

# 传感器网络通信延迟的实时预测

陈建荣, 朱培栋, 齐治昌

(国防科技大学计算机学院, 长沙 410073)

**摘要:** 针对使用 CSMA/CA 通信机制的无线传感器网络提出一种通信延迟估计算法。利用链路质量具有时间和空间相关性的特点建立组合的链路模型, 预测链路可靠性, 使用简化的冲突概率模型预测信道竞争延迟, 结合预测的链路可靠性得到通信延迟。分析了算法的开销, 在 ns2 模拟器上对信道竞争延迟预测算法进行了模拟, 结果表明预测算法能够比较准确地预测网络的竞争延迟。

**关键词:** 传感器网络; 通信延迟; 链路可靠性; 竞争延迟

## Real-time Estimation for Sensor Network Communication Delay

CHEN Jian-rong, ZHU Pei-dong, QI Zhi-chang

(College of Computer Science, National University of Defense Technology, Changsha 410073)

**【Abstract】** This paper proposes an algorithm that estimates the communication delay in sensor networks with CSMA/CA communication mechanism. It estimates the link probability with a model based on temporal correlation and spatial correlation of link quality, and estimates the contention delay using a simple collision probability model. The simulation of the algorithm on ns2 simulator shows that the contention delay estimation of the algorithm is accurate and the time cost is always low regardless of network contention.

**【Key words】** sensor network; communication delay; link probability; contention delay

### 1 概述

对于实时无线传感器网络, 单跳通信延迟是进行实时路由转发决策的基础。但是传感器网络的通信延迟很难计算或者测量, 原因主要包括: (1)通信链路不稳定, 无线信号随距离增加而衰减且容易受到干扰, 成功发送一次数据需要尝试多次重复传输, 增加了不确定的通信延迟; (2)共享信道的通信方式经常会采用随机的退避机制如 CSMA/CA, 使得抢占信道控制权的时间开销难以确定; (3)监控事件的发生导致网络流量突增, 使用消息交换来测量通信延迟的方法会进一步增加网络拥塞, 导致测量的时间开销增加; (4)传感器节点具有很小的内存和很低的计算速度, 节点上运行的算法不能太复杂。

### 2 相关工作

SPEED协议<sup>[1]</sup>是第一个面向传感器网络提出的实时路由协议。它的工作方式是根据转发节点的位置和链路间的通信延迟来选择下一跳的节点。利用数据报文和ACK中的时间戳来确定当前链路的通信时间, 通过EWMA模型平滑得到通信延迟。这种通信延迟的评测方法没有区分竞争延迟和重传延迟, 所以只是对于数据稀少的网络比较有效。在网络通信较为繁忙的情况下, 测量的误差较大。而且通信延迟的评估必须经过消息交换, 对于链路具有突变的情况, 很难快速有效。

在链路可靠性评估方面的工作有很多。在传统的 Internet 网络中, 可用资源比较丰富, 通常的链路评估都是使用 IGRP 协议发送探针报文, 通过收到报文和期待报文数的比值得到链路质量。由于报文传输路径固定以及数量较大, 因此这种方式不适合传感器网络。文献[2-3]研究了传感器网络的链路质量和行为, 利用无线信道的广播特性, 通过监听或者偷听

特定节点的报文来评估链路质量。定义链路质量为  $t$  时间内收到的报文和  $\max(t)$  时间内期待收到的消息的比值。这种链路可靠性的评估需要进行一定数量的消息交换, 能量和时间开销较大。

在竞争延迟评估方面的方面, 文献[4]把 IEEE802.15.4 协议的 CSMA/CA 机制建模成马尔科夫离散事件模型, 分析有限非饱和流节点在有限队列长度下处理消息的服务时间, 这里的服务时间包括竞争延迟和传输延迟。但是这种计算方法过于复杂, 并且很难精确得到公式中需要的参数。

### 3 通信延迟评估器设计

#### 3.1 网络基础和假设

无线传感器网络的工作前提是底层通信协议是基于 CSMA/CA 机制的 MAC 协议能够提供包括发送每个报文尝试的次数和网络的忙闲状态等参数; 另外每个节点的位置固定。

