

中图分类号: TP393

论文编号: 10006RY0806119

北京航空航天大学  
硕士学位论文

无线传感器网络数据处理与能耗  
优化策略研究

作者姓名 熊君武  
指导教师 李欢 副教授  
培养院系 计算机学院

# **Research of Energy Efficiency Optimization Strategy with Data Aggregation for Wireless Sensor Networks**

A Dissertation Submitted for the Degree of Master

**Candidate: Xiong Junwu**

**Supervisor: Associate Prof. Li Huan**

School of Computer Science and Engineering

Beihang University, Beijing, China

## 关于学位论文的独创性声明

本人郑重声明：所提交的论文是本人在指导教师指导下独立进行研究工作所取得的成果，论文中有关资料和数据是实事求是的。尽我所知，除文中已经加以标注和致谢外，本论文不包含其他人已经发表或撰写的研究成果，也不包含本人或他人为获得北京航空航天大学或其它教育机构的学位或学历证书而使用过的材料。与我一同工作的同志对研究所做的任何贡献均已在论文中做出了明确的说明。

若有不实之处，本人愿意承担相关法律责任。

学位论文作者签名：\_\_\_\_\_

日期： 年 月 日

## 学位论文使用授权书

本人完全同意北京航空航天大学有权使用本学位论文（包括但不限于其印刷版和电子版），使用方式包括但不限于：保留学位论文，按规定向国家有关部门（机构）送交学位论文，以学术交流为目的赠送和交换学位论文，允许学位论文被查阅、借阅和复印，将学位论文的全部或部分内容编入有关数据库进行检索，采用影印、缩印或其他复制手段保存学位论文。

保密学位论文在解密后的使用授权同上。

学位论文作者签名：\_\_\_\_\_

日期： 年 月 日

指导教师签名：\_\_\_\_\_

日期： 年 月 日



## 摘 要

物感网的兴起和蓬勃发展使得无线传感器网络的应用更加广泛，并且成为当前学术界和工业界关注的焦点。在无线传感器网络中，传感器节点资源受限尤其是能量资源受限问题成为目前制约传感器大规模部署与应用的主要瓶颈之一。通过对目前国内外无线传感器网络数据处理与能耗优化策略研究进行总结和分析发现，尽管不同数据处理策略对无线传感器网络生命周期的延长有着不同的影响，能够在一定程度上延长网络的生命周期，但是大多数方法，对于解决大规模同构网络下的能耗问题存在一定局限，数据在多跳累积导致靠近汇聚节点区域的部分节点过早死亡，网络过早瘫痪、网络节点能量利用率低下的问题依然存在。针对大规模的无线传感器网络的能耗问题，本文做了如下几个方面的工作：

首先，基于多跳分簇机制，结合对网络拓扑、多跳簇间距离、簇头数目及尺寸分布、数据融合等影响网络数据处理与能耗的因素进行定量分析，寻找最优的网络跳数以提高簇头节点的能量利用率，同时求解各跳子网内簇头数以保证全网所有总簇头耗能均衡，并在此基础上进一步分析了为提高网络簇头节点能量利用率簇尺寸分布所应满足的规律以及数据融合比簇头数分布的影响。

其次，对本文设计的无线传感器数据处理与能耗优化策略进行了仿真，并结合已有的被证明为有效的动态建簇、动态路由以及簇头轮转策略来设计了仿真协议，并对网络生命周期、网络节点剩余能量以及全网的数据融合比对网络能量利用率的影响等方面进行了大量仿真。通过对仿真结果进行了对比和分析，本文提出的数据处理和能耗优化策略，与之前同类能耗优化协议 UCR(Unequal Cluster-based Routing)协议相比，能够在不同的网络规模和不同的网络生命周期定义下，在很大程度上保持网络节点能耗均衡的同时提高网络能量利用效率，对不同网络规模具有可扩展性。同时发现，不同的数据融合比能够在一定程度上影响网络能耗均衡及网络能量利用率。

最后，论证了本文提出的数据处理及能耗优化策略建模分析过程中涉及基于矩形的各种子网拓扑原型方式的合理性及其在任意大规模实际应用部署场景下的扩展和组织部署的方案，以期无线传感器网络大规模部署和实际应用提供理论上的依据和实际应用的指导。

**关键词：**大规模无线传感器网络，能耗优化，多跳分簇机制，数据处理

## **Abstract**

The flourishing of wireless sensor network (WSN) has become the focus of both academic and industrial nowadays with the rapid development of CPS (Cyber-Physical System). However, sensor nodes' constrained resources, especial limited energy budget is still a bottle-neck blocking in the application of large-scale WSN. And how to reduce the consumption of sensors' precious energy resources through sensor data aggregation and offset or avoid the uneven energy depletion problem to extend the lifetime and promote the application of WSN in large-scale scenarios is a hot topic currently in the research of WSN.

It has been found in related work that although many existing data aggregation and energy optimization techniques can prolong the network lifetime to some extent, problems of low energy efficiency, uneven energy depletion among area near the sink node in large-scale homogeneous WSN are still challenging us. Besides, the impact of different data aggregation ratios on network topology, even energy consumption and energy efficiency has not been considered carefully. However, it has been concluded that clustering and multi-hop are comparable effective strategies to improve network scalability and energy efficiency by optimizing network structure and data aggregation among different nodes and have been adopted as the basic techniques of this thesis. And to address these problems for a given large-scale homogeneous WSN, the main contributions of this thesis lie in the following several aspects.

Firstly, based on clustering techniques, a theoretical model has been given to fully analyzing the effect of network topology, inter-cluster multi-hop relay distance, cluster sizes distribution and data aggregation ratio on the energy efficiency. An effective unit-based network topology prototype has been adopted and verified to qualify these parameters. Cluster heads in each unit with no communication with each another has the same hop number and experienced the same hops to the sink node. And the inter-cluster hop number has been found by optimizing the total energy consumption of all clusters in a round. Meanwhile, cluster number and sizes have been obtained to achieve even energy depletion among different areas of the network. Besides, the effects of different data aggregation ratios on the cluster size distribution and network energy consumption have been given.

Secondly, the performance of the data aggregation and energy consumption strategy in this thesis has been simulated with adopting other existing but effective techniques, such as cluster

heads rotation, dynamic clustering and dynamic routing strategies etc to further improve the energy efficiency among all nodes in the whole network. Extensive simulations have been conducted and the results demonstrated our model is very energy saving, scalable and robust compared with UCR in terms of lifetime for large-scale homogeneous WSN. What's more, it has been simulated that different data aggregation ratios play a great role on even energy depletion and energy efficiency.

Finally, different rectangle-based sub-network topologies and a much more general deployed scenario are discussed. It has verified the network prototypes adopted in this thesis and provides a helpful guide for the practical deployment in large-scale homogeneous WSN scenarios.

**Key words:** Large-scale WSN, Energy consumption optimization, Multi-hop clustering techniques , Data aggregation

# 目 录

第一章 绪论 .....	1
1.1 研究背景 .....	1
1.1.1 物感网的发展 .....	1
1.1.2 无线传感器网络应用前景 .....	2
1.1.3 大规模无线传感器网络能耗问题 .....	3
1.1.4 课题来源 .....	4
1.2 国内外研究现状 .....	4
1.2.1 网络数据聚集 .....	4
1.2.2 网络数据汇聚模式 .....	5
1.2.3 网络数据融合模式 .....	5
1.3 论文研究内容和目标 .....	6
1.4 论文组织结构 .....	6
第二章 基于数据处理的无线传感器网络能耗优化协议 .....	8
2.1 基于平面结构 .....	8
2.2 基于分簇结构 .....	8
2.2.1 LEACH 协议 .....	9
2.2.2 HEED 协议 .....	11
2.2.3 UCR 协议 .....	11
2.2.4 SCA 协议 .....	12
2.3 基于链结构 .....	14
2.4 基于树结构 .....	15
2.5 本章小结 .....	17
2.5.1 平面网络与结构化网络 .....	18
2.5.2 单跳与多跳数据传输 .....	19
2.5.3 网络数据融合 .....	19
2.5.4 结论 .....	20
第三章 无线传感器网络数据处理与能耗优化策略建模分析 .....	21
3.1 系统模型 .....	21
3.1.1 参数说明 .....	21
3.1.2 无线传感器网络模型和传感器节点模型 .....	23
3.1.3 传感器节点传感数据处理能耗模型 .....	24
3.1.4 网络生命周期定义 .....	25

