

中华人民共和国国家标准

GB/T 41818—2022

信息技术 大数据 面向分析的 数据存储与检索技术要求

Information technology—Big data—Technical requirements for
analysis-oriented data storage and retrieval

2022-10-12 发布

2023-05-01 实施

国家市场监督管理总局 发布
国家标准化管理委员会

目 次

前言	Ⅲ
1 范围	1
2 规范性引用文件	1
3 术语和定义	1
4 缩略语	1
5 概述	2
6 数据表结构	2
6.1 数据表	2
6.2 子表	2
6.3 行组	3
6.4 列组	3
6.5 数据页	4
7 索引技术要求	4
8 数据存储技术要求	4
8.1 基本要求	4
8.2 存储布局实现	4
8.3 数据处理	5
9 数据检索技术要求	6
附录 A (资料性) 面向分析的数据存储与检索应用示例	7
参考文献	9

前 言

本文件按照 GB/T 1.1—2020《标准化工作导则 第 1 部分：标准化文件的结构和起草规则》的规定起草。

请注意本文件的某些内容可能涉及专利。本文件的发布机构不承担识别专利的责任。

本文件由全国信息技术标准化技术委员会(SAC/TC 28)提出并归口。

本文件起草单位：华为技术有限公司、北京大学、中国电子技术标准化研究院、中国人民大学、北京航天智造科技发展有限公司、中汽研汽车检验中心(天津)有限公司、成都中科大旗软件股份有限公司、浪潮软件科技有限公司、江苏赛西科技发展有限公司、深圳赛西信息技术有限公司、浪潮电子信息产业股份有限公司、北京易华录信息技术股份有限公司、杭州中奥科技有限公司、方正国际软件(北京)有限公司、山东黄河河务局山东黄河信息中心、山东省计算中心(国家超级计算济南中心)、四川大学华西医院、湖南财政经济学院、杭州市第七人民医院、北京工业大学、北京理工大学、中山大学、桂林电子科技大学、天津大学、中冶赛迪工程技术股份有限公司、北京能科瑞元数字技术有限公司、北京启迪数字科技集团有限公司、中国人民解放军国防科技大学、西北工业大学、西安电子科技大学、腾讯科技(深圳)有限公司、河南云政数据管理有限公司、特斯联科技集团有限公司、深圳龙岗智能视听研究院、星环信息科技(上海)有限公司。

本文件主要起草人：梅宏、杜小勇、符海芳、陈亮、范科峰、张群、赵华、赵俊峰、王亚沙、刘驰、尹卓、王为中、曹幼林、卞昊穹、金国栋、陈跃国、李民东、刘哲、邹萍、殷晋、贺可勋、冯谦、李武鸿、郑申俊、李冰、张亮、石征、袁玥、赵斌、李威、王树良、洪江、黄先芝、胡清、王凌、陶智敏、李振东、张煜、逢锦山、张媛、潘彬、孙光、章俊航、杨绍武、史殿习、杨震、于海阳、万海、何倩、雷建军、李斐、潘兆庆、彭勃、刘国杰、张超超、贾晓杰、沈丽丽、张星星、邓乔、黎方学、车伟伟、肖学文、毛尚伟、杨刚、姚远、杨钰、张大鹏、杨洪山。

信息技术 大数据 面向分析的 数据存储与检索技术要求

1 范围

本文件规定了支撑多应用融合分析的大数据列式存储与检索技术要求。
本文件适用于面向分析的数据存储与检索的系统或子系统的设计、开发和使用。

2 规范性引用文件

下列文件中的内容通过文中的规范性引用而构成本文件必不可少的条款。其中，注日期的引用文件，仅该日期对应的版本适用于本文件；不注日期的引用文件，其最新版本（包括所有的修改单）适用于本文件。

GB/T 35295—2017 信息技术 大数据 术语

3 术语和定义

GB/T 35295—2017 界定的以及下列术语和定义适用于本文件。

3.1

元数据 metadata

定义和描述其他数据的数据。

[来源：GB/T 18391.1—2009,3.2.16]

