

数据融合处理算法在 ZigBee 中的应用

王 平,程明传,翁宗煌

(福建师范大学 物理与光电信息科技学院,福州 350007)

(pwang@fjnu.edu.cn)

摘 要:在标准 ZigBee 协议中没有设计相关的数据融合规范,使其只能用在低数据冗余的应用场合。针对大规模网络,网络中的数据冗余度很大,网络中的数据冗余会引起节点频繁地争抢信道,网络时延增加甚至出现网络瘫痪。为了解决这个问题提出了两层数据融合方法:第一层设计了终端节点到路由节点之间的统计融合,第二层设计了路由节点到网关的神经网络数据融合。实验证明,数据融合有利于降低数据冗余、改善查询效率、降低能量消耗,也使 ZigBee 协议适合高数据冗余的应用场合。

关键词:ZigBee 协议;数据融合;人工神经网络;智能温室

中图分类号:TP18 **文献标志码:**A

Application of data confusion algorithm in ZigBee protocol

WANG Ping, CHENG Ming-chuan, WENG Zong-huang

(College of Physics and OptoElectronic-Infomational Technology, Fujian Normal University, Fuzhou Fujian 350007, China)

Abstract: Related specialization of data fusion is not considered in Standard ZigBee protocol and this limits its application to low data redundancy. Data redundancy of mess network is high, which will lead to frequent channel competition, time delay of network and network paralysis. Two-layer data confusion was proposed to solve this problem; the first layer is on statistical confusion between terminal nodes and router nodes, and the second layer is on artificial neural network between router nodes and gateway node. Experiments prove data confusion is helpful to reduce data redundancy, improve query effect and reduce energy loss, and make ZigBee protocol suitable for high data redundancy.

Key words: ZigBee protocol; data fusion; Artificial Neural Network (ANN); intelligent greenhouse

0 引言

ZigBee 网络是资源受限的网络^[1],集中表现在电池能量、数据整合能力、存储空间及通信带宽等非常有限。终端节点采集的原始数据直接传送到路由节点是现在普遍采用的方法,但由于原始数据信息冗余性的存在会浪费有限的电池能量和通信带宽;频繁的信道竞争也会降低信息收集的效率。在只关注监测结果或只关注特征数据的场合中,传送大量原始数据给网关节点是不必要的,采用数据融合来降低数据冗余度能很好的满足这种需求。

ZigBee 的应用领域主要为:智能家居、工业自动化、医疗设备和环境监测系统^[2]。在智能家居、智能建筑和医疗设备中,数据的冗余度比较低,但是在工业自动化和监测系统中,数据的冗余度较高。数据融合是降低数据冗余度,提高数据收集效率有效方式。射频收发一个比特的数据消耗的能量是执行一条指令消耗能量的数百倍,使用数据融合可以显著较少射频收发的次数。ZigBee 协议只解决如何形成自组织网络的问题,其他问题都留在应用层自行解决。本文讨论了在 ZigBee 协议的应用层实现两层的数据融合算法,此方法有利于降低网络的信息冗余度,以提高数据的准确性,延长网络的寿命,在一定程度上扩大了 ZigBee 的应用范围。

1 ZigBee 网络的组成

无线传感网络^[3-4]理论的不断成熟,以及对低功耗、低速

率、低成本、远距离无线通信的强烈需要,导致了 ZigBee 协议的产生。和以太网相似,ZigBee 网络也包含三种设备:网关、路由器和终端设备。网关是整个网络的“大脑”,其主要功能包括分配网络地址、建立网络、维护网络和管理网络等。路由器和终端设备一般都是使用低成本的控制单元,只是软件上有所区别,路由节点具有路由功能,而终端节点没有。值得注意的是,对于小型网络,ZigBee 网络中的三种设备可以使用相同的控制器。对于大规模 ZigBee 网络^[5],网关的负担加重,需要强有力的硬件支撑。ZigBee 星形网络如图 1 所示。

对于数据冗余度比较高的应用,比如环境监测、安全检测、军事应用等,如果直接使用 ZigBee 协议,当网络规模达到一定程度时,网络中的数据冗余度很高,数据由于对射频信道的争夺,网络的传输效率大大下降,终端节点重发的几率加大,网络时延恶化。

