# A Semantic Approach to Avoiding Fake Neighborhoods in Collaborative Recommendation of Coupons through Digital TV

Manuela I. Martín-Vicente, Alberto Gil-Solla, Manuel Ramos-Cabrer, Yolanda Blanco-Fernández, and Martín López-Nores

**Abstract — ***Consumers are flooded with amounts of discount coupons, oftentimes for products that are far from their interests. This marketing custom is already rising on the Internet and is imminent in Digital TV, where the massive sending of coupons leads to their devaluation and consumer indifference. The computing capabilities of these media permit to alleviate this problem by means of recommender systems, which are very useful tools in application domains that suffer from information overload. However, current recommender systems overlook the diversity of products and services available in the market, which gives rise to forming fake neighborhoods in collaborative filtering strategies. In this paper, we apply semantic reasoning techniques to avoid such fake neighborhoods and, thereby, improve the recommendation process. Furthermore, taking advantage of the Digital TV medium, we propose matching the recommended coupons to TV contents semantically related with them, in order to increase their redemption*[1].

**Index Terms — Recommender systems, semantic reasoning, collaborative filtering, Interactive Digital Television, e-commerce.**

## I. INTRODUCTION

An important change is occurring in television, namely the analog to digital transition. Some countries have already made the analog switch-off and many others are at the gates of doing so. One of the most appealing features of the new medium is the ability to provide interactive digital contents, by sending software applications along with the audiovisual stream. Thus, since the TV is accessible to the vast majority of the population, a large number of users can be reached with services that were already available to the PC, and by others that challenged the technology so far. Among the many interesting services that can benefit from the new medium are those related to e-commerce. Closely connected to this field is the well-established marketing mechanism of distributing coupons attached to the promotional campaigns of many products. The massive and indiscriminate sending of coupons through newspapers, magazines and the Internet leads to their devaluation and consumer indifference [1]. Through the TV, coupons could be distributed in real time, whenever desired and they could be attractively presented to the user, relating them to the most appropriate programs broadcast at any moment. More importantly, it is also possible to personalize the delivery of coupons in the interests of users, thereby increasing their effectiveness –a user who receives few coupons tailored to his/her preferences is more likely to redeem them than another one flooded with coupons, most of them offering products that do not interest him/her at all.

Selecting which coupons best match each user's interests is a particular problem of recommendation. Recommender systems [2], [3] arose during the last years to save the users from dealing with the overwhelming amounts of information present in numerous domains, providing them with personalized suggestions. Specifically, recommenders for e-commerce websites have been increasingly explored in literature in the recent years due to their business interest, since selecting and offering to each user products he/she may be interested in can considerably increase sales revenue [4]-[6]. This aspect is even more relevant in the distribution of coupons, since the role of the users in e-commerce websites is active (they usually initiate the searching of what they want and explore the possibilities) whereas it is passive with coupons (users are deluged with coupons of a huge range of possible products).

One of the most popular recommendation strategies is collaborative filtering [7], [8], based on offering to the user who will receive the recommendation (hereafter, the *active user*) items that were appealing to other individuals with similar preferences (the so-called *neighbors*). Such similarity is estimated from the ratings registered in their personal profiles, considering two users alike if they have assigned similar ratings to the same items. Then, in order to predict the interest in an item unknown by a user (the *target item*), collaborative recommenders consider the ratings his/her neighbors have given to that item.

In an e-commerce application domain such as the distribution of coupons, the great diversity of products and services available to the consumer makes the current collaborative techniques subject to the formation of *fake neighborhoods* [9]. This happens because, when measuring the similarity between users, the ratings assigned to *all* the products included in their profiles are considered. The problem arises when users share likings for lots of products (for which they are reckoned to be neighbors) but they differ in products like the one that is being considered for recommendation at any given moment. For example, two users who have the same liking for books, cars, sports equipment and music but differ in food are fake neighbors whenever the target product is foodstuff.

[1] Work supported by the Ministerio de Educación y Ciencia research project TSI2007-61599, and by the Consellería de Educación e Ordenación Universitaria (Xunta de Galicia) incentives file 2007/000016-0.

Autors are with the Department of Telematics Engineering, University of Vigo, 36310 Vigo, Spain (e-mail: mvicente@det.uvigo.es, agil@det.uvigo.es, mramos@det.uvigo.es, yolanda@det.uvigo.es, mlnores@det.uvigo.es).

Manuscript received January 13, 2010 

In recent years, reasoning techniques borrowed from the Semantic Web have been adopted in recommender systems to successfully alleviate recurring weaknesses [10]-[12]. Traditional strategies employ syntactic matching techniques in the recommendation process, usually based on comparing text chains. Then, they are only capable to detect similarities between products identified with exactly the same terms. However, considering semantics, users' preferences and available items can be compared in a flexible and enhanced way and a great amount of useful knowledge can be included in the personalization process.

In this paper, we present two strategies based on semantic reasoning that aim to upgrade the distribution of coupons in a twofold manner: (i) overcoming the fake neighborhood problem, and (ii) relating coupons and TV contents in a meaningful way. The former means an improvement in the collaborative recommendations in e-commerce involving heterogeneous products and, thence, in the selection of the most suitable coupons for each user. The latter intends to present the recommended coupons in an attractive way, within the TV programs the user watches, in order to encourage to their redemption. For instance, a documentary about Italy may be considered by our strategy a propitious context to offer a discount coupon for a pizzeria.

The paper is organized as follows: Firstly, section II summarizes the technologies involved with our strategy in the IDTV field. Next, section III describes our reasoning-based coupon selection strategy, exposing how semantics can drive a neighborhood formation process free of fake neighbors. Section IV describes our strategy to present coupons to the users, relating them with TV contents by a semantic matching process. Afterwards, section V explores our system architecture and section VI describes a sample scenario to illustrate our strategies. Finally, section VII concludes the paper and points out the improvements achieved with our approach.

