

面向海量数据应用的物联网信息服务系统研究综述^{*}

周开乐^{1,2}, 丁 帅^{1,2}, 胡小建^{1,2}

(1. 合肥工业大学 计算机网络系统研究所 合肥 230009; 2. 过程优化与智能决策教育部重点实验室, 合肥 230009)

摘 要: 物联网信息服务系统整合了基础设施层获取的海量数据,是面向各类物联网应用提供跟踪、监测和控制等信息服务的关键平台。阐述了物联网海量数据处理技术,包括存储、预处理和数据挖掘;总结了物联网信息服务系统在系统安全、隐私保护和可信性方面的安全保障机制;从系统架构、实现技术和系统实践三个方面,探讨了物联网信息服务系统的设计。最后,展望了这一领域面临的挑战和有待进一步研究的方向。

关键词: 物联网; 信息服务系统; 海量数据; 系统安全; 系统设计

中图分类号: TP18 文献标志码: A 文章编号: 1001-3695(2012)01-0008-04

doi:10.3969/j.issn.1001-3695.2012.01.002

Survey of massive data applications oriented
Internet of things information service system

ZHOU Kai-le^{1,2}, DING Shuai^{1,2}, HU Xiao-jian^{1,2}

(1. Institute of Computer Network System, Hefei University of Technology, Hefei 230009, China; 2. Key Laboratory of Process Optimization & Intelligent Decision-Making, Ministry of Education, Hefei 230009, China)

Abstract: Internet of things information service system integrates the massive data acquired from infrastructure layer, and it is the key platform to provide variety of information services for Internet of things applications, such as tracking, monitoring, controlling and so on. This paper summarized the massive data processing technologies of Internet of things from storage, pre-processing and data mining. Then, described the security mechanism from system security, privacy protection and reliability, and then analyzed the design of Internet of things information service system from architecture, techniques and applications. Finally, gave some future prospects of the challenges and research directions in this area.

Key words: Internet of things; information service system; massive data; system security; system design

0 引言

美国麻省理工大学 Kevin Ashton 教授于 1999 年提出了以标示为特征的物联网概念 (Internet of things, IoT)^[1]。2005 年,国际电信联盟 (International Telecommunications Union, ITU)在突尼斯举行的信息社会世界峰会 (World Summit on Information Society, WSIS)上正式确定了“物联网”的概念,并随后发布了物联网报告《ITU Internet reports 2005—the Internet of things》^[2]。ITU 物联网报告指出,连接现实物理世界与虚拟信息世界的物联网使得任何时间、任何地点以及任何物品的连接成为可能,而不再仅仅是互联网时代人与人的连接。物联网通过射频识别、传感器以及网络互连等技术,实现对物理世界各种物体的标志、感知、监测、跟踪、控制和管理,在人与物、人与环境以及物与物之间实现高效、智能、和谐的信息交互,最终通过智能决策控制技术实现物理世界与信息世界的融合。物联网已经成为学术界和商业领域关注的焦点,甚至成为很多国家的发展战略之一。美国国家职能委员会 (US National Intelligence Council, NIC) 预测,到 2025 年物联网节点将存在于任何物体之内,包括食品包装、家具、纸质文件等^[3]。

物联网通过基础设施层的各种感知设备感知物体信息,然后利用网络技术将感知获取的海量数据传输至物联网数据中

心。物联网信息服务系统整合了海量数据,为各种物联网应用提供信息服务并接收反馈控制。物联网的层次架构如图 1 所示。

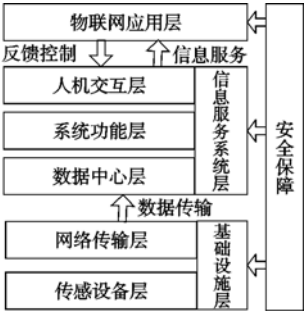


图1 物联网的层次架构

物联网通过各种传感设备获取的数据具有海量、异构、高维、冗余、时间序列相关和空间位置分散等特点。对物联网海量数据的存储、预处理和挖掘是物联网应用面临的关键问题。同时,物联网的开放架构要求更高的信息服务安全保障机制,包括信息服务系统的安全保障、用户隐私的保护和信息服务可信性的要求。物联网应用通过各种信息服务实现对物体的感知、跟踪、监测和智能控制等,物联网信息服务系统的设计是实现物联网信息服务的前提。