2 数据融合的实现方式

为了解决大规模网络中,数据冗余问题和网络效率降低问题,根据 ZigBee 网络的特点,设计一个针对 ZigBee 协议的数据融合架构。数据融合是一个多级、多层面的数据处理过程,主要完成对多个信息源的数据进行自动检测、关联、相关、估计及组合等的处理。所设计的数据融合架构包含两层:

1)第一层使用统计函数融合算法,在终端节点和路由器之间实现数据融合;

2)第二层使用神经网络算法,在路由器和网关之间实现

收稿日期:2009-02-19;修回日期:2009-03-23。 基金项目:福建省科技厅资助项目(2007F5039)。

作者简介:王平(1955-),男,福建福州人,副教授,主要研究方向:通信网络、无线传感网络;程明传(1982-),男,湖北谷城人,硕士研究生,主要研究方向:无线传感网络、嵌入式软件;翁宗煌(1985-),男,福建泉州人,主要研究方向:神经网络、现场可编程门阵列。

数据融合。

第一层使用的统计函数融合算法,主要实现终端节点和路由节点之间的统计查询功能,路由节点发送给终端节点的 5 种统计基本操作(计数、最大、最小、求和、平均),终端节点能够正确解释和执行。第二层使用神经网络算法,这一层是路由器和网关之间的数据融合。根据一定的准则和决策的可信度做最优决策,以最大限度的数据融合。数据融合架构的核心任务是:降低网络中的数据冗余度,减少数据重发的次数或数据发送失败的次数,以提高网络的运行效率。

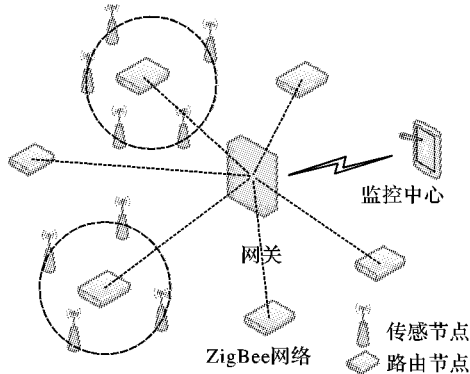


图1 ZigBee网络的星形拓扑结构图

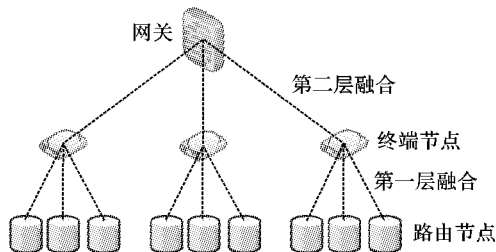


图2 两层数据融合的架构

在 ZigBee 网络中,数据融合起着十分重要的作用,主要体现在降低节点功耗、提高信息收集效率和准确性。在工业自动化和检测领域,为了提高 ZigBee 网络的鲁棒性和准确性,需要大量节点的监测区域相互重叠,这些节点的信息可能非常相近或者相同。针对这种情况,对冗余信息在网内进行数据融合,可以使在满足需求的前提下使需要传输的数据量最小化。减少射频通信的开销对降低节点功耗起重要作用。低成本的 ZigBee 网络由大量廉价的节点组成,从节点得到的信息存在较多不确定因素,比如精度低、通信抗干扰能力弱和瘫痪节点的脏数据等。使用数据融合对监测同一物理量的多个传感器所采集的数据进行数据融合,也能提高信息的精度和可信度。

3 Zigbee 协议应用层实现数据融合

3.1 数据融合的位置

ZigBee 协议符合 OSI 模型^[6],包含物理层、MAC 层、网络层、安全服务和应用程序支持层,在分层上和 TCP/IP 协议有相似之处,整个协议只解决最基本最核心的自组织联网问题。物理层和 MAC 层是 ZigBee 协议射频芯片硬件实现的,所以数据融合只能与网络层和应用层相结合^[7]。进行数据融合涉及到数据融合和应用数据之间的语义间隔问题。在应用层可以直接对应用数据进行数据融合;如果数据融合在网络层实现,由于语义间隔的问题,则需要跨协议层理解应用层数据的含义。数据融合的主要思想是:删除冗余、无效和可信度差的数据,同时将来自不同节点的信息结合起来进行处理,以提高通信效

率。基于应用数据的数据融合,根据应用要求可以实现最大限度的数据压缩。由于数据融合和网络层相结合带来的跨层理解语义的问题给协议栈的实现带来很多困难,故下面主要讨论数据融合如何与应用层相结合,整体框架如图 3 所示。