## II. Technological Landscape

Our approach brings together solutions from diverse areas of information and communications technology:

- *Formalization of domain knowledge*. Descriptions of products (promoted by the coupons) and TV programs must be formalized to accomplish the semantic reasoning processes that lead to our recommendation of coupons avoiding fake neighborhoods and their matching with TV programs. For that purpose, we exploit the mechanisms developed in the Semantic Web.
- *Manipulation of multimedia contents*. The presentation of coupons within TV programs implies manipulating the broadcast multimedia contents. We exploit the available technologies to carry out such presentation minimizing intrusiveness, i.e. avoiding interfering with the viewing of the programs.
- *IDTV platforms*. It is also necessary to bear in mind the characteristics of the IDTV platforms for the domestic receivers, considering the technical aspects of the available networks.

### A. Background in Formalization of Domain Knowledge

Ontologies are used in the Semantic Web as a form of knowledge representation about the world or some part of it. In this field, an ontology is defined as a *formal, explicit specification of a shared conceptualization* [13]. Basically, the role of ontologies is to facilitate the construction of a domain model, providing a vocabulary of terms and relations with which to model the domain. Concepts are represented by *classes* and their relationships by *properties*. Both entities are hierarchically organized in the conceptualization, which is populated by including specific instances of both the classes and properties. For instance, in the context of a recommender system, the instances of classes represent the available products and their attributes, whereas the instances of properties link the products and attributes to each other.

Ontologies have become the cornerstone in the Semantic Web due to two reasons. On the one hand, as these conceptualizations represent formally a specific domain, they permit to employ inference processes to discover new knowledge from the formalized information. On the other hand, ontologies facilitate automated knowledge sharing, by allowing an easy reusing between users and software agents. To this aim, several ontology languages have been developed in the Semantic Web which provide different facilities. OWL (Web Ontology Language [14]), standardized by W3C, has been the most expressive language to date. OWL 2, an extension and revision of the language, became a W3C recommendation in October 2009.

### B. Background in Manipulation of Multimedia Contents

Manipulation of broadcast multimedia contents has been traditionally limited by their transmission as binary flows, with hardly more structure than video frames or audio streams. Thereby, the composition capabilities of the IDTV receivers could not go beyond overlaying images with some color blending, mixing audio track while controlling volume or fading, and other simple effects. These restrictions are bound to disappear with the consolidation of MPEG-4 [15], a global standard for the creation, delivery and playback of interactive multimedia contents for an unrestricted range of networks and media. MPEG-4 allows decomposing a scene into individual pieces of audio and video, corresponding to the different objects (people, furniture, etc.) or sounds that appear in it. Specifically, MPEG-4 provides a language for describing and dynamically changing the scene, named the *Binary Format for Scenes* (BIFS). BIFS commands can not only add objects or delete them from the scene, but also change visual or acoustic properties. In fact, BIFS can be used to animate objects and to define their behavior in response to user input at the decoder, thus allowing the user to interact with objects within the scene. Furthermore, to give even more support for interactive applications, MPEG-4 includes a specific subset of the object-oriented Java language called MPEG-J (*MPEG-4 Java*), MPEG-J defines interfaces to elements in the scene, network resources, terminal resources and input devices, permitting not only to create highly interactive multimedia content, but also to make optimal use of terminal and network resources.

### C. *Background in IDTV Platforms*

IDTV stands for the broadcasting of a digital transport stream of traditional audiovisual contents mixed with binary data, so making it possible to deliver multimedia software applications to be executed in a DTV receiver or a Set-Top Box (STB).

Regarding the transmission of the audiovisual contents, the most successful initiatives for domestic receivers have been those proposed by the Digital Video Broadcasting consortium (DVB) for terrestrial, cable and satellite forms of broadcast (DVB-T, DVB-C and DVB-S, respectively). These standards have been widely adopted in Europe and in practically all continents, as well as in the United States and Japan, where they coexist with ATSC and ISDB standards, respectively.

As for the interactive applications, Europe, Japan and the United States all produced different standards for middleware specially designed to work with their own data broadcasting standards. In particular, the possibilities for the domestic market include open standards like Multimedia Home Platform [16] (MHP) in Europe, Advanced Common Application Platform (ACAP) and OpenCable Application Platform (OCAP) in the United States, or Association of Radio Industries and Business (ARIB) in Japan. In order to provide interoperability in several markets, the Globally Executable MHP [17] (GEM) initiative specifies those elements of the MHP standard that may be replaced by functional equivalents, thus defining a common core, forming in turn the basis of harmonization of international standards, setting the foundations for a world-wide market.

## III. COUPONS SELECTION

In this section, we describe our semantics-enhanced recommendation strategy, based on collaborative filtering, that will be employed to select the most suitable coupons for each user. The main contribution here is our neighborhood formation approach, where we use semantics in order to avoid the problem of fake neighbors disclosed in section I.

Before focusing on the details of our strategy, we describe the two principal components of the involved reasoning framework: the domain ontology and the user profiles.

### A. *The Domain Ontology*

We use an OWL ontology that formally represents the knowledge of the e-commerce application domain, capturing the semantic descriptions (metadata) about the commercial products promoted by the coupons. A brief excerpt from this ontology is shown in Fig. 1, where white ellipses, white squares, gray squares and gray ellipses denote product classes, products, attributes and attribute classes, respectively.

