< Previous   Next >

Dissertations

Search Using Social Media Structures

Download

Jangwon Seo, *University of Massachusetts - Amherst*

SHARE

Follow

Date of Award
9-2011

Document Type
Open Access Dissertation

Degree Name
Doctor of Philosophy (PhD)

Degree Program
Computer Science

First Advisor
W. Bruce Croft

Second Advisor
James Allan

Third Advisor
Andrew McCallum

Keywords
forum search, social media search, social media structures

Subject Categories
Computer Sciences

Abstract
Social applications on the Web have appeared as communication spaces
for sharing knowledge and information. In particular, social applications
can be considered valuable information sources because information in
the applications is not only easily accessible but also revealing in that the
information accrues via interactions between people. In this work, we
address methods for finding relevant information in social media
applications that use unique properties of these applications. In
particular, we focus on three unique structures in social media:
hierarchical structure, conversational structure, and social structure.
Hierarchical structures are used to organize information according to
certain rules. Conversational structures are formed by interactions within
communities such as replies. Social structures represent social
relationships among community members. These structures are designed
to organize information and encourage people to participate in
discussions in social applications. Accordingly, contexts extracted from
these structures can be used to improve the effectiveness of search in
social media relative to representations based solely on text content. To
exploit these structures in retrieval frameworks, we need to address

three challenges as follows. First, we should discover each structure because it is often obscure. Second, we need to extract relevant contexts from each structure because not all the contexts in a structure are relevant for retrieval. Last, we should represent each context or their combinations in a representation framework so that they can be encoded as retrieval components such as documents. In this work, we introduce an effective representation framework for multiple contexts. We then discuss how to discover or define each structure and how to extract relevant contexts from the structure. Using the representation framework, these relevant contexts are integrated into retrieval algorithms. To demonstrate that these structures can improve search in social media, the retrieval models and frameworks incorporating these structures are evaluated through experiments using data collections gathered from a variety of social media applications. In addition, we address two minor challenges related to social media search. First, it is not always easy to find relevant information from relevant objects if the objects are large. Accordingly, we address identification of relevant substructures in such objects. Second, text reuse structures are important since these structures have the potential to affect various retrieval tasks. In this thesis, we introduce text reuse structures and analyze text reuse patterns in real social applications.