identify these groups of users based on their preferences. In most cases, collaborative recommenders use for this purpose different types of ratings of specific items. Ratings are mostly binary, or express a preference for the wider range [14].

Content-based recommendation is not based on the similarity of users as collaborative filtering but on the similarity of items. Therefore, there are items recommended to the user, that they liked in the past [19,4]. For example, this is the case when the user rated or purchased the item. The similarity of items is determined based on the attributes assigned to it.

Demographic recommendations are based on demographic information about users, e.g. information related to personal attributes of the user. Classic examples are age, language or country, which create demographic profiles of users [28]. One of the first uses of this approach was the Grundy system [29], where books were recommended based on personal interview. Source of information for recommendations may also be part of various marketing research and customer segmentation [9].

Knowledge-based recommendations recommend items based on information about user needs and preferences [8]. This principle is, however, principle of many techniques but recommendation systems based on knowledge also have information on how the properties of items meet the users' needs [9]. In e-commerce, there are areas where users purchase certain items only once in a few years (electronics, bicycles, etc.). In these cases, it is appropriate to use knowledge-based recommendations.

Hybrid recommendations combine several techniques to achieve optimal recommendations to exploit their advantage and to eliminate the disadvantages [28,9,1]. They often combine collaborative recommendations with recommendations based on the content. Thus, we can remove problems with a small number of ratings in collaborative filtering.

The role of the recommendation technique is to generate suitable objects for a user. Therefore, the suitability of the recommendation depends on the quality of the recommendation technique. Unfortunately, there are plenty of different recommendation techniques (often based on latent features computation), which make white-box explanations impossible. In some domains, e.g. movies, users are not able to perceive the suitability of recommended item from some of its characteristics, e.g., name. Thus, our intent is to show the recommendations in the best way, provide reasons for the recommendations and in this way, convince the user to use the presented recommendations. Specific form of presentation and visualization of items of recommendations is the amount of information, structure of information, color, position or explanations of recommendations.

3. Explanation of Recommendations

Explanations have multiple definitions mainly because of the wide range of their application. We adopted following definition: recommendation explanation is an information that is designed to clarify why the item was recommended [2]. Explanations are also defined as a description that helps to determine the suitability of the item for a specific user [2]. However, generally speaking, they provide information about the recommendations and support the objectives defined by the creators of the system [32].

A very important point in the context of explaining the recommendations is to realize what the purpose of the explanation is. Recommender systems with explanations help make quick decisions and verify the suitability of buying a product [2]. But neither explanation can fix the problems with bad recommendations even though they can somehow compensate it [2].

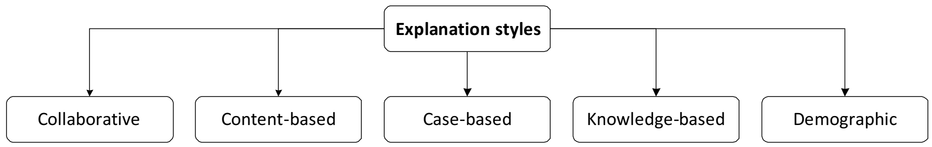

Fig. 1. Explanation styles categorization by [2].

Recommendations without explanations often act as a black box. This means that users have no clue how these systems get special knowledge about them and how they can know what is appropriate for them [26]. This is why the explanation itself is closely related to the recommendation technique. This means that the recommendation technique (eg, collaborative filtering, content-based) affects the explanation style of a specific recommended item. Therefore, the standard approach is when we use collaborative recommendations and the explanation will be based on the same technique.

Following the idea of the transparency, the explanations itself try to explain how the recommendation approach works. Unfortunately, the complexity and novel approaches to recommendation (e.g., based on some latent features [13]) cannot be explained to the average user. In such case, explanations aims as explaining the high level idea rather then the specific algorithm. Following the standard terminology, we will refer to it as the black-box explanation [37]. Pushing it forward, sometimes we have to abstract from the recommendation approach and generate explanations separately (e.g., we cannot explain the idea of neural network based recommender). On the contrary, if the explanations try to reveal the principle of the recommender, we will refer to it as the white-box explanations.

3.1. Explanations Styles

Since individual explanations are based on the recommendation technique which was used, they can also be categorized in this way (Figure 1). This approach relates to the fact that each technique can affect the style or form of explanation.

To make it clear, we provide short characteristic of different approaches to explanation styles referring to each style [32,2]:

Collaborative – Input for this technique are the ratings of items given by individual users. Such an approach is used e.g. in Amazon. An example of this style are explanations such as: *"Users like you positively rate this item."*

Content-based – In this case, an explanation tries to make clear to the user that the item was recommended based on the similarity with other items. An example is the explanation: *"This movie was recommended to you because it contains features of movies that you positively rated in the past."*

Another type of content-based explanation is case-based explanation, which clarifies that the item is similar to another item that has already been used as a recommendation and therefore the user liked it in the past. An example is the explanation: *"Recommendation was generated based on your most viewed items."*

Knowledge-based – This style of explanation is based on the description of user's needs or interests in the context of a recommendation. A classic example is the explanation: *"This destination has higher average temperature, which is better for sunbathing."*

Demographic – It clarifies the use of demographic data and its connection to the recommendations. One example is the explanation: *"This movie was recommended to you according to your age."*

3.2. Attributes of Explanations

The explanations can address certain goals, i.e., benefits we want to bring to the system [32]. The individual attributes cannot be achieved all at once. It is often a compromise. When we increase one attribute, the second one will be reduced. Important "beyond accuracy" attributes of recommender systems are:

Transparency or Justification – This is an explanation of how the system works, or explain how the recommendation item was generated [16].

Scrutability – This is particularly the possibility for users to tell the recommender system that something is wrong in the system or something is not working as it should.

Trust or Confidence – As the name implies this is about increasing user confidence in the system. Like the ability to respond to errors in the system (scrutability), confidence is closely linked with the transparency of the system.

Persuasiveness – This one is related to the effort to persuade the user to test or bought an item. However, it is necessary that the system did not push the user or force him to buy or choose any possibly unsuitable items.

Effectiveness or Education – This means helping the user to make good decisions. Explanation should help the user to evaluate the suitability of an item in terms of his preferences [16].

Efficiency – This is the combined effort to make a selection of items faster. In this case, explanations try to simplify the selection of items in order to reduce the time of decision making.

Satisfaction – User satisfaction with the recommendations may be necessary and useful also in the context of other attributes. Generally, this is an attempt to do the work with the system more enjoyable for user [2].

Explanations may vary in recommendation technique. However, they may also differ in the attributes or benefits that we can achieve by their use as part of a recommender system. Purpose of explanation then must make the recommendations appropriate to user goals or characteristics, because the particular user may prefer different presentation and explanation style.

4. Real-world Explanations Applications

There are a number of different approaches and studies that deal with recommendations and which bring new methods in the area of personalization models or explanations [16,24]. But especially explanations are still quite new areas of actual research. We describe several recent approaches to explanations that are included in real systems and applications.