Owing to the great diversity in the nature of commercial products, our ontology contains multiple hierarchies of classes that represent the domain concepts (*Foodstuff*, *Books*, *Music*, *Sports Equipment*, and so on), as well as specific instances of them (*Murder on the Orient Express, Japanese Melodies,* $Coffe_A$*, etc.*). Besides classes and instances of commercial products, the ontology includes their semantic attributes, which are also identified by instances belonging to diverse hierarchies. It is worth noting that some attributes will be typical of a specific hierarchy of products and other will be general. For example, *Foodstuffs* have some taste –*Sweet*, *Bitter*, *Sour*…–, *Books* do not, whereas the country of origin or the price range –*Luxury*, *Offer*, etc.– are linked to all kind of products. Each one of these attributes is related to the corresponding commercial products by conveniently labeled properties (*hasTaste*, *hasPriceRange*, *hasCountryOfOrigin*, *hasCreator*, etc.).

### B. *The User Profiles*

User profiles typically store the ratings provided by the users in relation to each item registered on them. However, our approach requires a formal representation of the users' preferences with which we would be able to reason about them and discover additional knowledge about the users' interests. To this aim, we model the users' preferences by reusing the knowledge formalized in the ontology of commercial products. Thus, in addition to the ratings provided by any user, his/her profile contains unique references or IDs to the corresponding products and related attributes. By means of an ID we can consult in the ontology all the semantic descriptions of the referenced product directly. For instance, a profile storing the title *Naruto* references an instance of the class *Manga* –in the *Books* hierarchy– with attributes like *Japan* (its country of origin) and *Masashi Kishimoto* (its creator).

Semantics permits us to transfer product ratings to their attributes. A product rating entails (initially) its rating for all its attributes. Continuing the last example, if *Naruto* is appealing for a user, the semantic characteristics *Japan* and *Masashi Kishimoto* would have an appealing rating too. However, the same attributes can be related to several products stored in a profile at once. In that case, the attribute rating is the average of the ratings provided for all the products in the profile sharing that attribute. The more products contributing to an attribute rating, the more informative it is about the user's preferences. Returning to the ongoing example, the more Japanese products (in addition to *Naruto*) are stored in the user's profile, the more meaningful would be his/her liking for the attribute *Japan*; and the more books written by *Masashi Kishimoto* he/she has rated, the more meaningful would be his/her liking for this author.

### C. *Our reasoning-based strategy*

The key factor when making a recommendation based on collaborative filtering is the formation of the neighborhood,

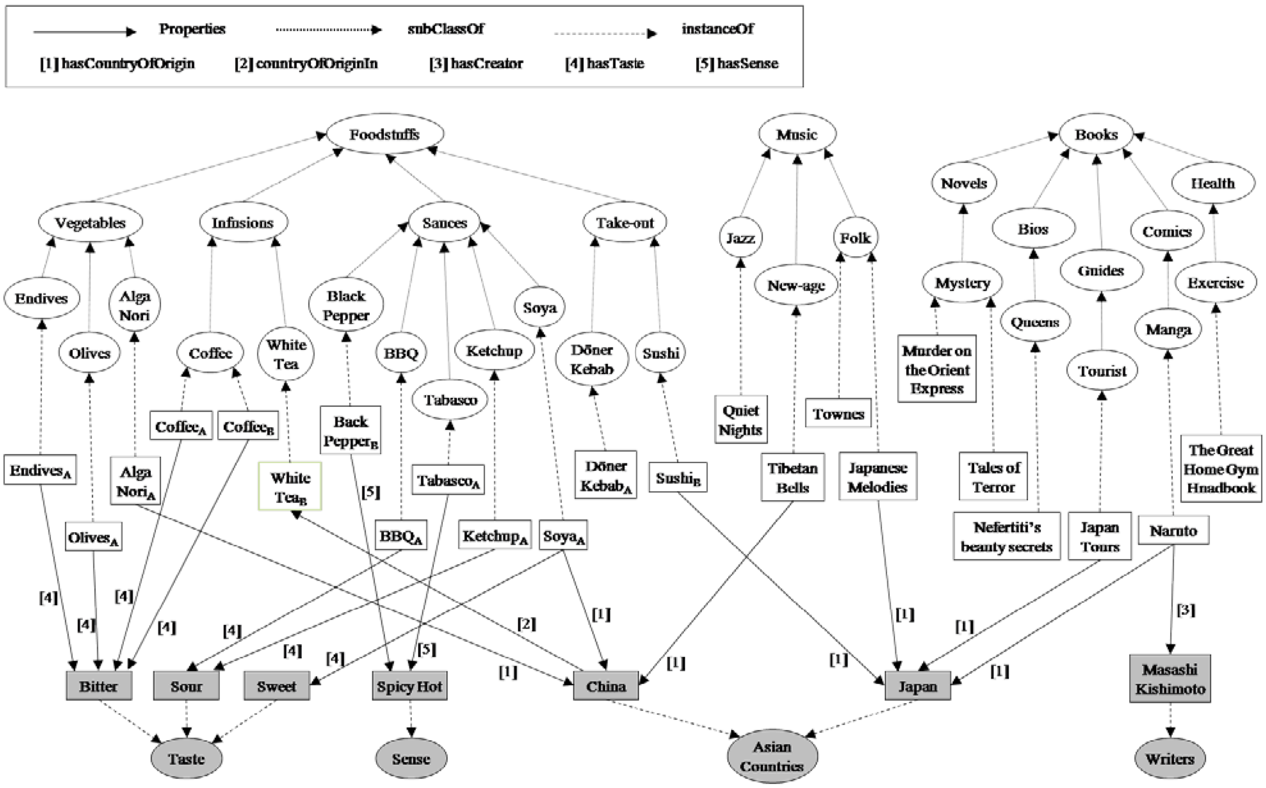



**Fig. 1. A micro-excerpt from our ontology of commercial products.**

because it determines the hit in the recommendation to the active user. Our goal is to avoid the fake neighbors problem in that selection. To achieve it, we only consider the group of users whose preferences are more similar to the active user's as regards the target product. That is to say, we form the neighborhood *dynamically*, considering a different neighborhood depending on the kind of product we want to recommend to a user (books, music, sports equipment, etc.). In the following lines we delve into the foundations of our semantic approach:

- Firstly, we search for similarities among user profiles just considering those products that belong to the *same hierarchy* in the ontology as the target. We benefit here from using an ontology in our strategy, since ontologies are hierarchical structures themselves.
- Secondly, we exploit the semantics when comparing product ratings, taking into account the class that each product belongs to instead of the product instance itself. Thereby, we detect overlaps among users' preferences even if their profiles do not store exactly the same products. Then, for example, we infer a common preference between two users who like coffee even if they have stored in their profiles different brands of coffee (*Coffe*$_A$ and *Coffe*$_B$). We also infer a common preference between a user who likes the book titled *Murder on the Orient Express* and a user who likes the book titled *Tales of Terror*, since both share likings for mystery novels.
- We do not limit ourselves to comparing product ratings. Instead, thirdly, we go one step further by *utilizing the attributes* we have at our disposal to enhance the neighborhood selection: the semantic features stored in a user's profile give us further information about his/her interests and consumption habits. Then, two users are more likely to be considered neighbors in a hierarchy of products, the more similar ratings they have assigned to the same classes of products in that hierarchy and the more semantic features they share about it. As described in section III-A, we distinguish two types of attributes in the ontology of commercial products: those typical of a specific hierarchy, and the general ones. The former give us clues about the users' likings in a particular kind of product (e.g. flavors they prefer if we are talking about food, or their preference for certain writers in case of books). The latter reflect users' interest in several kinds of products (e.g. an appealing rating of the attribute *Japan* may be due to a user's interest in a variety of Japanese products: food, music, books, cars, etc.). General attributes are significant regardless of the product we want to recommend, as they contain wealthy information about the users' general interests (e.g. preference for luxury items).

Once the neighborhood has been formed, it is time to select which coupons could potentially interest the active user, by considering the preferences of the corresponding neighbors. To this aim, we have adapted the semantics-enhanced collaborative strategy of the AVATAR system, presented in [18]. Specifically, we modified that strategy, originally designed for recommending TV programs, to deal with heterogeneous products. Briefly, instead of taking into account all the ratings included in the neighbors' profiles during the process, we just utilize those regarding to the kind of product being considered for possible recommendation. Anyhow, we benefit from the ability of the original strategy to consider the preferences of all the neighbors, regardless of whether or not their profiles contain the target item. In case they do, the rating for such item is considered. Otherwise, AVATAR estimates the degree of interest for each neighbor by a content-based strategy, which computes the semantic similarity between the target item and his/her preferences (see [18] for details).

## IV. PRESENTATION OF COUPONS THROUGH DTV

Next, we set out how we exploit the Digital TV medium to attractively offer to each user those coupons that, as explained in the previous section, have been selected for him/her. Our proposal consists in: (i) discovering relationships between the products on promotion and the TV contents in a meaningful way, and (ii) presenting the coupons within the corresponding programs minimizing intrusiveness. First of all, as our strategy includes semantic reasoning in the matching process between products and TV programs, we present the required TV ontology, that will be interlinked with the product ontology previously described.

### *A. TV Ontology*

In order to formalize the TV domain we make use of the ontology presented in [19], in which semantic characterizations of TV contents were extracted from TV-Anytime metadata specifications [20]. The TV programs represented in the ontology are identified by specific instances belonging to a hierarchy of genres that is organized in several levels (e.g. *Fiction*, *Music*, *Leisure*, *Sports*, etc.). However, the elements that most interest us for our purpose –relating TV contents with coupons by semantic reasoning– are the attributes contained in the ontology, which are also identified by instances belonging to diverse hierarchies. These structures organize information referred to the topics of the programs, places related to them and credits involved, among others. Each one of these attributes is related to the TV programs by conveniently labeled properties (e.g. *isAbout*, *hasPlace*, *hasActor*, etc.).

We establish the relationships between the TV ontology and the ontology of commercial products through shared attributes between their instances, that is, between programs and specific products presenting common semantic features. In Fig. 2 we present a small sample of how instances of the TV ontology (on the left hand side, classified under "TV contents") and commercial products (on the right hand side, classified under "Books" and "Foodstuff") interrelate.

### *B. Semantic matching between coupons and TV contents*

We start from a set of potentially interesting coupons for each user, which will be sent to their corresponding STBs along with the TV contents broadcast at any time. Our goal is to discover which TV programs have content that may encourage the redemption of every coupon, i.e. identifying the programs whose subject is related to the product being promoted in each case.

To this end, we propose a metric that quantifies the *semantic similarity* between the product promoted in each coupon and the contents of the TV programs that will be broadcast shortly. The TV-Anytime specification considers not only the characterization of the semantic descriptions for an entire TV program, but also for any segment in the audiovisual stream. Then, we are able to relate coupons to either programs as a unity or specific scenes contained in them. This allows exploring a much larger number of contexts in which coupons could be suitably presented.

Our semantic similarity metric is based on discovering implicit relations between the coupons and TV contents to match. Then, the stronger the connections between two of them, the greater will be the value of their similarity. We discover relations among programs and products that share semantic characteristics (e.g. topic, places, etc.). Particularly, our metric distinguishes two types of connections: the ones established through *common attributes*, if TV program and product share some attributes (the same instance in the ontology), and the ones established through *sibling attributes*, if they have attributes belonging to the same class in the ontology. For example, between a TV program whose scene is set in Japan and a product whose country of origin is Japan there is a relation through a common attribute, whereas between that program and a Chinese product there is a relation through sibling attributes (*Japan* and *China* belong to the same class in the ontology, *Asian Countries*). Our metric considers the relations through common attributes stronger than those through sibling attributes. Consequently, the more relations we can infer between a TV program and a product (and the more relations from that set are established through common attributes), the greater their semantic similarity will be.

