This paper (or a similar version) is not currently under review by a journal or conference, nor will it be submitted to such within the next three months

Mediation of User Models for Enhancing Personalized Services Delivery

Shlomo Berkovsky[1], Tsvi Kuflik[1], Francesco Ricci[2]

[1] University of Haifa, Haifa,
slavax@cs.haifa.ac.il, tsvikak@is.haifa.ac.il

[2] Free University of Bozen-Bolzano, Italy
fricci@unibz.it

Abstract. Provision of personalized services to users requires accurate modeling of their interests and needs. This paper proposes a general framework and some precise methodologies to enhance the accuracy of current user modeling in personalization systems by importing and aggregating data collected by other personalization systems. Such a process is defined as user models mediation. The paper discusses the details of such a generic user modeling mediation framework. It provides a generic user modeling data representation model, demonstrates its compatibility with existing personalization techniques, and discusses the general steps of the mediation. Specifically, four major types of mediation are presented: cross-user, cross-item, cross-context, and cross-representation. Finally, the paper applies the proposed framework and illustrates it with two practical mediation scenarios. Evaluations of these scenarios show that the mediation of user modeling data is practical and highly beneficial, as it allows upgrading the quality of the personalization provided to the users.

1 Introduction

During the last decade, the quantity of potentially interesting items, products or information services available on the Web has been growing rapidly and now exceeds human processing capabilities [34]. Moreover, there are many information search situations where the users would like to choose among a set of alternative items or services, but do not have sufficient knowledge, capability or time to make such decisions. As such, there is a pressing need for intelligent systems that advice users while taking into account their personal needs and interests, and deliver tailored service in a way that will be most appropriate and valuable to the users. This type of system is referred to in the literature as a personalization system [41].

Whatever the specific technology exploited by a personalization system, it can provide high quality personalized services to users only after having modeled their preferences. This information is referred to in the literature as the User Model (UM) [29]. The task of collecting the user modeling data can be performed in two ways: (1) explicit – through provision of the required information by the user, or (2) implicit –

through applying various reasoning mechanisms that infer the required information based on the user's observable behavior [24]. The explicit collection of user modeling data is considered to be a time- and effort-consuming task, typically avoided by the user. Alternatively, the implicit collection involves automated reasoning mechanisms, which can misinterpret user behavior. In practice, the explicit and implicit approaches may also be combined [32].

In general, the quality of personalized services provided to the user depends largely on the characteristics of the UMs, e.g., how accurate it is, what amount of information it stores, and whether this information is up to date. Hence, as a general rule, the more information is stored in the UM, i.e., the more knowledge the system has obtained about the user, the better the quality of the personalization will be. In this context, quality refers to the capability of the system to suggest exactly those products or services that the user will select and purchase, or to predict correctly those items that the user will like. In practice, obtaining sufficient user modeling data to deliver high quality personalization is difficult. This issue is especially important at the initialization stages, where all the existing personalization techniques face the bootstrapping problem, i.e., a situation where the information about the user and/or items does not suffice to provide high quality personalized services [33].

For example, consider recommender systems [46], as a typical example of personalization systems. These systems provide users with recommendations about items or products they may like. They generate personalized recommendations (i.e., recommendations tailored to the user), by exploiting various knowledge sources, including information collected during past interactions with a community of users searching or providing recommendations, and the evaluations of those recommendations. Research of recommender systems started over a decade ago and yielded a wide variety of recommendation techniques, such as content-based filtering [40], collaborative filtering [22], knowledge-based [14] and utility-based [35] recommendations and their multiple hybridizations [15]. These techniques are widely discussed in the literature and in several surveys of the state-of-the-art recommender systems [51], [39], [3].

When analyzing current recommender systems, one can show that in the vast majority of these systems, every service provider runs an independent recommender system, building and maintaining a proprietary collection of UMs [39]. In particular, this means that the collected user modeling data are tailored to: 1) the specific content (product categories) offered by the recommender system, e.g., movies [23], music [4], news items [16], tourism planning [47], and so on, and 2) the recommendation technique being exploited by the system, e.g., collaborative filtering [26], content-based filtering [40], demographic filtering [31], or one of their hybridizations [15]. Thus, a large amount of heterogeneous (and possibly partially overlapping) user modeling data are scattered among various systems.

In general, practical personalization systems (especially commercial ones) neither allow other external personalization systems to access them, nor share their proprietary user modeling data. However, it can be hypothesized that personalization systems could benefit from enriching its user modeling data by importing and aggregating user modeling data collected by other, possibly related, systems, and therefore provide better personalized services to the users. This can be achieved through a process that referred to as the *mediation* of UMs [6].

The idea of UMs mediation raises a number of issues. The first one refers to the nature of the information market, and its business models. Due to commercial competition, real-life personalization systems currently do not cooperate, and do not share their user modeling data. In [2], the authors point out that typical personalization systems are either provider-centric (i.e., each provider has its own personalization engine to tailor its content to consumers) or market-centric (i.e., providing personalization services for a specific marketplace in a particular industry or sector). The authors claim that the lack of technical data sharing solutions in the existing personalization systems is mainly explained by business limitations to the exchange of user-related information among competing parties in the same market.

The second issue is guaranteeing customer privacy. UMs built by a certain personalization system may contain customers' private and sensitive information, that the user would not like to be disclosed to other systems, and possibly to untrusted parties [17]. For this reason, many personalization services that store sensitive information about their users declare in their privacy policies that no personal information stored by the system will be transferred to other parties [55]. As a result, they are not allowed to transfer information stored in their UMs to other systems.

The third issue is related to the structural heterogeneity and incompleteness of the user modeling data. As mentioned earlier, every personalization system refers to a specific application domain, and the services are provided using a specific personalization technique, which implies also specific UM representations. The lack of a standard representation for the UMs, jointly with specific representation, storage and access requirements imposed by various personalization techniques, results in the systems collecting user modeling data in different, ad-hoc forms. This heterogeneity causes several problems: various techniques may store the preferences of the same user in different forms; the information in different systems may be conflicting or outdated, and may be influenced by various cross-lingual and cross-culture dependencies; and so on. All these heterogeneities aggravate the mediation task, since it must support the integration and resolution of inconsistencies and conflicts among the user modeling data obtained from various systems [19].

This work focuses on resolving only the heterogeneity issue. The reader is referred to [8] for a discussion of privacy-preserving data exchange methods in personalization systems and to [44] for a discussion of integrated business models for personalization systems.

This paper discusses the details of a generic framework for the UMs mediation. Initially, it provides a generic data representation model, allowing various aspects of the personalization process to be expressed: users, items, contexts, and the personalization evaluation. Then, it describes several instantiations of the model, demonstrating its compatibility with some state-of-the-art personalization techniques. The paper proceeds with a discussion of one of the main problems of personalization systems: the sparseness of the UMs, which foils provision of accurate personalized services to the users. To overcome the sparseness problem and enrich the UMs stored by the personalization system, it is proposed to apply the mediation of UMs [6], i.e., to import user modeling data collected by other systems and to aggregate them using various reasoning and inference mechanisms.

Then, the UMs mediation process is comprehensively discussed. In here, the general steps of the mediation are described and four major types of mediation that can potentially be applied are defined: cross-user, cross-item, cross-context, and cross-representation (including two subtypes of the latter) mediations. Then, the paper show two practical examples and the results of empirical evaluations of two mediation mechanisms: the generalized cross-items mediation [11] and one subtype of cross-representation mediation [7]. Both results clearly show that the mediation of user modeling data is practical and highly beneficial, as it allows the quality of the personalized services provided to the users to be upgraded.

Finally, the paper presents current conclusions that can be drawn and discusses future veins of research on cross-context UMs mediation. Since state-of-the-art personalization systems mostly ignore the context-awareness issue (and are therefore very limited when required to provide context-aware personalized service), this mediation type cannot be applied. However, provision of context-aware personalized services is nowadays becoming one of the important research directions in personalization systems [48]. The feasibility of context-aware personalization immediately leads to a need for context-aware user modeling and cross-context mediation of UMs. Hence, the paper advocates the importance and future potential of cross-context mediation and describes several practical scenarios for performing this mediation.

2 Data Representation Model

Nowadays personalization systems base the warehousing of their UMs on a general two-dimensional representation typically referred to as the ratings matrix. The two *generalized* dimensions of this representation are the *users* and the *items*. These dimensions are referred to as generalized because they may be described by a number of specific features. Hence, if the user is described by *n* features and the item by *m* features, the space of all possible user and item pairs is described by an $n \times m$ dimensional space. When the users and the items are described by their unique identities only, one typically refers to a matrix to represent the ratings (evaluations) given by the users to the items. Ratings are given on a predefined scale and could be given in an explicit or implicit way. Explicit ratings are typically provided by the users, while implicit ratings are inferred by the system through observing user behavior indicators. For example, if the user bought the recommended product, the system implicitly interprets it as a positive rating.

When the users and the items are described by a set of features, the ratings matrix is still referred to as the description of the ratings given by users to items. However, in fact, such a matrix is a high dimensional matrix, i.e., it is a function R'_{gen} from the $n \times m$ dimensional space of $User_{feat} \times Item_{feat}$ to a set of *ratings*:

$$R'_{gen}: User_{feat} \times Item_{feat} \rightarrow rating.$$

In the above definition $User_{feat}$ represents the users, $Item_{feat}$ represents the items, and *rating* represents the ratings given by the users to the items.

