

# 基于区域化压缩感知的无线传感器网络数据收集方法

杨 浩<sup>1),(2)</sup> 王喜玮<sup>3)</sup>

<sup>1)</sup>(盐城师范学院信息工程学院 江苏 盐城 224002)

<sup>2)</sup>(清华大学软件学院 & 无锡清华信息科学与技术国家实验室物联网技术中心 北京 100084)

<sup>3)</sup>(东北伊利诺伊大学计算机科学系 芝加哥 美国 60625)

**摘 要** 在无线传感器网络的数据收集过程中,传统的传输方式会使靠近 Sink 节点的中继节点接收和发送大量数据,形成一个以 Sink 节点为中心的区域.相对于其他节点来说,该区域内传感器的能量会更快地被耗尽,导致整个网络负载不均衡.目前,压缩感知(Compressive Sensing,CS)已经逐步应用于无线传感器网络领域用以减少传感器的传输量,以达到延长网络生命周期的目的.为实现能量有效的传输,该文提出一种基于区域化压缩感知(Regionalized Compressive Sensing,RCS)的数据收集方法.首先,网络拓扑被随机划分成多个区域,划分原则不依赖于所采集数据的特征或相关性;然后,每个区域选择选出一个区域中心节点以接收其他节点的采样值;随后,区域中心节点使用 CS 方法获得区域测量值;最后,这些节点将它们发送到 Sink 节点用于数据重构.为保证 RCS 的有效性,该文进一步考虑了实际应用中的两个问题,区域测量矩阵的设计和采样停止时机的选择.通过与当前典型的传输方法比较,RCS 方法的性能优于直接传输、分布式压缩感知以及混合算法,同时实验验证了所提出的区域测量矩阵设计原则和采样停止原则是可行的.

**关键词** 压缩感知;区域化压缩感知;数据收集;无线传感器网络;物联网  
中图法分类号 TP393 DOI号 10.11897/SP.J.1016.2017.01933

## Data Gathering Based on Regionalized Compressive Sensing in WSN

YANG Hao<sup>1),(2)</sup> WANG Xi-Wei<sup>3)</sup>

<sup>1)</sup>(School of Information Engineering, Yancheng Teachers University, Yancheng, Jiangsu 224002)

<sup>2)</sup>(School of Software & TNLIST, Tsinghua University, Beijing 100084)

<sup>3)</sup>(Department of Computer Science, Northeastern Illinois University, Chicago 60625 USA)

**Abstract** On the basis of traditional transmission methods in wireless sensor networks, a vital problem is that a center region near the sink will be created and the sensors in it require vast energy to receive or relay data packages from other sensors. In this case, their own energy will rapidly be exhausted. Recently, compressive sensing (CS) has been employed to reduce the number of data transmissions and extend the lifetime of networks. This paper proposes an energy-efficient data gathering approach based on regionalized compressive sensing (RCS). Firstly, the network is randomly divided into smaller regions without the requirement for any specific relationships or features among sensors in the same block. Subsequently, a center sensor for each region is elected and is utilized to receive original readings of all other sensors in its region. Furthermore, CS is implemented on each center sensor in individual region respectively. The regional measurement matrix is given to generate regionalized samplings. Finally, the regional measurements in every region are transported to the sink for reconstruction. To further guarantee the viability of our approach, two issues related to practical applications are considered: (1) how to design an available

收稿日期:2015-11-10;在线出版日期:2016-03-22. 本课题得到本课题得到国家“八六三”高技术研究发展计划项目基金(2015AA01A201)、国家自然科学基金(61402394,61379064,61303209)、江苏省自然科学基金(BK20140462)、江苏省高校自然科学基金项目(14KJB520040)、江苏省博士后科研资助计划(1601162B)、江苏省自然科学产学研前瞻项目(BY2016066-05)、江苏省产学研合作项目(BY2016066-04)、江苏省“青蓝工程”资助. 杨 浩,男,1980年生,博士,副教授,中国计算机学会(CCF)会员,主要研究方向为压缩感知、无线网络. E-mail: anysuc@163.com. 王喜玮,男,1984年生,博士,助理教授,主要研究方向为数据分析与挖掘.

regional measurement matrix, and (2) when the sampling procedure should be terminated. Experimental results on a real-world dataset indicate that the proposed RCS outperforms the existing methods in terms of efficiency, and the feasibility of the two issues in practice is also verified.

**Keywords** compressive sensing; regionalized compressive sensing; data gathering; wireless sensor network; Internet of Things; Cyber-Physical System

## 1 引言

在无线传感器网络(Wireless Sensor Network)的数据传输过程中,越靠近终端(即 Sink 节点)的中继节点需要转发的数据包量越多,这会形成一个以 Sink 节点为中心的区域,即所谓的中心区域<sup>[1]</sup>.在该区域内,传感器需要耗费巨大能量用以接收和发送数据.中心区域的形成导致整个网络负载极其不均,从而极大影响网络的生命周期.虽然利用一些特定传输方法或策略可以减小中心区域的能耗,但这些方法本身可能需要复杂操作,会消耗较多能量;同时,它们的性能依赖特定的应用环境和条件,进一步限制了在实际应用中的普适性.

为实现能量有效的数据通信,压缩感知(Compressive Sensing, CS)理论<sup>[2-3]</sup>被逐渐应用于无线传感器网络领域,以减小数据收集过程中的传输量.然而,研究者发现在实际应用中该技术并不是总是优于传统的传输方法<sup>[4]</sup>.换句话说,压缩感知的优势取决于具体的数据传输策略.以往文献<sup>[5-7]</sup>已经指出在某些情况下简单的压缩感知方法(Plain CS)可能性能不如非压缩感知(no-CS)方法.

为实现能量有效的传输,本文提出一种基于区域化压缩感知(Regionalization Compressive Sensing, RCS)的数据收集方法.该方法的基本思想是将压缩感知的采样过程区域化,以分散中心区域的负载,达到节约能耗的目的.具体来说,首先将网络划分成若干区域,划分时不需要考虑区域内节点感知数据的相关性(例如数据的空间相关性或时间相关性)或者特征(例如数据间的相似特性)等,然后每个区域选出一个区域中心节点以收集同一区域内其他节点的原始采样值,随后每个区域中心节点应用 CS 方法对区域内节点采样进行区域化测量,最后区域中心将所产生的区域测量值发送到 Sink 节点.这样, Sink 节点就得到完整的采样.

为保证 RCS 方法的实用性,本文进一步考虑了实际应用中遇到的两个问题并提出相应的解决方

案:一个是提出区域测量矩阵的设计原则以减小区域中心节点的空间需求及其传输次数,另一个是提出采样停止原则以避免不必要的采样.

本文利用真实数据进行对比实验,并基于 Zigbee 网络比较了数据包结构的不同参数情况.根据实验结果,本文方法性能优于当前典型的传输方法——直接传输、DCS 和混合算法,同时验证了区域测量矩阵设计原则和采样停止原则的有效性.

本文的贡献如下:

(1) 提出一个基于区域化压缩感知的能量有效的数据收集方法 RCS. 该方法将无线传感器网络进行区域化,在每个区域进行独立采样测量,有效地解决了中心区域载荷过重的问题,同时充分结合了直接传输和压缩感知的优势,极大地降低了整个网络的传输能耗.

(2) 针对 RCS 方法在实际应用中遇到的两个问题(区域中心节点进行数据处理时的空间需求过大和 Sink 节点采样停止的时机),提出相应的解决方案,进一步减少了数据收集过程的能量消耗.

(3) 基于 Zigbee 环境和实际数据进行实验,验证了本文方法性能优于当前典型的数据传输方法,同时展示了上述两个解决方案的有效性.

## 2 相关背景

压缩感知是一种有效的信号获取和压缩方式,应用于数据同步采样和压缩.假设  $X$  是  $N$  维信号,  $\Phi$  是  $M \times N$  的测量矩阵( $M < N$ )且信号  $Y = \Phi \times X$ , 则当  $X$  是  $k$  个基向量( $k \ll N$ )的线性组合时,利用  $Y$  和  $\Phi$  可以以极大概率恢复该信号.为了高概率精确恢复原始信号,矩阵  $\Phi$  的行数  $M$  需要满足  $M \geq c \cdot k$ , 其中  $k$  为信号稀疏度,  $c$  为采样率函数.

