

基于 PLE 的有确定解的端到端网络链路时延推测方法

梁永生¹ 邹 粤² 张基宏¹

(深圳信息职业技术学院可视媒体处理与传输深圳市重点实验室 深圳 518029)¹

(深圳大学计算机与软件学院 深圳 518060)²

摘 要 网络时延是重要的网络性能指标,端到端网络时延推测能够克服传统的基于路由器或者路由器协作的网络测量技术的弊端。在网络拓扑已知且稳定和链路性能时空独立性的假设前提下,给出了网络链路时延推测模型,提出了一种基于伪似然估计(PLE)的有确定解的端到端网络链路时延推测方法。在应用期望最大化算法的伪似然估计的基础上,控制背靠背发包方式,确定可以求解的探测单元,解决了不满足有确定解拓扑下的求解问题,且有效降低了计算复杂度。最后利用基于模型的计算验证了该方法的准确性和有效性。

关键词 伪似然估计,期望最大化,网络链路时延,推测方法

中图分类号 TP393 **文献标识码** A

Research on Approach to End-to-end Network Link Delay Inference Based on PLE with Definite Solution

LIANG Yong-sheng¹ ZOU Yue² ZHANG Ji-hong¹

(Shenzhen Key Lab of Visual Media Processing and Transmission, Shenzhen Institute of Information Technology, Shenzhen 518029, China)¹

(College of Computer and Software, Shenzhen University, Shenzhen 518060, China)²

Abstract Network delay is one of the important network performance parameters. End-to-end network delay inference could deal with the difficulties caused by other network measurements based on internal routers or router cooperation. Under the condition of two assumptions, network topology structure is gotten and stable, link performance is temporally and spatially independent, network delay inference model was presented, a new approach to network internal link delay inference based on Pseudo Likelihood Estimation(PLE) with definite solution was proposed in this paper. Based on PLE solved with Expectation Maximum(EM) algorithm, inference units with definite solution were determined via back-to-back packet sending way. This approach could solve the problem of indefinite solution and lower the computation complexity. Experimental study was performed based on model computation. The experimental results show that the approach is accurate and effective.

Keywords Pseudo likelihood estimation(PLE), Expectation maximum(EM), Network link delay, Inference approach

1 引言

网络时延是重要的网络性能指标,其本身有严格的定义,一般被认为是数据从发送节点到接收节点的传输时间。不同的网络配置,采取不同的网络协议,就有不同的网络时延^[1]。网络时延按其发生地点,分为链路时延与节点时延;按其成因,分为传播时延、传输时延、处理时延和排队时延 4 种。对于发生在不同地点的时延,统称为链路时延^[2]。

网络时延的准确掌握和预测,已成为网络性能测量与分析研究领域的重要目标之一。端到端网络链路时延推测就是基于端网络时延测量数据,应用一定的方法和技术,推测出网络内部链路时延的过程。很显然,端到端网络时延推测能够克服传统的基于路由器或者路由器协作的网络测量技术的弊端。

国内外学者应用极大似然估计^[3,4]、基于多播依赖树的 EM 估计^[5]、基于贝叶斯定理^[6]、基于累积生成函数的估计^[7]和基于矩的估计^[8]等算法进行时延推测研究。研究最多的还是应用极大似然估计的方法,但是基于似然函数的时延推测存在以下两个问题。

①根据观察数据得出的似然函数表达式不是解析的,一般采用 EM 算法迭代求解且计算复杂度较高;

②在探测过程中,发包的方式与网络拓扑决定了最后求解的确定性,不当的发包方式可能导致错误的推测结果。

基于上述情况,本文在两个假设——网络拓扑已知且稳定和链路性能时空独立性的前提下,应用已经采集的端到端网络时延数据,提出了一种基于伪似然估计的、有确定解的端到端网络链路时延推测方法,解决了不满足有确定解拓扑下的求解问题,且有效降低了计算复杂度。最后利用基于模型

到稿日期:2010-10-26 返修日期:2011-05-04 本文受国家自然科学基金(60672150),广东省自然科学基金(7008733),深圳市基础研究计划-杰青基金(JC201005280613A)资助。

梁永生(1971—),男,博士,教授,CCF 高级会员,主要研究方向为计算机网络与数据通信、信号处理与模式识别, E-mail: liangys@szit.edu.cn; 邹 粤(1984—),男,硕士生,主要研究方向为 IP 网络关键技术;张基宏(1964—),男,教授,博士生导师,主要研究方向为图像处理与传输。