Most of the techniques focus only on recommendations and their personalized models. Often they start using different technologies in order to make the personalization better. To evaluate impact of explanations, the eye-tracking seems to be a promising technology which can show us if the user perceive them the way we expected. In [40] authors used gaze and incorporates it into the personalization models. Their results showed that eye-tracking data even from small number of users can significantly boost the accuracy of the recommendations [40].

One of the shortcoming of such studies is that they focus only on one part of the problem and that is accuracy of the recommendations. There are other important aspects (metrics), which describe the performance of the recommender system. In this context, the very important aspect is trust of users [41], their loyalty and increase of their satisfaction. Explanations can play an important role in this problem and make it easier for users to reach objects of their interest or even persuade them [17].

Some approaches also benefit from the way of creating the recommendations. System RecExp [17] is able to generate semantic recommendations by utilizing meta paths and thus to find various similar users. This type of recommender system makes possible to generate personalized explanations adapted to a specific user. In this case, explanations are presented as a fan chart, which shows weight of meta path representing user preferences [17]. Explanations are in this case presented in the form of three most similar users according to actual user.

Explanations itself are important to users for decision making. However, in case of recommendations, it is also important to present explanations in the right way so they attract user attention. One way how to present the explanations is by use of a natural language [11,25]. Powerful way is to use crowd-sourcing as an approach to generate natural language explanations [11]. However, in order to have good explanations, it is important to use domain experts as a crowd to generate these explanations. Therefore, this approach [11] builds on existing algorithms that generate personalized content-based explanations and combines them with texts from reviews. Thus, the authors were able to generate natural language explanations containing richer information than standard content-based explanations.

Framework ExpLOD [25] also generates natural language explanations, based on the Linked Open Data cloud as a source of information. The method is based on building a graph, which connect items liked by user and recommended items according to the Linked Open Data (LOD). ExpLOD framework thus outperforms baseline approaches in accuracy and also offers more interesting explanations.

There are also works, which specifically focus on explanations and comparing different explanation types [15]. This is also part of our research, however in different environment and domain. Another studies are trying to focus on trust of users towards the recommendation [39].

Nowadays, recommendations are used by large number of websites and applications. For related work here we present several different approaches. Specifically, we focus on popular websites such as Amazon, IMDB and Last.fm. Among these, we have analyzed the reasons for using recommendations and especially the presentation of recommendations.

4.1. Amazon

Amazon is globally known and popular e-commerce. It offers a huge variety of items and goods and therefore needs to direct its offer to the user. Amazon for this purpose use recommendations and their wide range of application. The main technique for generating

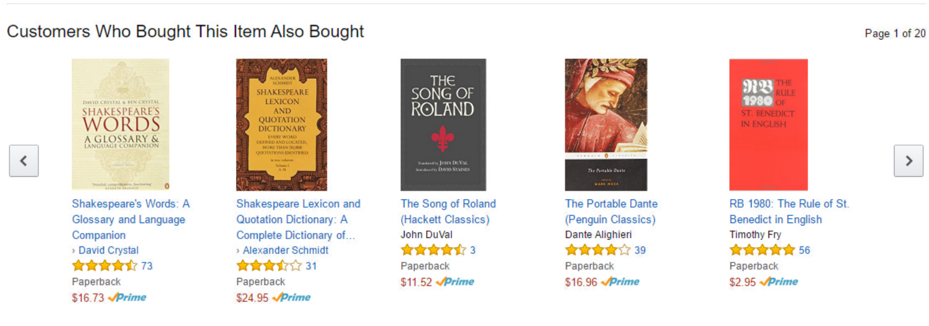

Fig. 2. Amazon: Explanation based on collaborative filtering.

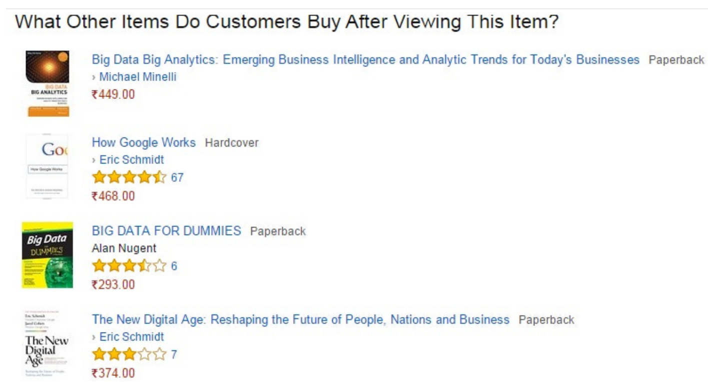

Fig. 3. Amazon: Explanation based on item-to-item recommendation.

recommendations is a content-based technique based on the similarity of items (item-to-item) [23].

The basic form of presentation of recommendations is a list of items along with an explanation of why they were recommended. Amazon uses especially the explanations based on the collaborative filtering (Figure 2) and on the recommendations based on the similarity of items (Figure 3).

4.2. Internet Movie Database - IMDB

IMDB is one of the largest film database and contains many items like Amazon. However, unlike the Amazon, IMDB is not try to sell these items. It is just a movie database that uses recommendations, specifically the item-to-item recommendation. Another difference is that the recommendations are not so promoted as in the case of Amazon.

The basic form of presentation of recommendations is a list of items along with an explanation of their recommendations (Figure 4). When a user hovers the cursor over one of the smaller images he will be able to see a fuller description on the right.

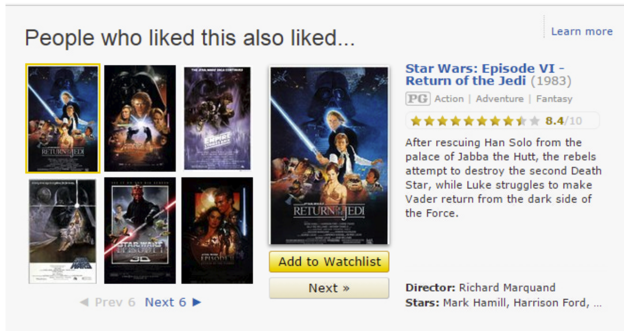

Fig. 4. IMDB: Explanation of recommendations.

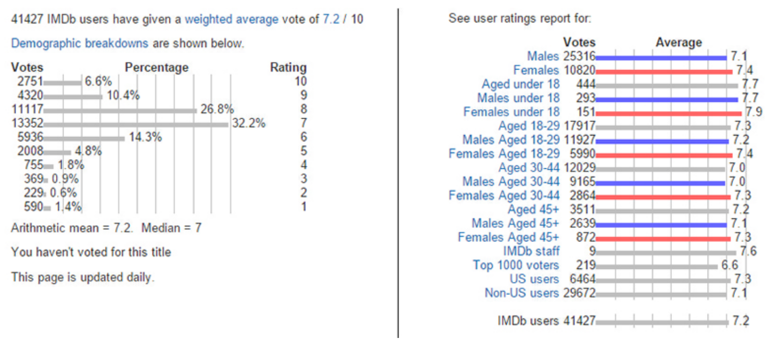

Fig. 5. IMDB: information about ratings.

