rogress as appro imation to the truth: a defence of the verisimilitudinarian approach

Published version:

DOI:10.1007/s10670-012-9362-y

Multimedia Recommendation and Delivery strategies

Ruggero G. Pensa, Antonio Penta, and Maria Luisa Sapino

1 Introduction

In the last decade, the spread of broadband internet connections even for mobile devices has contributed to an increased availability of multimedia information on the web. At the same time, due to the decrease of storage cost and the increasing popularity of storage services in the cloud, the problem of information overload has become extremely serious even in personal/company archives. The need of managing, retrieving and presenting all these data has promoted the development of advanced multimedia information systems, which include recommendation modules to account for the requests of personalised data selection and presentation.

Recommendation systems estimate *ratings*, or *utilities*, which quantify users' degree of interest for the different available data, so that the data can be offered to the users in a personalised way, in decreasing order of interest. Multiple approaches have been proposed in the literature to estimate such degrees of interest. In *Content-Based Filtering* [31], the utility (for a user) of a given item is estimated as a function of the ratings given by the same user to other similar items. For example, in a cultural heritage recommendation application, in order to recommend a monument to a user, content-based filtering relies on the similarity between that monument and the monuments the user has rated highly in the past (do they come from the same historical period? were they designed by the same architect? Do they have the same style? etc.). Then, only the monuments that have a high degree of similarity to the user's preferred ones are recommended. Obviously, the effectiveness of content-based filtering methods strongly depends on the feature extraction algorithms, and on the

Ruggero G. Pensa
University of Torino, C.so Svizzera 185 – I-10149 Torino, Italy, e-mail: pensa@di.unito.it

Antonio Penta
University of Torino, C.so Svizzera 185 – I-10149 Torino, Italy, e-mail: penta@di.unito.it

Maria Luisa Sapino
University of Torino, C.so Svizzera 185 – I-10149 Torino, Italy, e-mail: mlsapino@di.unito.it

similarity based retrieval engine. Content-based techniques only take into account users' past experience and the features of the objects they ranked highly, while ignoring the feedback provided by other similar users. A possible drawback of these method is *overspecialization*, since the system can only recommend items that are similar to those already rated by the user.

A dual approach is *Collaborative Filtering* [2], in which filtering (i.e, or estimating the object's utilities) for a given user is computed by referring to the opinions of other users. Unlike content-based recommendation methods, collaborative systems focus on the similarity among users. Thus, a major challenge faced by collaborative filtering is the need to associate each user to a set of other users having similar profiles: in order to make any recommendations, the system collects data either by asking for explicit users' ratings, or through non intrusive profiling algorithms which implicitly log users' actions. Passive filtering uses aggregates computed on the gathered data (such as the average rating for an item) to make predictions. As a result, each user (similar to the ones whose data have been collected and analysed) will be given the same predictions for a particular item. Active filtering instead uses patterns in user history to make predictions, thus obtaining user-specific and context-aware useful recommendations. An important limitation of collaborative filtering systems is the so called *cold start problem*, i.e., the inability for a recommender to make meaningful recommendations for an object in the absence of ratings by other similar users, thus degrading the filtering performance.

Content-based filtering and collaborative filtering may be manually combined by the end-user specifying particular features, essentially constraining recommendations to have certain content features. More often they are automatically combined in the so called *hybrid approach* [9, 5, 8, 34] that helps overcome some limitations of each method. Different ways to combine collaborative and content-based methods into a hybrid recommender system can (*i*) implement collaborative and content-based methods separately and then combine their predictions; (*ii*) incorporate some content-based characteristics into a collaborative approach; (*iii*) incorporate some collaborative characteristics into a content-based approach, or (*iv*), construct a general unifying model that incorporates both content-based and collaborative characteristics. In this chapter we first present the *co-clustering based recommendation techniques*, which allow to combine heterogeneous multimedia content information and data about the users' preferences and rankings, thus overcoming some of the content based filtering drawbacks, as well as some collaborative filtering weaknesses. Then, we briefly discuss the challenges in multimedia delivery and the most common strategies adopted in the context of cultural heritage media delivery.

2 Grouping of related objects and users through co-clustering

In this section we introduce the co-clustering techniques, which group together related objects and users potentially interested in them.

Each object subject to recommendation may be represented in different and heterogeneous feature spaces. For instance, the picture of a monument may be described by annotations concerning history of the monument, the materials it has been built with, low-level image features, experts' descriptions, visitors' descriptions and reviews, and so on. Each of these sets of features contributes to the characterisation of the objects to different extents. Hence, it is important to consider congruently each type of descriptor during the recommendation process.

"Similar" objects are clustered together, according to a similarity notion that should consider all (or subsets of) the different spaces of features. To this purpose, high-order star-structured co-clustering techniques [29, 14, 18, 20, 22] can be employed to address the problem of heterogeneous data clustering. In this context, the same set of objects is represented in different feature spaces. Such data represent objects of a certain type, connected to other types of data, the features, so that the overall data schema forms a star structure of inter-relationships.