3.2 大规模无线传感器数据处理与网络能耗问题 .....	26
3.2.1 大规模无线传感器网络的能耗问题 .....	26
3.2.2 大规模无线传感器网络数据处理与能耗优化策略 .....	29
3.3 无线传感器网络簇头节点总能耗最小策略 .....	30
3.3.1 子网内簇头数据及能耗负载分析 .....	30
3.3.2 大规模无线传感器网络多跳传输距离期望 .....	31
3.3.3 网络簇头总能耗最小及最优总簇间跳数 .....	33
3.4 大规模多跳无线传感器网络簇头能耗均衡策略 .....	35
3.4.1 网络簇头能耗均衡影响因素分析 .....	36
3.4.2 网络簇头数与能耗均衡 .....	36
3.4.3 网络簇尺寸分布、数据融合比与能耗均衡 .....	37
3.5 本章小结 .....	42
<b>第四章 模型仿真及性能分析 .....</b>	<b>43</b>
4.1 仿真平台和系统参数设置 .....	43
4.1.1 仿真平台 .....	43
4.1.2 系统参数设置 .....	43
4.1.3 COCA 协议参数设置 .....	43
4.1.4 UCR 协议参数设置 .....	44
4.2 模型与协议仿真设计 .....	45
4.2.1 网络模型初始化建立 .....	45
4.2.2 子网内部簇头选取、簇结构形成及簇内数据处理 .....	46
4.2.3 子网簇间数据处理动态能量感知路由 .....	46
4.2.4 子网簇间数据多跳转发 .....	47
4.3 大规模多跳网络生命周期 .....	47
4.3.1 簇头最佳轮转周期选取 .....	47
4.3.2 大规模多跳网络生命周期 .....	49
4.3.3 协议稳定性分析 .....	51
4.4 网络节点能量利用率 .....	53
4.4.1 不同网络规模的节点剩余能量 .....	53
4.4.2 数据融合比与网络节点剩余能量 .....	54
4.5 本章小结 .....	55
<b>第五章 模型与协议扩展及实际部署应用 .....</b>	<b>56</b>
5.1 不同子网拓扑分析 .....	56
5.1.1 递增子网模型 .....	56
5.1.2 递减子网模型 .....	56
5.1.3 均匀子网模型 .....	57

5.2 大规模网络实际部署建议方案 .....	58
5.3 本章小结 .....	59
总结与展望 .....	60
本文主要工作总结 .....	60
后续工作展望 .....	61
参考文献 .....	62
附录 A .....	66
附录 B .....	68
攻读硕士学位期间取得的学术成果 .....	70
致 谢 .....	71

## 图 目

图 1 NSF 关于物感实际应用的构想.....	1
图 2 大规模多跳无线传感器网络简单应用部署场景 .....	4
图 3 Directed Diffusion 的简单数据路由模式 .....	8
图 4 LEACH 协议执行流程.....	9
图 5 全网簇头的总耗能和最优子网数目之间的关系 .....	13
图 6 网络规模和最优的子网数目之间的关系 .....	13
图 7 不同子网区域内最优的簇头数目分布规律 .....	14
图 8 PEGSIS 协议网络部署场景.....	14
图 9 PEDAP 协议网络部署场景.....	16
图 10 网络节点的死亡时间(100*100m <sup>2</sup> 的网络, 基站远离网络部署区域) .....	17
图 11 网络节点的死亡时间(100*100m <sup>2</sup> 的网络, 基站处于网络部署区域内部)... ..	17
图 12 大规模无线传感器网络矩形部署区域 .....	23
图 13 传感器节点能耗模型 .....	24
图 14 多跳分簇结构化无线传感器网络部署场景 .....	29
图 15 同构无线传感器网络传感器节点部署场景 .....	32
图 16 不同数据融合比时簇头分布规律 .....	39
图 17 不同数据融合比时簇头分布规律 .....	40
图 18 数据融合比对单个子网内簇头数量增量的影响 (N=400) .....	41
图 19 数据融合比对单个子网内簇头数量增量的影响 (N=800) .....	42
图 20 第一个节点死亡定义下的簇头轮转周期与网络生命周期之间的关系 .....	48
图 21 部分节点死亡定义下的簇头轮转周期与网络生命周期之间的关系 .....	49
图 22 第一个节点死亡定义下网络生命周期的比较 .....	50
图 23 部分节点死亡定义下网络生命周期的比较 .....	51
图 24 COCA 两种不同定义下的生命周期的比较 .....	52
图 25 UCR 两种不同定义下的生命周期的比较.....	52
图 26 COCA 和 UCR 在头结点死亡和部分节点死亡情形下的网络剩余能量.....	54
图 27 COCA 在不同数据融合比时网络剩余能量 .....	55
图 28 递增子网模型 .....	57

图 29	递减子网模型 .....	57
图 30	正方形均匀子网模型 .....	58
图 31	短高与长宽不等的均匀子网模型 .....	58
图 32	长高与短宽不等的均匀子网模型 .....	58
图 33	大规模无线传感器网络实际应用部署场景 .....	59

## 表 目

表 1	经典算法协议总结 .....	19
表 2	模型参数说明 .....	21
表 3	模型参数说明（续） .....	22
表 4	模型参数说明（续） .....	23
表 5	系统模型相关参数设置 .....	44
表 6	无线传感器网络模型参数设置 .....	44
表 7	COCA 协议参数设置 .....	44
表 8	UCR 协议参数设置 .....	45



# 第一章 绪论

## 1.1 研究背景

### 1.1.1 物感网的发展

物感网，又称物联网或传感网。“传感网”在国际上通称“物联网”(IOT, Internet of Things)，即把所有物品通过射频识别等信息传感设备与互联网连接起来，实现智能化识别和管理，是继计算机、互联网与移动通信网之后的又一次信息产业浪潮，是一个全新的技术领域。物感网的构想最初是由美国军方提出的，美国国防部高级研究所计划署(DARPA, The Defense Advanced Research Projects Agency)于1978年开始资助卡耐基—梅隆大学进行分布式传感器网络的研究，这被看成是无线传感器网络的雏形。1999年，美国召开的移动计算和网络国际会议提出，物联网是下一个世纪人类面临的又一个发展机遇<sup>[1][2]</sup>。

传统物感网络应用非常广泛，包括航空航天、自动化、化学处理过程、民用设施、能源、医疗、建筑、运输、娱乐等领域及各种消费者群体中，如图1所示。这类系统通常被称为“嵌入式系统”，在该类系统中，更多强调计算设备本身的计算处理能力，对于相互间通信没有做出应有的关注<sup>[1][2]</sup>。

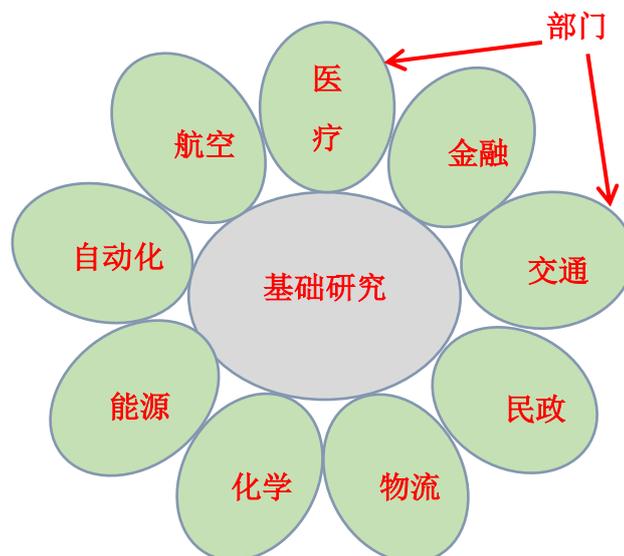


图1 NSF关于物感实际应用的构想

与传统物感网络不同，新一代物感网络强调所有计算设备组成的相互通信进行计算的网路整体，而不是仅仅着重于网络中独立设备。借助于未来几年内科学研究与工程领域内相关难点问题的攻克，物感网期待提升网络设备间相互通信质量，从而极大提升网