Any rating of individual items (10 stars scale) is also available with more detailed information. These include a graph with the number of users who have rated the film and graph with demographic information about ratings (Figure 5).

IMDB does not use recommendations as much as Amazon but it still has some interesting solutions such as the list of recommended films (Figure 4). Explanations are of two types:

– Explanations based on collaborative filtering (Figure 4)
– Explanations of ratings (Figure 5)

The most interesting form are demographic explanations of ratings that should be very helpful for the users.

4.3. Last.fm

Last.fm is actually a music database that focuses on personalized recommendation based on the music the user is listening to. The most basic feature of this site is a social approach to the recommendations [10].

The main type of presentation is a list of similar artists. The basic form of this list is in the form of pictures with artist name. However, after clicking on the artist a user will

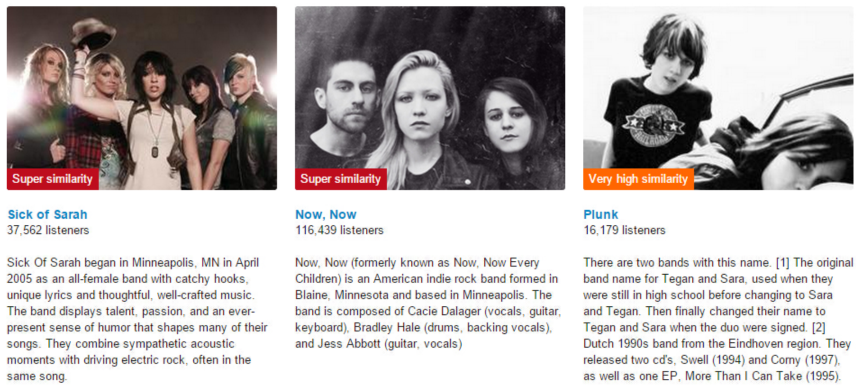

Fig. 6. Last.fm: List of similar artists.

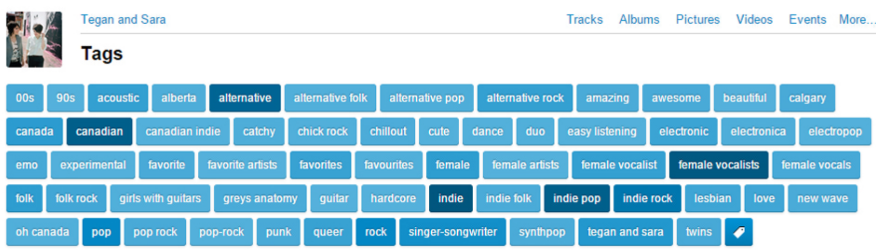

Fig. 7. Last.fm: Tags connected with artist.

be able to see a more specific list with a description and also with the level of similarity (Figure 6).

The strong feature of Last.fm is a labeling function (tag), which allows to add the artist or song to a certain group (Figure 7). These tags then form a new source of explanations for machine-generated and social-generated recommendations [10].

The advantage of Last.fm is fast and simple but very interesting recommender in the form of similar artists. Explanations are shown mainly by the level of the similarity (Figure 6). Tags are in this case also a form of explanations, which illustrates the similarity of the two artists.

4.4. Summary

Explanation of recommendations are widely used in many different ways. Amazon uses them as a part of their recommender system and also uses other features like star rating to make them more useful and usable for users. Main types of explanation are graphs with numbers of ratings, collaborative filtering based explanations and item-to-item based explanations. This approach is very easy to use and together with incorporating other features of the system (star rating, etc.) it is very interesting approach. However, the explanation is always the same and thus do not take into account the differences between users.

IMDB type of explanation is similar since they are using explanations together with ratings of items to make explanations more reliable for users. However, they are also providing a deeper analysis of these ratings with demographic breakdowns. This type of explanation can be very useful for users but there is harder to find it and thus the whole idea loses its meaning. In the same way, explanations itself are not personalized, which we see as a barrier.

Last.fm uses explanations to a lesser extent than other two systems before. However, their approach is one of the most interesting. They are showing the whole list of similar items together with the level of similarity and thus they make it easier to choose. However, personalization aspect is missing here too.

Most of the real world applications, web sites or recommender systems itself using the content-based recommendations. Subsequently, from this technique is derived the explanation style. This is the reason why the content-based explanations are that widely used. Nowadays, the most popular use of explanation is one general sentence, for example: *"People who liked this also liked ..."*. This approach is standard on recent systems and works quite well but it have some disadvantages. The biggest one is that the sentence is the same in every moment and for every user. However, every person is different and we need to somehow adjust even the explanation to these personal preferences of users. Moreover, these sentences are telling the truth on one hand, but just on high level of detail. For the users, it is hard to trust the system more, when they get this type of explanation on high level of detail without any personalization.

5. Hybrid Explanation Method

To answer our research questions in real-world scenario, we use a recommender service in the news domain. Our aim is to recommend news articles to users and present them on a website. For the purpose of presentation of the recommendations, our effort is to find optimal settings in the context of explanation of recommendations. According to our analysis, there are many recommended systems, which use explanations. However, in most cases, these explanations are fully based on recommendation technique (white-box). Such approach seems to us quite bonding, since it cannot be used for some current recommender techniques (e.g., latent features based). When focusing on specific domain, i.e. news, it is even more challenging to find explanation approach for the news recommender. In [6,5] authors proposed a news recommender systems, which uses a white-box explanations generated based on user previous feedback to similar articles. Therefore, we bring a novel, *black box* approach to explanations. For this purpose, we have proposed a method for hybrid explanation of recommendations – it generates content and collaborative explanations, independent of the used recommendation approach and tailored to user needs.

According to our method, *personalized explanation* is an approach which generates explanations with regard to user preferences. Thus, each user will be given an explanation adapted to what most impressed him (i.e., explanation style which he/she prefers). Another advantage is that proposed method is not dependent on recommendation technique that was used to recommend item. On the other hand, the recommendation techniques are used in the process of explanations generation. Based on the information about the

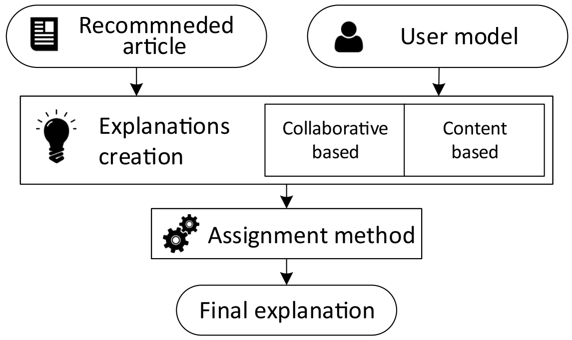

Fig. 8. Idea of proposed method for personalized back box explanation consisting of three steps: 1. Data acquisition; 2. Explanation creation; 3. Explanation assignment.

recommended items and about the user, we generate and present an explanation that is useful and interesting to the user.

