

Chapter 11

Exploring Virtual Communities with the Internet Community Text Analyzer (ICTA)

Anatoliy Gruzd
Dalhousie University, Canada

ABSTRACT

The chapter presents a new web-based system called ICTA (<http://netlytic.org>) for automated analysis and visualization of online conversations in virtual communities. ICTA is designed to help researchers and other interested parties derive wisdom from large datasets. The system does this by offering a set of text mining techniques coupled with useful visualizations. The first part of the chapter describes ICTA's infrastructure and user interface. The second part discusses two social network discovery procedures used by ICTA with a particular focus on a novel content-based method called name networks. The main advantage of this method is that it can be used to transform even unstructured Internet data into social network data. With the social network data available it is much easier to analyze, and make judgments about, social connections in a virtual community.

INTRODUCTION

In the age of cheap digital data storage more and more online interactions among people are being captured and stored for posterity. This treasure trove of data represents a unique opportunity for social scientists and Internet researchers to study and better understand the inner workings of virtual communities. Researchers can now easily scrutinize these recorded interactions and

answer questions like: how and why one virtual community emerges and another dies, how people agree on common practices and rules in a virtual community, and how they share knowledge and information among group members. Answers to these and other related questions will allow us to understand basic processes such as how people meet, communicate, and establish social relationships. It will also help practitioners to develop new technologies to better serve the information needs of different communities. For instance, social networking websites like Facebook and

DOI: 10.4018/978-1-60960-040-2.ch011

MySpace are good examples of how advancements in information technology can help people form and support a much larger number of online relationships than has ever been possible before. However, this avalanche of online data can be overwhelming for both researchers and the public at large. Thus, it is not surprising that there is an increasing interest in the ability to retrieve and analyze online data generated by virtual communities automatically.

This chapter presents a web-based system for automated analysis and visualization of online conversations called the Internet Community Text Analyzer (ICTA). The system is available at <http://netlytic.org>. The main goal of ICTA is to automate the process of analyzing and visualizing text-based communal interactions and provide researchers and other interested parties with effective automated methods to study virtual communities. This system was primarily tested with online learning communities, but it can also be used to analyze a wide variety of other types of text-based virtual communities. This chapter describes the development of ICTA, its infrastructure and user interface. In particular it focuses on a new method called 'name network' that allows users to automatically extract social networks from text-based computer mediated communication. Once discovered, social networks can provide researchers with an effective mechanism for studying collaborative processes in virtual communities such as shared knowledge construction, information sharing, influence, and support. In addition to being useful for researchers, social networks can also help web developers to improve online recommendation systems by analyzing the preferences of other users with similar interests (e.g., Amazon.com, Netflix.com) or provide new browsing capabilities for online information (e.g., Silobreaker.com). For example, with a social network representation of news it is now possible to trace explicit or implicit connections between events and individuals involved in the news (Pouliquen et al, 2007; Taney, 2007). Information about online social networks can also

provide a more secured and easier way to share private content with trusted individuals within the so-called Web of Trust (e.g., Golbeck, 2008; Matsuo et.al., 2004). Finally, companies can use online social networks to recruit talented individuals (e.g., Leung, 2003), find experts (Ehrlich et al, 2007; Li et al, 2007), organize more effective virtual marketing campaigns (e.g., Domingos, 2005), or build brand loyalty using customer networks (Thompson & Sinha, 2008).

RELATED WORK

Research on automated analysis and visualization of online conversations can be grouped according to the various types of computer-mediated communication (CMC) technology presently in use. Below is a brief overview of some of the research on the four most popular CMC and online media types: emails, online forums, blogs, and twitter. This review is not meant to be exhaustive; its primary purpose is to provide a general overview of common methods used for analyzing and visualizing online conversations and to give the reader a starting point for further reading. Social networking (SN) websites such as Facebook and MySpace are not reviewed separately since these sites often utilize one or more of the popular CMC types that will be discussed below.

Emails

Social network analysis (SNA) is one of the most common methods for studying email-based interactions. Email data contains characteristics that naturally fit with the network model; for example, senders and receivers form uniquely identifiable nodes within a network, and email traffic can be used to establish links between the nodes in the network. Among scholars who used SNA to study email data are Diesner and Carley (2005) and Lim et al. (2007). Both of these research teams used SNA to discover and analyze

17 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/exploring-virtual-communities-internet-community/50341

Related Content

A Virtual-Reality Approach for the Assessment and Rehabilitation of Multitasking Deficits

Otmar Bock, Uwe Drescher, Wim van Winsum, Thomas F. Kesnerusand Claudia Voelcker-Rehage (2018). *International Journal of Virtual and Augmented Reality* (pp. 48-58).

www.irma-international.org/article/a-virtual-reality-approach-for-the-assessment-and-rehabilitation-of-multitasking-deficits/203067

Performance Measurement of Computer Networks

Federico Montesino Pouzols, Angel Barriga Barros, Diego R. Lopezand Santiago Sánchez-Solano (2008). *Encyclopedia of Networked and Virtual Organizations* (pp. 1216-1222).

www.irma-international.org/chapter/performance-measurement-computer-networks/17746

Online Group Facilitation Skills

Nancy White (2006). *Encyclopedia of Virtual Communities and Technologies* (pp. 354-358).

www.irma-international.org/chapter/online-group-facilitation-skills/18100

Seeking Accessible Physiological Metrics to Detect Cybersickness in VR

Takurou Magakiand Michael Vallance (2020). *International Journal of Virtual and Augmented Reality* (pp. 1-18).

www.irma-international.org/article/seeking-accessible-physiological-metrics-to-detect-cybersickness-in-vr/262621

Evolutionary Framework for Cognitive Building Information Modelling and Augmented and Virtual Reality

Rakshit Kothari, Darshan Nyati, Mansvee Shah, Gayatri Mohantyand Mayank Patel (2024). *Modern Technology in Healthcare and Medical Education: Blockchain, IoT, AR, and VR* (pp. 216-236).

www.irma-international.org/chapter/evolutionary-framework-for-cognitive-building-information-modelling-and-augmented-and-virtual-reality/345890