When the matching process finishes, each coupon will be linked to zero, one or more TV contents. Once in the receiver, when a user is viewing a program, his/her DTV receiver presents him/her the most appropriate coupons. Our point is that we can increase coupon redemption by reminding the users they have certain coupons at their disposal and by presenting them attractively in a propitious context.

### *C. A non-intrusive presentation of coupons on TV*

Our system makes use of the advanced capabilities of the MPEG-4 standard for multimedia to *minimize intrusiveness* in the presentation of coupons, i.e. that such presentation does not interfere with the viewing of TV programs. In order to indicate to the user that a coupon (related with the contents he/she is currently watching) is available, an *MPEG-4 processor* in the receiver is in charge to place a small symbol on the screen, by means of the corresponding BIFS commands. MPEG-4 allows distinguishing the relevant objects in a scene (e.g. characters) from those which are not (e.g. the background of the image). Thus, to avoid

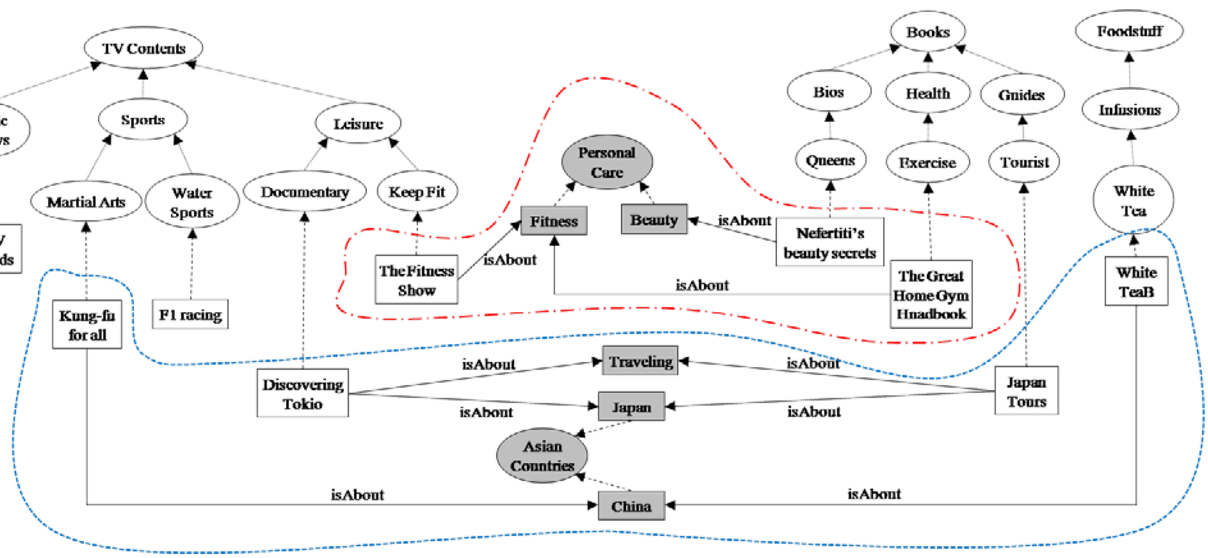



Fig. 2. Some relations between the ontology commercial products and the TV ontology

intrusiveness, the processor selects the most appropriate position (free of relevant objects) in the scene to place the symbol –preferably, one of the corners of the screen.

In the presentation of coupons, we also exploit the *interaction* capabilities offered by MPEG-4. The aforementioned symbols, that indicate the availability of a coupon in a scene, will be interactive objects. Then, when the user shows interest in a certain coupon, by clicking a button in his/her remote control or keypad, the system allows him/her to simply flag it –for future revision– or access the *coupons management application*. This simple MHP application, located in the receiver, permits the user to consult his/her coupons –stored in a repository– or redeem them, to indicate ratings for the promoted products, to inform about products he/she is interested in receiving coupons, and so on.

## V. System Architecture

In Fig. 3, we show the architecture of our system. The personalization scheme we present is server-side because it can provide the computing capability (limited in STBs) needed to handle large ontologies. To address privacy concerns, users and service provider sign a service-level agreement. In it, the service provider agrees not to disclose the content of the users' personal profiles, storing them on trustworthy servers. On the other hand, the users let the system use that information out of their receivers.

The server is composed of two main blocks: the *coupons recommender system* and the *coupons and TV contents matching module*.

The *coupons recommender system* is responsible for selecting, from among the heap of coupons available, those that most likely could interest to each user. There are four agents in charge of this process:

*Neighborhoods formation and selection agent*: It is responsible for creating and dynamically selecting the neighborhoods (avoiding fake neighbors) for the active user. According to what we have exposed in section III-C, the neighborhood varies depending on the kind of product to be considered in the recommendation. Since we can compute the neighborhood formation off-line, this agent forms as many neighborhoods for the active user as hierarchies we have defined in the ontology. Then, during the recommendation process, it is responsible for selecting the corresponding neighborhood to each target product.

*Collaborative filtering agent*: This agent receives as input every coupon to be considered as target item of the recommendation, associated in each case with the appropriate neighborhood for the active user. By means of the collaborative filtering strategy based on AVATAR (as discussed in section III-C), it decides whether the target product might be interesting for the user. Once the process is complete, it obtains the group of coupons to be sent to the user's DTV receiver.