3.2

存储布局 storage layout

数据在存储系统中的存储规划和安排。

注：通常包括各数据单元在存储系统中相对的存储顺序和存储位置的偏移等。

3.3

检索 retrieval

使用一定方法或工具，从数据集合中找出用户所需数据或数据相关信息的过程。

4 缩略语

下列缩略语适用于本文件。

CSV：逗号分隔值(Comma Separated Values)

EB：百亿亿字节(Exa Byte)

JSON：脚本对象(JavaScript Object Notation)

OLAP：联机分析处理(On-line Analytical Processing)

PB：千万亿字节(Peta Byte)

SQL：结构化查询语言(Structured Query Language)

WiFi:无线保真(Wireless Fidelity)

5 概述

在大数据场景中,数据分析是通往大数据应用的首要步骤,需要简便、快捷和准确的数据组织和管理方法,这种组织和管理主要体现在存储和检索方面。检索活动与数据分析中的数据提取直接关联,存储活动为数据分析提供便捷的数据组织。数据存储表的示例见附录 A。

本文件描述的数据表为大数据场景下面向分析的数据存储和检索提供数据组织和管理方面的支持。

6 数据表结构

6.1 数据表

数据表是对数据进行组织和管理的一种数据分层次编排结构和多层结构,其组成部分包括:数据子表(以下简称“子表”)、数据行组(以下简称“行组”)、数据列组(以下简称“列组”)和数据页,如图 1 所示。一张数据表包含 1 张或多张子表,主要用于对数据进行面向分析的组织和管理,数据表的元数据独立于数据表外存储。这些组成部分统称为数据单元。

