

无线传感器网络应用简单 Reed-Solomon 编码的低能耗和低时延可靠数据收集方案

朱艺华, 徐骥, 田贤忠, 池凯凯

(浙江工业大学 计算机科学与技术学院, 杭州 浙江 310023)

摘 要 无线传感器网络的数据链路层和物理层通常采用低功率低数据率的 IEEE 802.15.4 标准, 其节点由能量极为有限的电池供电, 且无线链路易丢包, 而且在无线传感器网络的一些应用中, 需要在给定的时间内将数据汇集到基站, 因此, 设计一种节能、低时延且可靠的数据收集方案对无线传感器网络的应用极为重要。本文利用 Reed-Solomon (RS) 编码以提高数据传递可靠性, 致力于将能耗、数据收集率、数据收集时延、数据包编码方案整合为一个优化问题; 给出了简单 RS 码字即 S-RS 码字集合, 证明了该码字集合中任何 n 个码字是线性独立的; 详细叙述了基于 S-RS 码字集合以及段内编码和段间编码的数据收集策略即 S-RS 数据收集策略的具体步骤; 利用概率论, 导出了 S-RS 数据收集策略的总能耗、数据收集率和时延公式, 并依此建立了时延和数据收集率约束下的能耗最小化这一优化问题。该优化问题可行解空间很小, 可以通过穷举法获得最优解。通过求解该优化问题, 节点能够设置最优的数据包编码和重传参数, 使得在时延和数据收集率约束下能耗达到最小。数值分析与仿真表明, S-RS 数据收集策略能够以较低的能耗达到较高数据收集率和较低时延要求。此外, 可以利用导出的总能耗、数据收集率和时延, 通过优化 S-RS 数据收集策略的参数, 最大化数据收集率, 最小化时延, 也可以同时对总能耗、数据收集率和时延中的两者或三者进行优化。

关键词 无线传感器网络; 数据收集; 节能; 可靠性; 时延; Reed-Solomon 编码

中图法分类号 TN92

论文引用格式

朱艺华, 徐骥, 田贤忠, 池凯凯, 无线传感器网络应用简单 Reed-Solomon 编码的低能耗和低时延可靠数据收集方案, 计算机学报, 2015, Vol.38: 在线出版号 No.13

ZHU Yi-Hua, XU Ji, TIAN Xianzhong, and CHI Kaikai, Energy-efficient and low-delay Reliable Data Gathering Scheme Applying Simple Reed-Solomon Code for Wireless Sensor Network, Chinese Journal of Computers, 2015, Vol.38: Online Publishing No.13

Energy-efficient and low-delay Reliable Data Gathering Scheme Applying Simple Reed-Solomon Code for Wireless Sensor Network

ZHU Yi-Hua, XU Ji, TIAN Xianzhong, and CHI Kaikai

(School of Computer Science and Technology, Zhejiang University of Technology, Hangzhou, Zhejiang, China 310023)

Abstract Wireless Sensor Network (WSN) usually adopts the low-power and low-rate IEEE 802.15.4 standard in its Medium Access Control (MAC) and Physical (PHY) layers. In the WSN, the nodes are powered by battery with very limited energy and wireless links are prone to losing packet. Moreover, in some WSN

本课题得到国家自然科学基金重点项目 (No. 61432015) 和面上项目 (No. 61472367, 61379124) 资助。朱艺华, 男, 1961 年生, 博士, 教授, CCF 高级会员 (E20-0012269S), 主要研究领域为物联网、无线网络、网络编码, E-mail: yhzhu@ece.ubc.ca, 手机号: 13588111446。徐骥, 男, 1990 年生, 学士, 硕士生, 主要研究领域为无线传感器网络, E-mail: cornuer@gmail.com。田贤忠, 男, 1968 年生, 博士, 教授, CCF 会员 (No. E200017096M), 主要研究领域为网络编码、移动计算、无线网络协议及优化, E-mail: txz@zjut.edu.cn。池凯凯, 男, 1980 年生, 博士, 副教授, 主要研究领域为无线通信网络, E-mail: kkchi@zjut.edu.cn

applications, it is required that the collected data are delivered to the sink node within a given time. Hence, it is extremely significant to develop an energy-efficient, low-delay, and reliable data gathering scheme for WSN applications. In this paper, Reed-Solomon (RS) code is applied to improve the reliability of packet delivery. It aims to form an Optimization Problem (OP) that integrates with energy consumption, Data Gathering Ratio (DGR), data gathering delay, and coding schemes. The set that contains the Simple Reed-Solomon (S-RS) codes is presented, and it is proved that any n members of the set are linearly independent. The data gathering scheme, called the S-RS data gathering scheme, which applies the S-RS codes, intra-segment coding, and inter-segment coding, is described in detail. In addition, based on probability theory, the Overall Energy Consumption (OEC), the DGR, and the data gathering delay under the S-RS scheme are derived, which are used to build the OP that minimizes the OEC with the constraints of DGR and data gathering delay. The number of the feasible solutions of the OP is small such that it can be easily solved using enumeration method. Through solving the OP, the nodes are able to find the optimal values of the parameters for packet encoding and packet retransmissions such that energy consumption in the S-RS scheme is minimized while the constraints of DGR and data gathering ratio are met. Numeric analysis and simulation results show that the S-RS scheme achieves a higher DGR and a lower delay with a lower energy consumption. Moreover, the derived OEC, DGR, and data gathering delay can be used in optimizing the parameters used in the S-RS scheme so as to maximize the DGR, minimize the data gathering delay, or optimize two or three of the OEC, DGR, and data gathering delay.

Key words Wireless Sensor Network; data gathering; energy conservation; reliability; delay; Reed-Solomon code

1 引言

无线传感器网络(Wireless Sensor Network, WSN)已经被广泛地用于环境监测、医疗保健、交通运输等领域。IEEE 802.15.4 标准^[1]定义了低功耗低速率无线网络的介质访问层(Medium Access Control, MAC)和物理层(Physical, PHY)。当代 WSN 在 MAC 层和 PHY 层采用上述标准,因而其无线链路传递数据包具有不可靠和易丢失等特点。由于低功耗无线链路的无线电覆盖范围小,数据包需要经过多跳传输才能完成从一个传感器节点到达 sink 节点,这会导致端到端的数据包传递成功率非常低。因此,需要设计提高数据包传递成功率的切实可行的方案。

在 WSN 中,节点通常是由能量非常有限的电池供电。因此,研究如何使节点能够在保证可用性的情况下尽量降低能量消耗,对 WSN 长时间工作显得尤为必要。此外,在 WSN 的一些应用尤其是实时应用中,传感器节点捕获的数据需要在短时间内送达 sink 节点,因此,降低 WSN 中数据包收集时延也是非常重要的。

在提高传递数据包可靠性方面,虽然 MAC 层的重传与确认机制有助于提高可靠性,但是它存在着以下弱点: 1) 在 MAC 层设置了数据帧的最大传

输次数,这会导致发送次数达到最大次数但依然发送失败的帧被 MAC 层丢弃; 2) 已有研究表明^[2],在节点移动的网络中,重传机制开销特别大,而且在能量受限的 WSN 中这种方法几乎无法实现数据包可靠传递。鉴于此,本文借助于编码技术来提高数据包传递可靠性,将来自网络层的数据包进行编码,并结合数据链路层的重传和确认机制,以能耗、数据收集率、时延等要求为目标来确定编码方案所采用的参数值。

诚然,编码方案需要借助一些冗余数据。增大冗余数据,可以增强数据包传递可靠性,但冗余数据包的传输会带来额外的能量消耗,也会增加原数据包的传递时延。反之,如果减少冗余数据,则会削弱数据包传递可靠性。因此,在设计一个低能耗、低时延且可靠性高的数据包传输协议时,需要在能耗、时延和可靠性三者之间作出权衡。

事实上,在基于 IEEE 802.15.4 标准的 WSN 中,节点的处理能力、无线电覆盖范围、可以通信的数据包长度、内存空间等均为捉襟见肘。例如,IEEE 802.15.4 标准是针对低数据率、低功耗的无线个域网,其物理层可以携带的来自数据链路层数据帧的最大长度只有 127 字节^[1];当 MAC 层使用安全头部时,数据链路层能够携带的来自网络层数据长度只剩下 81 字节^[3]。又如,CrossBow 公司生产的 TelosB 传感器节点,只有 10 kB 内存,且兼容 IEEE

802.15.4 标准的射频收发器的通信距离仅为数十米。遗憾的是，绝大多数已有研究成果未能充分考虑 WSN 节点的上述资源瓶颈，因而难以应用于实践。

我们注意到 Reed-Solomon 编码^[4]具有以下显著的优点：1) 接收方一旦收到固定个数的编码包，就能解码得到原始信息；2) 传输开销小，编码方式简单。这些特点易于在资源极度受限的 WSN 上实现。虽然在编码、数据包传递可靠性、数据收集时延、能耗等研究领域已有大量成果，但这些成果存在着缺陷：未能将编码、可靠传递、时延、能耗这四者综合起来统筹优化。因此，需要一个将四者作为一个整体来统筹优化的解决方案。此乃本文的主要动机所在。

本文应用 Reed-Solomon 码进行数据包的可靠传输，主要贡献如下：

(1) 详细描述了利用简单 Reed-Solomon (Simple Reed-Solomon, S-RS) 码以及段内编码或段间编码的“S-RS 数据收集策略”；证明了 S-RS 码中任意 n 个码字 (Code Word) 均为线性独立；推导了 S-RS 数据收集策略在一个数据收集周期内的总能耗、数据收集率和数据收集时延。

(2) 设计了数据收集率和数据收集时延约束下能耗最小化这一优化问题，通过优化 S-RS 码的冗余度、每个编码包携带的码字个数和 IEEE 802.15.4 标准 MAC 层数据包的重传次数这三个参数，可以达到：在事先给定的数据收集率和数据收集时延要求下，通过最优编码方案，使整个网络数据收集产生的能耗最小这一目标。

(3) S-RS 数据收集策略所采用的 S-RS 码具有属性：其前几个码字为原始信息符 (Original Information Symbol, OIS)，使接收方不用解码就能获得 OIS，当接收方收到的码字不足以解码出全部 OIS 时仍可以得到部分数据。段内编码和段间编码策略在每个编码包中携带多个 S-RS 码字，以此减少编码包的发送个数，提高 MAC 帧的利用率，从而降低能耗。

本文其余部分内容编排如下：第 2 节分析相关研究成果；第 3 节介绍 Reed-Solomon 编码和 S-RS 码；第 4 节叙述 S-RS 码的段内编码和段间编码方案，以及 S-RS 数据收集策略的具体步骤；第 5 节推导了使用段间编码的 S-RS 数据收集策略的能耗、数据收集率和数据收集时延，给出时延、数据收集率约束下的能耗最小化这一优化问题；第 6 节给出

在实际的传感器节点中进行实验的结果；第 7 节通过仿真分析 S-RS 数据收集策略的性能；第 8 节结束本文。

2 相关研究

在提高数据包传递可靠性方面，常见的方法有：在数据链路层采用重传和确认机制^[5]，在网络层采用多路径路由^{[6][7]}，机会路由^[8]等。在无线网络中利用编码技术来实现数据包的可靠传递方案如雨后春笋，不断涌现，所用的编码有 LT 码(Luby Transfor)^[9]、Tornado 码^[10]、Reed-Solomon 码^{[2][4]}等。

文献[11]研究了在使用多输入多输出(MIMO)情况下如何最小化能耗的问题。文献[12]提出了 RMER 和 RMECR 两种路由，其中 RMER 找到一条端到端的总能耗最小的路径，而 RMECR 则选择满足可靠性要求下的能耗最小路径。文献[13]研究了如何在低占空比 WSN 中使用网络编码来降低瓶颈区域的能耗以此延长网络生存时间这一问题。

依数据相关性，传感数据的收集方案可以分为数据可融合的收集^{[14][15][16]}和数据不可融合的收集^{[17][18]}。在数据可融合的收集应用中，可以把多个节点所感知的数据采用诸如最大值、平均值等多种方法进行整合，以此降低网络中节点转发数据包的数量，进而降低带宽的开销和节点的能耗。例如，MERIG 协议^[14]使用数据融合技术，自适应地传输冗余数据包，使得在保证可靠性的前提下能耗得以最小化。然而，在 WSN 的大量应用中，传感器节点所捕获到的数据相关性甚微，甚至是不相关的，因此，这类数据是不可融合的。对于此类应用，节点不能对数据进行融合，而需要原封不动地把来自其它节点的数据转发给邻居节点。在数据不能融合的 WSN 节能问题研究领域，也取得了一些成果。例如，考虑到流过离 sink 节点越近的节点的数据量越大，这些节点会越早因能量耗竭而死亡，MITT^[17]根据节点剩余能量构造一棵最小生成树使得瓶颈节点的子节点数最小，从而平衡 sink 附近节点的能耗，以最大化网络生存时间。文献[19]提出了能耗经济的延长网络生存时间的算法 ERAPL，通过构造数据收集序列 DGS(Data Gathering Sequence)让每个节点按照比例将数据发送给邻居节点，以此达到节能的目的。文献[20]提出了一种能量高效的基于线性回归的无线传感器网络分布式数据采集优化

