International Journal of Science and Research (IJSR)
ISSN (Online): 2319-7064
Index Copernicus Value (2013): 6.14 | Impact Factor (2015): 6.391

Personalized Image Search using Ontological Knowledge base User Profiles and Ontology Mining

A. N. R. Latha Kumari[1] , G. Pranitha[2]

[1, 2]Department of CSE, ANITS, Sangivalasa, India

Abstract: ***As the Internet users and the number of accessible Images grows, it is becoming increasingly difficult for users to find images that are relevant to their particular needs. Personalized search serves as one of such examples where the image search experience is enhanced by achieving the returned list according to the user interests. Personalized image search aims at improving the retrieval process by considering the particular interests of individual users. In this paper, ontology is proposed to create user profiles to capture individual user interests, and use themto produce moderate improvements when applied to search results. Image retrieval is based upon visual context of image and user profiles.***

Keywords: Image Retrieval, Ontology, Personalized Search, Ontology Mining, Specificity, Image search

1. Introduction

With the emergence and spread of digital cameras in everyday use , the number of images in personal and online databases with huge volumes. For example, the Google photos repository now consists of more than five billion images. Such kind of large image databases require adequate approaches for navigating, labeling, and searching. Gathering useful information from such large databases, becomes challenging to Web users. The current searching system concepts not able to entertain network users, as they are mostly based on keywords and tags -matching mechanisms and suffer from the problems of image mismatching and overloading. In this work we focus on the goal of selecting relevant images given a query term, i.e. Finding images showing content that associate with the query term. Personalized image access aims at enhancing the image retrieval (IR) process by accompaniment accurate user requests with definite user preferences, to better meet individual user needs. Personalization is being currently envisioned as a major research trend to relieve information overload, since Image Retrieval consistently tends to recapture the results likely even users may have different intentions for the query, e.g., searching for “OWL” by a web ontology researcher has a completely different meaning from searching by an animal specialist. Personalized search is one solution to address these problems, where user-specific information is to analyze intentions of the user queries exactly and to provide weights to user intentions. Because of flourishing importance of search engines, personalized search has the probable to significantly enhance searching experience.

The capture of user specific information requires the understandings of user’s personal interests and preferences. Were users profiles are constructed to preserve user Interesting domains. When gathering web information and images, User profiles represent the perception models consumed by users. In personalized searching user profile or user perception models are constructed by using ontologies. Such ontologies are termed as ontological user profiles or personalized ontologies.

For searching web documents and web images Search engines play an essential role . The search engine gathers, analyzes, organizes the data from the web and produces result to the user .The major components of a Search Engine are the Crawler, Indexer, and Query Processor. A Web Crawler is a simple, automated program, or script that absolutely scans or ―crawls through web documents to construct an index of the web documents it is looking for. The Search Engine which uses general web Crawler returns links. It may produce millions of web pages in response to a query and user interests. Examining all the resultant sets is difficult to user. So search engine makes use of ranking algorithm to display the resultant pages in a ranked order. The search engines commitment to be more efficient because there are large volumes of web documents.

The proposed framework shown in Fig 1 contains two components: 1) personalized user profiles. We introduce ontologies to learn user preferences from ontology domains and expanded preferences from user behavior.2)Ontology Mining to derive topic Specificity.

1.1 Framework

The proposed ontology model objective is to discover user background knowledge and learns personalized ontologies to represent user profiles. Figure 1 illustrates the framework of the ontology model. A personalized ontology is constructed, To capture user preferences, the domain ontology and the user’s local instance repository (browsing history), are utilized by this model. The domain ontology provides the structure for the personalized Ontology. Against the given topic, the specificity of topics are investigated for user background knowledge discovery(personalized preferences).

2. Personalized User Profiles

This paper is concerned with exploiting contextual information and smoothly integrating it into the personalization of user preferences retrieval. In this field, contextual information can be proven to be very helpful when dealing with image search. Most existing IR(image retrial) systems retrieval procedure based on queries and document collections; information about actual users and search context is largely ignored, and as result a significant number of misclassifications occurs.