本文对面向海量数据应用的物联网信息服务系统的现有研究成果进行了归纳和总结,分析了物联网海量数据的处理技术,包括海量数据的存储、预处理和数据挖掘,主要从系统安全、隐私保护和信息服务可信性三个方面总结和论述了物联网信息服务系统的安全保障机制;同时探讨了物联网信息服务系统的设计问题,主要包括物联网信息服务系统的架构、实现技术和系统实践三个方面。最后展望了物联网信息服务系统研究面临的挑战和有待进一步研究的方向。

1 物联网海量数据处理

物理世界的物体数量庞大、形式多样、不断运动变化,分布在各种不同地点,且易受外界环境影响,这就导致物联网通过各种感知设备获取的数据具有海量、异构、高维、冗余、时间序列相关和空间位置分散等特点。数据存储是物联网海量数据处理面临的首要问题,传统的数据存储技术和数据库管理系统(DBMS)已经无法完全适用于物联网海量数据的存储管理。此外,为了节省存储空间、提高存储效率和便于数据挖掘,必须对物联网海量数据进行有效的预处理。物联网根据从海量数据中发现的知识为用户提供各种信息服务,研究适用于物联网海量数据特点的数据挖掘技术是实现物联网信息服务的前提。

对物联网海量数据的存储是物联网海量数据处理面临的首要问题。在无线传感器网络和 RFID 海量数据的存储管理方面,Yu 等人^[4]提出了无线传感器网络数据存储的一对一和多对多模型,并在此基础上提出了可实现全局数据存储位置最优的最优数据存储(ODS)算法和可实现局部数据存储位置最优的近似最优数据存储(NDS)算法。为解决有效管理物联网中不同设备产生的异构数据,Fan 等人^[5]设计了一种基于 Web Service 的物联网数据管理框架。在数据库管理系统方面,传统物联网数据库管理系统将不再适用于物联网海量数据的管理^[6],Gonzalez 等人^[7]提出了一个存储 RFID 海量数据的新模型 RFID-Cuboids。

物联网感知层获取的海量数据具有大量冗余,在进行数据挖掘之前,必须对粗糙杂乱的海量数据进行预处理。Fan 等人^[8]将本体论作为数据描述和语义数据存储的统一方式,提出了一个具有开放框架的数据管理平台,可将物联网海量杂乱数据转换为精益数据,并封装为服务包。RFID 数据是物联网海量数据的重要组成部分,大量文献研究了 RFID 数据的清理过滤问题,如 Jeffery 等人^[9]提出 SMURF 模型,可以对 RFID 数据流进行自适应平滑过滤清洗;Bai 等人^[10]研究了过滤 RFID 数据的多种有效方法,包括消除噪声和去除重复数据。

从物联网海量数据中发现有用的知识是实现物联网信息服务的前提,而对物联网海量数据特征的数据挖掘方法与传统的数据挖掘方法不完全相同。物联网海量数据的挖掘形式主要包括数据流分析、基于路径的分类和聚类分析、频繁模式和序列模式分析以及孤立点分析等。邓维维等人^[11]研究了普适环境下无线网络和移动设备产生的移动数据流挖掘问题,分析了数据流管理和数据流挖掘的特点,总结了移动数据流挖掘领域现有的研究成果,并对该领域的研究前景进行了展望。Masciari^[12]研究了 RFID 数据流的离群点挖掘问题,并用于解决货物装运的低效率或欺诈问题。Lee 等人^[13]采用基于区域的分层和基于轨迹的聚类,提出一种轨迹分类方法——TraClass。在传感器数据的知识发现方面,George 等人^[14]提出一种用于

传感器数据挖掘的时空传感器图(STSG)模型,STSG 模型可以发现异常模式、不同时间的位置和将来热点的节点等不同类型的模式。Rashidi 等人^[15]提出了一种新的自适应挖掘框架,用于传感器数据的模式挖掘。

通过上述对物联网海量数据处理技术现有成果的总结,可以看出对物联网海量数据处理的研究已经取得了许多成果。但是现有的物联网海量数据处理技术研究大部分是从某一方面展开的,如 RFID 数据的处理或传感器网络数据的处理,没有全面考虑物联网数据的多源、异构和分布式等特点。现有的数据处理技术主要用于特定格式的数据,而非对所有数据集均有效,同时对现有物联网海量数据处理技术的算法复杂度和效率分析较少。对物联网海量数据处理技术的研究缺乏系统性,这是有待进一步研究的问题。