使用 CS 技术时,若信号  $X$  的稀疏基为  $\Psi$ , 当  $X = \Psi \cdot \Theta$ , 其中  $\|\Theta\|_0 = k$ ,  $\|\cdot\|_0$  表示  $l_0$  范数, 则此类信号称为  $k$ -稀疏( $k$ -sparse)信号.对一些不是严格稀疏的信号来说,如果其  $\Theta$  用前  $k$  个最大稀疏替代,得到  $X'$  可近似表示  $X$ , 那么这类信号称为  $k$ -可

压缩( $k$ -compressible)信号. 也就是说, 这前  $k$  个最大系数包含了原始信号最重要的信息.

将 CS 应用于无线传感网络领域时, 考虑以下场景: 整个网络包含  $N$  个传感器和 1 个 Sink 节点, 其中 Sink 节点收集所有传感器的采样数据. 假设传感器  $i$  采集的数据为  $x_i$ , 所有传感器采集的数据表示为向量  $\mathbf{X}=[x_1, x_2, \dots, x_N]$ , 则 Sink 节点将收到  $M$  次测量值, 表示为  $\mathbf{Y}=[y_1, y_2, \dots, y_M]$ , 其中  $y_j = \sum_{i=1}^N \varphi_{j,i} x_i$ . 根据 CS 理论, 当 Sink 节点接收到的测量值达到一定数量时, 可以精确恢复出每个传感器采集的原始数据. 这意味着重构精度与传输方式无关. 然而, 在实际应用中不合适的数据传输策略将会导致传感器产生不必要的能量消耗. 为了更好地应用该理论, 需要结合无线传感网络的特性, 设计更加有效的数据收集方法.

### 3 研究目的和现状

根据 CS 相关理论, 发送者在采样时对原始数据进行简单的加权求和得到相应的测量值, 这就是所谓的测量过程. 该过程避免了复杂的压缩和编码操作, 因此它可以降低传感器在数据通信时的能量消耗. 在无线传感器网络领域, CS 技术已经引起了国内外研究人员的关注, 主要从信号的采样、传输和重构 3 个方面展开<sup>[8-11]</sup>.

在信号采样方面, 由于无线传感网络是通过多个传感器节点共同感知环境的, 需要将研究对象扩展到对信号集的处理, 要考虑信息间的关联. 目前, 研究人员基于信号关联关系提出了许多有效的采样理论和方法, 典型的有基于分布式编码的分布式压缩感知理论<sup>[12]</sup>和基于信号相关性的采样方式<sup>[13]</sup>.

信号传输时, 压缩感知方法要求在传输过程中实现对采样数据的有效测量. 当前的解决方法可以简单分为面向随机测量<sup>[14-15]</sup>和面向网络结构的传输方法<sup>[16]</sup>. 前者根据网络结构或信号的稀疏基对信息进行随机测量, 以避免遍历所有节点和减小对网络稳定性的依赖, 而后者主要基于不同网络拓扑提出的传输改进算法, 以减少边缘节点不必要的传输.

终端节点的重构算法是基于对信号的采样和传输方式所设计的. 终端节点接收到不同区域信息, 且每个区域的信息都有自身的特性. 为此, 研究人员提出了许多重构算法, 主要包括基于信号支撑的凸优化算法<sup>[17-18]</sup>、基于分段匹配的贪婪追踪算法<sup>[19-20]</sup>和

多任务贝叶斯算法<sup>[21-22]</sup>等.

本文关注的是压缩感知的性能在一般情况下是否总是优于传统的传输方法. 例如, 图 1 分别使用 CS 方法和 no-CS 方法进行传输. 网络由 20 个传感器构成, 其中 1 个是 Sink 节点. 使用 CS 方法时, 假设所需测量数为 7 次, 则整个网络需传输  $7 \times 19 = 133$  次测量值, 而 no-CS 方法的传输总数为 53, 显然比 CS 方法要少的多.

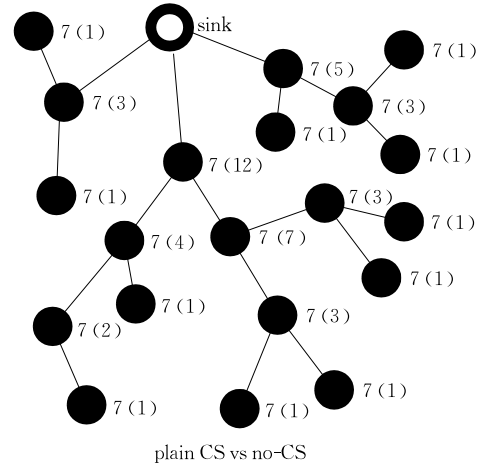


图 1 对比 plain CS 和 no-CS 的传输次数

Luo 等人<sup>[5]</sup>也提及这一现象, 指出当网络规模并不是很大的时候, 简单地应用 CS 并不比 no-CS 更节约能耗, 原因是相对于传统的传输方法来说, CS 会导致网络边界节点耗费更多能量. 根据这一现象, 他们设计了一种混合方法以减少边界节点的能耗. 如图 2 所示, 每个靠近边界的节点向其父节点直接传输自身采样. 以此类推, 直到节点的子孙数不少于 6 时改用 CS 方法进行. 该方法减少了边界节点的传输次数, 节约了网络能耗.

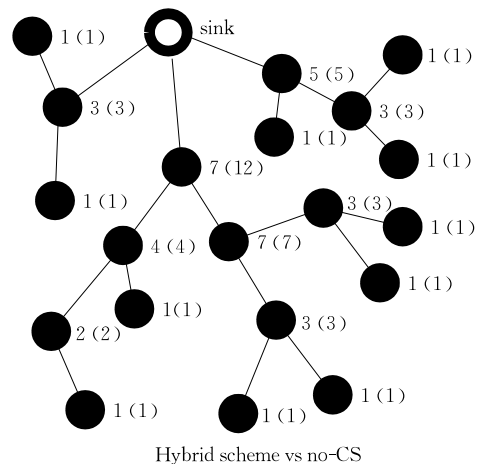


图 2 对比 Hybrid CS 和 no-CS 的传输次数

随后, Caione 等人<sup>[6-7]</sup>进一步基于 Zigbee 网络

验证了这一情况。他们针对高密度传感器网络提出将包转发(Pack and Forward, PF)传输方式和分布式压缩感知(Distributed Compressed Sensing, DCS)<sup>[12]</sup>相结合。该方法的核心思想是每个节点在数据传输过程中通过对比 PF 和 DCS 方法所需能耗的大小来选择合适的传输方式,因此它结合了这两种方法的优势,减少了传感器的传输能耗。

上述这些方法原则上都是混合方法,它们使整个网络节点被分成了两部分,其中一部分是使用 no-CS 方法的边界节点,另一部分是使用 CS 方法的中间节点。然而,网络在进行数据收集时,仍然无法避免中心区域的形成,而该区域的能量消耗占整个网络能耗的大部分。为此,需要在结合 CS 和传统传输方法优势的同时,尽可能地减少中心区域的负载,以延长网络寿命。

## 4 区域化压缩感知(RCS)方法

### 4.1 基本思想

本文提出一种基于区域化压缩感知(Regionalization Compressive Sensing, RCS)的数据收集方法,其基本思想就是不同区域内分别执行区域化的压缩感知,以分散中心区域的负载,达到节约整个网络传输能耗的目的。举例来说,假设一个无线传感网络由 3 个区域  $R = \{r_i | i = 1, 2, 3\}$  组成,每个区域有若干传感器节点,如图 3 所示。网络中存在一个 Sink 节点,收集每个区域内节点的感知数据。在区域  $r_i$  内,选择一个节点作为区域中心节点  $c_i$  (region center node) 负责压缩和转发数据。区域中心节点的

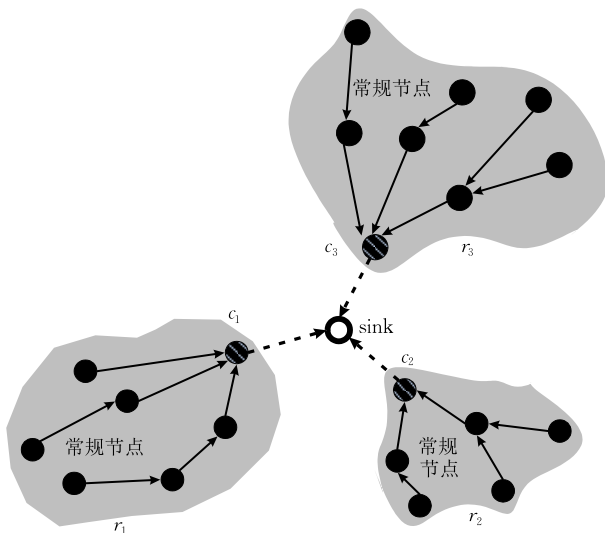


图 3 区域化压缩感知 RCS 的基本思想。整个网络有 3 个区域,每个区域包含一个区域中心节点和一些常规节点

选择一般采用对候选集的随机调度,其轮转标准可以是基于时间、能量、实时性、服务质量保证等。候选集的构成为每个区域中靠近 Sink 端的边界节点,区域内的其他节点称为常规节点  $n_j^{(i)}$  (normal node),其中  $n_j^{(i)} \in r_i, 1 \leq j \leq |r_i| - 1$ 。