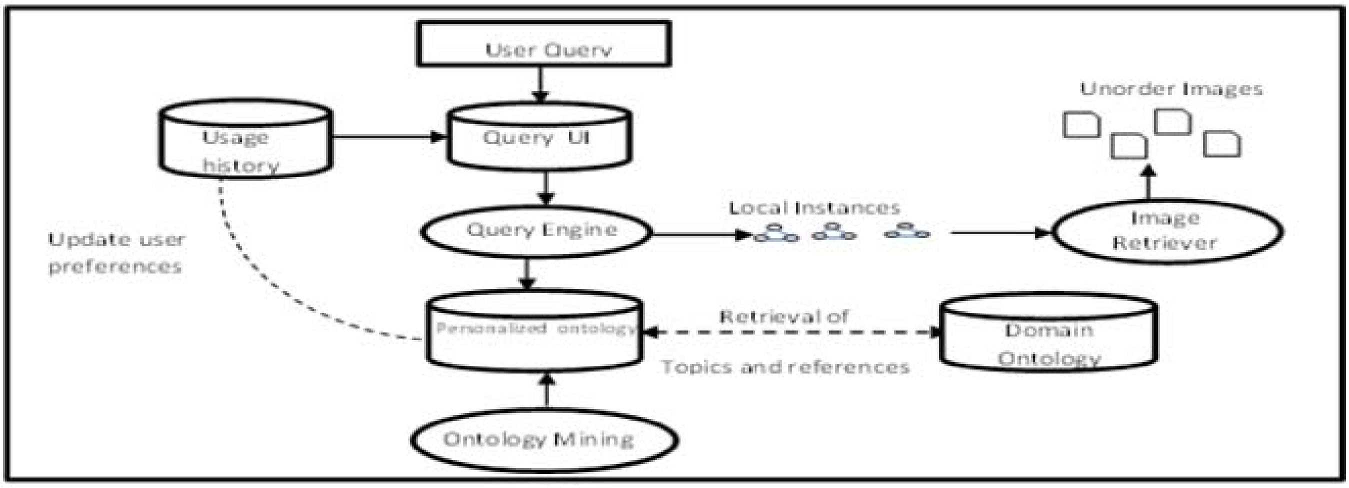

Figure 1: Framework of Personalized Image Search

To improve the retrieval accuracy one common solution is the use of relevance feedback . However, the effectiveness of relevance feedback is considered to be limited in real systems, because users are often uncertain to provide the required information. For this reason, implicit feedback has recently attracted greater attention. For a complex information request, the user profiles may need to update his/her preferences and view ranked documents in many iterations before the multimedia information need is satisfied. In such an interactive retrieval schema, the information naturally available to the retrieval system is more than just the current user query and the document collection in general, arbitrary interaction history can be made available to the retrieval system, including previous queries history, the multimedia documents that the user has visited.,

Our research aims at reinforce the efficiency and performance of approaches by a) using an enriched representation of the semantics of contents in the retrieval space, and b) adding information from the short-term retrieval context with a representation of long-term user interests, to gain a topic improvement for an individual searcher.

The model presented here is prepared on an ontology-based personalization framework. Building Ontology-based semantic structures and semantic raw data, the personalization system builds and exploits an explicit awareness of (meta) information about the user, either furnished by the user, or essential manifest along the history of his/her actions. Each user profile will be created from user implicit feedback and interaction history with weighted domains .

Creation of User profile acquisition techniques can be categorized into three groups: The Interviewing, Non-Interviewing, and Semi-Interviewing techniques. The interviewing user profiles are exclusively acquire using manual approach, such as questionnaires, interviews. Users inspect documents and cast positive or negative perception to the documents against given topics based upon the interpretation of user's interests and preferences explicitly. However, this form of User profile acquisition system is costly. The semi interviewing techniques comprise limited user involvement. The users were given set of topics or domains from that user have to specify the interesting and non-interesting topics. Based upon the feedback and user activity and behavior, user profiles were created. The non-interviewing techniques do not comprise users directly but ascertain user interests instead. Such user profiles are usually acquired by observing and mining knowledge from user image retrievals history and behavior. In our research semi interviewing method is used to create user profiles.

User Profiles can be represented using Ontologies. Ontologyproperly represents knowledge as a set of concepts within a domain, and the relationships between those concepts. Developing an ontology-based semantic structures and semantic metadata, the personalization system builds and exploits an explicit awareness of information about the user, either directly provided by the user, or implicitly evidenced along the history of his/her actions.