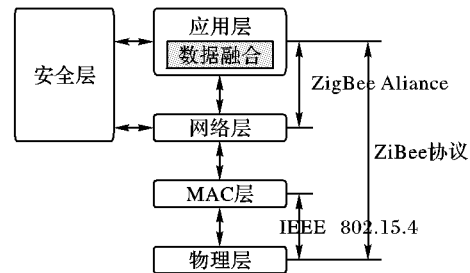


图3 ZigBee协议的框架

3.2 基于统计函数的第一层数据融合算法的实现

终端节点在网络中是最多的,是原始数据最集中的地方,如果能降低终端节点的数据冗余度,对降低整个网络的数据冗余度有重要意义。另外,充分利用本地计算能力,降低数据冗余,可以较少射频通信次数,提高网络的生存时间。于是在终端节点和路由节点之间构成数据融合的第一层,这一层主要通过统计函数实现数据融合。统计函数融合算法只实现最基本的 5 种统计函数: COUNT, MIN, MAX, SUM 和 AVERAGE。COUNT 操作统计元素个数;MIN 和 MAX 操作计算集合中的最小值和最大值;SUM 操作计算集合中所有元素的总和;AVERAGE 操作计算集合中元素的平均值。

在 5 种基本操作中,SUM 和 AVERAGE 操作容易受到节点失效的影响。下面以 AVERAGE 操作的算法实现为例,介绍在部分节点失效的情况如何得到准确的 AVERAGE。假设在 ZigBee 光照强度的监控应用中,反映数据分布结构的参数主要有中位数 TM 、上四分位数 TU 、下四分位数 FD 和四分位数离散度 DS 。设某一区域光照传感器收集的数据从小到大依次为: $V_1, V_2, V_3, V_4, V_5, V_6, V_7, V_8, V_9, V_{10}$, 则中位数为: $TM = (V_5 + V_6)/2$,四分位数离散度为 $DS = TU - FD$ 。疏失误差数据判别区间为:

$$|V_i - T_M| > DS; i = 1, 2, 3, \dots, 10 \quad (1)$$

满足式(1)的测量数据是无效的、不一致的数据,在数据融合时应将其剔除。AVERAGE 操作可以由贝叶斯算法实现。由贝叶斯定理:

$$P(L | V_1 V_2 \dots V_n) = \frac{P(L)P(V_1 V_2 \dots V_n | L)}{P(V_1 V_2 \dots V_n)} \quad (2)$$

由于 n 次对同一物理量的测量是相互独立的,即有:

$$P(V_1 V_2 \dots V_n | L) = \prod_{i=1}^n P(V_i | L) \quad (3)$$

由式(2)和(3)可得:

$$P(L | V_1 V_2 \dots V_n) = \frac{\prod_{i=1}^n P(L | V_i) \prod_{i=1}^n P(V_i)}{[P(L)]^{n-1} P(V_1 V_2 \dots V_n)} \quad (4)$$

在式(4)中, $P(L | V_1 V_2 \dots V_n)$ 和 $\prod_{i=1}^n P(V_i)$ 均与 L 无关,故光照的 n 次检测平均值为:

$$\hat{L}(V) = \frac{\sum_{i=1}^n \text{con}(L_i, V_i) V_i}{\sum_{i=1}^n \text{con}(L_i, V_i)}$$

其中, $\text{con}(L_i, V_i)$ 是随机变量的 L_i 和 V_i 的协方差。一个路由

器最多可以带 254 个终端节点,假设物理环境的实际值为 380,节点数从 2 到 254 变换,仿真数据由实际值和 3 倍高斯随机数相加构成仿真数据,分别求算术均值和贝叶斯均值。仿真结果如图 4、5 所示。

