Personalization Based on Domain Ontology

Mehdi Adda¹, Wu Lei², Yi Feng³

¹Department of Computer Science and Operations Research, University Of Montreal, QC Canada ²Department of Software Engineering, Rochester Institute of Technology, NY USA ³Computer Science Department, Algoma University, ON Canada

Abstract

As a consequence of the proliferation of multimedia contents, users are nowadays frustrated with the huge amount of available video information whose content is not targeted to their needs and preferences. Its challenging to analysis video content for video personalization due to the lack of semantic video summarization and retrieval techniques. In fact, most of current video personalization systems are using low-level features. However, users identify and select video content using high-level semantics. This creates a gap between user preferences and video content representation that must be bridged for video personalization systems. In this paper we present a new approach for video personalization based on domain knowledge. We first introduce an ontology based indexation approach to enhance retrieval performance. Then, we present a personalization strategy based on fine grained sequential pattern discovery. The proposed approach is based on both user and content personalization. The performance study and experiments show that the use of ontologies to index and represent video contents enhance running time and memory performances. This paper also describes VideoMiner, a system prototype that implement the proposed approach for video personalization.

Key-Words : Video personalization, Data mining, Domain knowledge, Ontologies.

1 Introduction

One of the most interesting topics in video research, as well as one of the most important challenges in video mining is the reasoning on video content and users at a semantic level and then deliver for each user a personalized content. It is essential to the success of a video based system such us interactive TV, video on demand, etc,. Indeed, from an e-commerce perspective, it is vital to have the ability to retrain visitors and turn casual browsers into potential users and customers. Video personalization systems, generally based on an unsupervised discovery of patterns in a defined input data [13], aims at tailoring video content retrieval based on user's past behavior and inference from other like-mind users to anticipate the needs of a user and provide a customized content [15]. The basis of video personalization includes modeling of video content and users, characterization of user content browsing and access experiences, matching between and across videos and/or users, and determination of the set of videos and/or video-parts to be recommended.

Traditional approaches to personalization include contentbased and user-based techniques [1]. In a content-based system, the objects are recommended according to the similarity of their associated features. User-based system compares current users past actions with the historical records of other users in order to find the users with similar interests and then recommend the objects they have visited. However, some problems arise from those techniques. On one hand, content-based systems suffer from the new user problem, consisting of the difficulty to find objects of interest for new user. On the other hand, user-based systems suffer from the new item problem, consisting on the difficulty to recommend new objects that has not been visited or has not had many ratings. To deal with those problems, and improve performance, hybrid techniques combining contentbased and user-based approaches are proposed. The scalability problem, generally encountered in personalization systems due to the huge amount of users and objects, is tackled by the application of data mining techniques.

Traditional personalization techniques can not be directly applied in video personalization systems. Indeed, it is challenging to apply such techniques due to the complexity level of video objects and contents. Representing a video as a whole and single object is not suitable for capturing complex relationships among video content at a deeper semantic level.

In this paper we present a video personalization system where videos are recommended according to the objects contained within. Indeed, in the proposed system, a video is represented according to its content by using domain concepts and the chronological order of apparition of such concepts is respected. The approach is based on a hybrid recommendation method where both content-based and userbased recommendation techniques are combined. The work is rooted in practical concerns from the data mining: we have studied the extraction of domain related behavioral patterns from server log files, and their application to video personalization. The data is therefore made of transformed video sequences. In the sequence transformation process, videos have to be represented according to fine grained concepts to preserve their semantics and ease interpretation.

However, videos are generally partitioned into shots and frames, and by then, low-level features such as motion, color, and texture are extracted to form MPEG7 descriptions [12]. In the next two paragraphes, we discuss some drawbacks and limitations of using low-level features and MPEG7 file descriptors.

Considering only low-level similarities ignores the domain related similarities among video content. Consequently, from the recommendation perspective, the proposed videos may have considerable variances both in semantics and visual content and, therefore, do not make much sense to human perception. Moreover, low-level features are very difficult to comprehend and interpret. As a result the semantic gap between user needs and the personalization system is increased.