络的可适性，自组性，效率，功能，可靠性，安全性及可用性，将物感网的研究前景进一步扩展，包括干预（如冲突避免），精确性研究（如机器人手术和纳米制造工业），危险的或不可接近的环境施工（如，搜寻，火险救援，深海探索），协作（如航空调度，战场），效率（绿色建筑），及人员能力提升（如，医疗监控和物流）等<sup>[1] [3] [4] [5]</sup>。

2005 年国际电信联盟发布的《ITU 互联网报告 2005:物联网》（ITU, International Telecommunication Union），正式提出了“物联网”概念。报告指出，无所不在的“物联网”通信时代即将来临，世界上所有的物体从轮胎到牙刷、从房屋到纸巾都可以通过因特网主动进行交换。射频识别技术（RFID, Radio Frequency Identification）、传感器技术、纳米技术、智能嵌入技术将到更加广泛的应用，人类将进入“智慧地球”时代。

《福布斯杂志》2008 年预测：未来传感网比现在的因特网大得多，美国权威咨询机构 Forrester 预测，到 2020 年，世界上物物互联的业务，跟人与人通信的业务相比，将达到 30 比 1，因此，“物联网”被称为是下一个万亿元级的通信业务。根据预测，到 2035 年前后，我国的传感网终端将达到数千亿个；到 2050 年，传感器将在生活中无处不在。此外，著名的美国橡树岭国家实验室（OAK, Oak Ridge National Laboratory）断言：IT（Information technology）时代正从“计算机即网络”转变为“传感器即网络”<sup>[2] [6] [7]</sup>。

奥巴马就任美国总统后，对 IBM（International Business Machines）首席执行官彭明盛首次提出的“智慧地球”概念（建议政府投资新一代的智慧型基础设施）给予了积极的回应并上升至美国的国家战略，在世界范围内引起轰动。该战略认为 IT 产业下一阶段的任务是把新一代 IT 技术充分运用在各行各业之中，具体地说，就是把感应器嵌入和装备到电网、铁路、桥梁、隧道、公路、建筑、供水系统、大坝、油气管道等各种物体中，并且被普遍连接，形成“物联网”。此外，日本、韩国等也分别提出“U-Japan”（Ubiquitous-Japan）、“U-Korea”（Ubiquitous-Korea）战略<sup>[6] [7]</sup>。

### 1.1.2 无线传感器网络应用前景

无线传感器网络作为物感网产业的主要内容，在 2003 年分别被“MIT 技术评论”（MIT, Massachusetts Institute of Technology）列为改变世界的 10 大新技术之一和被美国商业周刊评为全球未来的四大高技术产业之一，在《国家中长期科学与技术发展规划（2006—2020 年）》和“新一代宽带移动无线通信网”重大专项中均将无线传感器网络列入重点研究领域。2009 年 8 月，温家宝总理在中科院无锡高新微纳传感网工程技术研发中心考察明确指示，“要早一点谋划未来，早一点攻破核心技术”，力争主导制定无线传

感器网络国际标准。现在,无论是美国提出的智慧地球,还是中国提出的感知中国,都把无线传感器网络产业放到一个宽广的坐标系中来考量<sup>[6] [7] [8]</sup>。

### 1.1.3 大规模无线传感器网络能耗问题

无线传感器网络由部署到目标区域的大量具有信息传感、数据采集、一定距离内无线通信功能的传感器节点组成,由汇聚节点(Sink node)对这些传感节点的数据进行收集,具有低成本、低功耗、部署灵活等特点<sup>[5]</sup>。同时,无线传感器网络具有这些优点的同时,传感器节点自身资源非常有限,主要表现在:

- 1) 传感器节点能量有限;
- 2) 传感器节点数据处理能力有限(常用节点的处理器速度为 4MHZ)和存储空间有限(128MB)<sup>[9]</sup>;
- 3) 无线传感器网络通信带宽有限,而且传感器节点的部署通常都比较密集且节点采集的数据量很大;
- 4) 传感器节点的通信半径有限,只能在一定距离内进行无线通信。

在无线传感器网络节点受限的资源中,能耗问题显得最为突出:因为无线传感器节点一般都采用电池作为能源,而网络大多部署在各种复杂环境中,且一经部署后节点不易再充电,这使得传感器节点的能耗问题成为制约无线传感器网络的应用和发展的瓶颈。具体来说,无线传感器节点的耗能主要在于三个方面:传感数据的采集、数据融合处理和数据传输三方面,其中用于数据传输的能耗将近大约占了节点总能耗的 80%<sup>[9]</sup>。尤其在大规模多跳(Multi-hop)无线传感器网络中,如图 2 所示,传感器节点采集的原始传感数据经过通过中继节点处理后由远及近地通过多跳才能传输到汇聚节点(Sink node)<sup>[10]</sup>,在这个过程中,由于网络传输数据不断累加,使得靠近汇聚节点区域内的节点要处理和传输的平均数据负载量激增,从而使得这部分区域的节点能耗加剧、导致这些区域内的节点过早死亡(或能量耗尽),使得整个网络瘫痪,无法满足用户的需求,出现常见的“热点区域(Hot-Spot)”问题<sup>[11]</sup>。因此,对于一个给定的大规模无线传感器网络中,在传感器节点硬件资源受限的条件下,如何更有效地利用传感器节点的宝贵资源(尤其是节点的能量资源),减少整个网络所有节点的能耗,提高多跳无线传感器网络传感器节点的能量利用效率以及保持全网能耗均衡、减缓或避免多跳无线传感器网络中的“热点区域(Hot-Spot)”问题,延长无线传感器网络的生命周期和增加网络的可扩展性,促进无线传感器网络大规模地部署和应用,成为当前无线传感器网络领域的主要问题之一,也是本

文的主要研究目标。

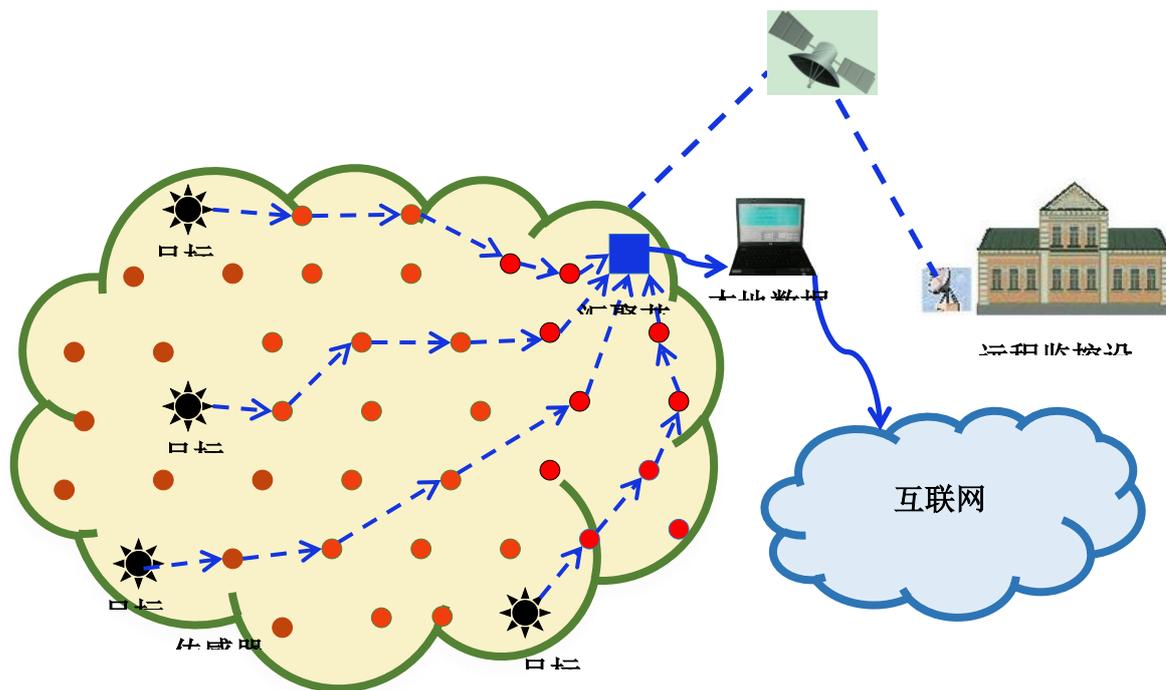


图 2 大规模多跳无线传感器网络简单应用部署场景

#### 1.1.4 课题来源

本课题受到国家自然科学基金（编号：60703056）、教育部留学回国人员科研启动基金资助项目（SRF for ROCS, SEM: Scientific Research Foundation for the Returned Overseas Chinese Scholars, State Education Ministry)和北航新进教师条件项目的支持。

