

封面：

分类号_____

密级_____

U D C _____

编号_____

厦 门 大 学
博 士 后 研 究 工 作 报 告

提供物联网传感器数据的私有云分析服务的研究

陈思恩

工作完成日期 2017年3月5日

报告提交日期 2017年3月15日

厦门大学

年 月

题名页

提供物联网传感器数据的私有云分析服务的研究

The research of the private cloud analysis service
for the Internet of things sensor data

博 士 后 姓 名 陈思恩

流动站（一级学科）名称 统计学

专 业（二级学科）名称 应用统计学

研究工作起始时间 2015 年 2 月

研究工作期满时间 2017 年 3 月

厦 门 大 学

厦门大学博士后研究工作报告

著作权使用声明

本人完全了解厦门大学有关保留、使用博士后研究工作报告的规定。厦门大学有权保留并向国家主管部门或其指定机构送交该报告的纸质版和电子版,有权将该报告用于非赢利目的的少量复制并允许该报告进入学校图书馆被查阅,有权将该报告的内容编入有关数据库进行检索,有权将博士后研究工作报告的标题和摘要汇编出版。保密的博士后研究工作报告在解密后适用本规定。

本研究报告属于: 1、保密 (), 2、不保密 ()

纸本在 年解密后适用本授权书;

电子版在 年解密后适用本授权书。

(请在以上相应括号内打“√”)

作者签名: 日期: 年 月 日

导师签名: 日期: 年 月 日

摘要

内 容 摘 要

结合企业业务层的经营目标和信息化建设与规划，研究提供物联网传感器数据的私有云分析服务，其中涉及技术架构、搭建技术平台、建立分析服务的实施方法和模型，帮助集团性企业实现自己的物联网数据分析方案。

私有云分析服务专注于处理传感器及其它类型机器产生的时间序列数据，剔除大量来自传感器及连接设备的时间戳数据，通过云数据库提供数据分析方案。

服务可以提供给通过传感器生成数据流的公司，其中包括物联网和工业互联网、可穿戴设备、医疗、交通、石油、天然气、金融服务及制造业、新兴的技术创业领域。研究成果也可为作为学术成果，作为集团性企业信息系统建设和项目实施的参考和指导。

关键词： 物联网、传感器、数据、私有云、分析服务

英文摘要

Abstract

Based on the enterprise business layer's business objectives and information construction and planning, the research provides networking sensor data analysis of private cloud services, which relates to the technical architecture, build a technology platform, the establishment of the implementation of the analysis method and model of service, help enterprise group to achieve things their own data analysis scheme.

Private cloud analysis services focus on processing the time series data generated by sensors and other types of machines, eliminating a lot of time stamp data from sensors and connected devices, providing data analysis programs through the cloud database

The service can be provided to generate data flow through the sensor of the company, including networking and Internet industry, wearable devices, medical, transportation, petroleum, natural gas, financial services and manufacturing, emerging technology entrepreneurship. Research results can also be as academic achievement, as a group of enterprise information system construction and project the implementation of reference and guidance.

Keywords: Internet of things, sensor, data, private cloud, analysis service

目 录

目 次

目录

1. 课题研究的来源和预期目标.....	1
1.1. 课题研究的来源.....	1
1.2. 预期目标.....	1
2. 课题研究的科学意义与学术价值.....	1
2.1. 研究的科学意义.....	1
2.2. 学术价值.....	2
3. 主要研究内容和研究工作.....	2
3.1. 主要研究工作.....	2
3.2. 预期主要研究成果.....	2
3.2.1. 预期主要研究成果（平台）.....	3
3.2.2. 预期主要研究成果（分析）.....	3
3.3. 主要研究工作计划.....	5
4. 技术难点、解决预案和可能创新点.....	5
4.1. 技术难点或关键问题.....	5
4.2. 可能的创新点.....	7
5. 研究进展情况.....	8
5.1. 技术平台特性.....	8
5.2. 服务体系支撑实例.....	10
5.2.1. 江苏移动 MDMC 监控系统.....	10
5.2.2. 变电站监控管理系统.....	11
5.2.3. 油耗监控管理系统.....	11
5.2.4. 国电热网监控系统.....	12
5.3. 技术平台原理附图.....	13
参考文献.....	15