The co-clustering task consists in simultaneously clustering the set of objects and the set of values in the different feature spaces. In this way we obtain a partition of the objects influenced by each of the feature spaces and at the same time a partition of each feature space. Similarly, co-clustering allows to simultaneously group objects and users potentially interested in them.

The recommendation process leverages the clustering results to select a set of candidate objects by using the user's profile, which is modeled as sets of descriptors in the same spaces as the objects' descriptors.

We now provide the formalization of our problem. Let $\mathscr{O} = \{O^1, \ldots, O^M\}$ be a set of M multimedia objects and $\mathscr{F} = \{F^1, \ldots, F^N\}$ be a set of N feature spaces. A dataset can be viewed under the different views given by the different feature spaces F^k. Therefore, the view k is associated with each feature space F^k. Let $\mathscr{R} = \{R^1, \ldots, R^N\}$ be a star-structured relation over $\mathscr{O}$ and $\mathscr{F}$. For each relation matrix R^k, each value $r^k_{st} \in R^k$ corresponds to the counting/frequency/presence of feature $f^k_t \in F^k$ in object $O^s \in \mathscr{O}$. Without loss of generality, we assume that $r^k_{st} \in \mathbb{N}$. An example of two-views star-structured data is given in Figure 1(a).

In this recommendation problem, a user is represented as a set of vectors $U = \{u^1, \ldots, u^N\}$ in the same N feature spaces describing the objects. Each vector u^k is updated each time the user visits (or re-visits) an object, by considering the object features in each space at the instant of the visit. Let $O^U \subset \mathscr{O}$ be the set of objects visited by the user represented by U. Hence, the component of vector $u^k \in U$ related to feature f^k_t is computed as:

$$u^k_t = \sum_{O^s \in O^U} r^k_{st}$$

Clearly, the action of updating the vectors in U can be performed incrementally, as the user visits new objects. Notice that, thanks to this approach, users are not described by sets of objects, but by sets of features that characterize the objects they visit, like or browse.

The first step consists in identifying clusters of similar objects in $\mathscr{O}$ by leveraging all feature spaces by means of a star-structured data co-clustering approach. Its goal is to find a set of partitions $\mathscr{Y} = \{Y^1, \ldots, Y^N\}$ over the feature set $\mathscr{F} =$

	f_1^1	f_2^1	f_3^1	f_4^1
O^1	r_{11}^1	r_{12}^1	r_{13}^1	r_{14}^1
O^2	r_{21}^1	r_{22}^1	r_{23}^1	r_{24}^1
O^3	r_{31}^1	r_{32}^1	r_{33}^1	r_{34}^1
O^4	r_{41}^1	r_{42}^1	r_{43}^1	r_{44}^1
O^5	r_{51}^1	r_{52}^1	r_{53}^1	r_{54}^1

	f_1^2	f_2^2	f_3^2
O^1	r_{11}^2	r_{12}^2	r_{13}^2
O^2	r_{21}^2	r_{22}^2	r_{23}^2
O^3	r_{31}^2	r_{32}^2	r_{33}^2
O^4	r_{41}^2	r_{42}^2	r_{43}^2
O^5	r_{51}^2	r_{52}^2	r_{53}^2

(a)

	Y_1^1	Y_2^1	
X_1	t_{11}^1	t_{12}^1	p_1^1
X_2	t_{21}^1	t_{22}^1	p_2^1
	q_1^1	q_2^1	

	Y_1^2	Y_2^2	
X_1	t_{11}^2	t_{12}^2	p_1^2
X_2	t_{21}^2	t_{22}^2	p_2^2
	q_1^2	q_2^2	

(b)

Fig. 1 An example of a star-structured relation consisting of two feature spaces F^1 and F^2 (a) and the contingency tables associated with a related star-structured co-clustering (X,Y^1) and (X,Y^2) (b). Each t_{ij}^k represents the contingency value of cocluster denoted by i and j, p_i^k are marginals for row clusters denoted by i and q_j^k are the marginals for column clusters denoted by j.

$\{F^1,\ldots,F^N\}$, and a partition X of the object set $\mathscr{O}$ by optimizing a certain objective function. To solve the high-order star-structured co-clustering problem, several algorithms have been proposed based on different approaches.

For relations involving the set of objects and a unique feature-space (such as documents-words data), many co-clustering approaches have been proposed. Co-clustering has been studied in many different application contexts including text mining [17], gene expression analysis [16, 32] and graph mining [13] where these methods have yielded an impressive improvement in performance over traditional clustering techniques. The methods differ primarily by the criterion they optimize, such as minimum loss in mutual information [17], sum-squared distance [16], minimum description length (MDL) [13], Bregman divergence [6] and non-parametric association measures [35, 22]. Among these approaches, only those ones based on MDL and association measure are claimed to be parameter-free [25]. However, methods based on MDL are strongly restricted by the fact they can only handle binary matrices. Association measures, such as Goodman and Kruskal τ, are internal measures of the quality of a co-clustering based on statistical considerations. They have also another advantage: they can deal with both binary and counting/frequency data [22, 35]. From an algorithmic point of view, the co-clustering problem has been shown to be NP-hard [4] when the number of row and column clusters are fixed. Therefore, proposed methods so far are based on heuristic approaches.