策略,以降低网络的总能耗。文献[21]将无线传感器网络的传感数据以树的形式收集到根。文献[22]提出了一种负载均衡感知的无线传感器网络容错分簇算法。

考虑到当代传感器网络节点可以是采用 Wi-Fi 使能的传感器 (Wi-Fi enabled sensor) [23] 节点, 本文作者改进了基于定时器的电源管理策略 TPM (Timer-based Power Management) [24], 提出了增强的基于定时器电源管理策略 E-TPM [25], 使得传感器节点能够经济地消耗能量。此外, 我们利用网络编码技术, 降低 WSN 的路由能耗 [26]; 利用非前缀信源编码技术, 降低纳米传感器网络发送方的能耗 [27]。

在利用编码提高传输可靠性方面, 可以应用 LT 码 [9]、Raptor 码 [28] 及其它随机系数的编码。但 LT 码 [9]、Raptor 码 [28] 以及采用随机系数的编码存在着一些不足之处, 如接收方一旦接收到一个数据包, 就尝试解码, 而且接收方可能因接收到的数据包的系数矩阵不满秩而无法解码。目前, 在应用 Reed-Solomon 编码于无线传感器网络可靠数据传输领域, 也有一些研究成果。文献 [29] 提出一种应用于 WSN 的 Reed-Solomon 码解码器算法以降低能耗; 在有多个 sink 节点的 WSN 应用中, 文献 [30] 利用 Reed-Solomon 码将数据包分片编码并在多条路径上把编码包传递给 sink 节点; 文献 [31] 提出了 Reed-Solomon 码的硬件实现方法; 文献 [32] 提出了可靠数据传输方案 RDTS (Reliable Data Transfer Scheme), 节点采用 Reed-Solomon 编码逐跳将数据传输到信宿。文献 [33] 提出了利用 Reed-Solomon 编码的可靠数据收集方案, 给出了数据收集率约束下的能耗最小化问题。

本文也采用 Reed-Solomon 码提高数据传递的可靠性, 本文策略与之前采用 Reed-Solomon 码的数据收集策略的明显异同之处在于: 将数据包编码方案、能耗、数据收集率、数据收集时延整合为一个优化问题, 对所涉及的指标进行统筹优化。

3 一种简单的 Reed-Solomon 码

Reed-Solomon 码是基于符号的编码 [4]。以 $GF(q) = \{0, 1, \alpha, \alpha^2, \dots, \alpha^{q-2}\}$ 表示伽罗华有限域, 其中 $q \geq 2$ 是整数, α 是一个素数, 满足 $\alpha^{q-1} = 1$ 和 $\alpha^i \neq \alpha^j$ ($1 \leq i < j < q-1$)。

在有限域 $GF(q)$ 中, 对于给定的 n 个 ($n < q$)

信息符 (即 OIS): m_0, m_1, \dots, m_{n-1} , 构建如下 $n-1$ 阶多项式 [4]

$$f(x) = \sum_{i=0}^{n-1} m_i x^i \quad (1)$$

利用式(1), 可以得到一个 q 维 Reed-Solomon 码 [4]: $(f(0), f(1), f(\alpha), f(\alpha^2), \dots, f(\alpha^{q-2}))$ 。这个 Reed-Solomon 码的 q 个码字有一个特性: 任意 n 个码字的信息符系数向量是线性无关的。例如, 对于前 n 个码字 $f(0), f(1), f(\alpha), f(\alpha^2), \dots, f(\alpha^{n-2})$, 从式(1)可知, 这些码字对应的信息符 m_0, m_1, \dots, m_{n-1} 的系数向量依次是: $(1, 0, \dots, 0), (1, 1, \dots, 1), (1, \alpha, \alpha^2, \dots, \alpha^{n-1}), (1, \alpha^2, \alpha^4, \dots, \alpha^{2(n-1)}), \dots, (1, \alpha^{n-2}, \alpha^{2(n-2)}, \dots, \alpha^{(n-1)(n-2)})$, 而且, 这 n 个向量组成了一个其值不为 0 的范德蒙行列式, 因此, 它们是线性无关的。

Reed-Solomon 码的上述性质说明: 接收方一旦收到 Reed-Solomon 码中任意 n 个码字, 就可以通过解一个满秩的线性方程组得到原来的 n 个信息符。这是因为在 q 维 RS 码 $(f(0), f(1), f(\alpha), f(\alpha^2), \dots, f(\alpha^{q-2}))$ 中所选取的 n 个码字中, 原来的 n 个信息符 m_0, m_1, \dots, m_{n-1} 的系数行列式是一个 n 阶范德蒙行列式, 其值不为 0。于是, 在 WSN 的数据包传递过程中, 我们可以选择一个合适的 q , 将原始 n 个信息符 m_0, m_1, \dots, m_{n-1} 编码生成一个 q 维 RS 码, 并将编码之后的数据包发送给下一个节点。这样, 接收方只要收到其中任意 n 个编码包就可以恢复出原始 n 个信号符。

与文献 [33] 一样, 我们引入向量: $v_i = (1, \alpha^i, \alpha^{2i}, \alpha^{3i}, \dots, \alpha^{(n-1)i})$, $i = 1, 2, \dots, q-2$ 。此外, 引入具有 n 个元素的行向量 e_i , 它满足条件: 第 i 个元素为 1, 其余 $n-1$ 个元素均为 0, 即 $e_i = (\overbrace{0, \dots, 0}^{i-1}, \overbrace{1, 0, \dots, 0}^{n-i})$, $i = 1, 2, \dots, n$ 。显然, e_1, e_2, \dots, e_n 构成一个 $n \times n$ 单位矩阵。然后, 选取自然数 M , 使它满足 $n < M \leq q-1$, 并构建 $M \times n$ 矩阵如下 [33]:

步骤 2. 初始化 $\text{Num_trial} = 0$, $\text{Idx_word} = 1$, $\text{Num_rcv} = 0$.

步骤 3. 在 M 个 S-RS 码字中取出第 Idx_word 个码字, 将之封装成一个 IEEE 802.15.4 帧, 把所组成的帧发送给父节点, 并设置帧超时定时器的值为 ω_{ack} .

步骤 4. 如果在 ω_{ack} 时间间隔内未能收到来自父节点的 ACK 包(确认包), 则重传数据包, 直至发送次数达到 $\sigma+1$. 否则(说明父节点已经正确接收到所发送的数据包), 将已经被接收的数据包个数 Num_rcv 加 1, 并判断等式 “ $\text{Num_rcv} = n$ ” 是否成立. 若成立, 则说明父节点已经收到 n 个码字, 足以解码, 因而停止发送码字; 否则(说明父节点因未收足 n 个码字而无法解码), 取出下一个码字继续发送.

从图 1 可见, 在 S-RS 数据收集策略中, 树枝节点的 MAC 层采用 IEEE 802.15.4 的重传和确认机制. 发送方只须收到其父节点的 n 个确认包就停止发送剩余的码字.

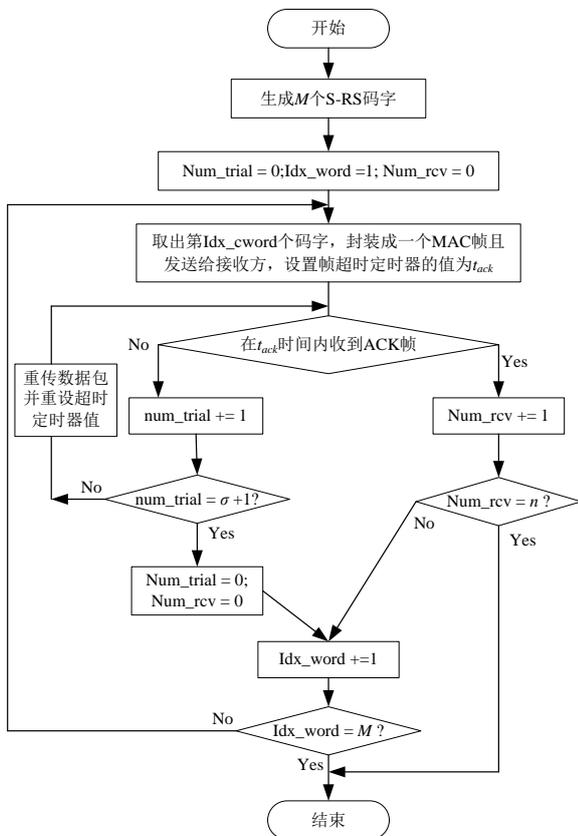


图 1 树枝节点发送数据包流程图

从式(4)可以看出: 码字 c_1, c_2, \dots, c_n 就是源信息符号 m_0, m_1, \dots, m_{n-1} . 因此, 接收方在接收到包含码字 c_1, c_2, \dots, c_n 的数据包时, 不需要解码. 在接收到

n^* 个原始信息符号 ($n^* \leq n$) 后, 接收方将之代入到式(4), 就可以得到包含其余 $n-n^*$ 个信息符的 $n-n^*$ 元满秩线性方程组, 通过高斯消元法解码就可以得到未知的 $n-n^*$ 个信息符. 因此, 在理想的链路状态下(如链路数据包传递率接近 1), 本文的 S-RS 数据收集策略只需要传递 n 个码字 c_1, c_2, \dots, c_n 即原始信息符号 m_0, m_1, \dots, m_{n-1} , 不需要增加额外的编码开销.

在 S-RS 数据收集策略中, 取

$$\begin{cases} M = \lambda n; \\ n \leq (q-1)/\lambda; \\ r \geq \log_2(\lambda n + 1) \end{cases} \quad (6)$$

其中 λ 为正整数, 称为冗余度, 且令 $q = 2^r$.

这样, 有限域 $\text{GF}(q)$ 变为 $\text{GF}(2^r)$, 它的每个元素占 r 比特. 我们称有限域 $\text{GF}(2^r)$ 的任意一个元素为一个信息单元 IU (Information Unit). 此外, 设节点要发送的传感数据为 X 个, 每个传感数据占 θ 个 IU, 因而, 每个节点需要发送的传感数据量为 $X\theta$ 个 IU.

S-RS 数据收集策略采用两种编码方案: “段内编码” 和 “段间编码” [33]. 这两种方案的共同点在于: 首先对要发送的 X 个传感数据进行分段, 使得每段包含 x 个传感数据, 共产生 $s = \lfloor X/x \rfloor$ 段, 记为 D_1, D_2, \dots, D_s . 显然, 每个段有 $f = x\theta$ 个 IU, 用 w_{ij} 表示段 D_i 的第 j 个 IU, $i = 1, 2, \dots, s, j = 1, 2, \dots, f$. 在 X/x 不是整数时, 产生 $s+1$ 段, 如图 2 所示. 为了减小发送的比特数, 在末段所含传感数据不足 x 时, 我们不采用末尾填补 0 的方式, 而是将该段单独使用 ARQ 发送给父节点.

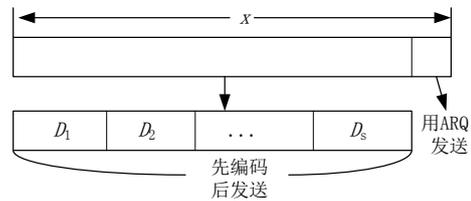


图 2 分段方案

在段内编码方案中, 编码操作是逐段进行的, 即在每个段内的 IU 之间进行, 它将 D_i 的 f 个信息符 $w_{i1}, w_{i2}, \dots, w_{if}$ 作为 S-RS 的编码输入, 使用式(4)编码, 生成 M 个 S-RS 码字. 具体地说, 对于 D_i , 用 c_{ij} 表示由 D_i 编码得到的第 j 个码字, 则段内编码方案生成的 M 个 S-RS 码字如下:

$$c_{ij} = \begin{cases} w_{ij}, & (1 \leq j \leq f); \\ w_{i1} + w_{i2}\alpha^{(j-f)} + w_{i3}(\alpha^{(j-f)})^2 + \dots + w_{if}(\alpha^{(j-f)})^{(f-1)}, & (f+1 \leq j \leq M) \end{cases} \quad (7)$$

其中 M 的值由式(6)确定, 且取 $n = f, i = 1, 2, \dots, s$ 。段内编码方案的编码结果如图 3 所示。

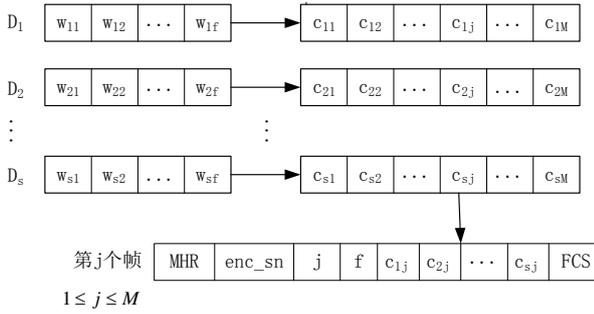


图 3 段内编码

当 X 个传感数据的 s 个段都编码完成后, 式(7)产生由 S-RS 码字组成的矩阵 $(c_{i,j})_{s \times M}$, 编码节点把这个矩阵的每个列向量封装成一个 IEEE 802.15.4 帧发送出去。例如, 第 j 个列向量 $(c_{1j}, c_{2j}, \dots, c_{sj})^T$ 封装成的 MAC 帧格式如图 3 的最后一行所示, 其中 MHR 和 FCS 分别为 IEEE 802.15.4 MAC 帧头部和校验和, enc_sn 为数据包序号(长度为 1 B), j 为编码向量的下标(长度为 1 B), f 为段的信息符个数(长度为 1 B)。