致谢	17
博士生期间发表的学术论文、专著	18
博士后期间发表的学术论文、专著	19
个人简历	20
联系地址	24

厦门大学博硕士学位论文摘要库

提供物联网传感器数据的私有云分析服务的研究

1. 课题研究的来源和预期目标

1.1. 课题研究的来源

本课题结合厦门信达股份信息化建设的实际需求，在中国博士后基金会、厦门市人事局等相关部门审批通过“福建厦门信达股份有限公司博士后科研工作站招收博士后研究项目 ”后设立的。

1.2. 预期目标

结合企业业务层的经营目标和信息化建设与规划，研究提供物联网传感器数据的私有云分析服务，其中涉及技术架构、搭建技术平台、建立分析服务的实施方法和模型，帮助集团性企业实现自己的物联网数据分析方案。

私有云分析服务专注于处理传感器及其它类型机器产生的时间序列数据，剔除大量来自传感器及连接设备的时间戳数据，通过云数据库提供数据分析方案。服务可以提供给通过传感器生成数据流的公司，其中包括物联网和工业互联网、可穿戴设备、医疗、交通、石油、天然气、金融服务及制造业、新兴的技术创业领域。研究成果也可作为学术成果，作为集团性企业信息系统建设和项目实施的参考和指导。研究成果在博士后在站期间（一般为两年）完成，成果符合博士后研究的标准和要求。

2. 课题研究的科学意义与学术价值

2.1. 研究的科学意义

在物联网时代，面对 PB 级的数据，企业将难以以一己之力完成基础设施的建设。物联网所产生的大量数据不仅会驱动现在的数据中心发生根本性的变化，同时也会驱动相关企业采用新的大数据策略。由于缺乏相关技能以及持续增长的数据对基础设施采购的需求，企业将逐步放弃 DIY 模式，转而使用 PaaS 和托管的解决方案，借助于数据库即服务（例如 Amazon 的 Redshift、Hortonworks 和

Cloudera 的企业级 Hadoop)、托管的大数据服务 (例如 Treasure Data) 以及矩阵式的数据中心服务 (例如 GoGrid) 实现自己的物联网数据分析方案。

物联网的价值在于数据。企业对数据的分析工作启动地越快,挖掘出的业务价值就越多。而云服务提供商的目的就是通过加大相关的投入,消除数据收集、管理的风险以及复杂性,让客户能够专注于分析。

2.2. 学术价值

课题研究成果,可作为其他跨行业的集团性公司进行信息化建设时参考与借鉴,探讨信息化建设的新方法和新手段领域,同时为国内学术中“大数据”、“云计算”、“数据分析与挖掘”、提供实证案例和理论补充,具有重要的意义。

未来的场景时代是建立在移动设备、社交媒体、大数据、传感器、定位系统之上的整合式体验——它可以是一个产品,也可以是一种服务,还可以一种无时无刻不在的私人助理。未来场景时代的要素归结为五大原力(分别是移动设备、社交媒体、大数据、传感器和定位系统),这五种所谓原力并非单独存在,而是作为一个综合体,提供物联网传感器数据的私有云分析服务可以作为这个综合体的支撑平台。

3. 主要研究内容和研究工作

3.1. 主要研究工作

1. 结合目前信达物联科技公司的商业模式,研究时间序列数据存储、分析、监测、预测的分析即服务平台的理论方法。
2. 研究平台的数据可视化、分析工具、存储、主机托管和分析报告的并行大数据分析工具箱,允许物联网社区开发人员运用其他应用程序访问和集成平台的功能,其中 API 实现包括按时间检索及分析数据。
3. 建立相应的模型和方法使其能提供给通过传感器生成数据流的公司,其中包括物联网和工业互联网、可穿戴设备、交通等领域。

