Chapter 11
Analysis of Content Popularity in Social Bookmarking Systems

Symeon Papadopoulos
Aristotle University of Thessaloniki, Greece Informatics & Telematics Institute, Thermi, Thessaloniki, Greece

Fotis Menemenis
Informatics & Telematics Institute, Thermi, Thessaloniki, Greece

Athena Vakali
Aristotle University of Thessaloniki, Greece

Ioannis Kompatsiaris
Informatics & Telematics Institute, Thermi, Thessaloniki, Greece

ABSTRACT

The recent advent and wide adoption of Social Bookmarking Systems (SBS) has disrupted the traditional model of online content publishing and consumption. Until recently, the majority of content consumed by people was published as a result of a centralized selection process. Nowadays, the large-scale adoption of the Web 2.0 paradigm has diffused the content selection process to the masses. Modern SBS-based applications permit their users to submit their preferred content, comment on and rate the content of other users and establish social relations with each other. As a result, the evolution of popularity of socially bookmarked content constitutes nowadays an overly complex phenomenon calling for a multi-aspect analysis approach. This chapter attempts to provide a unified treatment of the phenomenon by studying four aspects of popularity of socially bookmarked content: (a) the distributional properties of content consumption, (b) its evolution in time, (c) the correlation between the semantics of online content and its popularity, and (d) the impact of online social networks on the content consumption behavior of individuals. To this end, a case study is presented where the proposed analysis framework is applied to a large dataset collected from Digg, a popular social bookmarking and rating application.

INTRODUCTION

The emergence of Web 2.0 technologies and the widespread use of applications integrating such technologies have transformed the way people experience and act in online settings. In the first days of the Web, people were excited to browse through and consume online content (mostly static web pages) that was prepared and published by website owners or administrators. Nowadays, digital content consumption – e.g. online article reading, picture viewing and video watching – still appears to be one of the main activities for most internet users[†]. However, the advent of the Web 2.0 application paradigm has transformed the established "browsing-based" online content consumption behavior of users. This change was possible by means of offering users a host of rich interactivity features, such as content sharing, rating as well as online community building. Thus, users of today's Web 2.0 applications are empowered to share, organize, rate and retrieve online content. In addition, users are exposed to the content-related activities of other users and can even form online relations to each other. Consequently, online content consumption within a modern Web 2.0 application constitutes an overly complex phenomenon with interesting dynamics which have not been thoroughly investigated yet.

Social Bookmarking Systems (SBS) hold a prominent place among Web 2.0 applications with respect to content consumption since they provide a platform where users are provided with two significant features:

- Submitting and sharing bookmarks (links) to online resources, e.g. articles, photos or videos, which they consider interesting.
- Indicating their preference or disapproval to bookmarks submitted by other users, by voting for or against the interesting-ness/appeal of online resources and by commenting on them.

In addition to these two features which are fundamental for an SBS, there are two other optional groups of features, namely Taxonomic and Social-Community features. Taxonomic features pertain to the possibilities offered to users for assigning bookmarks to a predefined topic-scheme or for tagging them with freely chosen keywords.

Table 1. Taxonomic and social/community features of several social bookmarking services

Application name / URL	Type of Bookmarks	Taxonomic Feats		Social/Community Feats	
		Tags	**Topics**	**Friends**	**Groups**
Digg.com	News		+	+	
www.**StumbleUpon**.com	General	+		+	
www.**propeller**.com	News	+	+	+	+
www.**reddit**.com	News		+	+	
www.**fark**.com	News		+		
www.**newsvine**.com	News	+	+	+	+
delicious.com	General	+		+	
www.**blinklist**.com	General	+		+	
www.**clipmarks**.com	General	+		+	
furl.nt	General	+	+		+
www.**citeulik**.org	Ciations	+			+
bibsonomy.org	Citations	+			+

23 more pages are available in the full version of this document, which may
be purchased using the "Add to Cart" button on the publisher's webpage:
www.igi-global.com/chapter/analysis-content-popularity-social-
bookmarking/38226

Related Content

Partially Supervised Classification: Based on Weighted Unlabeled Samples Support Vector Machine
Zhigang Liu, Wenzhong Shi, Deren Liand Qianqing Qin (2008). *Data Warehousing and Mining: Concepts, Methodologies, Tools, and Applications (pp. 1216-1230).*
www.irma-international.org/chapter/partially-supervised-classification/7695

Web Usage Mining Data Preparation
Bamshad Mobasher (2008). *Data Warehousing and Mining: Concepts, Methodologies, Tools, and Applications (pp. 2551-2557).*
www.irma-international.org/chapter/web-usage-mining-data-preparation/7782

Ensemble Data Mining Methods
Nikunj C. Oza (2008). *Data Warehousing and Mining: Concepts, Methodologies, Tools, and Applications (pp. 356-363).*
www.irma-international.org/chapter/ensemble-data-mining-methods/7650

Data Mining for Intrusion Detection
Aleksandar Lazarevic (2005). *Encyclopedia of Data Warehousing and Mining (pp. 251-256).*
www.irma-international.org/chapter/data-mining-intrusion-detection/10602

Active Learning with Multiple Views
Ion Muslea (2005). *Encyclopedia of Data Warehousing and Mining (pp. 12-16).*
www.irma-international.org/chapter/active-learning-multiple-views/10557