网络进行数据收集时,常规节点  $n_j^{(i)}$  将采集到的原始信号  $x_j^{(i)}$  直接发送到相应的区域中心节点  $c_i$ ,随后  $c_i$  进行区域化压缩感知测量。具体测量步骤是:首先,  $c_i$  生成一个  $M \times |r_i|$  的区域测量矩阵  $\Phi_i$ ,其中  $M$  为整个网络准确重构所需要的采样量;然后,  $c_i$  在区域  $r_i$  内利用 CS 方法融合接收到的数据值,产生  $M$  个区域测量值,即  $Y_l^{c_i} = \sum_{j=1}^{|r_i|} \phi_{l,j} x_j, l = 1, \dots, M$ 。显然,  $Y_l^{c_i}$  只是网络整体测量值的一部分。当 Sink 节点收到所有区域测量值,就形成一次完整测量  $Y = \sum_{i=1}^3 Y_l^{c_i}$ 。

根据上述过程,区域内的数据传输次数为  $\sum_{c \in S}$

$\sum_{i \in R(c) \setminus c} hop(i)$ ,其中  $S$  表示网络的区域中心节点集合,  $R(c)$  表示区域中心节点为  $c$  的区域,  $hop(i)$  表示节点  $i$  到其中心节点的跳数。同时,区域中心节点的传输次数为  $\sum_{c \in S} Hop(c)$ ,其中  $Hop(c)$  表示区域中心节点  $c$  到 Sink 节点的跳数。由上可知,整个网络传输次数为

$$\sum_{c \in S} \sum_{i \in R(c) \setminus c} hop(i) + \sum_{c \in S} Hop(c) \quad (1)$$

该方法充分利用了直接传输方法的优势,使得在每个单独区域内传输能耗低于单纯使用 CS 方法时的能耗。换句话说,RCS 方法减少了边界节点的传输能耗,弥补了 CS 方法的不足。需要注意的是,RCS 中的边界节点指的是区域内的边界节点。由于它扩展了边界节点的内涵,因此在数据传输过程中网络不会形成较大的中心区域,使得它的性能在整体上优于混合压缩感知方法。

需要说明的是,RCS 方法不同于基于块的压缩感知(Block Compressed Sensing, BCS)方法<sup>[23]</sup>。BCS 要求处理的信号是二维信号且信号的行和列都是相关的,分块后的二维信号需要按行调整为一维信号且不能改变信号序列位置。同样地,RCS 也不同于 DCS,因为使用 DCS 时需要根据信号集满足的关联特征使用不同模型进行相应的处理。显然,RCS 方法没有相关性条件的约束且区域采样的处理方式也是截然不同的。

RCS 方法的数据传输过程如下:

1. 初始化, 将整个网络进行随机区域化, 得到若干区域  $R$ , 并分配相应的区域中心节点  $C$
2. 收集区域化数据. 对于每个区域  $r_i$ , 常规节点  $n_j^{(i)}$  直接向其中心节点  $c_i$  传输采样读数  $x_j^{(i)}$
3. 生成区域化测量. 区域中心节点  $c_i$  生成大小为  $M \times |r_i|$  的区域测量矩阵  $\Phi_i$ , 并融合所有接收到的数据  $X_i = \langle x_1^{(i)}, x_2^{(i)}, \dots, x_{|r_i|}^{(i)} \rangle$ , 得到区域测量值  $Y_{r_i} = \Phi_i \times X_i$
4. 传输区域测量值. 所有区域中心节点向 Sink 节点传输其区域测量值  $Y_r$ , 形成一次完整的测量
5. while 采样次数  $> \Gamma$   
    执行数据重构  
End
6. while 重构精度不满足采样停止规则  
    转向步骤 3  
End

其中,  $\Gamma$  是预先分配的阈值, 目的是为减少开始阶段不必要的重构时间, 其取值依赖实际环境.

## 4.2 RCS 方法分析

本文提出的 RCS 方法避免了大规模中心区域的形成且不同于一般的分簇方法, 其优势在于:

(1) 通过使用区域中心节点执行测量, 将传统数据传输方法所形成的中心区域的载荷压力分解到各个局部区域, 消除了形成大面积中心区域的可能. 同时, 区域中心节点间可以相互融合, 即当某中心节点接收了其他中心节点的测量值时, 不是简单地直接转发.

(2) 综合利用传统传输方法和 CS 方法的优势. 网络被分解成了多个区域使得 RCS 最大程度地利用了直接传输的优势. 这是因为区域化使得更多的节点被看作为边界节点, 因此在每个区域内尽可能地节约了网络的传输能耗.

(3) 区域大小不取决传输路由, 而是基于 CS 方法和具体的数据包结构, 因此不会因为链路的改变而调整区域大小.

进一步分析, 虽然 RCS 的网络划分是随机的, 但合适的分区将会减少更多的传输能耗, 因此本文给出区域划分的参考标准. 传感器进行通信时, 其数据包结构可以认为有 3 个部分<sup>[24]</sup>, 分别为标识 (Identification) 部分  $D_{id}$ 、数据装载部分  $D_{data}$  和相关信息部分  $D_{Rl}$ , 如图 4 所示.

Data Package Format		
Node ID Cluster ID	Transmission Sequence Number Data_Length Data Value	Physical address Destination address ...
ID	Data	Relative Information

图 4 数据包结构

若采用直接传输的方式, 则网络中的节点需要发送的数据包数量为

$$(N(\text{children}) + 1) \cdot (D_{id} + D_{data} + D_{Rl}) / P_{size} \quad (2)$$

相应地, 采用 CS 方法时, 网络中的节点需要发送的数据包数量为

$$M \cdot (D_{data} + D_{Rl}) / P_{size} \quad (3)$$

其中:  $N(\text{children})$  表示节点的子孙数量;  $P_{size}$  表示数据包在网络中的最大负载量;  $M$  为测量次数. 显然, 当网络节点数较少时, 式(2)的值小于式(3), 说明此时直接传输的性能优于压缩感知. 反之, 则式(2)大于式(3).

显然, 区域划分及中心节点选取影响了网络传输能耗. 通过对比上述两式可知, 直接传输方法的性能在传输规模较小时优于 CS, 但随着网络中的数据包不断向 Sink 节点汇聚, 前者的优势将逐渐被后者取代. 因此, 在进行区域划分时, 为了尽可能地减少数据传输量, 每个区域内节点数应使得式(2)与式(3)的值近似相等, 才能保证最大限度节约传输能耗. 因此, 区域大小取决于测量值和数据包结构. 在实际应用中, 由于网络的动态性 (链路中断、节点加入等), 对节点子孙的判断只能是一个大致结果. 网络连通性的变化会影响到区内和区间的传输. 为减少 RCS 的传输次数, 中心节点的选取原则是尽量减少  $\sum_{c \in S} \sum_{R(c) \setminus c} hop(i) + \sum_{c \in S} Hop(c)$  的值. 实际上, 中心节点既要收集区域内数据, 也要转发到 Sink 节点, 因此在实际使用中要根据所构建的区域选择靠近 Sink 端的边界节点作为中心节点的候选集, 以保证中心节点可以与 Sink 节点直接通信, 候选集的节点经过一段时间进行轮转调度. 若某些区域无法满足条件, 则需为其中心节点构建路由链接到靠近 Sink 节点的其他中心节点, 构建原则是尽量减少转发节点跳数. 这是因为当某中心节点将产生的测量值发送给另一个中心节点后, 后者可以根据 CS 方法与自身的测量值进行融合, 而不需要单独转发.

由此可知, RCS 方法的能量消耗依赖于  $N(\text{children})$ 、区域中心节点数  $|R(c)|$  和中间节点的跳数  $\sum_{c \in S} Hop(c)$ .

## 5 基于 RCS 的优化

### 5.1 实际应用中的两个问题

为保证 RCS 方法的有效性, 本节进一步考虑了

实际应用中的两个问题.

问题 1. 区域中心节点需要使用区域测量矩阵对接收到的数据值进行加权并求和,以产生一个区域测量值.随着网络规模的增加,该测量值的数值会较大,增大了区域中心节点的存储负载.同时,所生成的数据包可能超过通信协议能够允许的最大载荷,导致数据包不得不被分割为两次或多次传输.因此,需要设计合适的区域测量矩阵,以限制区域中心节点的存储及数据包长度.