2.1 Domain Ontology Representation

Domain Ontology contains concepts and the relations between them, domain Ontology must cover an exhaustive range of topics, since users may come from different backgrounds. For this reason, the Library of Congress Subject Headings -LCSH system is an ideal world knowledge base. The LCSH was developed for organizing and retrieving information from a massive amount of library collections. The LCSH represents the natural growth and distribution of intellectual work, and covers exhaustivetopics of world knowledge. The LCSHsystem contains four types of notations: ***UF(Used For***), ***BT(Broader Topics***), ***RT(Related Topics),and NT (Narrower Topics)***to express relations between topics.

UF(Used For) are terms synonymous with the topic heading . During the investigation, we found that these references are often used to describe a similarity between topics. UF notation can be encoded as the ***similar_To*** relations. Thus, the complex usages of UF references simplified in the LCSH and encodethem only as the ***similar_To***relations in the Domain Ontology.

The reference BT (Broader Topic) provides the user with allowed headings in the list which are more generic in concept than the essential heading. BT may include some or all aspects of topic. They are encoded as ***instance_Of*** relation

NT(Narrower Topics) LCSH has a *thesaural* format with a hierarchy of terms ranging from those that cover very broad topics down to very narrow topics . The BT references are for two subjects describing the same topic, but at distinct levels of abstraction .In our model, they are encoded as the ***subclass_Of*** relations in the Domain Ontology.

RT(Related Topics) that provide access to information in some way relevant to our subject heading. The RT references are for two subjects relevant in some manner other than by hierarchy. They are encoded as the ***contains*** relations in our Domain Ontology.

The topics in the Domain Ontology are linked to each other by the reference relations of *similar_To*, *instance_of*, *containsandsubclass_Of*.

Definition: Let R be a set of reference relations, an element r ∈R is a two-tuple r <edge; type> where

- An edge connects two topics that hold a type of relation;
- A type of relations is an element of *similar_To*, *instance_Of* , *contains* and *subclass_Of*.

With This definition Domain Ontology can be represented as directed graph.

2.2 Personalized User Preferences

The Personalized User Preferences presented here is based on an ontology-based personalization framework. Building ontology-based semantic structures and semantic metadata, the personalization system generates and exploits an explicit awareness of (meta)information about the user, either the user provides his preferences directly or from the history of his/her actions the user preferences implicitly evidenced.

The search system assumes that the topics in a search result space S are annotated with weighted semantic metadata which describe the meaning carried by the item content, in terms of domain ontology O. That is, each result r∈S is associated with a value M(r)∈[0,1] of topic weights, where for each topic *x*∈O, the weight M*x*(*r*) indicates the degree to which the concept *x* is important in the meaning of r.The taxonomies of domain ontology O is a graph linked by semantic relations(reference terms defined in above section).

***Definition 1*:** Let U be the set of all users, Let O(T)be the Set of all topics in Domain Ontology, and P the universe of all user preferences. Since each user will have different preferences, let P(u) : U →O(T) map each user to his/her preference from domain ontology. Similarly, each user is related to a different context topic at each step in a session(N) with the system, which we shall represent by a mapping p(u):U × N → O(T), since we assume that
The topics evolve over time.

***Definition 2*:** *p* ≡*q* means that any consequence that could be inferred from *q* could also be inferred from *p*. A user *u* ∈U, if P(*u*) = *q* implies that *u* "p was preference to user u" (whatever this means), then "*x also be a* preference to user u". Which we shall represent by a mapping p(u) ×q → p(u).

Now we can specify the above definition for a specific representation of preference and context. In our model weighted domain ontology concepts are used to represent user preferences by a set of which the user interest, where the intensity of the interest can range from 0 to 1.

Definition 3: Given a personalized ontology p_o≡Φ(t,r) ∀r∈R and ∀ t∈T ,where T is set of topics and R is references we define the *set of all preferences* over O as p_o(x) = [0,1] The value p_o(x)represents the preference intensity for a topic *x*∈O(T) in the Domain ontology.