Star-structured (co-)clustering, often referred to as high-order heterogenous star-structured (co-)clustering is an emerging topic whose importance is attested by an increasing number of works. Notice also that, in the machine learning community, this field of research is sometimes defined as multiview clustering. Since the topics is quite new, there is no classification for the proposed techniques. However, the existing approaches may be grouped into four main classes: *factorization-based approaches*, *information-theoretic approaches*, *probabilistic approaches* and *association-based approaches*.

2.1 Factorization-based approaches

Long et al. [29] use factorization to iteratively embed each type of data objects into low dimensional spaces in a way that takes advantage of the interactions among the different feature spaces. A partitional clustering approach (e.g., k-means) is then employed to obtain the final clustering computed on the transformed spaces. In the above formulated problem, the approach in [29] tries to minimize

$$L = \sum_{k=1\dots N} w^k \| R^k - C^O A^k (C^k)^T \|^2$$

where $C^O \in \{0,1\}^{M\times m}$ is a cluster indicator matrix for $\mathscr{O}$ such that $c^O_{pq} = 1$ denotes that pth object in O is associated with the qth cluster in X. Similarly $C^k \in \{0,1\}^{|F^k|\times n_k}$ is the cluster indicator matrix for Y^k. $A^k \in \mathbb{R}^{m\times n_k}$ is the cluster association matrix such that A^k_{pq} denotes the association between cluster p of X and cluster q of Y^k. Finally, $w^k \in \mathbb{R}_+$ is a weight associated to the kth relation. To compute the clustering X and $\mathscr{Y}$, the proposed algorithm first computes matrices C^O and C^k ($k = 1\dots N$) by solving a matrix factorization problem. Then, it uses k-means to transform each matrix into an indicator matrix. This method is quite difficult to adopt in practice, since it requires too many parameters: the number of clusters for the object set (m) and for each feature space (n_k, $k = 1\dots N$) and the weights w^k ($k = 1\dots N$).

Chen et al. [14] also propose a factorization method that performs multi-view co-clustering. The method is an extension of the Non-Negative Matrix Factorization approach that deals with multi-view data. The authors formulate the task as an optimization problem with non-negative matrix trifactorization of $\mathscr{R} = \{R^1,\dots,R^N\}$:

$$J = \min_{G^O\geq 0, G^k\geq 0, S^k\geq 0} \sum_{k=1}^{N} \| R^k - G^O S^k G^k \|^2$$

where $G^O \in \mathbb{R}^{M\times m}$, $G^k \in \mathbb{R}^{|F_k|\times n_k}$ ($k = 1\dots N$) are the cluster indicator matrices, and $S^k \in \mathbb{R}^{m\times n_k}$ is the cluster association matrix providing the relation between the clusters of objects and the clusters of each feature space. The factorization algorithm consists in an Expectation-Maximization approach that iteratively updates matrices G^O, G^k and S^k ($k = 1\dots N$).

Additionally, the approach computes new word-document and document-category matrices by incorporating user provided constraints through simultaneous distance metric learning and modality selection. This method is shown to be effective, but its formulation is not flexible. In fact, the number of clusters for each feature space is given as a parameter. Furthermore, the number of parameters grows with the number of feature spaces.

2.2 Information-theoretic approaches

The Information-theoretic co-clustering problem on star-structured data was first considered in [18] where Gao et al. propose to adapt the Information Theory co-clustering approach [17] to star-structured data. It consists in optimizing a weighted combination of mutual information evaluated over each feature space, where weights are chosen based on the supposed reliability/relevance of their correlation.

In the Information-theoretic approaches partitions X and Y^k ($k = 1 \dots N$) are defined as discrete random variables. Each variable $Y^k \in \mathscr{Y}$ has n_k categories $Y_1^k, \cdots, Y_{n_k}^k$, corresponding to n_k feature clusters, with probabilities $q_1^k, \dots, q_{n_k}^k$ and X has m categories $X_1, \cdots, X_m$ corresponding to m object clusters. However, for each variable Y^k, the m categories of X have different probabilities $p_1^k, \cdots, p_m^k$, $k = 1 \cdots N$. Probabilities p_i^k and q_j^k are computed as follows:

$$p_i^k = \frac{\sum_{O^s \in X_i} \sum_t r_{st}^k}{\sum_s \sum_t r_{st}^k}, \qquad q_j^k = \frac{\sum_{f_t^k \in Y_j^k} \sum_s r_{st}^k}{\sum_s \sum_t r_{st}^k}$$

The joint probabilities between X and any $Y^k \in \mathscr{Y}$ are denoted by t_{ij}^k, for $i = 1 \cdots m$ and $j = 1 \cdots n_k$ and are computed as follows:

$$t_{ij}^k = \frac{\sum_{O^s \in X_i} \sum_{f_t^k \in Y_s^k} t_{st}^k}{\sum_s \sum_t t_{st}^k}$$

Figure 1(b) provides an example of co-clustering computed on the two-space star-structured data depicted in Figure 1(a).