问题 2. 在实际环境中,采样对象的复杂性导致所需的测量次数难以事先预知,因此具体的采样次数应该取决于对原始信号的重构精度.换句话说,当重构准确度已经足够高或者达到预期要求时, Sink 节点就不需要再进行测量了.为了避免网络传输多余的测量值,需要考虑一个判断规则来决定采样过程何时可以终止.

显然,上述两个问题影响了 RCS 方法在实际应用中的性能.本文首先考虑区域测量矩阵的设计原则,同时给出一个合适的稀疏矩阵.然后,本文基于该矩阵给出重构精度的理论估计,并提出相应的采样停止原则.

## 5.2 区域测量矩阵设计原则

根据 CS 相关理论,当信号稀疏程度难以预先确定时,随机矩阵往往有较大概率满足与信号稀疏基的不相干性要求,因此它们常常被用作压缩感知的测量矩阵<sup>[25-26]</sup>,例如高斯随机矩阵.然而,当传感器网络规模较大时,面向整个网络构建满足高斯分布的高斯矩阵显然是难以实现的,这是因为无线网络结构的不稳定很容易破坏所构建矩阵的特性.另一方面,根据 RCS 方法,区域中心节点会生成一个区域测量矩阵,然后用矩阵参数作为权值来加权所接收到的采样值,并求和得到区域测量值.由于感知对象的不确定导致需要传输较多的测量数,这就增大了区域测量矩阵的规模,从而导致区域中心节点存储需求的增加.同时,当生成的测量值较大时,产生的数据包也会相应增大.在进行数据通信时,数据包所允许的最大载荷长度是受限的,即一个特定的通信协议的数据包的最大载荷是固定的.当区域中心生成的数据包超过最大载荷长度时,就会被分成两次传输,甚至多次传输,产生不必要的能耗浪费.因此,RCS 方法的区域测量矩阵应该是在满足压缩感知理论的前提下,使区域中心节点在生成区域测量值的过程中尽量减少所需存储空间,并要避免多次传输区域测量值.

根据上述原因,本文提出基于 RCS 的区域测量矩阵设计原则:

(1) 区域中心节点进行测量时,测量结果不能大于原始信号的数值,即  $|X_j^{(i)}| = |\varphi_i \cdot x_j^{(i)}| \leq |x_j^{(i)}|$ ,其中  $x_j^{(i)}$  表示区域  $i$  内的第  $j$  个节点的采样.

(2) 生成区域测量值时,区域中心节点产生的各节点的测量值应尽可能相互抵消.这意味着测量值的加权和绝对值  $|Y_l^c| = \left| \sum_{j=1}^{|\varphi_l|} \varphi_{l,j} x_j^{(i)} \right|$  应该尽可能小于其加权绝对值之和  $\sum_{j=1}^{|\varphi_l|} |\varphi_{l,j} x_j^{(i)}|$ ,其中  $l=1, \dots, M$ .

根据上述原则,本文使用一个稀疏随机矩阵作为区域测量矩阵,来验证 RCS 方法的有效性.众所周知,对实际环境进行采样时,由于传感器放置的地理位置相近,其感知结果会有一定程度的关联.对具体的采样对象和采样需求,区域内传感器采样值的关联程度是不同的.本文以同一区域内的传感器采样值是近似的情况为例,因此所使用的区域测量矩阵行参数值的概率均值为 0.

本文的测量矩阵  $\Phi$  维度为  $M \times L$ ,其中  $M$  是测量次数, $L$  是区域节点数.矩阵  $\Phi$  的参数  $\varphi_{i,j}$  是从集合  $A = \{0, B\}$  元素以概率  $p$  随机选取且集合  $B = \{-1, 1\}$ .矩阵  $\Phi$  及其参数  $\varphi_{i,j}$  具体构造如下:

$$\Phi = [\varphi_{i,j}]_{M \times L},$$

$$\text{其中, } \varphi_{i,j} = \begin{cases} 0, & p=1-1/\delta \\ B, & \text{otherwise} \end{cases}, \quad B = \begin{cases} -1, & p=1/2\delta \\ 1, & p=1/2\delta \end{cases}.$$

上式中,参数  $\delta$  依赖区域采样集的稀疏度.

首先,我们讨论矩阵  $\Phi$  满足 CS 理论中不相干性的要求.由于  $\Phi$  的非零部分实际上是 Bernoulli 矩阵,它满足不相干性要求<sup>[2]</sup>,可令其不相干系数最大值为  $q$ .同时,  $\Phi$  中非零值以概率  $1/\delta$  出现,因此该矩阵的不相干系数的范围是  $[0, q]$  且均值为  $q/\delta \leq q$ ,因此  $\Phi$  满足不相干性的要求.

进一步来说,令同一区域内的传感器采样值是近似的,设为  $x$ ,因此区域测量值的均值为  $(0 \times (1-1/\delta) + (-1 \times 1/2\delta) + (1 \times 1/2\delta)) \times x = 0$ ,这意味着在进行测量的过程中矩阵  $\Phi$  可以尽可能限制测量空间的增加.在实际应用中,为保证重构的精度, $\delta$  的取值要使实际测量值在对应的稀疏基上有足够的投影.根据 CS 理论,  $\Phi$  中每行元素的非零值数量要不小于  $\log n$  才能保证 Sink 节点能得到精确重构,即  $1 \leq \delta \leq n / \log n$ ,其中  $n$  是原始信号数量.当满足该条件时,理论上可以保证精确重构,但不同的  $\delta$  值所需的测量值不同.换句话说,若使用不同的  $\delta$  值和

相同的测量值,则得到的重构结果不同。

基于上述测量矩阵,本文进一步分析测量值的期望长度与区域内节点数的关系。

**定理 1.** 假设区域内节点数为  $n$ ,区域中心节点使用上述测量矩阵  $\Phi$  进行测量. 为简单化,令每个节点的采样值均为  $x$ ,则区域测量值存储空间  $S$  的期望长度为

$$E(S) \approx \begin{cases} \left\lceil \log_2 \left( \frac{\sqrt{2\pi n}}{\delta\pi} \times x \right) \right\rceil, & n \text{ 是奇数} \\ \left\lceil \log_2 \left( \frac{n-1}{n} \cdot \frac{\sqrt{2\pi n}}{\delta\pi} \times x \right) \right\rceil, & n \text{ 是偶数} \end{cases}.$$

证明. 令区域测量值的期望长度为  $E(x)$

(1) 当  $n$  是奇数时,

$$\begin{aligned} E(x) &= \frac{1}{\delta} \left( \frac{1}{2} \right)^n (n + C_n^1 \cdot (n-2) + \\ & C_n^2 \cdot (n-4) + \cdots + C_n^{\frac{n-1}{2}} \cdot 1) \times 2 \times x \\ &= \frac{1}{\delta} \left( \frac{1}{2} \right)^{n-1} (n + C_n^1 \cdot (n-2 \times 1) + \\ & C_n^2 \cdot (n-2 \times 2) + \cdots + \\ & C_n^{\frac{n-1}{2}} \cdot [n-2 \times (n-1)]) \times x \\ &= \frac{1}{\delta} \left( \frac{1}{2} \right)^{n-1} (n + n \cdot (C_{n-1}^1 - C_{n-1}^0) + \\ & n \cdot (C_{n-1}^2 - C_{n-1}^1) + \cdots + \\ & n \cdot (C_{n-1}^{\frac{n-1}{2}} - C_{n-1}^{\frac{n-1}{2}-1})) \times x \\ &= \frac{n!}{\delta \cdot 2^{n-1} \left[ \left( \frac{n-1}{2} \right)! \right]^2} \times x. \end{aligned}$$

根据 Stirling 定理<sup>[27]</sup>,

$$n! \approx \sqrt{2\pi n} (n/e)^n,$$

得到

$$\begin{aligned} E(x) &\approx \frac{\sqrt{2\pi n}}{\delta\pi e} \left( \frac{n}{n-1} \right)^n \times x \\ &= \frac{\sqrt{2\pi n}}{\delta\pi e} \left( 1 - \frac{1}{n-1} \right)^n \times x \\ &\approx \frac{\sqrt{2\pi n}}{\delta\pi} \times x, \end{aligned}$$

因此

$$E(S) \approx \left\lceil \log_2 \left( \frac{\sqrt{2\pi n}}{\delta\pi} \times x \right) \right\rceil.$$