2 物联网信息服务安全保障

物联网信息服务系统安全稳定的运行离不开有效的安全保障机制。物联网的架构开放性、感知的无处不在和交互的频繁性都对信息服务的安全、隐私保护和可信性提出了更高的要求。确保物联网信息服务的安全、可靠,是物联网技术大规模应用的前提,也是促使用户接受物联网信息服务的必要条件。

传统的信息服务安全保障机制有访问控制、防火墙、密钥机制、数字签名和身份识别、病毒扫描和查杀、入侵检测等,这些只能作为物联网信息服务安全保障机制的一部分,必须研究更加适合物联网信息服务系统的全面有效的安全保障机制。在物联网信息服务系统的安全保障方面,Wu 等人^[16]总结了物联网感知层和网络层的安全问题,包括传感器袭击、传感器异常、无线电干扰、网络内容安全、黑客入侵和非法授权等。Dlamini 等人^[17]分析了未来物联网应用可能面临的各种安全威胁,以及这些安全威胁出现的场景。Leusse 等人^[18]利用大规模安全系统、信息管理和自治系统中的相关知识,设计了一种物联网自我管理安全单元的架构模型。杨庚等人^[19]归纳了物联网的安全特征,提出了物联网安全技术架构,在此基础上讨论了密钥管理机制、认证与访问控制和入侵检测等物联网安全的关键技术。

物联网中无处不在的传感设备和频繁且广泛的信息服务交互使得物联网用户甚至非物联网用户的隐私安全面临很大的威胁。Sen^[20]提出了保护物联网中用户隐私数据的技术方法,这种方法考虑不同类型的安全问题,基于多方计算,采用不同等级的保护策略。Weber^[21]则指出了立法的重要性,强调通过立法保护物联网中的用户隐私。Feng 等人^[22]指出了物联网中尚未完全解决且普遍存在的隐私安全问题,并归纳了现有的保障物联网隐私安全的各种方法。

物联网信息服务系统的可信性是确保物联网信息服务有效实现的重要条件。传统的可信软件评估方法可以应用于物联网信息服务系统的可信性动态评估,杨善林等人^[23]在分析了复杂系统软件的可信性评估需求的基础上,设计了一种基于效用和证据理论的软件可信性评估推理模型。丁帅等人^[24]提出了应用于软件可信性评估指标系统自主配置的自适应重构期,并设计了一种需求驱动的软件可信性评估及演化模型。在物联网的可信性研究方面,Xiong 等人^[25]提出了基于物联网的可信安全系统架构,设计了包括可信用户模块、可信感知模块、可信终端模块、可信网络模块和可信代理模块的物联网可信系

统。Liu 等人^[26]根据敏感值将服务分类,建立服务分类评估表,同时将可信运营商作为所有服务提供者可信认证的一个代理,建立了保护用户安全的可信模型。Boukerch 等人^[27]针对无线传感器网络提出了一种基于 agent 的可信和声誉管理机制(ATRM),其目标是本地化管理可信和声誉,从而使额外资源和时间延迟开销最小化。

现有研究成果已经比较全面地指出了物联网信息服务面临的安全、隐私保护和可信性问题,而对具体安全保障技术的研究主要集中在物联网感知层和网络层,对物联网信息服务系统层的安全、隐私保护和可信问题研究较少。同时,现有的安全实现策略主要是将传统安全保障技术移植到物联网环境中,没有充分考虑传统安全技术 in 物联网环境中应用的局限性。针对物联网信息服务的特点,有待于进一步深入研究更加系统全面的物联网信息服务系统安全保障机制。

3 物联网信息服务系统设计

物联网信息服务系统是提供物联网信息服务的平台,也是实现物联网应用的载体。物联网信息服务系统底层面向海量数据、顶层面向庞大用户群,而且服务形式多样,因此对物联网信息服务系统的设计有别于对传统企业信息系统的设计。

