Chapter 9

Tag Cloud Reorganization: Finding Groups of Related Tags on Delicious

Alberto Pérez García-Plaza
UNED, Spain

Arkaitz Zubiaga
City University of New York, USA

Victor Fresno
UNED, Spain

Raquel Martínez
UNED, Spain

ABSTRACT

Tag clouds have become an appealing way of navigating through Web pages on social tagging systems. Recent research has focused on finding relations among tags to improve visualization and access to Web documents from tag clouds. Reorganizing tag clouds according to tag relatedness has been suggested as an effective solution to ease navigation. Most of the approaches either rely on co-occurrences or rely on textual content to represent tags. In this chapter, the authors explore tag cloud reorganization based on both of them. They compare these clouds from a qualitative point of view, analyzing pros and cons of each approach. The authors show encouraging results suggesting that co-occurrences produce more compelling reorganization of tag clouds than textual content, being computationally less expensive.

INTRODUCTION

Social bookmarking sites allow to collaboratively annotate Web pages, relying on the social tagging philosophy. Social tagging is a Web 2.0 application based phenomenon where users can describe Web contents by adding tags or keywords as metadata in an open, collaborative, and non-hierarchical way (Smith, 2008). Social bookmarking is a popular way to store, organize, comment on and search links to any Web page, and it has emerged as one of the most important Web applications that eases information sharing. Popular collaborative
Tagging sites have aggregated a vast amount of user-created metadata in the form of tags, providing valuable information about the interests and expertise of the users. Because of this, it becomes a fertile area to scientific research on social media (Gupta, et al., 2010).

In order to facilitate access to tagged resources, and to enable visual browsing, social bookmarking tools typically provide an interface model known as the tag cloud. A tag cloud is an appealing way to enable users to navigate through the most popular tags of a social bookmarking site. When users access the information contained in these structures, it is presented in the form of a cloud consisting of the most popular tags, where the bigger is the font size of a tag, the more popular it is on the site. Typical tag clouds include between 100 and 200 tags, and tag weights are represented by different font sizes, or other visual clues. In addition, tags can be sorted in alphabetical, size-based, or random order, and users can sometimes customize clouds with different fonts, layouts, and color schemes. These structures are particularly useful for browsing and for information discovery, because they provide a visual summary of the content in the collection. However, related tags do not appear in nearby spaces of the tag cloud, and it is not easy to find the tags of one’s interest. To solve this problem, research in the field has pointed out that grouping related tags, and showing them close to each other can help enhance navigation through tag clouds.

In order to enhance browsing phase in a tag cloud, an effective way is to identify inter-related tags and relations among contents. This book chapter aims to discuss the tag grouping task so that it enables an enhanced visualization and improved navigation through the tag cloud. To this end, several methods of representing tags have been proposed in earlier research. Most of them consider co-occurrences among tags to group related tags into clusters, but do not pay special attention on the algorithm employed to weight such co-occurrences. In this work, we focus on the reorganization of a tag cloud based on the identification of groups of inter-related tags, and compare different methods for weighting tag co-occurrences. We rely on a well-known clustering algorithm for this purpose.

Recently, there has been an increasing interest on tag clustering tasks; most of them tackle the problem from the point of view of tag co-occurrences (Specia & Motta, 2007; Mika, 2007; Sbodio & Simpson, 2009). Other works have followed a content-based approach, such as Zubiaga et al. (2009). All of them performed a qualitative evaluation of their results, finding appealing groupings for human users. Nevertheless, these works did not compare content-based methods with those based on tag co-occurrences, widely used in the literature.

In this book chapter we further explore several state-of-the-art weighting functions to represent co-occurrences among tags. After clustering tags with these weightings, we compare the results with those obtained by the content-based approach. Going further, we analyze and discuss the appropriateness and performance of each approach.

Next, in Section Background we cover some basic ideas about social tagging systems and present the related work. In Section Tag Cloud Reorganization, we explain the settings of our experiments, including dataset, tag representation approaches, and tag clustering algorithm. To conclude the section, we analyze the resulting clouds and discuss some possible applications. Finally, we summarize the future research directions and conclusion.

**BACKGROUND**

Tagging is an open way to assign tags or keywords to resources or items (e.g., a Web page), in order to describe the characteristics of them. This enables later retrieval of resources in an easier way. As opposed to a classical taxonomy-based categorization system, they are usually non-hierarchical,
Related Content

The Use of Social Media by Nonprofit Organizations: An Examination from the Diffusion of Innovations Perspective
www.igi-global.com/chapter/use-social-media-nonprofit-organizations/39796?camid=4v1a

Gender Segregation and ICT: An Indo-British Comparison
www.igi-global.com/article/gender-segregation-ict/51350?camid=4v1a

Developing Marketing Strategy on Social Networks
Boris Milovi (2018). Social Media Marketing: Breakthroughs in Research and Practice (pp. 73-89).
www.igi-global.com/chapter/developing-marketing-strategy-on-social-networks/203293?camid=4v1a

Social Aware Cognitive Radio Networks: Effectiveness of Social Networks as a Strategic Tool for Organizational Business Management
www.igi-global.com/chapter/social-aware-cognitive-radio-networks/201244?camid=4v1a