Cognitive Visualization of Popular Regions Discovered From Geo-Tagged Social Media Data

Yunzhe Wang, The Hong Kong Polytechnic University, Kowloon, Hong Kong
George Baciu, Department of Computing, The Hong Kong Polytechnic University, Kowloon, Hong Kong
Chenhui Li, The Hong Kong Polytechnic University, Kowloon, Hong Kong

ABSTRACT

This article focuses on the cognitive exploration of photo sharing data which contain information about the location where the photo was taken and potentially some description about the photo. Therefore, the features of photo-spots can be deduced. Spots with similar features constitute a region of cognitive interest. The objective is to identify these regions and allow users to explore into regions of interest by cognitive understanding of their features. The authors propose an approach that makes use of semantic analysis, data clustering, and cognitive visualization. In this article, the authors introduce the design of an interactive visualization interface which projects photo sharing data to cognitive social activity map components. The contributions are two-fold. First, the authors put forward a novel social-media data classification method. Second, the authors suggest a new method to explore social activity maps by discovering regions of cognitive interest. Experiments are performed on the Flickr dataset.

KEYWORDS
Cognitive Visualization, Data Mining, Machine Learning, Region Discovery, Semantic Analysis, Social Media

INTRODUCTION

Typically, the human brain lacks the computational power necessary for processing large amounts of data. This involves both the limitation in bearing large volumes of data and gaining insight into the data patterns, especially when the data appears as raw numbers or text. Often, the authors seek to obtain general patterns, trends, and structure of components in the dataset in order to perform further analysis if needed. Furthermore, the authors often need to analyze the data by means of data mining, machine learning, and other statistical means. Then, the authors must project the results into visually discernible patterns for users to perform cognitive explorations of the dataset structure, as clusters, links, shapes and density. As proposed by Wang (Wang, 2014), eyes are not only responsible for attaining vision, but also work as the browser of internal memory in thinking and perception. Visual elements such as color, shape, and contrast etc. are generally easier to perceive than numbers and text in very large datasets.

In this paper, the authors aim at obtaining information from tagged datasets extracted from social media applications, and represent this information by cognitive visualization techniques.

DOI: 10.4018/IJCINI.2018010102

Copyright © 2018, IGI Global. Copying or distributing in print or electronic forms without written permission of IGI Global is prohibited.
Social media has now become a pervasive global communication channel. Many applications and platforms have become available for users to post messages, follow friends and share experiences. Due to the high frequency with which users update their states, a large amount of data is being generated around the world every second. Among all social media applications, photo sharing platforms contain more information as they have the potential to establish virtual links between different locations and users. By analyzing this data, valuable patterns can be extracted such as the distribution of users, their common interests, activities, locations visited, etc. More specifically, different from pure text messages, photo posts conceal extra information, such as the photo location and the photo description. The description is often attached by users and commonly appears in the form of a sequence of tags. These tags take the form of ‘#’ mark as a prefix. The tag can be a word or a phrase without spaces. This makes it easier for users to convey the key point of their posts. Based on these tags, the authors are capable of measuring relatedness between multimedia such as images and videos. The obtained relatedness further support applications including clustering, searching, recommendation, and annotation (Xu et al., 2013).

For users who are interested in a specific theme, searching a keyword will quickly present all the posts that have been tagged with keywords in that theme. According to common experience, users tend to post photos of several commonly-seen objects, such as delicious food, beautiful scenes, and selfies. In other words, user tags usually share the same topic of interest. A cognitive classification of such tags leads to photo locations that are related to these tags and can also fall into different groups. Each group of locations contributes to a region of cognitive interest. For example, if many photos of food are collected within a region, then this region has a high probability of having a food court, a food market, or a cluster of restaurants. As the authors may refer to some terms frequently, the authors list their definitions below:

- **Photo-Spot**: The location where the photo is taken. It is denoted by S;
- **Region of Interest**: A geographical area which covers a set of photo spots sharing the same class label of tags. The region of interest is represented by R;
- **Class Label**: Predefined labels used to classify tags. The label indicates the theme of tags, for example, a label list (food, entertainment, transportation) separates tags into three categories. The label is denoted by L;
- **Region Feature**: Photo-spots that fall into the same region. With the region feature, the authors may deduce the region to be a food court, a transportation center, or an entertainment place.

On traditional maps, regions are divided by the border of districts, cities, countries or continents. This divisions are relatively static, because a city or a country seldom expands or shrinks its territory. However, within these borders, many facilities keep changing. For example, with the fast, changing economy, large areas of land may be opened up for new residential communities. Besides, shops selling similar types of products may group together to form a business district. As time goes by, the district itself may expand or shrink within the city border.

**Contributions**

The contribution of this work lies in the analysis of tags. Usually, users are allowed to tag their posts with multiple words. Some of these words can be found in the dictionary, while some are in short form or misspelled. With semantic analysis, it becomes possible to gain insight into the context of the photo-spot description of tags. The highlights of the contributions can be summarized as follows:

- **Tag Classification**: Given the original tag data, users will just have different impressions of independent photo-spots. It is difficult to relate one photo-spot with another. Based on the context learnt from semantic analysis, the authors can classify a large number of tags in order to reduce
Ordering: A Reliable Qualitative Information for the Alignment of Sketch and Metric Maps


[www.igi-global.com/article/ordering/114661?camid=4v1a](www.igi-global.com/article/ordering/114661?camid=4v1a)