#### 3.2 链路评估器

研究表明 2 个节点之间的链路质量随时间和空间变化具有一定的统计规律<sup>[2]</sup>。时间相关性的链路评估模型将比较有效地预测节点之间的链路状态, 但是需要周期性交换消息。当网络发生突发事件, 链路质量会有剧烈的变化, 例如移动目标对节点通信的阻碍、或者事件导致节点移动甚至失败等。在这种情况下, 基于时间相关性的链路评估方法将失效。基于空间的链路评估方法认为地理上分布很近的节点具有链路相关性, 当移动目标对节点通信产生阻碍导致链路失败, 和

**基金项目:** 国家“863”计划基金资助项目(2006AA01z213); 国家自然科学基金资助项目(60673169)

**作者简介:** 陈建荣(1977 -), 男, 博士研究生, 主研方向: 传感器网络; 朱培栋, 教授; 齐治昌, 教授、博士生导师

**收稿日期:** 2007-11-30 **E-mail:** chenjianrong301@163.com

这条链路地理很近的节点也都会受到影响。使用空间相关性模型可以从当前的链路情况预测其他相邻节点的情况，而不需要进行新的信息交换。但是这种方式计算量较大，需要保存较多的状态。组合这2种方法，节点使用较长的周期交换消息，并通过时间相关性模型来更新链路信息，维护空间相关性模型的系数，当网络流量变大，需要更新链路状态，就使用空间相关性模型，只需要少量节点的链路信息即可。时间相关性模型使用Window Mean with Exponentially Weighed Moving Average(WMEWMA)<sup>[5]</sup>，空间相关性模型使用简单的衰减模型。

WMEWMA 作为有效的链路评估模型，是通过消息交换得到链路可靠性，并作为下一次计算之间的预测值。在 WMEWMA 中每个节点维持邻居节点的一组链路历史状态， $C$  为当前节点，链路信息可表示为  $P = (p_1, p_2, \dots, p_i, \dots)$ ， $i \in neighbor(C)$ 。节点周期性维护邻居信息时，伴随维护链路信息，每个邻居节点尝试发送的消息次数为  $Num_{try}$ ，成功发送的消息数量为  $Num_{succ}$ ，那么节点和邻居之间的当前链路状态将更新为

$$P_i' = P_i \cdot \alpha + (1 - \alpha) \cdot Num_{succ} / Num_{try} \quad (1)$$

其中， $\alpha (0 < \alpha < 1)$  为衰减系数，根据链路质量评估精度的需求取值， $\alpha$  越小链路评估敏感性越强。

衰退模型是一种用来进行模式识别、分类和预测的典型技术，建立一组输入和输出的函数，在这里输入是节点的地理位置，输出就是节点链路状态。衰退模型可以使用当前节点的链路状态来预测其他节点的链路状态。对于一组链路状态  $P = (P_1, P_2, \dots, P_n)$  和节点的位置  $\langle x, y \rangle$ ，建立衰退模型为

$$P = \beta_0 + \beta_1 x + \beta_2 x^2 + \beta_3 xy + \beta_4 y + \beta_5 y^2 \quad (2)$$

其中， $\beta$  为衰退系数，可以在链路维护中通过如下公式得到：

$$\beta = (\Phi(s)^T \Phi(s))^{-1} \Phi(s)^T P = (\Phi(x, y)^T \Phi(x, y))^{-1} \Phi(x, y)^T P \quad (3)$$

$$\text{其中，} \Phi(s) = \Phi(x, y) = \begin{pmatrix} 1, x_1, x_1^2, x_1 y_1, y_1, y_1^2 \\ 1, x_2, x_2^2, x_2 y_2, y_2, y_2^2 \\ \dots \\ 1, x_n, x_n^2, x_n y_n, y_n, y_n^2 \end{pmatrix}$$