In fact, this function is not defined for all the possible user and item pairs, i.e., the system may not know the rating value given by a user to all the items. This is typically denoted by adding to the rating values a special symbol, such as '?', to denote a missing rating value. In fact, the goal of a personalization system is, given a user *u* that requires a personalized service, e.g., a recommendation: (1) to guess the rating value for some item *i*, which the user has not previously rated, and (2) to suggest those items having maximal predicted rating [1]. Moreover, it should be stressed that, although the fundamental R'_{gen} function is originally defined as a two-dimensional matrix, both of its basic dimensions $User_{feat}$ and $Item_{feat}$ can be described using a multidimensional representation by a set of features. Given this definition of UMs warehousing, a single UM, i.e., the model of a concrete user, is considered as a set of the R'_{gen} contents restricted to values of the features of this user.

This general representation R'_{gen} is applicable to a wide variety of state-of-the-art personalization techniques. Now, the warehousing of UMs for the most common and well-known personalization techniques will be briefly presented. Note that the rest of this section mostly focuses on recommendation techniques, as typical representatives of personalization techniques.

- In collaborative filtering [26], the two-dimensional ratings matrix R_{CF} is represented by:

  $$R_{CF}: User_{id} \times Item_{id} \rightarrow rating,$$

  where $User_{id}$ and $Item_{id}$ are the unique one-dimensional identifiers of, respectively, users and items, and *rating* is the rating given by the user to the item. In this case, the UM is represented by a set of ratings given by the users, and is referred to in the literature as the ratings vector. For example, consider the following ratings matrix R_{CF}*={ ((Alice, "The Lord of The Rings"),1), ((Alice, "The Matrix"), 0.8), ((Bob, "Psycho"), 0.2), ((Bob, "Friday the 13th"), 0) }*, representing the ratings of two users, Alice and Bob, given on a continuous scale of ratings between *0* and *1*. Note that in this example all the pairs, where the rating is unknown *((u,i),?)* have been omitted. Typically, collaborative filtering systems do not store any additional information about the features and content characteristics of users and items, besides their identities.
- In content-based filtering [40] the two-dimensional ratings matrix R_{CB} is a simplified variant of the matrix, represented by:

  $$R_{CB}: User_{id} \times Item_{feat} \rightarrow rating,$$

  where $User_{id}$ represents the unique identifier of the users, $Item_{feat}$ represents a feature space describing the item's features, and *rating* represents again the user's ratings to the items containing the feature. In this case, the UM is represented by a set of ratings given by $User_{id}$ to certain $Item_{feat}$ values, originated by the descriptions of items. For example, content-based ratings matrix R_{CB} from the previous example is R_{CB}*={ ((Alice, "The Lord of The Rings", science-fiction),1), ((Alice, "The Matrix", science-fiction), 0.8), ((Bob, "Psycho", horror), 0.2), ((Bob, "Friday the 13th", horror), 0) }*. In this example, each movie is described by an additional content-related feature, the genre, taking in these examples the two values "horror" and "science-fiction". It can be observed that in content-based personalization systems the *raw* UMs comprise the ratings of the users to the items described by a set of features, but this information is then typically used to build a refined model that depends on the classifier technique. For instance, if the system is based on a Naïve Bayes classifier, the UMs are given by a set of probabilities linking the feature values to the ratings, and if it is based on a centroid, the UMs are described by the set of weights for the features describing the item [43].

- In demographic filtering [31], the two-dimensional matrix R_{dem} is represented by:

  $$R_{dem}: User_{feat} \times Item_{feat} \rightarrow rating,$$

  where $User_{feat}$ is a set of features describing certain demographic characteristics of a group of users to which the user belongs, $Item_{feat}$ represents either the unique identity of the item or a set of features reflecting item's content, and *rating* represents the respective rating. In this case, the UM is represented by a combination $(user_{feat}, item_{feat})$ and the ratings provided by the user described by $user_{feat}$ to the items, containing $item_{feat}$. For example, R_{dem} of the users Alice and Bob from previous examples is R_{dem}*={(female, science-fiction), 0.9), ((male, horror), 0.1)}*. In this example it should be pointed out that the very notion of user and item can depend on the personalization technology. Here, for instance, the user is represented by a user prototype and is defined only by the gender feature. Similarly, the item is described only by the genre feature. All the previous personalization techniques could, in principle, adopt such generalized user and item representation, even if this has been exploited mainly by demographic approaches that generalize the user description, and content-based methods that generalize the item description. In fact, ephemeral systems that provide personalized services based only on session data exploit the generalized user description and consider all users having the same description as equal [49].

In addition, also the way the ratings of the users are represented is highly heterogeneous across different personalization systems [24]. Some systems store explicit numeric ratings given on a predefined, but not standard, scale (e.g., the scale may be discrete or continuous, the range of possible values may vary from one system to another, and so on), some store explicit symbolic ratings (e.g., positive or negative ratings, thumbs-up or thumbs-down, and so on), some store implicit system-specific feedback derived from user behavior (e.g., examining or not the recommended item, purchasing or not the recommended product, and so on), some store the resultant navigation history of the users (e.g., opening or not the recommended Web-link, period of time spent viewing the recommended Web-page, and so on), and others store the free-text feedback provided by the users. To resolve this heterogeneity and to refer to the wide variety of user ratings to the provided personalization in a uniform manner, all the possible types of feedback are generalized and denoted by *evaluation* [37].

In conclusion, from the short overview presented above it is clear that in order to make an interoperable personalization system, exploiting the data originating from systems using various technologies, the heterogeneity of the user modeling data representations needs to be dealt with.

To start, the output of the R'_{gen} function, i.e., the rating assigned by or predicted for the user to an item, which is exploited by most state-of-the-art personalization systems, is generalized into the definition of the experience of a user for an item. An experience is defined as an evaluation function that maps a pair, the user that experienced the experience and the item experienced by the user, to an evaluation. An experience evaluation details how the user and the item are linked together. Formally, the experience is represented by:

$$Exp: User_{feat} \times Item_{feat} \rightarrow evaluation,$$

where $User_{feat}$ and $Item_{feat}$ represent the feature spaces describing user and item features, and *evaluation* represents the feedback given by the user described by $User_{feat}$ for the item described by $Item_{feat}$.

For example, consider an experience *e* described verbally by "Alice likes science-fiction movies". Using a simple object-oriented-like notation, this experience can be represented by *Exp(user.name=Alice, item.movie.genre=science-fiction)=like*. This representation of the experience shows that the *evaluation* of a user, whose feature *name* is assigned the value Alice for the *item movie*, whose feature *genre* is assigned the value *science-fiction* is *like*.

This allows generalizing further the above representation R'_{gen} of the ratings matrix to the *Exp* representation for UMs warehousing comprising user- and item-dependent representation of experiences and evaluations. Also over the *Exp* representation, a single UM, i.e., the model of a concrete user, is considered as a set of the *Exp* contents restricted to values of the features of this user. In other words, the user model for user *u* is the range of the *Exp* function restricted to the user *u*. Moreover, the *Exp* representation of experiences allows devising the following formulation: "the personalization (recommendation) task is a task of predicting future evaluation of the new experience for a specific combination of $(user_{feat}, item_{feat})$ values, based on a set of past experiences". In other words, the personalization is aimed at predicting evaluations of the new experiences using the knowledge obtained from past experiences.

However, the above generalization of the classical personalization problem does not overcome a severe limitation of the majority of personalization techniques: ignoring the context of the experience [13]. The term *context* has been defined several times in the literature. For example, it can refer to the user location, the time or the temperature of the day, and so on [12], or it can be considered as the subset of physical and conceptual states of interest to a particular entity [42]. In fact, one of the most comprehensive definitions of context is given in [18]: "*Context is any information that can be used to characterize the situation of an entity. An entity is a person, place, or object that is considered relevant to the interaction between a user and an application, including the user and applications themselves*".

With respect to personalization systems, [20] defines context as a description of aspects of a situation and splits the generic user context into five components: (1) environment context – captures the entities that surround the user;(2) personal context – captures the state of the user and consists of two sub-components, the physiological context and the mental context; (3) task context – captures what the persons (actors) are doing in this user context; (4) social context – captures the social aspects of the current user context, such as friends, enemies, neighbors, co-workers and so on; and (5) spatiotemporal context – captures aspects of the user context relating to the time and spatial extent for the user context.

In fact, user preferences represented by the UM are generally valid only within specific contextual conditions, such as spatial, temporal, emotional, and other conditions. That is, a user's preferences stored in the UM may change as a function of various contextual conditions. Nonetheless, the generalized ratings matrix representation *Exp* considers an experience as a user- and item-dependent entity only, not influenced by the contextual conditions, which may actually affect the evaluation of the experience. For example, two experience evaluations may be defined as "*Alice likes to see comedy movies with her friends*" and "*Alice does not like to see comedy movies with her parents*". In this example, if the companion of a user is treated as a contextual condition, the evaluation of the same experience is positive in one contextual condition and negative in another. Hence, to facilitate provision of high quality context-aware personalized services, the above two-dimensional (user- and item-dependent) representation *Exp* should be extended by a third general dimension, reflecting various contextual conditions and features considered by the personalization system.

