

doi: 10.3969/j.issn.1007-7162.2014.03.002

大数据基元的 HBase 数据库存储模型与实现

李桥兴^{1,3,4}, 强保华², 杨春燕³

(1. 兰州理工大学 经济管理学院, 甘肃 兰州 730050; 2. 桂林电子科技大学 广西可信软件重点实验室, 广西 桂林 541004;
3. 广东工业大学 可拓学与创新方法研究所, 广东 广州 510006; 4. 兰州大学 管理学院, 甘肃 兰州 730000)

摘要: 大数据将在未来经济、社会和生活等领域产生深远影响,而大数据的整合存储研究为促进和深化其应用范围有重要的理论和实践意义. 本文利用分布式文件系统 HBase 的数据存储结构特征,采用可拓学基元对异构数据集进行整合处理并存储在 HBase 数据库. 通过提取数据特别是半结构化和非结构化数据的典型特征和属性取值并转换为基元后生成一个新的数据集,不但为数据分析和数据解释提供一种新的实现方式,而且为领域问题的策略生成提供大数据的研究思路 and 解决方案.

关键词: 大数据; 基元; 数据模型; 分布式文件; 可拓学

中图分类号: TP333

文献标志码: A

文章编号: 1007-7162(2014)03-0008-06

The Storage Model of Big Data Basic-elements in HBase Database and Its Realization

Li Qiao-xing^{1,3,4}, Qiang Bao-hua², Yang Chun-yan³

(1. School of Economics and Management, Lanzhou University of Technology, Lanzhou 730050, China;
2. Guangxi Key Lab of Trusted Software, Guilin University of Electronic Technology, Guilin 541004, China;
3. Institute of Extenics and Innovation Method, Guangdong University of Technology, Guangzhou 510006, China;
4. School of Management, Lanzhou University, Lanzhou 730000, China)

Abstract: Big data will have a profound impact on economics, society and life in the near future, and the research on integration and storage of big data may play an important theoretical and practical role in promoting and deepening the application scope of big data. It utilized the data storage structure of the distributed file system named as HBase and the basic-element of Extenics to integrate the heterogeneous data sets, and then stored the processed data set in HBase database. The new data set, which was obtained by extracting the typical characteristics as well as their value of data, especially the semi-structured and un-structured data, provides not only a new way for the analysis and interpretation of data but also the research ideas and strategy generation for the professional issues from the perspective of big data.

Key words: big data; basic-element; data model; distributed file; Extenics

随着移动互联网、物联网、社交网络 and 智能终端等迅速成长而涌现大数据,要求计算机具有处理大数据的技术体系,包括大数据采集与预处理、大数据存储与管理、大数据计算模式与系统、大数据分析与挖掘及大数据可视化计算等^[1-2]. 显然,数据采集、预处理和存储等为后续分析与计算提供必要前提和有效保障. 大数据管理需要根据计算机架构调整新的数据整合与存储模式. 文献 [2] 指出,现有技术已经能够将融合处理后的数据转换为规范的数据结构

并存储在专门的数据管理系统中,如文件或数据库等,形成专门的数据集. 由于数据集的多源异构特征,给大数据的分析和计算等带来了不确定和多样性,导致数据质量存在差异. 另外,异构数据集需要做进一步集成处理或整合处理生成一个新的数据集,为后续查询和分析处理提供统一的数据视图. 在实际应用中,结构化数据、半结构化数据和非结构化数据是对存储形式的一种数据类型分析,有助于企业细分行业案例,更好地为合作伙伴提供应用实施

收稿日期: 2014-07-07

基金项目: 国家自然科学基金资助项目 (61273306)

作者简介: 李桥兴 (1973-), 男, 副教授, 博士, 主要研究方向为系统科学与复杂性管理、资源经济与管理 and 城市化管理等.

方案. 因此,研究大数据的整合存储技术有重要的理论意义和使用价值.

可拓学采用形式化模型研究事物拓展的可能性和开拓创新的规律与方法并用于处理矛盾问题. 基元就是描述物、事和关系的形式化工具,包括物元、事元和关系元^[3]. 采用基元可以把客观事物及其相互关系形式化表达,并经过计算机编程实现智能化. 文献[4-6]对利用计算机实现可拓学基元的存储与管理作出了有益的尝试. 一些学者采用可拓学在自助游、数据挖掘、工业设计、企业自主创新和客户管理等领域开始探索实现计算机智能化处理^[7-12]. 鉴于可拓学在智能化处理矛盾问题的可行性和大数据计算机存储与管理的实践需要,利用基元形式化体系研究处理大数据,特别是半结构和非结构化数据有很好的理论意义与实践价值.