3.2. 预期主要研究成果

1. 研究报告 1 份《厦门信达物联网传感器数据的私有云分析服务的研究》

2. 发表学术论文 4 篇左右，其中国际刊物论文或 EI 检索论文占 1-2 篇。
3. 建立物联网和工业互联网、可穿戴设备、交通等领域相应的数据分析模型和方法。
4. 提供并行大数据分析工具箱原型。

3.2.1. 预期主要研究成果（平台）

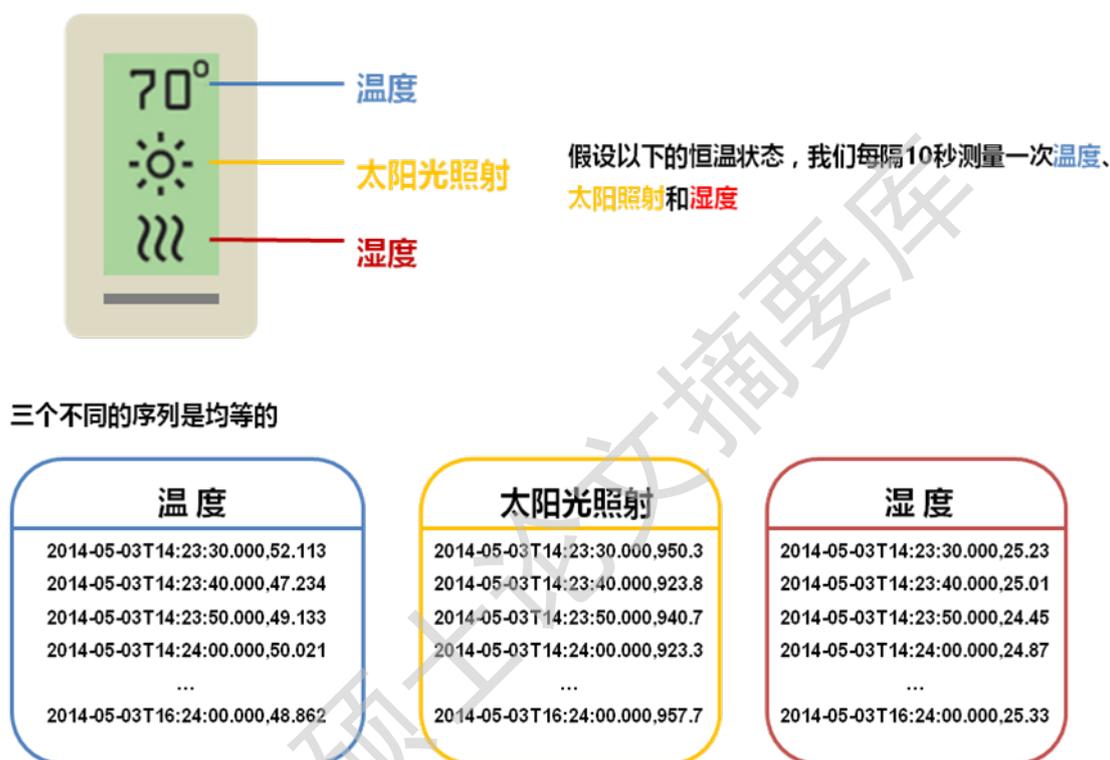


图 1：提供物联网传感器数据的私有云分析服务数据应用场景

3.2.2. 预期主要研究成果（分析）

以交通为例，以城市历史和实时数据为基础，以使用者个人偏好和政府限制规定为依据，向使用者反馈路线信息。为城市交通提供新的导航系统，有效平衡个人偏好和公众兴趣。使用者可以利用智能设备个性化的定制城市导航路线，如选择最短路径，避免犯罪区域、污染区域、交通拥堵区域，或者避免日程计划事件区域等。

