

基于网络编码的无线传感网多路径传输方法^{*}

李姗姗⁺, 廖湘科, 朱培栋, 肖 依

(国防科学技术大学 计算机学院, 湖南 长沙 410073)

A Method for Multipath Routing Based on Network Coding in Wireless Sensor Network

LI Shan-Shan⁺, LIAO Xiang-Ke, ZHU Pei-Dong, XIAO Nong

(School of Computer, National University of Defense Technology, Changsha 410073, China)

⁺ Corresponding author: E-mail: littlegege@263.net; shanshanli@nudt.edu.cn

LI SS, Liao XK, Zhu PD, Xiao N. A method for multipath routing based on network coding in wireless sensor network. *Journal of Software*, 2008,19(10):2638-2647. <http://www.jos.org.cn/1000-9825/19/2638.htm>

Abstract: Reliability is crucial in many wireless sensor network (WSN) applications. Most of existing approaches are redundancy-based, such as employing multi-path or retransmission schemes. However, those designs often waste energy, and thus shorten the network lifetime. To address this issue, this paper proposes an energy aware method which employs network coding scheme based on multi-path routings. By encoding a group of data into independent new packets and transmitting them along multiple paths, this paper offsets the effect of link failure with a little extra overhead. The other strength of this design is that it only needs small-scale linear operations. An approximate method to effectively estimate the number of paths needed is also employed. Comprehensive simulations and results verify the validation of the theoretical results in the paper.

Key words: WSN (wireless sensor network); reliable data delivery; network coding; multipath routing

摘 要: 可靠性在无线传感器网络中是非常重要的。传感器网络主要通过增加传输冗余来提高数据传输的可靠性,如多路径或重传。然而,这些方法会造成能效降低,缩短网络生命周期。因此,提出了一种能量有效的方法,将一种新型的网络编码与多路径结合在一起,通过将同组数据编码产生的相互独立的多份数据沿多条路径进行传输,有效地降低了对单份数据的依赖,减少了链路失效带来的影响。在保证数据传输可靠性的同时,显著地减少了通信量,而代价仅仅是少量的冗余数据传输和小规模的线性运算。此外,还就其中的关键问题——每组数据所需的最小路径数问题提出了一种低开销的近似方法。详细的模拟实验验证了该方法的有效性。

关键词: 无线传感器网络;可靠数据传输;网络编码;多路径路由

中图分类号: TP393 文献标识码: A

无线传感器网络是当前在国际上备受关注的、涉及多学科高度交叉、知识高度集成的前沿热点研究领域。它以分布式、自组织的方式灵活地对环境进行感知和监控,有着传统网络无法企及的优势^[1-3]。在许多紧急事件,如煤矿瓦斯泄漏^[4]、军事战场上敌方情况的监测等,监控方对网络数据传输的可靠性提出了比较高的要求,一旦

^{*} Supported by the National Basic Research Program of China under Grant No.2006CB303000 (国家重点基础研究发展计划(973)); the National High-Tech Research and Development Plan of China under Grant No.2006AA01Z213 (国家高技术研究发展计划(863))

Received 2007-02-15; Accepted 2007-06-18

发生异常事件,则要求报文可靠、实时地传回,否则就会造成大灾难.然而,无线通信本身并不稳定,受周围环境、传输距离、传输速率等多方面的影响,传感器节点能源受限,网络动态性强.因此,在传感器网络中保障可靠的数据传输是一个关键而棘手的问题.

当前的研究大都采用冗余传输的方式来保证数据传输的可靠性,如多路径或者重传^[5-9].多路径方式在源和目的(sink)之间构建多条路径,通过将(同一份)数据报文同时沿着多条路径传输来增加数据传输的可靠性.重传方式计算为达到一定的传输可靠性,路径上每跳至少需要的传输次数.多路径相对于重传其最大的优势是延时小,考虑到高可靠性要求的报文一般也对实时性要求也较高,所以,本文使用多路径作为保障可靠性的基本手段.另外,传感器网络一般都大量部署在野外和无人管理区,电池是唯一的能耗来源,因此,有效地利用能量非常重要^[10].一旦部分节点能量耗尽导致网络覆盖区域减小,则整个网络的监控能力也就减弱了.因此,必须高效地解决传输可靠性问题,减少多路径方法的冗余传输量.

网络编码^[11-14]是一种新型的数据编码方式,它可以在网络中提高网络吞吐率、降低能耗、减少传播延迟等等.本文将网络编码与多路径数据传输相结合,通过将同组数据编码产生的相互独立的多份数据沿多条路径进行传输,有效地减少了链路失效带来的影响,提高了传输可靠性.与传统的多路径方法相比,在保证同等数据传输可靠性的同时,显著地减少了通信量.传感器网络中主要的能耗用于节点间通信,因此,减少传输量是节约能耗的主要方法.网络编码可以减少为保证一定的期望可靠性所需要的多路径数量,为解决可靠数据传输问题提供了一种高效的方法.