### 1.2 国内外研究现状

针对大规模无线传感器网络中的能耗不均“热点”问题和能量利用率低下问题，国内外学者开展了广泛的学习和研究，提出了一些列基于能耗优化的传感数据处理策略<sup>[22]</sup>，主要有以下三个方面：

#### 1.2.1 网络数据聚集

数据聚集主要是指数据的采集、数据路由、传输和聚集方式，通过对数据聚集策略的研究，可以优化网络结构，使得网络能耗均匀和减少网络耗能，增加网络规模的可扩展性<sup>[24]</sup>。到目前为止，在无线传感器数据采集和数据路由方面，无线传感器网络的拓扑结构和数据路由已经由传统的平面网络结构发展到分簇、链状、树、网格等结构化的网络。国内外研究者们通常将原来的平面网络划分成多个较小规模的子区域，

即将网络进行分簇或形成链状、树状或网格等，从而形成结构化的网络。此外，通过改变无线传感器网络节点本身的部署密度也是一种避免能耗不均、提高网络能量利用率的有效方案<sup>[21]</sup>。

### 1.2.2 网络数据汇聚模式

在结构化的传感器网络中，传感器节点的数据传输和汇聚模式分为单跳传输模式(这里的一跳，指的是采集传感器的节点将数据传输到下一个数据转发节点或汇聚节点)和多跳传输模式<sup>[13]</sup>。在单跳的数据传输模式下，离汇聚节点(Sink node)较远的节点采集的数据无法直接传输到汇聚节点(Sink node)或者对于距离汇聚节点较远的节点由于较大的传输距离，节点能量会较离汇聚节点近的区域节点先耗尽，导致出现“热点”(Hot-spot)问题(此时的热点区域出现在距离汇聚节点较远的网络区域)。而且，由于传感器节点只能在一定距离内传输数据信息，直接制约了网络的规模。此外，由于单跳数据传输模式下采用的多对一的汇聚方式，增加了信道的冲突，不利于信道利用率的提高。

为了解决单跳传输方式的不足，目前许多学者提出了不同的解决方法，最典型的方法就是将单纯单跳的网络改进为多跳的网络。在多跳数据传输模式下，传感器节点采集的数据经过中间节点或簇头进行转发，最终汇集到汇聚节点。这种方式对于大规模的网络优势尤为明显。但是，在多跳数据传输模式下，靠近汇聚节点区域内的节点要转发的平均数据负载量较大，使得这部分区域的节点能耗加剧，导致这些区域内的节点过早死亡，网络生命受限，同样不利于保证实际应用中网络的生命周期及服务质量。

### 1.2.3 网络数据融合模式

数据融合(Data fusion)主要是指数据的压缩、合成、过滤等加工方式，通过数据融合技术，可以减少中转节点传输的数据量和基站处理的数据量，从而减少网络总能耗，增强数据的可靠性<sup>[23][40]</sup>。

在结构化的多跳无线传感器网络中，传感器中转节点将自己所管辖的成员节点的数据采集和上一跳数据接收后，进行一定的数据融合处理，然后再传输给下一跳。其中，对中转节点所管辖区域的节点所采集的数据进行融合称为网内数据融合，将接收到上一跳的节点或其它中转节点转发的数据进行融合称为网间数据融合。而具体的数据融合方式，又分为基于成员节点和基于中转节点本身两种不同的数据融合模式。基

于成员节点的数据融合模式，中转节点将以成员节点的数量和成员节点采集的数据作为处理依据，而基于中转节点的数据融合模式，主要根据中转节点自身的特征作为处理依据。

### 1.3 论文研究内容和目标

从目前国内外研究成果来看，目前国内外无线传感器网络数据处理与能耗优化策略研究，尽管对无线传感器网络生命周期的延长有着不同的影响，能够在一定程度上延长网络的生命周期，但是大多数方法，对于解决大规模同构网络下的能耗问题存在一定局限，数据在多跳累积导致靠近汇聚节点区域的部分节点过早死亡，网络过早瘫痪、网络节点能量利用率低下的问题依然存在。而且，现存大多数协议或算法，当全网采用统一的融合比时，没有全面地考虑不同的数据融合比对无线传感器网络拓扑、网络能耗均衡及网络能量利用率的影响。本文的研究内容就是针对大规模的无线传感器网络的能耗问题，结合已有的被证明为有效的数据处理策略（如分簇机制及多跳机制以等），综合各种不同的因素(网络拓扑、簇头数目及簇头尺寸分布规律、数据融合等)，建立网络数据处理与能耗优化网络模型，对这些因素进行定量分析，对传感器采集和传输的数据进行局部融合，以减少数据传输过程中的能耗，使整个网络所有节点的能耗最小化和均匀，从而提高多跳无线传感器网络传感器节点的能量利用效率，增加网络的可扩展性，最大限度地延长网络的生命周期，从而为人们研究无线传感器网络提供理论上的依据和实际应用的指导。

### 1.4 论文组织结构

本文的主要研究内容是针对大规模的无线传感器网络，设计相关的数据处理与能耗优化策略。首先，根据无线传感器网络数据处理与能耗优化的研究现状，分析了国内外相关工作的优势与不足，基于已有的分簇及多跳数据处理机制，综合网络拓扑、全网的簇头分布、数据融合等数据处理与能耗优化策略进行建模分析；其次，对本文提出的模型结合已有的被证明为有效的数据处理与能耗优化策略，设计相关的仿真协议，并与国内外相关工作进行了仿真模拟和性能对比分析；最后，指出了本文提出的模型与协议在一般情形下的扩展与实际应用部署方案。

本论文组织结构如下：

第一章介绍了本文的主要研究背景，对国内外相关工作的研究现状进行了总结，阐

明了本文的研究目标和主要研究内容。

第二章首先阐述了关于无线传感器网络数据处理与能耗优化问题最新重要研究成果，对比分析各个协议和算法实现的主要机制和原理，归纳出已有的适于大规模无线传感器网络中的应用的数据处理与能耗优化策略，从中总结出其中优势以及存在的不足，由此引申出本文的研究方案。

第三章针对大规模无线传感器网络的能耗问题，结合对网络拓扑、簇头数目及尺寸分布、数据融合等影响网络数据处理与能耗的因素进行分析，寻找最优的网络跳数以提高簇头节点的能量利用率，同时，求解各跳子网内簇头数以保证全网所有总簇头耗能均衡。此外，进一步定量分析了为提高网络簇头节点能量利用率簇尺寸分布所应满足的规律以及数据融合比簇头数分布以及网络能耗均衡的影响。

第四章对本文设计的无线传感器数据处理与能耗优化策略进行了仿真，并结合已有的被证明为有效的动态建簇、动态路由以及簇头轮转的策略来设计了仿真协议 COCA(Constructing Unequal Clustering Architecture)协议，并对 COCA 协议与 UCR 协议在网络生命周期、网络节点剩余能量以及全网的数据融合比对网络能量利用率的影响等方面的仿真结果进行了对比和分析。