1 基本概念

根据研究需要,本文首先介绍基元和分布式文件系统 HBase 数据模型的基本知识.

可拓学认为,客观世界的任何物、事和关系都可以采用基元形式化描述. 文献[3]介绍了基元(包括物元、事元和关系元)理论. 在文献[3]的基础上,文献[3]定义了多元多维基元,如下:

定义1 设 $O = \{O_1, O_2, \dots, O_p\}$ 为研究对象集(包括物的集合、事的集合和关系的集合等), $C = \{c_1, c_2, \dots, c_q\}$ 为研究对象的特征集合,对象 O_i 关于特征 c_j 的取值为 v_{ij} ($i = 1, 2, \dots, p; j = 1, 2, \dots, q$), 则称

$$B_{pq} = \begin{bmatrix} c_1 & v_{11} \\ O_1 & \vdots & \vdots \\ & c_q & v_{1q} \\ \cdots & \cdots & \cdots \\ & c_1 & v_{p1} \\ O_p & \vdots & \vdots \\ & c_q & v_{pq} \end{bmatrix} \text{ 或 } B_{pq} = \begin{bmatrix} & c_1 & c_2 & \cdots & c_q \\ O_1 & v_{11} & v_{12} & \cdots & v_{1q} \\ O_2 & v_{21} & v_{22} & \cdots & v_{2q} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ O_p & v_{p1} & v_{p2} & \cdots & v_{pq} \end{bmatrix}$$

为 p 元 q 维基元,其中研究对象称为基元的元;对象的特征称为基元的维. 若某个研究对象 O_i 不具备或不需要考察某个特征 c_j , 则令 $v_{ij} \equiv NA$, 表示取值无意义或不存在.

显然,当 $p = q = 1$ 时, B_{11} 就是基元的最初定义,即一元一维基元.

定义2 基元 B_{pq} 如定义1, 称

$$B_{pq}^i = \begin{bmatrix} O_i & c_1 & v_{i1} \\ & c_2 & v_{i2} \\ & \vdots & \vdots \\ & c_q & v_{iq} \end{bmatrix} \text{ 或 } B_{pq}^i = \begin{pmatrix} O_i & c_1 & c_2 & \cdots & c_q \\ & v_{i1} & v_{i2} & \cdots & v_{iq} \end{pmatrix}$$

为 B_{pq} 的第 i 元 q 维基元部 ($i = 1, \dots, p$).

事实上,第 i 元 q 维基元部 B_{pq}^i 即为 q 维基元.

数据需要存储在计算机,而 HBase 是一个分布式、面向列且适合非结构化数据存储的开源数据库^[4]. HBase 以表的形式存储数据,每个表由行和列组成,每个列属于一个特定的列族. 表中由行和列确定的存储单元称为一个元素,每个元素保存了同一份数据的多个版本,由时间戳来标识. 表1举例说明了网站 www.cnn.com 的数据存放逻辑视图^[4].

表1 数据存放逻辑视图

Tab.1 The logic view of date storage

行关键字	时间戳	列 contents	列 anchor	列 mime
com. cnn. www	T9		"anchor: cnnsi. com"	"CNN"
	T8		"anchor: my. look. ca"	"CNN. com"
	T6	"< html >..."		"text/html"
	T5	"< html >..."		
	T3	"< html >..."		

显然,表1中仅有一行数据,行关键字为 com. cnn. www. 在 HBase 数据库中,每一行相当于传统数据库中的一个表,而行关键字就是表名. 对表中行数据的每一次逻辑修改都有一个时间戳关联对应. 表1中共有3个列定义: < contents >、< anchor > 和 < mime >. 列定义为: < 列族 >: < 标签 >. 通过这两部分可以唯一指定一个数据的存储列. 对列族的定义和修改需要管理员权限,而标签可以在任何时候添加.

除逻辑存储模型外,表2给出 HBase 的物理存储模型. HBase 的物理模型是按列存储的稀疏行/列矩阵,实际上就是把概念模型中的一个行进行分割并按照列族存储^[4].