本文第1节简要介绍网络编码和多路径路由的相关工作.第2节讨论多路径数据传输模型.第3节首先分析实用网络编码的基本理论,然后提出网络编码与多路径数据传输具体的结合方式.第4节提出一种低开销的近似方法计算所需的最少路径数.第5节进行详细的模拟实验.第6节总结全文并提出今后的研究方向.

1 相关工作

网络编码理论最早由 Ahlswede 等人开始研究^[11],他们提出,通过使用网络编码,网络的组播能力可以达到(组播树的)最大流最小割流量,而这种网络吞吐率是常规路由无法达到的.文献[12]表明,线性网络编码足以达到这种组播数据传输能力.文献[13]提出了一种线性网络编码的几何框架,并证明了这种编码方法的时间无关性.这些工作大都基于理论研究,并假定已知网络全局拓扑和数据源的位置等等,而这些假设在分布式自组织网络中是不太现实的.文献[14]提出了一种实用网络编码方法,可以用于 P2P 网络中支持大规模高效的流媒体数据分发^[15].文献[16]在无线网络的两节点间使用网络编码进行数据交换,充分利用了无线广播信道,提高了网络吞吐率.文献[17]将网络编码应用在传感器网络的数据存储中,提高了数据备份的鲁棒性.文献[18]将网络编码与 directed diffusion^[19]相结合,选择性地加强有利数据的路径.目前,尚很少有实用的网络编码方法应用在传感器网络的可靠数据传输中,在保证数据传输的可靠性的同时减少数据传输量.

本文采用多路径的方式来保证数据传输的可靠性,但并不依赖于具体的多路径路由算法,下面对多路径的相关工作进行介绍.文献[5]基于 directed diffusion 的框架,使用多路径方式迅速从失败路径中恢复.EAR(energy aware routing)^[6]从能量的角度出发,按照能耗比例选择传输路径,EAR 只是基于路径能耗从多条路径中选择一条传输路径,并未使用多路径进行传输.文献[7]提出一种轻型的端到端多路径协议 ReInForM,用于保证数据传输的可靠性.因为无需在节点缓存信息,所以比较适合节点内存受限的传感器网络.文献[8]提出一种按需构造多路径方法,通过调整路径的相对距离消除路径间的干扰等.文献[9]基于地理位置^[20]构建多路径以缓解拥塞.这些多路径的方法都只对原始数据进行多路径传输,尚很少有文章将其与网络编码进行结合.

2 多路径数据传输模型

在多路径数据传输中,数据源节点根据应用的期望可靠性 r 以及信道的状况选择将数据传输到 sink 所需的路径数.

定义 1(期望可靠性). 数据源产生的包被成功传输到采集点的概率.

定义 2(路径). 在一个特定的无线传感器网络中的一系列首尾相连的有向链路集合.

假定网络中的节点随机分布,源节点周期性地感知数据传送给 sink.传统多路径方法将数据的副本同时沿多条路径传输.为简化问题,假定网络节点密集分布,可以很方便地从源到 sink 构建多条路径,路径长度与最短路径长度基本一致,平均跳步数为 k ,信道失效率为 e ,为达到期望可靠性 r ,每份数据至少需要的路径数为 N ,则 N 的计算方法见公式(3).如果数据源的邻居数少于所需的路径数,一些路径将转发多份数据副本,担负起多条路径的作用,即这里的多路径不一定是物理存在的多条路径.

$$N = \frac{\log(1-r)}{\log(1-(1-e)^k)} \quad (1)$$

在常规的多路径数据传输中,要传送 h 份不同的数据报文,需要 $h \times N$ 条路径,这种方法虽然达到了期望可靠性,但是严重浪费了带宽.本文将多路径数据传输与一种新型的网络编码结合在一起,通过将多份数据编码在一起,生成相互无关的 m 份数据沿多路径传输,只要 sink 收到其中的一部分就能恢复原始数据.依据概率统计理论,显然, m 小于 $h \times N$,也就是说,对于每份数据,平均只需要 $\frac{m}{h}$ 而不是 N 条路径.这种方法可以消除对单份数据的依赖性,减少总体所需的路径数,非常适合传感器网络节点通信易失效的弱点.在无线传感器网络中,通信开销是能耗的主体,减少的传输路径数就意味着降低了通信能耗,延长了网络生命周期.本文下面给出一种使用网络编码的方法,进行多路径传输以及相应的参数选择.

