

Title	個人の興味を利用したコンテンツ集約モデルに基づく 情報取得に関する研究
Author(s)	高橋, 朋之
Citation	
Issue Date	2015-03
Type	Thesis or Dissertation
Text version	author
URL	http://hdl.handle.net/10119/12701
Rights	
Description	Supervisor:長谷川忍, 情報科学研究科, 修士

A Study for Information Retrieval based on Contents Aggregation Model with Personal Interests

Tomoyuki Takahashi (1110037)

School of Information Science,
Japan Advanced Institute of Science and Technology

February 12, 2015

Keywords: Personal Interest, Information Retrieval, Contents Aggregation, Advance Organizer, Interest Equalizer.

In modern societies, ICT (Information and Communication Technology) is used in every situation of our life, and the amount of information created have increased dramatically. Because everyone can transmit information as UGCs (User Generated Contents) at low cost by evolution of ICT like Web 2.0 technology and can obtain massive information from diverse resources. In these surroundings, we suffer from "information overload / explosion" that we cannot find necessary information since the pace of increasing information overtakes the improvement of information retrieval technology. Especially, form of watching video contents is greatly changed. High quality video platforms via network like YouTube, niconico.com become rapidly popular in the Web. For this reason, it is required to develop new functions to retrieve the video contents effectively.

The purpose of this research is to propose a method to retrieve the video contents that a user has interested from niconico.com. Niconico.com provides a video retrieval function which narrows the results by keywords, categories, submitted date, number of plays, evaluated rate and ranking. If the user has targeted information clearly, it would work well. On the other hand, it is difficult to add serendipity to the retrieval results. Niconico.com also has a function which provides similar or related videos. But, it is not enough to find interesting videos from vast amounts of database.

As video search, a folksonomy approach with tags possesses great potential to support user's new discovery of the video contents from multifaceted classifications such as matter of the contents and sympathy for the contents. However, there are some difficulties in search engines like tag spelling inconsistency and tag multi meanings. In this paper, I first propose a distillation model, which consists of 'Interest Organizer', 'Interest Equalizer' and 'Contents Aggregation', to represent user's interests clearly. Interest Organizer provides the user with a tag list from niconico.com and the user is required to select interesting tags from the list. Interest Equalizer gives the users a way of parameterization of his/her degree of interest for each tag selected in Interest Organizer. These parameters are used for the search engine ranking. Contents Aggregation is a model to gather and store the contents retrieved by Interest Equalizer. This is like 'my list' in niconico.com. I have also developed a prototype system with a video retrieval method reflected on the degree of user's interest and tag co-occurrence relations and consider a visualization method for retrieval results. In addition, I conducted a small pilot study by myself to confirm that the prototype system runs as planned and to improve the procedure of the future case study. I selected 5 data sets included in 'sports with any fields', 'sports with baseball, football, and figure skating', 'animations', 'music', and 'animals'. From the results, the prototype system roughly worked well. But, it took much time for the assumed subjects to finish the case study. I also found there may be some conditions to yield a good retrieval results from different types of data sets of niconico.com. In the near future, I will have to improve the proposed functions and to conduct a couple of case studies.