2 基元的 HBase 数据库存储模型

可拓学基元使矛盾问题的处理能够被计算机智能化实现,而实现智能化处理的前提是解决基元数据的存储问题. 根据可拓学基元的形式化表示方法

表2 数据在物理上的存储方式
Tab.2 The physical storage form of data

行关键字	时间戳	列 contents
com. cnn. www	T5	"<html>..."
	T4	"<html>..."
	T3	"<html>..."
行关键字	时间戳	列 anchor
com. cnn. www	T8	"anchor: cnnsi. com" "CNN"
	T7	"anchor: my. look. ca" "CNN. com"
行关键字	时间戳	列 anchor
com. cnn. www	T6	"text/html"

表3 基元存放逻辑视图

Tab.3 The logic view of basic-element storage

行基元	时间戳	列对象 O				列特征 c_1		...	列特征 c_q	
B_{pq}	T_n	O_1^n	...	O_p^n	v_{11}^n	...	v_{p1}^n	...	v_{1q}^n	v_{pq}^n
	...	...	...	...	...	...	...	...	...	...
	t_2	O_1^2	...	O_p^2	v_{11}^2	...	v_{p1}^2	...	v_{1q}^2	v_{pq}^2
	t_1	O_1	...	O_p	v_{11}	...	v_{p1}	...	v_{1q}	v_{pq}

元存放的逻辑视图中,列定义有两种形式,分别为:
<列对象族>: <对象标签>和<列特征族>: <特征标签>. 在表3中, t_1 时刻对应的对象标签为基元的初始对象 O_i , 并且特征标签为各对象 O_i 关于特征 c_j 的初始取值 v_{ij} ($i=1, \dots, p; j=1, \dots, q$); 其他时间戳 $t_2, \dots, T_n$ 对应的对象标签和特征标签分别为初始对象和特征初始值在各时刻的改变值, 即在 HBase 数据基元库中 t_1 时刻存储的基元为初始基元 B_{pq} , 其他时刻对应的基元为经过修改后存储的基元 B_{pq} . 同理, 根据表2和表3可以得到 HBase 基元数据库的物理存储方式, 如表4所示.

3 大数据的基元描述及在 HBase 数据库的存储实现

多源异构是大数据的重要特征之一. 大数据的异构性体现在结构化、半结构化和非结构化. 大数据应用的关键是要对数据进行融合并有效表示, 因此本文首先探讨异构大数据的融合表示. 鉴于可拓学基元能够形式化描述客观世界的任何物、事及事物间的关系, 以及 HBase 数据库的存储模型与基元形式化模型的相似性, 本文将分别以数据案例探讨结构化、半结构化和非结构化数据的基元存储模型.

3.1 结构化数据的 HBase 基元数据库存储方式

结构化数据指存储在数据库里的行数据, 可以用二维表结构来逻辑表达实现, 如医疗 HIS 数据库、企业 ERP 数据和财务数据等. 结构化数据的基元表

和 HBase 数据库的存储模型, 表3给出了基元在 HBase 数据库存放的逻辑模型.

显然, 表3中 HBase 表格的行标识对应于基元名称 B_{pq} , 而列族则对应于基元的对象及其特征 c_j , 并且每一个标签对应于基元的对象名 O_i 和对象 O_i 关于特征 c_j 的取值 v_{ij} ($i=1, \dots, p; j=1, \dots, q$). 本文称基元在 HBase 数据库的存储集合为 HBase 基元数据库, 基元名称 B_{pq} 即为行关键字, 称为行基元; 把基元对象所在的列和特征 c_j 作为列族, 分别称为列对象和列特征; 对象名称 O_i 和 O_i 关于 c_j 的特征值 v_{ij} 为标签, 分别称为对象标签和特征标签. 由此, 在基

表4 基元数据在物理上的存储方式

Tab.4 The physical storage form of basic-element data

行基元	时间戳	列对象 O			
B_{pq}	T_n	O_1^n	...	O_p^n	
	...	...	...	...	
	t_2	O_1^2	...	O_p^2	
	t_1	O_1	...	O_p	
行基元	时间戳	列特征 c_1			
B_{pq}	T_n	v_{11}^n	...	v_{p1}^n	
	...	...	...	...	
	t_2	v_{11}^2	...	v_{p1}^2	
	t_1	v_{11}	...	v_{p1}	
行基元	时间戳	列特征...			
B_{pq}	T_n	...	...	...	
	...	...	...	...	
	t_2	...	...	...	
	t_1	...	...	...	
行基元	时间戳	列特征 c_q			
B_{pq}	T_n	v_{1q}^n	...	v_{pq}^n	
	...	...	...	...	
	t_2	v_{1q}^2	...	v_{pq}^2	
	t_1	v_{1q}	...	v_{pq}	