(2) 当  $n$  是偶数时,

$$\begin{aligned} E(x) &= \frac{1}{\delta} \left( \frac{1}{2} \right)^n (n + C_n^1 \cdot (n-2) + \\ & C_n^2 \cdot (n-4) + \cdots + C_n^{\frac{n-2}{2}} \cdot 1) \times 2 \times x + \\ & \frac{1}{\delta} \left( \frac{1}{2} \right)^n \times C_n^{\frac{n}{2}} \times 0 \times x \end{aligned}$$

$$= \frac{n-1}{n} \cdot \frac{n!}{\delta \cdot 2^{n-1} \left[ \left( \frac{n-1}{2} \right)! \right]^2} \times x.$$

根据上述证明,可得

$$E(S) \approx \left\lceil \log_2 \left( \frac{n-1}{n} \frac{\sqrt{2\pi n}}{\delta\pi} \times x \right) \right\rceil. \quad \text{证毕.}$$

根据定理 1,参数  $\delta$  影响存储空间的大小. 当  $\delta$  取值较大时,所需的存储空间可能变少,但会影响原始信号的重构准确度. 这是因为  $\delta$  的增加导致测量矩阵稀疏化,从而导致测量值不能包含足够的原始信号稀疏基的投影. 随后,我们将评估在使用矩阵  $\Phi$  时, Sink 节点收到不同测量数后的重构精度。

### 5.3 采样停止原则

根据压缩感知理论,测量次数  $M$  与信号稀疏度  $k$  应满足  $M \geq c \cdot k$ ,才能以极大概率实现精确重构. 其中,  $c$  为采样系数,取决于实际应用环境. 一般来说,如果采样信号的稀疏程度无法预先被获得或者准确估计,则只能根据经验来决定测量次数. 由于实际环境的复杂性,采样信号稀疏度难以准确获得,所以需要尽量多地测量以保证 Sink 节点能够精确重构. 另一方面,在进行信号重构时,当重构结果达到一定精度后,通过多增加一次测量对精度并不会会有明显的提高. 这意味着所增加的采样是无效的,然而其传输过程所消耗的能量却不能被忽略. 因此,为了保证精确重构将使区域中心节点传输大量冗余采样,这显然浪费了大量不必要的能耗。

实际上,当信号的恢复结果达到某一精度或某一预先设定的阈值时, Sink 节点应该停止采样以避免网络产生多余的能耗浪费. 然而,压缩感知方法在实际应用中的采样往往较多,原因是接收端一般难以判断重构信号的准确程度,因此不得不尽可能多地采样. 为解决这一问题, Malioutov 等人<sup>[28]</sup> 针对 Gauss 矩阵和 Bernoulli 矩阵设计了一定的采样停止规则,以避免不必要的采样。

本文基于上述设计的区域测量矩阵提出一个采样停止原则,旨在通过已接收到的测量值高准确率地估计信号的重构精度,以决定 Sink 节点是否需要停止采样。

**定理 2.** 假设  $N$  维信号  $\mathbf{S} = [s_1, s_2, \dots, s_N]$  为原始信号,若 Sink 节点使用测量矩阵为  $\Phi$ , 接收到  $M$  次测量,测量值为  $\mathbf{Y}_M = [y_1, y_2, \dots, y_M]$ , 且重构信号为  $S^{(M)}$ , 然后再进行  $l$  次测量,并得到  $l$  个重构信号  $S^{(M+1)}, S^{(M+2)}, \dots, S^{(M+l)}$ . 如果这些重构信号值是相等的,则重构信号  $S^{(M)}$  等于原始信号  $S$  的概

率满足

$$P\{S^{(M)} = S\} \geq 1 - (1 - 1/2\delta)^{-l}.$$

证明. 假设重构信号  $S^{(M)}$  与原始信号  $S$  不相等, 则存在零空间 (null space)  $S - S^{(M)}$ , 因此在第  $M$  次测量后, 得到等式  $a_M \cdot S = a_M \cdot S^{(M)}$ , 其中  $a_M \in W$ ,  $W = \{a_M | (S - S^{(M)}) \cdot a_M = 0, a_M \neq 0\}$ . 显然,  $W$  既是零空间, 也是  $N$  维实数域  $R^N$  的  $N-1$  维的子空间. 第  $M+1$  次测量后, 得到测量值  $Y_{M+1}$ . 若满足  $(S - S^{(M)}) \cdot a_{M+1} = 0$ , 则第  $M$  次重构信号  $S^{(M)}$  作为第  $M+1$  次的重构结果也是适用的, 并且  $P\{a_{M+1} \in W\} \approx 1/2\delta$ . 以此类推, 当满足  $S^{(M)} = S^{(M+1)} = S^{(M+2)} = \dots = S^{(M+l)}$  时,  $S^{(M)}$  与  $S$  相等的概率下限则为  $1 - (1 - 1/2\delta)^{-l}$ . 证毕.

根据上述定理, 本文提出采样停止原则: (1) 首先评估重构结果等于原始信号的概率; (2) 然后根据预先给定的约束条件 (例如重构准确概率、采样时间、能耗上限等) 来判定是否继续或停止采样.

对 RCS 方法来说, 该原则避免了对整个网络的冗余采样, 极大地节约了区域中心节点能耗, 尤其是扩展了该方法在大规模网络中的适用性.

## 6 实验与分析

### 6.1 数据集

为了验证 RCS 方法的有效性, 本文实验除了使用 CitySee 系统采样到的不同位置的温度, 还使用了太平洋 (北纬 7.0 度, 西经 180 度) 里收集到的不同深度的温度. 我们将两组数据分别进行验证, 然后将实验结果取平均值. 在实验时, 选择稀疏基前若干个最大系数以满足 CS 要求. 两个数据集均包含 1000 个数据值且每个数据集都执行 40 次实验.

### 6.2 稀疏变换

由于数据集的曲线是较为平滑的, 没有剧烈的起伏和跳变, 因此可以用一些信号变换方法对其进行稀疏呈现, 诸如离散小波变换和离散傅里叶变换等. 然而, 它们在进行稀疏变换时需要一些额外的限制条件<sup>[29]</sup>, 例如原始信号的维度必须满足 2 的幂, 这使得在稀疏变换前需要对信号进行周期性延拓或者增加若干个无效的“0”, 而离散余弦变换 (Discrete Cosine Transform, DCT) 则没有这些限制, 且 DCT 变换所需的计算能耗较小. 因此, 本文使用 DCT 作为数据集的稀疏变换.

### 6.3 重构准确度的测度标准

为了验证 RCS 的重构准确度, 本文实验使用

的测度标准是归一化均方误差 (Normalized Mean Square Error, NMSE), 其定义如下:

$$NMSE = \frac{\sum_{i=1}^n (x'_i - x_i)^2}{\sum_{i=1}^n x_i^2},$$

其中,  $x$  和  $x'$  分别表示原始信号和重构信号.

## 6.4 实验结果

### 6.4.1 RCS 与当前方法

本节对比 RCS 与当前典型的传输方法, 分别是直接传输 (Direct Transmission, DT)、DCS 和混合算法 (Mixed Algorithm, MA). 通过调整网络节点数、数据包的标识部分和数据部分的大小, 对比它们的传输次数和能量消耗. 在实验中, 数据包中标识部分  $D_{id}$  的大小分别设为 4 Byte、6 Byte 和 8 Byte, 数据部分  $D_{data}$  的大小分别设为 1 Byte、2 Byte 和 3 Byte, 相关信息部分  $D_{RI}$  的大小固定为 10 Byte. 数据包最大负载量为  $P_{size}$  为 127 Byte. 网络区域大小的划分标准是通过比较式 (2) 和式 (3).