当地政府也可以针对社区做出限制性规定，如减少学校地区的噪音污染，是否需要改变公共汽车开始时间，极端天气条件下带来的气象数据使运营商能够限制对乘客的影响。

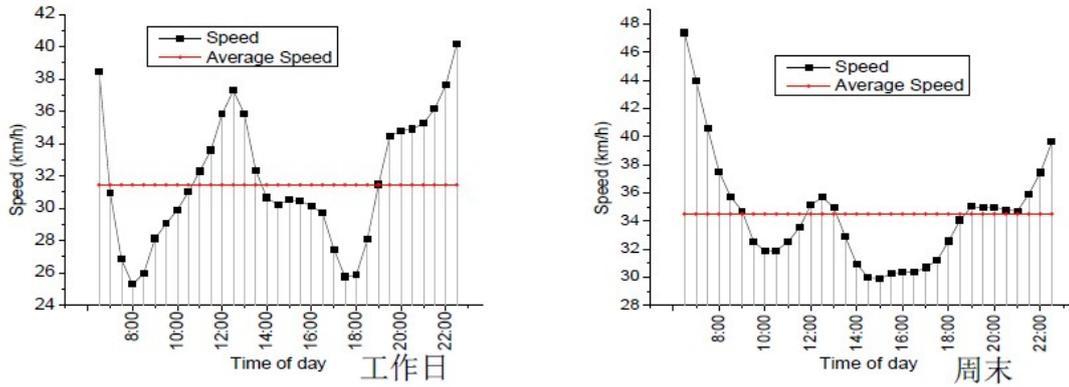


图 2：提供物联网传感器数据的私有云分析服务数据分析场景

以停车为例，数据整合（data aggregation），数据分析（data analytics）和停车信息分享。通过数据整合将繁琐的历史停车记录整合成有用的数据。大部分停车场都有各种各样的 IT 系统，例如停车收入控制系统，车牌识别系统，车位库存记录设备，车位传感器，停车自助缴费机，智能咪表等等。但是，这些停车场却没有很好的将从这些 IT 系统中获取的数据整合起来并加以利用。由平台将这些已有的繁琐的历史停车数据进行了整合并通过可视化处理展现给用户。

通过数据分析为用户的决策提供数据支持。基于平台提供的预测分析和定价方案则是建立在理性的数据分析之上的，因此更加准确有效。“以停车定价为例，有些新的停车场只是简单的根据周边停车场的经验做定价，这样的定价往往是不合理的。例如周中和周末的定价完全相同，但实际上两个时段的停车需求则是完全不同的。这就常常导致了停车场有空位而停车者由于收费过高而找不到合适停车位的现象。

平台数据分析包括对停车历史的统计分析，例如停车的历史趋势，季节性走势等；实时监控则主要针对停车场的运营做分析支持；预测分析则是对未来停车需求的预测分析。预测分析可以帮助停车场做停车的收费定价。

通过停车信息分享将所有通过数据分析得到的停车信息分享给广大用户和消费者。例如让消费者在去停车场之前就预先知道当时的车位可用性信息。有效避免了消费者开去停车场才发觉没有停车位的尴尬状态。信息分享则可以通过多种形式得以实现，例如手机 APP，网站查询等等。

3.3.主要研究工作计划

1. 根据下达的研究课题，逐步实现一个具体的数据策略解决方案。有针对性地调查和收集企业有关信息，调查研究对象的信息化现状，修订研究框架与研究方案，确定具体研究内容、关键和难点，制定详细的研究方案。
(2015.03-2015.09)
2. 根据现状和 IT 技术发展趋势与方向，研究和收集前沿资料，整个过程可以分为三个阶段：产生数据并通过 Internet 传递、中央系统收集并组织数据、持续的数据分析与使用。(2015.10-2016.3)
3. 第一阶段需要决定数据创建的标准以及如何通过网络进行传递。第二阶段则需要根据设备、网络以及功耗的限制决定是实时地收集数据还是在某个时间批量收集，同时还需要决定如何存储数据。(2016.3-2016.09)
4. 完成数据的收集与存储基础上进行分析。使用 Spark Streaming 满足近乎实时的要求，将这些技术与 Cassandra 这样的 NoSQL 数据库结合在一起使用，处理并分析大规模、快速移动的数据集。(2016.10-2017.01)
5. 完成出站研究报告和考核。(2017.02-2017.03)