示比较显而易见, 其属性特征一般在表格中已经给出. 表5即为典型的结构化数据, 其 HBase 基元数据库存储的逻辑视图和物理模式分别见表6和表7.

3.2 非结构化数据的 HBase 基元数据库存储方式

非结构化数据一般指不方便用数据库二维逻辑表来表现的数据, 其字段长度可变且每个字段的记录又可以由可重复或不可重复的子字段构成. 非结

结构化数据包括视频、音频、图片、图像、文档、文本等形式。现实中的典型案例如医疗影像系统、教育视频点播、视频监控和科学文献等。对于非结构化数据的基元描述,一般是以数据库题名为行基元,然后根据非结构化数据的复杂程度决定是否要分成若干组成部分(若需要划分,则提取列对象族,并以组成部分名称为对象标签;若无需划分则略列对象族),再提取非结构化数据的重要属性作为列特征族,最后对数据库资料进行内容分析并根据列特征提取其特征

征标签。

表5 基于不同方法的语义匹配实验效果表

Tab.5 The semantic matching talbe based on different methods

方法	平均查准率/%	平均查全率/%	平均训练时间/s
方法 [1]	83.7	98.3	16.243
方法 [2]	92.6	100	18.126
方法 [3]	100	67.9	<0.100

表6 结构化数据基元存放逻辑视图

Tab.6 The logic view of struatred basic-element data storage

行基元	时间戳	列对象 O	平均查准率 c_1 /%	平均查全率 c_2 /%	平均训练时间 c_3 /s
基于不同方法的语义匹配实验效果表	t_1	方法 [1] 方法 [2] 方法 [3]	83.7 92.6 100	98.3 100 67.9	16.243 18.126 <0.100

表7 结构化数据基元在物理上的存储方式

Tab.7 The physical storage form of structured basic-element data

行基元	时间戳	列对象 O
基于不同方法的语义匹配实验效果表	t_1	方法 [1] 方法 [2] 方法 [3]
行基元	时间戳	平均查准率(%) c_1
基于不同方法的语义匹配实验效果表	t_1	83.7 92.6 100
行基元	时间戳	平均查全率(%) c_2
基于不同方法的语义匹配实验效果表	t_1	98.3 100 67.9
行基元	时间戳	平均训练时间 c_3
基于不同方法的语义匹配实验效果表	t_1	16.243 18.126 <0.100

本文以文献 [5] 为例说明非结构化数据的基元存储方式。一般来说,科学研究文献有一定的书写格式,如文章题名、作者名、作者单位、摘要、关键词、正文、参考文献等。科学文献的题名可以作为行基元,其

各组成部分可作为列特征族,而组成部分的具体内容作为其特征标签。表8和表9分别为文献 [5] 的 HBase 基元数据库存放的逻辑视图和物理模式。

表8 非结构化数据基元存放逻辑视图

Tab.8 The physical view of unstructured basic-element data

行基元	时间戳	作者名	作者单位	出版刊物	出版年	摘要内容 c_5	关键词 c_6
		c_1	c_2	c_3	c_4		
可拓学的基础理论与方法体系 O	t_1	蔡文杨春燕	广东工业大学	科学通报	2013	人类的历史,是一部处理矛盾问题的发展史。根据一定的程序解决矛盾问题,利用计算机和网络,帮助人们处理矛盾问题,称为矛盾问题智能化处理。为了实现矛盾问题智能化处理,必须研究如何用形式化表示矛盾问题及其解决过程,研究处理矛盾问题的规律,建立处理矛盾问题的方法体系,研制处理矛盾问题的计算机软件,这就产生了一门新的学科——可拓学。本文介绍可拓学的基础理论、方法体系及其应用情况,并讨论其科学意义与未来发展。	可拓学 可拓论 可拓创新方法 基元 可拓集 可拓逻辑

