

构建深度智慧物联网

朱清波 高 瞻 宋庭新 编著



科学出版社

构建深度智慧物联网

朱清波 高 瞻 宋庭新 编著



科学出版社

北京

内 容 简 介

本书较为全面地介绍物联网相关的基本概念、智慧体系结构,并对智能传感器、无线传感网络、嵌入式终端平台等物联网基础硬件进行比较系统的介绍。同时还对智慧物联网所依靠的理论和实现软件部分,如复杂网络、云计算平台、智慧数据爬虫、智慧数据挖掘和智慧语义搜索等进行简要介绍。另外,本书通过实用系统的设计过程,使读者比较全面地掌握物联网的应用和设计方法。

本书可作为高等院校计算机、系统工程、测控和机电等专业高年级本科生和研究生的教材,也可供相关领域的技术人员参考。

图书在版编目(CIP)数据

构建深度智慧物联网/朱清波,高瞻,宋庭新编著. —北京:科学出版社, 2015.5

ISBN 978-7-03-044274-1

I. ①构… II. ①朱…②高…③宋… III. ①互联网络—应用②智能技术—应用 IV. ①TP393.4②TP18

中国版本图书馆 CIP 数据核字(2015)第 097349 号

责任编辑:王 哲 纪四稳 / 责任校对:桂伟利

责任印制:张 倩 / 封面设计:迷底书装

科 学 出 版 社 出 版

北京东黄城根北街 16 号

邮政编码:100717

<http://www.sciencep.com>

三河市骏立印刷有限公司 印刷

科学出版社发行 各地新华书店经销

*

2015 年 5 月第 一 版 开本:720×1 000 1/16

2015 年 5 月第一次印刷 印张:15

字数:289 000

定价:76.00 元

(如有印装质量问题,我社负责调换)

作者简介

朱清波（1982 年～ ）男，湖北工业大学机械工程学院副教授。2009 年毕业于湖北工业大学，获工学硕士学位。曾于 2008～2009 年在新加坡管理大学信息系统学院（School of Information System, Singapore Management University）从事信息系统研究工作，目前主要从事物联网与物流相关技术研究，主持和参与政府纵向、企业横向课题 8 项，发表论文 8 篇，获得 2 项软件著作权和 1 项专利。

高瞻（1979 年～ ）男，湖北交通职业技术学院讲师，软件开发高级工程师。2001 年毕业于华中师范大学，获工学硕士学位，目前主要从事信息化相关技术研究，共主持省部级课题 2 项，主持厅级课题 5 项，参与政府纵向、企业横向课题 8 项，发表论文 11 篇，出版学术专著 1 部。

宋庭新（1972 年～ ）男，湖北工业大学机械工程学院教授，华中科技大学博士，清华大学博士后，美国佐治亚理工学院访问学者。兼任湖北省机械工程学会工业工程专业委员会副理事长，湖北省人工智能学会常务理事，湖北省运筹学会理事。目前主要从事物联网与物流工程、云计算与大数据、企业信息系统集成等相关技术研究，共主持国家和省部级科研项目 12 项，发表论文 38 篇，出版学术专著 1 部，教材 1 部，获得 4 项发明专利和 9 项软件著作权，获得中国电子学会科技进步奖二等奖 1 项，湖北省高等学校优秀教学成果奖二等奖 1 项。

前言

物联网是新一代信息技术的重要组成部分，这里包含两层意思：其一，物联网的核心和基础仍然是互联网，它是在互联网基础上延伸和扩展的网络；其二，其用户端延伸和扩展到了任何物品与物品之间，进行信息交换和通信。物联网就是“物物相连的互联网”。物联网通过智能感知、识别技术与普适计算，广泛应用于网络的融合中，也因此被称为继计算机、互联网之后世界信息产业发展的第三次浪潮。物联网是互联网的应用拓展，与其说物联网是网络，不如说物联网是业务和应用。因此，应用创新是物联网发展的核心，以用户体验为核心的创新 2.0 是物联网发展的灵魂。

每当清晨，窗帘会适时自动打开，将第一缕阳光洒满你房间；当你将手掌贴在衣柜门上时，衣柜就会检测到你的体温，并综合气温为你搭配着装；家里的冰箱会对其中的食物进行监控，提示哪些食物即将过期，如果某些食物短缺，也会自动地联系超市帮助订购。这些近似科幻的生活场景正是物联网时代带给人们的便利生活。