在段间编码方案中, 编码操作是跨段进行的, 即把各段处于相同位置的 IU 进行 S-RS 编码。用 h_{ij} 表示由段 D_1, D_2, \dots, D_s 中第 j 个 IU 编码得到的 S-RS 码的第 i 个码字, 那么段间编码方案所生成的码字满足:

$$h_{ij} = \begin{cases} w_{ij}, & (1 \leq i \leq s); \\ w_{1j} + w_{2j}\alpha^{i-s} + w_{3j}(\alpha^{i-s})^2 + \dots + w_{sj}(\alpha^{i-s})^{(s-1)}, & (s+1 \leq i \leq M) \end{cases} \quad (8)$$

其中 M 的值由式(6)确定, 且取 $n = s, j = 1, 2, \dots, f$ 。段间编码方案的编码结果如图 4^[33]所示。

当 X 个传感数据的 s 个段都编码完成后, 式(8)产生由 S-RS 码字组成的矩阵 $(h_{i,j})_{M \times f}$, 编码节点把这个矩阵的每个行向量封装成一个 IEEE 802.15.4 MAC 帧发送出去。例如, 第 i 个行向量 $(h_{i1}, h_{i2}, \dots, h_{if})^T$ 封装成的 MAC 帧格式如图 4 的最后一行所示。其中, i 为编码向量下标(长度为 1 B), s 为段数(长度为 1 B)。

上述两种编码方法有各自的特点。段内编码的特点: 对于给定的 x , 每个帧携带的码字个数为 s , 它随 X 变化而变化, 但可发送的编码包的个数 M 不变; 每个段内前 f 个码字是原始信息符(即 OIS)。由于一个传感数据包含 θ 个 IU, 并且每个帧中连续存

放的码字来自不同段的数据, 所以不能保证对于某个收到的段, 其前 f 个码字是连续的。段间编码的

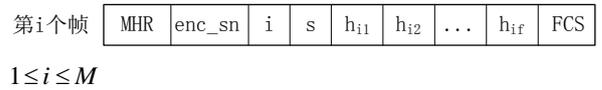
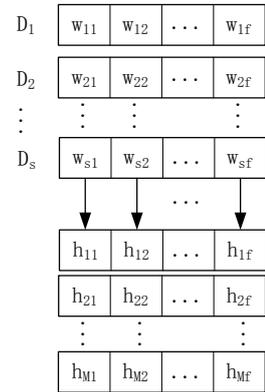


图 4 段间编码

特点是: 对于给定的 x , 每个帧携带的码字个数是固定的, 这是因为一个帧携带的码字个数为 $f = x\theta$ (θ 一般是一固定常数), 但可发送的编码包的个数 M 与 s 有关; 每个帧中的码字是连续的。

5 S-RS 数据收集策略的能耗、时延及其优化

下面以采用段间编码方式为例, 推导本文提出的 S-RS 数据收集策略的总能耗、数据收集率、数据收集时延(在采用段内编码时, 推导过程类似, 此处不予赘述)。考虑到在一些 WSN 应用中, sink 节点每隔一段时间收集数据, 以下称 sink 一次收集数据时间段为一个“数据收集周期”。

我们先分析节点将自己缓存的传感数据采用上述 S-RS 数据收集策略发送给其父节点收发双方产生的总能耗, 并由此导出整个网络的总能耗、sink 节点的数据收集率、sink 节点收集数据的时延。考虑到目前基于树的数据收集协议如 CTP (Collection Tree Protocol) ^[36]被广泛应用于无线传感器网络中, 本文做如下假设: 在所研究的 WSN 中, 已经用现成的协议构建了一棵树, 节点沿着这棵树将收集到的数据传递到 sink 节点。

此外, 我们假设^[33]: 每个节点配置了相同型号的传感器, 每个传感器在每一数据收集周期内产生

的数据量相同, 均为 $L_0 \leq \Gamma$, 其中, Γ 为 IEEE 802.15.4 MAC 帧的有效载荷, $L_0 = \theta r$, Γ 和 L_0 的单位均为比特。也就是说, 每个节点在一个数据收集周期内产生的数据量可以由一个 MAC 帧携带。

本文用 H 表示每个帧的 MAC 协议层头部(包括 FCS)长度(单位:比特)。

5.1 S-RS数据收集策略的能耗

本文采用文献[37]的能耗模型, 即当一个节点发送 l 比特数据给相距为 d 的另一节点时, 发送方的能耗为

$$E_{tx}(l, d) = l(\varepsilon_0 + \varepsilon_1 d^\gamma) \quad (9)$$

接收方的能耗

$$E_{rx}(l) = l\varepsilon_0 \quad (10)$$

其中, γ 为路径衰减指数, 在 2~4 之间取值, ε_1 为发送单位比特无线电放大器的能耗, ε_0 为编码、调制单位比特的能耗[37]。

以 Ω 表示 WSN 全体节点的集合, T 表示数据收集树, Ω_0 表示全体树叶节点的集合。对于节点 $k \in \Omega$, 以 p_k 表示它与其父节点间的无线链路传递数据包成功概率, d_k 表示它与其父节点的距离, L_k 表示它发送的数据包长度; X_k 表示它发送的传感数据个数, 特殊的, 当 k 为 sink 节点时, X_k 表示即为一个数据收集周期内到达 sink 的传感数据个数。此外, 以 $CN(k)$ 表示节点 k 的子节点集合, σ 表示 IEEE 802.15.4 MAC 层设定的帧的最大重传次数。

首先, 分析节点 k 使用自动重传机制即 ARQ 将 IEEE 802.15.4 MAC 帧(以下 MAC 帧与数据包不加区别)发送给父节点的能耗。易知, 节点 k 在给其父节点发送(包括重发)一个数据包时, 它的发送次数及其对应的概率如表 1 所示。

表 1 节点 k 发送次数及其对应概率

发送次数	对应的概率	成功与否
1	p_k	成功
2	$(1-p_k)p_k$	成功
3	$(1-p_k)^2 p_k$	成功
...	...	成功
σ	$(1-p_k)^{\sigma-1} p_k$	成功
$\sigma+1$	$(1-p_k)^\sigma p_k$	成功
$\sigma+1$	$(1-p_k)^{\sigma+1}$	失败

由表 1 可知, 在 MAC 层的最大重传次数为 σ 时,

节点 k 把一个数据包成功传递给父节点的概率为

$$P_k^{(S)} = 1 - (1-p_k)^{\sigma+1} \quad (11)$$

根据表 1 以及式(9)和(10)可以得到, 节点 k 在发送一个长度为 L_k 比特的 IEEE 802.15.4 MAC 帧给它的父节点时, 收发双方所产生的平均能耗为[33]:

$$E_k^{ARQ}(L_k) = P_k^{(S)} E_k^{(S)}(L_k) + (1-P_k^{(S)}) E_k^{(F)}(L_k) \quad (12)$$

式中, $E_k^{(S)}(L_k)$ 表示在传输成功的情况下收发双方产生的平均能耗:

$$\begin{aligned} E_k^{(S)}(L_k) &= [E_{tx}(L_k, d_k) + E_{rx}(L_k)] \sum_{i=1}^{\sigma+1} \frac{i(1-p_k)^{i-1} p_k}{P_k^{(S)}} \\ &\quad + E_{tx}(\beta_{ACK}, d_k) + E_{rx}(\beta_{ACK}) \\ &= L_k(2\varepsilon_0 + \varepsilon_1 d_k^\gamma) \frac{1 - (\sigma+2)(1-p_k)^{\sigma+1} + (\sigma+1)(1-p_k)^{\sigma+2}}{p_k [1 - (1-p_k)^{\sigma+1}]} \\ &\quad + 40(2\varepsilon_0 + \varepsilon_1 d_k^\gamma) \end{aligned} \quad (13)$$

其中, β_{ACK} 表示 ACK 包的长度, 占 5B[1], 即 40 比特。在式(13)中, 第一行是收发数据时双方产生的能耗, 求和符号后面的式子表示在发送成功条件下所需要的平均发送次数; 第二行是双方收发 ACK 包产生的能耗。此外, $E_k^{(F)}(L_k)$ 表示节点 k 在发送数据包失败的情况下(此时共发送了 $\sigma+1$ 次), 所产生的能耗:

$$\begin{aligned} E_k^{(F)}(L_k) &= (\sigma+1)[E_{tx}(L_k, d_k) + E_{rx}(L_k)] \\ &= (\sigma+1)[L_k(2\varepsilon_0 + \varepsilon_1 d_k^\gamma)] \end{aligned} \quad (14)$$

其次, 分析采用 S-RS 策略时节点 h 发送数据包收发双方所产生的能耗。根据段间编码方式, 节点 h 将 X_h 个传感数据分成 s 段, 每段包含 x 个传感数据即 $f=x\theta$ 个 IU。这样, 剩余 $X_h^* = X_h \bmod x$ 个传感数据, 其中, \bmod 表示整数的求模数运算。由于采用段间编码, 此时 $n=f$, 根据式(6), 节点 h 最多可发送 $M_h = \lambda f$ 个编码包, 其中前 s 个包中包含的码字是未经编码的源信息符(以下称之为“明文”), 即第 i 个编码包的数据为原始第 i 个段的数据($i \leq s$)。节点 h 发送的编码包长度为 $L_h = L_{phy} + H + fr + 24$ 。其中, L_{phy} 是物理层头部(包括同步前缀)长度, 它依赖于所采用的物理层; 数值 24 是编码序号 enc_sn (1 字节)、编码向量下标 i (1 字节)和段数 s (1 字节)所占的比特数(见图 4)。

由于节点 h 将 s 段数据编码封装成 M_h 个编码

包，每个编码包采用 ARQ 发送，使得其父节点只需成功接收到 s 个编码包就能够解码出原始 s 段数据。为了使父节点收到 s 个编码包以成功解码，节点 h 发送的编码包个数及其对应的概率如表 2 所示，其中， $P_h^{(s)}$ 是由式(11)确定的节点 h 成功发送一个编码包给其父节点的概率。

表 2 节点 h 发送编码包个数及其对应概率

发送编码包个数	对应的概率	父节点是否成功解码
s	$[P_h^{(s)}]^s$	是
$s+1$	$C_s^1(1-P_h^{(s)})[P_h^{(s)}]^s$	是
$s+2$	$C_{s+1}^2(1-P_h^{(s)})^2[P_h^{(s)}]^s$	是
...		
M_h-1	$C_{M_h-2}^{M_h-s-1}(1-P_h^{(s)})^{M_h-s-1}[P_h^{(s)}]^s$	是
M_h	$C_{M_h-1}^{M_h-s}(1-P_h^{(s)})^{M_h-s}[P_h^{(s)}]^s$	是
M_h	$(1-P_h^{(s)})^{M_h}$	否

由于节点 h 所发送的 M_h 个编码包中有 s 个编码包到达其父节点的概率 G_h 等于随机事件“在父节点成功接收到 s 个编码包时，节点 h 恰好发送 i 个数据包” ($i = s, s+1, \dots, M_h$) 的概率之和，由表 2 可得

$$G_h = \sum_{i=s}^{M_h} C_{i-1}^{s-1} (P_h^{(s)})^s (1-P_h^{(s)})^{i-s} \quad (15)$$

且节点 h 发送给父节点的编码包个数均值为

$$a_h = \sum_{i=s}^{M_h} i C_{i-1}^{s-1} [P_h^{(s)}]^s [1-P_h^{(s)}]^{i-s} + M_h [1-P_h^{(s)}]^{M_h} \quad (16)$$

因此，在节点 h 发送编码包给父节点时收发双方产生的平均能耗(数学期望)为

$$ET_h = \left[G_h \sum_{i=s}^{M_h} i C_{i-1}^{s-1} \frac{(P_h^{(s)})^s (1-P_h^{(s)})^{i-s}}{G_h} + (1-G_h)M_h \right] E_h^{ARQ}(L_h) \quad (17)$$

其中，方括号内第 1 项表示在父节点成功接收到 s 个段这个条件下节点 h 发送的编码包个数(均值)，第 2 项是在父节点未能成功接收到 s 个数据段条件下节点 h 发送的编码包个数(这时 M_h 个编码包都被发送)。 $E_h^{ARQ}(L_h)$ 是发送一个编码包的能耗，由式(12)确定。