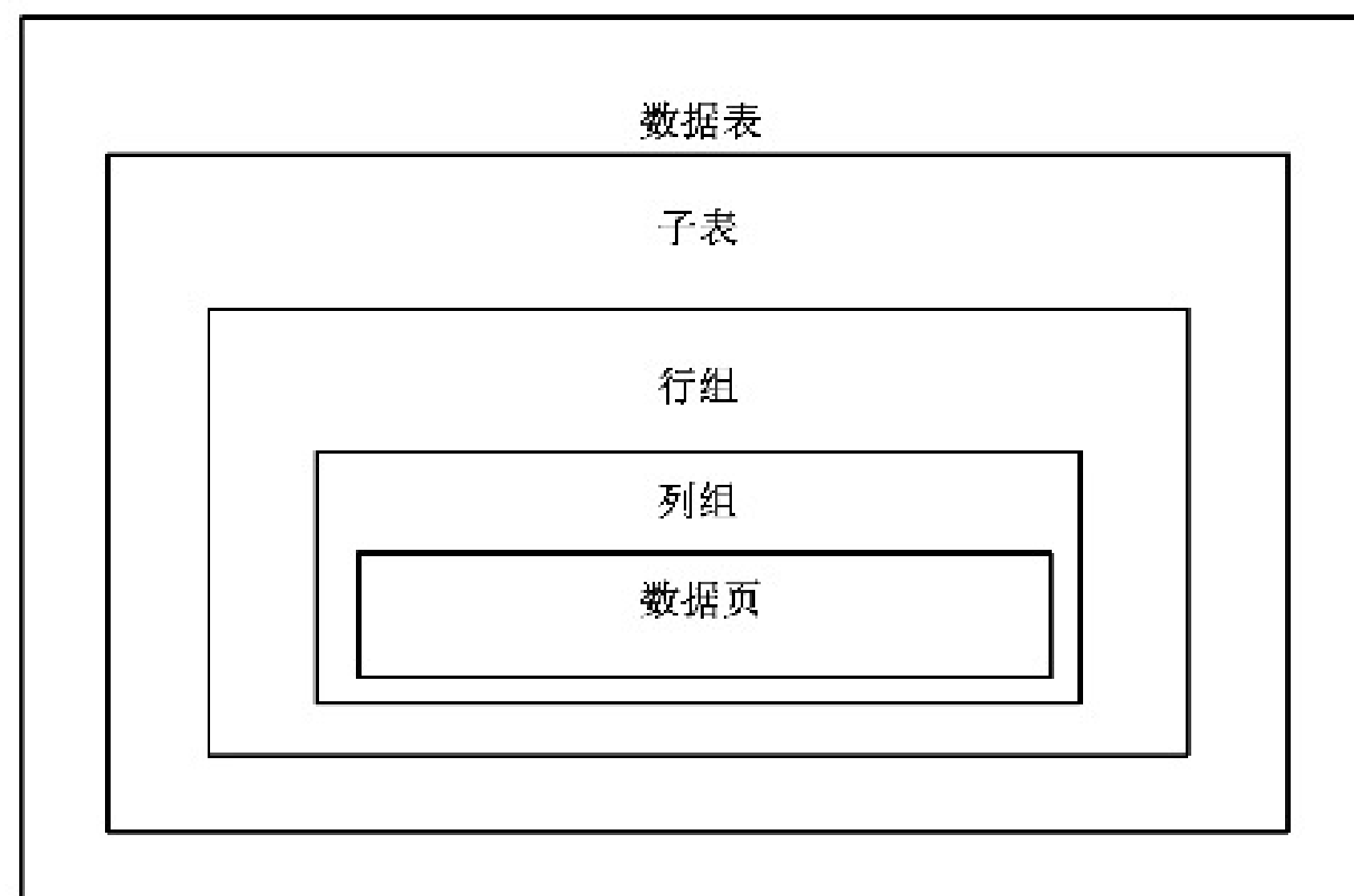


图 1 数据表结构图

6.2 子表

子表是组织和管理数据的最大数据单元,由子表头部、行组、子表尾部组成,一张子表内的数据可划分为 1~ n 个行组。子表的一般结构如图 2 所示。

子表头部给出子表的标识、版本信息以及子表的元数据。子表尾部给出子表的行组/列组索引,以及行组位置信息。行组索引由多个列组索引组成。

子表索引由行组索引组成。首次使用时,由行组索引记录汇聚而成,可缓存在内存中。