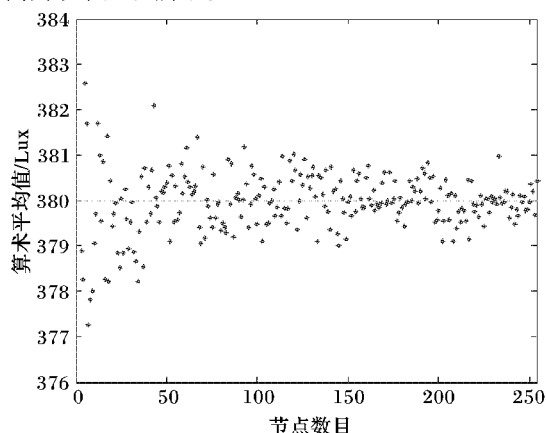


图4 算术平均值和节点数的关系

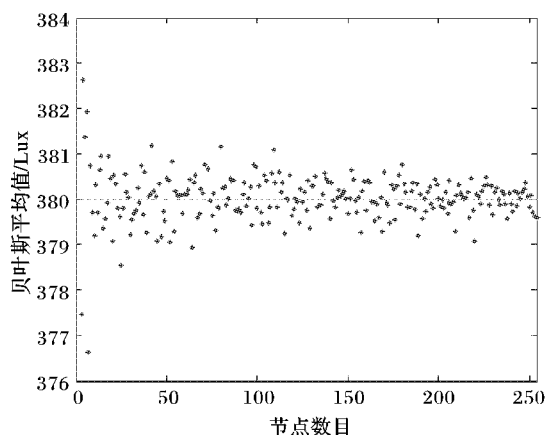


图5 贝叶斯平均值和节点数的关系

从图4可以看出,使用算术平均值的方法得到均值随着节点个数的增加,平均值趋于稳定。从图5可以看出,当节点数目很较少时(<50),贝叶斯均值比算术平均值更接近380,可以体现出贝叶斯均值的优势。当节点数据较多时(>50),两种算法的效果差不多,但是贝叶斯均值有更多的计算量,可能出现网络延时的情况。而现在 ZigBee 网络的实际情况是,路由节点一般带比较少的节点(<20),这样更加适宜采用贝叶斯均值法。采用贝叶斯有利于消除节点失效或其他异常数据对融合结果的影响,在节点较少效果更加明显。其他统计 MAX、MIN、SUM 和 COUNT 实现比较容易,这里不再累述。

3.3 基于神经网络的第二层数据融合算法的实现

对于只关心监测结果和报警数据的场合中,虽然路由器在第一层数据融合中通过本地计算已经去掉了终端节点中的数据冗余,但是在路由器和网关的通信过程中,也存在一定的数据冗余。所以在路由器和网关之间进行第二层数据融合,网关将数据融合到最简,将整个网络的数据压缩到最小。

第二层数据融合如图6所示,在专家数据库中保存了某种应用的经验特征,但对于在数据库中无法查询到的数据,无法给出准确的概率。 N 种不同性质的物理信号由阈值决策器送入数据库,得到一个相应应用的特征概率 O_2 ;原始数据进入具有趋势反馈的神经网络,得到信号的数据拟合特征概率 O_1 。 O_1 与 O_2 同时送到模糊推理融合系统进一步的融合,最终得出事件发生概率 U 。神经网络^[8-9]输入层有 N 个输入量和

网络反馈趋势信号 FP ;中间层选取3个节点;输出层有1个节点,对应该网络的输出的概率 U 。输入层和中间层的权值矩阵为 W_1 ,中间层和输出层的权值矩阵为 W_2 。

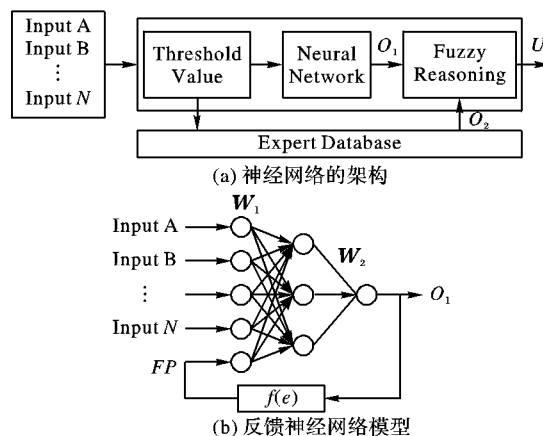


图6 基于神经网络的数据融合

网络模型搭建好后,需要使用已知的数据对网络进行训练,得出网络的权值矩阵。这部分工作可在 Matlab 上计算完成。通过实验采集 100 组数据,其中 90 组数据用于训练神经网络,10 组数据用于测试训练完毕的神经网络。训练结果如图7所示,性能 0.099 115 3,目标是 0.1。可见,经过 16 次训练后,网络误差达到设定的最小值。接下来用其余 10 组数据来检验网络的输出是否满足要求,经测试,完全满足要求。这样,就可以把训练好的神经网络内部的权值和阈值数据提取出来,将其直接用于网关的神经网络的程序中,用于处理专家数据库不能处理的数据,最后得到一个合理的输出。