The context-awareness issue has lead to a multidimensional warehousing of the UMs that captures the dependencies between the ratings and a generalized user-, item- and context-dependent model [1]. This model forces the representation of experiences by the two-variable function *Exp*, ignoring the context-awareness issue, to be modified, to a three-variable function Exp_{CA}, incorporating a third dimension of context. Given the above generalization of ratings to the experiences *evaluation*, context-aware experience is defined by:

$$Exp_{CA}: User_{feat} \times Item_{feat} \times Context_{feat} \rightarrow evaluation.$$

This representation, in addition to the standard $User_{feat}$ and $Item_{feat}$ features, also includes $Context_{feat}$ that represents the contextual conditions (or the values of the contextual features) of the experience. Similarly to $User_{feat}$ and $Item_{feat}$, $Context_{feat}$ is also described using a multidimensional representation by a set of features. Hence, a specific contextual condition of the experiences is referred to as a subspace of this multidimensional contextual space. For instance, in the above mentioned example, only one contextual feature of *companion* out of a large $Context_{feat}$ set of contextual features is mentioned. When the *companion* feature is assigned the value of *friends*, then the evaluation is positive, and when it has the value *family*, the evaluation is negative.

It should be stressed that modifying the representation of *Exp* function to Exp_{CA}, i.e., incorporating the contextual features, does not have any effect on UMs warehousing. Exp_{CA} is still referred to as a collection of past experiences, whereas the experiences are now context-aware. This representation of context-aware experience, Exp_{CA}, is illustrated in Figure 1.

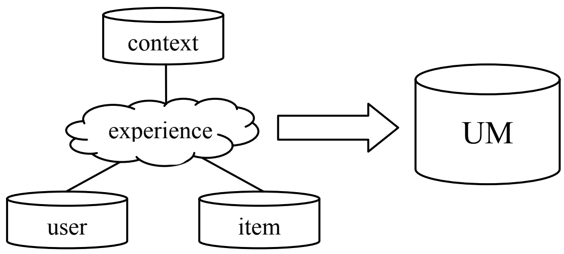

Fig. 1. Representation of Context-Aware Experiences and User Model

Note that although the definition and representation of experience was modified to capture the context-awareness issue, the definition of the personalization task remains unchanged. This still consists of predicting the evaluations of the new experiences based on a set of past experiences and their evaluations. However, since the experience is now modified to include the $context_{feat}$, the personalization should be provided in a context-aware manner, i.e., they should refer to a certain combination of $user_{feat}$, $item_{feat}$ and $context_{feat}$ values, and not of $user_{feat}$ and $item_{feat}$ values only.

3 Mediation of User Models for Context-Aware Personalization

The main problem in providing high quality context-aware personalization using the context-aware Exp_{CA} representation for UMs warehousing is the sparseness of data in the UM, i.e., the lack of sufficient user modeling data about the user in certain contextual conditions [9]. The problem of data sparseness is a well-known problem of traditional personalization systems, which rely on only the two-dimensional representation R'_{gen} and ignore the contextual dimensions [33]. This problem is even harder when the contextual information is considered, as the initially sparse two-dimensional experiences are 'sliced' among multiple contexts, reflecting the specific contextual conditions of the experiences. Hence, the amount of user modeling data referring to a specific contextual condition significantly decreases when the context-awareness issue is taken into account.

This work proposes to overcome the sparseness problem by using a *mediation* of UMs, which is the primary focus of this work. Hence, the definition of UMs mediation can be formulated as follows: "*mediation of UMs is a process of importing the user modeling data collected by other (remote) personalization systems, aggregating them and generating an integrated user model for a specific goal within a specific context*". In this definition, the term aggregation implies resolving the heterogeneities and inconsistencies in the obtained data. The primary goal of the mediation is to instantiate the sparse context-aware UMs through inferring the required user modeling data from past experiences and evaluations in a three-dimensional context-aware representation space. Hence, it enriches the existing UMs (or bootstraps empty UMs) using the user modeling data collected by the remote systems and facilitates the provision of better context-aware personalized services.

Two parties are involved in the mediation process. On the one hand, there is a target personalization system, i.e., the system that was requested to provide personalized services to the user. Formally, this system is the initiator of the mediation process, requesting the available user modeling data. On the other hand, there are numerous remote personalization systems that may provide valuable user modeling data (i.e., past experiences) to the target personalization system.

The two parties are interconnected via the UMs mediator, which, in fact, is the core element of the mediation process. As mentioned earlier, the main difficulty of the UMs mediation is the heterogeneity of the user modeling data. For example, personalization systems from different application domains imply different user modeling data stored in the UMs. Even within the same application domain, different systems may store different information in their partial UM, according to the specific personalization technique being exploited (e.g., ratings vector in collaborative filtering UMs [26] vs. a feature vector of interest topics in content-based UMs [40]). Moreover, even the UMs of two systems from the same application domain exploiting the same personalization technique may use different terms to describe equivalent semantic concepts. Thus, successful mediation of UMs requires developing and applying reasoning and inference mechanisms using semantically-enhanced knowledge bases, which allow commonalities between the user modeling data representation of different systems to be identified. Hence, the mediator consists of two principal components:

- **Aggregation Mechanism.** This component is responsible for resolving conflicts and heterogeneities in the obtained user modeling data using various reasoning and inference mechanisms. The obtained past experiences may be represented in different ways, e.g., using different ontologies, domain- and application-specific terminology, or even in different languages. Hence, these heterogeneities need to be resolved and the data need to be converted to a standardized form. In addition, the evaluations of the same experience in different systems may be contradictory. This requires the aggregation mechanism to implement and apply certain policies for conflicts resolution in the obtained data[1].
- **Knowledge Base.** This component is in fact an auxiliary component, required by the aggregation mechanism. It contains a number of semantically-enhanced inter- and intra-domain knowledge bases representing dependencies and relationships between various user, item and context features. The information stored in the knowledge bases actually facilitates the resolution of the heterogeneities in the obtained user modeling data. For example, it allows different ontologies exploited by various personalization systems to be reconciled, and the terms used by a certain system to be converted to a standard representation, and may even provide means for translation, resolving cross-lingual dependencies.

Although the mediation process is formally initiated by the target personalization system, practically it is initiated by the user issuing a personalized service request to the target system. The envisioned scenario of the user modeling data mediation process consists of the following stages (as illustrated in Figure 2):

1. The user's personalization request is treated by the target system as a request for a prediction of the user's evaluation of the new experience, i.e., for the specific combination of $user_{feat}$, $item_{feat}$, and $context_{feat}$. To provide a better service, the target personalization system queries the mediator for the UMs, containing past experiences valuable for predicting the evaluation of the new experience. The query contains the required $(user_{feat}, item_{feat}, context_{feat})$ combination.
2. The mediator analyzes the received $(user_{feat}, item_{feat}, context_{feat})$ combination and determines the set of remote personalization systems, which may store potentially relevant past experiences. In fact, this analysis is performed using the semantic data provided by the knowledge base (the criteria for selecting the remote systems and specific past experiences will be detailed later in this section, where specific mediation approaches will be discussed).
3. The mediator forwards the query to the set of remote personalization systems that was determined in the previous stage.

[1] It is reasonable to assume that different personalization systems will provide user modeling data with different levels of accuracy, and up-to-date information. Although the importance of resolving such conflicts and inconsistencies is highlighted in the paper, developing conflict resolution policies is beyond the scope of the current work.

4. Remote personalization systems, which actually store the relevant experiences, respond to the query and send to the mediator their locally collected UMs and/or the relevant experiences only.
5. The aggregation mechanism of the mediator aggregates the obtained experiences using the semantic data provided by the knowledge base. Clearly, different combinations of UMs representation in remote and target systems will imply exploiting different aggregation mechanisms (several aggregation mechanisms will be described in the following section, where several mediation approaches that have already been implemented and evaluated will be presented).
6. The generated user modeling data are sent to the target personalization system (several types of this user modeling data will be discussed later in this section). Since the user modeling data of the target system is enriched in comparison to the locally collected data stored before the mediation, the system is capable of providing better personalized services to the user.