首先, 固定  $D_{id}$  和  $D_{data}$ , 验证 4 种方法的传输次数随着节点数的变化情况. 本实验将  $D_{id}$  和  $D_{data}$  分别设为 8 Byte 和 1 Byte, 并且节点数从 20 增加到 1000. 在图 5 中, 横坐标表示节点数, 纵坐标表示传输次数. 令  $T_n(\cdot)$  表示网络节点数为  $n$  时的传输次数. 对 DT 来说,  $T_n(\text{DT})$  与节点数  $n$  是近似线性关系, 这是因为 DT 的传输方式是由中间节点直接转发数据. 对 DCS 来说, 当节点数为 20 时,  $T_n(\text{DCS})$  多于  $T_n(\text{DT})$ . 随着节点数增加至 100 时,  $T_n(\text{DCS})$  开始小于  $T_n(\text{DT})$ , DCS 的优势逐渐显现. 然而, 当节点数达到 421 个, 数据包的传输次数突然急剧增大, 原因是测量数  $M$  与  $D_{id} + D_{data}$  的乘积大于  $P_{size}$ , 此时的节点数称之为临界点<sup>[6]</sup>. 在这种情况下, 区域节点必须使用 2 个数据包来装载测量值. 尤其是当

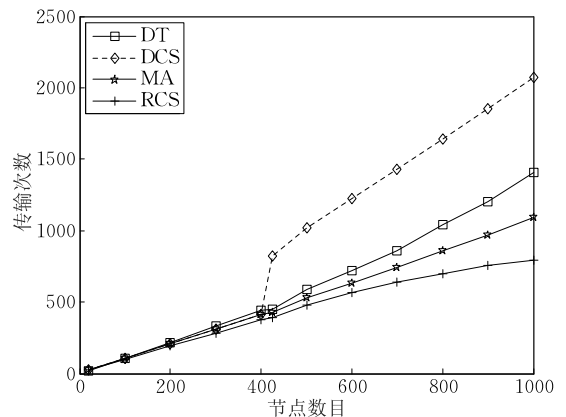


图 5 对比 DT、DCS、MA 和 RCS 的传输次数



节点数达到 421 时, MA 的传输次数并没有增大一倍, 而是在线性增长的基础上仅有少量增加. 由于 RCS 利用区域化传输, 在每个区域内都采用 DT 和 CS 相结合, 因此它的性能优于其他 3 种方法. 随着网络规模的扩大, 该方法的性能优势更加明显. 根据图 5, 当节点数大于 800 时, RCS 的传输次数曲线开始变得平滑, 这说明随着网络节点继续增加, 基于 RCS 方法的传输次数不会剧烈增长. 更重要的是, 随着网络密度的增加,  $T_n$ (RCS) 会逐渐少于节点数, 即整个网络的平均传输次数小于 1.

然后, 固定数据包中  $D_{data}$  的值, 验证  $D_{id}$  对传输能耗的影响. 令  $D_{data} = 1$  Byte,  $D_{id}$  从 4 Byte 增加到 8 Byte. 实验中, 假设对单个节点来说, 其传输能耗与数据包的传输数量和大小正相关. 在图 6 中, 横坐标表示节点数, 纵坐标表示能量消耗. 根据实验结果, 当  $D_{data}$  变大时, DT 能量消耗的增加最大, 这是因为该方法需要所有节点在发送数据包时加上自身的节点标识. 由于每个节点的数据包都增加, 导致中间节点的转发能耗急剧增加. 对 DCS 来说, 能量消耗并没有随  $D_{id}$  的增长而增加, 这是因为使用 DCS 进行数据收集时, 数据包不需要包含  $D_{id}$ . 根据压缩感知理论, 只要事先设定传感器的序号, 则 Sink 节点接收到足够的测量值后就能精确恢复出每个传感器的采样值. 这意味着  $D_{id}$  的变化对基于 DCS 方法的传输能耗没有影响. 对 MA 来说, 中间节点的传输方式是通过对比式 (2) 和式 (3) 的值进行选择的, 显然, 后者不受  $D_{id}$  变化的影响, 因此当  $D_{id}$  增加时,  $N(\text{children})$  变小, 使形成的中心区域增大, 导致网络传输能耗变大. 实验结果显示 RCS 性能总体上优于其他方法. 当  $D_{id}$  值较小时,  $N(\text{children})$  较大, 划分的区域数较小. 此时, 区域中心节点数较少, 因此其传输能耗也较少. 随着  $D_{id}$  变大, 区域数量变多, 消

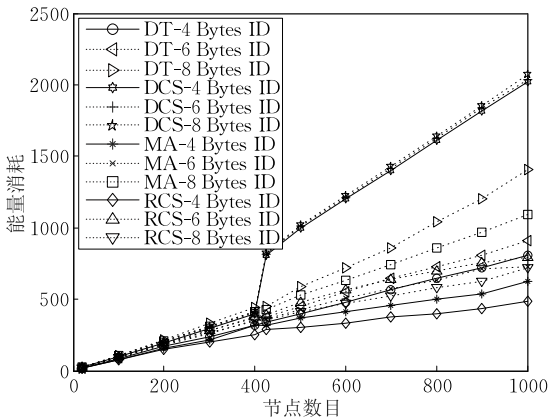


图 6 对比 4 种方法在  $D_{data} = 4, 6, 8$  字节时的传输能耗

耗也逐步增加, 但增加量是有限的. 从图中可以看出, 并不是  $D_{id}$  的值越大, 能量消耗越多. 根据上文, RCS 的能量消耗依赖于  $N(\text{children})$ 、 $|R(c)|$  和  $\sum_{c \in S} Hop(c)$ . 当  $D_{id}$  的值从 6 Byte 增加到 8 Byte 时, 虽然  $N(\text{children})$  减小, 使  $|R(c)|$  增加, 但  $\sum_{c \in S} Hop(c)$  却变少了, 因此网络的整体能耗并没有增加. 总的来说, 当  $D_{id}$  增加时, RCS 的传输能耗低于其他方法.

最后, 固定  $D_{id}$ , 通过改变  $D_{data}$  的大小, 对比 4 种方法的传输能耗. 图 7 显示的是  $D_{data}$  值分别为 1 Byte、2 Byte 和 3 Byte 时的结果. 其中, 横坐标表示节点数, 纵坐标表示能量消耗. 对于 DT,  $D_{data}$  值的改变对网络的传输能耗影响较小, 因为该值在整个数据包的比重较小. 相反地, 对 DCS 和 MA 来说, 由于  $D_{data}$  的变化会导致临界点的改变, 因此会极大影响传输过程的能耗. 尤其当  $D_{data} = 3$  Byte 时, MA 与 DT 相比没有较大优势. 结果显示  $D_{data}$  变化时, RCS 的性能总体上仍然优于 DT、DCS 和 MA.

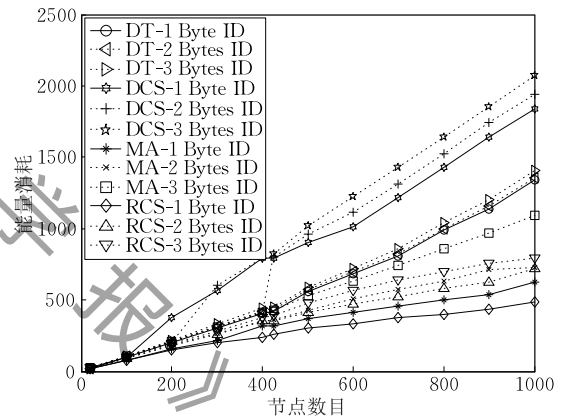


图 7 对比 4 种方法在  $D_{id} = 1, 2, 3$  字节时的传输能耗

根据上述实验, 随着节点数的增加, RCS 的传输量增长趋势逐渐趋于平缓, 其增长幅度是稳定的. 同时, 在  $D_{id}$  和  $D_{data}$  发生变化时, 基于 RCS 方法的传输能耗均低于其他方法.

#### 6.4.2 区域测量矩阵设计原则

为验证本文区域测量矩阵设计原则以及所使用的矩阵的有效性, 我们对比了 CS 应用中典型的测量矩阵, 分别是高斯 (Gauss) 随机矩阵、均匀 (uniform) 分布矩阵和伯努利 (Bernoulli) 矩阵<sup>[3, 18-19]</sup>.

实验中, 令矩阵  $\Phi$  的参数  $\delta = 2$  和 3, 高斯随机矩阵和均匀分布矩阵的期望值  $E(\text{Gauss}) = E(\text{uniform}) = 0$  且方差  $D(\text{Gauss}) = D(\text{uniform}) = 10$ , 伯努利矩阵的元素从集合  $\{-1, 1\}$  中选取且选择概率均等. 如图 8 所示, 横坐标表示节点数目, 纵

标则显示了网络进行一次测量所需传输次数。根据实验结果,使用矩阵  $\Phi$  ( $\delta=2$  和  $\delta=3$ ) 作为测量矩阵时网络的传输次数近似相等,均少于其他矩阵,并且  $T_n(\Phi)$  约等于节点数。这意味着在数据传输过程中节点生成的数据包没有超过最大载荷量,即每个节点只需要发送一个数据包。对伯努利矩阵和均匀分布矩阵来说,  $T_n(\text{Bernoulli})$  和  $T_n(\text{uniform})$  略大于  $T_n(\Phi)$ 。随着网络规模的增加,它们所增加的传输次数超过节点的增长,其中  $T_n(\text{uniform})$  的增幅更大。对高斯矩阵来说,  $T_n(\text{Gauss})$  最大。随着节点数的增加,甚至达到  $T_n(\Phi)$  的 2~3 倍。