Definition 4: Dynamically personalized user preferences can be upgrade from the interaction of a user *u* with the system during a retrieval session. Therefore, at each point of time, the retrieval topic z∈O(T) is considered as user preference and the retrieval topic is joined in personalized Ontology as z → p_o. We represent the retrieval topicp_o(z) as a value in [0,1] of concept weights;

3. Personalized Ontology Mining

Ontology mining discovers interesting and topic knowledge from the concepts, semantic relations, and instances in an personalized ontology. In this section, a ontology mining method is introduced: Specificity . Specificity (denoted Spec) describes a subject's focus on a given topic. This method aims to investigate the subjects and the strength of their associations in an ontology.

3.1 Specificity

The specificity of topic is explored based on thestructure of O inherited from LCSH(the Library of Congress Topic Headings). The strength of such a focus is determined by thetopic's locality in the t structure (Topics, Reference terms)of O.As stated above in II, the structure T_x(t,r) where r∈R and t∈T of O is a graph linked by semantic relations. The topics located at upper levels toward the root aremore abstract than those at lower levels toward the "leaves." The upper level topics have more descendants, and thus refer to more concepts, compared with the lower level topics. Thus, in terms of a concept being referred to by both an upper level and lower level topics, the lower level topic has a stronger focus because it has fewer concepts in its space .Hence, the specificity of a lower level topic is greater than that of an upper leveltopic.The semantic specificity is measured based on the hierarchical reference relations (instance_Of , subclass_Of, contains,similar_To) held by asubject and its neighbors in T_x. Specificity is also called absolute specificity and denoted by Spec.The determination of a subject's Spec is described in following algorithm.

The determination of a topic Spec is described in Algorithm 1. The instance_Of(a) , subclass_Of (a), contains(a) and similar_To(a) are four functions in the algorithm. The instance_Of(a) returns a set of topics $t^{'} \in$ T that satisfy $edge(t^{'} \to a) = TRUE$ and reference_relation($t^{'} \to a$)=instance_Of. The subclass_Of (a) returns a set of topics $t^{'} \in$ T that satisfy $edge(t^{'} \to a) = TRUE$ and reference_relation($t^{'} \to a$)= subclass_Of. The contains(a) returns a set of topics $t^{'} \in$ T that satisfy $edge(t^{'} \to a) = TRUE$ and reference_relation($t^{'} \to a$)= contains. The similar_To(a) returns a set of topics $t^{'} \in$ T that satisfy $edge(t^{'} \to a) = TRUE$ and reference_relation($t^{'} \to a$)=similar_To.

Algorithm: Evaluation of Topic specificity

*Input:*Personalized Ontology p_o=<T,R>; a coefficient Ω between (0,1).

Output: calculates Spec(t) for each topic.

1. Take Q=1, get set of all leaves $\boldsymbol{l_o}$For all $l_o \in p_o$(T) ; assign Spec($\boldsymbol{l_o}$) = $\boldsymbol{Q}$;
2. Get $\boldsymbol{L}^{'}$ Which is set of upper level topics(parent topics) of Leaves $\boldsymbol{l_o}$Andthe related reference terms.
3. If($\boldsymbol{L}^{'}$ ***is empty***) then terminate the algorithm.
4. For each ($\boldsymbol{a} \in \boldsymbol{L}^{'}$) do
   a) If(contains(a)==ϕ) then s_{cont} (a)=Q;
   b) Else s_{cont} (a)= Ω ×min{Spec(a) | a ∈ contains(a);
   c) If(subclass_Of(a)==ϕ) the $s_{subclass}$ (a)=Q;
   d) Else $s_{subclass}$ (a)= Ω × $\frac{\sum Spec\ (b)}{|subclass\ (a)|}$Where b∈subclass_Of (a);
   e) If (similar_To (a)==ϕ) the $s_{similar}$ (a)=Q;
   f) Else $s_{similar}$ (a)= Ω × max{Spec(a) | a ∈similar_To(a)};
   g) If (instance_Of (a)==ϕ) the $s_{instance}$ (a)=Q;
   h) Else $s_{instance}$ (a)= Ω × $\prod Spec(b)$Where b ∈instance_Of (a);
   i) Spec(a)=min{s_{cont} (a),$s_{subclass}$ (a),$s_{similar}$ (a),$s_{instance}$ (a)};
5. End;
6. Q=Q × Ω; $\boldsymbol{l_o} = \boldsymbol{l_o} \cup \boldsymbol{L}^{'}$; goto step 2