*Information agent*: This agent accesses to the ontology of commercial products and to the database where the users' profiles are stored. It serves as an interpreter, providing the required information by the *neighborhoods formation and selection agent* and by the *collaborative filtering agent*. Besides, it manages the

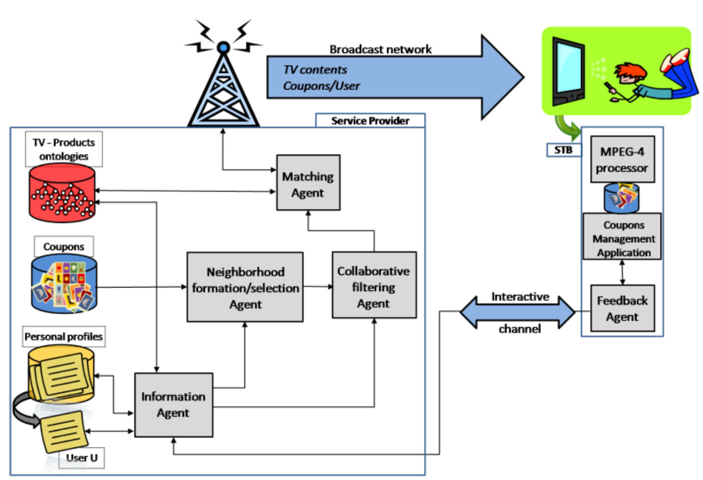



Fig. 3. System architecture.

information received from the *feedback agent* through the interactive channel, with which it updates the users' profiles.

*Feedback agent*: In the receiver, this agent compiles the information provided by the users through their interaction with the system: implicitly, by observing their behavior (coupons consulted, coupons redeemed…), and explicitly, through the product ratings they want to provide. This process allows refining the knowledge about the users' preferences and, therefore, guaranteeing updated recommendations.

Receiving as input the coupons selected for each user and the programming schedule, the *matching agent* accesses to the products and TV ontologies and carries out the semantic matching process described in section IV-B. As a result, the group of coupons that are sent to a user's receiver is enriched with information that links each coupon to zero, one or more TV programs on the schedule.

In the receiver, the STB stores the received coupons in a repository. The user may access his/her coupons at any time through the MHP *coupons management application*, previously mentioned. Furthermore, during the viewing of TV, when the context is appropriate, the *MPEG-4 processor* presents the corresponding coupons, by incorporating interactive objects to the scene, as discussed in section IV-C.

## VI. A Sample Scenario

In this section, we describe an application scenario to illustrate our reasoning-based strategies. Firstly, we present an example of coupons recommendation, highlighting how the use of semantics in the neighborhood formation allows us to avoid fake neighbors. Then we illustrate the matching of coupons and TV programs.

Due to space limitations, we assume some simplifications in these examples. So, we just consider brief excerpts from the ontologies, including a reduced number of attributes and class hierarchies. We also consider a small number of users in the recommendation process, as well as a subset of programs and coupons instances in the matching process. Nevertheless, these limitations do not prevent us from showing the benefits of our strategies, emphasizing the usefulness of the semantic inference in both.

### A. Coupons recommendation

To illustrate our strategy, we rely on the excerpt from the ontology of commercial products shown at the top of Fig. 1 and on Table I, that contains partial profiles of some users. The target item is *White Tea*, which belongs to the *Foodstuff* hierarchy. We consider that the active user is U, whose positive and negative preferences are described in the first column of Table I. The remaining columns correspond to the profiles of the other users (candidates to be neighbors of U): $N_F$, $N_1$ and $N_2$.

For simplicity, we do not develop the neighborhood formation quantitatively, but we will discuss it qualitatively.

Current neighborhood formation mechanisms would select the user $N_F$ as a neighbor of U given that, although they do not share likings in food, they have similar preferences in all the rest (here, *Books* and *Music*). In fact, even a syntactic approach could obtain their similarity concerning *Music*, since they have rated similarly exactly the same products (*Quiet Nights, Japanese Melodies* and *Tibetan Bells*). In relation to *Books*, their common dislike for mystery novels could be inferred solely with a semantic approach. In any case, that would be a fake neighbor to U as regards foodstuffs. The most reliable neighbors for recommending *White Tea* to U are those who have similar likings with regard to *Foodstuff*, although the likings in the rest of the hierarchies differ considerably. That is the case of $N_1$ and $N_2$.

$N_1$ has similar ratings to those of U in quite a few *Foodstuff* products and, consequently, they share liking in many attributes too. Therefore, we can see their relation forthwith. In contrast, $N_2$ does not have as many products in common with U but, looking at the attributes, we notice that they share a lot of likings. For instance, we get that bitter taste is appealing for both, since U loves *Coffee* and $N_2$ likes *Olives* and *Endives* (see Fig. 1). In the same way, we get that sour taste is unappealing for both, since U dislikes *Tabasco* and *Ketchup* and $N_2$ does not like *BBQ* sauce. Similarly, we can see that they both really dislike spicy hot sense, thanks to the U's unappealing rating for *Tabasco* sauce and the $N_2$'s unappealing rating for *Black Pepper* sauce. Considering semantics, we are able to discover similarities undetected with a syntactic approach.

Then $N_1$ and $N_2$ form U's neighborhood and, in this case, is immediate see how *White Tea* will be recommended to U, since both neighbors have such product stored in their profiles with an appealing rating. Therefore, a coupon for that product will be sent to U's DTV receiver.

Despite the simplicity of the example, we can see on it the importance of the general attributes. For instance, U's liking for Japanese products is evident not only by the food (*Sushi*), also by books (*Naruto*) and music (*Japanese Melodies*). As already mentioned, these attributes provide much information about the general preferences of the users, hence the importance of taking them into account regardless of the target item of the recommendation.

### B. Coupons and TV programs matching

Now consider that, in addition to a coupon of *White Tea*, our recommender system has selected for U, from among all the available coupons, a coupon of *Sushi*, one of the guide *Japan Tours*, one of the book *The Great Home Gym Handbook* and another of the book *Nefertiti's Skin Health Secrets*. Also, we have at our disposal the programme scheduling that is going to be broadcast shortly. It includes the programs: *Kung-fu for all*, a sports program on the Chinese martial art; *Discovering Tokyo*, a travelogue through the Japanese city; and *The Fitness Show*, a program that puts science to exercise.