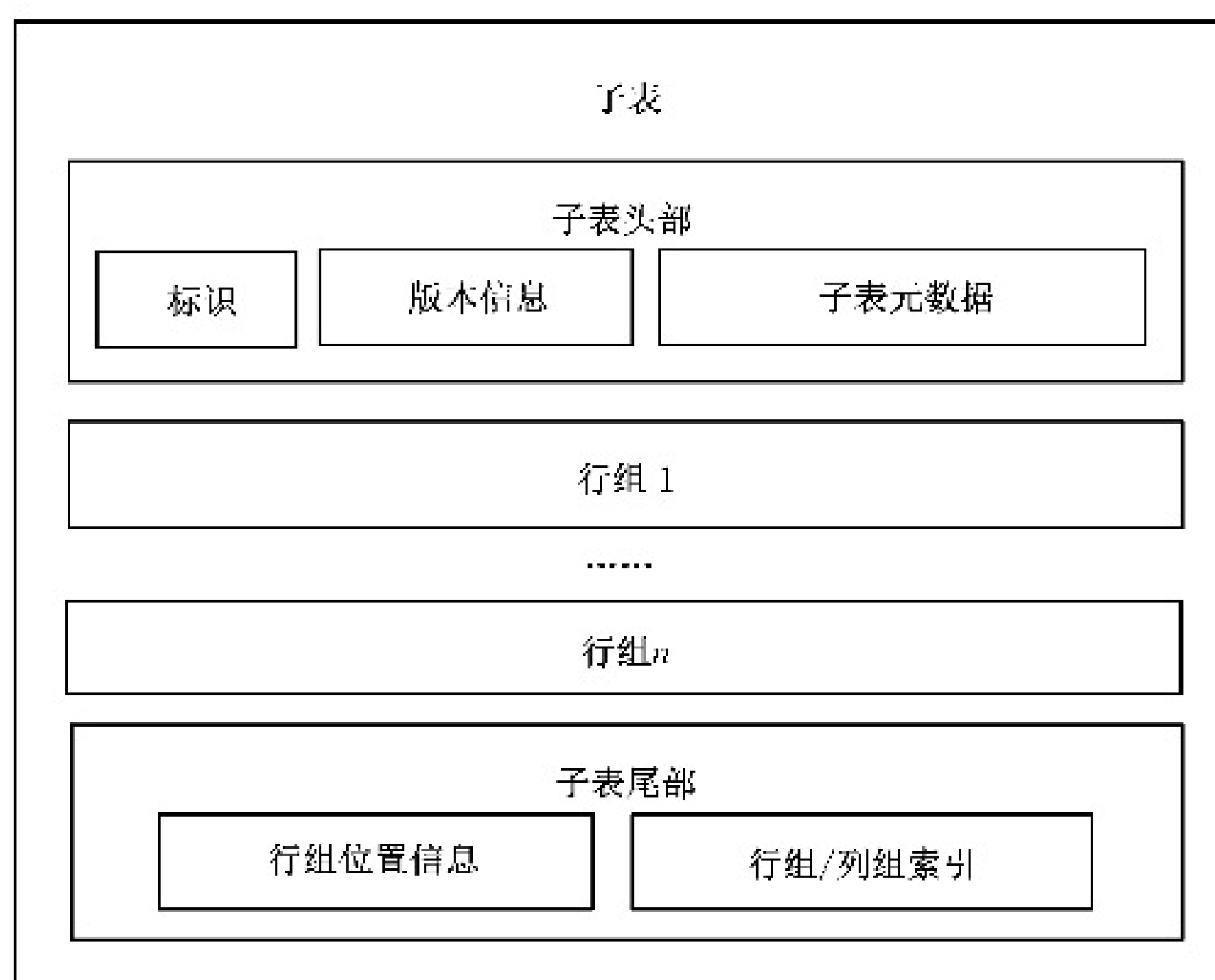


图 2 子表结构图

6.3 行组

行组是子表横向划分的最小单元,一个行组包含 $1 \sim L$ 行的数据 ($L \geq 1$),可划分为 $1 \sim M$ 个列组 ($M \geq 1$)。行组由行组元数据、数据页索引和列组构成,其中,行组元数据包含:列组元数据、列组位置和列组内数据页位置等信息,列组元数据包含了数据页数据的大小、编码方式等信息。一般行组结构如图 3 所示。