Table 1: User Preferences (first Iteration)

Topic	Spec
Car	1.0
City	1.0
Sea	1.0
Tobby	1.0
Vegetation	1.0

Table 2: User Preferences (After some Iterations)

Topic	Spec
Car	1.0
City	1.0
Road	0.5
Construction	0.7
Sea	1.0
Dog	0.3
Lake	0.8

4. Personalized Image Search

Personalized image search framework is consisting of two sections image retrieval and query search engine. An image retrieval system is a computer system for browsing, searching and retrieving images from a large database of digital images. Image retrieval techniques are Image metasearch, Content-based image retrieval (CBIR), also known as content-based visual information retrieval (CBVIR).

Image meta search is conventional and simple methods of image retrieval utilize some method of enumerating metadata like captions , keywords, tags or images descriptions so that retrieval can be performed over the annotation words. User search for images by providing query such as keyword, meta data, or click on some image, and the system will return result set images "similar" to the query. The image meta searching method uses meta tags, color distribution in images, region/shape attributes, etc. Keywords based image search may produce inaccurate images for the query, because that a set of keywords, tags, descriptions cannot describe accurately the contents of images;

CBIR intention is to avoiding the use of textual descriptions and instead retrieves images based upon correlation in their contents textures, objects, colors, shapes etc. using image processing techniques to a user query or user-specified image features. The search system consider the contents of the image rather than keywords, tags, or descriptions associated with the image. In **Content-based image retrieval** adopts image features, it might refer to colors, shapes, textures, or any other context that can be interpreted from the image itself. CBIR is desirable because searches that rely purely on metadata are vulnerable on annotation quality and completeness.

Query search engine retrieves images, whose context are matched with the given query .The retrieved images context might refer to colors, shapes, textures, or any other information that can be derived from the image itself . Query engine exploit personalized ontology to retrieve user interests and user preferences are compared with the image retrieval context and returns the images that are matched with highest specificity of topics.

5. Conclusions

In this paper, an ontology model is proposed for representing user preferences for personalized image searching. The model constructs personalized ontologies for each user by extracting domain knowledge from the LCSH system and discovering user interests from user local instance repositories such as user visited histories. A ontology mining method specificity, is also introduced for user background knowledge discovery. The model was compared against benchmark models by applying it to a common systems for image searching. The experiment results demonstrate that our proposed model is promising. In this investigation, we found that the combination of global and local knowledge works better than using any one of them. In Addition, the personalized ontology model using knowledge with instance_Of, contains, subclass_Of, similar_To relations works better than using only few of them. When using only global knowledge, these relations have the same contributions to the achievement of the ontology model. The proposed ontology model in this paper provides a Solution to emphasizing global and local knowledge in a Single computational model. The findings in this paper can be applied to the design of web image retrieval system. The model also has extensive contributions to theFields of Information Retrieval, web Intelligence, Recommendation Systems, and Information Systems.

References

[1] "Webseer: An Image Search Engine for the World Wide Web" Charles Frankel, Michael J. Swain, and Vassilis Athitsos.
[2] Learn to Personalized Image Search From the Photo Sharing Websites Jitao Sang, Changsheng Xu, Senior Member, IEEE, and Dongyuan Lu
[3] "Harvesting Image Databases from the Web" Florian Schroff, Antonio Criminisi, and Andrew Zisserman, IEEE transactions on pattern analysis and machine intelligence, vol. 33, no. 4, april 2011
[4] "Towards a Multiple Ontology Framework for Requirements Elicitation and Reuse".
[5] International Entrepreneurship research (1989–2009): A domain ontology and thematic analysis Marian V. Jones a, ,Nicole Coviello b , Yee Kwan Tang c.
[6] Mining Fuzzy Domain Ontology from Textual Databases Raymond Y.K. Lau and Yue Xu 2007 IEEE/WIC/ACM International Conference on Web Intelligence.
[7] Ontodm: An Ontology of Data Mining 2008 IEEE International Conference on Data Mining Workshops.
[8] Semantic Web Mining State of the art and future directions Gerd Stumme a,∗, Andreas Hotho a, Bettina Berendt elesevier.
[9] Tag-based social image search with visual-text joint hypergraph learning Y Gao, M Wang, H Luan, J Shen, S Yan