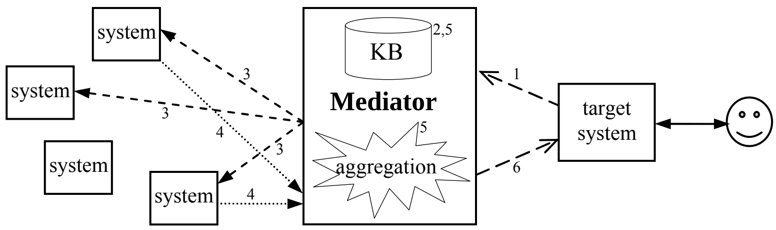

Fig. 2. Architecture and Stages of UMs Mediation

In principle, any available user modeling data (i.e., any past experience) may be relevant to some extent as an input to the mediation process, since they may help predicting the evaluations of the new experience. For example, consider a target personalization system that is supposed to predict the new experience evaluation for a specific combination of $(user_{feat}, item_{feat}, context_{feat})$ values. The possible groups of $(user_{feat}, item_{feat}, context_{feat})$ combinations in past experiences stored by other personalization systems are as follows (the respective mediation methods will be elaborately discussed later in this section):

- Experiences where the values of all three features refer to the same concept. These experiences refer to the same combination $(user_{feat}, item_{feat}, context_{feat})$. Hence, they represent past experiences of the same $user_{feat}$ for the same $item_{feat}$ in the same $context_{feat}$, where the values of certain experience features, actually referring to the same concept, may be represented in different ways.
- Experiences where the values of two features are the same and the value of one feature differs. Three possible combinations are possible:
  - $(user'_{feat}, item_{feat}, context_{feat})$ – past experiences of another $user_{feat}$ for the same $item_{feat}$ in the same $context_{feat}$.
  - $(user_{feat}, item'_{feat}, context_{feat})$ – past experiences of the same $user_{feat}$ for another $item_{feat}$ in the same $context_{feat}$.
  - $(user_{feat}, item_{feat}, context'_{feat})$ – past experiences of the same $user_{feat}$ for the same $item_{feat}$ in another $context_{feat}$.
- Experiences where the values of one feature are the same and the values of two features differ. Three possible combinations are possible:

- $(user_{feat}, item'_{feat}, context'_{feat})$ – past experiences of the same $user_{feat}$ for another $item_{feat}$ in another $context_{feat}$.
- $(user'_{feat}, item_{feat}, context'_{feat})$ – past experiences of another $user_{feat}$ for the same $item_{feat}$ in another $context_{feat}$.
- $(user'_{feat}, item'_{feat}, context_{feat})$ – past experiences of another $user_{feat}$ for another $item_{feat}$ in the same $context_{feat}$.

- Experiences where the values of all three features are different. These experiences refer to $(user'_{feat}, item'_{feat}, context'_{feat})$ and represent past experiences of another $user_{feat}$ for another $item_{feat}$ in another $context_{feat}$.

Clearly, the first group of experiences is the most valuable for UMs mediation, as it provides past evaluations of the target user for the required item in the relevant context. However, such experiences require aggregation mechanisms for resolving possible heterogeneities in the representations of feature values or experience evaluations to be applied. The second group of experiences (with only one feature different from the required combination) is also valuable for mediating the required user modeling data. Such experiences represent past evaluations, where the values of two out of three features match the values of the features in the new experience. Mediation of user modeling data in this case requires applying inference mechanisms, which are capable of identifying the relationships between the different values of the one feature that differs and projecting the available evaluations to the new experience.

In the third group of experiences, the values of two features out of three features are different, and only one feature matches the values of a feature in the new experience. Hence, mediation of such experiences requires applying more complicated inference mechanisms (e.g., several consecutive inferences similar to the inferences from the previous group of experiences, where the value of only one of the features was different). Although this user modeling information may be valuable and may enrich the user modeling data in the target system, applying complicated inference mechanisms may 'deteriorate' the original data represented by the past experiences. Therefore, such experiences are currently not used in the UMs mediation. Obviously, the situation is even worse for the fourth group of experiences, where the values of all three features are different, and the mediation requires three inference mechanisms to be applied. Hence, these experiences are also not considered for mediation at the moment.

In summary, two groups of experiences should be considered as input for the mediation process: (1) experiences having the required values of all three features, and (2) experiences having the required values of two features and a different value of one feature. This yields four types of UMs mediation over the context-aware three-dimensional representation of experiences. The first type of mediation is conducted between experiences having the required values of all three features, i.e., between heterogeneous representations of the same experience. Such mediation is referred to later as **cross-representation mediation**. The other three mediation types are so-called **cross-dimension mediations**. These mediations are conducted over the experiences having the required values of two features and a different value of one feature. This means that the values of two out of three dimensions in the space are fixed and the mediation is performed across a different value of the third dimension. Hence, three types of mediation are possible: (1) **cross-user mediation**, where the values of item and context features are fixed and the user in the experiences is item and context features are fixed and the user in the experiences is allowed to be modified; (2) **cross-item mediation**, where the values of user and context features are fixed and the item in the experiences is allowed to be modified; and (3) **cross-context mediation**, where the values of user and item features are fixed and the context in the experiences is allowed to be modified. In the rest of this section, these four possible types of UMs mediation are elaborately dicsussed.

3.1 Cross-Representation Mediation

This mediation is aimed at resolving the heterogeneity in the representations of the experiences of the same user for the same item in the same context. In other words, it incorporates past experiences of the same user for the same item in the same context, but expressed in different ways. For example, consider the following representation of the same item, a movie "*Gone with the Wind*", in two datasets: EachMovie [36] and MovieLens [26]. In EachMovie, it is classified as a *classic* movie, while in MovieLens it is classified as a *drama, romance* and *war* movie. In addition, a movie evaluation in EachMovie is a number between *0* and *1*, while in MovieLens it is expressed by a number of stars on a *5*-star scale. To implement mediation between these two systems, the mediator should be able to cope with such heterogeneities.

Hence, cross-representation mediation can be considered as an aggregation and inference of user modeling data between heterogeneous representations of the values of the experience features. This means that, although different representations of the features refer to the same (semantic) concept, they are expressed in different (syntactic) ways. Hence, the mediation is conducted between past experiences, where the values of all the experience components $user_{feat}$, $item_{feat}$ and $context_{feat}$ refer to the same concept, although represented in different ways. This mediation is divided into two groups:

a. Different representation of $user_{feat}$, $item_{feat}$ and $context_{feat}$ values. This mediation deals with a situation where the representation of one (or several) of the experience components is heterogeneous. This means that although the $user_{feat}$, $item_{feat}$ and $context_{feat}$ are semantically identical and reflect the same user modeling data, one (or several) of them is (are) syntactically expressed in different ways. For example, collaborative filtering systems represent an item using its unique identifier only, while content-based systems represent the same item using the set of its features. This mediation requires applying inference mechanisms identifying commonalities and dependencies between various representations of semantically identical $user_{feat}$, $item_{feat}$, or $context_{feat}$ using an external domain-specific knowledge base. Hence, this particular case of cross-representation mediation is referred to as *cross-technique mediation* [7].
b. Different representations of the evaluation values. In this mediation, the representations and the values of $user_{feat}$, $item_{feat}$ and $context_{feat}$ features are identical, but the evaluations of the experiences are expressed in different ways, i.e., the heterogeneity is in the representation of the evaluations. For example, the target personalization system represents the evaluation as a discrete numeric rating on a scale between *1* and *10*. However, the remote system represents it as positive or

negative evaluation only. In this case, there is a need to map and reconcile the evaluation representation values between the two scales.

3.2 Cross-User Mediation

Although the term mediation is not explicitly mentioned in collaborative filtering recommender systems, cross-user mediation and inference actually constitute the basis of this popular personalization technology [26]. In fact, collaborative filtering is based on the assumption that people with similar tastes (i.e., people who agreed in the past) will prefer similar items (i.e., will agree in the future) [52]. Or, in a simplistic view, collaborative filtering recommends items liked by similar users[2]. In order to generate a recommendation, collaborative filtering systems initially create a neighborhood of users with the highest similarity to the active user, and then generate a prediction by aggregating the ratings of these users. Hence, this process can be considered as a cross-user inference, or a particular case of cross-user mediation, where the mediated user modeling data are the ratings of similar users. In addition, other variants of cross-user inference are applied in several existing personalization techniques, e.g., demographic filtering [31] and some hybrid recommenders systems, such as [53].

It should be stressed that the existing implementations of cross-user mediation in the state-of-the-art personalization systems mostly disregard the context-awareness issue. This means, they project the collected experiences to the two-dimensional representation R'_{gen}, not reflecting the contextual conditions of the experience. Hence, these personalization systems apply inference mechanisms assuming that the collected experiences were recorded for the same contextual conditions. Thus, the prediction of the new experience evaluation can be considered as an inference process incorporating past experiences of other users for the same item in the same (in fact, undefined) context, i.e., the prediction generation process is pure cross-user inference.

3.3 Cross-Item and Cross-Domain Mediation

Cross-item mediation is also applied in various existing personalization techniques, such as content-based filtering [40], item-to-item collaborative filtering [50], utility-based recommendations [35] and some hybrid recommender systems, such as [43]. In general, these personalization techniques assume that also the similarity of items may be used for providing personalized services, i.e., items which are similar to the items the users liked in the past should be recommended to the users (in fact, most cross-user similarity metrics discussed in [26] may be also applied for computing cross-item similarity).

Also in these practical systems the context-awareness issue is mostly disregarded, and the collected experiences are represented using the two-dimensional representation R'_{gen}, such that the collected experiences are considered as if recorded in the

[2] The reader is referred to [26] for a discussion on collaborative filtering similarity metrics.

same contextual conditions. Thus, the prediction of the new experience evaluation can be considered as the result of an inference process, incorporating past experiences of the same user for other items in the same (in fact, undefined) context. This means that the prediction generation process is performed through cross-item inference.

However, cross-item mediation requires cross-item similarity and/or relationships, which cannot be easily defined for any arbitrary pair of items, to be defined explicitly. For example, in item-to-item collaborative filtering [50], cross-item similarity is computed by means of user ratings to items. Hence, computation of the similarity between two items in this case requires the items to be rated by a non-empty set of overlapping users. This requirement may be too restricting for sparse ratings matrices. Moreover, in many conditions, the available past experiences do not necessarily reflect the user's evaluation for an individual item, but rather on a generalized group (or category) of items. For example, a user may express his opinion not on a single movie, but on a genre of movies, or on movies directed by a certain director.