的计算验证了此方法的准确性和有效性。

2 时延推测模型

时延推测的目标网络可以抽象成树状逻辑网络模型 $T=(V, L)$, 如图 1 所示, V 为节点集, L 为链路集。节点 0 一般为发出探测包的节点, 节点 3、4、5、6 构成接收节点集 V_R , 节点 1、2 构成内部节点集 V_{in} 。假设一共发送了 n 个探测包, 用 $P_{i,j}$ 表示节点 i 到 j 的一条路径, $d_{P_{i,j}}(m)$, $m=1, 2, \dots, n$ 是第 m 个探测包在该路径测得的时延; 用 L_i 表示以节点 i 结尾的链路, $d_{L_i}(m)$, $m=1, 2, \dots, n$ 是第 m 个探测包在该链路的时延。时延推测就是根据已测得端时延向量 D_P , 推测出链路时延向量 D_L 的特性。

$$D_P = A D_L \quad (1)$$

式中, A 是网络拓扑固定的前提下发包路径对应的 0-1 矩阵。时延是网络质量服务中的加性参数, A 中的一行对应一探测路径。对于第 i 个接收节点 r , 有

$$d_{P_{0,r}} = \sum_{l \in P_{0,r}} d_l \quad (2)$$

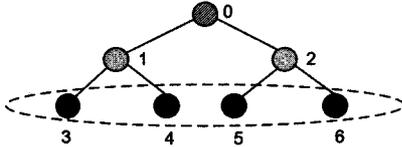


图 1 树状逻辑网络

3 PLE 推测方法

3.1 PLE 原理

端到端时延推测, 本质上是一统计逆问题。在实际应用中, 目标路由矩阵 A 通常不是满秩的, 无法直接求解 D_L , 即由接收节点收集到的探测包的路径时延无法推算出该探测包在构成这一路径的链路上的时延。所以一般采用统计推断的方法, 推测出链路时延的统计特性, 如概率分布。

无论网络链路采用何种模型, 时延推测的过程都类似。

首先确立推测参数 θ , 并将端测量路径时延表示成链路时延的函数。

$$Y_{\text{path}} = f(X_{\text{link}}(\theta)) \quad (3)$$

通常地, 如果 X_{link} 指链路时延本身, 则 $f(\cdot)$ 被定义成一累加函数; 如果 X_{link} 是链路时延的某种分布, 则 $f(\cdot)$ 常被定义成一联合分布函数。

然后将路径测量数据的统计特性表示成 X_{link} 的似然函数, 即

$$L = P(Y_{\text{path}}) = P(f(X_{\text{link}}(\theta))) \quad (4)$$

似然函数的构造一般采用两种条件概率: 一种是父链路时延与子链路时延的条件概率, 一种是路径时延与组成该路径的链路时延的条件概率。

最后, 最大化似然函数, 得到目标参数, 即取 θ , 使得

$$\frac{\partial L}{\partial \theta} = 0 \quad (5)$$

一般 L 都是非解析的, 直接求其偏导数非常困难, 一般采用期望最大算法 (Expectation Maximum, EM) 迭代求解。

现有求解逆问题的方法都是通过将原始似然极大化问题变换为伪似然估计问题, 来降低计算复杂度。伪似然估计的主要思想是将原来复杂的整体求解过程分解为一系列较为简单的子过程。具体做法是每次取路由矩阵的两行, 构造一子

问题, 视每个子问题的解为边缘分布, 所有子问题的解构成整体分布。设路由矩阵为 $A^{[X]}$, 子问题集 $S = \{S = (i_1, i_2) : 1 \leq i_1 < i_2 \leq I\}$ 。对每一子问题 $s \in S$, 有

$$Y^s = A^s X^s \quad (6)$$

这样, 原问题的似然函数如式 (7) 所示

$$L^p(y_1, \dots, y_n; \theta) = \prod_{i=1}^n \sum_{s \in S} L^s(y_i; \theta) \quad (7)$$

通过最大化 L^p 得到参数 θ 的最大似然估计, 但直接通过 $(L^p; \theta)' = 0$ 求得 θ 仍是非常困难的, 这里采用 EM 算法来迭代求解。先假设 X^s 是可以观察到的完整数据, 且当前参数为 $\theta^{(k)}$, 则目标函数 Q 在此基础上第 $k+1$ 次的取值为

$$Q(\theta, \theta^{(k)}) = \sum_{i=1}^n \sum_{s \in S} E_{\theta^{(k)}} (L^s(x_i^s; \theta^s | y_i^s)) \quad (8)$$