衰退系数的计算量相对于传感器网络资源有限的节点来说，计算量过大。但是邻居节点不轻易失败，那么  $\Phi(s)$  确定，可以得到  $\beta = MP$ ， $M$  计算一次可使用多次。算法额外的空间开销  $M$  矩阵。根据节点的 K-coverage 覆盖原则，传感器节点邻居的数目控制在很小的范围之内。假定有  $n$  个邻居并且每个数据保存需要 2 Byte。 $M$  矩阵的空间开销就为  $10n$  Byte 的存储空间。其他链路状态和位置信息在任何情况下都需要保持。衰退系数更新以后需要进一步得到其他节点的链路估计，分别需要  $10n$  次乘法和  $10n$  次加法。对于特定节点，计算的时间开销固定。当消息交换的代价较大时使用空间相关性模型评估链路 ( $Delay_{con} + Delay_{tran} > \bar{T}$ ， $\bar{T}$  为预设的值，也可以根据预测时间要求或者预测精度要求得到)。

### 3.3 信道竞争时间评估器

得到链路可靠性可以预测发送一次消息需要的重传次数，但是影响消息通信延迟的不确定因素还包括延迟大小与底层通信协议的参数如竞争窗口、退避次数等以及参与竞争的节点个数。文献[6]提出的 Kamuler 滤波方案，是针对 IEEE 802.11 协议在饱和流情况下评估竞争节点的方法。通过监控信道得到条件碰撞概率  $p$ ，利用条件碰撞概率  $p$  和发送

饱和数据流的节点数量之间的函数关系，可以得到竞争终端的个数。

采用相似的方法通过对信道监控得到发送消息的条件碰撞概率，使用 Auto Regressive Moving Average(ARMA)滤波算法来计算运行时的条件碰撞概率  $p$ 。通过条件碰撞概率可以得到发送一次消息需要退避的次数。设定成功发送概率阈值为  $P_{resh}$ ，当发送概率超过这个值将认为发送成功。节点发送成功需要进行退避的次数为  $k = \lceil \log_p(1 - P_{thresh}) \rceil$ ，最差情况下一跳的竞争延迟就为

$$Delay_{con}(P, P_{resh}) = [k / MaxBackoff] \sum_{i=0}^{MaxBackoff} \frac{2^{\max(i, maxBE)} - 1}{2} + \sum_{i=0}^{k \% MaxBackoff} (2^{\max(i, maxBE)} - 1) / 2 \quad (4)$$

其中， $MaxBackoff$  为 CSMA/CA 退避机制最大的退避次数； $maxBE$  是最大的退避窗口指数；% 表示取模运算。

为了得到条件发送概率，节点需要对信道进行持续监控，能量消耗很大。认为存在一些应用对通信延迟的需求重于节省能源，如对紧急情况的监控，如果时间需求不能得到满足，网络更长时间的工作寿命是没有意义的。另外，可以使用低功率监听的新技术，通信模块使用反向否定的方法判断信道是否空闲，可以大大减少能量的使用。同时，算法也可以在无数据发送时停止工作，节点可以进行睡眠/工作调度，减少空监听带来的能源浪费。

### 3.4 通信延迟评估

节点  $i$  到节点  $j$  消息的通信延迟  $Delay_{ij}$  可以分为竞争延迟  $Delay_{con}$ 、传输延迟  $Delay_{tran}$  和重传延迟  $Delay_{retr}$ 。传输延迟可以根据发送消息的长度和节点通信模块的传输速率决定 ( $Delay_{tran} = L_{packet} / V$ )。标记  $p_{ij}$  为节点  $i$  到节点  $j$  的链路可靠性，成功传输一次消息需要重传的次数将是  $1/p_{ij}$ 。竞争延迟通过式(4)来得到。那么通信延迟预测为

$$Delay_{ij} = (Delay_{con} + L_{packet} / V) / p_{ij} \quad (5)$$

## 4 模拟评估

为了验证本文的预测模型，在 ns2 模拟器上进行 2 个模拟实验。分别对信道竞争延迟评估和链路可靠性组合模型进行验证。模拟的底层通信协议为 IEEE 802.11 协议，信道传输速率为 2 Mb/s，报文大小统一为 200 Byte，不考虑到链路失败的情况，2 个节点之间以每秒 200 个报文的速度发送数据，每 5 s 增加一对竞争节点，竞争节点的数量从 2 个逐渐增加到 6 个，竞争节点的发送速度也为每秒 200 个报文。发送成功概率  $P_{resh}$  为 0.95，图 1 为实际的竞争延迟和预测的延迟之间的关系。预测延迟基本上能够反应网络的竞争情况，具有较强的灵敏性。