3 基于网络编码的多路径传输方法

3.1 实用网络编码方法

实用网络编码^[14]的基本思想是,将分组数据编码成多份相互无关的数据,接收端只要能够收到其中的一部分,就可以恢复出原始数据,具体产生的编码数据数量与应用的需求相关.假定一组含有 h 份报文,当节点要发送这些报文时,首先产生 h 个随机数并组成 h 维向量 r_1, r_2, \dots, r_h ,随后,按向量 (r_1, r_2, \dots, r_h) 以及数据块 B_1, B_2, \dots, B_h 的内容生成一个新的数据块 E ,并且也按向量 (r_1, r_2, \dots, r_h) 及该组中的系数向量 $\bar{e}_1, \bar{e}_2, \dots, \bar{e}_h$ 生成一个新的系数向量 $\bar{\alpha}$, E 和 $\bar{\alpha}$ 的定义分别如下:

$$E = r_1 B_1 + r_2 B_2 + \dots + r_h B_h \quad (2)$$

$$\bar{\alpha} = r_1 \bar{e}_1 + r_2 \bar{e}_2 + \dots + r_h \bar{e}_h \quad (3)$$

每次传输都需产生一组新的随机数.除了数据源以外,中间节点如果收到多份数据,还可以再次编码.这种编码方法的好处在于,把不同的数据糅合在一起,减少对单份数据的依赖,增加了数据的传输可靠性,非常适合传感器网络报文易丢失的特点.只要 sink 接收到 h (或大于 h) 份数据,就可以恢复出原始数据.假定 sink 收到的 h 份数据分别是 E_1, E_2, \dots, E_h ,则 t 进一步判断相应的 h 个系数向量 $\alpha_1, \alpha_2, \dots, \alpha_h$ 是否线性无关,即 $\alpha_1, \alpha_2, \dots, \alpha_h$ 所组成的 $h \times h$ 维矩阵是否为满秩的,如果是,则原始数据可按图 1 所示方式恢复.这种方法与传统的前向纠错编码(forward error correction,简称 FEC)^[18]在思想上有些相似,但网络编码比 FEC 更灵活、更鲁棒,这一点可以在第 3.2 节中描述的网络编码与多路径传输的具体结合方法中非常明显地体现出来.

$$\begin{bmatrix} E_1 \\ E_2 \\ \vdots \\ E_h \end{bmatrix} = \begin{bmatrix} r_1^1 & \dots & r_h^1 \\ \vdots & \ddots & \vdots \\ r_1^h & \dots & r_h^h \end{bmatrix} \begin{bmatrix} B_1 \\ B_2 \\ \vdots \\ B_h \end{bmatrix} \Rightarrow \begin{bmatrix} B_1 \\ B_2 \\ \vdots \\ B_h \end{bmatrix} = \begin{bmatrix} r_1^1 & \dots & r_h^1 \\ \vdots & \ddots & \vdots \\ r_1^h & \dots & r_h^h \end{bmatrix}^{-1} \begin{bmatrix} E_1 \\ E_2 \\ \vdots \\ E_h \end{bmatrix}$$

Fig.1 Recovering of original data

图 1 原始数据恢复

3.2 网络编码在多路径传输中的应用方法

当使用网络编码进行多路径传输时,源节点不像传统多路径方法那样对每份数据沿多条路径传输,而是首

先将数据分组,每个组包含 h 个等长的数据段 B .每组按其产生的先后顺序赋予一个组标识,第 1 组标识为 0,第 2 组标识为 1,并依此类推,当组标识增长到某个上限值时再让其重新回归为 0.每个数据段都分配一个 h 位的单位系数向量 e_i ,对于数据段 $B_i(1 \leq i \leq h)$,其相应的 e_i 值定义为:除第 i 个分量的值为 1 外,其余分量的值均为 0.这样,一个组内的所有向量就构成了一个 $h \times h$ 单位矩阵.对每个数据段而言,其报文格式如图 2 所示.在下文中,除非特殊说明,所有数据段均属于同一组.

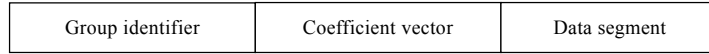


Fig.2 Packet format using network coding

图 2 使用网络编码的报文格式

在多路径中结合网络编码的主要方法如下:

1. 对于源节点,首先计算所需的路径数 m ,并产生 m 组随机数;再用每组随机数对组内原始数据和随机数按照公式(2)、公式(3)方式进行编码,如图 3 所示,分别产生 m 份编码数据,然后将编码数据沿 m 条路径传输.

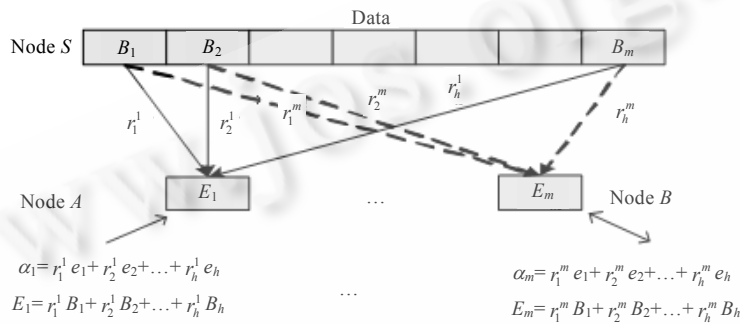


Fig.3 Source node encodes group data into m new packets

图 3 源节点将组内数据编码成 m 份新数据

2. 对于中间节点,在给定的时间 t 内检查所收到的编码数据,假定数量为 c ,则产生 c 组随机数,并按照公式(2)、公式(3)方式对编码数据进行再次编码,进一步减少数据的相关性.

