

RFID 中间件数据处理研究

孙 红, 厉彦刚, 陈世平

(上海理工大学 光电信息与计算机工程学院, 上海 200093)

摘要: RFID 中间件位于读写器和企业应用系统之间, 管理它们之间的信息数据流, 对企业应用至关重要. 为研究中间件的数据处理, 介绍了 RFID 技术及 RFID 中间件的功能、特点和架构设计. 结合国内外研究现状对 RFID 中间件数据处理作了详细介绍, 重点从数据存储、数据过滤、生产数据源和安全控制等几个方面介绍了 RFID 中间件数据处理的一般流程. 从企业实际应用的角度出发, 提出了一种利用集合思想和反馈思想来改善 RFID 中间件的数据处理流程, 为 RFID 中间件数据处理提出了一种新的思路. 对改进后的处理流程进行了分析, 发现可以改善 RFID 中间件的数据处理过程.

关键词: 射频识别; 中间件; 数据处理; 集合; 反馈

中图分类号: TP 391

文献标志码: A

Data Processing of RFID Middleware

SUN Hong, LI Yan-gang, CHEN Shi-ping

(School of Optical-Electrical and Computer Engineering, University of Shanghai for Science and Technology,
Shanghai 200093, China)

Abstract: RFID (radio-frequency identification) middleware lies between readers and enterprise application systems. It manages the information dataflow between them and it is very important to the enterprise application. In order to learn the data processing of RFID middleware, the RFID technology and RFID middleware's functions, features and architecture design were introduced, which summarized the current research results both at home and abroad and the data processing of RFID middleware in detail. The data processing procedure was described mainly from the aspects of data storing, data filtering, data source producing, security control, etc. From the perspective of enterprise application, a method to improve the data processing procedure was provided by introducing the ideas of gather and feedback. The improved process was analyzed and its advantages were pointed out.

Key words: *RFID; middleware; data processing; gather; feedback*

收稿日期: 2013-06-04

基金项目: 国家自然科学基金资助项目(61170277); 上海市教委科研创新(重点)资助项目(12ZZ137); 上海市一流学科建设资助项目(S1201YLXK); 上海市研究生教育创新基金资助项目(SHGEUSST1301)

第一作者: 孙 红(1964-), 女, 副教授. 研究方向: 计算机网络通信与云计算、计算机科学与技术、控制科学与工程. E-mail: sunhong_sh@sohu.com

近年来物联网技术的发展非常迅速,有关物联网技术的研究也成为热点^[1]. 射频识别(radio-frequency identification, RFID)技术可应用于物联网感知层,它是一种应用于信息采集系统的非接触式自动识别技术,通过无线射频方式自动识别目标对象,获取相关数据信息^[2]. 随着电子、通信与信息技术的迅速发展,RFID技术已经进入商业化广泛应用的阶段,已经被广泛应用于物流、交通运输、门禁和医疗等领域,被誉为21世纪最有发展前景的技术之一. 由于通过标签和读写器感知的数据具有海量、异构、高维、冗余等特点,为了更好地处理这些数据流,自动识别产品技术中心(Auto ID Center)推出了一种分层、模块化的Savant(即RFID中间件). RFID中间件介于前端读写器和后端应用程序之间,它使得RFID应用系统开发变得更容易,提高了软件的可移植性,增强了系统的可维护性和可靠性^[3]. RFID中间件可以对数据流进行缓存、过滤、挖掘等操作,从而减少由阅读器传往企业应用的海量原始数据,增加抽象出有意义的信息量,提高效率.

1 RFID中间件概述

RFID中间件是实现RFID硬件设备与应用系统之间数据存储、过滤、传输等的一种中间程序,将RFID读写器读取的各种信息数据,经中间件处理后导入企业的信息数据管理系统,并通过应用系统供用户操作.