Exploring in Fig. 1 (bottom) the *common* and *sibling attributes* we find several semantic relations between programs and coupons. On the one hand, we find the program *The Fitness Show* connected with both the book *The Great Home Gym Handbook* and the one entitled *Nefertiti's Skin Health Secrets*. The relation with the former is stronger, as it is established through the common attribute *Fitness*, whereas with the latter it shares the sibling attributes *Fitness* and *Beauty*, which belong to the class *Personal Care*. On the other hand, we see that there is relation between the programs *Kung-fu for All* and *Discovering Tokyo* with the products *Japan Tours*, *Sushi* and *White Tea*. Studying these relations we can

TABLE I
PARTIAL PROFILES

| User U | User $N_F$ | User $N_1$ | User $N_2$ |
|---|---|---|---|
| Quiet Nights ++ | Quiet Nights ++ | Quiet Nights -- | Quiet Nights -- |
| Japanese Melodies ++ | Japanese Melodies + | Townes -- | Townes -- |
| Tibetan Bells ++ | Tibetan Bells + | Tibetan Bells + | |
| Naruto ++ | Naruto + | | Naruto ++ |
| Tales of terror -- | Murder on the Orient Express -- | Tales of Terror ++ | Murder on the Orient Express ++ |
| The Great Home Gym Handbook + | The Great Home Gym Handbook + | Nefertiti's beauty secrets ++ | Nefertiti's beauty secrets + |
| | *White Tea*$_B$ - | *White Tea*$_B$ ++ | *White Tea*$_B$ ++ |
| *Coffee*$_A$ ++ | *Coffe*$_A$ -- | *Coffee* + | |
| | *Olives* -- | | *Olives* ++ |
| | | | *Endives* + |
| *Sushi*$_B$ ++ | *Sushi*$_B$ - | *Sushi*$_B$ ++ | *Sushi*$_B$ ++ |
| *Döner Kebab*$_A$ -- | *Döner Kebab*$_A$ - | | *Döner Kebab*$_A$ -- |
| *Tabasco*$_A$ -- | *Tabasco*$_A$ + | *Tabasco*$_A$ - | *Black Pepper*$_B$ -- |
| *Ketchup*$_A$ -- | *BBQ*$_A$ ++ | *Ketchup*$_A$ -- | *BBQ*$_A$ - |
| *Soya*$_A$ + | | *Soya*$_A$ ++ | *Soya*$_A$ + |
| *Alga Nori*$_A$ ++ | *Alga Nori*$_A$ - | *Alga Nori*$_A$ + | *Alga Nori*$_A$ ++ |
| Spicy Hot -- | Spicy Hot + | Spicy Hot - | Spicy Hot -- |
| Japan + | Japan + | Japan + | Japan ++ |
| China + | China - | China ++ | China ++ |
| Sweet + | | Sweet ++ | |
| Bitter ++ | Bitter -- | Bitter + | Bitter ++ |
| Sour -- | Sour ++ | Sour -- | Sour - |

Numerical data is omitted for simplicity: ++ indicates highly appealing, + moderately appealing, - moderately unappealing and – highly unappealing.
Products in italics belong to *Foodstuff* hierarchy, the rest of them belongs to either *Music* or *Books* hierarchies.
The shading part at the bottom contains the corresponding attributes (see Fig. 1).

conclude that:

- *Kung-fu for All* will be a more propitious context to offer *White Tea* than any of the other two products, as it shares with *White Tea* the attribute *China*, so that the relation is stronger than those established through *China* and *Japan* as sibling attributes (both of them belong to the class *Asian Countries*).
- *Discovering Tokyo* is related through common attributes with the *Japan Tours* guide and *Sushi*. However, note how *Japan Tours* is the product that is more semantically related to the program, as they share two common attributes, *Traveling* and *Japan*.

Thus, the coupons are sent to the U's DTV receiver together with the TV broadcast, indicating for every coupon which programs can be an appropriate context to encourage its redemption. We also indicate a ranking of preference of the TV contents, starting with the program that has the strongest semantic relation with the coupon, i.e. the one sharing most attributes (preferably common ones) with it.

## VII. CONCLUSION

Digital TV offers many possibilities, especially in terms of interactivity, bringing to the living room not only services that are already available via a computer but also others that are outside the technology at present.

In this paper, we have focused our attention on improving the distribution of coupons attached to the promotional campaigns of many products. This distribution, conducted through the Digital TV, can be done in real time, whenever desired and related to the TV programs broadcast at any moment. Furthermore, the coupons can be delivered selectively, according to individual preferences.

With this goal in mind, we have presented a new approach that uses semantic reasoning techniques to solve the fake neighborhood problem in e-commerce recommender systems, consequence of the great diversity of products offered to the consumers. Our proposal overcomes typical limitations of the conventional syntactic methods and, thanks to the way we handle neighborhoods, it improves the recommendation results. Specifically, we improve the precision of the recommendations, since we just consider the preferences of those users with similar likings in relation to the kind of product being considered for possible recommendation.

We have also proposed linking the recommended coupons with the TV contents broadcast at any time. The purpose is to encourage the redemption of coupons, presenting them to the users within the TV programs whose content may be an appropriate context. To do this, we have defined a semantic similarity metric based on inferring semantic relations between commercial products and TV programs. These relations are discovered between programs and products that share semantic features, so that the larger the number of common semantic features, the greater the similarity measurements. Coupons are presented during the most similar TV contents minimizing intrusiveness.

## REFERENCES

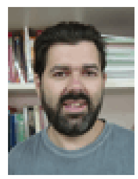
[1] H. K. Cheng, and K. Dogan, "Customer centric marketing with Internet Coupons," *Decision Support Systems*, vol. 44, no. 3, pp. 606-620, 2008.
[2] M. Montaner, B. López, J. L. de La Rosa, "A taxonomy of recommender agents on the Internet," *Artificial Intelligence Review*, vol. 19, no. 4, pp. 285-330, 2003.