Our method combines different approaches for explanation (explanation styles) to find suitable explanation for the actual user. Thus, we present an explanation that fits the user's preferences. For each recommended item, it is necessary to find explanation that is suitable in the context of characteristics of this item and which is also suitable for specific user. Proposed method is thus a hybrid and personalized type of explanation.

Our method follows three basic steps (Figure 8):

1. **Data acquisition**
   In the first step, there are two inputs: the first input is a list of recommended items together with their characteristics (id, title, author, content, etc.) and keywords (basic metadata). For the item representation we use only keywords extracted from the article content. Extraction of keywords from the content of the recommended items (articles) was accomplished by the following steps:
   - Removing special characters from text (html tags, punctuation, numbers, etc.)
   - Conversion to lowercase
   - Removing stop words
   - Getting individual words from the text - tokenization
   - Assessment of the number of occurrences of individual words

   The second input is the information about the users represented by their user model. The user model is determined by the keywords from the liked items (articles) read by the user previously. In other words, we use the bag of the words user model. With these two data sources, we can generate personalized explanation of recommended items (e.g., articles).
2. **Explanation creation**
   In the second step, the actual explanations are generated based on both methods of personalized explanations, which use these two approaches:
   - *Explanation based on similar users* – Explanation based on similar users uses the user's neighbors and items that they read. The basic procedure for generating explanations is as follows:
     (a) Selection of recommended items and their characteristics or keywords.

(b) Selection of the *k* - neighbors of the user and their preferences. Neighbors are other users who interact with (e.g., read, bought, clicked on) similar items as the actual user. Preferences consist of items that neighbors interact with in the past.
(c) Find a correlation between the recommended item and the items that have been read by neighbors of actual user. This comparison is based on the calculation of cosine similarity (based on users user models).
(d) If the correlation is above the selected threshold, then we will generate the explanation (in the news domain): *"This article was recommended to you because the similar user XY read this article too"*.

– *Explanation based on the content of items* - Explanation based on the content of items uses items that the user interact with in the past. The basic procedure for generation of the explanations consists of these steps:
  (a) Selection of recommended items and their characteristics or keywords.
  (b) Select items that user interact with in the past, together with their characteristics or keywords.
  (c) Find a correlation between the recommended item and items that the user interact with in the past. This comparison is again based on the calculation of cosine similarity (based on items representation - keywords).
  (d) If the correlation is above the selected threshold then we will generate explanation (in the news domain): *"This article was recommended to you because it is similar to article that you read before and his title was: XY"*.

3. **Explanation assignment**
The third step of proposed approach is the assignment of appropriate type of explanation for the specific user. Initial setting for a new user assigns to both explanation types the same weight. Results of these two explanations are presented in a standard way called the interleaving list. With this, we create a list of recommended items together with explanations in order to identify if users prefer a certain explanation style or approach. The learning phase is performed by continuous monitoring the user clicks on each item explained by different approaches and adjusting their weights based on user's preferences (Equation 1). This refers to the online learning, specifically test-then-train approach (idea similar as bandit algorithms). As a result, the users obtain more explanations generated by their preferred type (based on the item content or based on the similar users).

$$No_{cont} = Min\left(Round\left(\frac{\frac{Clk_{cont}+1}{Clk_{coll}+1}}{Clk_{cont}+Clk_{coll}+2} \times No_{to\ rec}\right), No_{to\ rec}-1\right), \quad (1)$$

where No_{cont} is the number of content explanations to generate, $No_{collab} = No_{to\ rec} - No_{cont}$, Clk_{cont} and Clk_{coll} is the number of content and collaborative explanations user clicks in his/her history and $No_{to\ rec}$ is the number of explanations to generate. The idea is to present to the user more explanations by the type hi/she likes. On the contrary, at least one of the explanations differ in the style and thus the users have a chance to adjust their preferences over the time (at last one explanation is generated by different method as the rest).

Fig. 9. Interleaved list for the new users (User interface of system ExplORe (setting ”with recommendation”).

6. Evaluation

As part of the evaluation of our method, we conducted several experiments where we used our method to explain recommendations to specific users in the news domain. In a live uncontrolled experiment with real users we recorded their activities while reading articles in our system. This system has been developed for the purpose of displaying the recommended news articles and related explanations.

6.1. ExplORe – a system for article recommendation

For the purpose of experimentation with our explanation method, we developed a system, which provides the place for displaying recommended articles and explanations generated by our method. System interface (Figure 9) contains news articles with explanation why they are recommended (setting ”with explanations”). Every user has displayed 10 most suitable articles and after he/she click on some of them, this article is no more part of the list of recommended articles. The system will replace it with the next most suitable item from the recommendation list.

ExplORe is the recommender system that uses recommendations to identify suitable articles for a particular user. We analyzed advantages and disadvantages of several options of recommendation approaches. Consequently, we decided to use the library Apache Mahout, which is quite widespread [35,30] and reported to be well performing. Moreover, it provides a number of features for a recommendations and evaluation of the accuracy of the recommendations.

The system itself also implements our hybrid explanation method. It was implemented according to description in Section 5 and consist of two approaches. Both approaches work with the keywords associated with individual articles and compare them in order to find a suitable explanation. Specifically, the method works with these two basic models:

– Model of news articles – each article has assigned 10 keywords
– User model – each user is identified by keywords of articles that he/she read

These models help us to generate appropriate explanations for users. However, every user can perceive the explanations in different way even if they are equally suitable. Therefore, it is also important to know where and how to display explanations as a part of user interface. Thus as a part of design procedure, we conducted user eye-tracking study. The aim was to identify optimal position of explanations and amount of information which should be presented to users.

6.2. Eye-tracking user study

In this first experiment, we focused on obtaining basic information on the location and structure of information displayed in our system. We subsequently used this information to design the user interface of our system (Figure 10). We focused on:

– Presentation of recommendation: amount and structure of information
– Explanation of recommendation: positioning and visualization of explanation

Settings For each of the areas listed above, we created a few designs and showed them to participants of our eye-tracking study. This user study was conducted with 10 participants (8 men and 2 women), which is standard eye-tracking sample size and should eliminate major UX problems [33]. All participants were university students in age from 20 to 28. The participants should choose one article which was the most interesting for them. This showed us how difficult is for people to find this specific article in different conditions according to eye movements.

For the purpose of tracking the eye movements of participants, we used eye-tracker Tobii TX-300. This device has recorded both eyes at a rate of 300 Hz during our experiments. For evaluation of results from this study, we used basic gaze metrics: Fixation duration, Total fixation duration, Fixation count and Time to first fixation. According to I-VT fixation filter [27], eye-movement is considered as a fixation if its duration is above the minimum fixation duration (60 ms). Similarly, two fixations are merged when the duration between them is smaller than maximum time between fixations (75 ms).