1.1 RFID中间件发展阶段

从发展趋势看,RFID中间件大体可分3个阶段:应用程序阶段,该阶段多为RFID读写器厂商提供简单应用程序接口(API),供用户把后端系统与RFID读写器相连;基础架构阶段,该阶段中间件具有数据搜集、过滤等功能,满足企业多对多连接要求,同时具备平台管理与维护功能;解决方案阶段,该阶段中间件可以为不同用户提供更多不同功能,具有更多新应用的解决方案.

1.2 RFID中间件特点

RFID中间件是RFID系统的神经中枢,其特点主要有:

a. 独立性. 中间件提供一组通用的API,即使存储RFID标签数据的数据库软件或后端应用程序增加,甚至RFID读写器种类增加等情况发生,应用端不需修改也能处理,省去多对多连接的维护复杂性问题^[4].

b. 数据流. RFID中间件具有对数据采集、过滤、传递等功能,将准确的对象信息传递到企业后端的应用系统中.

c. 处理流. 提供顺序的消息流,具有数据流设计与管理能力,维护数据传输路径、安全性、路由和数据分发规则.

1.3 RFID中间件架构设计

根据数据采集、数据处理和数据发布3个层次设计RFID中间件架构. RFID中间件结构如图1所示.

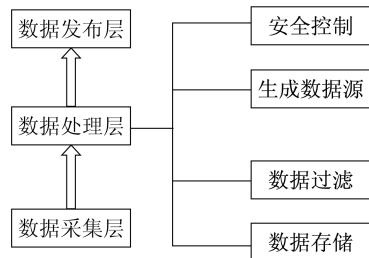


图1 RFID中间件架构

Fig.1 RFID middleware architecture

a. 数据采集层. 含有读写器接口(RS 232, RS 485, USB 2.0等),负责为整个系统提供准确的实时数据. 由于不同型号的读写器采集到的数据格式不同,该层用读写适配器把不同格式的数据转化为系统需要的统一格式,最后输出的数据为格式化的RFID及其它现场数据. 数据采集层读写器接口如图2所示.

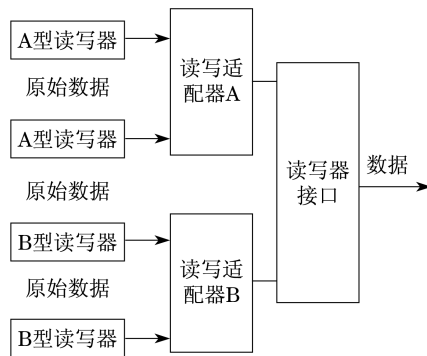


图2 数据采集层读写器接口图

Fig.2 Reader interface chart at data collection layer

b. 数据处理层. 该层是RFID中间件核心,对数据进行存储、过滤、生产数据源等处理,为后面提供更加准确和精简的数据,同时进行安全控制.

c. 数据发布层. 为用户提供分类数据源,供不同需求的用户使用. RFID中间件为应用程序提供多个服务接口,可为用户提供RFID数据,提供用户添加、删除、编辑等RFID数据操作.

2 RFID 中间件数据处理

从标签得到的数据有异构、高维、冗余和空间位置分散等特点^[5],需要对这些数据进行处理后才能供后端企业应用。

a. 数据存储. 多个读写器同时工作的时候,每秒就可以读几百个标签,产生的数据流速度极快,而现在数据库不具有这么快的存储能力.因此有必要在数据源和数据库之间添加一个有速度协调能力的单元,一般通过消息队列缓存数据来实现^[6].消息队列缓存机制见图3。

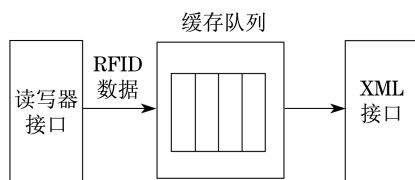


图3 消息队列缓存机制

Fig.3 Message queue cache mechanism

来自 RFID 的数据通过验证后,传递到消息队列服务器.消息队列服务器对数据进行缓存,同时在网络暂时不可用或不稳定时,还可以保证数据传输的可靠性.然后通过 XML (extensible markup language)接口将消息队列服务器上的数据消除,得到数据存储为 XML 文件.XML 是通用数据格式,可以通过生成数据源模块将数据移植到后端数据库。