Using independent file descriptors for each video is a basic component in video mining and personalization frameworks. In most practical cases the number of MPEG7 description files is huge, and they are individually accessed and parsed. On one hand, handling different files makes the task of unifying terms and semantics above a video collection very difficult, and thus the semantic gap is going to be increased, especially in cases where the video content is heterogenous and videos are indexed by different persons without a predefined and unified set of domain terms and vocabulary. In those cases, different descriptors may be used to describe the same content of different videos. On the other hand, using one file for metadata description per video is time consuming. Since accessing and handling files in huge data is time consuming-the more of its video description files is, the higher the processing cost will be. In fact, elementary video operations such as indexing, retrieval, management require the manipulation of the associated metadata, which if dispersed in many files -event thousands- affect the system efficiency. One suitable solution for those two problems consists of using unified vocabulary and descriptors for all videos, and use one file to represent and index all videos. Such a unified vocabulary is provided by domain Ontologies. A domain ontology offers the necessary conceptualization model of video content for the specific application area. With means of ontologies, the semantic gap between video content and human perception can be bridged, and domain ontology can be used as a unique source for video content indexation and retrieval. Our paper, first, addresses issues of developing a new hybrid approach for video recommendation that offers a fine grained, and domain related personalization by the mean of an OWL domain ontology [8]. Secondly, a personalization system prototype based on the proposed approach is presented.

Our main goal, when developing the hybrid approach, is to create a model of user behaviors through the seamless integration of semantic knowledge with user access sequences (provided by means of log files). The profiling process consist of discovering fine grained video sequential patterns. Once frequent patterns are extracted, a recommendation algorithm seeks for most relevant videos to be presented to the current user by taking into account the last visited videos.

The design framework and system prototype is a three-tier architecture of server, middleware, and client. The server maintains the video content and metadada descriptors consisting of the domain ontology. The client communicates user query to retrieve content, and displays the personalized content. The middleware consists of a set of modules that aim at mining and maintaining usage patterns, index videos according to the domain ontology, select a personalized content that correspond to user behavior and past experiences, and offer an alternative browsing technique : ontology based video browsing and retrieval.

In this paper we assume that videos are indexed by an expert in the domain of application. Ideally the indexation process should be produced automatically. To this end, we have to achieve automated analysis of the semantic contents of videos. However, since highly reliable analysis of such contents is very difficult at present, we assume the process is done manually. In our experiments, we use the open platform Protégé [16] for video indexation based on an OWL domain ontology.

The rest of the paper is organized as follows. Section 2 provides the theoretical basis and definitions underlying our approach. Section 3 presents architecture and design of VideoMiner. Section 4 presents experimental results. Section 5 gives an overview of related work on video personalization. Concluding remarks and discussions on future work are given in Section 6.

2 Ontology-Based Video Personalization Strategy

The problem of Ontology-Based Personalization can be divided into three subproblems as follows :

- 1. Ontology-Based Video Indexation,
- 2. Mining fine grained video access patterns,
- 3. Video recommendation based on domain knowledge and fine grained access patterns.

2.1 Ontology-Based Video Indexation

Video indexation is important for video retrieval and personalization. Ontologies are typically represented using natural language domain terms, and thus are suitable for video annotation and indexation. In fact, linguistic and domain terms are appropriate to semantically represent video content (events and objects). VideoMiner system uses an ontology to describe semantic contents such as objects and events occurring in a video. Video clips and images are index using the concepts and relations of the OWL ontology. Ideally the indexation process should be produced automatically. To this end, we have to achieve automated analysis of the semantic contents of videos. However, since highly reliable analysis of such contents is very difficult at present, we assume the process is done manually. In our experiments, we use Protégé for video indexation according to a domain ontology.