After we finish eye tracking part of our experiment, we spoke with participants and ask them about their opinions on particular designs to find their subjective views.

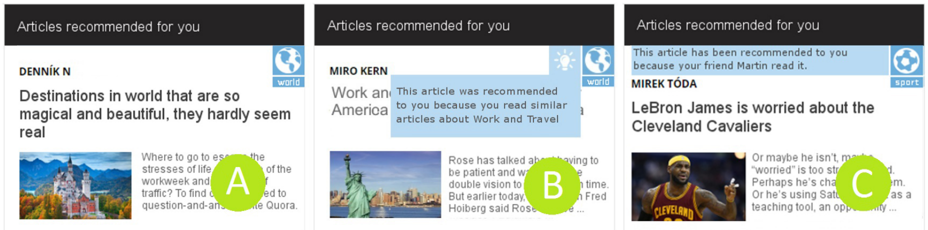

Fig. 10. Designs of explanation visualization presented to participants.

Results In case of presentation of recommendation, we found out that participants prefer articles described with short description, title, category and image. However, our main focus was to find out what is the proper way to visualize explanations of recommended articles. We created three designs of explanation visualization (Figure 10):

– Explanation next to image (Fig. 10 part A)
– Explanation in pop-up menu (Fig. 10 part B)
– Explanation above the title of article (Fig. 10 part C)

We analyzed fixation metrics as time to first fixation to determine how much time users took to find the explanation. Moreover we analyzed number and duration of fixations on explanations, along with the duration of viewing recommended items to uncover the difficulty of getting explanation information. According to these eye-tracking metrics and interview with participants, we decided to implement third option which displaying explanations above the title of articles. This options was easiest to find according to eye fixations (Figure 11). Participants also like the second option but we find out that some of them had a problem to find out that they have to point cursor on specific icon to open the pop-up menu with explanation.

6.3. Recommendations explanation method

In order to evaluate our hybrid method of explanation of recommendations, we conducted an experiment by use of our ExplORe system. Our goal was to find out what is the quality of explanation method in terms of impact on users while choosing articles and which approach will be more suitable in which specific conditions.

Settings To explore the features of proposed approach a long-term experiment with simulating the real media with news articles was conducted. Duration of the experiment was 18 days. Experiments were attended by a total of 17 people and 13 of them were university students. Up to 15 participants were aged 20-30 years. 17 participants were randomly divided into two groups without any specific criteria for the division: Group α (group consists of 8 participants), Group β (group consist of 9 participants). As the experiment toked 18 days, the number of participants is similar to other academia studies in recommender systems [34,22,40]. On the contrary, we are aware of limitations of the conclusions we address in this paper, which provide a chance for further exploration.

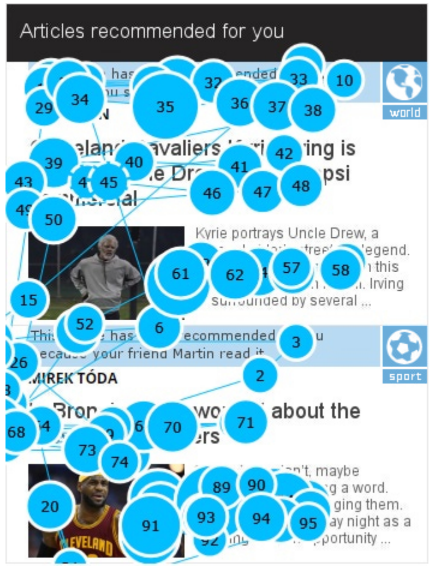

Fig. 11. Fixations on the explanation component of the ExplORe system.

	Group α	**Group β**
Phase 1	without explanations	with explanations
Phase 2	with explanations	without explanations

Table 1. Phases of the experiment and configuration settings.

Each group started with a different setups of the system. Group α started with newspaper articles without explanation. Group β started with newspaper articles along with explanations. In the middle of the experiment, the groups exchange their setups (Table 6.3). We also let participants complete two questionnaires. The first in the middle of the experiment, before the change of the setups. Second at the end of the experiment.

For the evaluation, we report the number of clicks on recommended articles and explanations. Article to which the participant clicked was always replaced by another article in the list. In further evaluation, we also report the Precision of the recommendation with and without explanation as a part of indirect evaluation.

Results We conducted the evaluation of the experiment in respect of all participants in the experiment but also in terms of individual groups so that we were able to point to individual differences in the data. The results can be divided into two main groups based on observed metrics:

– Impact of personalized explanations
– Comparison of different approaches

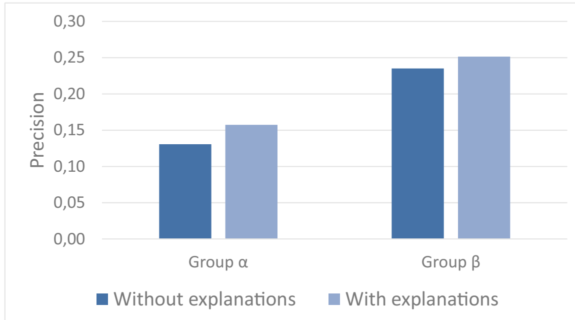

Fig. 12. Precision metric for both groups (settings).

RQ1: Can we increase the understandability and precision of recommendations with use of explanations?

Distrust of users in recommendations is relatively serious problem that the recommender system must face. During the experiments, we let participants to fill two questionnaires. The questionnaires were filled in by all participants and we find out that large part of them (14 out of 17 participants) reading news articles almost every day. The same number of participants also know the concept of recommendations.

In these questionnaires, we also asked participants about their experience with recommender systems. For many participants (12 out of 17) was unpleasant to know that recommender systems are watching them and that these systems know a lot of information about them. This confirms our assertion that users do not trust recommender systems too much.

However, we also asked them if it possible to reduce their distrust by using the explanations (as a tool for better understanding). A large majority of participants (15 out of 17) said that they will trust recommender systems more if they will use explanations. The results of our experiments and following questionnaires showed that our hybrid method of explanation of recommendations can reduce the distrust of user in recommender systems and thus to also increase transparency of these systems.

Basic statistics in the context of further evaluation is the frequency of clicks on articles without explanations and articles with explanations. We discovered that articles with explanations were clicked 1.24 times more (1 196 times) compared to articles without explanations (1 064 times).

Moreover, we looked at the same ratio in terms of precision between different groups (Figure 12). The graph clearly shows that in both cases there is an increase of the precision on articles with explanations. That means that users click more on articles with explanations (supported also by the questionnaire). Our assumption was that personalized explanations increase the number of clicks on articles. This was shown in both groups (Table 6.3) of participants. Thus, this result shows that explanations increase understandability and attractiveness of recommended articles.