通过最大化 Q 可以得到相应的 θ 的最大似然估计。

3.2 PLE 推测过程

假设推测过程中链路 L_i 可能出现的最大时延为 $d_{\max}(i)$, 根据精度 q 的要求, 令 $K_i = d_{\max}(i)/q$, 则链路时延的可能取值有

$$d_{L_i}(m) = \{kq | k=0, 1, 2, \dots, K_i\}, m=1, 2, \dots, n \quad (9)$$

式中, $k=0$ 表示该包在节点 i 上遇到的转发队列长度为 0, 且这里不考虑丢包的情况, 即探测过程中如果发现丢包, 则该次探测无效。对于 d_{L_i} 的每一取值, 有

$$P_{L_i}(k) = P(d_{L_i} = kq) \quad (10)$$

式中, $P_{L_i}(k)$ 表示链路 i 上时延为 kq 的概率。假设包时延是时序上独立的, 且路径上的链路时延也相互独立, 由此问题转化为由边接收节点集的路径时延 D_P 的分布, 从而推断出内部链路时延 D_L 的分布。

假设探测过程中链路 i 上经过 n_i 个探测包, 且时延为 kq 的包数为 $a_i(k)$, 则

$$P_{L_i}(k) = \frac{a_i(k)}{n_i} \quad (11)$$

实际上, n_i 可由选取探测单元直接观测得到, 而 $a_i(k)$ 是无法观测得到的。在所有链路的时延分布 P_L 已知的情况下, $a_i(k)$ 会有一最大期望值。设在一背靠背探测单元 $u \in U$, 且链路 i 包含在其探测路径中, 即 $L_i \in L(u)$, rx 和 ry 为 u 的两个接收节点, 某次探测得到的时延二元组 $D_u(m) = (d_{P_{0,rx}}(m), d_{P_{0,ry}}(m))$, 则

$$a_i(k) = \sum_{m=1}^{n_i} \sum_{u \in U, L_i \in L(u)} P(d_{L_i}(m) = kq | D_u(m) = (d_{P_{0,rx}}(m), d_{P_{0,ry}}(m))) \quad (12)$$

其中, 每一个探测单元的两个时延构成一种输出模式。设该模式在探测过程中出现的次数为 $C(D_u)$, 则式 (12) 改写为

$$a_i(k) = \sum_{u \in U, L_i \in L(u)} C(D_u) P(d_{L_i} = kq | D_u) \quad (13)$$

式中, $C(D_u)$ 可由观测数据统计得出。对于某一探测单元的 $P(d_{L_i} = kq | D_u)$, 有

$$P(d_{L_i} = kq | D_u) = \frac{P(D_u | d_{L_i} = kq) P_{L_i}(k)}{P(D_u)} \quad (14)$$

探测单元 u 中的 $P(D_u | d_{L_i} = kq)$ 可采用如下公式递推。对于以节点 i 结尾的链路 L_i , 设其父节点 $f(i)$ 到 rx, ry 的时延分别为 $d_{P_{f(i),rx}}, d_{P_{f(i),ry}}$, 则

$$\begin{aligned} P(D_u | d_{L_i} = kq) &= P(D_{L_i} = kq) P(d_{P_{f(i),rx}} - kq) P(d_{P_{f(i),ry}} - kq) \\ &= P_{L_i}(k) P(D_{i,k,x}) P(D_{i,k,y}) \end{aligned} \quad (15)$$

对于 $P(D_{i,k,x})$, 设 i 的直接子节点为 $d(i)$, 则

$$P(D_{i,k,x}) = \sum_{k=0}^{K_{d(i)}} P_{L_{d(i)}}(k) P(d_{P_{i,rz}} - kq)$$

$$= \sum_{k=0}^{K_{d(i)}} P_{L_{d(i)}}(k) P(D_{d(i),k,x}) \quad (16)$$

$P(D_{i,k,y})$ 同理求得。当 i 为接收节点时,显然有 $P(D_{i,k,x}) = P_{L_i}(k), i \in V_R$ (17)

计算时,可先随机选择一初始参数集 P_L ,即可通过式(13)迭代求解,直到某次更新 $|P_L^{\text{update}} - P_L^{\text{best}}|$ 小于设定值,则迭代停止。