第五章首先论证了本文提出的数据处理及能耗优化策略建模分析过程中涉及基于矩形的各种子网拓扑原型方式的合理性及其在任意大规模实际应用部署场景下的扩展和组织部署的方案。

总结和展望部分总结了本文的主要工作，分析了本文提出的各种数据处理和能耗优化策略的优势和不足，并对未来的研究设想进行了展望。

## 第二章 基于数据处理的无线传感器网络能耗优化协议

本章主要依据数据聚集的方式对目前最新重要研究成果进行分类总结，对比分析各个协议和算法所采用的数据聚集方式，数据的汇聚模式以及数据的融合模式，分析其实现的主要机制和原理，归纳出已有的适于大规模无线传感器网络中的应用的数据处理与能耗优化策略，从中总结出其中优势以及存在的不足，并以此引申出本文的研究方案。

### 2.1 基于平面结构

在数据收集策略方面，对于传统的平面网络主要的典型协议有：SPIN<sup>[28]</sup> (Sensor Protocols for Information via Negotiation)，Directed Diffusion<sup>[29]</sup> 等协议。

Directed Diffusion 协议是以数据为中心的协议，节点本身并不重要。首先由汇聚节点(sink node)对感兴趣的事情提出请求，然后传感节点感知相应的兴趣事件，找到事件源，利用了数据的属性，当数据的属性由与感兴趣事件匹配的源节点生成时，网络便建立识别该数据源的梯度向量，选择和建立最优的路由，最后由中转节点路由将感知数据路由给汇聚节点，如图 3 所示。在数据传输过程中，节点路由能够局部修补和增强。协议能够对多个汇聚节点、多个数据源和多个请求建立多跳路由。

仿真显示，Directed Diffusion 协议平均耗能只有 omniscient multicast 协议的 60%；以数据为中心，不需要节点定位机制，每个节点具备数据融合和缓存功能。但是，协议是基于需求请求驱动的，不适合具有连续数据请求的应用，而且基于属性命名的策略机制局限于相应的具体应用。

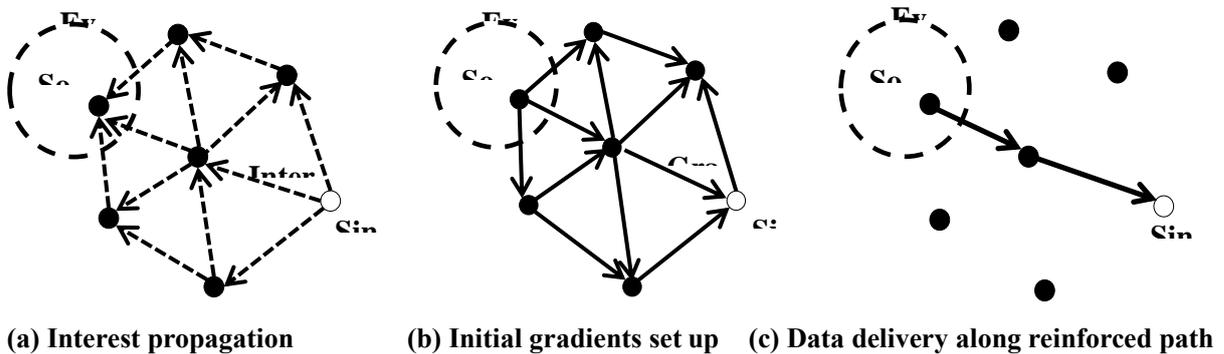


图 3 Directed Diffusion 的简单数据路由模式

### 2.2 基于分簇结构

本节主要对近年来基于分簇机制在无线传感器网络的数据处理与能耗的主要经典协议和算法进行综合比较分析。

### 2.2.1 LEACH 协议

LEACH<sup>[31][32]</sup> (Low Energy Adaptive Clustering Hierarchy) 是最早提出簇结构的协议, TEEN<sup>[33]</sup> (Threshold sensitive Energy Efficient sensor Network protocol), UCS<sup>[34]</sup> (Unequal Clustering Size) 等协议中的分簇基本思想都借鉴了 LEACH 的分簇机制。

LEACH 是一个基于局部簇头随机轮转来达到网络能耗均衡的协议, 通过局部的簇节点随机轮转协调和控制建簇使得动态网络具有可扩展性和稳健性。而且, 协议采取了数据融合技术, 使得传输到基站的数据量大大减少。

在 LEACH 协议研究的传感器网络中, 所有网络节点是同构而且能量有限, 网络只有一个位置固定的汇聚数据的基站, 而且基站远离网络节点的部署区域, 所有的成员节点以恒定的速率向簇头汇报数据以保证网络总是有数据向基站发送。

LEACH 协议的运行过程被划分成多轮, 每轮包括簇头建立阶段(簇结构的形成阶段)和稳定阶段(数据从簇头传输至基站的阶段), 而且簇头建立之后紧接着就是稳定阶段。为了减少网络中额外的开销, 稳定阶段要比簇结构建立时间长得多。

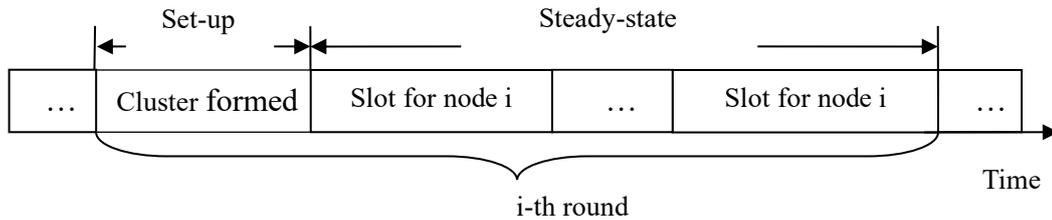


图 4 LEACH 协议执行流程

整个协议执行一轮的基本流程如下:

#### 1、簇头选取的广播阶段

在创建簇头的阶段, 每个节点基于预先一个给定的概率决定其是否成为簇头节点, 然后这些簇头将它们的状态使用 CSMA MAC 协议广播给网络中其它传感器节点。其中, 预先给定的概率根据以下公式来计算<sup>[31]</sup>:

$$T(n) = \begin{cases} \frac{P}{1 - P * (r \bmod \frac{1}{P})} & \text{if } n \in G \\ 0 & \text{otherwise} \end{cases}$$

这里 P 是指定的簇头在整个网络所占的百分比, r 是当前的轮转数, G 是在最后的 1/P 轮里没有被选为簇头的节点的集合。

当这个阶段完成后, 每个非簇头节点决定了在该轮中它所属的簇头。

#### 2、簇结构形成的成簇阶段

---

当每个节点决定它所属的簇头后，它又使用 CSMA (Code-Division Multiple Access, 码分多路复用) MAC 协议反馈告知相应的簇头它将成为该簇的成员节点，形成簇结构。

### 3、成员数据汇报时间戳分配

当簇结构形成之后，簇头基于 TDMA(Time-Division Multiple Access,时分多路复用) 机制给簇内成员节点的数目给簇内每个节点分配一个传输数据的时间戳告诉成员向簇头汇报数据的时间段。

### 4、数据传输阶段

簇结构形成和 TDMA 调度机制确定之后，非簇头成员节点便在自己所分配的时间戳内传输数据给所属的簇头，在其它时间保持休眠状态以减少能耗。LEACH 假设所有的节点在自己总是有数据发送，在所分配的时间戳内将传感数据传输给所属的簇头。簇头接受完所有成员节点的数据后采用一定的数据处理策略将所有成员的传感数据压缩为单个信号数据包。这是 LEACH 一轮的稳定阶段(Steady-state)。

以上流程便是 LEACH 协议执行一轮的基本流程，当一轮执行完成后，协议便开始下一轮执行，如重新选取簇头等工作。

LEACH 是一个完全分布式的、不需要基站全局控制信息的协议，而且节点执行协议不需要不需要知道全网的信息。它通过轮转簇头使得网络所有节点能量负载均衡减少了整个网络的能耗，比静态的成簇性能大大提高，延长了整个网络的生命周期。