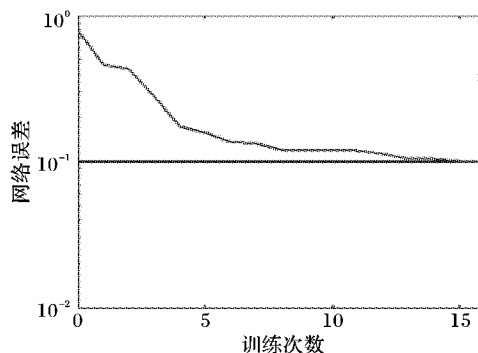


图7 神经网络的训练结果

4 智能温室环境监控实例

以一个实例证实数据融合在 ZigBee 网络中的作用。数据融合的应用层类 SQL 查询接口,主要包括 MIN、MAX、COUNT、SUM 和 AVERAGE5 个基本查询操作。ZigBee 温室环境监控应用中,节点的物理分布和地址分配如图8所示。采用点到点的网络拓扑结构,节点0、1、8、15、22是全功能的,具有路由功能,其他是精简功能节点,不具有路由功能。网内地址的配置参数为:最多子设备数 $Cm = 6$;最大网络参数 $Lm = 2$;最大网络路由 $Rm = 4$ 。节点上配置了温度传感器、湿度传感器和光照传感器。

如果要查询温室 D 区的最小光照度,则可以使用下面的查询请求:

```
SELECT Area, MIN(Light)
FROM Sensors
WHERE Area = D
```

执行这个查询时,汇聚节点发送查询请求,具有路由功能

的节点响应请求。查询到D区后,D区的路由节点即节点15转发请求,收集区域内的光照信息,收集到的数据如图4所示,分别为:(D,81)、(D,78)、(D,81)、(D,75)、(D,83)、(D,76)。路由节点15进行本地数据融合,得到 $\text{MIN}(\text{Light}) = 76$,接着将查询结果(D,76)发送到汇聚节点。从这个简单的例子可以看出,一次查询如果不经数据融合,射频发送的次数为46次;但是本地数据融合以后,射频发送的次数为7次。这将有助于降低功耗,提高网络生存期。通过数据融合,汇集节点通信压力大大减轻。温室环境软件监控画面如图9所示。

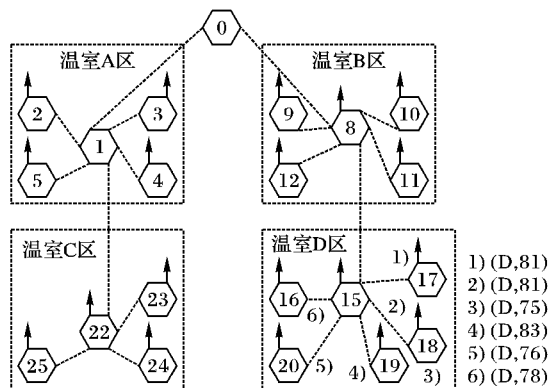


图8 温室环境 ZigBee 监控网络结构

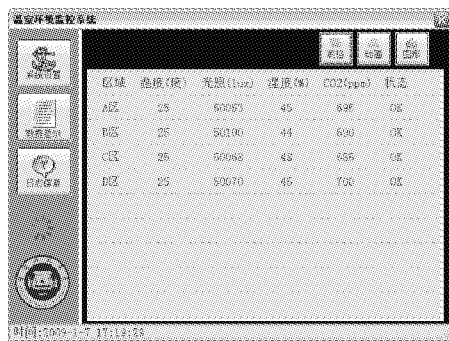


图9 温室环境软件监控画面

数据融合利用节点的分布式计算能力,对数据做了初步处理,去掉了数据冗余,降低了网关的通信压力。如果原始数据直接转发给网关,网关做所有的数据处理,它的性能将很大

程度上影响系统的性能,网关上软件监控软件的设计复杂度也会增加。

5 结语

使用两层的数据融合算法,有利于降低 ZigBee 网络的信息冗余度,网络中信息量的减少对网络的好处是显而易见的,可以减少节点信息重发的次数,节省网络有限的能量,提高信息准确度,延长网络的生存期。但是有一些问题也需要注意,如果数据融合算法过于复杂,会给网络带来一些性能损失,如网络时延,实时性恶化,路由节点计算压力过大而导致网络效率降低等。另外数据融合的应用相关性是比较强的,需要效率和复杂度上做一个权衡,由具体应用决定使用何种数据融合方式。