b. 数据过滤. 来自读写器的海量数据存在大量冗余,也存在一些错误信息数据,此时需要对数据进行过滤.在数据过滤方面,Jeffery 等提出了 SMURF 模型^[7],可以对 RFID 数据流进行自适应平滑过滤清洗.Bai 等研究了过滤 RFID 数据的多种有效方法,包括消除噪声和去除重复数据.分散的平方根信息过滤 (decentralized square-root information fliter,DSRIF)方法也可以提高数据的精确性、可靠性^[8].在这里选择哈希表来进行过滤,设标签身份 (TagID)为哈希表的关键字, t 为时间间隔.缓存的数据与哈希表里的数据进行比较,得到相应结论后再重新开始,具体算法如图4所示。

在 XML 文件的过滤中,要考虑如何减少过滤时的开销,提高过滤的效率,从而追求效益的最大化.这里选择哈希表进行数据过滤的优势在于可以减少无效元素的处理,减少文档的解析量.哈希表存储了元素的位置信息,可以快速判断文档中任意两个元素的层次关系,以便判断元素是否为无效元素,

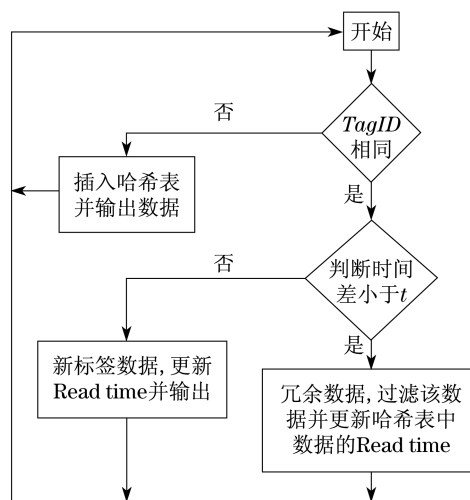


图4 过滤算法流程图

Fig.4 Filter algorithm flow chart

并跳过无效元素,从而提高过滤的效率.当输入的文档深度较大(如 RFID 海量数据流)时,本文中的方法要优于传统 XML 过滤方法的效率.SMURF 模型更倾向于数据的自适应平滑过滤,而 DSRIF 方法则以数据的精准性为主.从企业实际应用的角度看,哈希表法更具优势。

c. 生成数据源. RFID 数据经过存储和过滤后是对用户有用的数据,通过数据源生产模块将 XML 数据转移到数据库中,供后端应用程序使用.数据源生成模块包括 XML 编程接口和数据库编程接口,XML 编程接口对 XML 文档解析、定位和查询,数据库编程接口则访问数据库,如数据库的更新和检索等.该模块基本思想是判断是不是第一次生成数据源,若是则将 XML 表影射到数据库中,生成表结构进行数据移植;若不是只需进行数据移植即可。

d. 安全控制. RFID 数据的采集是通过无线射频方式得到的,在无线传输过程中也会受到内部热噪声和外部电磁波等干扰.在这里采用循环冗余码校验(CRC).CRC 是数据通信领域中最常用的一种差错校验码,其特征是信息字段和校验字段的长度可以任意选定,可以对同一数据块中字节互换校验.基本原理是:在 K 位信息码后再拼接 R 位的校验码,整个编码长度为 N 位,因此,这种编码也叫 (N, K) 码.对于一个给定的 (N, K) 码,可以证明存在一个最高次幂为 $N - K = R$ 的多项式 $G(x)$.根据 $G(x)$ 可以生成 K 位信息的校验码,而 $G(x)$ 叫做这个 CRC 码的生成多项式.校验码的具体生成过程为:假设发送信息用信息多项式 $C(x)$ 表示,将 $C(x)$ 左移 R 位,则可表示成 $C(x)$ 乘以 x 的 R 次

方,这样 $C(x)$ 的右边就会空出 R 位,这就是校验码的位置.通过 $C(x)$ 乘以 2 的 R 次方除以生成多项式 $G(x)$ 得到的余数就是校验码.

