Modeling betweenness for question answering

Conference Paper · January 2010

DOI: 10.1145/1871962.1871983 · Source: DBLP

1 author:

Brandeis Marshall
Spelman College

Modeling Betweenness for Question Answering

Brandeis Marshall
Computer and information Technology, Purdue University
West Lafayette, IN 47907 USA
brandeis@purdue.edu

ABSTRACT

The gap between the user's information need and query is expanding due to the pervasiveness of Web search. Multimedia question answering is restricted by disjoint infrastructures of text, image, video and audio content. We address the construction and integration of multiple sources of related information. In particular, we consider the challenging information need of betweenness and defining it for any form of multimedia. We motivate betweenness definition using four questions and suggest k nearest neighbor as a first-step solution.

Categories and Subject Descriptors

H.3.3 [**Information Search and Retrieval**]: Search Process; H.4 [**Information Systems Applications**]: Models and Principles

General Terms

Design, Management

Keywords

social networking graph model, betweenness data property

1. INTRODUCTION

The user's information need and query expressiveness of search portals are becoming further disjoint as Web growth and usage increases. For example, a more difficult question is: *is there a Starbuck's directly off the interstate?* An up-to-date GPS system can provide a good resolution; however it is restricted to location-based requests with direct routes. Many college campuses are an exception due to pedestrian-only walkways and closeness of buildings. This type of question answering does not extend to other forms of multimedia[1].

One challenge in question answering is how to handle images and the extensive amount of data contained within images [10]. Digital cameras allow for instant photographs; however, they typically auto-label photos according to a numerical scale while users rename photos based on context such as date, place and/or event. Kirk et. al. [6] identifies the user struggle to organize their photos since a download session usually consists of multiple images over a duration of time. Users could make more effective use of photo sharing applications, such as Flickr [2] and Picasa [3], to perform group tagging and annotations. However, the modular platform of these applications limit its portability and extensibility: (1) users have a username and password, (2) users have restricted storage capabilities and (3) photo annotations are non-transferable.

Tagging and annotations do have their drawbacks. Photo annotation relies on human intervention that can be error-prone with disjoint and unrelated tags. In addition, the desire to understand how pictures and objects within the pictures are related seems infeasible given the current data infrastructure of disjointness amongst text, metadata, image, video and audio content. Online photo sharing portals provide a partial solution; however user accessibility is limited with the reliance on remote servers as primary storage, indexing and retrieval.

Any form of multimedia question answering must construct and incorporate requests that need multiple sources of information. We consider a multimedia question answering solution through the concept of *betweenness*. Betweenness can be described in many ways, but we restrict our definition while considering four motivation questions. We also provide pseudocode of a k nearest neighbor approach to computing betweenness. By formulating betweenness queries rather than traditional SQL-style queries, the relatedness of multimedia can be exploited to provide improved portability and extensibility without dramatically impacting storage.

2. RELATED WORK

Traditionally, image retrieval approaches model one of two paths: content-based image retrieval (CBIR) or photo annotation. The CBIR approaches attempt to identify objects in a photograph through extracting the visual features of an image such as color, texture, shape, etc. However, they do not mainly focus on the other attributes of an image such as "where" the photograph was taken, "what" is occasion at which it was taken and the association between objects of "how" they are related [8].

Modern day online photo sharing applications like Flickr and Picasa allow users to tag their photographs with multiple keywords. To assist the tagging process, prior studies

[1]Our definition of multimedia includes text, metadata, image, video and audio content.

[2]http://www.flickr.com
[3]http://www.picasaweb.google.com

ESAIR'10, October 30, 2010, Toronto, Ontario, Canada.
ACM 978-1-4503-0372-9/10/10.

[2, 9] present methods to automate annotation. Golder [4] concentrates on using the people in photos to indicate the picture-taker's social relationships. We use the social networking graph to identify the set of attributes of *(who, what, when, how, where)* for a photograph and the betweenness property as a portal to understanding information for better knowledge management.

3. *SNG* FRAMEWORK

Digital image collections are usually managed as a hierarchy of directories and folders where the date and event name are used consistently to organize photos [6]. However, users must manually perform the creation and renaming of folders and sub-folders to make future viewing or retrieval accessible. To represent relationships amongst photos, we are developing the Social Networking Graph (SNG) framework [7]. The SNG framework incorporates the *who, how, where, what* and *when* attributes of a photograph. These five major attributes are named *object, association, place, event* and *time* respectively in the SNG framework.

A photo can be described using a combination of *any* of these five SNG attributes. Formally, the SNG attributes of a photo is a 5-tuple: $Photo(t, o, p, e, a)$ where t $\neq$ null. These SNG attributes can assist in categorizing images and identifying patterns amongst photos. A photograph is represented as an undirected social networking graph $SNG(V, E)$ where V denotes the set of all objects and E denotes the set of all relationships between any two objects in the graph. The time, place, event and association content, e.g., $\chi = \{$t, p, e, a$\}$, related to any particular photo is stored as an edge $edge(o_i, o_j)$ for a given pair of objects o_i, o_j.