Hence, in a broader view, the items need to be aggregated and grouped, which allows applying a more complex type of mediation, incorporating the evaluations of past experiences for a generalized group of items [38]. Generalizing individual items into groups and domains and then exploiting cross-domain dependencies and inferences introduces the issue of **cross-domain mediation**, where the evaluation of the new experience for a generalized group of items from a certain domain is inferred from past experiences for items in other domains [11]. In this sense, cross-domain mediation can be considered as a mediation incorporating past experiences of the same user in the same contextual conditions, but for another generalized group of items.

3.4 Cross-Context Mediation

The issue of cross-context mediation is a new research direction in user modeling. Such mediation is based on context-aware representation of the experiences, and its goal is to predict the evaluations of the new experiences in a given context using past experiences in other contextual conditions [9]. This means that cross-context mediation incorporates past experiences of the same user for the same item in other contextual conditions. For example, it can predict future evaluation for an item by a user in the *evening* given past evaluation of the same user for the same item in the *morning*, or, it can predict future evaluation for an item by a user when accompanied by a group of friends given past evaluations of the same user for the same item when accompanied by a parent.

Since the state-of-the-art personalization systems mostly disregard the context-awareness issue and are not capable of providing context-aware personalization, this type of mediation requires the definition of various novel cross-context reasoning mechanisms. Two simple mechanisms, exploiting semantically enhanced RDF/OWL [45] representations of $user_{feat}$, $item_{feat}$ and $context_{feat}$, were proposed in [9]:

- **Rule-Based Reasoning.** This reasoning mechanism exploits the semantically-enhanced representations of the experience components for the purpose of defining a set of rules that explicitly provide the relationships between the values of the features. For example, consider a rule defining that user's preferences regarding a certain item (e.g., stocks news) in the *evening* are the opposite to preferences in the *morning*. Or, another rule defines a projection of user's preferences at *4PM* to a more general *afternoon* time period. Applying these rules facilitates the inference of the required user modeling data across various contextual conditions.

- **Similarity-Based Reasoning**. This reasoning mechanism exploits the semantically-enhanced representations of the experience components for the purposes of defining an explicit similarity metric for the contextual conditions. For example, such a metric may computationally express similarity between Tuesday and Wednesday as mid-week days and dissimilarity between Tuesday and Sunday as mid-week and week-end days. Such a cross-context similarity metric allows various adaptation rules to be derived, similar to the rules used in Case-Based Reasoning [49], that facilitate reuse of the evaluations of past experiences. For example, this can be done by a collaborative-like aggregation of the evaluations of past experiences of the same user for the same item in similar contextual conditions.

Comparing the above discussed rule-based and similarity-based reasoning approaches shows that, on the one hand, rule-based reasoning may produce more accurate user modeling data, as the reasoning rules are typically defined by domain experts. On the other hand, defining and updating the inference rules in today's highly dynamic information world may hamper the scalability of the mediation process. Conversely, the typical scenario for similarity-based reasoning is fully autonomous and therefore gives a more flexible mediation process. However, similarity-based reasoning requires a large number of past experiences to bootstrap the reasoning process. Other machine learning approaches can be considered for this purpose.

4 Sample Experimental Results

This section presents and discusses experimental evaluations of two practical mediation implementations: **cross-technique mediation** [7], as a particular case of cross-representation mediation and **cross-domain mediation** [11], as a generalized form of cross-item mediation. The issue of cross-user mediation was studied intensively in many prior works on collaborative filtering [26]. It became one of the most widely used personalization techniques applied in a wide variety of domains and E-Commerce applications [51]. Collaborative filtering has been implemented and evaluated many times and proves to be an effective and relatively accurate personalization approach [27]. Hence, cross-user collaborative filtering is left out of the scope of this work. Conversely, the research of cross-context mediation is currently in its initial stages only [9], and will be discussed it in the future research section.

The main difficulty involved in conducting an experimental evaluation of UMs mediation is the lack of publicly available data, representing the user modeling data of the same users across several different application domains, containing heterogeneous representations of the experiences, or reflecting different contextual conditions. Currently available datasets are mostly collaborative filtering data, representing users' ratings to items from three application domains: movies (MovieLens [26] and EachMovie [36]), books (BookCrossing [56]), and jokes (Jester [21]). However, these datasets are not cross-linked, i.e., no users from one dataset can be identified in another dataset. Moreover, no datasets with data different from collaborative filtering UMs are publicly available. A limited dataset of context-aware collaborative filtering UMs was reported in [1], but no other context-aware datasets are publicly available.

Hence, in the experimental evaluation of the proposed mediation methods only the EachMovie dataset of movie ratings was used [36]. Cross-technique mediation (as a particular case of cross-representation mediation) mimicked a situation where a content-based recommender system is the target system and a collaborative filtering recommender system is the remote system. Hence, the collaborative filtering UMs stored in EachMovie dataset served as the input user modeling data for the mediation process, and it was converted to content-based data [7]. Cross-domain mediation (as a generalization of cross-item mediation) mimicked domain-related partitioning of UMs. For this, collaborative filtering UMs stored in EachMovie dataset were partitioned into different domains. As all the items in EachMovie dataset actually belong to the domain of movies, domain-related partitioning was assigned by artificially partitioning the movies into the domains according to the genres of the movies [11].

Note that both mediation methods require exploiting semantically-enhanced domain knowledge bases. For example, cross-technique mediation is aimed at generating content-based user modeling data. Hence, semantic information about the movies is required for extracting and mining the features of the movies rated in the collaborative filtering UMs. Conversely, in cross-domain mediation, semantic information about the movies is required for partitioning the movies according to their genres and also for devising inter-domain distances, which will be explained later in this section. Since both mediation methods exploited the data from EachMovie dataset of movie ratings [36], their implementations used movie information provided by IMDb (The Internet Movie Database) [28]. The details of these mediation methods are extensively discussed in the following subsections.

4.1 Cross-Technique Mediation

Cross-technique mediation deals with the heterogeneity of the experience representations in different personalization techniques. In [7], the paper focuses on cross-technique mediation of user modeling data, where the heterogeneity was encapsulated in the representations of the experience evaluations. The source of the user modeling data (i.e., the remote system) was a collaborative filtering [26] recommender system, whose user modeling information was mimicked by EachMovie dataset. In collaborative filtering systems the UMs are represented by a ratings matrix, containing the users' vectors of ratings to the items stored by the system. These data were mediated to a content-based recommender system [40], where the UMs are represented by a list of item's features and user's preferences on these features. Hence, the mediation was conducted between two recommender systems within the same domain of movies, using the movies metadata mined from the IMBd database [28]. The goal of the experimental evaluation was to determine the conditions in which cross-technique mediation improves the quality of the generated recommendations.

For example, consider the following set of collaborative filtering experiences of a user UM_{CF}*={ "The Lord of The Rings":1, "The Matrix":0.8, "Psycho":0.2, "Friday the 13th":0, "Star Wars":0.9, "The Nightmare":0.1, "Alien":0.9 }*, where the evaluations are given on a continuous scale between *0* and *1*. It can be easily recognized that the user strongly likes science-fiction movies, and strongly dislikes horror movies. Thus, the converted content-based UM of the user may be UM_{CB}*={ science-fiction:0.9, horror:0.1 }*, where the weights of the genres are computed as an average of the ratings given to the movies in this genre. Similarly, also the weights of other movie features, such as directors, actors and other can be computed.

However, applying such mediation requires external knowledge that, for instance, "*The Matrix*" is a *science-fiction* movie and "*Psycho*" is a *horror* movie. This information was mined from the IMDb. In general, the IMDb provides information in *49* categories of features, such as *genre*, *actors*, *directors*, *writers*, *composers*, *keywords*, *distributors*, *languages*, and many others [28]. For the sake of simplicity, only seven categories of features were initially used: *genres*, *keywords*, *actors*, *actresses*, *directors*, *production countries* and *languages*, as these categories seem to have the most influence on the user's evaluation for a movie. Later on, a wrapper feature selection algorithm [30] was applied to identify the set of categories that should be used by the context-based prediction mechanism in order to achieve the highest accuracy of the generated predictions. As a result, only five categories of features were finally used: *genres*, *keywords*, *actors*, *actresses*, and *directors*.

The mediation process converted the collaborative filtering user modeling data represented by the ratings vectors into content-based data represented by weighted vectors of features. Implicitly, the mediation was based on the assumption that users' collaborative filtering ratings implicitly reflect their preferences with respect to the content features of the movies, e.g., preferred movie genre, beloved topic, or favorite actors and/or directors. Hence, the mediation task could be described as a process of identifying and learning commonalities in the features of positively and negatively rated movies and generalizing these commonalities into a weighted list of features deriving the required weights of the features basing on these commonalities.

The mediation of collaborative data to content-based data received the user's ratings vector as an input. Since different users might express their ratings in different ways (e.g., rating *4*, provided by a user whose average rating is *2* should be treated differently from rating *4* provided by a user whose average is *3.5*), users' ratings were normalized by subtracting the average rating of the user from the provided ratings. For each movie rating in a collaborative UM of the user, a list of a movie features (from the above five categories only) was extracted from the IMDb. The weights of these features in the content-based UM of the user were updated according to the normalized rating of the movie. In other words, the normalized rating of the movie was assigned to the features of the movie: genres, actors and directors involved in the movie, movie languages, and so on. For example, consider the rating *0.9*, given by a user whose average rating is *0.6* to a movie *"Star Wars"*. According to the IMDb, the genres of *"Star Wars"* were *action*, *adventure*, *fantasy* and *science-fiction*. Thus, the existing weights of these four features were increased by *0.9-0.6=0.3*. Similarly, the weights of the movie director, of all the actors involved in the movie and the rest of the movies features were increased by *0.3*.