2.2 Fine grained video access patterns

Once enough user access sessions are collected, access sequences containing the consulted videos are extracted. Our objective is to capture the user interest by semantically exploring the objects forming those videos. In other words, we are looking to find domain related similarities on the video sequences. To achieve it, the initial sequences are transformed to fine grained sequences called video domain sequences (VDS).

In the next subsections we present a formal description of input data. Then, we present the transformation process, a formalism for VDS representation, with new notions and definitions. By then, we give the basis of sequential patterns mining method.

2.2.1 Description of data and patterns

The format for initial data is a usual video sequence S, i.e, a list of videos that have been consulted by a user. Let V =

 $\{v_1, v_2, ..., v_n\}$ be a set of videos. Each video v_i is composed of a list of clips such us $v_i = \langle cl_i^1, cl_i^2, ..., cl_i^m \rangle$, where a clip cl_i^j is composed of a list of images. In other words $cl_i^j = \langle im_{i,1}^j, im_{i,2}^j, ..., im_{i,p}^j \rangle$. An image $im_{i,k}^j$ is represented by a set of domain related ontology concepts. Let an ontology, say Ω , with its concepts set T_{Ω} , an image representation belongs to $T_{\Omega}^w = (T_{\Omega} \times T_{\Omega} \times T_{\Omega}...)$.

A video domain sequence S_c is obtained by replacing each image of each clip of each video of S by the list of domain concepts involved in this image. The set of concepts involved in an image are put together in a set called conceptset. A conceptset have to be non-empty. Thus, a video domain sequence is an ordered list of conceptsets. Figure 1 shows an example of transformation process from video based sequence to domain based sequence.

We denote a VDS by $\langle s_1, s_2, ..., s_n \rangle$, where s_i is a conceptset also designed by S(i). We also call s_i an element of the VDS sequence. The length of a VDS S is the number of its elements and is denoted by |S|. We denote an element of a sequence by $\{c_1, c_2, ..., c_m\}$, where c_i is a concept of T_{Ω} . A concept can occur many times in different conceptsets but can occur only once in a conceptset.



Figure 1: Domain ontology and VDS generation example.

Given a collection of sequences Γ , the support of a sequence S, denoted by sup(S), is the number of sequences in Γ that are generalized by S (see Definition 2.1). Conventionally, a sequence S is called a sequential pattern if $sup(S) \leq minsup$, where minsup is a user-fixed minimum support threshold.

Definition 2.1 Video sequence generalization : A sequence S_1 is a generalization of a sequence S_2 , denoted either by $S_1 \succeq_{\Omega} S_2$ or by $S_2 \preceq_{\Omega} S_1$, if there exist a set $\mathcal{I} \subseteq [1, S_1]$ and a surjective monotonously non-decreasing map $\psi : \mathcal{I} \to [1..|S_2|]$ such that: $\forall i \in \mathcal{I}, S_2(\psi(i)) \sqsubseteq_c S_1(i)$, where \sqsubseteq_c is a subsumption relation among conceptsets (see Definition 2.2 bellow).

Definition 2.2 Conceptset subsumption : A conceptset $s_1 = \{x_1, x_2, ..., x_n\}$ subsumes a conceptset $s_2 = \{y_1, y_2, ..., y_m\}$ denoted $s_2 \sqsubseteq_c s_1$ if : $\forall x_i \in s_1, \exists y_j \in s_2 : y_j \sqsubseteq_\Omega x_i$. Where \sqsubseteq_Ω is the symbol of subsumption between two concepts of Ω [3].

2.2.2 Pattern mining

To mine frequent patterns, an Apriori based algorithm is used. *Apriori* is a classical algorithm used for sequential pattern mining [2]. In the following we describe our method that, similarly to Apriori, performs a top down level-wise search through the pattern space. This means, at the (k+1)th level, the algorithm uses frequent patterns generated at level k to compose candidates and then check their frequency. At the first pass, candidates are the sequences of size one made of the maximum concepts in the domain ontology. At later steps, the algorithm extends a frequent pattern either by adding a new element or extending an existing element with a concept. The algorithm terminates when there are no frequent sequences at the end of a pass, or when there are no candidate sequences generated.