物联网通过各式各样的传感器获取信息，随之而来的就是如何处理急速膨胀的数据？又如何让这些数据产生智慧？这无疑对数据挖掘技术提出了新的挑战。在这里数据挖掘只是一个代表性概念，它是一些能够实现物联网“智能化”“智慧化”的分析技术和应用的统称。细分起来，包括数据挖掘和数据仓库、云计算、人工智能、专家系统、智慧搜索等技术和应用。在通常情况下，数据分析的步骤是：首先收集数据，将其存储在数据库中，再用模型对数据库进行搜索，这种高度结构化的方式通常需要耗费很多时间；其次要对收集的数据进行融合，该过程对多维数据或信息进行处理，组合出更有效、更符合用户需求的数据；最后用先进的软件算法在数据流入的同时对其进行分析，使程序在短时间内对环境变化做出智能的回应。

如何完成从收集融合后的数据到信息、到知识、再到智慧的转换呢？信息技术的发展成就已远远出乎人们的想象，只要给它提供一个平台，它就可以像精灵一样创造出无限的智慧。完成这一过程有很多方式，可以用机器学习的办法让机器本身具有一定的智能，也可以采用专家系统，建立一个内部含有大量的某个领域专家水平的知识与经验的数据库，使系统能够利用人类专家的知识和解决问题的方法来处理该领域的问题。从学习的方法来看有两类任务：有监督的学习和无监督的学习。前者利用已知的类别标志，对一组训练样本进行学习，发现样本特征与类别之间的关系，用于预测新样本的类别，这个过程称为分类。后者在样本类别未知的情况下，根据样本之间的相似性进行分组，这个过程称为聚类。在物联网数据挖掘中，同样存在这两种学习，由于所处理的样本数据蕴含着路径信息，称为基于路径的分类和聚类。基于路径的分类根据物品在一段时间内的移动路径，判断物品所属的类别。例如，在物流监控中，将货

物的运输轨迹分为正常和异常两种,发现这两类轨迹的特征可以用来监测可疑行为。对路径的分类通常需要先将其划分为子序列,然后利用决策树、神经网络、贝叶斯等方法进行分类。基于路径的聚类发现相似的路径群,揭示路径之间的关系以及路径与其他特征之间的关系。聚类的结果可以用来优化道路设计,减少交通阻塞,预测交通流量。路径聚类一般包括数据表示、相似度量定义、聚类、聚类描述和聚类评价几个步骤。通过以上步骤可以使得物联网具有一定的智能,能够自行地处理一些问题,使人们的生活更加便捷舒适。

物联网涉及多学科融合,若要设计一个完美的物联网项目,设计人员必须具备坚实的软硬件工程经验。本书力图从不同学科专业知识出发来理解物联网的搭建和设计过程,为此,分别对智能传感器、无线传感网络、嵌入式终端平台、复杂网络、云计算平台、智慧数据爬虫、智慧数据挖掘和智慧语义搜索进行介绍,既介绍本学科知识又讲解该学科与物联网的联系与应用,强调物联网工程实用性是本书的目的所在。

本书的研究内容得到了多个项目基金的资助。本书的问世是集体智慧的结晶,感谢赵大兴、刘幺和、宋庭新等为本书内容的写作付出的辛勤劳动,感谢湖北工业大学对本书出版的大力支持。最后,要向我温柔体贴的妻子吴祯祯在生活上、事业上对我的关心、支持和理解表示深深的谢意。

尽管全书内容是经过作者多次修改才定稿的,但由于时间仓促,难免存在一些疏漏或不足,希望广大读者批评指正。

作 者

2014年11月

目 录

前言

第 1 章 物联网概述	1
1.1 引言	1
1.2 物联网的特点	2
1.3 物联网的全面感知	3
1.4 物联网的智慧运算	3
1.5 物联网的智慧分析	5
第 2 章 智能传感器	9
2.1 引言	9
2.2 传感器的基本原理	9
2.2.1 传感器的定义	9
2.2.2 传感器的结构	10
2.2.3 传感器的应用模式	11
2.3 智能传感器的基本原理	14
2.3.1 智能传感器的定义	14
2.3.2 智能传感器的产生	15
2.3.3 智能传感器的功能	15
2.3.4 智能传感器的特点	16
2.4 智能传感器的系统结构	17
2.5 智能传感器的实现	19
2.5.1 集成化	20
2.5.2 软件化	23
2.6 智能传感器数据校正技术	23
2.6.1 非线性自校正技术	23
2.6.2 软件抗干扰技术	24
2.6.3 自补偿技术	27
2.6.4 自检技术	33
2.7 多传感器信息融合	34
2.7.1 多传感器信息类型及其融合方法	34
2.7.2 多传感器信息融合过程	36