The generated content-based UMs served as a basis for generating content-based recommendations. To evaluate the effect of the mediation, the accuracy of the collaborative recommendations built using the original collaborative filtering user modeling data was compared with the accuracy of the content-based recommendations built using the generated content-based data. The goal of this particular experimental evaluation was to assess the impact of mediation as a function of the sparseness of the original collaborative filtering user modeling data.

Hence, all the available users in the dataset were partitioned into *12* groups, according to the number of rated movies in their collaborative filtering UMs: users who rated below *25* movies, users who rated between *26* and *50* movies, and so on, up to the users who rated between *301* and *500* movies. Collaborative filtering data of each user was partitioned into *90%* used for the training set and *10%* for the testing set. In other words, the training set of *90%* of ratings served as a basis for UM mediation, while the recommendations were built on the testing set of *10%* of ratings. In summary, two types of recommendations, collaborative and content-based, were generated, and their accuracy was measured using the well-known MAE metric [27]:

$$MAE = \frac{\sum_{i=1}^{N} | p_i - r_i |}{N},$$

where N denotes the total number of generated recommendations, p_i is the predicted rating, and r_i is the real rating to the item number i. Figure 3 shows the MAE values. The horizontal axis reflects the number of rated movies in the collaborative filtering UMs, while the vertical axis stands for the MAE.

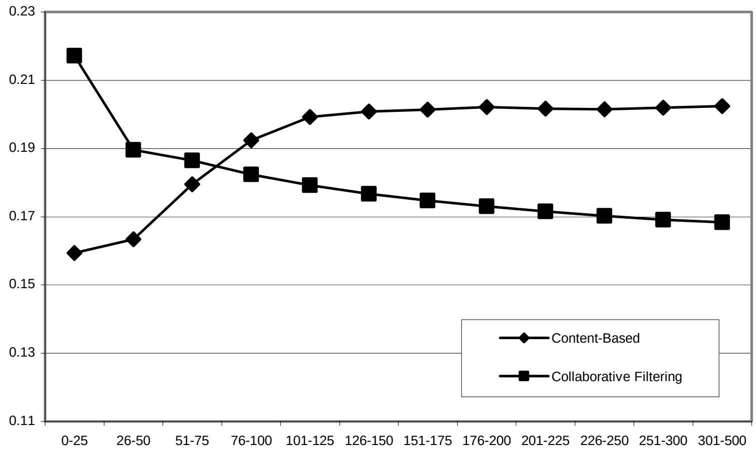

Fig. 3. Cross-Technique Mediation of User Models

The chart shows that the MAE of the content-based predictions built over content-based UMs, generated from collaborative UMs containing the ratings of less than *50* movies, is relatively low (approximately *0.16*). This is explained by the observation that for a low number of rated movies in the collaborative filtering UM, it is easy to find the important features for content-based prediction, i.e., the features that are strongly liked or disliked by the user. For larger UMs, between *50* and *150* rated movies, the MAE increases with the number of rated movies. This happens due to a larger number of content features that occur in the rated movies. Hence, it is harder to identify which features are important for accurate predictions and which are not, such that some of the features behave as noisy features, and hamper the accuracy of the generated predictions. Finally, for the collaborative UMs with over *150* rated movies, the MAE converges.

A comparison of the content-based and the collaborative filtering techniques shows that when the original collaborative filtering UMs contain less than *75* rated movies, the accuracy of pure content-based recommendations based on the generated (artificial) content-based user modeling data outperforms the accuracy of collaborative filtering recommendations based on the original collaborative filtering data.

In addition to this experiment, the distribution of users over the groups of numbers of rated movies was computed (i.e., what percentage of users belongs to each one of the above *12* groups of users, partitioned according to the number of movies rated by each user). The distribution shows that *78.4%* of the users rated less than *75* movies. Thus, improving the accuracy of the recommendations for the users with such a magnitude of rated movies is especially important. Since the accuracy of the collaborative filtering recommendations for this amount of user modeling data is quite low, mediation of UMs and further pure content-based recommendations provide a solid alternative. For a more comprehensive discussion of cross-technique mediation, fine-tuning of the mediation and other experimental results, the reader is referred to [7].

One important observation regarding the proposed approach should be stressed. The content-based recommendations are generated solely based on content-based user modeling data, derived from collaborative filtering data. As such, the recommendation mechanism is capable of generating content-based predictions regardless of the number of available movie ratings. Hence, this approach resolves the well-known *first-rater problem* [40] of collaborative filtering, where an item cannot be recommended unless it has already been rated by a sufficient number of users. However, being a pure content-based recommendation mechanism, it suffers from an inherent *serendipity problem* [26], i.e., it can recommend only movies that are similar to the movies that have already been rated by the user and cannot produce 'surprising' recommendations. Hence, while resolving one problem, the proposed approach introduces another one.

4.2 Cross-Domain Mediation

In cross-domain mediation, the user modeling data are imported from remote recommender systems providing recommendations for items from other application domains. This specific implementation of cross-domain mediation referred to a situation where both the remote and the target systems were collaborative filtering recommender systems [11]. Hence, all the systems involved in the user modeling data mediation represented their UMs using the same representation, i.e., as a ratings vector: a list of ratings provided by users to the domain items. In addition, it is assumed that a user can be identified in all the systems by the same unique identifier.

In this setting, an item can theoretically belong to several domains, and therefore three types of user modeling data could be imported and aggregated. These types of data, in fact, reflect the stages of prediction generation in collaborative filtering: similarity computation, neighborhood formation, and prediction generation [26]. For the similarity computation stage, the UMs (i.e., ratings vectors) were imported from the remote recommender systems. For the neighborhood formation stage, the list of nearest-neighbors, jointly with their similarities computed by the remote systems, was imported. Finally, for the prediction generation stage, complete predictions generated by the remote systems were imported. Now, the mediation process for each of the above three types of user modeling data being mediated will be briefly discussed:

a. UMs stored by the remote systems
   A collaborative filtering UM was represented by a ratings vector, i.e., a list of ratings provided by the user to some items managed by the system. In order to enrich the UMs stored by the target system, remote systems sent to the target system their local ratings vectors of the users. Upon receiving the ratings vectors from the remote systems (and given a policy for resolving conflicts in the ratings of the same user to the same item coming from different systems), the target recommender system constructed a unifying ratings matrix. This matrix was generated by integrating locally stored ratings vectors with the vectors imported from the remote systems. Then, a traditional collaborative filtering mechanism was applied over the unifying ratings matrix. Since the unifying ratings matrix can be considered as equivalent to a centralized ratings matrix (if the UMs were not partitioned across various recommender systems), this approach was referred to in the experimental results as *standard* collaborative filtering.
b. Lists of the nearest neighbors and degrees of their similarity
   Since the remote systems were also collaborative filtering recommender systems, they could autonomously compute the similarity between the active user and the other users based on their locally collected ratings vectors. A set of users with the highest similarity values was considered as the list of nearest neighbors in the remote system. Then, each remote system provided its list of nearest neighbors (and the similarity values of the neighbors) to the target system. Upon receiving the local sets of nearest neighbors, the target system computed the overall similarity between the active user and the other users as a weighted (according to inter-domain distances[3]) average of local similarity values in the remote systems. Hence, this improved the accuracy of the similarity computation, as it was not based only on the user modeling data collected by the target system, but also on the similarity values obtained from the target systems from other domains (i.e., genres). Then, standard collaborative filtering recommendations were generated using the ratings (collected by the target system) of the users with highest overall similarity. This approach is referred to as *cross-genre* collaborative filtering.
c. Complete predictions generated by the remote systems

---

[3] Inter-domain distances were computed in two ways: (1) according to the similarity of the domain contents (e.g., mined from IMDb), and (2) according to the ratings provided by the users to the domain items. The reader is referred to [11] for a detailed description of the inter-domain distances and users' overall similarity computation.

In domain-related partitioning an item might belong to several application domains. As a result, the ratings to the item could be stored by several recommender systems, and the predictions could be generated by any of these systems. Hence, in this case, the remote recommender systems actually storing the ratings to the required item, could generate their local predictions (i.e., predictions based on the locally collected user modeling data), and send them to the target system. Upon receiving the predictions from the remote systems, the target system can aggregate them into a single recommendation value. Since in this case, the aggregation was done by averaging the values of the predictions received from the target systems, this approach is referred to as *remote-average* collaborative filtering. Alternatively, the predictions could be generated in a weighted average using the inter-domain distances.

Experimental evaluation of the proposed mediation approaches involved EachMovie dataset of movie ratings [36]. Since the ratings in EachMovie belong to the same application domains, the existing data was artificially partitioned into different domains according to the movie genres (mined from the IMDb database of movies metadata [28]), instead of collecting ratings from different application domains. This mimicked domain-related partitioning of collaborative filtering ratings, where each domain is represented by a certain movie genre and by the appropriate collaborative filtering recommender system.