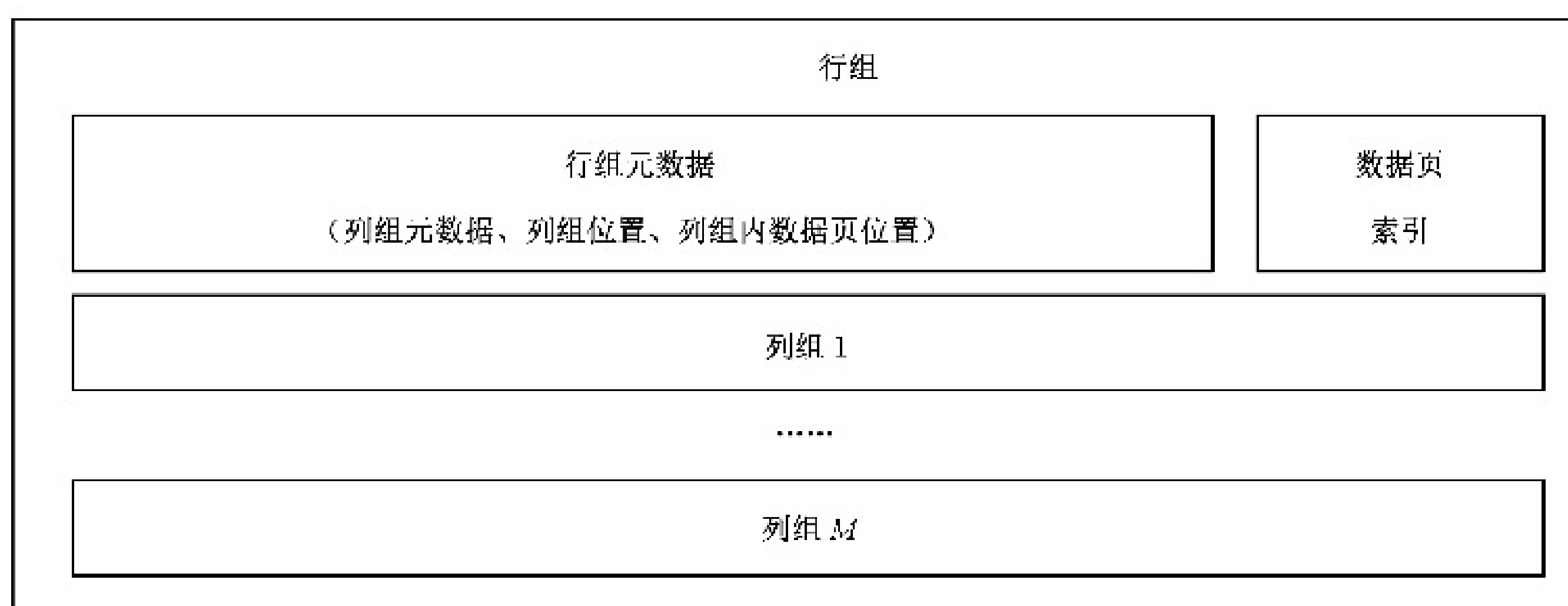


图 3 行组结构图

6.4 列组

列组是行组内数据纵向划分的最小单元,一个列组包含该行组内 $1 \sim P$ 列的数据 ($P \geq 1$),列组可划分为 $1 \sim Q$ 个数据页 ($Q \geq 1$)。列组的一般结构如图 4 所示。



图 4 列组结构图

6.5 数据页

数据页是列组内数据读写和数据过滤的最小单元,是一段裸数据,大小和格式等信息由行组中的列组元数据信息提供。

7 索引技术要求

本文件中的数据索引是快速定位所需数据的检索方式,包括:子表索引、行组或列组索引、数据页索引等。索引技术要求如下:

- a) 应支持行组或列组级别的索引,通过索引可获得需要的行组,如倒排索引;
- b) 应支持数据页级别的索引,通过索引可获得需要的数据页,如最大最小索引;
- c) 应支持多个列组的索引组合;
- d) 应支持主索引、时空索引、多值列索引等多种索引方式进行组合;
- e) 宜支持索引构建后数据的更新和删除;
- f) 宜支持子表级别的索引,通过索引获得需要的子表,如二叉树索引;
- g) 宜支持异步索引构建能力。

注:异步索引构建指的是用户先将数据入库,再利用业务闲时对数据构建索引。

8 数据存储技术要求

8.1 基本要求

数据存储的基本技术要求如下:

- a) 应支持数据按行组列组存储;
- b) 应支持数据持久化保存;
- c) 应支持文件存储、对象存储、内存存储等不同数据存储系统;
- d) 应支持依据不同维度对数据进行分类存储,如数据使用频率等;
- e) 宜支持总量达 EB 级,单表数据达万亿行级别的数据存储和检索;
- f) 宜支持对单条数据记录进行更新和删除;
- g) 宜基于第 6 章描述的数据表结构,设计和规划用于组织和管理数据存储的总体方案。

8.2 存储布局实现

存储布局实现的技术要求如下:

- a) 应随检索需求的重大变化及时调整存储布局;
- b) 应支持子表存储结构按行组来设定存储顺序;
- c) 应支持子表存储结构按行组来自动适应存储顺序;
- d) 应支持子表存储结构按列组来设定存储顺序;
- e) 应支持子表存储结构按列组来自动适应存储顺序;
- f) 应支持用户自定义行组存储布局中列组之间的存储顺序;
- g) 应支持用户自定义列组存储布局中行组之间的存储顺序;
- h) 应支持行组存储布局中自适应的列组分裂和合并;
- i) 应支持各数据单元大小的自适应调整;

- j) 宜支持 PB 级数据查询响应为秒级,并且百列索引组合查询下的点查询响应最优为秒级;
- k) 宜支持 PB 级数据导入延时(耗时)达到秒级,单机数据吞吐量大于每秒 10 万行。

