

首都师范大学信息工程学院 2019—2020 学年第二学期

2019 级硕士研究生智能应用技术专业期末考试试卷

课程名称：数据库新技术

考试形式：撰写论文

考试时间：2020 年 5 月 20 日

姓 名：屈原斌

班 级：学硕班

学 号：2191002033

任课教师：关永

成 绩：

时空数据库的研究进展概述

屈原斌¹⁾

¹⁾(首都师范大学信息工程学院 北京 100048)

摘 要 随着大数据时代的来临，大量的时空数据的存储需求推动了时空数据库的快速发展。本文主要分析了目前时空数据库的发展现状和存在的问题，并介绍了几种典型的时空数据模型，最后将通过简单介绍阿里云时空数据库的发展来对时空数据库的将来发展方向给出了一些建议和展望。

关键词 大数据；时空数据；时空数据库；时空数据模型；阿里云

中图法分类号 TP391

Overview of the Research Advances of Spatio-Temporal Databases

QU Yuan-Bin¹⁾

¹⁾(School of Information Engineering, Capital Normal University, Beijing 100048)

Abstract With the advent of the era of big data, the storage requirements of a large amount of spatio-temporal data have promoted the rapid development of spatio-temporal databases. The spatio-temporal data model as the basis of spatio-temporal databases has also experienced a long period of development and evolution. This article mainly analyzes the current development status and existing problems of spatiotemporal databases, and introduces several typical spatiotemporal data models such as sequence snapshot model, ground state correction model, spatiotemporal composite model, spatiotemporal cube model, non-one-paradigm model and object relational model. Finally, I will give some suggestions and prospects for the future development direction of the spatio-temporal database by briefly introducing the development of the Ali Cloud spatio-temporal database.

本课题得到华为技术有限公司委托项目(YBN2018065021)、国家自然科学基金(61825204, 61932016, 61972039)、国家重点研发计划课题(2018YFB0803405)、北京高校卓越青年科学家计划项目(BJJWZYJH01201910003011)、北京市自然科学基金(4192050)、鹏城实验室大湾区未来网络试验与应用环境项目(LZC0019)资助。屈原斌，男，1998年生，硕士研究生，主要研究领域为自然语言处理。E-mail: ybqu@cnu.edu.cn.

Key words Big data; spatiotemporal data; spatiotemporal database; spatiotemporal data model; Ali Cloud

1 引言

随着互联网发展进入到大数据时代,无线技术与定位技术相结合,共同发展,使现在的应用程序能够处理移动的对象的位置数据,如车辆、无线设备的用户、海上运输等。同时其他的一些非时空对象应用程序也可以依靠无线定位技术来获取一些连续的、多维的变量信息。这类应用程序都建立在大量时空数据的收集的基础之上,针对于时空数据的存储,从而产生了时空数据库(Spatio-Temporal DataBases, STDB)。

STDB 在 20 世纪 80 年代末开始受到人们的重视,现已成为数据库领域中备受关注的前沿方向。STDB 是时态数据库(Temporal DataBases, TDB)与空间数据库(Spatial DataBases, SDB)的统一体,即包括时间与空间要素,主要用于存储与管理位置或形状随时间而变化的各类空间对象,在现在出现的大量处理动态对象的应用程序中越来越重要(比如,交通控制、气象监测、移动计算等)。

时空数据模型作为时空数据库的基础,是一种有效组织和管理时态数据、空间、专题、时间语义完整的数据模型,它不仅强调时空对象的空间和专题特征,而且强调这些特征随时间的变化,既时态特征。建立合理、完善、高效的时空数据模型是实现时空数据库的基础和关键。

2 时空数据库研究现状

STDB 的研究内容相当丰富,主要涉及时空对象表达,时空数据建模、索引、查询,STDB 体系结构、原型系统等,同时时空推理、时空查询代价模型等也为 STDB 的研究带来了一定的挑战。

