Available online at www.sciencedirect.com

ScienceDirect

Procedia Computer Science 91 (2016) 452 – 459

Procedia Computer Science

Information Technology and Quantitative Management (ITQM 2016)

Combining Channel Theory, HowNet and Extension Model to Analyze Big Data

Weihua Li[a]*, Chunyan Yang[b]

[a]*School of Computer, Guangdong University of Technology, Guangzhou, 510006 China*
[b]*Research Institute of Extenics and Innovation Methods, Guangdong University of Technology, Guangzhou, 510006 China*

Abstract

Because the diversity of unstructured data has brought new challenges to big data analysis, this paper proposes to combine Channel theory, HowNet and extension model to improve big data analysis ability. The paper proposes a new method to process big data, which is based on the Channel theory idea and HowNet structure, in order to overcome the semantic conflicts of big data. In view of the problems that people are difficult to analyze their big data in order to get profits, the paper proposes a case study to show the effective of our method.

Peer-review under responsibility of the Organizing Committee of ITQM 2016

Keywors: Channel theory, HowNet, extension model, big data, analysis

1. Introduction

Big data does not only enable users to get information, but also cause difficulty for them to get really required information. This concerns semantics understanding of information. Heterogeneity, scale, timeliness, complexity, and privacy problems with big data impede progress at all phases of the pipeline that can create value from data [1]. The key to making big data initiatives a success lies within making the produced data more digestible and usable in decision making, rather than making it just 'more,' resulting in the creation of an environment wherein information is used to generate real impact[2]. Although there are a number of successful customer data analysis tools, but we cannot just analyze data without trying to improve the situation. In order to build analysis tools for companies to get value from big data, we focus not only on information volume, variety and velocity, but also on understanding big data and get value from it.

* Weihua Li. Tel.: +86-20-39322279; fax: +86-20-39322279.
E-mail address: lw@gdut.edu.cn.

Peer-review under responsibility of the Organizing Committee of ITQM 2016
doi:10.1016/j.procs.2016.07.118

This paper proposes to combine Channel theory, HowNet structure and extension model to analyze big data. We believe this will more effectively treat big data so that people can get values. As a case study, at the end of this paper, an example of out method is described to show the effects.

2. Channel Theory

Big data, with its increase in the quantity and variety of information, almost inevitably, these discrete resources use different terms to describe similar concepts, or even use identical terms to mean very different things, introducing confusion and error into their use. Barwise and Seligman proposed a very general qualitative theory of information flow (in distributed systems) [3]. It is a mathematic model that aims at establishing the laws that govern the flow of information. Information Flow Theory, which is also called Channel Theory, is a general theory of regularity that applies to the distributed information inherent in both natural world of biological and physical systems and the artificial world of computational systems [3]. It is based on the understanding that information flow results from regularities in a distributed system, and that it is by virtue of regularities among the connections that information of some components of a system carries information of other components.

In channel theory, each component of a distributed system is represented by a classification:

$$\mathbf{A}=<A, \sum_A, \vDash_A>$$

It consisting of a set A of objects to be classified, called tokens of A, a set $\sum_A$ of objects used to classify the tokens, called the types of A, and a binary relation, $\vDash_A$ between A and $\sum_A$ that tells one which token are classified as being of which types.

In channel theory, the notion of an infomorphism f: A⇆B gives a mathematical model of the whole-part relationship between instances of a whole, as modeled by a classification B, and that of a part, as modeled by a classification A.

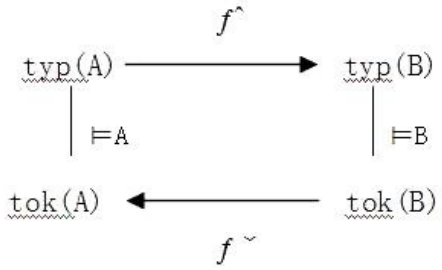

Fig. 1. Infomorphism

An information channel consists of an indexed family C = $\{f_i: \mathbf{A}⇆\mathbf{C}\}_{i∈I}$ of infomorphisms with a common codomain C, called the core of the channel.