8.3 数据处理

8.3.1 数据表管理

数据表管理的技术要求如下:

- a) 应支持创建表、删除表、修改表、添加列、删除列等数据单元操作能力;
- b) 应支持创建索引、删除索引、修改索引等数据索引操作能力;
- c) 应支持按数据表数据的更新和批量更新;
- d) 应支持数据表单条记录的删除和批量删除;
- e) 应支持对双精度浮点型、长整型等不同数据类型进行类型转换和列式数据压缩;
- f) 宜支持增加或删除数据表中的行组或列组、修改行组或列组的命名和数据类型等数据表结构的修改能力。

8.3.2 数据导入

数据导入的技术要求如下:

- a) 应支持数据导入数据表时同步构建索引;
- b) 应支持数据表历史数据的批量导入;
- c) 应支持数据表增量数据的导入;
- d) 应支持多个数据表的数据同时导入;
- e) 应支持将文本格式的数据导入数据表;
- f) 应支持其他列式存储格式的数据转换导入数据表;
- g) 宜支持实时流式数据导入;
- h) 宜支持数据库增量同步到数据表,数据立即可见。

8.3.3 数据压缩

轻量级数据压缩,即压缩状态下的数据可进行访问、运算等操作,其技术要求如下:

- a) 应支持增量编码、运行长度编码、位压缩和字典编码等轻量级压缩算法;
- b) 宜支持自适应选择轻量级压缩算法;
- c) 宜支持对轻量级压缩数据的查询。

注:轻量级压缩数据的查询是对于经过轻量级压缩的数据,在解压之前,直接在压缩数据上执行查询。

重量级数据压缩,即数据解压后才能对数据进行访问、运算等,其技术要求如下:

- a) 应支持用户选择重量级数据压缩方式;
- b) 应支持对轻量级压缩后的数据使用重量级数据压缩方法;
- c) 宜支持重量级压缩算法;
- d) 宜支持插件式添加重量级压缩算法。

8.3.4 数据合并

应支持将多个较小的子表合并成较大的子表,并将多个子表索引构建成一个较大的子表索引。

8.3.5 数据分区

数据分区的技术要求如下:

- a) 应支持数据表按照时间、地理位置、范围、列表、取值等不同方式进行分区；
- b) 应支持不同批次导入的数据划分到不同的分区；
- c) 宜支持将经常组合出现的列组排列在一起存储。

8.3.6 数据更新

数据更新的技术要求如下：

- a) 应支持一个更新语句包含一条或多条数据更新记录；
- b) 应支持根据查询结果更新数据表；
- c) 应支持带子查询的更新或删除语句；
- d) 应支持数据表清空；
- e) 宜支持异步更新数据表。

9 数据检索技术要求

数据检索的技术要求如下：

- a) 应支持基于规则、成本等优化器对 SQL 进行优化；
- b) 应支持分布式计算检索,对于较大的检索通过分布式计算提升检索效率；
- c) 应支持数据的即时检索；

注 1：即时检索指的是检索响应达到秒级的检索。

- d) 应支持交互式 OLAP 检索、过滤检索、多维分析检索、实时检索等多种场景的数据检索；
- e) 应支持基于列组做算术运算的计算检索；

注 2：算术运算指的是聚合函数与常量的运算。

- f) 应支持对全部或部分检索结果按照一列或多列值的大小进行分组,值相等的为一组；
- g) 应支持并操作、交操作、差操作等集合检索；
- h) 应支持从检索结果中进行检索；
- i) 应支持分组、排序、计数、子查询、关联查询等标准 SQL 查询功能；
- j) 宜支持检索的预聚合；

注 3：预聚合指的是常用的聚合检索提前聚合。

- k) 宜支持多种因素组合的过滤检索；
- l) 宜支持单个数据表的多个列组以及多个数据表间的关联检索；
- m) 宜支持 JSON 等对象数据的查询与检索；
- n) 宜支持跨分区并行检索；
- o) 宜支持基于大数据的准交互式检索；
- p) 宜支持子表检索功能,对指定多个子表进行数据检索。