确定物联网信息服务系统的架构是物联网信息服务系统设计的首要步骤。宁焕生等人^[28]设计了一种中国物联网信息服务系统架构,将中国物联网信息服务系统分为国家物联网管理中心、行业和专业物联网管理中心、本地物联网管理中心和 RFID 应用系统四个层级;同时设计了中国物联网信息服务系统服务器,并提出了中国物联网信息服务系统管理协议(RFID-MP)。Pu 等人^[29]提出了一种基于物联网上下文融合的智能交互架构,突破普适计算中虚拟与现实的界限,实现了更加智能的计算和服务。Michael 等人^[30]结合软件架构和工程学原理,广泛融合各种架构方案,提出了应用于某些特定案例的移动 RFID 服务(MRFID)架构方案 Quasi-Rich-Client。Zhong 等人^[31]考虑不同智能空间中各种物体之间的互操作性和协调性,指出了智能空间对象服务提供的特点,并给出了一个基于语义的自动框架。在框架中,对象服务提供、上下文识别、语义推理和普遍服务管理技术可以顺畅地合成。

国际电信联盟(ITU)在其《ITU Internet reports 2005—the Internet of things》报告中阐述了物联网的四个关键技术^[2],即标识事物的 RFID 技术、感知事物的传感器技术、思考事物的智能技术以及微缩事物的纳米技术。在物联网信息服务系统的实现技术方面,孙其博等人^[32]指出了计算与服务技术的重要性,同时管理与支撑技术也是实现物联网信息服务系统的重要技术。王保云^[33]论述了物联网信息服务系统中的智能技术、人工智能技术、人一机交互技术与系统、智能控制技术与系统和智能信号处理是实现物联网信息服务的重要技术。人一机交互技术是物联网信息服务系统实现信息服务的关键技术之一。Kranz 等人^[34]采用多种原型开发了各种人机交互接口(HCI),在日常物品中使用嵌入式交互技术,实现了日常物品与物联网的无缝通信,可以便捷而有效地使用物联网信息服务。Broll 等人^[35]提出了一种物联网信息服务的移动交互框架 Perci,它将 Web Service 和移动交互与物理世界的物体整合,通过信息输入增强互操作性。

物联网在实现“智慧地球”的广泛普及应用之前,已经在

局部特定领域得以成功实践。王粉花等人^[36]研究了以物联网中 ZigBee 无线通信技术为基础的人体运动状态监测系统的设计方案,通过对所测数据的分析,实现了对人体状态的实时监控判断,从而满足了老龄化环境下的老年人护理需求。Darian 等人^[37]设计了一种基于物联网的智能家居服务系统,该系统在减少费用、能源消耗和复杂性方面取得了效益。黄冬梅等人^[38]利用物联网技术构建了一种软硬件结合的救灾物资管理应用系统,该系统可以对整个救灾物资运输过程进行全面的辅助决策,并且可以对一些突发情况进行预警。Luo 等人^[39]设计了一种远程监测信息系统,通过人体监测传感器芯片和物联网远程获取病人的详细生理数据,并自动生成电子病历,在信息平台的支持下,系统会反馈相应的诊断和用药建议。

目前,对物联网信息服务系统设计的研究主要集中在面向某一特定领域的物联网信息服务系统设计,对面向不同应用的信息服务的共性问题研究较少;同时,现有的研究偏向于理论框架的研究,对具体实现技术的研究不够系统和深入,对系统实施的可行性论证不够完善。物联网信息服务系统实施中可能面临的应急保障、服务接口和大容量群体访问等问题也是需要进一步研究的内容。

4 结束语

随着 RFID 技术、传感器技术、网络技术以及智能技术的进一步发展,物联网技术将给整个社会生产生活方式带来深远的变革。物联网信息服务系统作为提供各种物联网应用服务的重要平台,在海量数据处理、系统安全和系统设计等方面还有许多需要进一步深入研究的问题。

在物联网海量数据处理方面,针对物联网数据海量的特点,需要进一步探讨简单、高效和稳定可靠的海量数据存储技术。对物联网海量数据的预处理要充分考虑在尽可能去除冗余的同时而又不影响其完整性和真实性。在对分布式异构融合的物联网海量数据挖掘方面,要进一步研究数据规约和提取技术,从海量数据中规约或提取,以高效、全面挖掘出所需知识的特征数据集。针对不宜简化或者难以简化的复杂海量数据的挖掘,要重点考虑算法的复杂度和效率问题。