Since objects may appear in multiple photos, the SNG framework can have multiple edges between the objects o_i and o_j of the form of $e^{\chi}(o_i, o_j)$. Hence, for a given photo, we may have a set of multiple objects, events, places and associations. Any photo $Photo_i$ can be transformed to a clique as represented in equation (1).

$$\text{subgraph } G(Photo_i) = Clique(V', E') \text{ s.t.} \tag{1}$$

$$V' = \{o_1, ..., o_w\} \text{ and } E' = \{edge_1, ..., edge_z\} \tag{2}$$

For each subgraph, data contents of place, event and association can be updated intermittently by using semi-automatic labeling techniques. By reusing annotations, storage usage can be minimized.

3.1 Betweenness Property

The betweenness centrality of a node [5] refers to the position of a node in the shortest path to other nodes in the network. In our SNG context, betweenness refers to the nearest neighbor relationship amongst more than two SNG attributes, both nodes and edges. More concretely, we are motivated by the following questions:

1. Having identified two SNG objects, *who* and *what* SNG attributes are between them?

2. Having identified two edge-based SNG attributes, *what* are the other edge-based SNG attributes and *who* are the other SNG objects between them?

3. Having identified an SNG object and an edge-based SNG attribute, *what* are the edge-based SNG attributes and *who* are the other SNG objects between them?

4. Given two photos, *what* are the edge-based SNG attributes and *who* are the other SNG objects between them?

Betweenness tends to be very computationally-expensive, especially in complex networks [1, 3]. Below we provide a code snippet of a betweenness function using the SNG framework. The objective is to construct a shortest path given the desired objects by concatenating the most frequently seen objects. The pairwise comparison in Lines 2-5 can be stored and updated on disk rather than in main memory.

1: **function** btw(graph:SNG,objects:$\{o_1, \ldots, o_n\}$, type:χ)
2: **for** $i = o_1$ to o_{n-1} **do**
3: **for** $j = o_2$ to o_n **do**
4: // ecount(i,j) stores the number of distinct photos between o_i and o_j with k distance
5: $ecount(i, j) = NN(k, o_i, o_j, \chi)$
6: // assume o_1 is source, o_x is target
7: $o_i = o_1$
8: $path = path \cup o_i$
9: **while** o_x is unseen **do**
10: $currmax = 0$, $id = x$
11: **for** $o_m \neq o_x$ **do**
12: **if** currmax $< max(ecount(i, m))$ **then**
13: currmax$= max(ecount(i, m))$; $id = m$
14: $path = path \cup o_{id}$
15: $o_i = o_{id}$
16: **return** path

4. REFERENCES

[1] D. A. Bader, S. Kintali, K. Madduri, and M. Mihail. Approximating betweenness centrality. In *Proceedings of the ACM international conference on Algorithms and models for the web-graph*, pages 124–137, 2007.

[2] H.-M. Chen, C. Ming-Hsiu, P.-C. Chang, M.-C. Tien, W. H. Hsu, and J.-L. Wu. Sheepdog - group and tag recommendation for flickr photos by automatic search-based learning. In *ACM Multimedia*, pages 737–740, 2008.

[3] S. Dolev, Y. Elovici, and R. Puzis. Routing betweenness centrality. *The Journal of the ACM*, 57(4), 2010.

[4] S. Golder. Measuring social networks with digital photograph collections. In *Proceedings of the ACM conference on Hypertext and hypermedia (HT)*, pages 43–48, 2008.

[5] R. A. Hanneman and M. Riddle. *Introduction to social network methods.* online at http://www.faculty.ucr.edu/~hanneman/nettext/, 2005.

[6] D. S. Kirk, A. J. Sellen, C. Rother, and K. R. Wood. Understanding photowork. In *In Proc. CHI 2006*, pages 761–770, 2006.

[7] B. Marshall and S. Pandey. Using the social networking graph for image organization. In *Proceedings of the International Conference on Advances in Future Internet*, July 2010.

[8] T. Rattenburg, N. Good, and M. Naaman. Towards automatic extraction of event and place semantics from flickr tags. In *Proceedings of the ACM SIGIR Conference on Research and development in information retrieval*, pages 103–110, 2007.

[9] B. Sigurbhornsson and R. van Zwoi. Flickr tag recommendation based on collective knowledge. In *Proceedings of ACM WWW*, pages 327–226, 2008.

[10] T. Yeh, J. J. Lee, and T. Darrell. Photo-based question answering. In *Proceedings of the ACM international conference on Multimedia*, pages 389–398, 2008.