Following [29], the optimal Information-theoretic star-structured co-clustering is the one that minimizes:

$$D = \sum_{k=1}^{N} \alpha_k \left(I(\hat{Y}^k, \hat{X}) - I(Y^k, X) \right)$$

where $I(\hat{Y}^k, \hat{X}) = \sum_i \sum_j r_{ij}^k \log(\frac{r_{ij}^k}{p_i^k q_j^k})$ is the mutual information, $\hat{X}$ and $\hat{Y}^k$ are partitions where each cluster contains exactly one object/feature, $\alpha_k \geq 0 \ \forall k$ and $\sum_k \alpha_k = 1$.

The optimization approach is an adaptation of the ITCC algorithm [17]. Beyond the parameters inherited from the original algorithm, the weight α_k involved in the linear combination also has to be fixed by the end-user. Another drawback of this approach is its complexity, that prevents its use on large-scale datasets. Greco et al. [20] propose a similar approach based on the linear combination of mutual information evaluated on each feature space, where the parameter of the linear combination is automatically determined.

2.3 Probabilistic approaches

In [30], a parametric probabilistic approach to cluster relational data is proposed. A Monte Carlo simulation method is used to learn the parameters and to assign objects to clusters. The problem of clustering images described by segments and captions is considered in [10]. The proposed algorithm is based on Markov random fields in which some of the nodes are random variables in the combinatorial problem. Ramage et al. [33], propose a generative clustering algorithm based on latent Dirichlet allocation to cluster documents using two different sources of information: document text and tags. Each source is modeled by a probability distribution and a weight value is used to weigh one vector space with respect to the other. During the learning step, the algorithm finds the distribution parameters, and models documents, words and tags. In addition to the weight parameter, the method has another drawback: it constrains the number of hidden topics in text and tag sources to be the same, which is a strong assumption on data that is not always true.

2.4 Association-based approaches

In [22], Ienco et al. present a parameter-less iterative algorithm that maximizes the Goodman-Kruskal τ, a statistical measure of association that automatically identifies a congruent number of high-quality co-clusters. We provide in-depth details of this approach because it is parameter-less, i.e., contrary to the other approaches, it does not require a user-defined number of clusters. Goodman and Kruskal τ measure [19] is one of them that estimates the association between two categorical variables X and Y by the proportional reduction of the error in predicting X knowing or not the variable Y:

$$\tau_{X|Y} = \frac{e_X - E[e_{X|Y}]}{e_X}$$

Evaluating the quality of the partition of objects, given the partitions of features, is formalized as follows. The partition of objects is considered as the dependent variable X, and the N partitions of the feature spaces are considered as many independent variables $\mathscr{Y} = \{Y^1, \ldots, Y^N\}$. X and Y are defined as for the Information-theoretic co-clustering setting.

The error in predicting X is the sum of the errors over the independent variables of $\mathscr{Y}$: $e_X = \sum_{k=1}^{N} \sum_{i=1}^{m} p_i^k (1-p_i^k) = N - \sum_{k=1}^{N} \sum_{i=1}^{m} (p_i^k)^2$. $E[e_{X|\mathscr{Y}}]$ is the expectation of the conditional error taken with respect to the distributions of all $Y^k \in \mathscr{Y}$:

$$E[e_{X|\mathscr{Y}}] = \sum_{k}^{N} \sum_{j}^{n_k} q_j^k e_{X|Y_j^k} = \sum_{k}^{N} \sum_{j}^{n_k} q_j^k \sum_{i}^{m} \frac{t_{ij}^k}{q_j^k} (1 - \frac{t_{ij}^k}{q_j^k}) = N - \sum_{k}^{N} \sum_{i}^{m} \sum_{j}^{n_k} \frac{({t^k}_{ij})^2}{q_j^k}$$

The generalized Goodman-Kruskal's $\tau_{X|\mathscr{Y}}$ association measure is then equal to:

$$\tau_{X|\mathscr{Y}} = \frac{e_X - E[e_{X|\mathscr{Y}}]}{e_X} = \frac{\sum_k \sum_i \sum_j \frac{({t^k}_{ij})^2}{q_j^k} - \sum_k \sum_i (p_i^k)^2}{N - \sum_k \sum_i (p_i^k)^2} \tag{1}$$

If we consider Y^k as a dependent variable, and X as an independent variable, the corresponding $\tau_{Y^k|X}$ is computed as follows:

$$\tau_{Y^k|X} = \frac{e_{Y^k} - E[e_{Y^k|X}]}{e_{Y^k}} = \frac{\sum_i \sum_j \frac{(t_{ij}^k)^2}{p_i^k} - \sum_j (q_j^k)^2}{1 - \sum_j (q_j^k)^2} \tag{2}$$

The adopted co-clustering approach for star-structured data is formulated as a multi-objective combinatorial optimization problem which aims at optimizing $N+1$ objective functions based on Goodman-Kruskal's τ measure. The main procedure of the algorithm is sketched in Figure 2. The reader may refer to [22] for further algorithmic details.

```
Input: a star-structured dataset 𝒮𝒟 and an integer N_iter
Output: a coclustering (X, 𝒴)
  Initialize Y^1, ··· , Y^N, X with discrete partitions
  i ← 0
  T ← ∅
  for k = 1 to N do
    T^k ← ContingencyTable(X, Y^k, SD^k)
    T ← T ⋃ T^k
  end for
  while (i ≤ N_iter) do
    [X, T] ← OptimizeMultiObjectCluster(X, 𝒴, T)
    for k = 1 to N do
      [Y^k, T^k] ← OptimizeFeatureCluster(X, Y^k, T^k)
    end for
    i ← i + 1
  end while
  return Y^1, ··· , Y^N, X
```

Fig. 2 Pseudo-code of the adopted star-structured co-clustering algorithm [22].