Direct transmission 容易产生距离基站较远区域的节点先死亡，而 MTE (Minimum-transmission-energy) 容易产生距离基站较近区域的节点先死亡的现象。文献 [31] 分析了协议在网络总能耗，协议运行一定时间后传感器存活节点数目，网络不同时间段的簇结构，网络标准化总能耗与簇头数目占整个网络的节点数的比例之间的关系，将协议的网络总能耗，网络部署的半径，协议运行一定时间后传感器存活节点数目，第一个节点和最后一个节点的死亡时间等方面与 Direct transmission、MTE 进行了比较。根据协议性能对比可知，与 Direct transmission、MTE (Minimum-transmission-energy) 这些传统的路由协议相比，LEACH 用于数据通信能耗效率提高了 8 倍。网络中第一个节点死亡和最后一个节点的死亡时间比 Direct transmission、MTE 和 Static Clustering 协议分别延长了 8 倍和 3 倍。

LEACH 假设网络中簇头采取一定的压缩算法将簇内所有的数据都融合成一个数据包信号包。此外，LEACH 从簇头到基站只有一跳，不适于大规模的网络应用。

### 2.2.2 HEED 协议

HEED<sup>[35]</sup> (Hybrid Energy-Efficient Distributed clustering) 协议所研究的网络中, 所有的节点一旦部署后不具备移动性能, 而汇聚节点可以是静止或者移动的。网络链路是对称的, 而且如果传输距离知道的前提下每个节点能够通过接受信号的强弱计算到达另外一个节点的大致距离。节点不具备定位功能。

所有的节点都是同构的, 具有相同的处理性能, 具有相同的初始能量和能量调控能力(包括可调的传输能耗性能和最大的传输距离), 且节点部署后不能再补充能量。簇头的选取基于每个节点的剩余能量和与邻居节点个数, 所有成员节点到达簇头的平均最小能耗代价被定义为 AMRP(Average minimum reachability power)。在 HEED 协议每轮的执行过程中, 每个节点成为簇头的概率为

$$P_{CH} = C \frac{E_{residual}}{E_{max}}$$

$C$  是用户初始定义的簇头所占的百分比,  $E_{residual}$  节点当前的剩余能量,  $E_{max}$  是节点的是初始能量。每个节点不断向候选簇头发送簇头选取信息, 直到概率  $P_{CH}$  为 1。每个选取所有候选簇头中具有最小的 AMRP 的节点为簇头, 然后每个节点在下一轮将自己的概率设置成  $\min(2 * P_{CH}, 1)$ , 这个过程反复执行直到所有的节点找到自己所属的簇头。

HEED 没有考虑簇间通信代价, 与 gen-LEACH 协议性能进行了比较。gen-LEACH 协议中由于簇头的随机选取导致某些节点很快死亡, 而 HEED 协议中簇头选取由于考虑到了最小的通信代价使得簇头选取是分布的很好, 而且成簇代价小于 gen-LEACH 协议。

### 2.2.3 UCR 协议

UCR<sup>[13]</sup> (Unequal Cluster-based Routing) 协议所研究的网络中, 所有节点均匀分布, 对一个大面积的环境进行监测。所有的节点都是同构的, 具有相同的处理性能, 具有相同的初始能量和能量调控能力(包括可调的传输能耗性能和最大的传输距离), 每个节点拥有一个唯一的 ID。所有的节点能够进行休眠模式和激活模式操作, 网络只有一个远离监测区域的汇聚节点。所有的节点和汇聚节点一旦部署后都不具备移动性能。网络链路是对称的, 而且如果传输距离知道的前提下每个节点能够通过接受信号的强弱计算到达另外一个节点的大致距离。UCR 协议包括不等分簇节能算法——EEUC(Energy-Efficient Unequal Clustering)和基于地理位置贪婪算法和能量感知的簇间路由协议组成。

EEUC 是一个自组织的基于竞争的算法，也如 LEACH 一样采用簇头轮转策略，簇头的选取基于局部的信息(如相邻节点的剩余能量等)。节点的竞争半径由下面的计算式来计算：

$$R_i = (1 - c \frac{d_{max} - d(s_i, BS)}{d_{max} - d_{min}}) R_0$$

这里  $c$  是  $0 \sim 1$  之间的一个常系数， $R_0$  是预定义的最大竞争半径， $d_{max}$ ， $d_{min}$  分别表示网络中的节点与基站最大和最小的距离， $d(s_i, BS)$  表示节点  $s_i$  与基站的距离。节点的竞争半径随着离基站的距离减小而减小，从而离基站越近簇结构越小，使得离基站近的簇头用于簇内数据处理的能耗越少、用于簇间数据传输的能量越多。

而数据路由协议综合考虑了节点之间数据传输代价和节点的剩余能量。

由于 UCR 同时考虑了簇内和簇间的数据传输的影响，因此有效地缓解了多跳网络的“热点”问题、延长了网络的生命周期。

## 2.2.4 SCA 协议

SCA<sup>[13]</sup> (Structuralized Clustering Algorithm) 协议研究的传感器网络中，假设所有网络节点是同构而且能量有限，网络只有一个位置固定的汇聚数据的基站，而且基站远离网络节点的部署区域。协议首先将一定面积的大规模的网络划分成若干个小规模的子网，然后对每个子网区域内部节点轮转成簇，簇内所有成员节点以恒定的速率向簇头汇报数据，然后由不同子网内的簇头之间经过多跳由远及近向汇聚节点汇报。

网络经过初始化设置之后，协议的运行被划分成很多轮，每轮包括：簇头的选取、簇结构形成、成员节点的数据汇报和不同子网内簇头之间的数据传输几个阶段。

协议基于簇头数量、网络拓扑与网络能耗之间关系建立了最优化模型。通过以下数学建模，使得整个网络所有簇头的总耗能最少，定量分析出整个网络的总耗能和最佳子网数目之间的关系（参见图 5），网络规模和最佳子网数目之间的关系（参见图 6），不同子网区域之间的最佳簇头分布规律（参见图 7）。

$$E_{CH_i} = \frac{[(m-i)na + n]}{c_i} q_0 E_{elec} + \frac{(m-i+1)na}{c_i} q_0 (E_{elec} + \epsilon d^\alpha) + \frac{n}{c_i} q_0 E_{DA}$$

$$E_{CHtotal}(m) = \sum_{i=1}^m c_i E_{CH_i}$$

$$= \{ (\frac{m-1}{2}a + 1) E_{elec} + \frac{m+1}{2}a (E_{elec} + \epsilon_{mp} (\frac{S}{m})^\alpha) + E_{DA} \} N q_0$$

这里， $c_i$ 、 $E_{CH_i}$  分别为第  $i$  个子网内的簇头数目和每个簇头的能耗负载， $m$  是整个网络被

划分后的子网区域数,  $E_{CHtotal(m)}$  为所有子网区域的总耗能 (其它参数设置请参考文献 [13])。

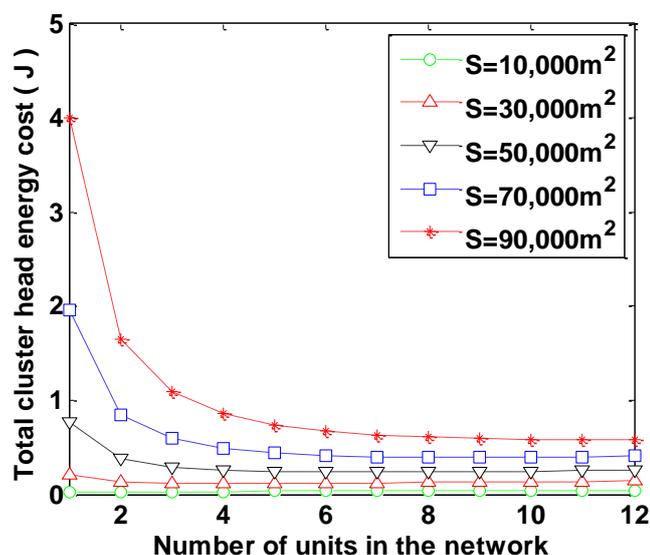


图 5 全网簇头的总耗能和最优子网数目之间的关系

协议通过定量分析网络拓扑、簇头数量和网络耗能之间的关系, 对第一个节点死亡和部分节点死亡两种不同的网络生命周期和协议运行一定轮数之后网络存活的节点总数几个不同方面与不同子网内等簇头数分布情形和 UCR 协议进行了对比, 使得网络能耗比这两种协议网络耗能更少、更均衡, 进一步延长了网络的生命周期。