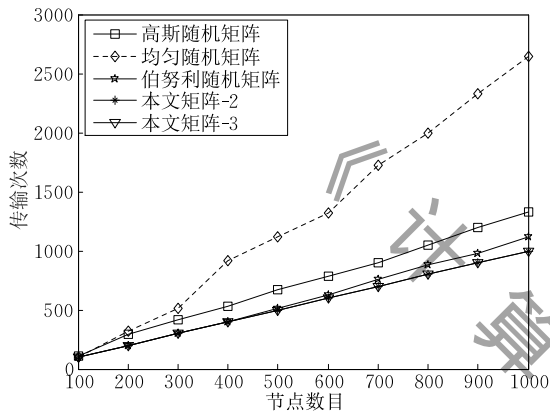


图 8 对比 4 种测量矩阵所需的传输次数

由于一次测量所需的传输次数不会少于网络中的节点数量,要进一步降低传输能耗,可以考虑通过增大  $\delta$  的值来减少数据包的长度。然而,较大的  $\delta$  可能使得产生的测量值没有足够的投影,导致重构精度不够。图 9 展示了这一现象,图中横纵坐标分别表示节点数和重构结果的归一化均方差。

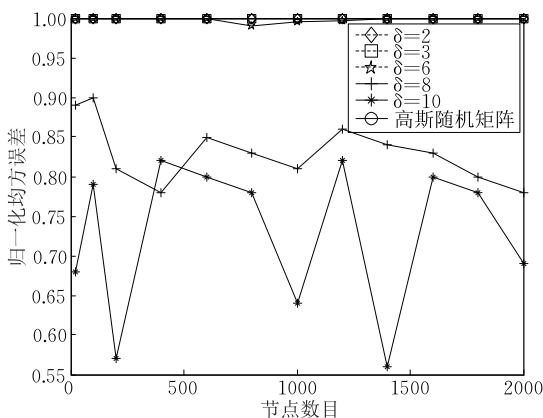


图 9  $\delta=2, 3, 6, 8$  和 10 的重构准确度(基于 Gauss 矩阵)

实验使用高斯随机矩阵且重构精度满足要求时所需的测量数,然后使用相同的测量数,得到  $\delta$  在不同取值时得到的重构结果,之后对比重构结果的 NMSE。其中,  $\delta$  分别为 2、3、6、8、10。图中显示当  $\delta$

值较小时,矩阵  $\Phi$  与高斯随机矩阵的重构精度是相似的。然而,当  $\delta=8$  和 10 时,矩阵  $\Phi$  的重构准确度急剧降低。这是因为随着  $\delta$  的增加,矩阵  $\Phi$  中的非零元素增多,使得区域中心节点在生成区域测量值时,很多采样值的系数为零,导致 Sink 节点得到的测量值在对应的稀疏基上没有足够的投影,因此这些测量值无法用于恢复信号。

由上述实验可知,虽然  $\delta$  增加可减少存储空间,却影响了重构精度,因此在实际应用中要根据具体要求来决定  $\delta$  的取值,既要保证 Sink 节点可精确重构,又尽可能降低节点所需空间以减少传输次数。

#### 6.4.3 采样停止原则

矩阵  $\Phi$  中参数  $\delta$  的取值还影响采样停止前相同重构的次数,即定理 2 中的参数  $l$ 。为验证采样停止原则的适用性,本文对比了不同  $\delta$  值时参数  $l$  的值与重构精度 NMSE 的关系,如图 10 所示。图中的横坐标表示有相同重构结果的节点数,纵坐标表示重构的归一化均方差。根据图 10 可知,  $\delta$  对重构准确度有很大影响:当  $\delta$  较小时,只要 10 ( $\delta=2$ ) 到 15 ( $\delta=3$ ) 个相同重构就可以得到准确的重构结果 ( $>90\%$ )。随着  $\delta$  的增加,所需要的  $l$  值越来越大。当  $\delta=5$  时,为保证重构精度达到 90%,  $l$  值需不低于 23,极大增加了传输能耗。这意味着增大参数  $\delta$  虽然可以减少发送的数据包长度,但所需的测量次数会相应增加。

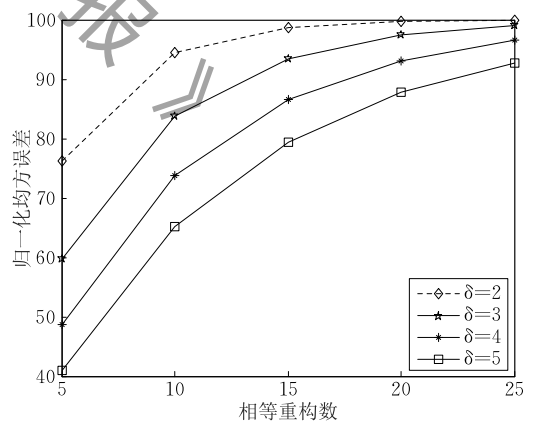


图 10  $\delta=2, 3, 4, 5$  时参数  $l$  与重构准确度的关系

由上述实验可知,参数  $\delta$  的大小会影响生成的数据包长度(进而影响节点发送数据包的个数),以及所需的测量次数。在实际应用中要根据不同的应用场景、环境和具体要求选择合适的参数构建区域测量矩阵时。由实验可知,一般  $\delta$  可为 2 或 3。

根据 CS 理论,测量数依赖原始信号的稀疏度 ( $M \geq c \cdot k$ ),因此一般需要事先给定测量数。当预设

的采样量较少时, 重构准确度较低; 反之, 虽保证了准确重构, 却增大了网络的传输能耗. 根据 CS 的相关应用<sup>[19]</sup>, 测量数需是信号稀疏度的 3~4 倍才可能实现准确重构, 即  $c=3$  和 4. 然而, 由于感知环境的复杂性和多变性, 实际上这仍然难以确保重构精度. 如图 11 所示, 本文使用高斯矩阵作为测量矩阵, 验证了  $c=2, 3, 4, 5$  时的重构精度, 同时展示了  $\delta=2, l=16$  时的重构结果. 其中, 横坐标表示网络节点数, 纵坐标表示重构结果的归一化均方误差. 为了显示  $c$  在不同取值下的重构准确度, 实验使用了多组信号. 从图中看出, 当  $c$  值越大, 准确重构的机率越大, 但即使  $c=5$  仍然无法保证对每组信号都能准确重构. 另一方面, 对不同信号来说, 所需的  $c$  值并不是等同的, 这说明预先固定测量次数会增加不必要测量次数, 从而增大传输能耗. 相应地, 当  $\delta=2, l=16$  时, 可以得到稳定的重构结果.

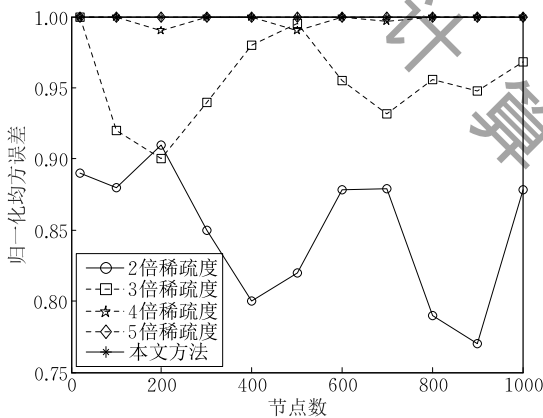


图 11  $c=2, 3, 4, 5$  的重构准确度(基于 Gauss 矩阵)

总的来说, 本文所提出的区域测量矩阵设计原则和采样停止原则是实用的, 可以在保证采样精度的同时, 有效地降低网络能耗. 尤其当网络规模较大时, 能够减少更多不必要的传输能耗, 延长整个网络生命周期.

## 7 结 论

本文提出了一个基于区域化压缩感知的数据收集方法 RCS, 以满足无线传感器网络环境下能量有效的采样. 根据该方法, 每个区域独立执行压缩感知过程, 避免在数据传输过程形成较大的数据中心, 且充分利用了直接传输和压缩感知方法的优势. 为进一步提高该方法的有效性, 本文针对实际应用中的两个问题给出了相应的解决方案, 包括区域测量矩阵的设计和采样停止的时机选择. 前者减少了区域

中心节点的存储需求和产生的数据包长度, 从而降低传输能耗. 后者通过当前重构结果来估计重构的精度, 避免了不必要的测量. 本文实验使用实际采样数据, 验证了 RCS 方法优于典型的传输方法, 同时也验证了本文提出的区域测量矩阵的设计原则和采样停止原则是有效的, 可以进一步优化 RCS 方法的传输能耗.

致 谢 审稿专家与编辑给出了宝贵的评语和建议, 在此表示感谢!