由于 STDB 在 GIS、多媒体应用、导航、生态环境系统、智能交通系统、战场信息分析系统、战争仿真系统、战争游戏系统等方面具有广泛的应用前景和潜在的经济价值,从而激发了世界上大量科研工作者及相关商家的浓厚兴趣,尤其是欧洲国家已经开展了许多相关项目的研究。例如,由欧盟资助的 CHOROCHRONOS 项目,主要研究时空对象表达、时空模型与语言、时空信息图形用户接口、时空查询处理、时空数据存储与索引以及 STDB 体

系结构等。Oracle、DB2 以及 Informix 等都提供了可扩展的对象关系数据库技术以支持时空应用。当前国际上一些权威期刊和重要学术会议都将 STDB 研究作为主题内容之一,为该领域的广大研究人员提供了更多的交流机会。

STDB 在最近几年里取得了很大的进展,在国内,阿里云时空数据库的正式商业化,也极大的推动了 STDB 的发展。

3 几种典型的时空数据模型介绍

构建好的时空数据模型对时空数据库的发展是至关重要的,合理的时空数据模型包含如下几方面的因素:如何节省存储空间,加快存取速度,确立时空语义诠释和模型体系结构的层次。表 1 给出一些经典的时空数据模型的优缺点对比,同时该部分也将对序列快照模型、时空立方体模型、时空复合模型、基态修正模型、面向对象的时空数据模型等五种模型进行展开分析。

3.1 序列快照模型

快照模型有栅格快照模型和矢量快照模型,是通过将一系列时间片段的快照存储起来,为反映地理现象的时空演变过程,要使每个切片都分别对应着不同时刻的状态图层,按照需要对所选择的时间片段进行播放,有一些 GIS 就是用该方式来慢慢逼近时空特征。该模型的优越之处在于:一是目前的数据库一直处于合理有效的状态;二是能够直接在目前的地理信息系统软件中实现。然而,快照将没有发生变化的全部特征进行储存,会出现许多数据冗余,如若模型变化较为频繁,而且数据量较大时,系统效率就会迅速降低,结果必然难以处理时空对象间的时空关系。

3.2 时空立方体模型

1970 年 Hagerstrand 提出时空立方体模型,由两个空间维和一个时间维组成三维立方体,给定一个时间位置值,即可从三维立方体中获得相应的截面。时空立方体单元是以规则的空间网格为底,相应的时间段为高组成的方形柱体,是数据存储的基本单元。当时空立方体单元中存储的物理表或逻辑表超出数据存储的上限时,数据的管理效率将会大幅下降,同时在时空立方体表达方面也较难实现。

表 1 时空数据模型对比分析

时空数据模型	核心思想	优点	缺点
序列快照数据模型	通过栅格数据表示某时刻状态	当前数据库总是处于有效状态	数据冗余量大
时空复合数据模型	由静态属性表示动态时空变化	数据冗余量小	不能区分同一时间段内的对象
面向对象的时空数据模型	记录时刻状态对象及具有时态信息的空间、属性和关系	支持对象嵌套和边长记录	实现困难
面向事件数据模型	在事件上添加事件标记	查询方便	必须结合状态变化才能得到对象的最近状态
时空立方体模型	形象直观地运用时间维的几何特性表现空间表现空间实体是一个时空体概念	直观，易于接受	冗余量大，操作复杂
GSM 数据模型	地理空间图。移动对象之间的社交联系图和移动对象轨迹图	稳定性较高	位置数据时间间隔不规律，分割方法产生的多边形数量较多

3.3 时空复合模型

在与矢量基态修正模型相似的时空复合模型中，时空复合式把空间分割成具有相同时空过程的最大公共时空单元，较容易实现时空拓扑分析。该模型保留了沿时间的空间拓扑关系，是基态修正模型和序列快照模型的折衷模型，所有更新的特征都被加入到当前的数据集中，新的拓扑关系也会伴随生成。但随着时间的延续，时空数据库中的数据会逐渐增多，影响时空复合模型的查询效率。

3.4 基态修正模型

当状态变化相对缓慢，为了降低数据冗余量，存储时只记录变化部分，可选择基态修正模型，又称为底层叠加模型。每变化一次的快照可通过叠加每次变化内容获得，每个对象只需存储一次，记录很小的数据量。虽然解决了数据冗余问题，但也增加了时空操作的复杂性。该模型设使用最为频繁的状态为基态，每当有状态变换时，将前一状态相对于最新状态的变化部分存入历史库。存储入库的内容会被变更数据不断修正。该模型处理矢量数据的效率较低，但处理栅格数据较为合适。