参考文献:

- [1] 李燕君,王智资.资源受限的无线传感器网络基于衰减信道的决策融合[J].软件学报,2007,18(5):1130-1137.
- [2] 汪祥莉,李腊元,王文波.无线传感器网络中的路由协议研究[J].计算机科学,2008,35(07):50-53.
- [3] COLLETT M A, COX M G, DUTA M, *et al.* The application of self-validation to wireless sensor networks[J]. Measurement Science and Technology, 2008, 19(12):125-201.
- [4] ZHU Y W, ZHONG X X, SHI J F. The design of wireless sensor network system based on Zigbee technology for greenhouse[J]. Journal of Physics, 2006, 48(1):1195-1199.
- [5] 成小良,邓志东.基于 ZigBee 规范构建大规模无线传感器网络[J].通信学报,2008,29(11):158-164.
- [6] 仇国庆,杨志龙,杨莉,等.基于 ZigBee 协议的无线传感器网络节点设计[J].自动化与仪器仪表,2008(3):7-8.
- [7] 雷听,鄢楚平,徐海川.无线传感器网络数据融合技术的研究与仿真[J].计算机工程与设计,2008,29(18):4669-4671.
- [8] MENG JOO ER, CHEN WEILONG, WU SHIQIAN. High-speed face recognition based on discrete cosine transform and RBF neural networks[J]. IEEE Transactions on Neural Networks, 2005, 16(3):679-691.
- [9] 刘玲,王耀南,况菲,等.基于神经网络和遗传算法的机器人路径规划[J].计算机应用研究,2007,24(2):264-265.

(上接第1879页)

在-20 dBm处,由0~4阶5条曲线的带宽在FCC允许的3.1~10.6 GHz范围内。以上分析及仿真可以发现,只要改变参数,总能在孤子脉冲的奇数阶导数中找到符合FCC规定的UWB脉冲信号。因此,在超宽带通信中可以选择电孤子信号的奇数阶导数作为发射脉冲。

4 结语

本文对电孤子脉冲信号用于超宽带发射脉冲进行了理论分析,并在FCC的要求和规范下,在带宽范围和发射功率限制方面做了数值仿真验证。仿真结果表明电孤子脉冲信号序列的时域超短对于频域的超宽,通过调节参数,可以用作超宽带无线通信系统的发射信号。

参考文献:

- [1] 葛利嘉,曾凡鑫,刘郁林,等.超宽带无线通信[M].北京:国防工业出版社,2005.
- [2] SIWIAK K, McKEOWN D. Ultra-Wideband Radio Technology[M]. 张中兆,沙学军,译.北京:电子工业出版社,2005.
- [3] HAM D, LI XIAOFENG, DENENBERG S A, *et al.* Ordered and

chaotic electrical solitons: Communication perspectives[J]. IEEE Communications Magazine, 2006, 44(12):126-135.

- [4] RICKETTS D S, LI XIAOFENG, HAM D. Electrical soliton oscillator[J]. IEEE Transactions on Microwave Theory and Techniques, 2006, 54(1):373-382.
- [5] RICKETTS D S, LI XIAOFENG, SUN NAN, *et al.* On the self-generation of electrical soliton pulses[J]. IEEE Journal of Solid-State Circuit, 2007, 42(8):1657-1668.
- [6] REMOISENET M. Waves called solitons: concepts and experiments[M]. New York: Springer-Verlag, 1999.
- [7] ET Docket 98-153. Revision of Part 15 of the Commission's Rules Regarding Ultra-Wideband Transmission Systems[S], 2002.
- [8] SINGER A C. Signal Processing and communication with solitons[D]. Massachusetts: Massachusetts Institute of Technology, 1996.
- [9] 陆曼琪.超宽带(UWB)无线通信系统的分析与研究[D].上海:上海大学,2005.
- [10] PROAKIS J G. Digital Communications[M]. 张力军,译.4版.北京:电子工业出版社,2003.
- [11] YANG LIUQING, GIANNAKIS G B. Ultra-wideband communications: An idea whose time has come[J]. Signal Processing Magazine, 2004, 21(6):26-54.