4 有确定解的探测方案

从路由矩阵中选取两行或更多来构成子问题虽然可行,但由此会带来计算复杂度的增长,极限情况下取整个路由矩阵就可还原成原来的整体问题;而且一次引入路由矩阵中的更多行,子问题间的相关性就更难忽略。因此需要划分有确定解的探测方案。

数据包的网络传送方式一般有单播与多播两种。在单播情况下,每次发包可得一条探测路径;在多播情况下,每次发包可得 $|V_R|$ 条探测路径。所以多播的探测效率较单播高。但由于经济和技术上的原因,IP 多播在实际网络中的支持有限。

探测方案的选取要满足有确定解的充要条件^[9]:

- ① 每一内部节点至少是一种 $k(k > 1)$ 播探测单元的分支点;
- ② 接收节点必须至少被一个探测单元覆盖。

为此,文献[10]提出一种发送背靠背探测包的方式来模拟多播。由于时空相关性,两数据包在经过的共同链路上有相同的时延,从而克服了单播探测效率低下、多播难于实现的缺点。

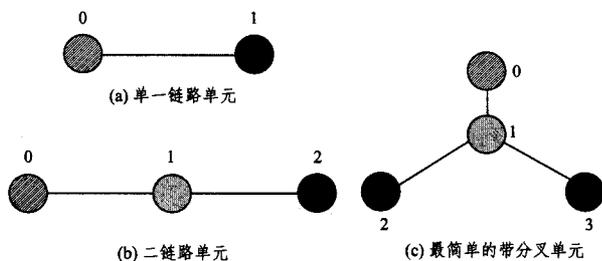


图 2 3种探测单元

探测方案有确定解的充要条件为:当且仅当组成该探测方案的探测单元都是有确定解的单元。本文采用背靠背探测的方式采集端时延数据。图 2 中(a)、(b)、(c)分别为 3 种探测单元,图 2(a)由单一链路构成,这是满足树状网络模型的最小形式。图 2(b)是常见的二链路单元。图 2(c)是最简单的有分叉网络。从非线性方程组求解的角度,可以确定单元图 2(a)和图 2(c)是有确定解的最简单形式。

对于图 2(b)来说,可以考虑首先推测链路 1 的时延 $D_{L(1)}$,然后推测包含节点 1,2 的路径的时延 $D_{P(1,2)}$,则链路 2 的时延 $D_{L(2)}$ 可通过下式求得:

$$D_{L(2)} = D_{P(1,2)} (1/*) D_{L(1)} \quad (18)$$

式中, $(1/*)$ 为反卷积运算。

因此,对于不同的推测网络,在链路时空相关性的假设前提下,可以通过划分上述 3 种探测单元进行求解。

5 基于模型的计算及结果分析

所谓基于模型的计算,即是设定链路时延离散分布的参

数,按照赌轮算法产生对应的时延分布,保存端到端的数据,并用于后续的链路时延分布推测。

本文采用图 3 所示网络进行基于模型的计算。在图 1 中的节点 0 上添加一父节点 f 作为发包节点,节点 0 变成一内部节点。探测单元分别选择 $\langle 3,4 \rangle$ 与 $\langle 5,6 \rangle$,链路 P_{L_0} 的分布可直接推测得到。链路时延参数随机产生,由赌轮算法产生目标时延分布,共发送 5000 个探测包,每个探测包随机选取某一探测单元的目标节点进行发送。设定总体误差 $|P_L^{\text{update}} - P_L^{\text{best}}| < 10^{-5}$ 时迭代停止。由于该网络拓扑不满足有确定解的条件,因此先选取图 2(c)所示的探测单元,抽取出链路 P_{L_1} 与 P_{L_2} ,其中 P_{L_2} 的推测结果如图 4(a)所示。对于图 2(b)所示的探测单元中的链路,如 P_{L_3} ,反卷积后得到推测结果如图 4(b)所示。