4. 技术难点、解决预案和可能创新点

4.1.技术难点或关键问题

什么是时间序列数据？均匀分布在时间轴上，且能成功测量的定量观察搜集

例子：

对一个心脏每秒心率的持续监控

对一个酒店所有房间每小时气温的读取

同步向量每秒读取 60 次 7x24 小时

每小时 4 次表的读数 7x24 小时

卡车和拖车的制动压

每小时每平方米土地上的温度变化

城市所有高速上的车速

换句话说——M2M 数据多半是时间序列

时间序列分析目标描述：相关数据中识别规律—趋势、季节性或其他变量

时间序列分析目标解释：理解数据，构建数据模型

时间序列分析目标预测：根据先前模型预测短期趋势

时间序列分析目标干预分析：单一事件如何改变时间序列

时间序列分析目标质量管理：指定大小的偏差预示着某个问题

数据量大：分布式存储数据，潜在的大量数据才可以高性能的方式被管理和访问

多样化/有效性：时间序列数据具有时间上的顺序，这使得它明显有别于其他常见的数据分析问题，很因为其他的数据观察起来是没有自然顺序的

速度：能够在极度高速输入速度下接收和处理数据，需要建立在大型集群上的快速和通用数据处理架构

价值：提供有效机制，处理分析搜集来的数据，尽可能快的出结果

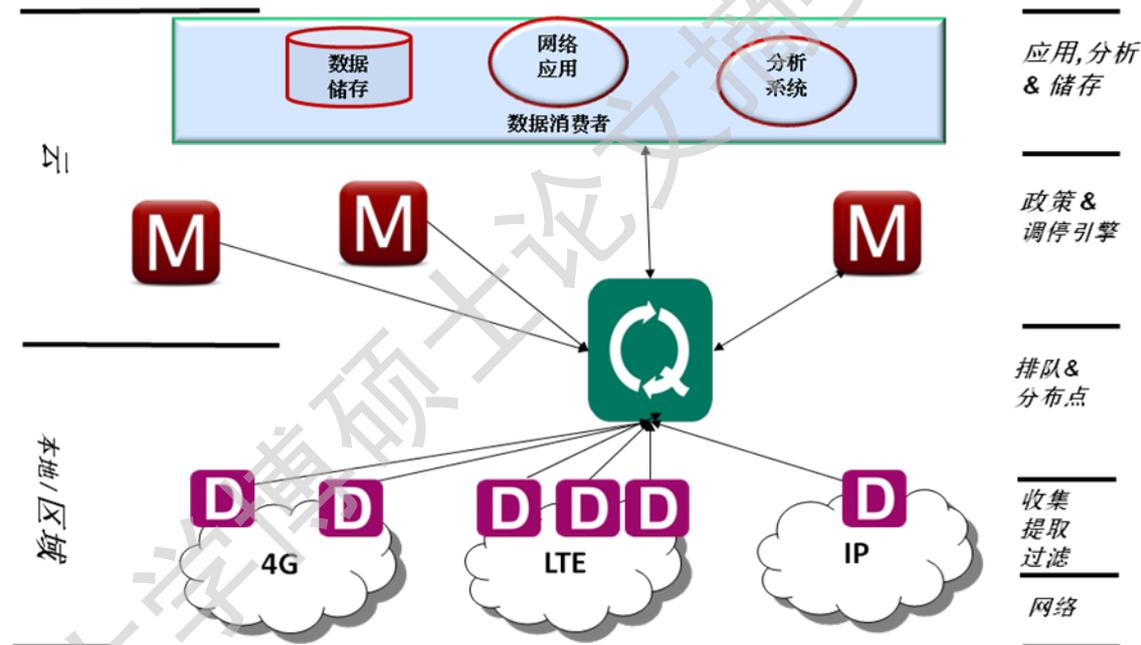


图 3：提供物联网传感器数据的私有云分析服务架构图（一）

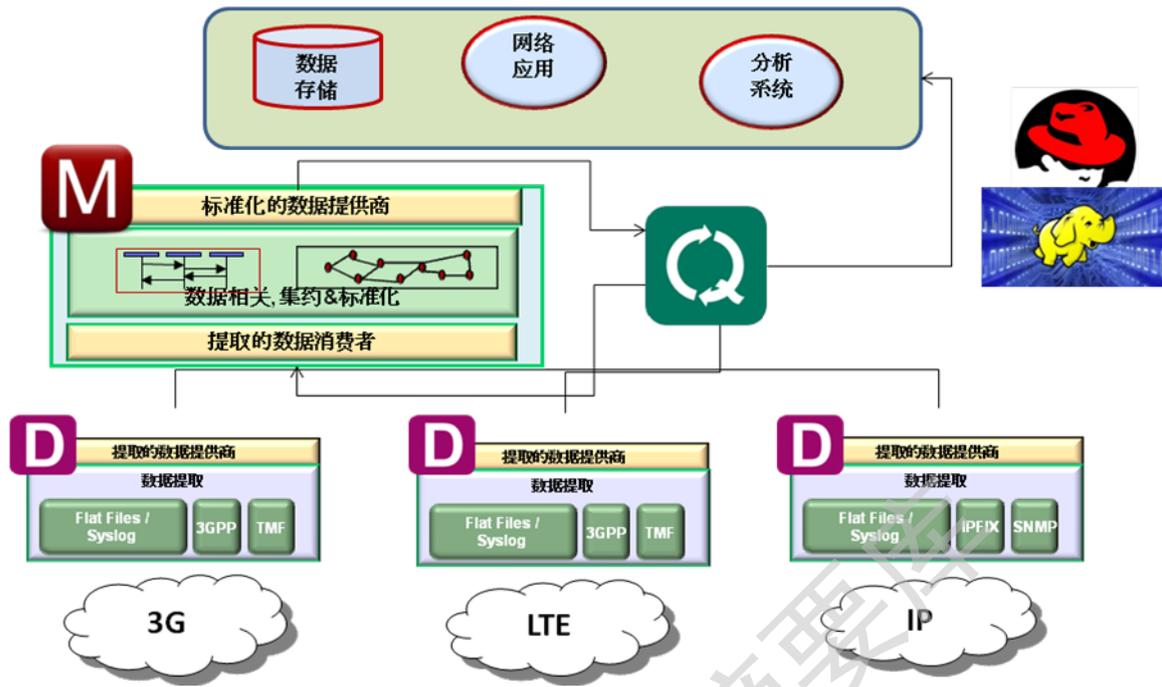