在物联网信息服务系统安全方面,可以在传统信息安全技术的基础上,研究更加适合物联网信息服务特点的安全保障机制。努力建立更加全面、有效和可靠的信息服务系统安全性、隐私保护和可信性方面的安全保障机制。

在物联网信息服务系统的设计方面,需要更加充分考虑物联网信息服务系统中的共性问题,即适用于各种不同类型物联网信息服务应用的统一系统架构和实现技术。同时,对实现物联网信息服务系统的关键技术的研究也有待进一步深入。此外,由于物联网具有系统开放性和应用形式多样性的特点,在设计物联网信息服务系统时,要充分考虑物联网用户群体类型的差异化和服务需求的多样性,物联网信息服务系统应该在面向普通大众用户提供一般共性服务的同时,具有满足少数特定用户群体个性化信息服务需求的差异化功能。

通过对面海量数据应用的物联网信息服务系统研究现状的总结归纳和对未来的展望,有助于促进物联网信息服务更好地实现和物联网技术更广泛地普及。

参考文献:

[1] CONTI P. The Internet of things[J]. Communications Engineer,

- 2006, 4(6):20-25.
- [2] International Telecommunication Union, Internet reports 2005: the internet of things[R]. Geneva: ITU, 2005.
- [3] ATZORI L, IERA A, MORABITO G. The Internet of things: a survey[J]. *Computer Networks*, 2010, 54(15): 2787-2805.
- [4] YU Zhao-chun, XIAO Bin, ZHOU Shui-geng. Achieving optimal data storage position in wireless sensor network[J]. *Computer Communications*, 2010, 33(1): 92-102.
- [5] FAN Tong-rang, CHEN Yan-zhao. A scheme of data management in the Internet of things[C]// Proc of the 2nd IEEE International Conference on Network Infrastructure and Digital Content. 2010: 110-114.
- [6] JEFFERY K G. The Internet of things: the death of a traditional database? [J]. *IETE Technical Review*, 2009, 26(5): 313-319.
- [7] GONZALEZ H, HAN Jia-wei, LI Xiao-lei, *et al.* Warehousing and analyzing massive RFID data sets[C]// Proc of the 22nd International Conference on Data Engineering. 2006: 83-92.
- [8] FAN Chun-xiao, SONG Jie, WEN Zhi-gang, *et al.* A scalable Internet of things lean data provision architecture based on ontology[C]// Proc of IEEE GCC Conference and Exhibition. 2011: 553-556.
- [9] JEFFERY S R, GAROFALAKIS M, FRANKLIN M J. Adaptive cleaning for RFID data streams[C]// Proc of the 32nd International Conference on Very Large Database. 2006: 553-556.
- [10] BAI Yi-jian, WANG Fu-sheng, LIU Pei-ya. Efficiently filtering RFID data streams[C]// Proc of CleanDB Workshop. 2006: 50-57.
- [11] 邓维维, 彭宏, 郑启伦. 基于数据流的移动数据挖掘研究综述[J]. *计算机应用研究*, 2007, 24(1): 5-9.
- [12] MASCIARI E. A framework for outlier mining in RFID data[C]// Proc of the 11th International Database Engineering and Applications Symposium. 2007: 263-267.
- [13] LEE J G, HAN Jia-wei, LI Xiao-lei, *et al.* TraClass: trajectory classification using hierarchical region-based and trajectory-based clustering[J]. *Proceedings of the VLDB Endowment*, 2008, 1(1): 1081-1094.
- [14] GEORGE B, KANG J M, SHEKHAR S. Spatio-temporal sensor graphs (STSG): a data model for the discovery of spatio-temporal patterns[J]. *Intelligent Data Analysis*, 2009, 13(3): 457-475.
- [15] RASHIDI P, COOK D J. An adaptive sensor mining framework for pervasive computing applications [C]// Proc of the 2nd International Workshop on Knowledge Discovery from Sensor Data. 2008: 41-49.
- [16] WU Chuan-kun. A preliminary investigation on the security architecture of the Internet of things[J]. *Bulletin of Chinese Academy of Sciences*, 2010, 25(4): 411-419.
- [17] DLAMINI M T, ELOFF M M, ELOFF J H P. Internet of things: emerging and future scenarios from an information security perspective [C]// Proc of Southern Africa Telecommunication Networks and Applications Conference. 2009: 6-11.