Within the comparison of articles with explanations and without explanation, we looked on the ability to persuade the user to read an article, even if the article was not fully appropriate and recommender system ranked it on lower positions in the list of rec-

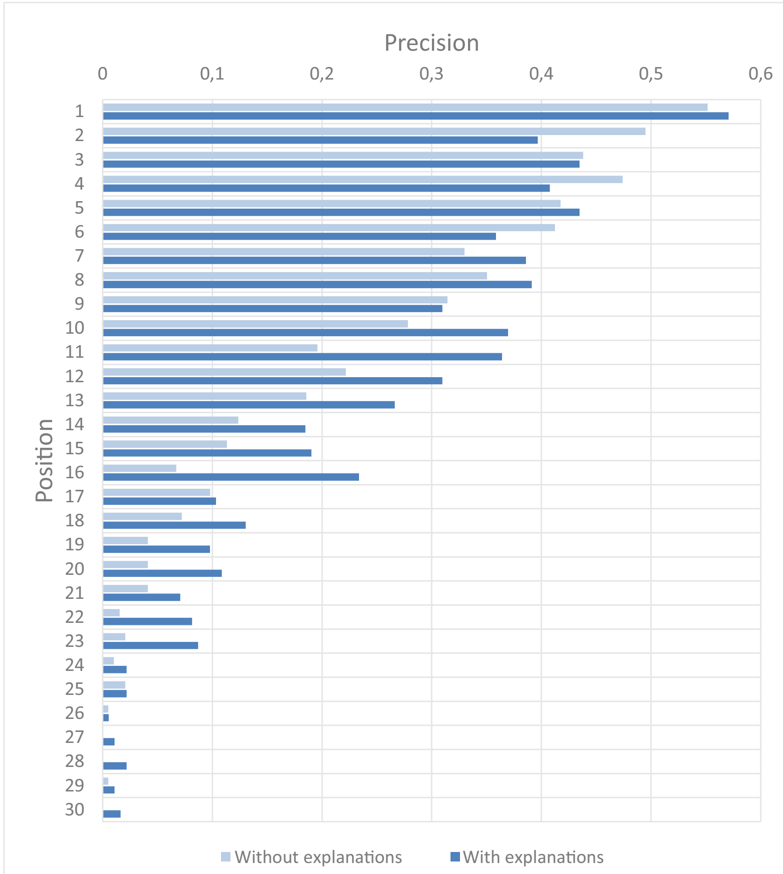

Fig. 13. Precision on positions in the list of recommended articles.

ommended articles. The following diagram (Figure 13) shows precision for each position in the list of recommended articles with and without explanations.

The figure shows that the articles in lower positions have higher precision when they are displayed with explanation. A paired t-test was conducted to compare precision over positions for two groups. On group had articles displayed with explanations and other without explanations. The difference is considered to be statistically significant ($t(29) = -3,363424$; $p = 0,002178$; $\alpha = 0,05$). This demonstrate that people click on articles which are less suitable for them more when they were explained by our explanation method. In other words, articles which are often ignored by users (presented on lower list positions), users found more interesting - while the explanations provide reasons for it. The figure thus clearly shows that explanation make articles more attractive and the recommendation more understandable.

Finally, we can conclude that participants clicked on the articles with explanations more than on the articles without explanations. This was supported in the context of precision, when we obtained higher values for individual users and also for whole groups.

Overall, we can evaluate the explanations in this case as a success in terms of impact on the understandably and precision of recommendations.

RQ2: Is there one explanation style that is preferred by users?

Firstly, we watched how often users clicked on various explanation approaches or styles. We discovered that the articles with explanations based on the content of articles had greater percentage of clicks (more than double) then articles with explanation based on similar users.

If we look at the precision in the context of each group we can see a similar result (Figure 14). In this graph, it is confirmed that the increase between the individual approaches are about twice as big. The first graph indicates that all users prefer one approach to explanation - content-based explanation.

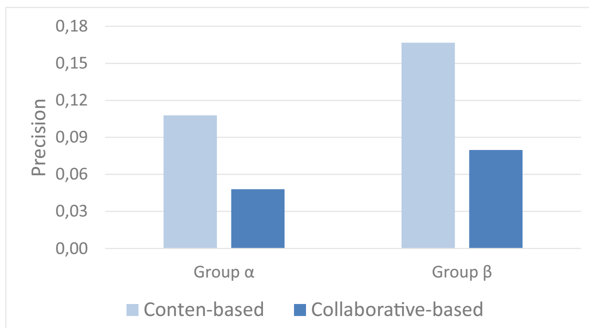

Fig. 14. Precision on articles with various explanations types for both groups (settings).

We hypothesize that the dominance of content based explanations was caused by two main reasons - Character of individual methods for explanation and Character of experiment. As our aim was to observe different preferences over various explanations types, the ratio between content and collaborative explanations was fixed to 0.5 (we ignored assignment mechanism from Equation1). First case was caused by nature of explanation based on similar users. In this approach, it was necessary to find someone else who reads the new article recommended for actual user. If there is nobody who reads this article, then we cannot use this type of explanation. This means that we explain a few more articles with content based approach then with approach based on similar users.

Second case was, in our opinion, caused by nature of experiment itself, which was attended by 17 participants. Problem was, that not all of these participants knew each other. Therefore, content based explanations could be more attractive for participants. If one does not know the person mentioned in the explanation based on similar users, he/she is less interested to click on the following article.

In this chapter we evaluate our approach with several experiments. Results of these experiments show that the idea of hybrid explanations make sense. We showed that different users prefer different type of explanations and that is first step towards the creation of the method of hybrid explanation.

7. Conclusions

In our work, we address the topic of recommendations presentation via web. Recommender systems are trying to solve the problems of information overload but often act as a black box. Users do not know how they recommend them each item or why these systems store a personal information about user preferences. In the context of these problems, we can speak about distrust of users.

To increase users' satisfaction and attitude to recommendations, the explanations are often used. Generally, the topic of the trust and human-computer interactions is a widely researched nowadays. In this paper, we proposed a method of presentation and explanation, which would be interesting for users while also address some of the problems of recommendation systems followed in our research questions.

To find out answers, we proposed an approach to explanation of the recommendations together with a suitable and transparent presentation of these explanations and recommendations themselves. This is the so-called hybrid method of personalized explanation of recommendations. The basic characteristics of this method are:

- The method is independent of the recommendation technique (black-box explanations)
- The method provides a personalized type of explanation that combines two approaches to explaining a) Content-based explanation; b) Collaborative-based explanations

We performed several experiments to evaluate proposed approach. We conducted a users study of system ExplORe using an eye-tracking tool (eye-tracker). We have mainly focused on obtaining basic information about the location and structure of presentation of explanations. Subsequently, we used this information to design the user interface of our system.