To provide a first candidate list of objects to be recommended, one can measure the *cosine similarity* of each user vectors associated to the k-th space, with the centroids of each object clusters in the k-th space. Let x_i^k be the centroid of cluster X_i in the feature space F^k. The t-th component of x_i^k is computed as:

$$x_i^k = \frac{\sum_{O^s \in X_i} d_{st}^k}{|X_i|}$$

and the cosine similarity between u^k and x_i^k is evaluated as

$$sim(u^k, x_i^k) = \frac{u^k \cdot x_i^k}{\|u^k\| \|x_i^k\|}.$$

For each space, the most similar object cluster is chosen leading to a set of N clusters $\mathscr{X}^c = \{X_1^c, \ldots, X_N^c\}$ of candidate objects. Then, two different strategies can be adopted to provide the pre-filtered list of candidate objects $\mathscr{O}^c$:

- **relaxed strategy**: the objects belonging to the union of all clusters are retained, i.e.,
$$\mathscr{O}^c = \bigcup_k X_k^c$$
- **strict strategy**: the most represented cluster in $\mathscr{X}^c$ is retained, i.e.,
$$\mathscr{O}^c = \underset{X_k^c \in \mathscr{X}^c}{\operatorname{argmax}} |X_l^c \in \mathscr{X}^c \;\; s.t. \;\; X_k^c \equiv X_l^c| \,.$$

The first strategy is suitable when user's vectors are associated to very small clusters (e.g., because the user likes very uncommon objects). In any other situation, the second strategy is the most appropriate. As an additional step, objects already visited/liked/browsed by the user can be filtered out. We do not filter-out these objects at the beginning of the pre-filtering stage because they are relevant for the co-clustering step. In fact they are likely to be involved in important cross-associations between sets of features and sets of objects.

Finally, provided that each object in $\mathscr{O}$ is geo-referenced, the set of candidate objects $\mathscr{O}^c$ issued by the above-described process can be further refined by an ordering step. To this purpose, we employ the route distance between the user's current position and the position of each object in $\mathscr{O}^c$. Closer objects are on top of the items' list, while more distant ones are on its bottom. In conclusion, at the end of the pre-filtering stage, we provide an ordered list of candidate objects $\hat{\mathscr{O}}^c$ grouped by the related cultural POI (in this manner a user can easily choose items coming from more different cultural POIs).

3 Delivery Strategies for Multimedia Recommendation

Content-based filtering, collaborative filtering, as well as co-clustering based recommendation techniques described in the previous section help identifying, the multimedia items a user might be interested in. Suitable delivery strategies are then applied to deliver the identified multimedia objects (ranked in decreasing order of expected interest for the user), in order to fit to the best the user's requests. Contents are adapted to the user's request relying on contextual information, such as the location the users are in, the device they are using to retrieve the information of interest, their profile and their search history. Delivery strategies can overcome several drawbacks of common approaches of the classical multimedia recommendation systems. In fact, although the users' requirements are often expressed in terms of high level descriptions of the desired contents, it is not always possible to automatically extract meaningful high level information from multimedia features, and directly use such features in the recommendation algorithms. Thus, using context information can help increase the performances of recommendation systems by filtering out those items that do not match the user needs, in the given context. Also, for some kinds of multimedia data there does not exist a precise correlation between high and low level features (e.g. in images the concept of "moon" is related to a region with a circular shape and white color with a given uncertainty). It is important to understand the semantic of the users query, and rely on it to strengthen or weaken the ranking of the objects identified as interesting by the content based recommendation systems.

Users' preferences and feedbacks are not always explicitly known and available. This is especially the case when the multimedia system does not require a registration – with the specification of profiling information – from the users. Contextual information can help estimating the users' preferences at the query time, maybe also taking into account the features of the objects the user is currently observing. For example, the main colors of the painting the user is watching can give hints on the corresponding artistic movement or school, and can be taken into account when identifying other paintings to suggest. In the next subsections, we will provide a brief survey of the most common approaches used to define the delivery strategies.

3.1 Context-Based delivery strategies

There are many definition of what is a context. [1] defines the context as " 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 for the interaction between a user and an application, including the user and the application themselves." .

In a recent survey ([11]) different context modelling approaches, including object-role based models, spatial models, and ontology-based models, are presented, together with a discussion on how context-based reasoning can enhance the quality of many services.