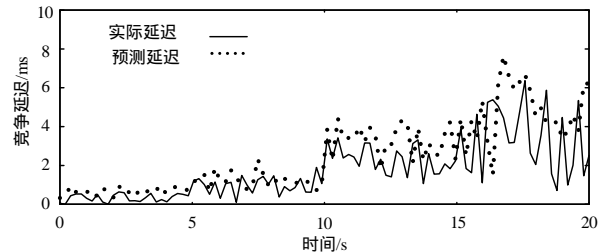


图 1 竞争延迟的评估

图 1 表明，在竞争较多的情况下，竞争延迟预测相对误差较小，而在竞争比较少少的情况下，这种评估相对误差较大。

对于有实时需求的网络，比实际大的估计不会丢失性能。而且因为竞争较少时，误差对竞争延迟的绝对值没有太大的影响，所以评估整体上比较有效。

然后进行链路可靠性的评估。WMEWMA 和关于空间链路相关性模型评估精度在以前的文献中都有所研究，这里只评估组合模型的时间开销。模拟场景如前面竞争延迟设定相同。在  $Delay_{con} + Delay_{tran} > 2$  ms 时，启用空间相关性模型，反之，使用 WMEWMA 模拟进行评估。WMEWMA 中衰减系数为 0.5，邻居节点数量保持在 10 个不变，图 2 所示是单独使用 WMEWMA、空间链路相关性模型(Regression)以及组合模型(Combined)的时间开销。从图 2 中可看出，WMEWMA 模型的时间开销随网络竞争程度的增加而增大，而使用综合模型可以保持比较稳定的时间开销。空间的衰退模型也能保持较好的时间开销，但其精度不高。通信开销的评估是计算所得，这里也不再进行模拟。

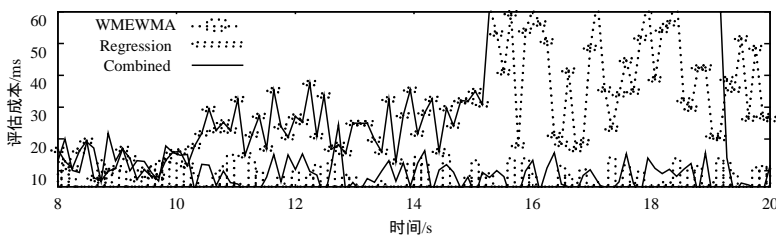


图 2 链路评估的时间开销

## 5 结束语

本文提出通信延迟的预测方法，对竞争延迟和重传延迟分别评估。利用链路的时间和空间相关性模型预测链路质量，使得预测更加迅速；利用信道忙闲状态的监控预测竞争延迟。以网络忙闲情况作为链路组合模型的调控机制，可以控制进

(上接第 14 页)

验，基于 SVM 算法的文本分类方法把标题和正文进行了分类特征词的计算，分类特征词选定使用的是目前较为流行的最大熵方法，两者的实验结果如表 3、表 4 所示。

表 3 基于 SVM 算法的文本分类结果

| 具体类别 | 网页数目 | 正确分类网页 | 准确率/(%) |
|------|------|--------|---------|
| 政治   | 216  | 187    | 86.6    |
| 财经   | 225  | 194    | 86.2    |
| 娱乐   | 305  | 286    | 93.8    |
| 军事   | 159  | 127    | 80.0    |
| 体育   | 247  | 209    | 84.6    |
| 教育   | 168  | 148    | 88.1    |
| 饮食   | 92   | 78     | 84.8    |
| 卫生   | 146  | 117    | 80.1    |
| 环保   | 96   | 79     | 82.3    |
| 法治   | 138  | 106    | 76.8    |