2.3 Video recommendation strategy based on fine grained domain patterns

Recommendation systems are the most employed tools in e-commerce businesses. Recommendation systems help the users to rapidly and efficiently find the content they would like to consult. The basis of recommendation systems is to apply data analysis techniques to generate a list of recommended content products for each user according to his/her past behavior.

The video recommendation algorithm we developed VRAl is hybrid : both usage-based and content-based. VRAl (see Figure 2) suggests videos according to last visited videos and either to the set of VDS patterns generated in the previous step, or the domain ontology. The steps of VRAl are as follows : (i) obtain a VDS from the last visited videos, say S, (ii) select the VDS patterns that contain subsequence of S according to Definition 2.3, (iii) if the selected patterns have at least one conceptset, than recommend videos indexed with concepts involved in those conceptsets, (iv) if no VDS pattern matches or no concepts tremains we suggest videos indexed by the same concepts than those contained is S and which are not yet visited.

Definition 2.3 Subsequence, super-sequence : A sequence $S_1 = \langle X_1, X_2, ..., X_n \rangle$ is called a subsequence of $S_2 = \langle Y_1, Y_2, ..., Y_m \rangle$, and S_2 a super-sequence of S_1 , if there exist a set $\mathcal{I} \subseteq [1, S_1]$ and a surjective monotonously non-decreasing map $\psi' : \mathcal{I} \to [1..|S_2|]$ such that: $\forall i \in \mathcal{I}, S_1(i) \subseteq S_2(\psi(i))$.

3 Proposed Framework

This section gives an overview of *VideoMiner*, a video personalization system we have developed. The goal is

```
Algorithm 1 VRA1 : Video Recommendation Algorithm
 1: Ω : domain ontology
 2: \mathcal{C} \leftarrow \emptyset: // set of videos to recommend
 3: \Gamma_{\Omega} : set of all frequent patterns;
       : set of last visited videos
 5: \Gamma_1 \leftarrow \oslash; // \text{ set of selected patterns}
 6: CS \leftarrow getIndexedConceptSets(\Omega, V); //get the list
         of conceptsets involved in the indexation of \mathcal{V} ele-
         ments.
 7: \Gamma_1 \leftarrow getInvolvedPatterns(CS); //get the set of pat-
         terns where conceptsets of CS are involved with re-
         spect to the apparition order
 8: if Γ
            \neq \oslash then //Apply usage-based recommendation
        CS \leftarrow getRemCptSet(\Gamma_1) - CS; //get the set of
             conceptsets that are involved in the selected pat-
             terns and that did not figure in the conceptsets
             of the visited videos.
        \mathcal{C} \leftarrow getIndexedVideos(\Omega, CS); //get a list of
10:
             videos indexed in \Omega by the set of conceptsets
             CS.
11: else//Apply content-based recommendation
         \mathcal{C} \leftarrow getIndexedVideos(\Omega, \mathcal{CS}) - \mathcal{V};
12:
13: end if
14: return C
```

Figure 2: Video recommendation algorithm.

to provide a general idea of what is each module is for. VideoMiner architecture is basically implemented as a plug-in the open-source Protégé platform [16]. We have extended Protégé for video indexation, visualization and personalization using the approach we have proposed in this paper.

VideoMiner adopts a three-tier architecture composed of : client view and interaction interface, a middleware and of back-end server. View A of Figure 3 illustrates the block diagram of these components. The client communicates user query to retrieve content, and display the personalized content. The server maintains the video content and medatada descriptors consisting of the domain ontology. The middleware consists of a set of modules that aims to mine and maintain usage patterns, index videos according to the domain ontology, select a personalized content that correspond to user behavior and past experiences, offer an alternative browsing technique : ontology based video browsing and retrieval. The browsing technique exploits the underlaying ontology structure to navigate through the concepts, and by then, show the videos where those concepts are present.