2.7.3	多传感器信息融合结构	37
2.7.4	多传感器信息融合方法	39
2.7.5	多传感器融合实例	45
2.7.6	网络化	46
2.8	应用实例	49
第3章	复杂网络	52
3.1	引言	52
3.2	网络的概念	53
3.3	复杂网络的基本特征量	54
3.3.1	平均路径长度	54
3.3.2	簇系数	54
3.3.3	度分布	54
3.3.4	介数	55
3.4	复杂网络的基本模型	55
3.4.1	规则网络	55
3.4.2	ER 随机网络	56
3.4.3	小世界网络	58
3.4.4	无标度网络	61
3.5	物联网信息传播机制	62
3.5.1	研究模型的拓扑与建模	63
3.5.2	信息传播模型搭建与分析	65
3.5.3	信息传播动力学分析	66
第4章	无线传感器网络	69
4.1	引言	69
4.2	无线传感器网络结构与特点	71
4.2.1	网络拓扑结构	73
4.2.2	传感器节点结构	74
4.2.3	无线传感器网络协议栈	74
4.3	无线传感器网络协议	76
4.3.1	洪泛式路由协议	76
4.3.2	以数据为中心的路由协议	77
4.3.3	层次路由协议	79
4.3.4	基于位置信息的路由协议	81
4.3.5	路由协议的比较	81
4.4	无线传感器网络与互联网融合	81

4.5	应用实例	83
4.5.1	工程背景描述	83
4.5.2	项目解决的关键问题	85
4.5.3	项目的研究方法与技术路线	86
第 5 章	嵌入式终端平台	91
5.1	引言	91
5.2	嵌入式系统简介	91
5.2.1	嵌入式系统	91
5.2.2	嵌入式系统的分类	94
5.2.3	嵌入式系统发展趋势	95
5.3	嵌入式系统硬件	98
5.3.1	微处理器	98
5.3.2	存储器	99
5.3.3	常用总线与接口	101
5.4	嵌入式系统软件	104
5.4.1	嵌入式实时操作系统	104
5.4.2	RTOS 体系结构	105
5.4.3	RTOS 的衡量标准	106
5.4.4	RTOS 的分类方法	106
5.4.5	RTOS 的关键技术	107
5.5	主流嵌入式系统	108
5.6	VxWorks 实时操作系统	109
5.6.1	VxWorks 的体系结构	109
5.6.2	VxWorks 内核分析	110
5.7	Linux 实时操作系统	111
5.8	Windows CE 实时操作系统	112
第 6 章	云计算平台	117
6.1	引言	117
6.2	云计算的概念	118
6.3	云计算发展现状	120
6.4	云计算实现机制	121
6.5	网格计算与云计算	123
6.6	MongoDB	126
6.6.1	NoSQL 简介	126
6.6.2	MongoDB 简介	128

6.6.3	MongoDB 安装与配置	130
6.6.4	体系结构	134
6.6.5	数据库操作	138
6.6.6	管理端 GUI 工具	141
第 7 章	智慧数据爬虫	142
7.1	网络数据爬虫简介	142
7.1.1	为什么要在 Web 上爬行	143
7.1.2	爬行过程	143
7.1.3	智能爬行与聚焦爬虫	149
7.1.4	网页搜索策略	150
7.1.5	网页分析算法	150
7.2	构建智慧爬虫	152
7.2.1	核心算法	152
7.2.2	系统需求分析	153
7.2.3	网络爬虫体系结构	153
7.2.4	系统构造分析	154
7.2.5	爬行策略分析	155
7.2.6	URL 抽取、解析与保存	155
7.3	开源爬虫	157
7.3.1	Lucene	157
7.3.2	Nutch	159
第 8 章	智慧数据挖掘	162
8.1	数据挖掘	162
8.1.1	概论	162
8.1.2	数据挖掘研究的内容与本质	163
8.2	数据挖掘流程	165
8.3	数据挖掘的典型方法	167
8.3.1	神经网络	169
8.3.2	遗传算法	170
8.3.3	决策树方法	171
8.3.4	粗集方法	172
8.3.5	覆盖正例排斥反例方法	173
8.3.6	统计分析方法	173
8.3.7	模糊集方法	173
8.4	数据挖掘软件	173