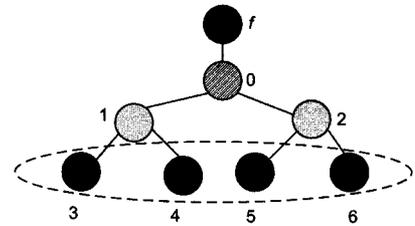


图 3 待推测网络拓扑

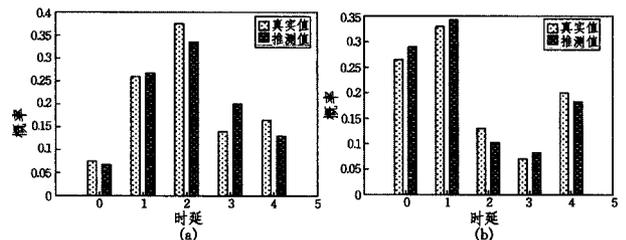


图 4 链路时延推测结果

由图 4 可知,推测值与真实值在不同的 k 值时接近,由此说明本文提出的推测方法是可行的。

在算法复杂度上,图 5 是令平均采样 $|\bar{K}|$ 每次增加 10 时测得的算法运行时间。从图中可知, $|\bar{K}|$ 的线性增长导致本文算法复杂度接近 3 次方增长,远低于指数级复杂度。与经典的基于 CGF 的网络链路时延推测算法比较,本文算法的计算复杂度为 $O(|U| |\bar{K}|^3)$,其中 $|U|$ 为探测单元链路数,且 $\max(|U|) = 3$,与 CGF 算法的复杂度 $O(mn^2) + O(n^3)$ 相当,其中 m 为链路数, n 为路径数。但基于 CGF 的方法在推算过程中需要构造满秩的路由矩阵,使原来的内部节点成为准端节点,有违端到端网络测量的初衷。

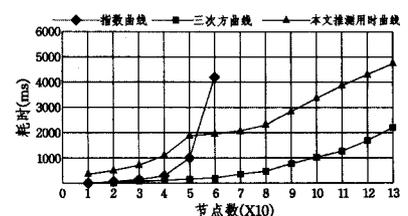


图 5 推测算法的时间复杂度曲线

结束语 本文在网络拓扑已知且稳定和链路性能时空独立性的假设前提下,给出了网络链路时延推测模型,提出了一种基于伪似然估计(PLE)的有确定解的端到端网络链路时延推测方法。应用期望最大化算法,控制背靠背发包方式,划分可以求解的探测单元,解决了不满足有确定解拓扑下的求解问题,且通过应用 PLE 和划分有确定解的探测单元有效降低

了计算复杂度。最后利用基于模型的计算验证了本方法的准确性和有效性。在复杂度相当的前提下,本文算法克服了经典的基于 CGF 的方法在推算过程中需要构造满秩路由矩阵的缺点。