3 RFID 中间件数据处理流程改进

现阶段 RFID 数据主要是面向企业应用的,由于不同企业之间差异较大,需要的信息数据也不尽相同.因此在数据处理之前将数据进行大体归类,企业在应用时可以根据自身情况进行选取,只选择企业需求的数据信息,不需对所有数据进行处理,这就大大提高了数据处理效率.

3.1 集合思想的应用

集合的概念是现代数学中最基本的概念之一,集合论已经是数学教育中一个不可或缺的部分,并已深入到各种科学和技术的领域中.集合里所含有的个体称为集合的元素,所有元素组成的集合叫做全集,部分元素组成的集合成为全集的子集.

将与标签数据有关的原始属性记作一个全集,如{读取地点,读取时间,标签类型,读取次数...}.该全集有不同子集,如{读取时间,读取地点}、{读取次数,读取类型}、{读取时间,读取地点,读取次数}等.不同企业可以根据自身实际情况选择子集,比如某企业只对标签读取数据时的时间和地点感兴趣,则该企业可只选择子集{读取时间,读取地点},而不需要其它数据信息,这样可大大减少数据处理量,加快处理效率.

3.2 反馈控制的应用

反馈是控制论的一个极其重要的概念.反馈就是由控制系统把信息输送出去,又把其作用结果返送回来,并对信息的再输出发生影响,起到控制的作用,以达到预定目的.原因产生结果,结果又构成新的原因、新的结果,如此循环.反馈在原因和结果之间架起了桥梁.

由于来自于标签的数据流信息量较大,不同企业经过中间件数据处理后得到的信息数据侧重点各不相同.在经过一段时间的选择后,中间件会把企业需求较大的信息数据记录.如果把这些信息数据反馈到标签数据流,就可以提高数据流的有效性,提高系统的实用性和效率.

3.3 改进后的数据处理流程分析

下面简要分析一下经过集合思想和反馈控制思想改进后的 RFID 中间件数据流程.

第一,RFID 数据流的基本单元是单个的标签,

具有编码、来源等原始属性.这些属性正是企业应用系统关注的核心信息.企业可选择符合自身需求的数据信息,而不是数据的全部信息,这样节省了开销,提高了效率,集合思想的使用恰好解决该问题.

第二,一方面会把企业需求较大的信息数据记录,把这些信息数据反馈到标签数据流;另一方面,由于企业在不同发展阶段会有不同的数据信息需求,这就使得对数据集合的选择不是静态的,而是随着企业的发展阶段动态变化,这就需要一个反馈来动态选择原始数据.被控量为信息数据,控制目标则是企业实现对动态数据的管理,反馈控制的应用则可解决该问题.

第三,稳定性分析.标签数据单元在系统流动的过程中会产生新的特征,如数据在系统中停留时间长短、流动路径等^[9].数据原始属性有所改动时,系统的应用便会受到影响,这时企业应用系统会检测到数据问题,从而通过反馈来重新调整原始数据的选择,满足企业应用要求,从而维持系统动态稳定.经过集合思想和反馈控制思想改进后的 RFID 中间件数据流程如图 5 所示.

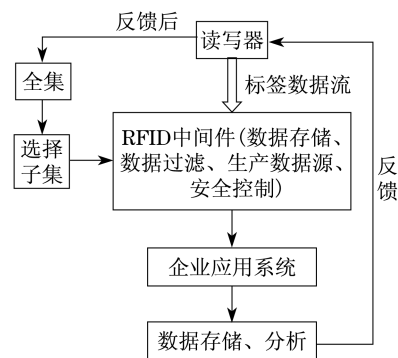


图5 RFID中间件数据处理流程改进

Fig.5 Modified flow chart of RFID middleware data processing

4 总结与展望

RFID 中间件研究仍旧是一个较新的研究内容,它在海量数据的处理、传输等方面有着举足轻重的作用.本文介绍了 RFID 技术及 RFID 中间件,对 RFID 中间件的数据处理分几个模块进行介绍,并从企业应用的角度,提出了利用数学集合思想以及反馈思想来改进数据处理的流程,为 RFID 中间件数据处理提供了一种思路.