8.4.1	IBM Intelligent Miner	174
8.4.2	SAS Enterprise Miner	175
8.4.3	SPSS Clementine	176
8.5	大数据及相关技术	180
8.5.1	大数据发展历程	180
8.5.2	大数据的代表技术	182
8.5.3	Hadoop 体系结构	184
第 9 章	智慧语义搜索	188
9.1	智慧搜索引擎工作原理	189
9.2	本体语义	191
9.2.1	语义 Web 的概念	191
9.2.2	语义 Web 的体系结构	192
9.2.3	Web 服务与语义 Web 服务	193
9.2.4	用 OWL 构建领域本体过程	195
9.2.5	本体构造工具 Protégé 介绍	196
9.2.6	本体评估标准	197
9.3	语义 Web 服务的智能机器人控制	198
9.3.1	智能机器人控制领域的 Web 服务体系结构	198
9.3.2	模型设计和开发的基本原则	199
9.3.3	基于机械手本体的语义 Web 服务模型	199
9.3.4	语义 Web 服务的参数匹配分析	200
9.3.5	服务在参数上的匹配	201
9.3.6	参数的语义可推导	203
9.3.7	考虑参数语义的匹配算法	204
9.4	语义 Web 服务组合的控制流程实现	207
9.4.1	总体技术路线	207
9.4.2	实现 Web Service	208
9.4.3	组合服务的条件	211
9.4.4	组合服务的匹配	213
9.4.5	语义推理	218
9.4.6	组合服务的执行	222
9.4.7	实验结果与评价	222
参考文献		225

第 1 章 物联网概述

1.1 引言

Internet 的发明已经解决了人与人之间的交流，但是对于人与机器、机器与机器之间的交互仍在探索中。早在物联网概念产生之前，人们已经在安全监测、自动抄表、维修服务、自动售货机、公共交通系统、车队管理、工业流程自动化、电动机械、城市信息化等环境中有了广泛的应用和解决方案实例。其实，M2M (Machine to Machine) 模型是物联网的一个雏形。物联网 (Internet of Things, IOT) 又名传感网，是指通过射频识别 (Radio Frequency Identification, RFID)、红外感应器、语音识别装置、条码识别、全球定位系统、激光扫描器等信息传感设备，按约定的协议，把任意物品与互联网连接起来，进行信息交换和通信，以实现智能化识别、定位、跟踪、监控和管理的一种网络。简而言之，物联网是通过在物品上嵌入电子标签、条形码等能够存储物体信息的标志，通过互联网将其信息发送到后台信息处理系统。各大信息系统可以互联形成一个庞大的网络，从而可达到对物品实施跟踪、监控等智能化管理的目的。在物联网概念产生之前，人们在自动化领域就提出了 M2M 通信领域的控制模型，如图 1.1 所示。M2M 表达的是多种不同类型通信技术的有机结合：机器之间通信、机器控制通信、人机交互通信、移动互联通信。M2M 技术综合了数据采集、全球定位系统、远程监控、通信、计算机、网络、设备、传感器等系统，使业务流程自动化，并创造增值，这种网络的特点是既有测控传感器的功能，又有分布式 Web Service 平台上数据查询功能以及各种在云计算平台上交互式应用服务。在这里，物联网的关键技术不仅是对物体实现操控，它通过技术手段的扩张，实现了人与物、物与物之间的相融与沟通。物联网并不是互联网简单的翻版，也不是互联网的接口，而是互联网的一种延伸。作为互联网的扩展，物联网具有互联网的特性，不仅能够实现由人找物，而且能够实现以物找人。