在 S-RS 数据收集策略中，当剩余的传感数据 $X_h^* > 0$ 时，节点 h 使用 ARQ 发送这 X_h^* 个传感数据，收发双方产生的平均能耗 ET_h^* 可以由式(12)获得 [33]；

$$ET_h^* = (1-p_h)^{\sigma+1} (\sigma+1)L_h(2\varepsilon_0 + \varepsilon_1 d_h^\gamma) + [1-(1-p_h)^{\sigma+1}] \left\{ 40(2\varepsilon_0 + \varepsilon_1 d_h^\gamma) + \tilde{L}_h(2\varepsilon_0 + \varepsilon_1 d_h^\gamma) \cdot \frac{[1-(\sigma+2)(1-p_h)^{\sigma+1} + (\sigma+1)(1-p_h)^{\sigma+2}]}{p_h[1-(1-p_h)^{\sigma+1}]} \right\} \quad (18)$$

其中， \tilde{L}_h 是 X_h^* 个传感数据所占的比特数与帧头部比特数之和，即 $\tilde{L}_h = L_{phy} + (X_h - sx)\theta + H$ 。

对于任意一个节点 z ，以 E_z 表示在采用 S-RS 数据收集策略时收集以节点 z 为根的子树上所有节点数据所产生的能耗。我们有下述递归式子 [33]：

$$E_z = \sum_{h \in CN(z)} [E_h + \tilde{E}_h], \quad z \in \Omega \setminus \Omega_0 \quad (19)$$

其中，当 $h \in \Omega_0$ 时， \tilde{E}_h 由式 (12) 计算，即 $\tilde{E}_h = E_k^{ARQ}(L_k)$ ；当 $h \in \Omega \setminus \Omega_0$ 时， \tilde{E}_h 由下式计算：

$$\tilde{E}_h = ET_h + ET_h^* \quad (20)$$

它表示节点 h 发送 X_h 个传感数据，收发双方产生的能耗均值。

特别的，当 z 为 Sink 节点时， E_{Sink} 即为一数据收集周期数据收集的总能耗。

最后，举一个例子来说明，如何利用递归式子(19)计算节点 z 为根的子树上所有节点传递数据所产生的能耗 E_z 。假设节点 z 的所有子孙节点 $a, b, c, a_1, a_2, c_1, c_2, c_3$ 构成的拓扑结构如图 5 所示。这样， z 的子节点集合 $CN(z) = \{a, b, c\}$ ，由式(19)，得到 $E_z = (E_a + \tilde{E}_a) + (E_b + \tilde{E}_b) + (E_c + \tilde{E}_c)$ ，其中， \tilde{E}_a 和 \tilde{E}_c 因 a 和 c 是树枝节点由式(20)计算，而 \tilde{E}_b 因 b 是叶子节点由式(12)计算。此外，由于 $CN(b)$ 是空集，由式(19)知， $E_b = 0$ 。由于节点 a 有两个子节点 a_1 和 a_2 ，即 $CN(a) = \{a_1, a_2\}$ ，利用式(19)得 $E_a = (E_{a_1} + \tilde{E}_{a_1}) + (E_{a_2} + \tilde{E}_{a_2})$ 。由于 a_1 和 a_2 均为叶子节点，因此， $E_{a_1} = E_{a_2} = 0$ ，从而， $E_a = \tilde{E}_{a_1} + \tilde{E}_{a_2}$ ，且 \tilde{E}_{a_1} 和 \tilde{E}_{a_2} 均由式(12)计算。同理， $E_c = \tilde{E}_{c_1} + \tilde{E}_{c_2} + \tilde{E}_{c_3}$ ，且 \tilde{E}_{c_1} 、 \tilde{E}_{c_2} 和 \tilde{E}_{c_3} 由式(12)计算。综上所述，有： $E_z = (\tilde{E}_{a_1} + \tilde{E}_{a_2} + \tilde{E}_a) + \tilde{E}_b + (\tilde{E}_{c_1} + \tilde{E}_{c_2} + \tilde{E}_{c_3} + \tilde{E}_c)$ ，其中， \tilde{E}_a 和 \tilde{E}_c 由式(20)计算，其它项均由式(12)计算。

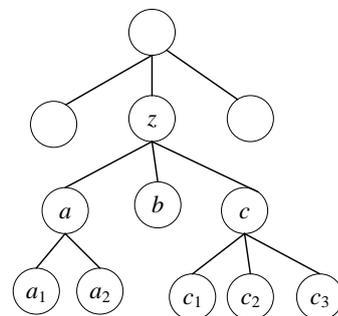


图 5 节点 z 为树根的子树

5.2 S-RS策略的数据收集率

我们把S-RS数据收集策略的数据收集率 ρ 定义为：在一个数据收集周期内，到达sink节点的数据包个数与网络中节点总数的比值，即平均每个节点传递到sink节点的数据包个数。

以 A_h 表示节点 h 所发送的 X_h 个传感数据中到达父节点的传感数据平均个数，则数据收集率可以表示为

$$\rho = \frac{\sum_{h \in CN(\text{sink})} A_h}{|\Omega|} \quad (21)$$

式中， $|\Omega|$ 是网络中节点个数。

下面推导 A_h 的计算公式。显然，在S-RS数据收集策略中，分段剩余下来的尾部 X_h^* 个传感数据在采用ARQ方式发送时发送成功的数据包平均个数为 $X_h^*[1 - (1 - p_k)^{\sigma+1}]$ 。对于参与编码的 xs 个传感数据，在采用S-RS数据收集策略时，发送成功的概率为 G_h ，故其发送成功的平均个数为 xsG_h 。由于在参与编码的 xs 个传感数据中，前 s 个帧中包含的码字是明文(即原来的信息符，不需要解码)，因此，即使接收方未能正确收齐 s 个帧而导致解码失败，接收方依然可以从收到的帧中获得部分不需要解码的原信息符。由于节点 h 产生的编码包个数为 $M_h = \lambda s$ ，因此，接收方在未能正确收齐 s 个帧导致解码失败的情况下，能够正确收到明文帧的所有可能情况为：接收方成功收到 i 个明文帧，其它 $M_h - i$ 个帧中最多被正确接收的帧数为 $s - i - 1$ 个($i = 1, 2, \dots, s - 1$)。考虑到在采用段间编码时，成功接收一个明文帧意味着成功接收到 x 个传感数据。于是，我们得到节点 h 所发送的 X_h 个传感数据中到达父节点的传感数据平均个数^[33]：

$$A_h = x \sum_{i=1}^{s-1} \left\{ i C_s^i (P_h^{(S)})^i \cdot (1 - P_h^{(S)})^{s-i} \cdot \sum_{j=0}^{s-i-1} C_{(\lambda-1)s}^j (P_h^{(S)})^j (1 - P_h^{(S)})^{(\lambda-1)s-j} \right\} \quad (22)$$

$$+ X_h^* [1 - (1 - p_k)^{\sigma+1}] + xsG_h$$

值得说明的是，对于任意一个节点 z ，它需要发送的传感数据平均个数为

$$X_z = \sum_{h \in CN(z)} A_h + 1 \quad (23)$$

其中，数字1表示节点 z 自己产生的一个传感数据。

5.3 S-RS策略的数据收集时延

为了支持WSN的一些实时应用，IEEE 802.15.4标准提出了超帧(superframe)结构^[1]，使节点能够在

超帧的免竞争时期(contention-free period, CFP)中使用其GTS(guaranteed time slots)发送数据。以下基于超帧结构，对S-RS的时延进行分析。

本文将节点A到节点B的数据包时延定义为：从节点A开始发送该数据包到节点B的MAC层接收完该数据包这段时间。自然地，sink节点收集数据的时延定义为，网络中任一节点与sink节点的端到端时延的最大者，以 τ_{sink} 表示。

首先，分析节点 k 使用ARQ发送传感数据给父节点所产生的时延。在IEEE 802.15.4标准中，发送方用参数 $\text{macAckWaitDuration}$ ^[1]设置接收方返回ACK包的定时器值，即 $\omega_{\text{ack}} = \text{macAckWaitDuration}$ 。也就是说，如果发送方在 ω_{ack} 时间耗完后尚未收到接收方的ACK包，那么节点 k 就认为所发送的数据包丢失从而重传该数据包。在一个帧被正常接收时，接收方必须等待时间间隔 t_{ack} 才能回复ACK帧，其中， t_{ack} 的取值范围为 $[12, 32]$ ，其单位为符号时间(Symbol Period)^[1]。于是，我们获得结论：节点 k 在发送长度为 L_k 的数据包时，每失败一次，耗时 $L_k/R + \omega_{\text{ack}}$ ，其概率为 $1 - p_k$ ；每成功一次，耗时 $L_k/R + t_{\text{ack}} + \beta_{\text{ack}}/R + t_{\text{ifs}}$ ，其概率为 p_k 。其中， R 是信道的数据率(单位：bps)， β_{ack} 是ACK包长度(单位：比特)， t_{ifs} 是帧间间隔。因此，节点 k 每次发送数据包的平均耗时(数学期望)为

$$t_k^{\text{pkt}} = (1 - p_k) \left(\omega_{\text{ack}} + \frac{L_k}{R} \right) + p_k \left(\frac{L_k}{R} + t_{\text{ack}} + \frac{\beta_{\text{ack}}}{R} + t_{\text{ifs}} \right) \quad (24)$$

设超帧的时间长度为 ω ，在超帧中分配给节点 k 的GTS时间长度为 ω_k 。若节点 k 在超帧的GTS时间内不能成功发送一个数据包，则该数据包推迟到下一个超帧的GTS时间内发送，而且，MAC层允许发送该数据包的最大次数为 $\sigma + 1$ 。因此，我们需要计算节点 k 发送一个数据包所需的超帧个数。

由于节点 k 在一个超帧内的GTS时间长度为 ω_k ，故在这段时间内节点 k 可以发送(包括重传)数据包的平均次数为

$$n_k = \frac{\omega_k}{t_k^{\text{pkt}}} \quad (25)$$

由表1可知，节点 k 使用ARQ发送传感数据给父节点的平均发送次数为

$$ETX_k = \sum_{i=1}^{\sigma+1} i (1 - p_k)^{i-1} p_k + (\sigma + 1) (1 - p_k)^{\sigma+1} \quad (26)$$

因此节点 k 使用ARQ发送一个数据包所需要的GTS平均个数为

$$b_k = \text{ceil} \left(\frac{ETX_k}{n_k} \right) \quad (27)$$

其中，函数 $ceil(x)$ 表示不小于 x 的最大整数。

考虑到节点 k 在发送一个数据包时，若它在所分配的 GTS 时间内重传多次仍不能成功发送该数据包，那么，该数据包在下一个超帧的 GTS 时间内重传。由于在最后一个超帧内节点所分配的 GTS 开始时间可以在超帧的任意位置，我们取节点的 GTS 的开始时间为 $\omega/2$ 。因此，节点 k 使用 ARQ 发送一个数据包的平均时延为

$$T_k = (b_k - 1)\omega + \omega/2 + t_k^{pkt} \\ = [ceil(\frac{ETX_k}{n_k}) - \frac{1}{2}]\omega + t_k^{pkt} \quad (28)$$

其次，我们分析节点 z 使用 S-RS 策略收集其子节点数据的时延。若 z 的子节点为树叶时，树叶节点没有收集数据的时延，只有发送数据包的时延。若 z 的子节点 h 不是树叶时，则节点 h 在完成收集其子节点数据之后，将收集到的传感数据分成 s 段，然后根据段间编码的方式得到 S-RS 码字，并组装成 M_h 个 IEEE 802.15.4 MAC 帧，而且 MAC 帧的重传次数为 σ 。

由式(16)，节点 h 发送数据包的总次数为 $a_h ETX_h$ ，所需要的 GTS 平均个数为 $ceil(a_h ETX_h / n_h)$ ，因此，树枝节点 h 把编码包发送给其父节点 z 产生的时延为 $ceil(a_h ETX_h / n_h - 1)\omega + \omega/2 + t_h^{pkt}$ 。此外，在 X_h 不被 x 整除的条件下，该分段剩余下来的尾部 X_h^* 个传感数据采用 ARQ 方式发送，其时延由式(28)计算。因而，节点 h 把编码包发送给其父节点 z 产生的时延为

$$T_h = \omega [ceil(\frac{a_h ETX_h}{n_h}) + ceil(\frac{ETX_h}{n_h}) - 1] + 2t_h^{pkt} \quad (29)$$

其中， ETX_h, n_h, a_h 和 t_h^{pkt} 分别由式(26)，(25)，(16)和(24)计算。

对于任一节点 z ，用 τ_z 表示在以节点 z 为根的子树中，将所有子孙节点的数据收集到节点 z 的时延。因此， τ_z 等于节点 z 所有子节点收集其对应的子孙节点数据所耗的时间与该子节点发送数据给节点 z 的时间之和的最大者。换言之， τ_z 满足以下递归式：

$$\tau_z = \text{Max}_{u \in CN(z)} \{ \tau_u + T_u \} \quad (30)$$

其中，当 u 为叶节点时， $\tau_u = 0$ ， T_u 由式(28)计算；当 u 为树枝节点时， T_u 由式(29)计算。值得说明的是，由式(30)递归计算 τ_z 的方法与 5.1 节递归计算 E_z 的过程类似，请参见 5.1 节的例子。