To answer our research questions, we conducted long term experiment in the domain of news. The results of the experiments have shown that users prefer articles with explanations compared to the article without explanations. We also found that users strongly preferred the content-based explanations compared to collaborative-based. The most interesting finding was that the explanations were able to convince participants to read less suitable articles (items that have been placed on the lower position in the list of recommendations (Fig. 15)). This is extremely useful from the business perspective, as we were able to increase the precision of recommendations, while the recommended method remains untouched. Last but not least, by using a questionnaire, we subsequently found that a large majority (15/17) of participants believe that the explanations can reduce their distrust in recommendation systems and thus actually increase the transparency of recommendations. This is a promising outcome, which can be used in further, large scale, online experiment.

As our aim was to explore the idea of personalized explanations, there are several other aspects, which should be addressed in following research. The idea of personalized explanations is to provide various explanations for various users. We proved, that users with similar demography prefers similar explanation approaches. However, this seems to be more important for users with different demographic (e.g., elderly users).

These results were obtained within the domain of news. However, we believe that the similar approach can be successful also in other domains, but this has to be proven

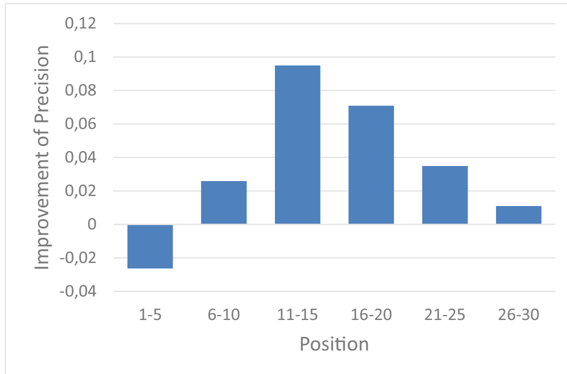

Fig. 15. The improvement of precision (recommendations "with explanations" compared to recommendations "without explanations") on recommendation list position.

by another experiment. Only big difference within these experiments will be the use of different characteristics of recommended items than with articles.

To sum up, we proposed a novel method of hybrid explanation of recommendations, which is independent on the recommender approach. Proposed method improved the precision of lower-positioned items in the recommendation list. Moreover, the participants express positive attitude to explanation as an approach for distrust reduce.

We see a great potential that our method brings into the area of recommender systems. This method has the potential of reducing the problems associated with the recommendations. However, in the future, the personalized explanation can be used in many other areas. The method is designed to include more explanation styles as explored in our experiments. This is useful for application in various domains. For instance the collaborative explanations can be considered as privacy issue in some domains (e.g., pharmacy e-shop). On the contrary, only users friends (after explicit agreement) should be used for such explanation style.

From our perspective, the most interesting area is to find which type of explanation is suitable for different type of articles. Thus, we can personalize explanation not only to user but also to items of recommendations (e.g., different type of articles). Second area is the context of user. Idea behind this concept is to show different types of explanation based on the context (e.g., season, weather, etc.) of the actual user.

Acknowledgments. This work is partially supported by the Slovak Research and Development Agency under the contract No. APVV-15-0508, the Scientific Grant Agency of the Slovak Republic, grants No. VG 1/0646/15 and VG 1/0667/18, and is the partial result of the Research & Development Operational Programme for the projects ITMS 26240220084 and ITMS 26240220039, co-funded by the European Regional Development Fund.

References

1. Aggarwal, C.C.: Recommender Systems: The Textbook. Springer (2016), `https://doi.org/10.1007/978-3-319-29659-3`

2. Al-Taie, M.: Explanations in recommender systems: overview and research approaches. In: Proceedings of the 14th Int. Arab Conf. on Information Technology, Khartoum, Sudan, ACIT. vol. 13 (2013)
3. Amatriain, X., Pujol, J.M., Oliver, N.: I Like It... I Like It Not: Evaluating User Ratings Noise in Recommender Systems, pp. 247–258. Springer, Berlin, Heidelberg (2009), http://dx.doi.org/10.1007/978-3-642-02247-0_24
4. Bielikova, M., Kompan, M., Zelenik, D.: Effective hierarchical vector-based news representation for personalized recommendation. Computer Science and Information Systems (21), 303–322 (2012), http://eudml.org/doc/252774
5. Billsus, D., Pazzani, M.J.: A personal news agent that talks, learns and explains. In: Proceedings of the Third Annual Conference on Autonomous Agents. pp. 268–275. AGENTS '99, ACM, New York, NY, USA (1999), http://doi.acm.org/10.1145/301136.301208
6. Billsus, D., Pazzani, M.J.: User modeling for adaptive news access. User modeling and user-adapted interaction 10(2-3), 147–180 (2000)
7. Brusilovsky, P.: Adaptive Hypermedia: From Intelligent Tutoring Systems to Web-Based Education, pp. 1–7. Springer, Berlin, Heidelberg (2000), http://dx.doi.org/10.1007/3-540-45108-0_1
8. Brusilovsky, P., Kobsa, A., Nejdl, W. (eds.): The Adaptive Web, Methods and Strategies of Web Personalization, Lecture Notes in Computer Science, vol. 4321. Springer (2007), https://doi.org/10.1007/978-3-540-72079-9_4
9. Burke, R.: Hybrid recommender systems: Survey and experiments. User Modeling and User-Adapted Interaction 12(4), 331–370 (Nov 2002), http://dx.doi.org/10.1023/A:1021240730564
10. Celma, O., Herrera, P.: A new approach to evaluating novel recommendations pp. 179–186 (2008), http://doi.acm.org/10.1145/1454008.1454038
11. Chang, S., Harper, F.M., Terveen, L.G.: Crowd-based personalized natural language explanations for recommendations. In: Proc. of the 10th ACM Conf. on Recommender Systems. pp. 175–182. RecSys '16, ACM, New York, NY, USA (2016), http://doi.acm.org/10.1145/2959100.2959153
12. Chen, Y.C., Shang, R.A., Kao, C.Y.: The effects of information overload on consumers' subjective state towards buying decision in the internet shopping environment. Electron. Commer. Rec. Appl. 8(1), 48–58 (Jan 2009), http://dx.doi.org/10.1016/j.elerap.2008.09.001
13. Costa, G., Ortale, R.: Model-based collaborative personalized recommendation on signed social rating networks. ACM Trans. Internet Technol. 16(3), 20:1–20:21 (Jul 2016), http://doi.acm.org/10.1145/2934681
14. Gedikli, F.: Recommender Systems and the Social Web: Leveraging Tagging Data for Recommender Systems. Springer Vieweg (2013)
15. Gedikli, F., Jannach, D., Ge, M.: How should i explain? a comparison of different explanation types for recommender systems. International Journal of Human-Computer Studies 72(4), 367–382 (2014)
16. Herlocker, J.L., Konstan, J.A., Riedl, J.: Explaining collaborative filtering recommendations. In: Proc. of the 2000 ACM Conference on Computer Supported Cooperative Work. pp. 241–250. CSCW '00, ACM, New York, NY, USA (2000), http://doi.acm.org/10.1145/358916.358995
17. Hu, J., Zhang, Z., Liu, J., Shi, C., Yu, P.S., Wang, B.: Recexp: A semantic recommender system with explanation based on heterogeneous information network. In: Proc. of the 10th ACM Conf. on Recommender Systems. pp. 401–402. RecSys '16, ACM, New York, NY, USA (2016), http://doi.acm.org/10.1145/2959100.2959112
18. Jannach, D., Zanker, M., Felfernig, A., Friedrich, G.: Recommender Systems: An Introduction. Cambridge University Press, New York, NY, USA, 1st edn. (2010)