3.5 面向对象的时空数据模型

在静态对象模型的基础上扩展时态信息表达，其核心思想是把所考察的系统看成一个对象，提取它们之间的共性和相互作用方式并进行规范化，使得逻辑过程的建模变得简单且容易实现。该模型能够较好地表达数据的时间、空间和语义属性信息，反映要素对象之间的多层次关联关系，有助于探索事物动态变化的内在规律。面向对象的数据库也将

是未来数据库的发展方向。

4 阿里云时空数据库介绍

2018 年 12 月 13 日，伴随阿里云 HBase 全新发布 X-Pack 全托管 NoSQL 数据库平台，HBase Ganos 时空数据库引擎正式上线。HBase Ganos 以阿里云飞天操作系统为强大底座，结合云 HBase 新一代 KV、时序、时空、图多模数据综合处理能力以及云上 Spark 大数据分析计算服务，为迎接在线时空全量大数据应用构筑 PaaS（Platform-as-a-Service）平台能力。2019 年 5 月 5 日，阿里云时空数据库正式开始免费公测，到 2019 年 9 月 10 日正式商业化售卖，极大的推动了 STDB 的发展。

阿里云时空数据库具备一下优势：

（1）易用：基于标准 PostgreSQL，支持 JDBC/ODBC 驱动访问。时空数据同其它业务数据一体化处理，兼容 OGC 空间计算函数；支持符合 OGC 规范的 WKT 和 WKB 格式数据输入和输出。

（2）写入性能强劲：单机不同规格下可以支持到数万到数十万的 TPS 写入。

（3）高效分析能力：阿里云时空支持查询条件自动选择分区，高效空间索引，并行的聚合操作等提升分析性能。

（4）自动扩展分区：支持自动扩展分区，减少用户的管理量，不需要人为的干预自动扩展分区。

（5）功能丰富：具有丰富的时间和空间处理查询函数；支持点（POINT）、线（LINESTRING）、

多边形 (POLYGON)、多点 (MULTIPOINT)、多线 (MULTILINESTRING)、多多边形 (MULTIPOLYGON) 和几何对象集 (GEOMETRYCOLLECTION) 等几何类型存储。

(6) 自动保留策略: 根据用户配置, 自动删除过旧数据, 极大降低用户使用成本, 减少用户管理工作。

(7) 自动 Failover: 提供全自动 Failover 机制, 一旦所在硬件发生不可恢复的故障, 会在非常短的时间内使用其他硬件替换故障硬件。

(8) 高可靠;

阿里云时空数据库可以广泛应用于交通监控与分析、物流配送、可穿戴设备监测、新能源车辆监测、LBS、地图服务等领域。

5 时空数据库的发展方向

从阿里云时空数据库的商业化, 我们可以清除的认识到 STDB 发展到现在已经在时空对象表达、时空数据建模、索引、查询, STDB 体系结构、原型系统等方面取得了一些令人可喜的成果, 但是仍有许多富有挑战性的问题有待广大研究者的进一步研究与解决。例如以下问题:

(1) 用规则与完全的语义来设计和实现时空数据模型与操作的问题;

(2) 时空数据索引与查询的优化问题;

(3) 用各种现有的体系结构 (如分层体系结构、集成体系结构等) 构建实验性的 STDBMSs 的问题。

另外, STDB 在移动与无线计算、数据仓库与挖掘以及语义 Web 等研究领域中也许多值得研究的问题。

6 结论

随着是时空应用的不断出现和大量的时空数据产生, STDB 已经成为了数据库中一个重要的研究领域。它在 GIS、多媒体应用、导航系统、生态环境系统、智能交通系统等方面的广泛应用前景引起了广大科研人员的浓厚兴趣。近来的许多研究成果更是证明了 STDB 在处理时空信息方面有着强大的能力, 并且许多研究者正逐步将其研究成果转化为实用的系统。STDB 在未来的研究中仍然有很大的发展空间和应用领域。

致谢 衷心感谢老师和同学们本学期的课程讲解, 希望老师可以对本文提出的宝贵意见和建议!

