Mobile Novelty Mining

Agus T. Kwee, Nanyang Technological University, Singapore
Flora S. Tsai, Nanyang Technological University, Singapore

ABSTRACT

Service-oriented Web applications allow users to exploit applications over networks and access them from a remote system at the client side, including mobile phones. Individual services are built separately with comprehensive functionalities. In this article, the authors transform a standalone offline novelty mining application into a service-oriented application and allow users to access it over the Internet. A novelty mining application mines the novel, yet relevant, information on a topic specified by users. In this article, the authors propose a design for a service-oriented novelty mining application. After deploying their service-oriented novelty mining system on a server, use case scenarios are provided to demonstrate the system. The authors’ service-oriented novelty mining system increases the efficiency of gathering novel information from incoming streams of texts on their mobile devices for users.

Keywords: Contextual Information, Implementation, Mobile Application, Novelty Mining, Service-Oriented, Design, Web Services

INTRODUCTION

Nowadays, people are becoming more and more mobile. Thus, people want to get necessary and important information fast and in any given time, regardless where they are. As mobile phone technology has become more advanced, applications are developed so that they can be executed on the mobile devices and provide information for mobile users.

With the fast growth of technology, the Web is also changing from a datacentric Web into a Web of semantic data and Web of services (Yee, Tiong, Tsai, & Kanagasabai, 2009). In fact, the demand for Web services that enable users to run offline standalone applications over the Internet has rapidly increased. The World Wide Web Consortium (W3C) defines a Web service as “a software system designed to support interoperable machine-to-machine interaction over a network” (Hugo & Allan, 2004). As a result, more software applications are available online and can be accessed on the remote system at the client side. To easily identify these online applications, each of them is assigned with the unique URI (Uniform Resource Indicator), which serves as the address of the individual application or service. Individual services can be built and combined with one another to create other services with more comprehensive functionality at little additional cost (Zheng & Bouguettaya, 2009). Nowadays, the use of these Web services has significance in the business domain. They primarily use these Web services to communicate or exchange data between businesses and clients.

DOI: 10.4018/japuc.2009100104
Another consequence of the rapid growth of technology is the information overload in the form of news articles, scientific papers, blogs (Chen, Tsai, & Chan, 2008), and social networks (Tsai, Han, Xu, & Chua, 2009). Information is abundantly available on the Internet, where people can find almost anything they desire. However, most of the time, people tend to suffer from information overload because of irrelevant and redundant information in these documents. Therefore, novelty mining (NM), or novelty detection, is a solution to this phenomenon. Novelty mining can be used to solve many solve many business problems, such as in corporate intelligence (Tsai, Chen, & Chan, 2007) and cyber security (Tsai, 2009; Tsai & Chan, 2007). A novelty mining process is a process of retrieving novel yet relevant information, based on a topic given by the user (Ng, Tsai, & Goh, 2007; Ong, Kwee, & Tsai, 2009).

To the best of our knowledge, no previous work has been reported in designing and developing a service oriented novelty mining system for the business enterprise. The purpose of creating a Web-based application for the novelty mining system is to offer an online system for users to ease their information burden. This saves time by allowing users to only read novel and relevant information in their topic of interest. For enterprise users, this service-oriented novelty mining system conveniently helps them to retrieve new information about certain events of interest. They do not need to go through all documents or passages in order to find the novel information. Creating Web services for the novelty mining components facilitates the rapid deployment and availability of these services for these diverse set of users, which can balance technical significance and business concerns in business processes and enterprise systems.

The objective of this article is to develop a Web service application i.e. novelty mining system that is used to detect and retrieve any new information of an incoming stream of text to be executed on mobile devices.

This article is organized as follows. First, we review recent work on novelty mining. Next, we describe our novelty mining framework, algorithm, and performance evaluation. Then, the architecture and design of the service-oriented novelty mining system are presented. The development of the new system is explained. Two test case scenarios are given to illustrate the service-oriented novelty mining system. Next, we describe the integration of Web service into mobile devices. Finally, the last section summarizes the entire article.

LITERATURE REVIEW

The pioneering work of novelty mining was proposed by Zhang et al. at the document level (Zhang, Callan, & Minka, 2002). In their work, “novelty” was defined as the opposite of “redundancy”. Given any set of documents, any document which was less similar to its history documents was regarded as “novel”. However, although users can retrieve all novel documents, they still need to read through each document to find the novel information because not all sentences are novel in these documents. Therefore, to serve users better, later studies of novelty mining were performed at the sentence level (Allan, Wade, & Bolivar, 2003; Kwee, Tsai, & Tang, 2009; Zhang, Xu, Bai, Wang, & Cheng, 2004; Zhang & Tsai, 2009b).

Many studies related to novelty mining originated from the TREC Novelty Tracks (Harman, 2002; Soboroff, 2004; Soboroff & Harman, 2003). The novelty track was introduced in the eleventh Text REtrieval Conference (TREC) in 2002 (Harman, 2002). This track was designed to investigate systems’ abilities in locating new and relevant information from a given document set which is categorized into topics. Other works on sentence and document-level novelty mining tend to apply some promising content-oriented techniques (Allan, Lavrenko, & Jin, 2000; Franz, Ittycheriah, McCarley, & Ward, 2001; Yang, Zhang, Carbonell, & Jin, 2002).

Recent studies suggested several ways to improve the performance of novelty mining by integrating various natural language processing (NLP) techniques, such as part-of-speech (POS)
Related Content

Strategic Study on Culture Brand in Shijiazhuang
[www.irma-international.org/article/strategic-study-culture-brand-shijiazhuang/68802/](www.irma-international.org/article/strategic-study-culture-brand-shijiazhuang/68802/)

Pervasive Computing: A Conceptual Framework
[www.irma-international.org/chapter/pervasive-computing-conceptual-framework/28447/](www.irma-international.org/chapter/pervasive-computing-conceptual-framework/28447/)

WLAN and Bluetooth Positioning in Smart Phones
Ruizhi Chen, Ling Pei, Jingbin Liu and Helena Leppäkoski (2012). *Ubiquitous Positioning and Mobile Location-Based Services in Smart Phones* (pp. 44-68).
[www.irma-international.org/chapter/wlan-bluetooth-positioning-smart-phones/67039/](www.irma-international.org/chapter/wlan-bluetooth-positioning-smart-phones/67039/)

The Information Construction of Wind Farm Based on SIS System
[www.irma-international.org/chapter/information-construction-wind-farm-based/72937/](www.irma-international.org/chapter/information-construction-wind-farm-based/72937/)

Dynamic Spectrum Auction and Load Balancing Algorithm in Heterogeneous Network
[www.irma-international.org/article/dynamic-spectrum-auction-load-balancing/64314/](www.irma-international.org/article/dynamic-spectrum-auction-load-balancing/64314/)