3.3 半结构化数据的 HBase 数据库基元存储方式
半结构化数据是指介于结构化和非结构化之间的数据,包括邮件、HTML 和各种报表等,其现实中的

典型场景如邮件系统、WEB 集群和档案系统等。根据具体数据类型,有的半结构化数据具有比较明显的属性特征,有的半结构化数据的属性特征需要提炼。

表9 非结构化数据基元在物理上的存储方式

Tab.9 The physical storage form of unstructured basic-element data

行基元	时间戳	作者名 c_1
可拓学的基础理论与方法体系 O	t_1	蔡文,杨春燕
行基元	时间戳	作者单位 c_2
可拓学的基础理论与方法体系 O	t_1	广东工业大学
行基元	时间戳	出版刊物 c_3
可拓学的基础理论与方法体系 O	t_1	科学通报
行基元	时间戳	出版年 c_4
可拓学的基础理论与方法体系 O	t_1	出版年 c_4
行基元	时间戳	摘要内容 c_5
可拓学的基础理论与方法体系 O	t_1	人类的历史,是一部处理矛盾问题的发展史. 根据一定的程序解决矛盾问题,利用计算机和网络,帮助人们处理矛盾问题,称为矛盾问题智能化处理. 为了实现矛盾问题智能化处理,必须研究如何用形式化表示矛盾问题及其解决过程,研究处理矛盾问题的规律,建立处理矛盾问题的方法体系,研制处理矛盾问题的计算机软件,这就产生了一门新的学科——可拓学. 本文介绍可拓学的基础理论、方法体系及其应用情况,并讨论其科学意义与未来发展.
行基元	时间戳	关键词 c_6
可拓学的基础理论与方法体系 O	t_1	可拓学、可拓论、可拓创新方法、基元、可拓集和可拓逻辑

本文以下面的网页内容为例说明半结构化数据的基元存储模式.

```
<? xml version = "1.0" encoding = "UTF-8"? >
<! DOCTYPE 应聘信息 SYSTEM "fourDTD.
dtd" >
```

```
<应聘信息>
```

```
<应聘者>
```

```
<姓名>张三</姓名>
```

```
<硕士>湖南大学计算机理学硕士<硕士>
```

```
<奖励>获得计算机竞赛一等奖1次</奖励>
```

```
<奖励>获得3次一等奖学金</奖励>
```

```
<处分>被警告处分1次</处分>
```

```
<性别>男</性别>
```

```
</应聘者>
```

```
<应聘者>
```

```
<姓名>李翠花</姓名>
```

```
<博士>武汉大学自动化专业工学博士</博
士>
```

```
<奖励>获得5次一等奖学金</奖励>
```

```
<性别>女</性别>
```

```
</应聘者>
```

根据上面的内容,可建立该网页的 HBase 基元数据库存储的逻辑视图和物理模式分别如表 10 和表 11 所示.

表10 半结构化数据基元存放逻辑视图

Tab.10 The logic view of semi-structured basic-element data storage

行基元	时间戳	姓名 c_1	性别 c_2	学位 c_3	奖励 c_4	处分 c_5
	t_2					无
应聘 信息	t_1	张三 李翠花	男 女	湖南大学 计算机理学硕士	武汉大学 自动化专业工 学博士	1次计算机竞赛 一等奖和3次一 等奖学金

4 结论与展望

在大数据背景下,数据已经成为类似人力资源、土地资源和能源资源等一样宝贵的资源财富. 利用

数据资源的前提是提供可靠的数据质量和整合元数据资料生成数据库. 由于可拓学基元描述物、事和关系的形式化方式与 HBase 数据库存储模式相似,本文探讨了大数据基元在 HBase 数据库的存储模型与

表 11 非结构化数据基元在物理上的存储方式

Tab. 11 The physical storage form of unstructured basic-element data

行基元	时间戳	姓名 c_1	
应聘信息	t_1	张三	李翠花
行基元	时间戳	性别 c_2	
应聘信息	t_1	男	女
行基元	时间戳	学位 c_3	
应聘信息	t_1	湖南大学计算机理学硕士	武汉大学自动化专业工学博士
行基元	时间戳	奖励 c_4	
应聘信息	t_1	1 次计算机竞赛一等奖和 3 次一等奖学金	5 次一等奖学金
行基元	时间戳	处分 c_5	
应聘信息	t_2	无	
	t_1	1 次警告处分	

实现. 在实际应用中, 大数据基元也可以在其他类型数据库中存储与实现. 大数据基元的 HBase 存储方式不但使异构数据集具有统一的模式, 而且可以根据时间戳了解数据的变化情况. 鉴于可拓学在解决矛盾问题中的独特作用, 大数据的基元处理及其在 HBase 中的存储实现不但为大数据的分析和计算等提供新的实现方式, 而且为后续探讨如何从大数据中获取可拓知识以及如何利用从大数据中获得的知识生成解决矛盾问题的策略等提供前提条件.