3. 对于 sink 节点,根据收到的报文数量 $Recvcnt$ 进行判断:

如果 $Recvcnt \geq h$,判断公式(1)对应的系数向量矩阵是否线性无关,或者公式(2)系数向量线性相关但满足秩大于或等于 h ,如果是,则 sink 按图 1 的方法恢复原始报文;

如果 $Recvcnt < h$ 或者 $Recvcnt \geq h$ 但对应的系数向量的秩小于 h ,判断当前事件的数据是否满足期望可靠性 r ,如果满足,则丢弃这些报文;如果不满足,则 sink 沿路径相反方向请求再次发送新的编码数据,要求新编码数据的系数向量和已有数据的系数向量线性无关.

整个编码数据传输的过程如图 4 所示.接下来再具体讨论网络编码相对 FEC 的优势,FEC 只允许在源节点编码,目标节点解码,而网络编码在中间节点也可以再编码和解码.例如,传输过程中某中间节点 A 收到了编码数据 $\bar{\alpha}_1$ 和 $\bar{\alpha}_2$,节点 B 收到了编码数据 $\bar{\alpha}_3$ 和 $\bar{\alpha}_4$,假定 $\bar{\alpha}_1$ 与 $\bar{\alpha}_3$ 的系数向量线性相关,如果使用 FEC, A, B 节点会把数据直接转发出去.如果在接下来的传输过程中,由于信道原因造成 $\bar{\alpha}_2$ 与 $\bar{\alpha}_4$ 丢失,则 sink 收到包括 $\bar{\alpha}_1$ 与 $\bar{\alpha}_3$ 在内的 h 份数据时,依然无法恢复出原始数据,因为这 h 份数据的系数向量组成的矩阵非满秩矩阵.如果使用网络编码, A, B 会再次对收到的编码数据编码,重新生成 4 份编码数据,那么,这 4 份数据收到的任意两份的系数向量都是线性无关的.实际上,编码产生的 4 份数据与原始 4 份数据的系数矩阵的秩相同,即其中任意 3 份的系数向量都是线性无关的.此外,网络编码采用了线性编码,编码数据只与产生的随机数有关,计算量很小,非常适合传感器节点资源受限的特点.

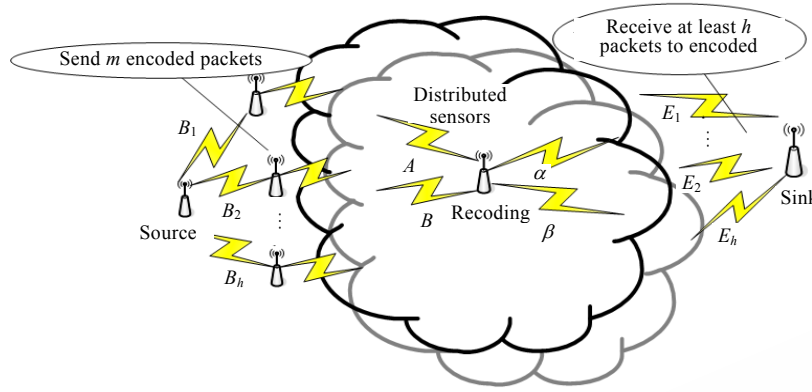


Fig.4 Multipath data delivery based on network coding

图 4 基于网络编码的多路径传输

3.3 参数设置

在这种方法中,主要有 3 个参数需要设置:

- 随机数的范围 Fq .如果太小,系数向量容易线性相关;太大,会引发大的存储开销.表 1 给出了随机数大小与系数向量线性无关的概率的对应关系^[21].可以看到,当随机数在 $[0,2^8]$ 范围内取值时,系数向量线性无关的概率为 99.6%,原始数据被恢复的概率非常高,而系数向量所需的存储开销仅为一个字节;
- 组内包含的报文数量 h .如果设置得太小,就无法充分发挥网络编码的潜在优势;如果太大,系数向量的存储开销会比较大. h 的设置与具体的节点有关,不同的报文长度有不同的最佳 h 值;
- 每组数据总共需要多少条路径,即 m 的计算.这个参数直接关系到网络编码为多路径数据传输减少的通信量.因此,这个参数的设置非常关键.

Table 1 The relation between the field size of random number and the linear dependency of coefficient vectors

表 1 随机数取值范围与系数向量线性无关的概率的对应关系

q	Probability	q	Probability	q	Probability
2^1	0.288 788	2^5	0.967 773	2^9	0.998 043
2^2	0.688 538	2^6	0.984 131	2^{10}	0.999 022
2^3	0.859 406	2^7	0.992 126	2^{11}	0.999 511
2^4	0.933 595	2^8	0.996 078	2^{12}	0.999 756

4 最小路径数计算方法

为计算满足期望可靠性所需要最小路径数,达到能耗最少,第 4.1 节讨论了路径数目的理想计算方法.但是,计算这个参数的理想值十分复杂,开销较大,而传感器节点计算能力受限,因此,第 4.2 节提出了一种低开销的近似方法,并能自适应地进行修正以逼近最优值.另外,本文提出的方法可以普遍适用于各种多路径路由方法,并不依赖于具体的多路径协议.目前,已有许多工作研究了多路径的构建^[4,6,7],因此,本文假定各数据源已按某种方法构建和维护了多条路径^[4].