[3] G. Adomavicius, and A. Tuzhilin, "Towards the next generation of recommender systems: a survey of the state-of-the-art and possible extensions," *IEEE Transactions on Knowledge and Data Engineering*, vol. 17, pp. 734-749, 2005.
[4] B. Sarwar, G. Karypis, J. Konstan, and J. Riedl, "Analysis of recommendation algorithms for e-commerce," *Proccedings of the Second ACM Conference on Electronic Commerce*, pp. 158-167, 2000.
[5] Y. H. Cho, and J. K. Kim, "Application of web usage mining and product taxonomy to collaborative recommendations in e-commerce," *Expert Systems with Applications*, vol. 26, no. 2, pp. 233-246, 2004.
[6] J. B. Schafer, J. A. Konstan, and J. Riedl, "E-commerce recommendation applications," *Data Mining and Knowledge Discovery*, vol. 5, no. 1-2, pp. 115-153, 2001.
[7] R. Burke, "Integrating Knowledge-based and collaborative filtering in recommender systems," *Proc. Of the Workshop on AI and Electronic Commerce*, pp. 69-72, 1999.
[8] W. Yang, Z. Wang, and M. You, "An improved collaborative filtering method for recommendations' generation," *Proccedings of the IEEE International Conference on Systems, Man and Cybernetics*, pp. 4135-4139, 2004.
[9] Y. Li, L. Lu, and L. Xuefeng, "A hybrid collaborative filtering method for multiple-interests and multiple-content recommendation in E-Commerce," *Expert Systems with Applications*, vol. 28, pp. 67-77, 2004.
[10] M. López-Nores, J.J. Pazos-Arias, J. García-Duque, Y. Blanco-Fernández, M. I. Martín-Vicente, A. Fernández-Vilas, M. Ramos-Cabrer, and A. Gil-Solla, "MiSPOT: dynamic product placement for digital TV through MPEG-4 processing and semantic reasoning," *Knowledge and Information Systems*, pp. 1-28, in press 2009.
[11] Y. Blanco-Fernández, J. J. Pazos-Arias, A. Gil-Solla, M. Ramos-Cabrer, M. López-Nores, J. García-Duque, A. Fernández-Vilas, and R. P. Díaz-Redondo, "Exploiting synergies between semantic reasoning and personalization strategies in intelligent recommender system: A case study," *Journal of Systems and Software*, vol. 81, no. 12, pp. 2371-2385, 2008.
[12] R. Wang, and Q. Kong, "Semantic-enhanced personalized recommender system," *Proccedings of the 6th International Conference on Machine Learning and Cybernetics*, pp. 4069-4074, 2007.
[13] T. R. Gruber, "A translation approach to portable ontology specifications," *Knowledge Acquisition*, vol. 5, no. 2, pp. 119-220, 1993.
[14] D. McGuinness, and F. van Harmelen. OWL Web Ontology Language Overview. W3C Recommendation, 2004.
[15] T. Ebrahimi, and F. Pereira. *The MPEG-4* book. Prentice Hall, 2002.
[16] S. Morris, and A. Smith-Chaigneau. *Interactive TV Standards*. Focal Press, 2005.
[17] DVB consortium, "Globally Executable MHP," ETSI TS 102 819 V1.3.1, June 2006.
[18] Y. Blanco-Fernández, J. J. Pazos-Arias, M. López-Nores, A. Gil-Solla, and M. Ramos-Cabrer, "AVATAR: An improved solution for personalized TV based on semantic inference," *IEEE Transactions on Consumer Electronics*, vol. 52, no. 1, pp. 223-231, 2006.
[19] Y. Blanco-Fernández, J. J. Pazos-Arias, A. Gil-Solla, M. Ramos-Cabrer, and M. López-Nores, "Providing entertainment by content-based filtering and semantic reasoning in intelligent recommender systems," *IEEE Transactions on Consumer Electronics*, vol. 54, no. 2, pp. 727-735, 2008.
[20] TV-Anytime Specification Series S-3 on Metadata, 2001 [on line]. Available in *http://www.tv-anytime.org*.

## BIOGRAPHIES

**Manuela Isabel Martín Vicente** was born in Barcelona, Spain in 1982. She received the Telecomunications Engineering degree from the University of Vigo in 2008. She is currently a Ph.D. student at the Department of Telematics Engineering from the same University, and a student member of the Interactive Digital TV Laboratory. Her major research interests are the design and implementation of services for Interactive Digital TV.

**Alberto Gil Solla** was born in Domayo, Spain in 1968. He is an associate professor in the Department of Telematics Engineering at the University of Vigo. He received the Ph.D. degree in Computer Science from the same University in 2000. Nowadays, he is involved with different aspects of middleware design, interactive multimedia services, and personalization applications for Interactive Digital TV.



**Manuel Ramos Cabrer** was born in Lugo, Spain in 1966. He received his degree in Telecommunications Engineering from Polytechnic University of Madrid, Spain in 1991, and his Ph.D. degree in Telematics from the University of Vigo, Spain in 2000. Since 2001, he is an associate professor in the Department of Telematics Engineering. His research topics are focused on recommender systems for Interactive Digital TV.

**Yolanda Blanco Fernández** was born in Orense, Spain in 1980. She received the Telecommunications Engineering degree from the University of Vigo in 2003, and her Ph.D. degree in Computer Science from the same University in 2007. She is an assistant professor in the Department of Telematics Engineering, and her research interests are focused on personalization services for Interactive Digital TV.

**Martín López Nores** was born in Pontevedra, Spain in 1980. He received the Telecommunications Engineering Degree from the University of Vigo in 2003 and his Ph.D. degree in Computer Science from the same University in 2006. He is an assistant professor in the Department of Telematics Engineering. His main research lines are focused on the design and development of interactive services for Digital TV.