参考文献:

- [1] 中国电子科学研究院学报编辑部. 大数据时代 [J]. 中国电子科学研究院学报, 2013(1): 27-31.
Editorial department of the Journal of Chinese Academy of Electronic Science. The era of big data [J]. The Journal of Chinese Academy of Electronic Science, 2013(1): 27-31.
- [2] 中国计算机学会大数据专家委员会. 中国大数据技术与产业发展白皮书 [M]. 2013.
- [3] 杨春燕, 蔡文. 可拓工程 [M]. 北京: 科学出版社, 2007.
- [4] 李立希, 杨春燕, 李铨汶. 可拓策略生成系统 [M]. 北京: 科学出版社, 2006.
- [5] 李杨, 谢光强. 可拓基元的形式化表示与实现 [J]. 软件导刊, 2013, 12(7): 37-39.
Li Y, Xie G Q. Formative representation and implementation of Extension basic-element [J]. Software Guide, 2013, 12(7): 37-39.
- [6] 李卫华, 杨春燕. Agent 识别矛盾问题核问题的方案研究 [J]. 计算机科学与工程, 2010(8): 127-129.
Li W H, Yang C Y. The scheme research for identifying the kernel problem of the contradiction problems with Agent [J]. Computer Engineering and Science, 2010(8): 127-129.
- [7] 叶广仔, 李卫华, 李淑飞. 可拓策略生成系统的构件化设计与实现 [J]. 智能系统学报, 2010, 5(4): 366-371.
Ye G Z, Li W H, Li S F. Components based design of an extensional strategy generation system [J]. CAAI Transaction on Intelligent Systems, 2010, 5(4): 366-371.
- [8] 杨春燕, 李小妹, 陈文伟, 等. 可拓数据挖掘方法及其计算机实现 [M]. 广州: 广东高等教育出版社, 2010.
- [9] 李承晓, 李卫华. 租房可拓策略生成系统 [J]. 智能系统学报, 2011, 6(3): 272-278.
Li C X, Li W H. Research on a tenement extension strategy generation system [J]. CAAI Transaction on Intelligent Systems, 2011, 6(3): 272-278.
- [10] 李兴森, 李萍, 朱正祥. 可拓思维辅助企业管理创新 [J]. 科技智囊, 2009(2): 40-43.
Li X S, Li P, Zhu Z X. Extension thinking aids enterprise management innovation [J]. Think Tank of Science and Technology, 2009(2): 40-43.
- [11] 赵燕伟, 苏楠. 可拓设计 [M]. 北京: 科学出版社, 2010.
- [12] 李兴森, 朱正祥, 刘艳彬. 预防客户流失的可拓转化策略研究 [J]. 广东工业大学学报, 2012, 29(3): 18-22.
Li X S, Zhu Z X, Liu Y B. Research on the extension transformation strategy for customer retention [J]. Journal of Guangdong University of Technology, 2012, 29(3): 18-22.
- [13] 李桥兴. 多元多维基元及高阶多元多维可拓集合构造 [J]. 广东工业大学学报, 2009, 26(4): 84-86.
Li Q X. The method to construct multivariate and multidimensional basic-element and extensible set with high order and multivariate and multi-dimensions [J]. Journal of Guangdong University of Technology, 2009, 26(4): 84-86.
- [14] 张德丰. 大数据走向云计算 [M]. 北京: 人民邮电出版社, 2014.
- [15] 蔡文, 杨春燕. 可拓学的基础理论与方法体系 [J]. 科学通报, 2013, 58(13): 1190-1199.
Cai W, Yang C Y. Basic theory and methodology on Extensionics [J]. Chinese Science Bulletin, 2013, 58(13): 1190-1199.