参考文献

- [1] 梁永生,张基宏,张乃通. IEEE 标准容限内以太网转发时延的测试与分析[J]. 电子学报,2008,36(1):46-50
- [2] 胡治国,张大陆,侯翠平,等. 基于随机神经网络多步网络时延预测模型[J]. 计算机科学,2009,36(7):85-87,112
- [3] Liang Gang, Yu Bin. Maximum Pseudo Likelihood Estimation in Network Tomography[J]. IEEE Transactions on Signal Processing,2003,8(51):2043-2053
- [4] 李贵山,蔡皖东. 网络链路时延分布估计方法研究[J]. 计算机工程与应用,2009,45(8):20-22,28
- [5] Yolanda T, Mark C, Robert N D. Network Delay Tomography [J]. IEEE Transactions on Signal Processing, 2003, 8 (51): 2125-2135
- [6] Guo Dong, Wang Xiao-dong. Bayesian Inference of Network Loss and Delay Characteristics with Applications to TCP Performance Prediction[J]. IEEE Transactions on Signal Processing, 2003,8(51):2205-2218
- [7] Shih M-F, Alfred H O. Unicast Inference of Network Link Delay Distributions from Edge Measurements [C]//Proceedings of the IEEE International Conference on Acoust Speech and Signal Processing. Salt Lake City, U. S. A, May, 2001, 6: 3421-3424
- [8] Lawrence E, Michailidis G, Vijayan N N. Fast, Moment-based Estimation Methods for Delay Network Tomography[J]. IEEE Transactions on Signal Processing, 2008(9):1-22
- [9] Lawrence E, Michailidis G, Vijayan N N. Network Delay Tomography Using Flexicast Experiments[J]. Journal of the Royal Statistical Society, Series B, 2006(68):785-813
- [10] Mark C, Robert N D. Network Tomography for Internal Delay Estimation [C]//Proceedings of the IEEE International Conference on Acoust Speech and Signal Processing. Salt Lake City, U. S. A, May 2001, 6: 3409-3412
- [11] 曾鹏,于海斌. 工业无线网络 WIA 标准体系与关键技术[J]. 自动化博览,2009(1):24-27
- [12] 陆晓希. 低功耗自适应无线传感网络路由设计[J]. 微计算机信息,2007,23(2-1):172-174
- [13] 姜波,张荣福. 基于节能的无线传感器网络路由协议研究[J]. 现代电子技术,2009(1):26-29
- [14] Kulik J, Heinzelman W R, Balakrishnan H. Negotiation-based protocols for disseminating information in wireless sensor networks[J]. Wireless Networks, 2002, 8(223):169-185
- [15] Intanagonwivat C, Govindan R, Estrin D, et al. Directed diffusion for wireless sensor networking[J]. IEEE/ACM Transactions on Networking, 2003, 11(1):2-16
- [16] Braginsky D, Estrin D. Rumor routing algorithm for sensor networks[C]// Proceedings of the 1st Workshop on Sensor Networks and Applications. 2002:22-31
- [17] Sohrabi K, Gao J, Ailawadhi V. Protocols for Self-organization of a Wireless Sensor Network[J]. IEEE Personal Communications, 2000, 7(5):16-27
- [18] Heinzelman W R, Chandrakasan A P, Balakrishnan H. An application-specific protocol architecture for wireless micro sensor networks[J]. IEEE Transactions on Wireless Communications, 2002, 1(4):660-670
- [19] Manjeshwar A, Grawal D P. TEEN: A Routing Protocol for Enhanced Efficiency in Wireless Sensor Networks[C]// Proceedings of the 15th International Parallel and Distributed Processing Symposium. San Francisco, USA; IEEE Computer Society, 2001:2009-2015
- [20] Lindsey S, Raghavendra C S. PEGASIS: Power-efficient gathering in sensor information systems [C] // Proceedings of the IEEE Aerospace Conference. 2002:1125-1130
- [21] Gupta I, Riordan D, Sampalli S. Cluster-head election using fuzzy logic for wireless sensor networks[C]// Proceedings of the 3rd Annual Communication Networks and Services Research Conference. Halifax, IEEE Computer Society, 2005:255-260
- [22] Heinzelman W, Chandrakasan A, Balakrishnan H. Energy efficient Communication Protocols for Wireless Sensor Networks [A]//IEEE Proceedings of the Hawaii International Conference System Sciences'00[C]. 2000:3005-3014

(上接第 102 页)

的要求,充分利用工业环境中部分节点有线供电的优势,减少关键位置节点因大量进行数据转发使网络出现能量黑洞的可能性,从而延长无线网络整体生存时间。

仿真结果证明,基于 VCR 的路由算法通过结合最小跳数路由算法与节能型路由算法的优点,针对 WIA-PA 网络中不同的 VCR 应用不同的路由算法,实现管理报文转发的延迟低、数据报文转发能耗低,符合工业无线应用的要求。进一步的研究将重点评价算法的收敛速度问题,以便适应大规模的工业无线网络应用。

参考文献

- [1] 曾鹏,于海斌. 工业无线网络 WIA 标准体系与关键技术[J]. 自动化博览,2009(1):24-27
- [2] 王太峰,林珂,范乐昊. IEEE802. 15. 4 标准的无线传感器网络自组网方案[J]. 信息技术,2009(2):30-35
- [3] Freescale semi-conductor [S]. 802. 15. 4 MAC/PHY software user's guide,2005
- [4] Freescale semi-conductor [S]. 802. 15. 4 MAC/PHY software reference manual,2005
- [5] Zheng Jiang-liang, Lee M J. Will IEEE802. 15. 4 Make ubiquitous networking a reality: A discussion on a potential low power, Low bit rate[J]. IEEE Communication Mag, 2004, 42(6):140-146
- [6] 梁炜,李强. 无线 HART 网络管理器的开发与应用[J]. 中国仪器仪表,2008(S1):153-156
- [7] 王俊. HART 通讯协议简介[J]. 气体分离,2006,(6):35-36
- [8] 王华,刘枫,杨颂华. 工业无线网络 WIA-PA 网络研究与设计[J]. 自动化与仪表,2009(7):17-21
- [9] 张丹,刘枫,杨颂华. WIA-PA 中自适应心跳协议的研究与设计[J]. 控制系统,2009(8):1-4
- [10] 冀翔宇,冯冬芹. 低功耗工业无线传感器网络的设计与实现[J]. 化工自动化及仪表,2008,35(4):47-50
- [11] 王磊,李增荣. 低功耗无线传感器网络节点的设计[J]. 电脑知识与技术,2009,5(1):74-76