H

Human Factors in the Development of Trend Detection and Tracking Techniques

Chaomei Chen
Drexel University, USA

Kaushal Toprani
Drexel University, USA

Natasha Lobo
Drexel University, USA

INTRODUCTION

Trend detection has been studied by researchers in many fields, such as statistics, economy, finance, information science, and computer science (Basseville & Nikiforov, 1993; Chen, 2004; Del Negro, 2001). Trend detection studies can be divided into two broad categories. At technical levels, the focus is on detecting and tracking emerging trends based on dedicated algorithms; at decision making and management levels, the focus is on the process in which algorithmically identified temporal patterns can be translated into elements of a decision making process.

Much of the work is concentrated in the first category, primarily focusing on the efficiency and effectiveness from an algorithmic perspective. In contrast, relatively fewer studies in the literature have addressed the role of human perceptual and cognitive system in interpreting and utilizing algorithmically detected trends and changes in their own working environments. In particular, human factors have not been adequately taken into account; trend detection and tracking, especially in text document processing and more recent emerging application areas, has not been studied as integral part of decision-making and related activities. However, rapidly growing technology, and research in the field of human-computer interaction has opened vast and, certainly, thought-provoking possibilities for incorporating usability and heuristic design into the areas of trend detection and tracking.

BACKGROUND

In this section, we briefly review trend detection and its dependence on time and context, topic detection and tracking, supported by instances of their impact in diverse fields, and the emerging trend detection especially for text data.

Trend Detection

A *trend* is typically defined as a continuous change of a variable over a period of time, for example, unemployment numbers increase as the economy enters a cycle of recession. Trend detection, in general, and topic detection techniques are groups of algorithmic tools designated to identify significant changes of quantitative metrics of underlying phenomena. The goal of detection is to enable users to identify the presence of such trends based on a spectrum of monitored variables. The response time of a detection technique can be measured by the time duration of the available input data and the identifiable trend; it is dependent on specific application domains. For example, anti-terrorism and national security may require highly responsive trend detection and change detection capabilities, whereas geological and astronomical applications require long-range detection tools. Other applications of this technology exist in the fields of business and medicine.

Much research has been done in the field of information retrieval, automatically grouping (clustering) documents, performing automated text summarization, and automatically labeling groups of documents.

Policymakers and investigators are, obviously, eager to know if there are ways that can reliably predict each turn in the economy. Economists have developed a wide variety of techniques to detect and monitor changes in economic activities. The concept of *business cycles* is defined as fluctuations in the aggregate economic activities of a nation. A business cycle includes a period of expansion, followed by recessions, contractions, and revivals. Three important characteristics are used when identifying a recession: duration, depth, and diffusion — the three Ds. A recession has to be long enough, from a year to 10 years; a recession has to be bad enough, involving a substantial decline in output; and a recession has to be broad enough, affecting several sectors of the economy.

Topic Detection and Tracking

Topic Detection and Tracking (TDT) is a sub-field primarily rooted in information retrieval. TDT aims to develop and evaluate technologies required to segment, detect, and track topical information in a stream consisting of news stories. TDT has five major task groups: (1) story segmentation, (2) topic detection, (3) topic tracking, (4) first story detection, and (5) story link detection. Topic detection focuses on discovering previously unseen topics, whereas topic tracking focuses on monitoring stories known to a TDT system. First story detection (FSD) aims to detect the first appearance of a new story in a time series of news associated with an event. Roy, Gevry, and Pottenger (2002) presented methodologies for trend detection. Kontostathis, Galitsky, Roy, Pottenger, and Phelps (2003) gave a comprehensive survey of emerging trend detection in textual data mining in terms of four distinct aspects: (1) input data and attributes, (2) learning algorithms, (3) visualization, and (4) evaluation.

TDT projects typically test their systems on TDT data sets, which contain news stories and event descriptors. The assessment of the performance of a TDT algorithm is based on *Relevance Judgment*, which indicates the relevancy between a story and an event. Take the event descriptor *Oklahoma City Bombing* as an example. If a matching story is about survivors' reaction after the bombing, the relevance judgment would be *Yes*. In contrast, the relevance judgment of the same story and a different event descriptor *U.S. Terrorism Response* would be *No*. Swan and Allan (1999) reported their work on extracting significant time varying features from text based on this type of data.

An interesting observation of news stories is that events are often reported in burst. Yang, Pierce, and Carbonell (1998) depicted a daily histogram of story counts over time. News stories about the same event tend to appear within a very narrow time frame. The gap between two bursts can be used to discriminate distinct events.

Kleinberg (2002) developed a burst detection algorithm and applied to the arrivals of e-mail and words used in titles of articles. Kleinberg was motivated by the need to filter his e-mail. He expected that whenever an important event occurs or is about to occur, there should be a sharp increase of certain words that characterize the event. He called such sharp increases *bursts*. Essentially, Kleinberg's burst detection algorithm analyzes the rate of increase of word frequencies and identifies the most rapidly growing words. He tested his algorithm on the full text of all the State of the Union addresses since 1790. The burst detection algorithm identified important events occurring at the time of some of the speeches. For example, *depression* and *recovery* were *bursty* words in 1930-1937, *fighting* and *Japanese* were bursty in 1942-1945, and *atomic* was the buzz word in 1947 and 1959.

EMERGING TREND DETECTION (ETD)

ETD for Text Data

Unlike financial and statistical data typically found in an economist's trend detection portfolio, ETD in computer science often refers to the detection of trends in textual data, such as a collection of text documents and a stream of news feed. ETD takes a large collection of textual data as input and identifies topic areas that are previously unseen or are growing in importance with in the corpus (Kontostathis et al., 2003). This type of data mining can be instrumental in supporting the discovery of emerging trends within an industry and improving the understanding

Related Content

Toward a Significant E-Voting Adoption Model: The Digital Divide
Mohammad Kamel Alomariand Habib Ullah Khan (2022). *International Journal of Technology and Human Interaction (pp. 1-16).*
www.irma-international.org/article/toward-a-significant-e-voting-adoption-model/300283

Understanding Cognitive Processes in Educational Hypermedia
Patricia M. Boechler (2006). *Encyclopedia of Human Computer Interaction (pp. 648-651).*
www.irma-international.org/chapter/understanding-cognitive-processes-educational-hypermedia/13188

Mobile Screen Media Practices in Korea
HyeRyoung Ok (2015). *Encyclopedia of Mobile Phone Behavior (pp. 1103-1109).*
www.irma-international.org/chapter/mobile-screen-media-practices-in-korea/130220

The Role of Human Factors in Cybersecurity Vulnerabilities
Harish Chandra Vermaand Saurabh Srivastava (2025). *Human Impact on Security and Privacy: Network and Human Security, Social Media, and Devices (pp. 1-12).*
www.irma-international.org/chapter/the-role-of-human-factors-in-cybersecurity-vulnerabilities/358317

Exploring Social Aspects Influence on Change in Network Relationships: A Case Study of Digital Innovation
Jesper Svenssonand Carina Ihlström Eriksson (2012). *International Journal of Social and Organizational Dynamics in IT (pp. 14-33).*
www.irma-international.org/article/exploring-social-aspects-influence-change/76383