4.1 理想计算方法

对于源节点而言,每条路径成功的概率为 $(1-e)^k$.那么根据概率统计的相关知识, m 应该满足如下条件: m 条路径中至少有 h 条成功传输的概率等于 r ,如方程(4)所示:

$$C_m^h(1-(1-e)^k)^{m-h}(1-e)^{kh} + C_m^{h+1}(1-(1-e)^k)^{m-h-1}(1-e)^{k(h+1)} + \dots + C_m^m(1-e)^{km} = r \tag{4}$$

显然,使用方程(4)计算 m 计算比较复杂,尤其是对于计算能力有限的传感器节点.在下一小节中,我们将提出一种近似方法来简化计算,同时保持较高的计算精度.

4.2 近似方法

在无线传感器网络中,路由为多跳传输方式,每条路径的成功传输概率是路径上所有跳步成功传输概率的累积.那么,对于 k 跳的路径,成功传输的概率是 $(1-e)^k$.假定每条路径的跳步数都为 k ,则 m 条路径中有 n 条成功传输的概率服从 Bernoulli 分布,且成功传输的路径数的期望值为公式(5):

$$E(n) = m(1-e)^k \frac{-b \pm \sqrt{b^2 - 4ac}}{2a} \quad (5)$$

根据中心极限定理,当 m 趋近无穷时,Bernoulli 分布逼近正态分布.为方便计算,这里也用正态分布近似,误差将在模拟实验中进行分析.目标是希望成功传输至少 h 份数据的概率为 r ,也就是期望 $P(n \geq h) \geq r$ 成立.我们利用正态分布的均值和方差求解 m .

$$\mu = E(n) = m(1-e)^k \quad (6)$$

$$\sigma^2 = m(1-e)^k(1-(1-e)^k) \quad (7)$$

不同的 m 与 $(1-e)^k$ 会产生不同的正态分布,因此可以转化成标准正态分布再进行计算,变量 n 转换为 n^* , n^* 服从 $N(0,1)$ -分布.

$$n^* = \frac{n - \mu}{\sigma} \quad (8)$$

在标准正态分布中,对于给定的 r ,满足 $P(n^* \geq x_r) \geq r$ 的 x_r 可以通过查表获知.表 2 给出了一些数值对,这些数值与路径数 m 无关.

Table 2 x_r in different desired reliability r

表 2 不同期望可靠性 r 下的 x_r 值

r (%)	70	75	80	85	90	95
x_r	-0.52	-0.67	-0.85	-1.03	-1.28	-1.65

根据 r 获得相应的 x_r 后,通过公式(8)变换回原始的正态分布公式(9),并将 $P(n^* \geq x_r) \geq r$ 变形成 $P(n \geq x_r \sigma + \mu) \geq r$.

$$\frac{n - \mu}{\sigma} \geq x_r \Rightarrow n \geq x_r \sigma + \mu \quad (9)$$

这样,就可以把 $x_r \sigma + \mu$ 作为期望成功传输的路径数,并通过方程(10)解出 m :

$$h = x_r \sigma + \mu = x_r \sqrt{m(1-e)^k(1-(1-e)^k)} + m(1-e)^k \quad (10)$$

4.3 自适应路径数修正

利用上述近似方法,可以很方便地解出所需的路径数.然而,由于 m 不趋近于无穷,解出的近似值与公式(4)中的理想值会存在偏差.图 5 分析了不同的期望可靠性(r)和信道失效率(e)下的实际误差.其中,组内包含的数据包数目 h 为 4,路径的平均跳数为 6.从图中可以看出,近似值略大于理想值,当期望可靠性很大、信道失效率很高的情况下($r=0.95, e=0.4$),路径数的偏差小于 10,不到实际最优路径数的 5%.因此,数据源可以对近似值作一些启发式修正,减少不必要的通信量.

$$m = \left\lceil \frac{m}{h} \right\rceil \times h \quad (11)$$

然而,这种静态调整并不精确,再加上传感器网络动态性强,路径可能发生变化,实际数据传输过程中,理想值会不稳定,最佳值将随网络状况动态变化.因此,本文提出了一种自适应的调整策略修正多路径的数量.Sink 周期性地检查接收到的数据的可靠性,并通告给数据源进行相应的调整.调整策略与应用相关,一些多路径算法^[7,8]通过调整转发概率来获得最佳值.本文使用公式(12)进行调整,直到收到的数据的实际可靠性等于期望可靠性为止.

$$m_{i+1} = \frac{r \times m_i}{r_i^r} \quad (12)$$

m_i, r_i^r : 第 i 个检测周期中的路径数与实际获得的可靠性, $m_0=m$.