附录 A

(资料性)

面向分析的数据存储与检索应用示例

A.1 应用场景

根据本文件中规定的数据库、数据索引、数据存储和数据检索技术要求,以旅游行业大数据为例,进行相应实践应用介绍。旅游数据存储表结构如图 A.1 所示。

旅游大数据是指与旅游行业的从业者及消费者有关的数据,不仅涵盖了景区、酒店、旅行社、导游、游客等旅游相关行业所产生的管理或业务数据,还涉及互联网数据、旅游气象环保数据、交通数据、网络舆情数据、运营商数据、旅游交易数据等旅游行业所产生的基础资源信息。在对旅游大数据进行梳理、挖掘清洗、分析计算、共享交换、数据安全保护等前提下,还需要对旅游大数据进行多维融合存储和检索,围绕游客数据进行分析,包括对游客身份、行为习惯、兴趣爱好、旅游消费喜好、驻留时长、行为轨迹等大量数据进行处理分析,输出分析大数据报告,让游客数据价值更直观。

A.2 数据存储

旅游大数据的游客相关数据包含游客行为轨迹、驻留时长数据、运营商数据和旅游景区景点的 WiFi 数据。其中运营商由于其特殊的安全性要求,对外统一提供的是离线数据文件,数据内容也是经过处理后的粗粒度结果数据;而 WiFi 数据为对接边缘设备收集的实时细粒度明细数据。针对这两类不同状态和类型的数据需要做不同方式的数据存储,具体的处理方法如下。

a) 数据导入与管理

运营商数据包含景区客流量、景区逗留时长、景区省内外客源、景区客流乘坐交通工具、基站客流量统计等多种类别的数据。由于运营商数据的提供方式为 CSV 文件,需要使用大数据的数据导入工具,自动根据数据结构创建原始数据表,根据时间维度创建数据分区。WiFi 数据包含移动终端信息、WiFi 所处景点信息、WiFi 终端厂商信息、实时信号强度、游客接入和离开时间信息。WiFi 数据为实时回传的数据,具有数据量较大、数据回传及时、数据信息内容繁多等特点。在创建原始数据表时偏向使用结构简单的宽表,数据存储以时间为主时序,以 WiFi 终端或景点作为分区维度。

b) 数据合并、分区与压缩

通过运营商数据中的景区逗留时长数据、景区客流量数据以及移动终端接入 WiFi 信号数据可以分析出游客在整个旅游过程中的行动轨迹和逗留时长,运营商数据可以描述游客的宏观轨迹行为,WiFi 数据可以描述游客的微观轨迹行为。把宏观与微观的游客轨迹行为连接在一起需要对数据做融合处理,首先,根据游客游览行为建立组织与管理数据表,并分别构建运营商与 WiFi 信号的元数据信息子表。然后,根据景区景点信息建立维度区分表,形成以时间维度为连接点的 WiFi 信号数据与运营商数据融合的数据表,依据时间粒度的差异性,动态更新子表数据内容。最后,在子表中构建基于时间和景点地理信息的索引,并压缩子表数据内容。

c) 分析数据建模

根据数据应用需求建立数据业务模型。如景区景点游客驻留分析场景,可建立景区景点驻留区域数据模型,例如游客景点驻留时长分布数据模型,游客行动轨迹数据模型等。这些业务数据模型可为后续的数据分析检索提供逻辑计算支持。