Many multimedia recommendation platforms use context information for delivering multimedia services. For example, in [42] the authors describe a system that can handle many context categories like user preferences, situation and capability context. Their model takes into account an ontology which delineates attributes like user situation and user media preferences using namespaces and concepts derived from MPEG-7 descriptors of the media metadata. Recently, in [7] a multimedia recommendation system is presented that extends the classical recommendation delivery strategies by supporting useful *context-aware* services (e.g. a multimedia touristic guide). Such services are developed to assist users when visiting cultural environments (indoor museums, archeological sites, old town centers) containing several *cultural Points Of Interest* (e.g. paintings of museum rooms, buildings in ancient ruins or in an old town center, etc.) correlated with a large amount of multimedia data available in multiple web repositories. It has been shown in real case studies both in outdoor and indoor scenario, that this approach is successful, in terms of both user's satisfaction and system accuracy. Among the hybrid solutions, the uMender system [39] exploits context information, musical content and relevant user ratings to perform music recommendations on mobile devices. A framework for recommendation of multimedia objects based on processing of individual ontologies with context information is proposed in [23]: the recommendation process takes into account similarities calculated both between objects' (metadata) and users' ontologies, which reflect the social and semantic features existing in the system. Smart TV system are multimedia systems which easily present different types of multimedia content to end-users. Recently, some of those systems have also developed personalization techniques to recommend the most suitable content to users, exploring approaches from content-based user modeling to group-based collaboration. [28] describes a smart TV system with several unique features such as a Kinect-based component to recognize human body gestures for TV control, social tags and various environmental situations to annotate multimedia items and mprove users recommendation according. In [40], a useful taxonomy for mobile multimedia recommender systems has been presented, which is based on context-aware services. The classification of those systems is based on the type of collected information(explicit or implicit feedback), the type of recommendation learning process, and the algorithms used to make prediction/recommendation.

3.2 Location-Based delivery strategies

Due to the recent increase in availability of powerful mobile devices, location-aware systems are becoming more widespread, and recommendation systems are used to find interesting events, places, objects that are close to users' locations. [12] provides a survey of different location-based recommendation systems. An example showing the need for location-aware delivery systems is the case in which users rate cultural points of interests using multiple different features, not including the distance at the time of voting between the users and the point of interest they are voting. However, users who are relying on mobile personalised touristic guides, expect that the ranking in the recommendation of a point of interest takes into account not only the good ratings from similar users, but also the distance between their current position and the considered point of interest. Location awareness might imply a reordering of the rankings of a set of interesting points in such a way that a driving path reaching all the potentially interesting nearby points can be identified. In the context of multimedia data recommendation, [24], presents a system to recommend music well suited for points of interest (POIs). The considered scenario consists of a mobile city guide based on an enhanced presentation of places of interest for tourists, in which music related to the each point of interest being described (i.e., music that is culturally or emotionally associated with the place) is played. Similar systems are presented in [15], where the proposed models take into account the influence of online music social trends on users' local preferences. [38] describes a geospatial model taking into consideration GPS coordinates and semantic locations (continent, country, and state) of the user. In [41] a recommender system is described that correlates viewable scene information from sensors with geographic contextual tags from OpenStreetMap. The co-occurence of geo-tags and mood tags is computed based on a set of categories of the web site "Foursquare.com" and a mapping from geo-tags to mood tags is obtained. The music retrieval component returns music based on matching mood tags.

A special case of location based recommendation is the one referring to the concept of "smart space", as defined in [37]. In this case multimedia systems that are delivering recommended content as the interactive TV applications encapsulate both the information in a physical space as well as the information about the access to this information. This kind of location becomes a dynamic environment that changes over time, reflecting the way the different entities interact with it to share information among them.

3.3 Delivery Strategies based on Devices Features

In the definition of the delivery strategy for the recommendation algorithms, the effect of the device on which the selected media will be actually played has a great importance. In fact, users access to multimedia systems using different devices, including desktop or laptop computers, smart phones, tablets, etc. Each devices has its own interface characteristics (e.g., display capability), its specific Internet connection parameters, including cost and upload/download speed , and different storage space and computational capability. These differences have in impact on the user behavior and his preferences; for example, when we use a cellular phone we could prefer to download a lighter multimedia content than when we use a cabled device. Thus, multimedia content delivery should be adapted to the different devices. To face this problem Rosaci et al. in [36] have proposed a multimedia web service whose architecture allows to compute multi-device context-aware recommendations using an agent-based system. On the other hand, it is possible to introduce device adaptation features in the specification of multimedia documents, to support the delivery of different versions of the same document on different devices, taking into account the device characteristics. In this line, in [27] the authors describe a framework for standard multimedia documents based on an abstract structure that captures the spatio-temporal and hypermedia dimensions of multimedia documents, and propose an algorithm which transforms (in a minimal way) such multimedia documents to satisfy the presentation device constraints.