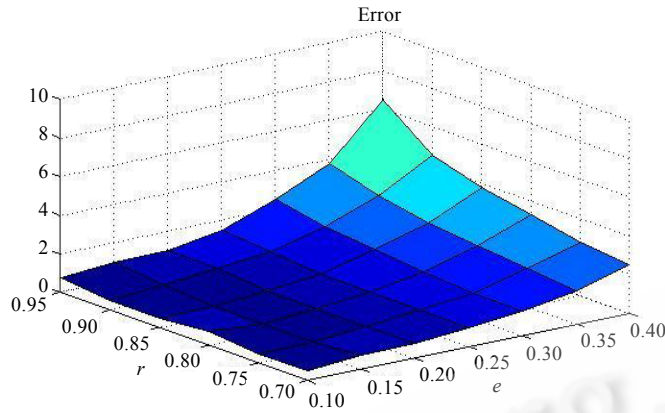


Fig.5 Error of the approximate method in different desired reliability and channel error rate
图 5 近似方法在不同可靠性、不同信道失效率下的误差

5 性能模拟

本文使用 Matlab 中的 Simulink 对基于网络编码的多路径数据传输进行模拟,分别考察网络编码为传统多路径传输带来的能效增加以及编码产生的通信开销.

模拟环境由 100m×100m 区域内统一分布的 200 个节点组成,在所有的模拟实验中,数据源位于区域的左下方,而 sink 不断地变换位置,以适应不同模拟的需求(例如,图 7 中的模拟需要满足从数据源到 sink 的不同 hop 数).标准的数据包长度设置为 100 字节,从数据源到 sink 的多路径使用文献[7]中的方法构建.

当传输报文时,数据源首先将感知数据分组,每组 h 个报文并产生 m 组随机数,通过编码产生 m 份编码数据沿多条路径传输.当 sink 收到 h 份报文时,就判断这些报文对应的系数矩阵是否满秩:如果是,则解码恢复原始数据;若不是,则继续等待其他报文,待给定时间后仍未收到足够的编码数据,则该组数据无法恢复.在模拟实验中,随机数的取值域 Fq 为 8 bit.

图 6 给出了在不同期望可靠性要求下使用网络编码与常规传输方式的路径数($e=20\%, k=6, h=4$).当期望可靠性增高时,网络编码带来的优势越发明显.通过使用网络编码,无需保证每份编码数据的可靠性,只需要保证所有编码数据的一部分传到 sink 即可.另外,所有编码数据相互独立,没有依赖关系,当期望可靠性增加时,需要的路径数迅速增加,使用网络编码所需的路径数与不需要时的差异就越来越大,显著地减少了通信量.图 7 给出了在不同的路径长度下所需的路径数($e=20\%, r=0.9, h=4$).随着路径 hop 数的增加,路径的可靠性呈指数降低,为了保证期望可靠性,所需的路径数也急剧增加.因此,路径越长,网络编码的优势越明显.

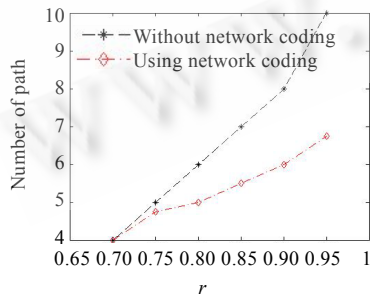


Fig.6 Number of path needed per packet under different r
图 6 不同期望可靠性下所需的路径数

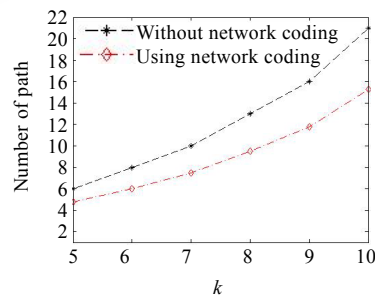


Fig.7 Number of path needed per packet
图 7 不同路径长度下所需的路径数

图 8 给出了在不同信道失效率下所需的路径数.同理,信道状况越差,则所需的路径数越多,网络编码节约的能量也越多.以上模拟实验表明,网络编码在期望可靠性越高,信道状况越差,网络规模越大时优势越明显,非常适合传感器网络动态性强、规模大等特点.

网络编码虽然具有诸多优势,但是也带来了额外的能耗开销.系数向量与随机数之间的计算是线性的,开销很小,与通信能耗相比几乎可以忽略不计.但系数向量增加了报文的长度,增加了通信开销,因此,必须考察这部分能耗是否会抵消了网络编码节约的通信能量.依照上面的假设,正常报文长度为 100 字节,每组包含 4 个数据包,则使用编码的数据包长度将为 104 字节(F_q 为 8 bit).由于系数向量无法用路径数来表达,为方便比较,使用字节数作为能耗度量.