但是, SCA 协议数据融合是全网采用统一的数据压缩模式, 未能从根本上解决多跳网络中“热点”问题。

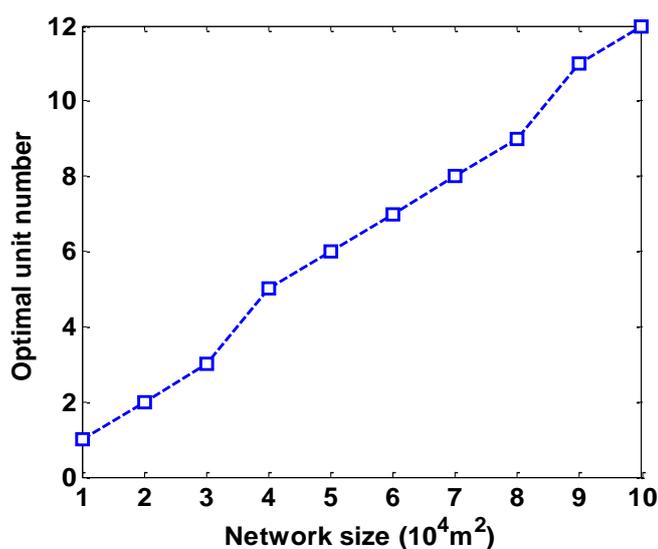


图 6 网络规模和最优的子网数目之间的关系

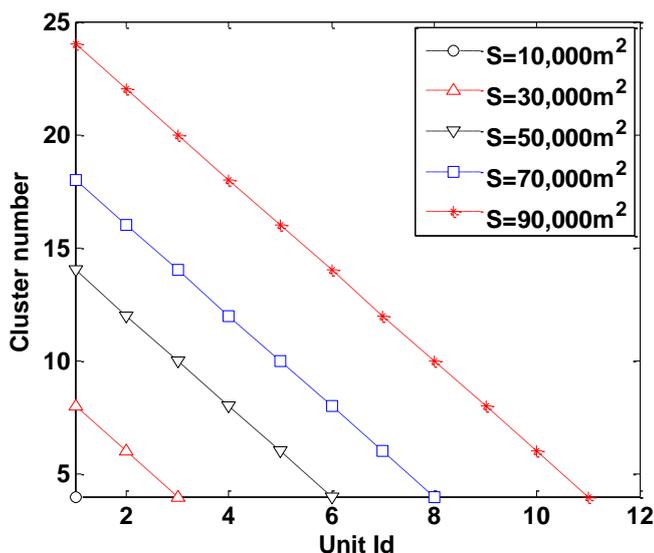


图 7 不同子网区域内最优的簇头数目分布规律

### 2.3 基于链结构

基于链状结构的典型协议有 PEGSIS<sup>[36]</sup> (Power-Efficient GATHERing in Sensor Information Systems)等。

PEGSIS 协议研究的网络中，基站是固定而且远离传感器网络的部署区域，所有的传感器节点是同构而且能量有限的，所有的节点知道全网的信息。PEGSIS 协议是一个基于链状路由的协议，在该协议中，所有的节点利用贪婪算法形成一条最小长度的链，每个传感器节点只与它相邻的一个节点进行通信，依次轮转向基站汇报数据。PEGSIS 协议假设每个节点将邻居节点和本节点的数据融合成一个同样大小的包后传输给它的另一个邻居节点或基站。

与 LEACH 相比，PEGSIS 协议每轮只有一个节点与基站进行通信，进一步减少了与基站通信的节点数；LEACH 协议中节点形成簇头的代价非常之大，PEGSIS 协议每个节点在数据传输节点只需要将数据传输给它的邻居节点，这样大多数节点传输能耗较 LEACH 小得多；PEGSIS 协议与基站交互的管理节点每次接受的消息较 LEACH 少得多。

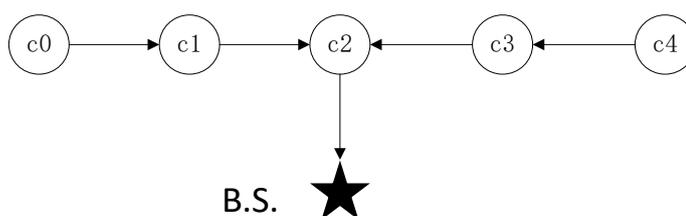


图 8 PEGSIS 协议网络部署场景

当基站远离网络部署区域时，分别以 1%,20%,50%,100%的节点死亡为网络生命周期定义时，PEGSIS 协议与 LEACH 协议相比，性能提高了 1~3 倍。

## 2.4 基于树结构

基于树状数据路由方式的典型协议有 PEDAP<sup>[37]</sup> (Power Efficient Data gathering and Aggregation Protocol), TTDD<sup>[38]</sup> (Two-Tier Data Dissemination)等。PEDAP 协议的主要目标是使得每轮全网的能耗最小和均匀化来延长网络的生命周期。PEDAP 协议研究的网络中，每个节点周期性地感应周围的环境，然后将感应的数据发送到一个固定的基站，所有的节点都是同构而且能量有限和固定的，协议假设将  $n$  个大小为  $k$  的数据包融合成一个大小为  $k$  的数据包。

PEDAP 协议中，节点采用和 LEACH 协议节点相同的能量模型，所有节点的数据的位置固定，而且基站预先知道每个节点的拓扑信息，每个节点能够直接与基站和其它节点进行通信。数据路由根据 Prim 最小生成树产生，时间复杂度为  $O(n^2)$ ，由基站负责计算全网的路由信息，这样，每个节点的用于数据接收到能耗和定期建立网络路由的能耗就远远小于 LEACH 和 PEGSIS 协议中节点的能耗。

PEDAP 协议包含两个算法：PEDAP 和 PEDAP-PA。PEDAP 算法的目标是减少整个网络的能耗来使得网络生命周期得以延长。在计算 Prim 最小生成树路由时，边长的权值为两节点之间的能耗代价。由于当网络中第一个节点死亡时，网络所消耗的能量最小，从而其余节点的剩余能量保留了下来。由于 PEDAP 算法没有考虑全网的能耗均衡问题，PEDAP-PA(Power Efficient Data gathering and Aggregation Protocol- Power Aware).是在 PEDAP 基础上进行了改进。在计算 Prim 最小生成树路由时，两节点之间的能耗权值改为：

$$C_{ij} = \frac{2kE_{elec} + kE_{amp}d_{ij}^2}{e_i}$$

$$C'_i = \frac{kE_{elec} + kE_{amp}d_{ib}^2}{e_i}$$

$e_i$  是节点  $i$  标准化后的剩余能量( $0 \leq e_i \leq 1$ )。

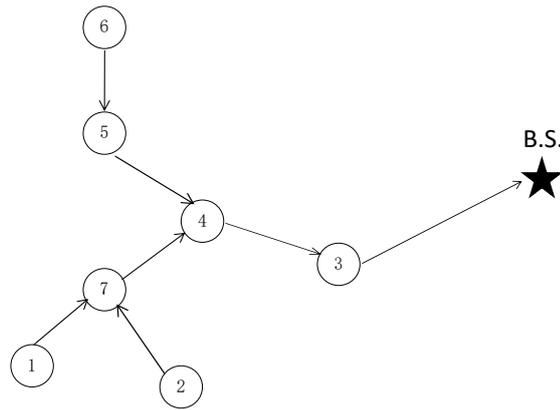


图 9 PEDAP 协议网络部署场景

这个新的能量计算公式仅仅推迟能量较低的节点加入到最小生成树中而不改变节点发送数据的邻居节点。节点越迟加入到生成树中，所接受的消息数就越少。这样，当 MST 算法执行一定的轮数之后，网络能耗将更加的均衡。PEDAP-PA 算法尽管稍微减少了最后一个节点的死亡时间，但延长了第一个节点的死亡时间。