- [18] De LEUSSE P, PERIORELLIS P, DIMITRAKOS T, *et al.* Self managed security cell, a security model for the Internet of things and services[C]// Proc of the 1st International Conference on Advances in Future Internet. 2009: 47-52.
- [19] 杨庚, 许建, 陈伟, 等. 物联网安全特征与关键技术[J]. *南京邮电大学学报: 自然科学版*, 2010, 30(4): 20-29.
- [20] SEN J. Privacy preservation technologies in Internet of things[J]. *International Journal BITM Transactions on EECC*, 2009, 1(4): 496-504.
- [21] WEBER R H. Internet of things: new security and privacy challenges [J]. *Computer Law & Security Review*, 2010, 26(1): 23-30.
- [22] FENG Hai-long, FU Wen-xiu. Study of recent development about privacy and security of the Internet of things[C]// Proc of International Conference on Web Information Systems and Mining. 2010: 91-95.
- [23] 杨善林, 丁帅, 褚伟. 一种基于效用和证据理论的可信软件评估方法[J]. *计算机研究与发展*, 2009, 46(7): 1152-1159.
- [24] 丁帅, 鲁付俊, 杨善林, 等. 一种需求驱动的软件可信性评估及演化模型[J]. *计算机研究与发展*, 2011, 48(4): 647-655.
- [25] XIONG Li, ZHOU Xuan, LIU Wen. Research on the architecture of trusted security system based on the Internet of things[C]// Proc of the 4th International Conference on Intelligent Computation Technology and Automation. 2011: 1172-1175.
- [26] LIU Yang, CHEN Zhi-kun, XIA Feng, *et al.* A trust model based on service classification in mobile services[C]// Proc of IEEE/ACM International Conference on Green Computing and Communications & International Conference on Cyber, Physical and Social Computing. 2010: 572-577.
- [27] BOUKERCH A, XU Li, EL-KHATIB K. Trust-based security for wireless Ad hoc and sensor networks[J]. *Computer Communications*, 2007, 30(11-12): 2413-2427.
- [28] 宁焕生, 张瑜, 刘芳丽, 等. 中国物联网信息服务系统研究[J]. *电子学报*, 2006, 34(12A): 2514-2517.
- [29] PU Hai-tao, LIN Jin-jiao, LIU Fa-sheng, *et al.* An intelligent interaction system architecture of the Internet of things based on context [C]// Proc of the 5th International Conference on Pervasive Computing and Applications. 2010: 402-405.
- [30] MICHAEL M P, DARIANIAN M. Architectural solutions for mobile RFID services for the Internet of things[C]// Proc of IEEE Congress on Services. 2008: 71-74.
- [31] ZHONG Dong, ZHU Yian, LEI Wang-bao, *et al.* Object service provision in Internet of things[C]// Proc of International Conference on e-Education, e-Business, e-Management, and e-Learning. 2010: 234-237.
- [32] 孙其博, 刘杰, 黎彝, 等. 物联网: 概念、架构与关键技术研究综述[J]. *北京邮电大学学报*, 2010, 33(3): 1-9.
- [33] 王保云. 物联网技术研究综述[J]. *电子测量与仪器学报*, 2009, 23(12): 1-7.
- [34] KRANZ M, HOLLEIS P, SCHMIDT A. Embedded interaction: interacting with the Internet of things[J]. *Internet Computing*, 2010, 14(2): 46-53.
- [35] BROLL G, RUKZIO E, PAOLUCCI M, *et al.* PerCi: pervasive service interaction with the Internet of things[J]. *Internet Computing*, 2009, 13(6): 74-81.
- [36] 王粉花, 年忻, 郝国梁, 等. 物联网技术在生命状态监测系统中的应用[J]. *计算机应用研究*, 2010, 27(9): 3375-3377, 3380.
- [37] DARIANIAN M, MICHAEL M P. Smart home mobile RFID-based Internet-of-things system and service [C]// Proc of International Conference on Advanced Computer Theory and Engineering. 2008: 116-120.
- [38] 黄冬梅, 方的, 张明华, 等. 物联网技术在救灾物资配送管理系统中的应用[J]. *计算机应用研究*, 2011, 28(1): 189-191.
- [39] LUO Jing-ran, CHEN Yu-lu, TANG K, *et al.* Remote monitoring information system and its applications based on the Internet of things [C]// Proc of International Conference on Future BioMedical Information Engineering. 2009: 472-475.