## 参 考 文 献

- [1] Al-Karaki J N, Kamal A E. Routing techniques in wireless sensor networks: A survey. *IEEE Wireless Communications*, 2004, 11(6): 6-28
- [2] Donoho D L. Compressed sensing. *IEEE Transactions on Information Theory*, 2006, 52(4): 1289-1306
- [3] Candès E J, Wakin M B. An introduction to compressive sampling. *IEEE Signal Processing Magazine*, 2008, 25(2): 21-30
- [4] Luo C, Wu F, Sun J, et al. Efficient measurement generation and pervasive sparsity for compressive data gathering. *IEEE Transactions on Wireless Communications*, 2010, 9(12): 3728-3738
- [5] Luo J, Xiang L, Rosenberg C. Does compressed sensing improve the throughput of wireless sensor networks?// *Proceedings of the 2010 IEEE International Conference on Communications*. Cape Town, South Africa, 2010: 1-6
- [6] Caione C, Brunelli D, Benini L. Distributed compressive sampling for lifetime optimization in dense wireless sensor networks. *IEEE Transactions on Industrial Informatics*, 2012, 8(1): 30-40
- [7] Caione C, Brunelli D, Benini L. Compressive sensing optimization for signal ensembles in WSNs. *IEEE Transactions on Industrial Informatics*, 2014, 10(1): 382-392
- [8] Wang J, Tang S, Yin B, et al. Data gathering in wireless sensor networks through intelligent compressive sensing// *Proceedings of the 2012 IEEE International Conference on Computer Communications*. Orlando, USA, 2012: 603-611
- [9] Liu Fang, Wu Jiao, Yang Shu-Yuan, et al. Research advances on structured compressive sensing. *Acta Automatica Sinica*, 2013, 39(12): 1980-1995(in Chinese)  
(刘芳, 武娇, 杨淑媛等. 结构化压缩感知研究进展. *自动化学报*, 2013, 39(12): 1980-1995)
- [10] Xu Zhi-Qiang. Compressed sensing: A survey. *Scientia Sinica Mathematica*, 2012, 42(9): 865-877(in Chinese)  
(许志强. 压缩感知. *中国科学 数学(中文版)*, 2012, 42(9): 865-877)

- [11] Dai Qiong-Hai, Fu Chang-Jun, Ji Xiang-Yang. Research on compressed sensing. *Chinese Journal of Computers*, 2011, 34(3): 425-434(in Chinese)  
(戴琼海, 付长军, 季向阳. 压缩感知研究. *计算机学报*, 2011, 34(3): 425-434)
- [12] Baron D, Wakin M B, Duarte M F, et al. Distributed compressed sensing. *IEEE Transactions on Information Theory*, 2006, 52(12): 5406-5425
- [13] Yang H, Huang L, Xu H. Distributed compressed sensing in vehicular ad-hoc network. *Ad Hoc & Sensor Wireless Networks*, 2015, 25(1/2): 121-145
- [14] Liu Rui, Liang Yao, Zhong Xiao-Yang. Poster: Compressed sensing inspired approaches for path reconstruction in wireless sensor networks//*Proceedings of the 16th ACM International Symposium on Mobile Ad Hoc Networking and Computing*. Hangzhou, China, 2015: 379-380
- [15] Khedr A M. Effective data acquisition protocol for multi-hop heterogeneous wireless sensor networks using compressive sensing. *Algorithms*, 2015, 8(4): 910-928
- [16] Zheng H, Yang F, Tian X, et al. Data gathering with compressive sensing in wireless sensor networks: A random walk based approach. *IEEE Transactions on Parallel and Distributed Systems*, 2015, 26(1): 35-44
- [17] Zhan J, Vaswani N. Time invariant error bounds for modified-CS-based sparse signal sequence recovery. *IEEE Transactions on Information Theory*, 2015, 61(3): 1389-1409
- [18] Jacques L. A short note on compressed sensing with partially known signal support. *Signal Processing*, 2010, 90(12): 3308-3312
- [19] Pati N, Pradhan A, Kanoje L K, et al. An approach to image compression by using sparse approximation technique. *Procedia Computer Science*, 2015, 48: 770-776
- [20] Sahoo S K, Makur A. Signal recovery from random measurements via extended orthogonal matching pursuit. *IEEE Transactions on Signal Processing*, 2015, 63(10): 2572-2581
- [21] Wu Q, Zhang Y D, Amin M G, et al. Complex multitask Bayesian compressive sensing//*Proceedings of the 2014 IEEE International Conference on Acoustics, Speech and Signal Processing*. Florence, Italy, 2014: 3375-3379
- [22] Wu Q, Zhang Y D, Amin M G, et al. Multi-task bayesian compressive sensing exploiting intra-task dependency. *IEEE Signal Processing Letters*, 2015, 22(4): 430-434
- [23] Mun S, Fowler J E. Block compressed sensing of images using directional transforms//*Proceedings of the 2009 IEEE International Conference on Image Processing*. Cairo, Egypt, 2009: 3021-3024
- [24] Vuran M C, Akyildiz I F. *Wireless Sensor Networks*. Hoboken, America: John Wiley & Sons, 2010
- [25] Candes E J, Tao T. Near-optimal signal recovery from random projections: Universal encoding strategies? *IEEE Transactions on Information Theory*, 2006, 52(12): 5406-5425
- [26] Baraniuk R, Davenport M, DeVore R, et al. A simple proof of the restricted isometry property for random matrices. *Constructive Approximation*, 2008, 28(3): 253-263
- [27] Romik D. Stirling's approximation for  $n!$ : The ultimate short proof? *The American Mathematical Monthly*, 2000, 107(6): 556
- [28] Malioutov D M, Sanghavi S R, Willsky A S. Sequential compressed sensing. *IEEE Journal of Selected Topics in Signal Processing*, 2010, 4(2): 435-444
- [29] Xu X, Ansari R, Khokhar A. Hierarchical data aggregation using compressive sensing (HDACS) in WSNs. *ACM Transactions on Sensor Networks*, 2015, 11(3): 45



**YANG Hao**, born in 1980, Ph. D., associate professor. His current research interests include compressive sensing, wireless network.

**WANG Xi-Wei**, born in 1984, Ph. D., assistant professor. His current research interests include data analyze and data mining.

## Background

In the process of data transmission in wireless sensor networks, a vital problem is that a center region near the sink will be created and the sensors in it require vast energy to receive or relay data packages form other sensors, especially in large-scale sensor networks. Although some particular methods can reduce consumption, complex operations usually cost huge

energy, and some special requirements also limit their universality in actual applications. To feature energy-efficient communications, compressive sensing (CS) has been utilized to extend the lifetime of networks.

As CS has drawn more and more attention in WSNs, however, it is still doubtful that whether this technique

always outperforms traditional methods. Previous literatures have demonstrated that in some cases, plain CS is not necessarily better than non-CS schemes, due to the fact that edge nodes cost more energy in CS than in non-CS.

In this paper, we propose a regionalized compressive sensing technique for energy-efficient data gathering. The basic idea is, the CS sampling procedure is carried out by partitioning the topology the network. More specifically, the network is firstly divided into several blocks in a random manner without requiring any relationships (with respect to spatial or temporal factors) or features (e. g. , data correlation) among sensors. On top of that, center sensors of each region are elected with randomization. Subsequently, CS is implemented in each region separately with its regional measurement matrix. In detail, every center sensor collects data of the same region and generates the weighted sums as part of samplings using CS technique. Lastly, samplings in all regions are sent directly to the sink for reconstruction. By doing this, the measuring operation based on CS is executed by center sensors respectively. Hence, the advantage of direct data transmission over plain CS can be taken in each region, which makes our scheme have a combined superiority for the entire network. Experiments that examine different package parameters illustrate

that our RCS outperforms direct transmission, DCS, and the mixed algorithm.

To further improve our scheme, we address two practical issues. One is that the design principle is proposed to reduce storage space of regional center sensors and the number of data transmissions. The other is that a sampling stop rule is established to avoid unnecessary samplings.

This work is supported by the National High Technology Research and Development Program (863 Program) of China (2015AA01A201), National Natural Science Foundation of China under Grant No. 61402394, 61379064, 61303209, National Natural Science Foundation of Jiangsu Province of China under Grant No. BK20140462, Natural Science Foundation of the Higher Education Institutions of Jiangsu Province of China under Grant No. 14KJB520040, China Postdoctoral Science Foundation funded project under Grant No. 2016M591922, Jiangsu Planned Projects for Postdoctoral Research Funds under Grant No. 1601162B, Prospective Joint Research project of Jiangsu Province under Grant No. BY2016066-05, Industry university research project of Jiangsu Province under Grant No. BY2016066-04, and sponsored by Qing Lan Project.