这样一来，利用式(30)就可以通过递推得出 sink 节点收集数据的时延 τ_{sink} (即从 sink 节点开始收集数据到所有传感数据到达 sink 这个时间段)。

5.4 数据收集率和时延约束下的能耗最小化

接下来我们研究如何选择 S-RS 的参数使得在满足事先给定的数据收集率和时延约束下数据收集的能耗最低这一问题。显然，S-RS 数据收集策略的能耗、数据收集率、时延受到参数 λ ， x 和 σ 的影响。在 S-RS 策略中，采用 S-RS 编码虽然可以提高数据传递的可靠性，但编码、解码以及其它相关运算均会产生额外的能耗。我们注意到，在传感器节点中，能耗主要来自于节点之间的通信^[38]。而且，在 WSN 中，给距离 10 米或 100 米的节点发送 1 比特的耗能是计算操作的几千倍乃至几百万倍^[39]。因此，本文仅考虑节点之间的通信能耗即数据收集过程产生的能耗，而忽略编码、解码及其它计算所产生的能耗。于是，我们提出如下优化问题：

$$\begin{aligned} & \text{Min } E_{sink} \\ & \text{w.r.t. } x, \sigma, \lambda \\ & \left\{ \begin{array}{l} x = 1, 2, \dots, \lfloor (992 - H) / (\theta r) \rfloor; \\ E_z = \sum_{h \in CN(z)} [E_h + ET_h], \quad z \in \Omega \setminus \Omega_0; \\ E_k = ET_k, \quad k \in \Omega_0; \\ \sigma = 0, 1, \dots, 7; \\ \lambda = 2, 3, \dots, 7; \\ \rho > \rho_0; \\ \tau_{sink} < T_0 \end{array} \right. \quad (31) \end{aligned}$$

其中 E_{sink} 是关于 λ ， x 和 σ 的函数； Ω 表示全体节点的集合， Ω_0 表示全体叶节点的集合；重传次数 σ 的取值范围 0~7^[1]； λ 是编码的冗余倍数，由式(6)知 $\lambda \leq (q-1)/n$ ，这里 λ 的取值范围设置为 2~7；在 IEEE 802.15.4 中 MAC 帧的最大长度为 127 字节^[1]，因此 $H + x\theta r + 24 \leq 127 \times 8$ ，即 $x \leq (992 - H) / (\theta r)$ ； ρ_0 表示一个事先给定的传感数据收集率约束； T_0 表示一个事先给定的时延约束。

在优化问题(31)中，目标函数 E_{sink} 是由式(20)递推而得的把整个网络节点的数据收集到 sink 所产生的总能耗（参见第 5.1 节）； λ ， x 和 σ 是待优化的参数。式(31)的第 1 个约束条件表示，网络中任意一个叶子节点 k 发送发送编码包给其父节点时收发双方产生的平均能耗 E_k ，由式(12)计算。式(31)的

第2个约束条件表示,任意一个树枝节点 z 从它的所有子孙节点收集数据所产生的能耗 E_z ,可以采用递推式(19)计算(参见第5.1节)。在式(31)的最后一个约束条件中, τ_{sink} 由式(30)递推计算而得。

由于优化问题(31)中3个参数的可能取值个数为: λ 为6个, σ 为8个, x 不足20个(对于表3所列的参数, x 为12个)。此外,式(31)约束条件中迭代计算 E_z 需要和 N 次(N 为网络中节点个数);对于判断约束条件 $\tau_{sink} < T$,用式(30)递推计算 τ_{sink} 时,所涉及的求最大值至多需要比较 N 次。因此,利用穷举法求解优化问题(31)的复杂度为 $O(N)$,在多项式时间内可以获得优化问题(31)的解。

6 S-RS 数据收集策略的实验

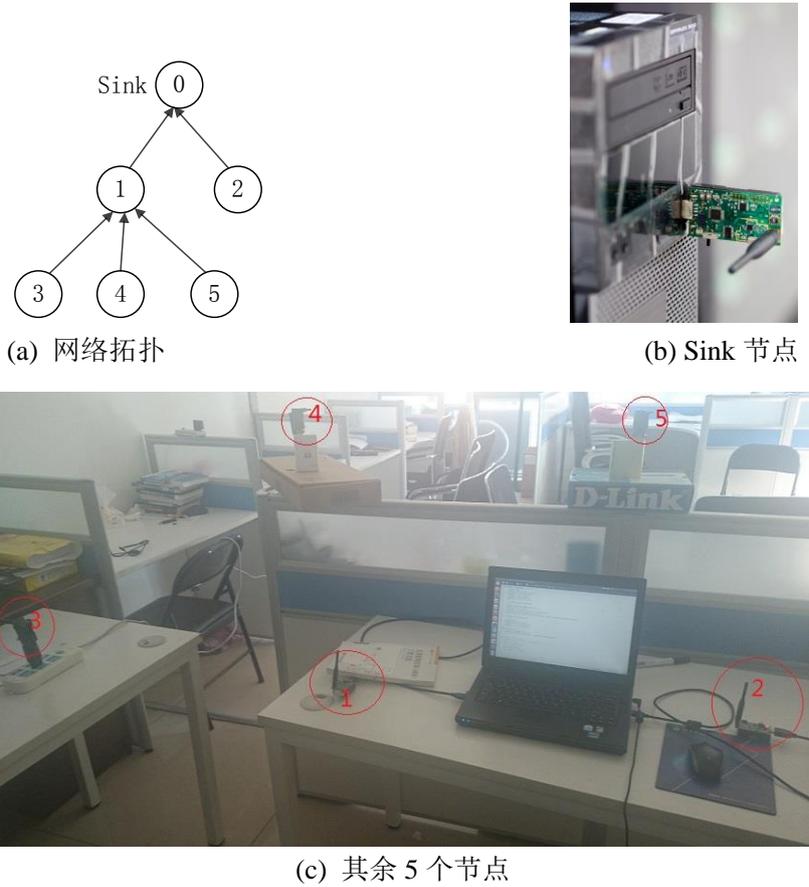
我们用 TinyOS 操作系统、NesC 程序设计语言和 TelosW 节点^[40]进行实验,以此说明本文提出的 S-RS 数据收集策略可以在资源受限的传感器节点上实现。TelosW 节点属于 Telos 系列,具有内置的能量表 (Energy Meter)^[41] 供程序读取。根据能量表的数值可以计算出节点的能耗^[42],不需要采用外接本身就耗能的设备来检测能耗。

将6个 TelosW 节点联成如图6(a)所示的拓扑结构,其中,节点0为 sink 节点。Sink 节点连接到台式机直接由台式机供电,其实物如图6(b)所示;其余5个节点的布置如图6(c)所示。每个 TelosW 节点每一轮产生10个传感数据,每个传感数据长度为10字节(即 $\theta = 10$)。每个段所包含的传感数据个

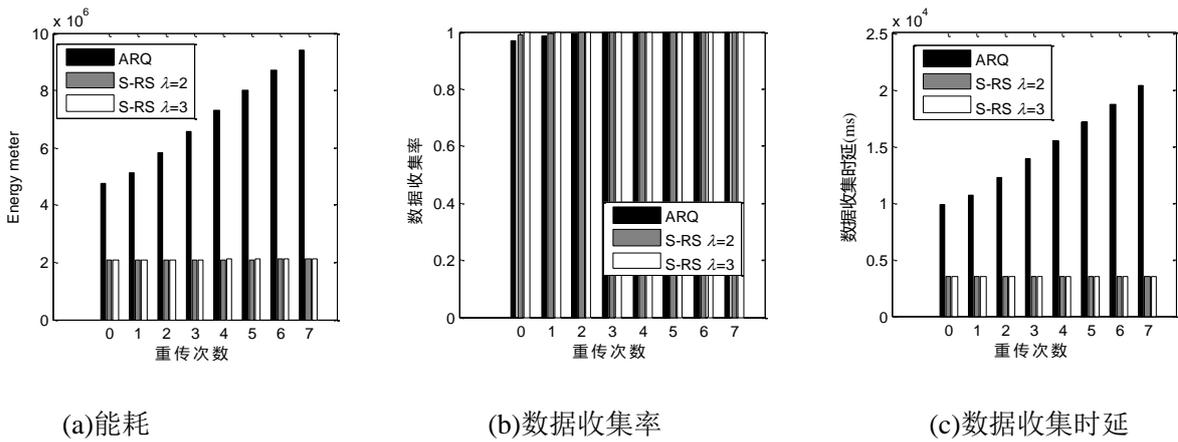
数 $x = 4$ 。图7是在不同重传次数下收集1000轮数据的平均值。从图7可以看出:1) S-RS 的能耗 (“Energy meter”) 远低于 ARQ 方案(见图7(a))。需要说明的是,Energy meter 的数值可以转化为能量,它与能量值存在着正比例关系:它的数值越大,对应的能量值也越大^[42]。2) 在重传次数(即 σ)的值比较大时(如 $\sigma \geq 2$)时,两者的数据收集率一致;但在 $\sigma < 2$ 时,S-RS 的数据收集率优于 ARQ 方案(见图7(b))。3) S-RS 的数据收集时延远低于 ARQ (见图7(c))。

7 性能分析

上一节我们已经在 TelosW 节点上对本文提出的 S-RS 方案进行了实验。事实上,真实环境的实验存在着困难之处,即难以对一些参数的变化进行模拟。例如,我们难以布置一对节点使它们之间链路的数据包传递成功率为一个事先给定的数值。因此,需要求助于仿真。本节,我们用 C++ 程序设计语言编制仿真程序对 S-RS 数据收集策略的性能进行分析。在仿真中,在 $1000 \times 1000 \text{ m}^2$ 区域里随机部署400个节点, Sink 节点放在区域中间;节点的通信范围为100m。数据收集树以 Sink 为根节点随机生成。本节所给出的结果是随机生成500棵树,每棵树上采用 S-RS 数据收集策略200次的平均值。



(a) 网络拓扑 (b) Sink 节点 (c) 其余 5 个节点
图 6 用 TelosW 节点的实验场景



(a) 能耗 (b) 数据收集率 (c) 数据收集时延
图 7 用 TelosW 节点的实验结果

与文献[37]一样，取 $\epsilon_1 = 10 \text{ pJ}/(\text{bit m}^2)$ ， $\epsilon_0 = 50 \text{ nJ}/\text{bit}$ ，且取 $\gamma = 2$ (对应于自由空间模型)。在 S-RS 编码中，选择信息单元 IU 的长度为 8 比特。每个传感数据包含 IU 的个数 $\theta = 8$ 。在 IEEE802.15.4 中，MAC 帧头部长度 $H = 200$ 比特(即 25 字节)，这样，在式(31)中 x 的最大值为 12。另外，假设物理层使用 O-QPSK(Offset- Quadrature Phase Shift Keying) 调制方案，其数据率 $R = 250 \text{ Kbps}$ ，此时一个符号时间为 $16 \mu\text{s}$ ， t_{lifs} 和 t_{ack} 分别为 0.64 ms 和 0.192 ms ，

接收 ACK 包的定时器时间值 ω_{ack} 为 0.704 ms ^[1]。

在 IEEE802.15.4 中，一个超帧结构由参数 $amacSuperframeOrder$ 和 $amacBeaconOrder$ 描述^[1]，当这两个参数取值相同时，超帧为活跃时间，一个超帧由 16 个时隙构成，每个时隙的长度为 $aBaseSlotDuration \times 2^{amacSuperframeOrder}$ ，其中 $aBaseSlotDuration = 960 \mu\text{s}$ 。为了保证 S-RS 策略在一个时隙中可以发送一个帧，我们取 $amacSuperframeOrder = 3$ 。在仿真中，超帧中不设

自由竞争周期 CAP(Contention Access Period)^[1], 即 16 个时隙全部用于分配 GTS。为清楚起见, 仿真采用的主要参数值列在表 3 中。

在进行性能分析之前, 我们通过仿真来验证第 5 节推导的分别以式(19)、(21)和(30)表示的总能耗 E_{sink} 、数据收集率 ρ 和数据收集时延 τ_{sink} 的正确性。图 8 是当 $x = 10, \lambda = 5, \sigma = 5$ 时在不同链路质量下仿真和计算结果及其误差图。图中, “S-RS-Sim” 和 “S-RS-Cal” 分别表示仿真和数值计算结果。在计算结果中, 总能耗 E_{sink} 、数据收集率 ρ 和数据收集时延 τ_{sink} 分别由式(19)、式(21)和式(30)计算而得。在图 8 中, 第二行分别是计算结果 v_c 与仿真结果 v_s 的相对误差, 定义为 $|v_c - v_s|/v_s$ 。从图 8 可见, 仿真结果与计算结果比较吻合且趋势一致, 因此, 在分析 S-RS 数据收集策略的性能时, 我们直接利用式(19)、式(21)和式(30)进行分析。