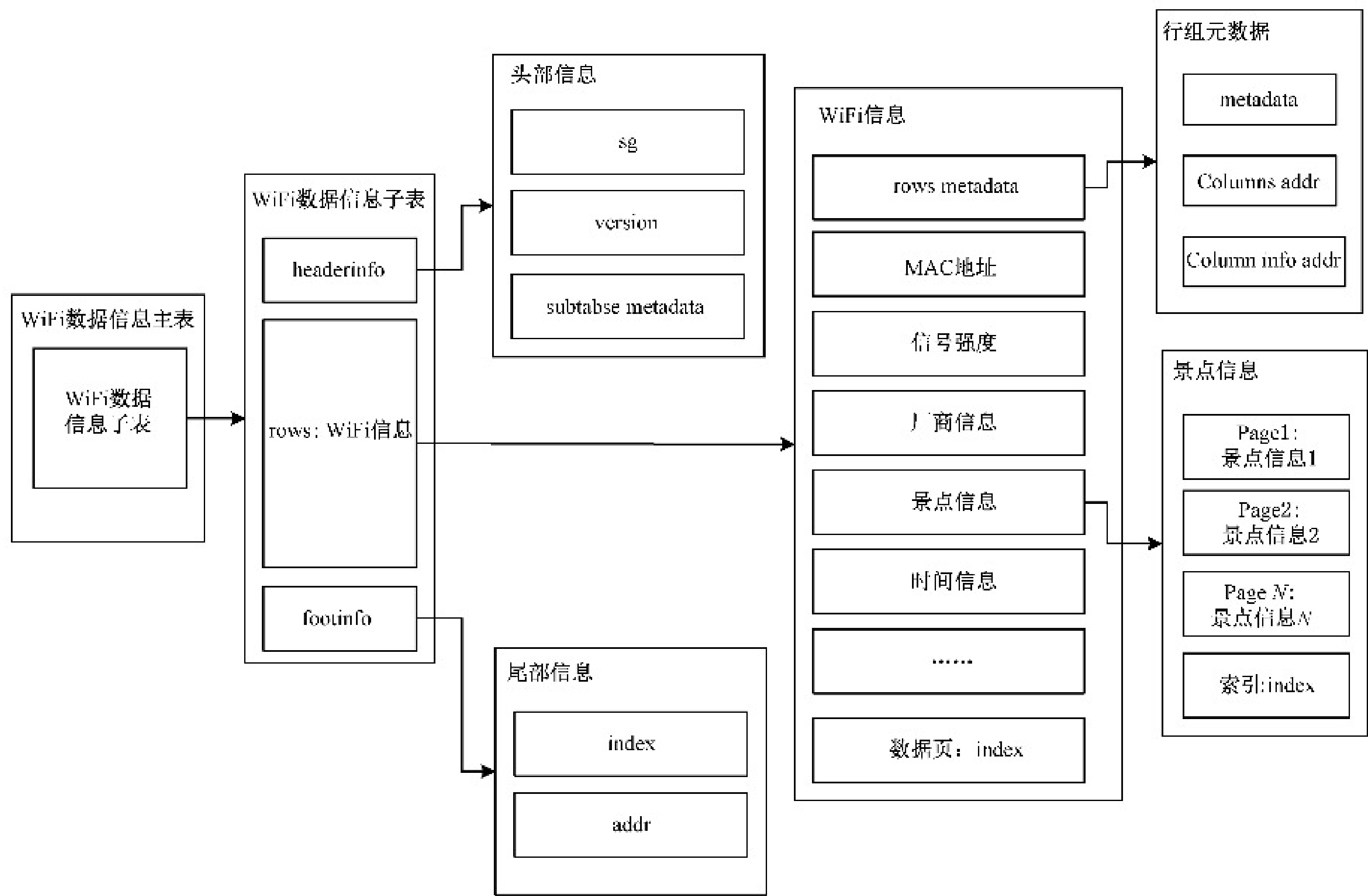


图 A.1 旅游数据存储表结构图

A.3 数据检索

将旅游大数据分类并创建数据表和子表，并将游客关联数据导入、合并、分区、压缩，结合数据索引，完成旅游大数据的游客数据存储。为了能快速检索合并压缩后的数据表中的数据，使用 OLAP 在线联机交互式检索数据模型，将高频的、高并发的、逻辑复杂度高的数据建立多维立体模型并联机预处理，实现数据快速检索提取。

在联机预处理数据时，根据数据模型的需求，基于列组做交、差、并运算和多列组数据过滤，创建数据的维度模型，加入时间、景点等维度索引，为上层的游客业务应用数据检索提供技术支持。

参 考 文 献

- [1] GB/T 18391.1—2009 信息技术 元数据注册系统(MDR) 第1部分:框架
-