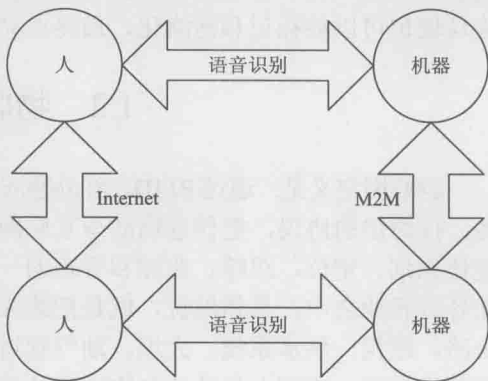


图 1.1 M2M 通信领域的控制模型

物联网的概念是在 1999 年提出的，顾名思义，物联网就是“物物相连的互联网”。它包含两层意思：第一，物联网的核心和基础仍然是互联网，其是在互联网基础上的

延伸和扩展的网络；第二，其用户端延伸和扩展到了任何物品与物品之间，进行信息交换和通信。

2005 年，国际电信联盟（International Telecommunication Union, ITU）发布了一份题为 *The Internet of Things* 的年度报告，正式将物联网称为 “The Internet of Things”，对物联网概念进行了扩展，提出了任何时间、任何地点、任何物体之间的互联（Any Time, Any Place, Any Things Connection），无所不在的网络（Ubiquitous Network）和无所不在的计算（Ubiquitous Computing）的发展蓝图。

物联网主要包括三个层次。首先是传感网络，也就是目前所说的包括 RFID、条形码、传感器等设备在内的传感网，主要用于信息的识别和采集；其次是信息传输网络，主要用于远距离无缝传输传感网所采集的巨量数据信息；最后是信息应用网络，该网络主要通过数据处理和解决方案来提供人们所需要的信息服务。

RFID 技术是物联网中非常重要的技术。以简单 RFID 系统为基础，结合已有的网络技术、数据库技术、中间件技术等，构筑一个由大量联网的读取器和无数移动的标签组成的，比 Internet 更为庞大的物联网成为 RFID 技术发展的趋势。物联网用途广泛，遍及智能交通、环境保护、政府工作、公共安全、智能家居、智能消防、工业监测、老人护理、个人健康等多个领域。物联网是继计算机、互联网与移动通信网之后的又一次信息产业浪潮。

物联网是技术变革的产物，它代表了计算技术和通信技术的未来，它的发展依靠某些领域的技术革新，包括 RFID 技术、无线传感技术和纳米技术。RFID 技术通过射频信号自动识别目标对象并获取物体的特征数据，将日常生活中的物体连接到同一个网络和数据库中。无线传感技术用于检测事物物理特征的变化并适时收集变化数据，并对之进行处理（变换）和识别。纳米技术的发展则使越来越小的物体实现连接和交流。以上主要技术的发展将创建连接所有物体的物联网。随着集成化信息处理技术的发展，工业产品和日常物体都将表现出智能化的特征，可以被远程识别或检测。甚至连垃圾也可以被标记和网络化，最终由完全静态的物体转变为新型动态的物体。

1.2 物联网的特点

物联网定义是：通过 RFID、红外感应器、全球定位系统、激光扫描器等信息传感设备，按约定的协议，把任意物品与互联网连接起来，进行信息交换和通信，以实现智能化识别、定位、跟踪、监控和管理的一种网络。物联网把新一代信息技术充分运用在各行各业之中，具体地说，就是把感应器嵌入和装备到电网、铁路、桥梁、隧道、公路、建筑、供水系统、大坝、油气管道等各种物体中，然后将物联网与现有的互联网整合起来，实现人类社会与物理系统的整合，在这个整合的网络中，存在能力强大的中心计算机群，能够对整合网络内的人员、机器、设备和基础设施实时地管理和控制。在此基础上，人类可以以更加精细和动态的方式管理生产与生活，达到 “智慧” 状态，提高资源利用率和生产力水平，改善人与自然的关系。