图 4：提供物联网传感器数据的私有云分析服务架构图（二）

4.2. 可能的创新点

1. 厦门信达物联网传感器数据的私有云分析服务理论方法。
2. 建立物联网和工业互联网、可穿戴设备、交通等领域相应的数据分析模型和方法。
3. 提供并行大数据分析工具箱原型。

不能使用现在的策略，因为可以被捕获、管理并利用的数据将更加多样化，同时用例也会更加丰富。附加到各种设备和对象上的传感器会产生各种类型的数据。这些数据将会用于各种响应式的、主动的或者创造性的目的。

提供物联网传感器数据的私有云分析服务的任务就是与业务用户一起工作，完全理解物联网方面的用例，然后寻找满足业务需求的技术。通过分析平台和工具，让业务用户能够获取到需要的数据，分析数据的含义并快速地做出响应。



图 5：并行大数据分析工具箱原型

5. 研究进展情况

5.1. 技术平台特性

建立相应的模型和方法使其能提供给通过传感器生成数据流的公司，其中包括物联网和工业互联网、可穿戴设备、交通等领域。整个过程可以分为三个阶段：产生数据并通过 Internet 传递、中央系统收集并组织数据、持续的数据分析与使用。结合企业业务层的经营目标和信息化建设与规划，研究提供物联网传感器数据的私有云分析服务，其中涉及技术架构、搭建技术平台、建立分析服务的实施方法和模型，帮助集团性企业实现自己的物联网数据分析方案。服务体系积累了一套稳定的物联网解决方案，可以高效的接入各类通讯协议设备，如 SMS, TCP, UDP, SOAP, RS-485, RS-232 等协议。