表 4 基于领域词语激活的文本分类结果

| 具体类别 | 网页数目 | 正确分类网页 | 准确率/(%) |
|------|------|--------|---------|
| 政治   | 216  | 209    | 96.8    |
| 财经   | 225  | 216    | 96.0    |
| 娱乐   | 305  | 298    | 97.7    |
| 军事   | 159  | 147    | 92.5    |
| 体育   | 247  | 247    | 100.0   |
| 教育   | 168  | 150    | 89.3    |
| 饮食   | 92   | 72     | 78.3    |
| 卫生   | 146  | 132    | 90.4    |
| 环保   | 96   | 73     | 76.0    |
| 法治   | 138  | 129    | 93.5    |

在具体进行测试后发现，这种方法比全文进行向量空间化后训练、测试速度要提高近 80%，分类平均准确率从 84.3% 提升到了 91%，提升近 6 个百分点，效率明显提高。观察具体的测试数据，发现基于领域词语激活的文本分类结果普遍

好于基于 SVM 算法的文本分类结果，但饮食和环保类的分类结果略差，究其原因，来自 HNC 理论的分类不同于普通的文本分类体系，它的分类体系更细于普通的文本分类体系，一般的文本分类体系类别可以包含多个 HNC 的具体领域类别，但是在饮食和环保领域，则 HNC 的具体领域类别具有和别的类别的交叉性，导致分类效果不明显。

## 参考文献

- [1] He Tian, Stankovic J A, Lu Chenyang, et al. SPEED: A Stateless Protocol for Real-time Communication in Ad Hoc Sensor Networks[C]//Proc. of the 23rd International Conference on Distributed Computing Systems. Rhode Island, USA: IEEE Computer Society, 2003: 46-55.
- [2] Xu Yingqi, Lee Wangchien. Exploring Spatial Correlation for Link Quality Estimation in Wireless Sensor Network[C]//Proc. of IEEE International Conference on Pervasive Computing and Communications. Pisa, Italy: [s. n.], 2006: 200-211.
- [3] Woo A, Tong T, Culler D. Taming the Underlying Challenges of Reliable Multihop Routing in Sensor Networks[C]//Proc. of the 3rd ACM Conference on Embedded Networked Sensor Systems. Los Angeles, California, USA: ACM Press, 2003: 14-27.
- [4] Mistic J, Mistic V B. Access Delay for Nodes with Finite Buffers in IEEE 802.15.4 Beacon Enabled PAN with Uplink Transmissions[J]. Computer Communications, 2005, 28(10): 1152-1166.
- [5] Woo A, Culler D. Evaluation of Efficient Link Reliability Estimators for Low-power Wireless Networks[D]. Berkeley, USA: EECS Department, University of California, 2003.
- [6] Bianchi G, Tinnirello I. Kalman Filter Estimation of the Number of Competing Terminals in an IEEE 802.11 Network[C]//Proc. of INFOCOM'03. San Francisco, CA, USA: IEEE Computer Society Press, 2003: 844-852.

好于基于 SVM 算法的文本分类结果，但饮食和环保类的分类结果略差，究其原因，来自 HNC 理论的分类不同于普通的文本分类体系，它的分类体系更细于普通的文本分类体系，一般的文本分类体系类别可以包含多个 HNC 的具体领域类别，但是在饮食和环保领域，则 HNC 的具体领域类别具有和别的类别的交叉性，导致分类效果不明显。

## 5 结束语

本文提出一种基于文章标题信息进行文本自动分类的新方法，有效减少了文本分类的计算量，降低了特征向量的维数，该方法在标题信息的分析过程中采用 HNC 句类分析系统，将标题信息中带有领域信息的词语或短语进行领域激活，通过领域词语的 HNC 映射符号来进行对应领域的分类，这种文本分类新方法大大减少了向量空间的维数，同时也提高了文本分类的速度和精度。

## 参考文献

- [1] 程传鹏. 中文网页分类的研究与实现[J]. 中原工学院学报, 2007, 18(1): 61-64.
- [2] 王金凤. 一种基于特征聚合理论和 LSI 的文本分类新方法[J]. 北京理工大学学报: 社会科学版, 2004, 6(5): 92-94.
- [3] 缪建明. 专业活动领域句类的设计与知识表示[D]. 北京: 中国科学院声学研究所博士论文, 2007.
- [4] 黄曾阳. 语言概念空间的基本定理和数学物理表示式[M]. 北京: 海洋出版社, 2004.