1.3 物联网的全面感知

传感器是机器感知物质世界的“感觉器官”，可以感知热、力、光、电、声、位移等信号，为网络系统的处理、传输、分析和反馈提供最原始的信息。随着科学技术的不断发展，传统的传感器正逐步实现微型化、智能化、信息化、网络化，正经历着从传统传感器（Dumb Sensor）、智能传感器（Smart Sensor）到嵌入式 Web 传感器（Embedded Web Sensor）的内涵不断丰富的发展过程。无线传感器网络（Wireless Sensor Network, WSN）是集分布式信息采集、信息传输和信息处理技术于一体的网络信息系统，以其低成本、微型化、低功耗和灵活的组网方式、铺设方式以及适合移动目标等特点受到广泛重视，它是关系国民经济发展和国家安全的重要技术。物联网正是通过遍布在各个角落和物体上的形形色色的传感器以及由它们组成的无线传感器网络，来感知整个物质世界的。传感器网络节点的基本组成包括如下几个基本单元：传感单元（由传感器和模数转换功能模块组成）、处理单元（包括 CPU、存储器、嵌入式操作系统等）、通信单元（由无线通信模块组成）和电源。此外，可以选择的其他功能单元包括定位系统、移动系统以及电源自供电系统等。在传感器网络中，节点可以通过飞机布撒或人工布置等方式，大量部署在被感知对象内部或者附近。这些节点通过自组织方式构成无线网络，以协作的方式实时感知、采集和处理网络覆盖区域中的信息，并通过多跳网络将数据经由汇聚节点（接收发送器）链路把整个区域内的信息传送到远程控制管理中心。另外，远程管理中心也可以对网络节点进行实时控制和操纵。总之，传感网的全面感知是物联网的特点之一。

1.4 物联网的智慧运算

众所周知，人类进入计算机时代已有 69 载，计算对人们的影响每时每刻都在加剧。如果哪天人们离开了芯片、标签、晶体管和海量的有用数据，那么人们熟悉而又美丽的世界将不复存在。21 世纪是数据爆炸的时代，数据内容的使用者将彻底改变消费模式，数据内容的提供者也将彻底转变服务模式。IDC（International Data Corporation）的数据显示，目前全球数据量每 18 个月就要翻一番，全球每年产生的数据量已达 1~2EB（1EB=1024PB，1PB=1024TB）。面对全球数据量的成倍增长，传统的数据运算模式将不能胜任需要，运算效率和资源优化将再次成为影响业务发展和人们日常生活的瓶颈问题。

针对此种形势，2015 年 3 月 31 日，IBM 宣布成立物联网事业部，并表示会在接下来的 4 年时间里投资 30 亿美元用于研究，超过 2000 名研究员、软件开发者和咨询师将工作于这一项目，陆续提供一些基于云计算的服务以及帮助开发者获取互联网连接设备数据的工具。结合早在 2011 年 3 月正式推出了“智慧的运算”（Smarter Computing）战略，并提出“智慧的运算”是帮助人们实现“智慧的地球”的重要战略。“智慧的运算”为数据而生，为工作而优化，通过云计算管理。

云计算是适应当前和未来信息化需要的一种新型交付和使用模式，通过云计算可以管理整个基础架构，也可以帮助实现“智慧的运算”。云计算的四种服务模型如图 1.2 所示。

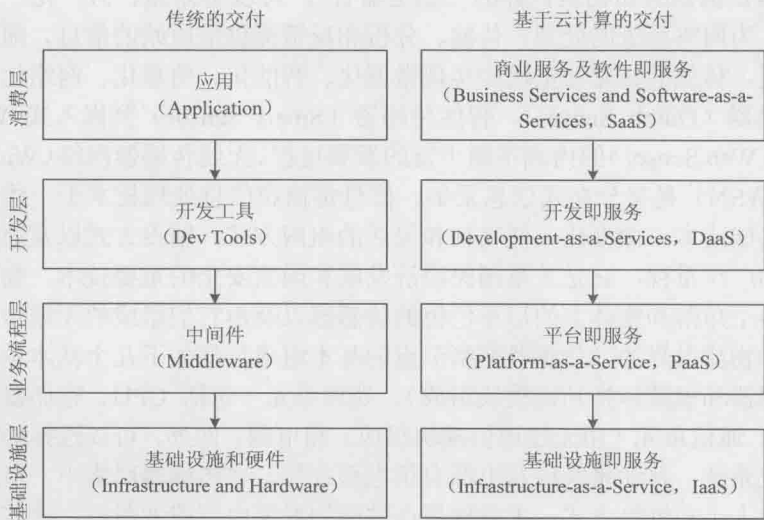


图 1.2 云计算的四种服务模型

据 IDC 统计，全球 IT 运维成本逐年攀升。与购买新服务器的成本增势平缓相比，服务器电力和冷却成本以及管理与维护成本在 IT 运维成本中的比重则逐年增加，而客户 IT 预算的增幅近年来却有不断放缓的趋势。