参考文献

- [1] Gao Yun-Jun, Chen Ling, Chen Gen-Cai, Zhang Xue-Jie. An overview of research progresses in spatio-temporal databases. Chinese Journal of Computers, 2004: 601-606.
(高云君, 陈岭, 陈根才, 张学杰. 时空数据库研究进展综述[C]. 第二十一届中国数据库学术会议论文集 (技术报告篇), 2004: 601-606.)
- [2] Wang Yang-Yang, Sun Wei. Research progress of spatiotemporal data model[J]. Geospatial information, 2016, 14(09): 29-31+6.
(王洋洋, 孙伟, 古丽米拉·克孜尔别克. 时空数据模型研究进展[J]. 地理空间信息, 2016, 14(09): 29-31+6.)
- [3] Liu Song. Alibaba: cloud database spatiotemporal engine empowers spatial data inclusive computing[J]. Chinese Surveying and Mapping, 2019(08): 36-37.
(刘松. 阿里巴巴: 云数据库时空引擎 赋能空间数据普惠计算[J]. 中国测绘, 2019(08): 36-37.)
- [4] Han Jian. Overview and research of spatio-temporal data model [J]. Science and Technology Communication, 2012(12): 54-55.
(韩剑. 时空数据模型概述及研究[J]. 科技传播, 2012(12): 54-55.)
- [5] Wang Tian-Ming, Li Ying, Liang Jian-Ping. Design and research of smart city spatio-temporal database [J]. Beijing Surveying and Mapping, 2017(06): 72-76.
(王天明, 李莹, 梁建平. 智慧城市时空数据库设计与研究[J]. 北京测绘, 2017(06): 72-76.)
- [6] Yang Fan. Research on update technology of geographic information spatio-temporal database [J]. Latitude and longitude, 2017(06): 44-49.
(杨帆. 地理信息时空数据库更新技术研究[J]. 经纬天地, 2017(06): 44-49.)
- [7] Chen Zhi-Bo, Lu Shou-Yi. Research progress of spatio-temporal data model in TGIS [J]. Research on Hebei Forest and Fruit, 2003(04): 395-400.
(陈志泊, 陆守一. TGIS中的时空数据模型的研究进展[J]. 河北林果研究, 2003(04): 395-400.)
- [8] Armstrong MP. Temporality in Spatial databases [J]. Proceeding: GIS LIS'88, 1988(2): 880-889.
- [9] Introduction to Big Data in Geography, <https://zhuanlan.zhihu.com/p/58545777> 2019, 3, 20.
(地理时空大数据概论 <https://zhuanlan.zhihu.com/p/58545777> 2019, 3, 20.)
- [10] Alibaba Cloud's spatio-temporal database engine HBase Ganos is online, with full analysis of scenarios, functions and advantages, <https://zhuanlan.zhihu.com/p/52969763> 2018, 12, 21.
(阿里云时空数据库引擎HBase Ganos上线, 场景、功能、优势全解析, <https://zhuanlan.zhihu.com/p/52969763> 2018, 12, 21.)

[11] Perfect unity of time and space! Alibaba Cloud time-space database officially commercialized, <https://zhuanlan.zhihu.com/p/83366519> 2019,9,20.

(时间和空间的完美统一！阿里云时空数据库正式商业化, <https://zhuanlan.zhihu.com/p/83366519> 2019,9,20.)

[12] Stefano Spaccapietra. Spatio-Temporal Data Models and Languages [J]. Geoinformatica, 2001(5).

Qu Yuan-Bin, Born in 1998, a master's degree candidate is currently studying, current research direction is natural language processing.

Background

The spatio-temporal database has made a huge contribution to the large-scale spatio-temporal data storage and all operations on these data in the era of big data. Currently, it is used in GIS, multimedia applications, navigation, ecological environment systems, intelligent transportation systems, battlefield information analysis systems, war simulation The system, war game system and other aspects have wide

application prospects and potential economic value. Some authoritative journals and important academic conferences at home and abroad take STDB research as one of the main contents, providing more researchers in this field with more Exchange opportunities. There will be more research directions in the future.