首先, 分析参数 x (每个编码包携带的传感数据个数) 和 λ (编码冗余度) 的变化对 S-RS 策略性能的影响。固定链路传递数据包成功率 $p_k = 0.6$, 让 x 变化, 得到图 9; 固定 $x = 10, \sigma = 3$, 让 p_k 变化, 得到图 10。从图 9(a)和图 9(c)可见: 随着 x 增大, 能耗和时延逐渐降低, 这是因为发送的编码包个数减小了。从图 9(b)可见: 随着 x 增大, 数据收集率逐渐降低, 这是因为, 在 x 增大之后每个编码包中包含的传感数据个数也增大, 这导致在出现丢包时丢

表 3 仿真参数及其值

参数名	值
区域	1000 × 1000 m ²
节点数	400
通信范围	100 m
ϵ_l	10pJ/(bit m ²)
ϵ_0	50nJ/bit
γ	2
r	8
θ	8
H	200 bit
R	250Kpbs
β_{ack}	40 bit
t_{ifs}	0.64 ms
t_{ack}	0.192 ms
ω_{ack}	0.704 ms
$amacSuperframeOrder$	3
$amacBeaconOrder$	3
一个 Slot 时间	7.68 ms

失的传感数据也增大, 进而导致数据收集率下降。从图 10 可见: 对于一个给定的 p_k , 编码冗余度 λ 增大到一定值之后, 继续增大 λ 不能带来数据收集率

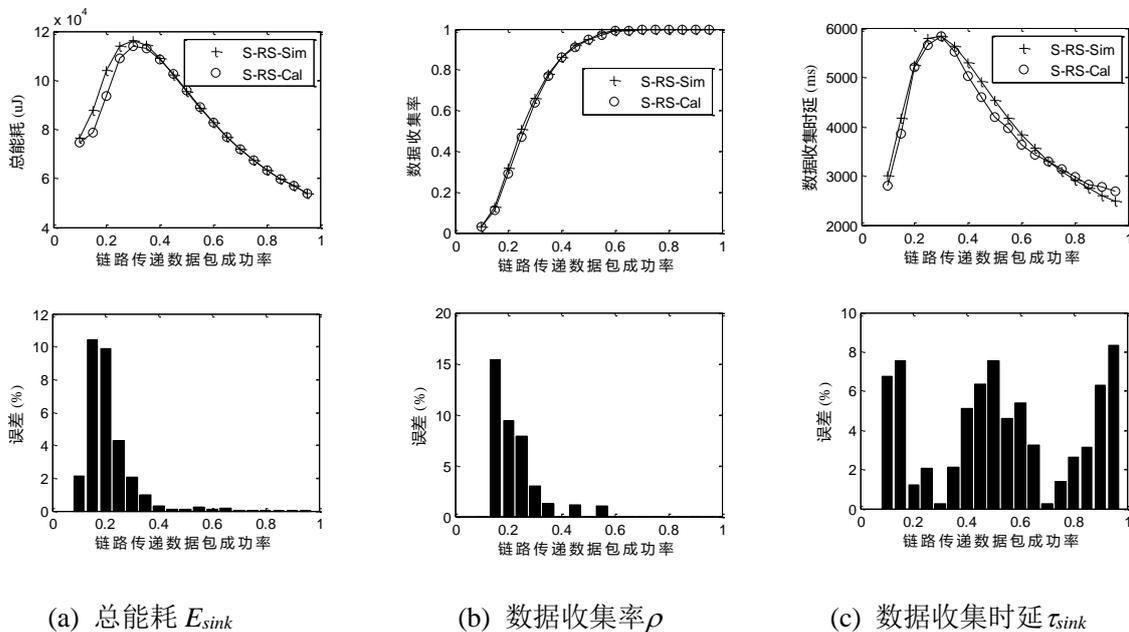


图 8 总能耗、数据收集率和数据收集时延仿真与计算结果对比

的增大（见图 10(b)），但也不增大能耗和时延（见图 10(a)和图 10(c)），这是因为在 S-RS 策略中，节点在收到足够解码的数据包个数之后，给发送方发送 ACK，使得发送方停止继续发送编码包。

其次，利用仿真比较 S-RS 数据收集策略和 IEEE 802.15.4 的 ARQ 策略。图 11 是在取 $\lambda = 7$ 和 $x = 10$ 时不同最大重传次数(即 σ)下两种策略的比较结果。从图 11 可以看出：1) 对于一个给定的最大重传次数，S-RS 数据收集策略的总能耗 E_{sink} 和数据收集率 ρ 均优于 ARQ(见图 11(a)和(b))；2) 在链路传递数据包成功概率 p_k 大于 0.2 时，S-RS 数据收集策略的时延 τ_{sink} 优于 ARQ(见图 11(c))。此外，从图 11 可以进一步发现两个策略的共同特点：1)数据收集率 ρ 随着链路质量的提高(即链路传递数据包成功概率 p_k 的增大)而提高，这和直观是吻合的；2)两种策略的总能耗 E_{sink} 和数据收集时延 τ_{sink} 随着链路质量提高，呈现先上升后下降之趋势，其理由如后所述。当链路质量比较差时， p_k 略微增大会导致节点的数据包经过更多的中间节点，但最终因链路质量差未能到达 sink。也就是说，在这种情形， p_k 的增大会引起网络中更多的中途夭折的数据包被发送，使得

整个网络的能耗和收集数据的时延增大。然而，当链路质量比较好时， p_k 的增加使得更多的数据包被传递到 sink 节点，而且在每条链路，数据包的重传次数也降低了，从而降低了整个网络的能耗和收集数据的时延。

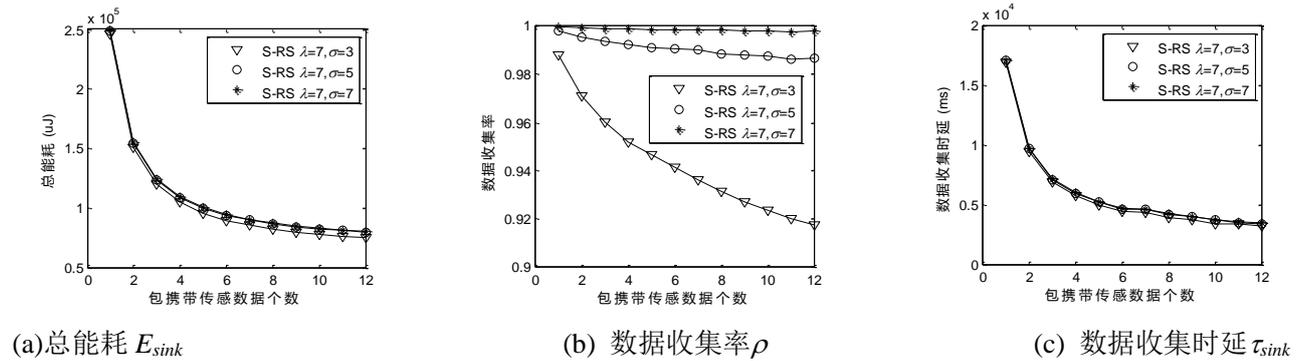


图 9 参数 x 对 S-RS 策略性能的影响

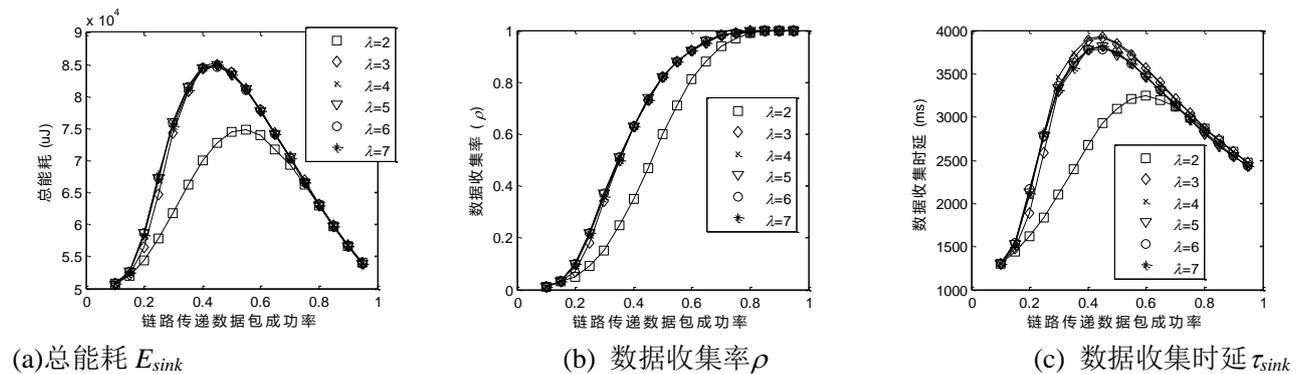


图 10 参数 λ 对 S-RS 策略性能的影响

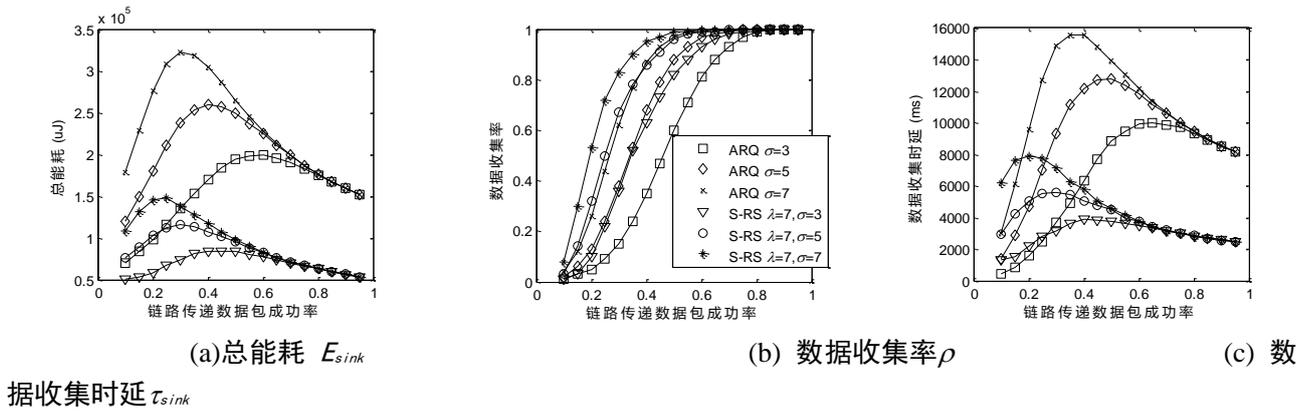
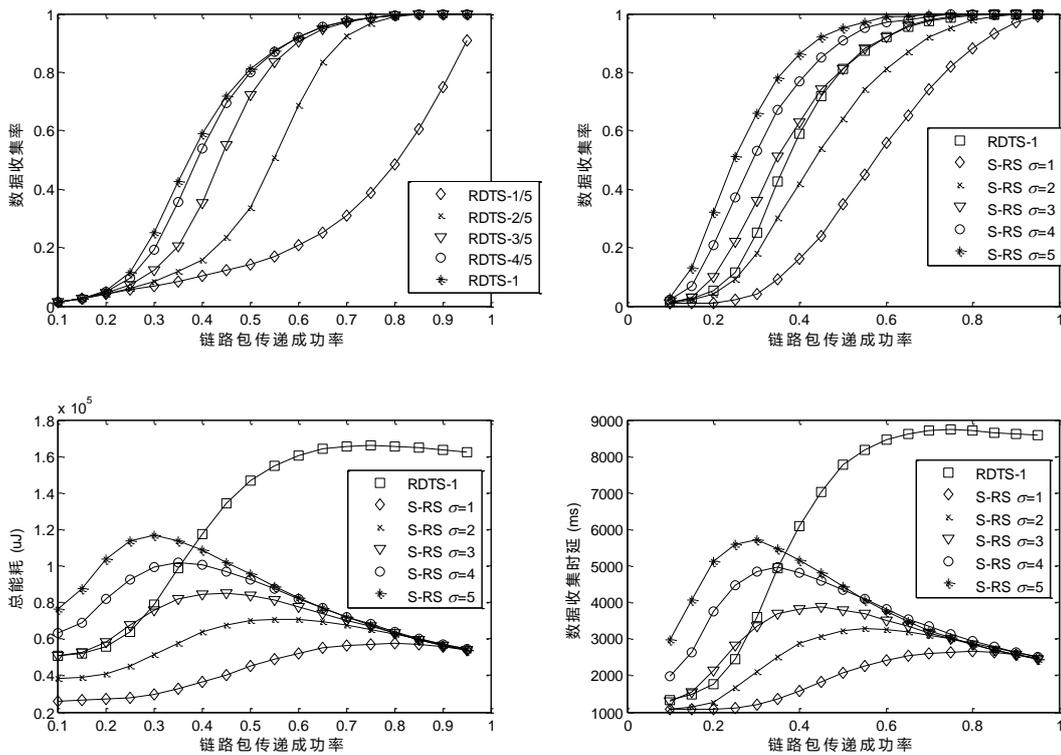