另外，随着企业的高速增长，IT 管理者发现他们的数据中心正面临巨大的挑战：越来越复杂的系统环境，越来越高的来自各个业务部门的数据服务需求，以及不断攀升的数据中心运营成本。面对这些挑战，企业越来越需要更加智能的数据中心来支持企业业务的快速发展。

“智慧的运算”提供了一种数字平台，用于整合硬件、基于分析的软件、网络管理服务和虚拟化，整个基础架构正在变得可以由软件定义，且工作负载可被所需分配到可编程的基础架构上。这一平台可通过三个主要特征来描述。

- (1) 为大数据设计 (Designed for Big Data)。大数据和信息集成功能可以从大量数据中获得洞察力，从根本上改变公司使用信息的方式。
 - (2) 由软件定义 (Software-Defined Environment)。为了应对当前的数据数量、多样性和速度，企业数据中心必须更动态、更灵活，在“由软件定义”的环境中，网络、存储和服务以及软件可以根据需求，随时对工作负载进行分配。
 - (3) 开放协作 (Open and Collaborative)。只有通过开放标准和平台，企业才能在当前丰富的创新生态系统中支持日益增多的非结构化数据、设备和服务，并且开展业务。
- 目前在 IT 领域的各个层面同样也发生着“智慧”的转变：得益于无处不在的智能

移动设备, IT 接入方法摆脱了原有的束缚, 海量数据应运而生; 通过改善流程生产力, IT 应用也更加注重改善流程生产力, 为每个独立用户提供实时的监测; 系统架构模式也由过去的异构孤岛向建立在工作负载优化系统之上的整合与更灵活的架构转变。

IT 的发展和应用正在以更快的速度和广度渗透和影响经济社会的方方面面。广大中小企业和各行各业用户, 都比以往更加重视信息化建设, 重视数据信息在个人和企业决策中所起的重要作用。而近年来兴起的物联网、云计算和虚拟化技术, 更是将 IT 发展推向了又一个高潮。

物联网是现代信息技术发展到一定阶段后出现的一种聚合性应用与技术提升, 它将各种感知技术、现代网络技术、人工智能与自动化技术聚合及集成应用, 创造了一个智慧的世界。物联网服务支持系统框架图如图 1.3 所示。

