Chapter 64
Big Data Analytics on the Characteristic Equilibrium of Collective Opinions in Social Networks

Yingxu Wang
University of Calgary, Canada

Victor J. Wiebe
University of Calgary, Canada

ABSTRACT

Big data are products of human collective intelligence that are exponentially increasing in all facets of quantity, complexity, semantics, distribution, and processing costs in computer science, cognitive informatics, web-based computing, cloud computing, and computational intelligence. This paper presents fundamental big data analysis and mining technologies in the domain of social networks as a typical paradigm of big data engineering. A key principle of computational sociology known as the characteristic opinion equilibrium is revealed in social networks and electoral systems. A set of numerical and fuzzy models for collective opinion analyses is formally presented. Fuzzy data mining methodologies are rigorously described for collective opinion elicitation and benchmarking in order to enhance the conventional counting and statistical methodologies for big data analytics.

1. INTRODUCTION

The hierarchy of human knowledge is categorized at the levels of *data, information, knowledge,* and *intelligence* [Debenham, 1989; Bender, 1996; Wang, 2006, 2014a]. Big data is one of the fundamental phenomena of the information era of human societies [Jacobs, 2009; Snijders et al., 2012; Wang, 2014a; Wang & Wiebe, 2014]. Almost all fields and hierarchical levels of human activities generate exponentially increasing data, information, and knowledge. Therefore, big data analytics has become one of the

fundamental approaches to embody the abstraction and induction principles in rational inferences where discrete data represent continuous mechanisms and semantics.

Big data are extremely large-scaled data in terms of quantity, complexity, semantics, distribution, and processing costs in computer science, cognitive informatics, web-based computing, cloud computing, and computational intelligence. Big data and formal analytic theories are a pervasive demand across almost all fields of science, engineering, and everyday lives [Chicurel, 2000; Snijders et al., 2012; Wang, 2003, 2009a, 2014a]. Typical applications of big data methodologies in sciences are such as mathematics, number theories, neuroinformatics, computing systems, IT and web systems, brain science, memory capacity, genomes, linguistics, sociology, and management science [Wang, 2015; Wang & Berwick, 2012]. Paradigms of big data applications in engineering are such as Internet technologies, web systems, search engines, telecommunications, image processing, cognitive knowledge bases, multi-media databases, data mining, online text comprehension, machine translations, cognitive informatics, and cognitive robotics [Wang, 2003, 2007b, 2009b, 2009c, 2010, 2012a, 2013a, 2013b, 2015]. Big data in the modern society are fast approaching to the order of Petabyte (10^{15}) per year [Wiki, 2012].

Big data analytics in sociology and collective opinion elicitation in social networks are identified as an important filed where data are often complex, vague, incomplete, and counting-based [Wang & Wiebe, 2014]. Censuses and general elections are the traditional and typical domains that demand efficient big data analytic theories and methodologies beyond number counting and statistics [Emerson, 2013; Saari, 2000]. Among modern digital societies and social networks, popular opinion collection via online polls and voting systems becomes necessary for policy confirmation in general elections.

This paper presents formal models and methodologies of big data analytics for collective opinions and representative equilibrium in social networks and electoral systems. The cognitive and computing properties of big data are explored in Section 2. Potential pitfalls of Conventional counting-based voting methods for collective opinion elicitation and the majority rule embodiment are analyzed in Section 3. The characteristic equilibrium of collective opinions in social networks is revealed via big data analyses and numerical algorithms in Section 4. A set of fuzzy models for collective opinion elicitation and aggregation are rigorously described in Section 5. Case studies on applications of the formal methodologies of big data analytics are demonstrated in big poll data mining, collective opinion elicitation, and characteristic equilibrium determination.

2. PROPERTIES OF BIG DATA

The sources of big data are human collective intelligence. It is noteworthy that the first principle of mathematics is abstraction [Bender, 1996; Wang, 2009c, 2014a]. The essence of data is an abstract qualification or quantification of a real world entity or its attributes against a certain scale or benchmark. Therefore, data are fundamental information and knowledge of human civilization. The human capability for manipulating big data indicates the development of sciences and technologies in information, knowledge, and intelligence processing. Typical human activities that produce big data are such as many-to-many communications, massive downloads of data replications, digital image collections, and networked opinion forming.

Data modeling and representation in mathematics has been advanced from nonquantitative (O), binary (B), natural number (N), integer (Z), real (R), fuzzy (F) numbers, to hyper numbers (H) in line with the development of human civilization as illustrated in Figure 1 [Wang, 2014e]. Although decimal

16 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/big-data-analytics-on-the-characteristic-equilibrium-of-collective-opinions-in-social-networks/150222

Related Content

Measuring Semantic-Based Structural Similarity in Multi-Relational Networks
Yunchuan Sun, Rongfang Bieand Junsheng Zhang (2016). *International Journal of Data Warehousing and Mining (pp. 20-33).*
www.irma-international.org/article/measuring-semantic-based-structural-similarity-in-multi-relational-networks/143713

SLOD-BI: An Open Data Infrastructure for Enabling Social Business Intelligence
Rafael Berlanga, Lisette García-Moya, Victoria Nebot, María José Aramburu, Ismael Sanzand Dolores María Llidó (2015). *International Journal of Data Warehousing and Mining (pp. 1-28).*
www.irma-international.org/article/slod-bi/130664

A Predictive Modeling of Retail Satisfaction: A Data Mining Approach to Retail Service Industry
M. Hemalatha (2013). *Data Mining: Concepts, Methodologies, Tools, and Applications (pp. 1276-1290).*
www.irma-international.org/chapter/predictive-modeling-retail-satisfaction/73495

Spatial Clustering in SOLAP Systems to Enhance Map Visualization
Ricardo Silva, João Moura-Piresand Maribel Yasmina Santos (2012). *International Journal of Data Warehousing and Mining (pp. 23-43).*
www.irma-international.org/article/spatial-clustering-solap-systems-enhance/65572

Uncertainty-Based Clustering Algorithms for Large Data Sets
B. K. Tripathy, Hari Seethaand M. N. Murty (2018). *Modern Technologies for Big Data Classification and Clustering (pp. 1-33).*
www.irma-international.org/chapter/uncertainty-based-clustering-algorithms-for-large-data-sets/185977