19. Jannach, D., Zanker, M., Felfernig, A., Friedrich, G.: Recommender Systems: An Introduction. Cambridge University Press, New York, NY, USA, 1st edn. (2010)
20. Jawaheer, G., Szomszor, M., Kostkova, P.: Comparison of implicit and explicit feedback from an online music recommendation service pp. 47–51 (2010), http://doi.acm.org/10.1145/1869446.1869453
21. Jiang, W., Wu, J., Wang, G.: On selecting recommenders for trust evaluation in online social networks. ACM Trans. Internet Technol. 15(4), 14:1–14:21 (Nov 2015), http://doi.acm.org/10.1145/2807697
22. Kowald, D., Seitlinger, P., Ley, T., Lex, E.: The impact of semantic context cues on the user acceptance of tag recommendations: An online study. arXiv preprint arXiv:1803.02179 (2018)
23. Linden, G., Smith, B., York, J.: Amazon.com recommendations: Item-to-item collaborative filtering. IEEE Internet Computing 7(1), 76–80 (Jan 2003), http://dx.doi.org/10.1109/MIC.2003.1167344
24. Muhammad, K., Lawlor, A., Rafter, R., Smyth, B.: Great Explanations: Opinionated Explanations for Recommendations, pp. 244–258. Springer, Switzerland, Cham (2015), http://dx.doi.org/10.1007/978-3-319-24586-7_17
25. Musto, C., Narducci, F., Lops, P., De Gemmis, M., Semeraro, G.: Explod: A framework for explaining recommendations based on the linked open data cloud. In: Proc. of the 10th ACM Conf. on Recommender Systems. pp. 151–154. RecSys '16, ACM, New York, NY, USA (2016), http://doi.acm.org/10.1145/2959100.2959173
26. O'Donovan, J., Smyth, B.: Trust in recommender systems. In: Proceedings of the 10th International Conference on Intelligent User Interfaces. pp. 167–174. IUI '05, ACM (2005), http://doi.acm.org/10.1145/1040830.1040870
27. Olsen, A.: The tobii i-vt fixation filter. Tobii Technology (2012)
28. Ricci, F., Rokach, L., Shapira, B., Kantor, P.B.: Recommender Systems Handbook. Springer-Verlag New York, Inc., New York, NY, USA, 1st edn. (2010)
29. Rich, E.: User modeling via stereotypes. In: Cognitive Science: A Multidisc. Journal. vol. Vol. 3, No. 4, pp. 329–354 (1979)
30. Schelter, S., Owen, S.: Collaborative filtering with apache mahout. In: Proc. of ACM RecSys Challenge (2012)
31. Sutcliffe, A.G., Wang, D., Dunbar, R.I.M.: Modelling the role of trust in social relationships. ACM Trans. Internet Technol. 15(4), 16:1–16:24 (Nov 2015), http://doi.acm.org/10.1145/2815620
32. Tintarev, N., Masthoff, J.: Designing and Evaluating Explanations for Recommender Systems, pp. 479–510. Springer US, Boston, MA (2011), http://dx.doi.org/10.1007/978-0-387-85820-3_15
33. Turner, C.W., Lewis, J.R., Nielsen, J.: Determining usability test sample size. International encyclopedia of ergonomics and human factors 3(2), 3084–3088 (2006)
34. Verbert, K., Parra, D., Brusilovsky, P.: Agents vs. users: Visual recommendation of research talks with multiple dimension of relevance. ACM Transactions on Interactive Intelligent Systems (TiiS) 6(2), 11 (2016)
35. Walunj, S.G., Sadafale, K.: An online recommendation system for e-commerce based on apache mahout framework. In: Proc. of the 2013 Annual Conference on Computers and People Research. pp. 153–158. SIGMIS-CPR '13, ACM, New York, NY, USA (2013), http://doi.acm.org/10.1145/2487294.2487328
36. Yang, C.C., Chen, H., Hong, K.: Visualization of large category map for internet browsing. Decis. Support Syst. 35(1), 89–102 (Apr 2003), http://dx.doi.org/10.1016/S0167-9236(02)00101-X
37. Yoo, K.H., Gretzel, U., Zanker, M.: Persuasive Recommender Systems: Conceptual Background and Implications. Springer Briefs in Electrical and Computer Engineering, Springer, New York (2013)

38. Yuksel, B.F., Collisson, P., Czerwinski, M.: Brains or beauty: How to engender trust in user-agent interactions. ACM Trans. Internet Technol. 17(1), 2:1–2:20 (Jan 2017), http://doi.acm.org/10.1145/2998572
39. Zhang, J., Curley, S.P.: Exploring explanation effects on consumers trust in online recommender agents. International Journal of Human–Computer Interaction 34(5), 421–432 (2018)
40. Zhao, Q., Chang, S., Harper, F.M., Konstan, J.A.: Gaze prediction for recommender systems. In: Proc. of 10th ACM Conf. on Recommender Systems. pp. 131–138. ACM, USA (2016), http://doi.acm.org/10.1145/2959100.2959150
41. Zimmerman, J., Kurapati, K.: Exposing profiles to build trust in a recommender. In: CHI'02 Extended Abstracts on Human Factors in Computing Systems. pp. 608–609. ACM (2002)

Martin Svrcek received an M.S. degree in Information Systems from Slovak University of Technology in Bratislava, Slovakia in 2016. He is currently a PhD student at Slovak university of Technology in Bratislava. His research interests include web technologies, data science, eye tracking the user behaviour and user experience in general.

Michal Kompan received his PhD. degree in 2014, from the Slovak University of Technology in Bratislava. He is an associate professor, presently at the Institute of Informatics, Information Systems and Software Engineering, Slovak University of Technology. His research interests are in the areas of personalized recommendation for single and group of users, user modeling and machine learning. He is a member of IEEE Computer Society and ACM.

Maria Bielikova is a full Professor at the Institute of Informatics, Information Systems and Software Engineering, Slovak University of Technology in Bratislava. She is a head of User Experience and Interaction Research Centre at the same university. Her research interests include web-based systems, emphasizing user modelling and personalization, context awareness, and usability. She is a member of the Editorial Board of the Int. Journal of Web Engineering (JWE) and User Modeling and User-Adapted Interaction (UMUAI). She is a senior member of IEEE, a member of the IEEE Computer Society, and a senior member of ACM.

Received: December 17, 2017; Accepted: June 30, 2018.