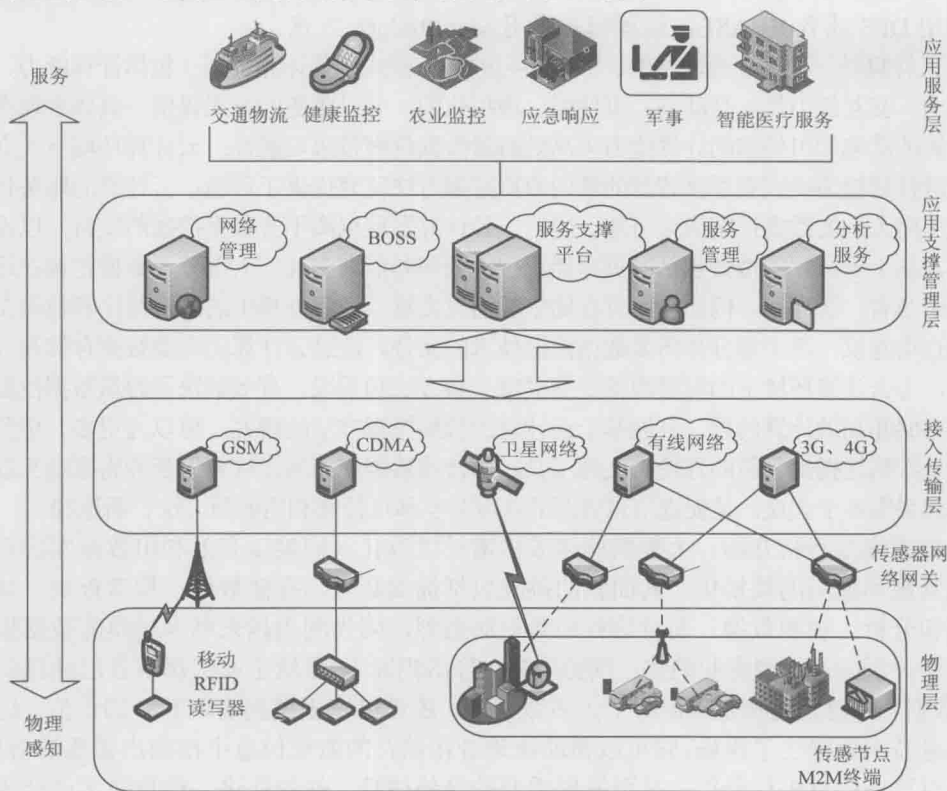


图 1.3 物联网服务支持系统框架图

1.5 物联网的智慧分析

21 世纪是数据爆炸的时代。由于竞争的本性和社会的发展, 很多企业倾向于获取足够多的、来自企业内部或者外部的数据, 以帮助企业做出准确决策, 实现持久永续

发展。这些数据通常与采集数据的工具、平台、分析系统一起被视为大数据 (Big Data)。数据挖掘 (Data Mining)，又称数据库中的知识发现，就是从大量数据中获取有效的、新颖的、潜在有用的、最终可理解的模式的过程。目前，数据挖掘技术已在金融、医疗、军事、管理等诸多领域的决策分析中广泛应用。

数据挖掘技术的发展经历了五代。第一代是单独算法、单个系统、单个机器，采用的是向量数据。第二代是与数据库相结合，支持多个算法。第三代是与预测模型相集成，支持 Web 数据、半结构化的数据，是一种网络化计算。第四代是分布式数据挖掘，是基于网格计算的多种算法，分布在多个节点上的方式。第五代是现在基于云计算的并行数据挖掘与服务的模式，同一个算法可以分布在多个节点上，多个算法之间是并行的，多个节点的计算资源实行按需分配，而且分布式计算模型采用云计算模式，数据用 DFS 或者 HBASE，编程模式采用 Map/Reduce 方式。

云计算是一种基于互联网的、大众参与的计算模式，其计算资源（包括计算能力、存储能力、交互能力等）是动态、可伸缩、虚拟化的，并以服务的方式提供。具体表现在：云计算的动态和可伸缩的计算能力为高效海量数据挖掘带来可能性；云计算环境下大众参与的群体智能为研究集群体智慧的新的数据挖掘方法研究提供了环境；云计算的服务化特征使面向大众的数据挖掘成为可能。同时，云计算发展也离不开数据挖掘的支持，以搜索为例，基于云计算的搜索包括网页存储、搜索处理和前端交互三大部分。数据挖掘在这几部分中都有广泛应用，例如，网页存储中的网页去重、搜索处理中的网页排序和前端交互中的查询建议，每个部分都需要数据挖掘技术的支持。通过云计算的海量数据存储和分布计算，为云计算环境下的海量数据挖掘提供了新方法和手段，有效解决了海量数据挖掘的分布存储和高效计算问题。开展基于云计算的数据挖掘方法的研究，可以为更多、更复杂的海量数据挖掘提供新的理论与支撑工具。而传统数据挖掘向云计算的延伸将推动互联网技术成果服务于大众，是促进信息资源的深度分享和可持续利用的新方法、新途径。

根据维基百科介绍，大数据就是成倍增长以致让人们难以有效利用数据库管理工具来管理和使用的数据集。其面临的挑战包括捕捉数据、存储数据、搜索查询、共享数据和分析、观测数据。而且这种趋势愈演愈烈，因为利用越来越大规模的数据集，有助于分析师“预测商业趋势，预防疾病，打击犯罪”。虽然企业的决策者已经意识到“大数据”中蕴含的价值，但对于大多数企业，还难以真正做到实现其中的价值。这时候信息技术就派上了用场，它可以帮助决策者在储存的海量信息中挖掘出需要的信息，并且对这些信息进行分析，从而发现重要的趋势信息。换句话说，信息技术已经成为“大数据”发挥作用的催化剂。在服务领域，“大数据”的重要性和价值越发明显地得到证明。与美国海洋和大气局 (National Oceanic and Atmospheric Administration, NOAA)、美国宇航局 (National Aeronautics and Space Administration, NASA) 等机构一样，一些制药企业和众多能源企业同样累积了大量的数据信息，现在这些企业想要将这些日常积累下来的数据转化为一种“大数据科技”，希望这些数据能够带来额外的价值。物联网智慧计算服务体系框架图如图 1.4 所示。