图 11 S-RS 策略和 IEEE802.15.4 自动重传机制的性能比较



再次，利用仿真比较 S-RS 数据收集策略和同样应用 Reed-Solomon 码的可靠数据传输策略 RDTS^[32]进行比较。在 RDTS 策略中， δ 是一个关键参数，但文献[32]未给出具体的计算方法。因此，我们分别取 $\delta=M/5, 2M/5, 3M/5, 4M/5$ 和 $\delta=M$ ，并把对应的 RDTS 策略分别记为 RDTS-1/5, RDTS-2/5, RDTS-3/5, RDTS-4/5 和 RDTS-1。在 S-RS 策略中，取 $\lambda=5, x=10$ 。1000 次仿真运算结果如图 12 所示。由图 12 左上方子图可见，在 RDTS 策略中， δ 越大（冗余数据越多），则数据收集率越高，在这些 δ 取不同值的 RDTS 策略中，以 RDTS-1 的

数据收集率最高。因此，我们选取数据收集率最高的 RDTS-1 策略与本文 S-RS 策略进行比较，比较结果显示在图 12 的其余 3 个子图中。从图 12 右上方子图可见，在 S-RS 策略中，在 MAC 帧重传次数 $\sigma \geq 2$ 时，S-RS 策略的数据收集率优于 RDTS 策略。而且，从图 12 左下方子图可见，在 $\sigma=3$ 时，S-RS 策略的能耗也优于 RDTS 策略。从图 12 可以进一步看出：在链路成功概率 $p_k \geq 0.35$ 时，只须选取 $\sigma=3$ ，S-RS 策略的数据收集率、能耗、时延均优于 RDTS 策略。

最后，考虑求解式(31)的优化问题。固定 $p_k = 0.6$ ，让事先给定的数据收集率取不同的值： $\rho_0 = 0.88, 0.90, \dots, 0.98$ 时，且取时延约束 $T_0 = 2900, 3000, 3100$ ms 时，通过求解式(31)的优化问题，我们得到图 13。图中，左上方是能耗 E_{sink} 的最优值，右上方、左下方、右下方分别是优化问题最优解对应的参数 x 、 σ 、 λ 值。从图 13 可以看出：1) 在 T_0 给定时，能耗 E_{sink} 最优值随着 ρ_0 的增大而增大，这

是因为 ρ_0 的增大意味着可靠性的提高，从而需要一个比较大的重传次数。例如，在 $T_0 = 3000$ ms 时，随着 ρ_0 从 0.88 逐渐增大到 0.92，重传次数 σ 从 1 逐渐增大到 3，这导致能耗增加；2) 式(31)的优化问题并非总是存在最优解。例如，在时延约束 $T_0 = 3000$ ms 时，当数据收集率约束 ρ_0 大于 0.94 时，上述优化问题无解。也就是说，在实践中很难做到：既要达到高数据收集率，又要数据时延足够小。其

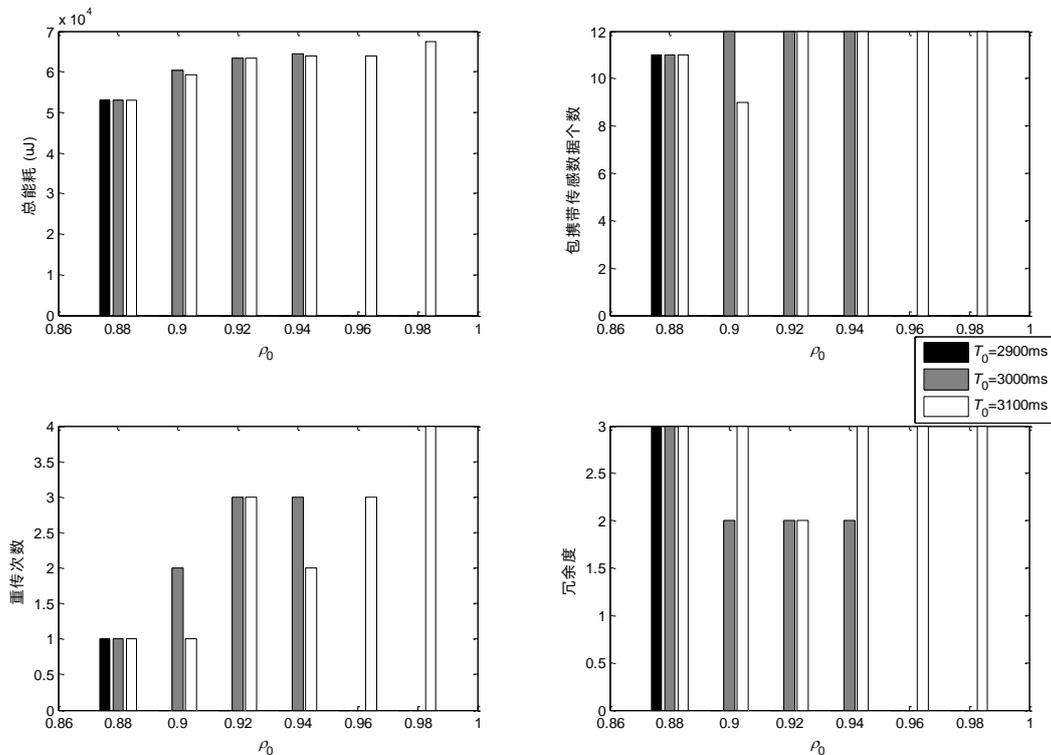
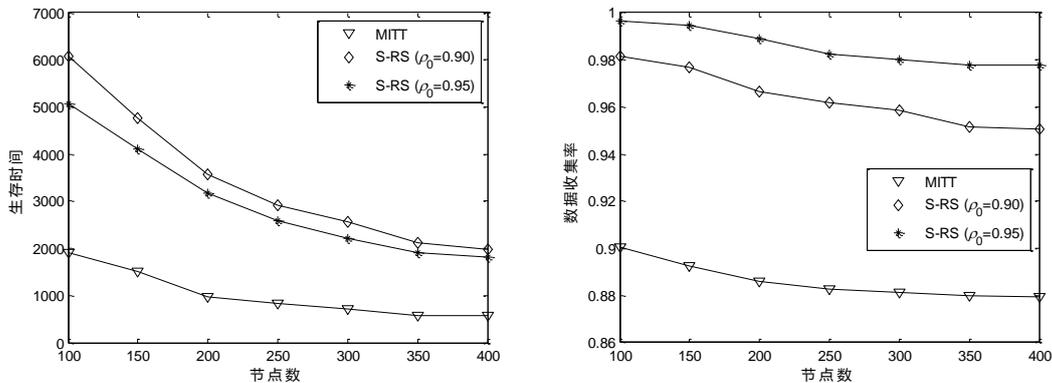


图 13 优化问题的解



(a) 网络生存时间

(b) 数据收集率 ρ

图 14 MITT 和 S-RS 比较

理由是, 如果我们需要 S-RS 策略达到比较高的数据收集率, 那么节点需要发送比较多的冗余编码包, 这会导致能耗和时延的增加。同样, 为了降低能耗和时延, 节点需要减少重传次数和编码包个数, 这会导致一些数据包无法到达 sink 节点, 从而降低数据收集率。3) 在优化问题(31)存在最优解的条件下, 通过求解优化问题(31)获得参数 x 、 λ 、 σ 的最优值, 从而使 S-RS 策略的能耗达到最低。例如, 在收集率约束 $\rho_0 = 0.94$ 和时延约束 $T_0 = 3000$ ms 时, sink 节点可以选择 $x = 12$, $\sigma = 3$ 和 $\lambda = 2$ 使得 $E_{sink} \approx 6.4 \times 10^4$ uJ。

考虑到 MITT^[17]策略构造一棵最小生成树使得瓶颈节点的子节点数最小, 从而平衡 sink 附近节点的能耗, 以最大化网络生存时间, 我们通过仿真将本文提出的 S-RS 策略在其参数 x 、 λ 、 σ 设置为优化问题(31)的最优解的条件下, 与 MITT 策略的生存时间进行对比。此处, 生存时间定义为网络自投入运行至出现第一个剩余能量为零的节点这一时间段(以 sink 收集传感数据的轮数度量)。在仿真中, 在随机构建了一颗数据收集树之后, 随机分配给每个节点 1.5-2 焦耳的初始能量。MITT 策略的 MAC 层重传次数取 3(此乃 IEEE 802.15.4 标准的默认值)。在网络节点数从 100 到 400, $T_0 = 3100$ ms 时, 分别取 $\rho_0 = 0.90$ 和 $\rho_0 = 0.95$, 解优化问题(31), 将最优参数的 S-RS 策略和 MITT 的结果进行比较, 如图 14 所示。从图 14(a)可以看出: 1) 本文提出的 S-RS 策略在数据收集率高于 MITT 策略的前提下(见图 14(b)), 网络生存时间也大大高于 MITT 策略(见图 14(a)); 2) S-RS 策略和 MITT 策略的网络生存时间都随着网络规模的增大而减小(见图 14(a)), 这是因为当网络中节点增多时, sink 每轮需要收集的数据就越多, 从而能耗也越大, 这会缩短网络生存时间; 3) S-RS 策略和 MITT 策略的数据收集率都随着网络规模的增大而减小(见图 14(b)), 这是由于随着网络中节点的增加, 数据包需要经过更多节点的中继才能被传递到 sink 节点, 这会降低数据包端端的可靠性。

顺便指出, 把式(31)的优化目标函数换成数据收集时延 τ_{sink} , 我们可以对时延最小化; 或者把式(31)的优化目标换成对数据收集率 ρ 进行最大化。也可以将能耗 E_{sink} 、数据收集率 ρ 和数据收集时延 τ_{sink} 中的两者或者三者同时作为优化目标, 构建多目标优化问题, 从而使能耗、数据收集率和数据收集时延均达到最优。有兴趣的读者可以自行对之进行研究。

8 结束语

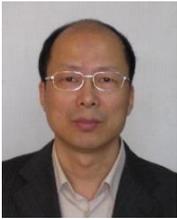
本文应用简单的 Reed-Solomon 码于可靠数据传递, 提出了 S-RS 数据收集策略。这个策略可以让节点采用段内编码或者段间编码将要传递的传感数据进行分段和编码, 通过优化每个段所包含的传感数据个数、编码冗余度、数据链路层重传次数这三个参数, 使得在事先给定的数据收集率和时延约束下收集数据的能耗达到最低。S-RS 数据收集策略适应于链路易丢包的低功耗的数据不可融合的无线传感器网络。

参考文献

- [1] IEEE Computer Society. IEEE 802.15.4 Standard for Wireless Medium Access Control (MAC) and Physical Layer (PHY) Specifications for Low-Rate Wireless Personal Area Networks (WPANs). 2011
- [2] Mutschlechner M, Li B, Kapitza R, Dressler F. Using Erasure codes to overcome reliability issues in energy-constrained sensor networks//Proceedings of the 11th Annual Conference on Wireless On-demand Network Systems and Services (WONS). Obergurgl, Austria, 2014: 41-48
- [3] Montenegro G, Kushalnagar N, Hui J, Culler D. Transmission of IPv6 packets over IEEE 802.15.4 Networks. 2007, RFC4944
- [4] Wicker S B, and Bhargava V K. Reed-Solomon codes and their applications: An introduction to Reed-Solomon Codes. New York City: Wiley-IEEE Press, 1999
- [5] Lin S, Costello D, Miller M. Automatic-repeat-request error-control schemes. IEEE Communications Magazine, 1984, 22(12): 5-17
- [6] Deb B, Bhatnagar S, Nath B. ReInForM: Reliable information forwarding using multiple paths in sensor networks//Proceedings of the 28th Annual IEEE International Conference on Local Computer Networks. Bonn, Germany, 2003: 406-415
- [7] Chakraborty Suchetana, Chakraborty Sandip, Nandi S, Karmakar S. RelBAS: Reliable data gathering from border area sensors//Proceedings of the 2013 IEEE Symposium on Computers and Communications (ISCC). Split, Croatia, 2013: 904-909
- [8] Xu Ming-Sen, Song Wen-Zhan, Zhao Yi-Chuan. Collaborative data collection with opportunistic network erasure coding. IEEE Transactions on Parallel and Distributed Systems, 2013, 24(10): 1941-1950
- [9] Luby M G. LT Codes//Proceedings of the 43rd Annual IEEE Symposium on Foundations of Computer Science. Vancouver, Canada, 2002: 271-280