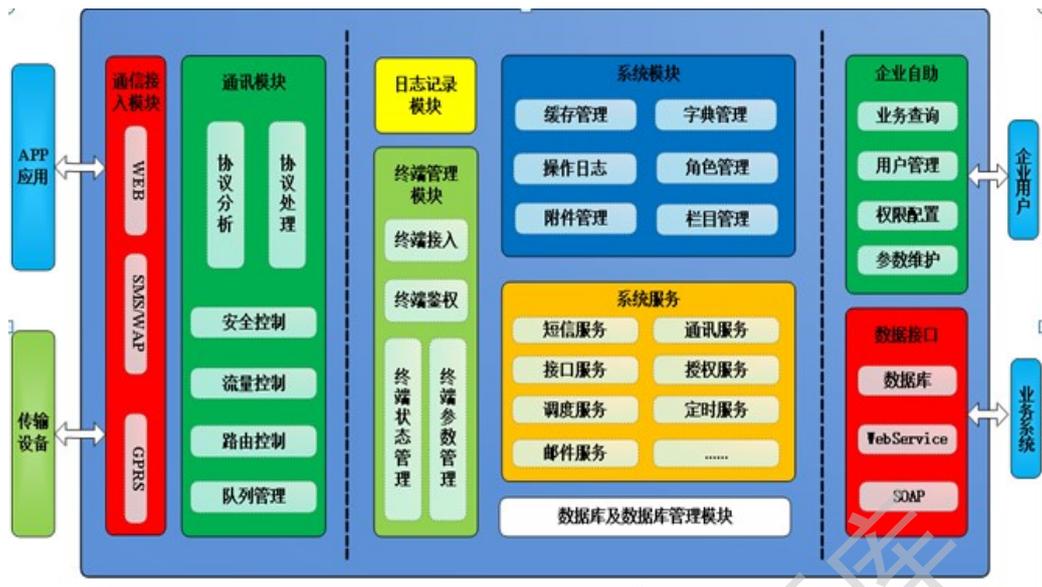


图6：提供物联网传感器数据的私有云分析服务架构图（三）

1. 整合和优化 Hadoop 生态圈技术，提供数据集成、数据存储、数据运算、数据调度、数据管理、数据访问、安全管理、集群管理等一站式的大数据企业级解决方案
2. 提供搭载数据级事务处理与数据级分析处理的数据应用架构，为海量数据提供业务应用支持和分析效果反馈等能力
3. 优化整合行业不同数据量级下数据标准的快速处理能力，特别针对行业 PB/TB/GB/MB 等不同级别数据的处理和整合能力，现在已经在民航、安全领域、政府、教育等行业具备优化整合能力
4. 自助研发的各项引擎，为数据的应用可供高度的支持。全文检索引擎，可理解全网的数据进行快速的检索，支持中文分词、同义词、快速匹配等检索方式；知识图谱引擎，以图论数据库为基础，结合机器学习算法和知识推导能力，在搜索的相似性推导、关系的挖掘、异常点检测、群体发现等领域有深入的应用；并发算法执行引擎，以 spark 执行内核为基础，为模型的并发执行、AB TEST 等提供了模型训练、模型评估、模型持久化、模型预测、模型访问等能力
5. 商业智能能力。平台具备完整的数据处理流程能力+商业智能展示能力，同时与 tableau、qlikview、pentaho 等商业智能工具深度集成。
6. 机器智能能力。平台自主研发和集成了大量的分布式算法库，并且自助研发了面向行业的模型库，同时具备画像分析的模型库和推荐引擎的推荐库。整

Degree papers are in the "[Xiamen University Electronic Theses and Dissertations Database](#)". Full texts are available in the following ways:

1. If your library is a CALIS member libraries, please log on <http://etd.calis.edu.cn/> and submit requests online, or consult the interlibrary loan department in your library.
2. For users of non-CALIS member libraries, please mail to etd@xmu.edu.cn for delivery details.

厦门大学博硕士学位论文摘要库