

# Video Recommendations for the Open Video Project

Johan Bollen  
Research Library  
Los Alamos National  
Laboratory  
Los Alamos, NM  
jbollen@lanl.gov

Michael L. Nelson and  
Raquel Araujo  
Department of Computer  
Science  
Old Dominion University  
Norfolk, VA  
{mln, raraujo}@cs.odu.edu

Gary Geisler  
Graduate School of Library &  
Information Science  
Simmons College  
Boston, MA 02115-5898  
gary.geisler@simmons.edu

## ABSTRACT

We describe a DL multimedia recommender system implemented for the Open Video project. Recommendations are generated by a spreading activation algorithm operating on a video network created from log download sequences. We compared the system's recommendations to those generated by a collaborative filtering technique.

## Categories and Subject Descriptors

H.3.7 [Information Systems]: Digital Libraries

## General Terms

Algorithms

## Keywords

Digital Library, recommender system, log data

## 1. INTRODUCTION

Collaborative Filtering (CF) algorithms [1] have shown to be effective tools for multimedia recommender systems. However, user behavior in a DL environment is often shaped by transitory information needs invalidating the use of long-term user profiles. Furthermore, many DLs can not practically or legally store user profiles. We implemented a recommendation system that generates document recommendations by a process of spreading activation on the basis of co-download patterns derived from download log data for the Open Video project's, [www.open-video.org](http://www.open-video.org), video repository [2].

## 2. IMPLEMENTATION

We analyzed 10,954 video downloads registered between July and November 2002 in the Open Video logs and extracted a set of video co-downloads. A co-download was defined as the same user downloading a pair of videos with a small temporal latency. A co-download network of 1320 video relationships was generated. A Java Servlet was implemented to receive a text query, map the query terms to the corresponding videos, and apply spreading activation to the generated video network. At the conclusion of this process, videos were returned to the user ranked according to their final activation values. Table 1 lists the relationships generated for the video entitled "San Francisco Earthquake Aftermath, 1906".

Copyright is held by the author/owner.  
JCDL'05, June 7–11, 2005, Denver, Colorado, USA.  
ACM 1-58113-876-8/05/0006.

San Francisco Earthquake Aftermath, 1906	
weight	Video Title
0.96	Hidden Fury, segment 09 of 11
0.49	Shock Troops of Disaster: Story of New England Hurricane
0.44	S. F. Earthquake Aftermath: Riding Down Market St.
0.17	Your Chance to Live: Earthquake
0.10	San Francisco Earthquake Aftermath

Table 1: Relationships generated for "San Francisco Earthquake Aftermath, 1906" video

The set of recommendations for the video "San Francisco Earthquake Aftermath, 1906" demonstrates the ability of the log analysis methodology to establish video relationship which reflect thematic and conceptual similarities. For example, the video "Shock Troops of Disaster: Story of New England Hurricane" does not share any title terms with the query video, but is strongly related to the common concept of "natural disasters".

## 3. EVALUATION

We implemented a base-line CF recommendation system which derived video profiles from the above mentioned Open Video log. A video profile consisted of the set of users that download the particular video. TFIDF weighing was used to determine the degree to which a user was associated with a particular video. Video similarities were then calculated on the basis of cosine similarities between video profile vectors [3] and used to generate CF recommendations. Precision and recall scores were determined for CF and SA recommendations on the basis of a set of selected videos. The SA system outperformed the CF system by a wide margin in terms of its precision as well as its recall.

## 4. REFERENCES

- [1] Jonathan L. Herlocker, Joseph A. Konstan, Al Borchers, and John Riedl. An algorithmic framework for performing collaborative filtering. In *Proceedings on the 22nd Annual International ACM SIGIR Conference on Research and Development in Information Retrieval*, pages 230–237, Berkeley, CA, August 15–19 1999.
- [2] Gary Marchionini and Gary Geisler. The Open Video Digital Library. *D-Lib Magazine*, 8(12), December 2002.
- [3] Badrul Sarwar, George Karypis, Joseph Konstan, and John Riedl. Item-based collaborative filtering recommendation algorithms. In *Proceedings of the tenth international conference on World Wide Web*, pages 285–295. ACM Press, 2001.