图 9~图 11 分别比较了在不同期望可靠性、不同信道失效率和不同路径长度下网络编码减少的能量与传输系数向量耗费的能量.从图中可以看出,系数向量耗费的能量与网络编码节约的能量相比几乎可以忽略不计,主要原因在于系数向量所占的字节数非常少.但在图 9 中有一个异常点,当期望可靠性为 70%时,系数向量耗费的能量比网络编码节约的能量要多.这是因为,当期望可靠性较低时,数据传输需要路径数较少,网络编码无法发挥优势.相比之下,系数向量所耗费的额外能量还要更多.因此,在网络状况较好、期望可靠性较低和网络规模较小时,没有必要使用网络编码.

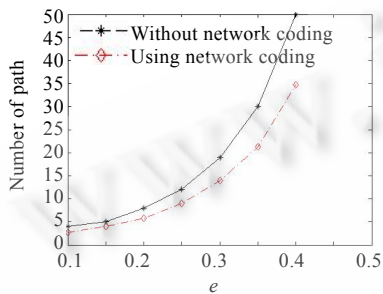


Fig.8 Number of paths needed per packet under different e
($h=4, k=6, r=0.9$)

图 8 不同信道失效率下所需的路径数
($h=4, k=6, r=0.9$)

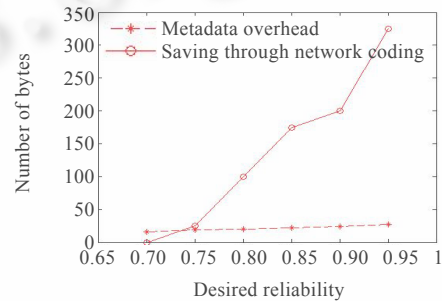


Fig.9 Number of bytes saved by network coding vs. consumed by metadata per packet under different r
($e=20\%, k=6, h=4$)

图 9 网络编码减少的能耗和系数向量耗费的能量随可靠性变化的关系($e=20\%, k=6, h=4$)

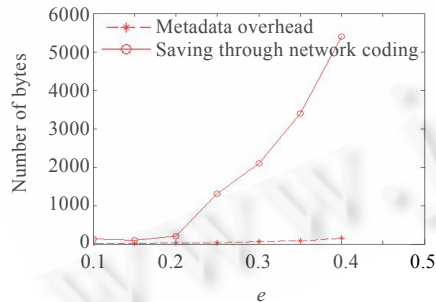


Fig.10 Number of bytes saved by network coding vs. consumed by metadata per packet under different e
($r=0.9, k=6, h=4$)

图 10 网络编码减少的能耗和系数向量耗费的能量随信道失效率变化的关系($r=0.9, k=6, h=4$)

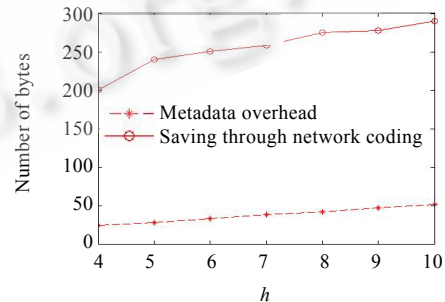


Fig.11 Number of bytes saved by network coding vs. consumed by metadata per packet under different h
($r=0.9, e=20\%, k=6$)

图 11 网络编码减少的能耗和系数向量耗费的能量随路径长度变化的关系($r=0.9, e=20\%, k=6$)

6 结论及展望

本文将网络编码技术结合进多路径数据传输中,降低了链路失效带来的影响,在有效保证传输可靠性的同时,进一步减少了能耗,所需的代价仅仅是少量的元数据(系数向量)通信开销.这对能量极其受限的传感器网络而言是至关重要的,它可以延长网络的生命周期.实验结果表明,网络编码在期望可靠性越高、信道状况越差、网络规模越大时,优势越明显,非常适合传感器网络动态性强规模大等特点.

然而,由于传感器网络大规模部署,很多数据都是冗余的,对所有数据都编码是没有必要的.在未来的工作中,我们将研究如何更好地利用数据之间的相关性,基于分簇技术,将压缩编码与网络编码结合起来,期望在能耗有效性与数据鲁棒性之间找到更好的折衷点.

References:

- [1] Akyildiz IF, Su W, Sankarasubramaniam Y, Cayirci E. Wireless sensor networks: A survey. *Computer Networks*, 2002,38(4):393-422.
- [2] Karl H, Willig A. A short survey of wireless sensor networks. Technical Report, TKN-03-018, Telecommunication Networks Group, Technical University Berlin, 2003.
- [3] Chen DZ, Varshney PK. QoS support in wireless sensor networks: A survey. In: Proc. of the Int'l Conf. on Wireless Networks (ICWN). Las Vegas, 2004.
- [4] Liu YH, Li M. Iso-Map: Energy-Efficient Contour Mapping in Wireless Sensor Networks. In: Proc. of the IEEE ICDCS. 2007. <http://ieeexplore.ieee.org/Xplore/login.jsp?url=/iel5/4268147/4268148/04268190.pdf?tp=&isnumber=&arnumber=4268190>
- [5] Ganesan D, Govindan R, Shenker S, Estrin D. Highly-Resilient, energy-efficient multipath routing in wireless sensor networks. *ACM SIGMOBILE Mobile Computing and Communications Review*, 2001,5(4):11-25.
- [6] Shah RC, Rabaey M. Energy aware routing for low energy ad hoc sensor networks. In: Proc. of the IEEE Wireless Communications and Networking Conf. (WCNC). 2002. 17-21.
- [7] Deb B, Bhatnagar S, Nath B. ReInForm: Reliable information forwarding using multiple paths in sensor networks. In: Proc. of the 28th Annual IEEE Conf. on Local Computer Networks. Los Alamitos: IEEE Computer Society, 2003. 406-415. http://ieeexplore.ieee.org/xpl/freeabs_all.jsp?arnumber=1243166
- [8] Voigt T, Dunkels A, Braun T. On-Demand construction of non-interfering multiple paths in wireless sensor networks. In: Proc. of the 2nd Workshop on Sensor Networks at Informatik. Bonn, 2005. <http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.14.4578>
- [9] Popa L, Raiciu C, Stoica I, Rosenblum DS. Reducing congestion effects in wireless networks by multipath routing. In: Almeroth CK, Higashino T, eds. Proc. of the ICNP. Santa Barbara: IEEE Press, 2006. 96-105. <http://portal.acm.org/citation.cfm?id=1317535.1318360&coll=&dl=>
- [10] Li M, Liu YB. Underground structure monitoring with wireless sensor networks. In: Proc. of the ACM/IEEE IPSN. Cambridge, 2007. http://ieeexplore.ieee.org/xpl/freeabs_all.jsp?isnumber=4379652&arnumber=4379666&count=83&index=13
- [11] Ahlswede R, Cai N, Li SYR, Yeung RW. Network information flow. *IEEE Trans. on Information Theory*, 2000,46(4):1204-1216.
- [12] Li SYR, Yeung RW, Cai N. Linear network coding. *IEEE Trans. on Information Theory*, 2003,49(2):371-381.
- [13] Koetter R, Medard M. An algebraic approach to network coding. *IEEE/ACM Trans. on Networking*, 2003,11(5):782-795.
- [14] Chou PA, Wu YN, Jain K. Practical network coding. In: Proc. of the 41st Annual Allerton Conf. on Communication Control and Computing. 2003. <http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.11.697>
- [15] Gkantsidis C, Rodriguez PR. Network coding for large scale content distribution. In: Proc. of the IEEE Inforcom. 2005. http://ieeexplore.ieee.org/xpl/freeabs_all.jsp?arnumber=1498511
- [16] Wu YN, Chou PA, Kung SY. Information exchange in wireless networks with network coding and physical-layer broadcast. Technical Report, MSR-TR-2004-78, Microsoft Research, 2004.

- [17] Dimakis AG, Prabhakaran V, Ramchandran K. Ubiquitous access to distributed data in large-scale sensor networks through decentralized erasure codes. In: Proc. of the Symp. on Information Processing in Sensor Networks (IPSN). Los Angeles, 2005. http://ieeexplore.ieee.org/xpl/freeabs_all.jsp?arnumber=1440909
- [18] Toledo OA, Wang XD. Efficient multipath in sensor networks using diffusion and network coding. In: Proc. of the 40th Annual Conf. on Information Sciences and Systems. 2006. <http://ieeexplore.ieee.org/iel5/4067758/4067759/04067782.pdf?tp=&isnumber=&arnumber=4067782>
- [19] Intanagonwiwat C, Govindan R, Estrin D. Directed diffusion: A scalable and robust communication paradigm for sensor networks. In: Proc. of the ACM Mobicom. Boston: ACM Press, 2000. 56–67. <http://portal.acm.org/citation.cfm?id=345920&dl=ACM&coll=portal>
- [20] Li M, Liu YB. Rendered path: Range-Free localization in anisotropic sensor networks with holes. In: Proc. of the ACM MobiCom. 2007. <http://portal.acm.org/citation.cfm?id=1287853.1287861&coll=ACM&dl=GUIDE&type=series&idx=SERIES395&part=series&WantType=Proceedings&title=MobiCom>
- [21] Medard M, Acedanski S, Deb S, Koetter R. How good is random linear coding based distributed networked storage? In: Proc. of the NETCOD 2005. 2005. <http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.62.4801>



李姗姗(1980—),女,江西丰城人,博士生,主要研究领域为无线传感器网络的可靠路由。



朱培栋(1971—),男,博士,副教授,主要研究领域为路由技术。



廖湘科(1963—),男,教授,CCF 高级会员,主要研究领域为并行计算,操作系统,普适计算。



肖依(1969—),男,博士,教授,博士生导师,CCF 高级会员,主要研究领域为体系结构,网格计算,传感器网络。