3.4 Profile-based delivery strategies

One of the classical modules in recommendation systems is the user profiling module, which learns (or at least estimates) users' interests over a long period of time, by analysing users' history, their inputs and/or their relationships . Most state-of-the-art user profiling approaches are based on the textual content of relevant documents to identify these interests. Hopfgartner et al. in [21] exploited the Linked Open Data Cloud to identify similar news stories that match the users interest to support the intelligent delivering of multimedia news. Albanese et al. in [3] described a multimedia recommendation system which combines the intrinsic features of multimedia objects, past behaviour of individual users, and overall behaviour of the entire community of users resembling the well-known PageRank ranking strategy. Konstas et al. in [26] used the additional relationships in a social network as user profile to develop a track recommendation system, thus taking into account both the social annotation and friendships inherent in the social graph established among users, items and tags, in order to create a collaborative recommendation system that effectively adapts to the personal information needs of each user.

4 Conclusion

The need of managing, retrieving and presenting multimedia information on the web has promoted the development of advanced multimedia information systems, which include recommendation modules to account for the requests of personalised data selection and presentation. Multiple approaches have been proposed in the literature to estimate users' degree of interest for the different available data. In this chapter we have presented the *co-clustering based recommendation techniques*, which allow to combine heterogeneous multimedia content information and data about the users' preferences and rankings, thus overcoming some of the content based filtering drawbacks, as well as some collaborative filtering weaknesses. Then, we briefly discussed the challenges in multimedia delivery and the most common strategies adopted in the context of cultural heritage media delivery.

References

1. Abowd, G.D., Dey, A.K., Brown, P.J., Davies, N., Smith, M., Steggles, P.: Towards a better understanding of context and context-awareness. In: Proceedings of the 1st International Symposium on Handheld and Ubiquitous Computing, HUC '99, pp. 304–307. Springer-Verlag (1999)
2. Adomavicius, G., Tuzhilin, A.: Toward the next generation of recommender systems: a survey of the state-of-the-art and possible extensions. IEEE Transactions on Knowledge and Data Engineering **17**(2), 734 – 749 (2005)
3. Albanese, M., d'Acierno, A., Moscato, V., Persia, F., Picariello, A.: A multimedia recommender system. ACM Transaction Internet Technology **13**(1) (2013)
4. Anagnostopoulos, A., Dasgupta, A., Kumar, R.: Approximation algorithms for co-clustering. In: Proceedings of the twenty-seventh ACM SIGMOD-SIGACT-SIGART symposium on Principles of database systems, PODS '08, pp. 201–210. ACM, New York, NY, USA (2008)
5. Anand, S.S., Kearney, P., Shapcott, M.: Generating semantically enriched user profiles for web personalization. ACM Trans. Internet Technol **7**(4) (2007)
6. Banerjee, A., Dhillon, I., Ghosh, J., Merugu, S., Modha, D.S.: A generalized maximum entropy approach to bregman co-clustering and matrix approximation. J. Mach. Learn. Res. **8**, 1919–1986 (2007)
7. Bartolini, I., Moscato, V., Pensa, R., Penta, A., Picariello, A., Sansone, C., Sapino, M.: Recommending multimedia visiting paths in cultural heritage applications. Multimedia Tools and Applications Journal pp. 1–30 (2014)
8. Bartolini, I., Zhang, Z., Papadias, D.: Collaborative filtering with personalized skylines. IEEE Trans. on Knowl. and Data Eng **23**(2), 190 – 203 (2011)
9. Basilico, J., Hofmann, T.: Unifying collaborative and content-based filtering. In: Proceedings of the Twenty-first International Conference on Machine Learning, ICML '04. ACM (2004)
10. Bekkerman, R., Jeon, J.: Multi-modal clustering for multimedia collections. In: Computer Vision and Pattern Recognition, IEEE Computer Society Conference on Vision and Pattern Recognition, pp. 1–8 (2007)
11. Bettini, C., Brdiczka, O., Henricksen, K., Indulska, J., Nicklas, D., Ranganathan, A., Riboni, D.: A survey of context modelling and reasoning techniques. Pervasive and Mobile Computing **6**(2), 161 – 180 (2010). Context Modelling, Reasoning and Management
12. Bobadilla, J., Ortega, F., Hernando, A., Gutiarrez, A.: Recommender systems survey. Knowledge-Based Systems **46**(0), 109 – 132 (2013)
13. Chakrabarti, D., Papadimitriou, S., Modha, D.S., Faloutsos, C.: Fully automatic cross-associations. In: KDD, pp. 79–88 (2004)
14. Chen, Y., Wang, L., Dong, M.: Semi-supervised document clustering with simultaneous text representation and categorization. In: ECML/PKDD (1), pp. 211–226 (2009)
15. Cheng, Z., Shen, J.: Just-for-me: An adaptive personalization system for location-aware social music recommendation. In: Proceedings of International Conference on Multimedia Retrieval, p. 185. ACM (2014)
16. Cho, H., Dhillon, I.S., Guan, Y., Sra, S.: Minimum sum-squared residue co-clustering of gene expression data. In: Proceedings of SIAM SDM 2004 (2004)
17. Dhillon, I.S., Mallela, S., Modha, D.S.: Information-theoretic co-clustering. In: Proc. ACM SIGKDD 2003, pp. 89–98. ACM Press, Washington, USA (2003)
18. Gao, B., Liu, T.Y., Ma, W.Y.: Star-structured high-order heterogeneous data co-clustering based on consistent information theory. In: ICDM '06: Proceedings of the Sixth International Conference on Data Mining, pp. 880–884 (2006)
19. Goodman, L.A., Kruskal, W.H.: Measures of association for cross classification. Journal of the American Statistical Association **49**, 732–764 (1954)
20. Greco, G., Guzzo, A., Pontieri, L.: Co-clustering multiple heterogeneous domains: Linear combinations and agreements. IEEE Transactions on Knowledge and Data Engineering **99** (2009)