The goal of this particular experimental evaluation was to assess the impact of different mediation types as a function of the sparseness of the user modeling data. Hence, the available users in the genre-related datasets were partitioned into *12* groups, according to the percentage of movies rated in the relevant genre: users who rated below *3%* of movies, users who rated between *3%* and *6%* movies, and so on, up to users who rated over *33%* of the genre movies. For every group, *1,000* predictions were generated for various combinations of the target user, movie to be predicted, and the target genre of the movie (since every movie typically belonged to several genres). Also in this experiment, the accuracy of the predictions for every group of users was measured using the MAE metric [27]. Figure 4 shows the experimental results: the horizontal axis stands for the percentage of rated movies in the target genre, and the vertical axis for the MAE.

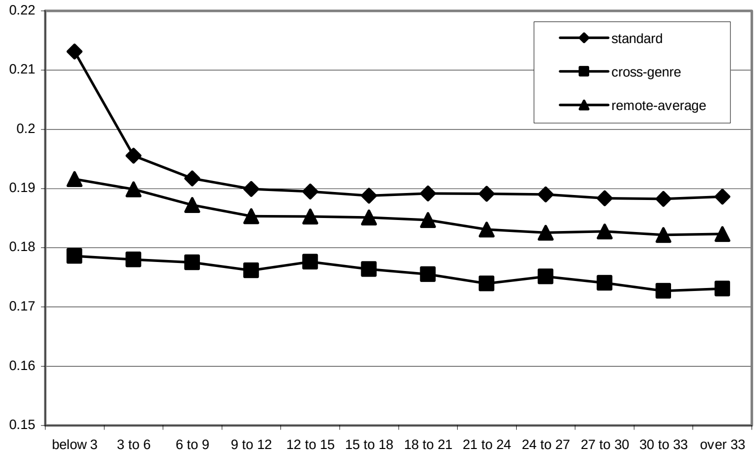

Fig. 4. Cross-Domain Mediation of Collaborative Filtering User Models

The baseline for the comparisons is *standard* collaborative filtering, as its accuracy is similar to the accuracy that would have been obtained in traditional centralized collaborative filtering. It can be clearly seen that for any percentage of movies rated in the target domain, both *cross-genre* and *remote-average* approaches outperform *standard* approach (both are statistically significant, *p=0.00058* and *p=1.63E-06*, respectively).

The results of the *cross-genre* approach can be explained by the observation that its weighted similarity computation is more accurate than *standard* similarity computation. The better accuracy is achieved due to the fact that the *cross-genre* approach aggregates users' domain-related similarities according to the inter-domain distances, while the *standard* approach disregards the inter-domain distance. Hence, also the predictions generated by the *cross-genre* approach are more accurate. The better accuracy of *remote-average* predictions can be explained by arguing that the similarity computation over the ratings from several highly relevant domains (to which the movie belongs) yields more accurate similarity values than *standard* similarity computation, performed over all the available ratings. This can be explained by the observation that the ratings from highly relevant domains are important for computing the similarity value in the target domain, whereas the other ratings insert noise into the computation. Hence, *remote-average* predictions are more accurate that *standard* predictions.

A comparison of *cross-genre* and *remote-average* approaches shows that the *remote-average* approach is more accurate for any percentage of rated movies (statistically significant, *p=4.13E-10*). However, in certain conditions the *remote-average* approach may be inapplicable. For example, for sparse user modeling data collected by the remote system, i.e., when a remote system does not have a sufficient number of ratings, the users' similarity cannot be reliably identified and accurate predictions cannot be generated. In this case, *cross-genre* approach should be applied, as its accuracy still outperforms the accuracy of *standard* collaborative filtering. For a more comprehensive discussion of cross-domain mediation, inter-domain distances computation and other experimental results, the reader is referred to [11].

In summary, these experimental results (both of cross-technique and cross-domain mediation) allow drawing a conclusion that cross-domain mediation of the user modeling data is beneficial. It allows enriching the user modeling data of the target system and therefore improving the accuracy of the generated predictions.

5 Conclusions and Future Research

In this work, a generic framework for the mediation of user modeling data, which can be applied in personalization systems, was presented. The mediation is aimed at resolving the data sparseness issue, which is one of the main problems of state-of-the-art personalization systems, and, as a result, at improving the quality of the personalized services provided to the users. The analysis started with discussing the representation of UMs in various personalization techniques. Then, the experience was defined as a fundamental unit of the user modeling data representation, encapsulating the representations of user, item and context, and a generic representation of the UMs was devised. Then, several mediation methods, using different types of past experiences, were discussed. In particular, four mediation types allowing user's evaluation of the new experiences to be predicted based on their evaluation of past experiences were derived: cross-user mediation, cross-item mediation, cross-context mediation, and two types of cross-representation mediation.

The paper presented sample experimental results, demonstrating two mediation techniques: cross-technique (as a particular case of cross-representation) mediation and cross-domain (as a generalized form of cross-item) mediation. Both evaluations showed that the mediation of user modeling data can practically improve the quality of the provided personalization (i.e., generated recommendations). For example, in cross-technique mediation, converting sparse collaborative filtering UMs to the appropriate content-based UMs increases the accuracy of the generated predictions. Also in cross-domain mediation of collaborative filtering recommender systems user modeling data, importing either domain-related users' similarity or complete recommendations increases the accuracy of the generated predictions. Hence, UMs mediation can improve the quality of personalized services provided to the users.

Another practical implementation of UMs mediation is discussed in [10]. This work deals with two systems from different (although relatively close) domains: tourism and cultural heritage. This mediation scenario is aimed at converting user modeling data from Trip@dvice trip planning system [47] to PIL, personalized museum visitor's guide [32]. In Trip@dvice, the user modeling data are represented by cases, containing a set of products selected and examined by the user while planning a trip (e.g., attractions, restaurants, accommodation, and so on). In PIL, the user modeling data are content-based, i.e., a user's preferences are represented by a weighted vector of terms reflecting the content of the presentations on the museum exhibits. The mediation is performed through extracting the terms from the descriptions of Trip@dvice user's case items and projecting them to the set of terms representing the museum exhibits' presentations. Currently, this mediation is being evaluated.

Although cross-user mediation was extensively studied in collaborative filtering research, and cross-item and cross-representation mediations were studied in the reported works, the research of cross-context mediation is in its very initial stages. A general motivation for and the terminology of cross-context mediation were defined in [9], where the primary inference methods were also briefly sketched. These include rule-based and similarity-based inference over semantically-enhanced representations of users, items and context using GUMO, generic user modeling ontology [25]. However, none of the cross-context inference methods has been experimentally evaluated yet, mainly due to the lack of publicly available extensive personalization systems datasets storing contextual information.

It is planned to evaluate cross-context mediation of user modeling data within several scenarios with existing demonstrators developed in the *SharedLife* [54] and *Passepartout* [5] projects. In both, the users are supported in collecting and exchanging experiences across various contextual conditions, allowing evaluation of various cross-context reasoning mechanisms. *SharedLife* is a multi-user shopping scenario completed by other everyday life activities, such as listening to music, cooking, and so on. A user's positive feedback observed in a certain contextual condition can be used to provide personalized services in other contextual conditions. However, this should be justified by a sufficient similarity between the contexts, since the positive feedback might actually relate to several contextual elements. *Passepartout* illustrates typical search, browse and viewing activities of individual and group users with a personalized digital TV program guide. In *Passepartout*, user modeling data typically appear in different daily, weekly, monthly or yearly contexts. Hence, it is important to identify the right granularity on which the cross-context reasoning should be applied (e.g., recommending in contexts based on hours, days, or weeks). Another situation of cross-context reasoning is recommending TV programs for the same user when alone and when accompanied by other users, which may affect his/her preferences.

Clearly, context-aware personalization is one of the most promising future research directions [48], since provision of such services is crucial for the dissemination and wide acceptance of personalization systems. Hence, it will immediately lead to a need for cross-context mediation, allowing users' preferences to be inferred across various contextual conditions. Future research of cross-context mediation is needed to extend the initial ideas proposed in [9]. In particular, it is planned to focus on formalizing the cross-context mediation model, integrating it with known data representation and reasoning techniques and demonstrating its advantages in everyday scenarios and practical personalization systems.

References

[1] G. Adomavicius, R. Sankaranarayanan, S. Sen, A.Tuzhilin, "*Incorporating Contextual Information in Recommender Systems using a Multidimensional Approach*", in ACM Transactions on Information Systems, vol. 23(1), p.103-145, 2005.

[2] G. Adomavicius, A. Tuzhilin, "*Personalization Technologies: a Process-Oriented Perspective*", in Communications of the ACM, vol. 48(10), p.83-90, 2005.