在 RFID 中间件数据处理方面,还有很多内容.如在数据传输时,可对数据进行压缩减少数据传输

量;可以使用编码和算法来保证传输过程的数据安全;可对整个数据处理过程建立模型来综合处理数据.相信在不久的将来,越来越多的数据处理方法会应用在RFID中间件中,更为有效地提高数据处理的效率.同时,RFID中间件技术的发展也会为物联网、云计算^[10]等技术提供更好的帮助.

参考文献:

- [1] 章刚,陈庆奎.一种物联网群体访问路由算法[J].上海理工大学学报,2012,34(4):337-342.
- [2] 杨刚,沈沛意,郑春红,等.物联网理论与技术[M].北京:科学出版社,2012.
- [3] 成修治,李宇成.RFID中间件的结构设计[J].计算机应用,2008,28(4):1055-1060.
- [4] 张春红,裘晓峰,夏海轮,等.物联网技术与应用[M].北京:人民邮电出版社,2011.
- [5] 周开乐,丁帅,胡小建.面向海量数据应用的物联网信

息服务系统研究综述[J].计算机应用研究,2012,29(1):8-11.

- [6] 邓海生,李军怀,刘红英.基于RFID的数据采集中间件[J].计算机技术与发展,2007,17(9):188-191.
- [7] Jefery S R, Garofalakis M, Franklin M J. Adaptive cleaning for RFID data streams[C]//Proceedings of the 32nd International Conference on Very Large Data Bases. San Francisco:IEEE Computer Society,2006.
- [8] 丁振华,李锦涛,罗海勇,等.RFID系统与传感器网络中的数据处理综述[J].计算机应用研究,2008,25(3):660-665.
- [9] 张明哲,张强,袁巍,等.嵌入式RFID中间件数据过滤模型研究[J].计算机工程与设计,2010,31(17):3743-3746.
- [10] 秦桂英,刘宇熹.基于云计算的供应链运营模式研究[J].上海理工大学学报,2013,35(3):245-250.

(编辑:丁红艺)

(上接第226页)

1 000,仿真结果显示这种混合算法具有更高的精度、更容易收敛到全局最优的优点.

参考文献:

- [1] Kennedy J, Eberhart R. Particle swarm optimization [C]//Proceedings of IEEE International Conference on Neural Networks. Perth,1995.
- [2] 吕丹,童创明,钟卫军.基于粒子群和模拟退火算法的混合算法研究[J].计算机工程与设计,2011,32(2):663-666.
- [3] 马慧民,叶春明,张爽.二进制改进粒子群算法在背包问题中的应用[J].上海理工大学学报,2006,28(1):31-34.
- [4] 马慧民,吴勇,叶春明.车辆路径问题的并行粒子群算法研究[J].上海理工大学学报,2007,29(5):435-439.
- [5] 李晓磊,邵之江,钱积新.一种基于动物自治体的寻优

模式:鱼群算法[J].系统工程理论与实践,2002,22(11):32-38.

- [6] 邹康,刘婷,鲍韦韦,等.人工鱼群算法研究综述[J].山西电子技术,2012(2):92-93.
- [7] 罗德相,周永权,黄华娟.粒子群和人工鱼群混合优化算法[J].计算机与应用化学,2009,26(10):1257-1261.
- [8] Maleki H R, Tata M, Mashinchi M. Linear programing with fuzzy variables [J]. Fuzzy Sets and Systems, 2000,109(1):21-33.
- [9] Shi Y H, Eberhart R. A modified particle swarm optimizer [C] // IEEE World Congress on Computational Intelligence. Anchorage,1998.
- [10] 杨光友.混合粒子群优化及其在嵌入式智能控制中的应用[D].武汉:武汉理工大学,2006.
- [11] 刘勇,马良,宁爱兵.函数优化的量子竞争决策算法[J].计算机工程与应用,2010,46(21):21-24.

(编辑:丁红艺)