21. Hopfgartner, F., Jose, J.M.: Semantic user profiling techniques for personalised multimedia recommendation. Multimedia systems **16**(4-5), 255–274 (2010)
22. Ienco, D., Pensa, R.G., Meo, R.: Parameter-free hierarchical co-clustering by n-ary splits. In: Proceedings of ECML/PKDD 2009, *Lecture Notes in Computer Science*, vol. 5781, pp. 580–595. Springer (2009)
23. Juszczyszyn, K., Kazienko, P., MusiaA, K.: Personalized ontology-based recommender systems for multimedia objects. In: Agent and Multi-agent Technology for Internet and Enterprise Systems, Studies in Computational Intelligence, pp. 275–292. Springer-Verlag (2010)
24. Kaminskas, M., Ricci, F., Schedl, M.: Location-aware music recommendation using auto-tagging and hybrid matching. In: Proceedings of the 7th ACM Conference on Recommender Systems, RecSys '13, pp. 17–24. ACM (2013)
25. Keogh, E., Lonardi, S., Ratanamahatana, C.A.: Towards parameter-free data mining. In: Proceedings of the tenth ACM SIGKDD international conference on Knowledge discovery and data mining, KDD '04, pp. 206–215. ACM, New York, NY, USA (2004)
26. Konstas, I., Stathopoulos, V., Jose, J.M.: On social networks and collaborative recommendation. In: Proceedings of the 32Nd International ACM SIGIR Conference on Research and Development in Information Retrieval, SIGIR '09, pp. 195–202. ACM (2009)
27. Laborie, S., Euzenat, J., Layaïda, N.: Semantic adaptation of multimedia documents. Multimedia tools and applications **55**(3), 379–398 (2011)
28. Lee, W.P., Kaoli, C., Huang, J.Y.: A smart {TV} system with body-gesture control, tag-based rating and context-aware recommendation. Knowledge-Based Systems **56**(0), 167 – 178 (2014)
29. Long, B., Zhang, Z.M., Wú, X., Yu, P.S.: Spectral clustering for multi-type relational data. In: ICML '06: Proceedings of the 23rd international conference on Machine learning, pp. 585–592 (2006)
30. Long, B., Zhang, Z.M., Yu, P.S.: A probabilistic framework for relational clustering. In: KDD '07: Proceedings of the 13th ACM SIGKDD international conference on Knowledge discovery and data mining, pp. 470–479 (2007)
31. Pazzani, M.J., Billsus, D.: Content-based recommendation systems. In: The adaptive web. Springer-Verlag (2007)
32. Pensa, R.G., Boulicaut, J.F.: Constrained co-clustering of gene expression data. In: Proceedings of SIAM SDM 2008, pp. 25–36 (2008)
33. Ramage, D., Heymann, P., Manning, C.D., Garcia-Molina, H.: Clustering the tagged web. In: WSDM, pp. 54–63 (2009)
34. Ricci, F., Rokach, L., Shapira, B.: Recommender Systems Handbook. Springer (2011)
35. Robardet, C., Feschet, F.: Comparison of three objective functions for conceptual clustering. In: Proceedings PKDD'01, *LNAI*, vol. 2168, pp. 399–410. Springer (2001)
36. Rosaci, D., Sarno, G.: Recommending multimedia web services in a multi-device environment. Information Systems **38**(2), 198 – 212 (2013)
37. Saleemi, M., Lilius, J.: Exploiting smart spaces for interactive tv applications development. The Journal of Supercomputing pp. 1–18 (2014)
38. Schedl, M., Vall, A., Farrahi, K.: User geospatial context for music recommendation in microblogs. In: Proceedings of the 37th International ACM SIGIR Conference on Research and Development in Information Retrieval, SIGIR '14, pp. 987–990. ACM (2014)
39. Su, J.H., Yeh, H.H., Yu, P.S., Tseng, V.S.: Music recommendation using content and context information mining. IEEE Intelligent Systems **25**(1), 16–26 (2010)
40. Xia, F., Asabere, N., Ahmed, A., Li, J., Kong, X.: Mobile multimedia recommendation in smart communities: A survey. CoRR **1**, 606–624 (2013)
41. Yu, Y., Shen, Z., Zimmermann, R.: Automatic music soundtrack generation for outdoor videos from contextual sensor information. In: ACM Multimedia, pp. 1377–1378 (2012)
42. Yu, Z., Zhou, X., Zhang, D., Chin, C.Y., Wang, X., men, J.: Supporting context-aware media recommendations for smart phones. Pervasive Computing, IEEE **5**(3), 68–75 (2006)