[3] G. Adomavicius, A. Tuzhilin, "*Towards the Next Generation of Recommender Systems: A Survey of the State-of-the-Art and Possible Extensions*", in IEEE Transactions on Knowledge and Data Engineering, vol. 17(6), p.734-749, 2005.
[4] S. Aguzzoli, P. Avesani, P. Massa, "*Collaborative Case-Based Recommender System*", in proc. of the ECCBR Conference, 2002.
[5] L. Aroyo, H. Schut, F. Nack, T. Schiphorst, M. Kauw-A-Tjoe, "*Personalized Ambient Media Experience: move.me Case Study*", in proc. of the IUI Conference, 2007.
[6] S. Berkovsky, "*Decentralized Mediation of User Models for a Better Personalization*", in proc. of the AH Conference, 2006.
[7] S. Berkovsky, T. Kuflik, F. Ricci, "*Cross-Technique Mediation of User Models*", in proc. of the AH Conference, 2006.
[8] S. Berkovsky, T. Kuflik, F. Ricci, "*Enhancing Privacy while Preserving the Accuracy of Collaborative Filtering*", in proc. of the Workshop on Recommender Systems, 2006.
[9] S. Berkovsky, L. Aroyo, D. Heckmann, G.J. Houben, A. Kröner, T. Kuflik, F. Ricci, "*Predicting User Experiences through Cross-Context Reasoning*", in proceedings of the ABIS Workshop, 2006.
[10] S. Berkovsky, A. Gorfinkel, T. Kuflik, L. Manevitz, "*Case-Based to Content-Based User Model Mediation*", in proc. of the UbiqUM Workshop, 2006.
[11] S. Berkovsky, T. Kuflik, F. Ricci, "*Cross-Domain Mediation of User Models in Collaborative Filtering*", submitted.
[12] P.J. Brown, J.D. Bovey, X. Chen, "*Context-Aware Applications: From the Laboratory to the Marketplace*", in IEEE Personal Communications, vol. 4(5), p. 58-64, 1997.
[13] L. Buriano, M. Marchetti, F. Carmagnola, F. Cena, C. Gena, I. Torre, "*The Role of Ontologies in Context-Aware Recommender Systems*", in proc. of the MOSO Workshop, 2006.
[14] R. Burke, "*Knowledge-based Recommender Systems*", in A.Kent (ed.), 'Encyclopedia of Library and Information Systems', vol. 69(32), 2000.
[15] R. Burke, "*Hybrid Recommender Systems: Survey and Experiments*", in User Modeling and User-Adapted Interaction, vol.12(4), p.331-370, 2002.
[16] M. Claypool, A. Gokhale, T. Miranda, P. Murnikov, D. Netes, M. Sartin, "*Combining Content-Based and Collaborative Filters in an Online Newspaper*", in proc of Workshop on Recommender Systems, 1999.
[17] L.F. Cranor, J. Reagle, M.S. Ackerman, "*Beyond Concern: Understanding Net Users' Attitudes about Online Privacy*", Technical Report, AT&T Labs-Research, 1999.
[18] A.K. Dey, G.D. Abowd, "*Towards a Better Understanding of Context and Context-Awareness*", in proc. of the HUC Symposium, 1999.
[19] L Francisco-Revilla, M. Shipman, "*Managing Conflict in Multi-Model Adaptive Hypertext*", in proc. of the Hypertext Conference, 2004.
[20] A. Goker, H.I. Myrhaug, "*User Context and Personalization*", in proc. of the ECCBR Conference, 2002.
[21] K. Goldberg, T. Roeder, D. Gupta, C. Perkins, "*Eigentaste: A Constant Time Collaborative Filtering Algorithm*", in Information Retrieval, vol.4(2), p.133-151, 2001.
[22] D. Goldberg, D. Nichols, B. Oki, D. Terry, "*Using Collaborative Filtering to Weave an Information Tapestry*", in Communications of the ACM, vol. 35(12), pp. 61-70, 1992.
[23] N. Good, J.B. Schafer, J.A. Konstan, A. Borchers, B. Sarwar, J. Herlocker, J. Riedl, "*Combining Collaborative Filtering with Personal Agents for Better Recommendations*", in proc. of the AAAI Conference, 1999.
[24] U. Hanani, B. Shapira, P. Shoval, "*Information Filtering: Overview of Issues, Research and Systems*", in User Modeling and User Adapted Interactions, vol. 11 (3), p. 203-259, 2001.
[25] D. Heckmann, "*Ubiquitous User Modeling*", PhD thesis, Saarland University, 2005.

[26] J.L. Herlocker, J.A. Konstan, A. Borchers, J. Riedl, "*An Algorithmic Framework for Performing Collaborative Filtering*", in proc. of the SIGIR Conference, 1999.
[27] J.L. Herlocker, J.A. Konstan, L.G. Terveen, J.T. Riedl, "*Evaluating Collaborative Filtering Recommender Systems*", in ACM Transactions on Information Systems, vol.22(1), p.5-53, 2004.
[28] Internet Movie Database, *http://www.imdb.com*
[29] A. Kobsa, "*Generic User Modeling Systems*", in User Modeling and User-Adapted Interaction, vol.11(1-2), pp.49-63, 2001.
[30] R. Kohavi, G.H. John, "*Wrappers for Feature Subset Selection*", in Artificial Intelligence Journal, vol.97(1-2), pp.273--324, 1997.
[31] B. Krulwich, "*Lifestyle Finder: Intelligent User Profiling Using Large-Scale Demographic Data*", in Artificial Intelligence Magazine, vol. 18(2), pp. 37-45, 1997.
[32] T. Kuflik, J. Sheidin, S. Jbara, D. Goren-Bar, P. Soffer, O. Stock, M. Zancanaro, "*Supporting Small Groups in the Museum by Context-Aware Communication Services*", in proc. of the IUI Conference, 2007.
[33] G. Linden, B. Smith, J. York, "*Amazon.com Recommendations: Item-to-Item Collaborative Filtering*", in IEEE Internet Computing, vol. 7(1), pp. 76-80, 2003.
[34] P. Maes, "*Agents that Reduce Work and Information Overload*", in Communication of the ACM, vol. 37 (7), pp. 31-40, 1994.
[35] N. Manouselis, D. Sampson, "*A Multi-Criteria Model to Support Automatic Recommendation of e-Learning Quality Approaches*", in proc. of the ED-MEDIA Conference, 2004.
[36] P. McJones, "*EachMovie Collaborative Filtering Data Set*", HP Research, 1997.
[37] S.M. McNee, S. Lam, J. Konstan, J. Riedl, "*Interfaces for Eliciting New User Preferences in Recommender Systems*", in proc. of the UM Conference, 2003.
[38] B. Mehta, C. Niederée, A. Stewart, M. Degemmis, P. Lops, G. Semeraro, "*Ontologically Enriched User Profiling for Cross System Personalization*", in proc. of the UM Conference, 2005.
[39] M. Montaner, B. Lopez, J.L. de la Rosa, "*A Taxonomy of Recommender Agents on the Internet*", in AI Review, vol.19(4), p.285-330, 2003.
[40] M. Morita, Y. Shinoda, "*Information Filtering Based on User Behavior Analysis and Best Match Retrieval*", in proc. of the SIGIR Conference, 1994.
[41] M.D. Mulvenna, S.S. Anand, A.G. Buchner, "*Personalization on the Net Using Web Mining*", in Communications of the ACM, vol.43(8), p. 123-125, 2000.
[42] J. Pascoe, "*Adding Generic Contextual Capabilities to Wearable Computers*", in proc. of the ISWC Symposium, 1998.
[43] M. Pazzani, "*A Framework for Collaborative, Content-Based and Demographic Filtering*", in Artificial Intelligence Review, vol. 13(5-6), pp. 393-408, 1999.
[44] U. Rabanser, F. Ricci, "*Recommender Systems: Do they have a Viable Business Model in E-Tourism?*", in proc. of the ENTER Conference, 2005.
[45] Resource Description Framework, available online at *http://www.w3c.org/RDF/*
[46] P. Resnick, H.R. Varian, "*Recommender Systems*", in Communications of the ACM, vol. 40(3), pp. 56-58, 1997.
[47] F. Ricci, B. Arslan, N. Mirzadeh, A. Venturini, "*ITR: a Case-Based Travel Advisory System*", in proc. of the ECCBR Conference, 2002.
[48] F. Ricci, "*Recommendations in Context*", in proc. of the EC-Web Conference, 2006.
[49] F. Ricci, D. Cavada, N. Mirzadeh, A. Venturini, "*Case-Based Travel Recommendations*", in 'Destination Recommendation Systems: Behavioural Foundations and Applications', CABI pubs., 2006.
[50] B. Sarwar, G. Karypis, J. Konstan, J. Riedl, "*Item-Based Collaborative Filtering Algorithms*", in proc. of the WWW Conference, 2001.

[51] J.B. Schafer, J. Konstan, J. Riedl, "*Electronic Commerce Recommender Applications*", in Journal of Data Mining and Knowledge Discovery, vol. 5(1/2), p.115-152, 2000.
[52] U. Shardanand, P. Maes, "*Social Information Filtering: Algorithms for Automating 'Word of Mouth'*", in proc. of the CHI Conference, 1995.
[53] M. Vozalis, K.G. Margaritis, "*Enhancing Collaborative Filtering with Demographic Data: The Case of Item-Based Filtering*", in proc. of the ISDA Conference, 2004.
[54] W. Wahlster, A. Kröner, D. Heckmann, "*SharedLife: Towards Selective Sharing of Augmented Personal Memories*", in 'Reasoning, Action and Interaction in AI Theories and Systems', Springer pubs., 2006.
[55] Y. Wang, A. Kobsa, "*Impacts of Privacy Laws and Regulations on Personalized Systems*", in proc. of the PEP Workshop, 2006.
[56] C.N. Ziegler, S.M. McNee, J.A. Konstan, G. Lausen, "*Improving Recommendation Lists through Topic Diversification*", in proc. of the WWW Conference, 2005.