3.1 System prototype

The system prototype is based on the system structure and architecture such as presented in View A of Figure 3. It mines maintains and stores fine grained sequential patterns by XML documents. The system contains 4 major components: (1) Video indexation, (2) Pattern extraction, (3) Video recommendation, and (4) Video browsing and retrieval. Figure 3 shows the graphical user interface of the implemented system.



Figure 3: Block diagram of VideoMiner.

4 Performance study and experimental results

To evaluate the effectiveness and efficiency of the use of ontologies to index and retrieve video content on personalization systems, we performed an experimental evaluation on real datasets composed of 1984 videos, same number of MPEG7 files, and an OWL file that indexes the whole videos. Two charts are reported : the first one related to the execution time necessary to parse description files, and the second one, presents the memory space allocated to video descriptors without the content. We consistently found that performances are better when OWL is used to describe and represent the video content than when MPEG7 files are used.

Indeed, on one hand, View A of Figure 4 shows the parsing time for different sets of MPEG7 files with the associated videos compared to the parsing time for the OWL file. On the other hand, View B of Figure 4 shows that the memory space dedicated to different sets of MPEG7 files compared to the memory space dedicated to the OWL file is higher. As it can be seen the cost of using OWL for video description and indexation compared to MPEG7 descriptors is pretty minimal.



Figure 4: Performance improvement.

5 Related works

Personalization and adaption of information to meet the users needs and interest is an active research area. Although

a lot of methods for personalization related to web personalization where content is either text or semi-structured data (e.g., XML) have been proposed. A survey of existing techniques and systems can be found in [15]. The proposed approaches cannot be directly applied to video personalization. Also some projects are research works have been investigating personalization of video and multimedia content. In [10] a personalization system based on static user preferences is presented. This system adapts the multimedia content using the user selected preferences. In [5] a dynamic personalization system is presented. User profiles are formed of weighted keywords reflecting user preferences. In [6] user preferences are dynamically acquired by investigating whether he/she is interested in the TV program or not. Babaguchi [4] discussed video abstraction based on its semantical content in the sports domain. A profile is formed of a pair of keywords and their weights. To select highlights of a game, an impact factor for a significant event in two-team sports was proposed. All the cited systems are using MPEG7 [12] standard for content and metadata descriptions. In [9, 11] ontologies are used to annotate videos. It can be pointed out that these methods are based on surface features of the video rather than on its semantical contents. In other words, instead of exploiting linguistic terms to form domain concepts, visual and low-level features extracted from video content are used.

To solve common problems related to scalability and efficiency, there had been a large utilisation of data mining techniques for video clustering and classification [14], pattern detection to characterize video events [17]video association mining for semantic indexing and event detection [18]. Some other works such us [7] use combinations of unsupervised and supervised learning techniques for highlights extraction.

6 Conclusion and future work

In this paper, we have proposed a new approach for video personalization, and we have presented an implementation of this approach. In this approach, we have used a domain ontology for both video indexation, retrieval, and pattern discovery, in order to enhance performances and reduce the semantic gap between user behavior and perception, and video content representation.

In the proposed approach, user preferences are extracted from the usage data, and are dynamically and automatically matched to domain knowledge. Such preferences are utilized together with the domain ontology towards the personalization of video content. The techniques discussed in this paper are based on a great extend on the utilization of sequential pattern techniques and ontology based knowledge representation. Furthermore, we have presented the design and architecture of VideoMiner prototype. VideoMiner is a video recommendation system based on fine grained sequential patterns. The system adopts a three-tier architecture composed of : client view and interaction interface, a middleware and of back-end server.

The proposed approach can be extended by taking into account ontology inte-concept relations in the mining process to capture complex relations among video content. By doing so, video domain sequences will have the structure of directed graphs for which new personalization methods have to be developed. The implemented prototype can be extended so that the system will be accessible via the Web with a light interface.