- [10] Luby M G, Mitzenmacher M, Shokrollahi M A, and Spielman D A. Efficient erasure correcting codes. *IEEE Transactions on Information Theory*, 2001, 47(2): 569-584
- [11] Li B, Wang W, Li H, Yin Q, et al. Performance analysis and optimization for energy-efficient cooperative transmission in random wireless sensor network. *IEEE Transactions on Wireless Communications*, 2013, 12(9): 1635-1639
- [12] Vazifehdan J, Prasad R V, Niemegeers I. Energy-efficient reliable routing considering residual energy in wireless ad hoc networks. *IEEE Transactions on Mobile Computing*, 2013, 13(2): 434-447
- [13] Rout R R, Ghosh S K. Enhancement of lifetime using duty cycle and network coding in wireless sensor networks. *IEEE Transactions on Wireless Communications*, 2013, 12(2): 656-667
- [14] Luo Hong, Tao Hui-Xiang, Ma Hua-Dong, Das S K. Data fusion with desired reliability in wireless sensor networks. *IEEE Transactions on Parallel and Distributed Systems*, 2011, 22(3): 501-513
- [15] Kui Xiao-Yan, Wang Jiang-Xin, Zhang Shi-Geng. A data gathering algorithm based on energy-balanced connected dominating sets in wireless sensor networks//*Proceedings of the 2013 IEEE Wireless Communications and Networking Conference (WCNC)*, Shanghai, China, 2013: 1139-1144
- [16] Baksi S, Sinha A, Khatua S, Das R K. Finding the maximum lifetime data-gathering tree in sensor networks//*Proceedings of the 2014 IEEE Ninth International Conference on Intelligent Sensors, Sensor Networks and Information Processing (ISSNIP)*, Singapore, 2014: 1-6
- [17] Liang Jun-Bin, Wang Jiang-Xin, Cao Jian-Nong, Chen Jian-Er, Lu Ming-Ming. An efficient algorithm for constructing maximum lifetime tree for data gathering without aggregation in wireless sensor networks//*Proceedings of the 2010 IEEE INFOCOM*, California, USA, 2010: 1-5
- [18] Imon S K A, Khan A, Francesco M D, Das S K. RaSMaLai: A randomized switching algorithm for maximizing Lifetime in tree-based wireless sensor networks//*Proceedings of the 2013 IEEE INFOCOM*, Turin, Italy, 2013: 2913-2921
- [19] Zhu Yi-hua, Wu Wan-Deng, Jian Pan, Tang Yi-Ping. An energy-efficient data gathering algorithm to prolong lifetime of wireless sensor networks. *Computer Communications*, 2010, 33(5): 639-647
- [20] Song Xin, Wang Cui-Rong. Linear regression based distributed data gathering optimization strategy for wireless sensor networks. *Chinese Journal of Computers*, 2012, 25(3):568-580 (in Chinese)
(宋欣, 王翠荣. 基于线性回归的无线传感器网络分布式数据采集优化策略, *计算机学报*, 2012, 25(3):568-580)
- [21] Liu Yun-Lu, Pu Ju-Hua, Fang Wei-Wei, Xion Zhang. A MAC layer optimization algorithm in wireless sensor networks. *Chinese Journal of Computers*, 2012, 35(3):529-539 (in Chinese)
(刘云璐, 蒲菊华, 方维维, 熊璋. 一种无线传感器网络 MAC 协议优化算法, *计算机学报*, 2012, 35(3):529-539)
- [22] Su Jin-Shu, Guo Wen-Zhong, Yu Chao-Long, Chen Guo-Long. Fault-tolerance clustering algorithm with load-balance aware in wireless sensor network. *Chinese Journal of Computers*, 2014, 37(3):445-456 (in Chinese)
(苏金树, 郭文忠, 余朝龙, 陈国龙. 负载均衡感知的无线传感器网络容错分簇算法. *计算机学报*, 2014, 37(3):445-456)
- [23] S. Tozlu, Feasibility of Wi-Fi enabled sensors for Internet of Things //*Proceedings of 7th International Wireless Communications and Mobile Computing Conference (IWCMC)*, wuhan, China, 2011:291-296
- [24] Zhu Yi-hua and Leung V. C.M., Efficient Power Management for Infrastructure IEEE 802.11 WLANs, *IEEE Transactions on Wireless Communications*, 2010, 9(7):2196-2205
- [25] Zhu Yi-hua, Luan Shenji, Leung V. C. M., and Chi Kaikai, Enhancing Timer-based Power Management to Support Delay-Intolerant Uplink Traffic in Infrastructure IEEE 802.11 WLANs, *IEEE Transactions on Vehicular Technology*, 2015, 64(1):386-399
- [26] Lu Wen-Wei, Zhu Yi-Hua, Chen Gui-Hai. Energy-efficient routing algorithms based on linear network coding in wireless sensor networks. *Chinese Journal of Electronics*, 2010, 38 (10): 2309-2314 (in Chinese)
(卢文伟, 朱艺华, 陈贵海. 无线传感器网络中基于线性网络编码的节能路由算法, *电子学报*, 2010, 38 (10): 2309-2314)
- [27] Chi Kai-Kai, Zhu Yi-Hua, Jian Xiao-Hong, Leung V.C.M. Energy-efficient prefix-free codes for wireless nano-sensor networks using OOK modulation. *IEEE Transactions on Wireless Communications*, 2014, 13(5): 2670 - 2682
- [28] Shokrollahi A, Raptor Codes, *IEEE Transactions on Information Theory*, 2006, 52:2551-2567
- [29] Qassim Y, Magana M E, Error-tolerant non-binary error correction code for low power wireless sensor networks //*Proceedings of 2014 International Conference on Information Networking (ICOIN)*. Phuket, Thailand, 2014: 23-27
- [30] Ali S, Fakoorian A, Taheri H, Optimum Reed-Solomon Erasure Coding in Fault Tolerant Sensor Networks //*Proceedings of 4th International Symposium on Wireless Communication Systems (ISWCS 2007)*. Trondheim, Norway, 2007: 6 -10
- [31] Tan Poh-Ling, Cheah Cheng-Lai, Ho Chee-Kit, Hardware implementation of Reed-Solomon error correction technique for wireless sensor network based on error pattern analysis //*Proceedings of 2014 IEEE Region 10 Symposium*. Kuala Lumpur, Malaysia, 2014: 347 - 350
- [32] Srouji M S, Wang Z., Henkel J, RDTS: A Reliable Erasure-Coding Based Data Transfer Scheme for Wireless Sensor Networks //*Proceedings of 2011 IEEE 17th International Conference on Parallel and Distributed Systems (ICPADS)*. Tainan, China, 2011: 481-488
- [33] Zhu Yi-Hua, Xu Ji, Li Er-Tao, Xu Lin. Energy-efficient reliable data gathering scheme based on enhanced reed-solomon code for wireless sensor networks//*Proceedings of the SMARTCOMP2014*, Hongkong, China, 2014: 275-280

- [34] Gnawali O, Fonseca R, Jamieson K, Moss D, Levis P. Collection tree protocol//Proceedings of the 7th ACM Conference on Embedded Networked Sensor Systems, New York, USA, 2009: 1-14
- [35] Jambli M N, Azlina A J, Farha Anati A M, Lenando H, Abdullah J, Sinarwati M S. Performance evaluation of CTP routing protocol for mobile Wireless Sensor Network//Proceedings of the 2013 IEEE Conference on Wireless Sensor, Kuching, Malaysia, 2013: 97-101
- [36] Omprakash Gnawali, Rodrigo Fonseca, Kyle Jamieson, Maria Kazandjieva, David Moss, and Philip Levis. CTP: An Efficient, Robust, and Reliable Collection Tree Protocol for Wireless Sensor Networks, ACM Transactions on Sensor Networks, 2013, 10(3):10-17
- [37] Heinzelman W B, Chandrakasan A P, Balakrishnan H. An application-specific protocol architecture for wireless microsensor networks. IEEE Transactions on Wireless Communications, 2002, 1(4): 660-670
- [38] Murugantha S D, Ma D C F, and Fapojuwo A Q. A centralized energy-efficient routing protocol for wireless sensor networks, IEEE Communications Magazine, 2005, 43(3): s8-13,
- [39] Al-Karaki J N, Kamal A E, Routing techniques in wireless sensor networks: a survey, IEEE Wireless Communications, 2004, 11(6): 6-28
- [40] Lu G, De D, Xu M, Song W-Z, and Cao J. TelosW: Enabling ultra-low power wake-on sensor network//Proceedings of 2010 Seventh International Conference on Networked Sensing Systems (INSS). Kassel, Germany, 2010: 211 – 218
- [41] Dutta P, Feldmeier M, Paradiso J, and Culler D. Energy Metering for Free: Augmenting Switching Regulators for Real-Time Monitoring//Proceedings of 7th International Conference on Information Processing in Sensor Networks (IPSN'08). St. Louis, United States, 2008: 283 – 294
- [42] Jin L, Zhu Y-H, Leung V C M, Song W-Z. An Adaptive Energy-conservation Scheme with Implementation Based on TelosW Platform for Wireless Sensor Networks//Proceedings of 2011 IEEE Wireless Communications and Networking Conference (WCNC2011). Cancun, Mexico, 2011:926-931



ZHU Yi-hua, born in 1961, Ph.D., Professor. His main research areas include the Internet of Things (IoT), wireless networks, and network coding. Email: yhzhu@ece.ubc.ca.

XU Ji, born in 1990, a candidate of M. S. degree. His research interest focuses on wireless sensor networks. Email: cornuer@gmail.com.

TIAN Xian-Zhong, born in 1968, Ph.D., Professor. His current research focuses on network coding, mobile computing, and the optimization of protocol for wireless networks. Email: txz@zjut.edu.cn.

Chi Kaikai, born in 1980, Ph.D., associate professor. His main research areas include wireless Ad Hoc networks, wireless sensor network, etc. Email: kkchi@zjut.edu.cn.

联系人：朱艺华，手机：13588111446，email: yhzhu@ece.ubc.ca; yhzhu@zjut.edu.cn

Background

Wireless Sensor Network (WSN) is used in gathering data from the environment. In a WSN, nodes are usually powered by battery with very limited energy. It is very important to design an energy-efficient data gathering scheme to prolong the WSN's runtime. Contemporary WSN usually adopts the low-power and low-rate IEEE 802.15.4 standard in its Medium Access Control (MAC) and Physical (PHY) layers. Therefore, wireless links connecting the sensor nodes in the WSN are prone to losing packet. This reduces Packet Delivery Reliability (PDR), which can be measured by Data Gathering Ratio (DGR) defined as the ratio of the total number of the packets received by the sink node to that of transmitted by the sensor nodes. Moreover, the sensor nodes have short radio ranges, causing a packet to be forwarded by multiple sensor nodes over lossy wireless links before it reaches the sink node, which further reduces PDR.

To improve PDR, the traditional scheme applies retransmission/acknowledgement (ACK) scheme in the MAC layer. But this may cause extra energy consumption in addition to packet delay. Up to date, researchers around the world have been seeking an efficient and effective way of improving PDR based on coding scheme in which redundant data are added, but transmitting the redundant data packets also

causes extra energy consumption. In the literature, there are many works that focus on reducing energy consumption and/or improving PDR of WSN. However, there is still a void in integrating coding scheme with energy consumption, PDR, and data gathering delay such that the optimal coding scheme is found to minimize the energy consumption under the constraints of given PDR and data gathering delay. This paper fills this void.

In this paper, the data gathering scheme called "S-RS scheme" is proposed, which uses a Simple Reed-Solomon (S-RS) coding. The Overall Energy Consumption (OEC), DGR, data gathering delay of the proposed S-RS scheme are derived, which are used to build the Optimization Problem (OP) that minimizes the OEC with the constraints of the preset data gathering delay and DGR. Through solving the OP, the nodes are able to find the optimal values of the parameters, which is then used in the S-RS scheme to achieve the minimal energy consumption under the given constraints.

This research is supported by Natural Science Foundation of China (NSFC) under grant numbers of 61432015, 61472367, and 61379124, which investigate some key problems arising in wireless networks, including reliable disseminating information, saving energy, improving throughput etc.

in wireless networks. The proposed S-RS scheme is applicable to reliably and energy-efficiently delivering packets over lossy wireless links. We have proposed several schemes to save energy, such as the Timer-based Power Management (TPM), the prefix-free coding scheme, and others, which were published in the journals including *IEEE Transactions on Wireless Communications*, *IEEE Transactions on Network and Service Management*, *IEEE Transactions on Vehicular Technology*, and more.

本课题得到国家 863 课题“地址驱动网络关键技术和验证”(SS2015AA010203)和国家发改委 CNGI 课题“一种新型网络体系结构: 地址驱动的网络体系结构、技术研发和试验”(CNGI-12-03-001)的资助。吴建平, 男, 1954 年出生, 博士, 教授, 主要研究领域包括计算机网络系统架构和协议设计, 等等。Email: jianping@cernet.edu.cn。李丹, 男, 1981 年出生, 博士, 副教授, 主要研究领域包括互联网架构、数据中心网络和云计算, 等等。Email: toolidan@tsinghua.edu.cn。毕军, 男, 1972 年出生, 博士, 教授, 主要研究领域包括新型互联网体系结构, 等等。Email: junbi@tsinghua.edu.cn。徐恪, 男, 1974 年出生, 博士, 教授, 主要研究领域包括计算机网络结构、高性能路由器和大规模 P2P 系统, 等等。Email: xuke@mail.tsinghua.edu.cn。李星, 1956 年出生, 博士, 教授, 主要研究领域包括信息科学技术, 等等。Email: xing@cernet.edu.cn。朱晶, 1989 年出生, 博士在读, 主要研究领域包括下一代互联网体系结构、分布式计算系统, 等等。Email: zjinn@aliyun.com。