References

- G. Adomavicius and A. Tuzhilin. 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*, 17:734 – 749, June 2005.
- [2] R. Agrawal and R. Srikant. Fast algorithms for mining association rules. In Proc. 20th Int'l Conf. on Very Large Data Bases, VLDB94, 1994.
- [3] F. Baader, D. Calvanese, D. McGuinnessand D. Nardi, and P. Patel-Schneider. The description logic handbook : Theory, implementation and applications. *Cambridge University Press*, page 574, 2003.
- [4] N. Babaguchi, Y. Kawai, T.Ogura, and T. Kitahashi. Personalized abstraction of broadcasted american football video by highlight selection. *IEEE Transactions on Multimedia*, 6:575 – 586, August 2004.
- [5] N. Babaguchi, K. Ohara, and T.Ogura. Effect of personalization on retrieval and summarization of sports video. Proc. of the IEEE Joint Conf. of the Fourth Int'l Conf. on Information, Communications and Signal Processing, and the Fourth Pacific Rim Conf. on Multimedia, 2:940 – 944, December 2003.
- [6] B.Smyth and P.Cotter. A personalized television listings service. *Communication ACM*, 43(8):107–111, 2000.
- [7] A. Divakaran, K. Miyahara, K. A. Peker, R. Radhakrishnan, and Ziyou Xiong. Video mining using combinations of unsupervised and supervised learning techniques. Special Session on Video Mining, SPIE Electronic Imaging Conf. on Storage and Retrieval for Mcdia Databases, San lose, CA,, 2004.

- [8] T. R. Gruber. A translation approach to portable ontology specifications. *Current issues in knowledge modeling, Academic Press Ltd. London, UK*, pages 199 – 220, 1993.
- [9] A. Hakeem and M. Shah. Ontology and taxonomy collaborated framework for meeting classification. *Proc. of the Pattern Recognition*, 17th Int'l Conf. on (ICPR'04), 4:219 – 222, August 2004.
- [10] Rune Hjelsvold, Subu Vdaygiri, and Yves Léauté. Web-based personalization and management of interactive video. *Proc. of the tenth Int'l Conf. on World Wide Web*, April 2001.
- [11] L. Hollink and M. Worring. Building a visual ontology for video retrieval. *MULTIMEDIA '05: Proc. of the 13th annual ACM Int'l Conf. on Multimedia*, 4:479 – 482, 2005.
- [12] J. M.Martinez. Overview of the mpeg-7 standard (version 6.0). ISOIIEC JTCIISC29IWGll N4509, 2001.
- [13] B. Mobasher, X. Jin, and Y. Zhou. Semantically enhanced collaborative filtering on the web. Semantically Enhanced Collaborative Filtering on the Web, Proc. of the European Web Mining Forum, - 2004.
- [14] J. Pan and C. Faloutsos. Videocube: A novel tool for video mining and classification. *Proc. IntŠl Conf. Asian Digital Libraries (ICADL)*, pages 194–205, 2002.
- [15] D. Pierrakos, G.O Paliouras, C. Papatheodorou, and C.D. Spyropoulos. Web usage mining as a tool for personalization: A survey. User Modeling and User-Adapted Interaction 13: 311-372, Kluwer Academic Publishers, Hollande, 2003.
- [16] Stanford University. Protégé. http://protege.stanford.edu, 2005.
- [17] L. Xie, S.-F. Chang, A. Divakaran, and H. Sun. Unsupervised mining of statistical temporal structures in video. *Video Mining, A. Rosenfeld, D. Doremann, and D. Dementhon eds., Kluwer Academic*, 2003.
- [18] X. Zhu, X. Wu, A. K. Elmagarmid, Z. Feng, and L. Wu. Video data mining: Semantic indexing and event detection from the association perspective. *IEEE TRANSACTIONS ON KNOWLEDGE AND DATA EN-GINEERING*, 17:5, MAY 2005.