协议两个算法对基站处于网络部署区域中心和离部署区域较远的地方都具有很好的性能，与 LEACH 相比，网络生命周期提高了 4~20 倍，与 PEGSIS 相比提高了 3 倍。当基站离网络部署区域较远时，PEDAP-PA 比 PEGSIS 的第一个节点死亡时间延长了 4 倍，PEDAP 比 PEGSIS 的最后一个节点死亡时间延长了 1.25 倍；当基站处于网络部署区域的中心时，PEDAP-PA 比 PEDAP 第一个节点死亡时间延长了 2 倍，PEDAP 和 PEDAP-PA 算法比 LEACH, PEGSIS 最后一个节点的死亡时间延长了 2 倍。因此，当网络采用第一个节点死亡时间定义时，PEDAP-PA 算法能够取得最好的性能；当网络采用最后一个节点的死亡来定义网络生命周期时，PEDAP 和 PEDAP-PA 算法都能取得较好的性能。当基站处于网络部署区域内部时，协议仍然取得较好的性能而 LEACH 和 PEGSIS 协议则不能。PEDAP 协议还指出，当节点初始能量加倍时，第一个节点的死亡时间增加约 2.5 倍。协议的算法采用的是集中式的算法，当基站不在网络节点的通信范围内时，可以采用分布式的最小生成树算法。

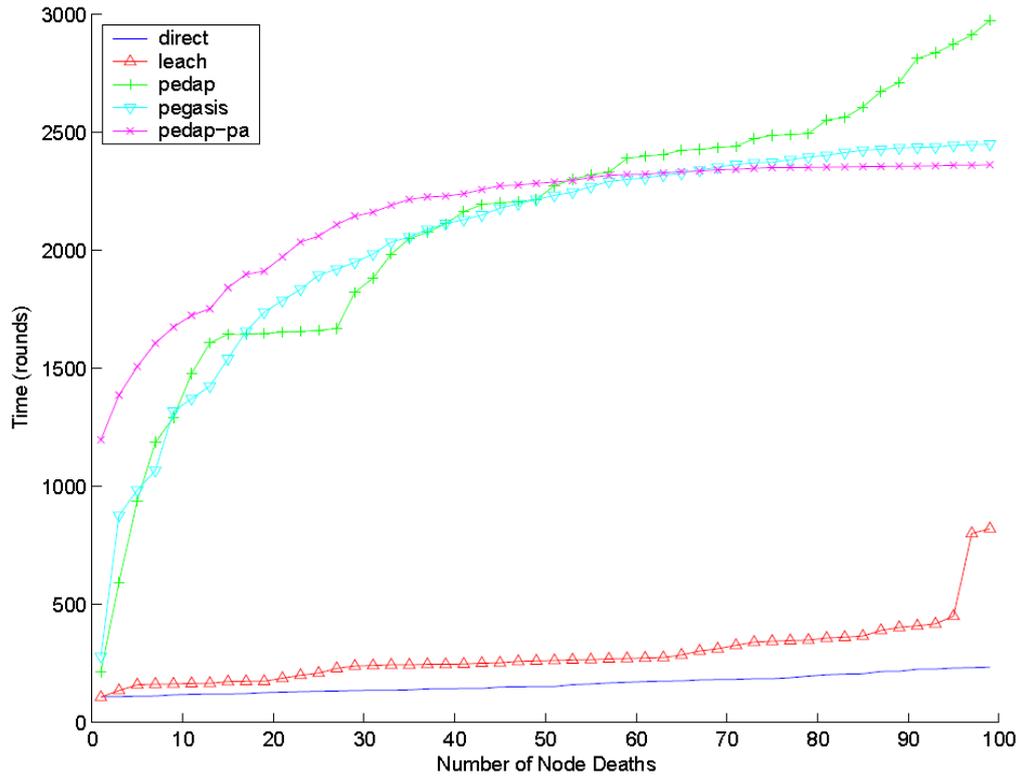


图 10 网络节点的死亡时间(100\*100m<sup>2</sup>的网络, 基站远离网络部署区域)

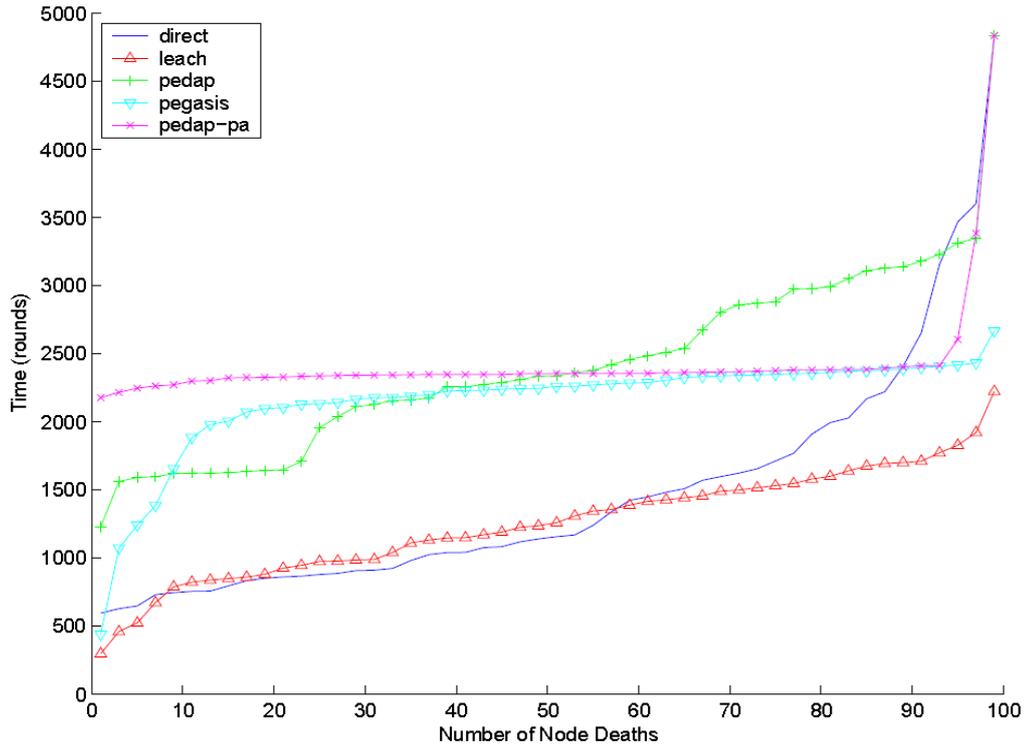


图 11 网络节点的死亡时间(100\*100m<sup>2</sup>的网络, 基站处于网络部署区域内部)

## 2.5 本章小结

本文通过如下几个方面, 对上面涉及到的所有协议中的数据处理与能耗策略研究

---

的协议和算法进行分类和总结，表 1 中各项字符的含义说明如下：

P(Protocols): 表示协议或算法的名词；

Ins(Intra-network data sending pattern ): 表示结构化网络中转结点之间数据发送模式，1 表示成员结点每次发送一个包，x 表示成员结点每次发送多个包；

Inf(Intra- network data fusion pattern ):表示中转节点内部数据融合模式，1 表示中转节点将其它节点或成员结点采集的数据融合成一个包，x: 表示中转节点将其它节点或成员结点采集的数据融合成多个包；

Intf(Inter- network data fusion pattern ): 表示结构化网络中转结点之间数据融合模式，1(fused to 1 packet and accumulated transfer): 表示将中转节点将中转节点之间数据融合成一个包而且不传输累加数据仅传输融合后的一个包，1+表示中转节点将中转节点之间数据融合成一个包并传输多跳累加的数据；x 表示将中转节点将中转节点之间数据融合成多个包而且不传输累加数据仅传输融合后的多个包，x+表示将中转节点将中转节点之间数据融合成多个包然后传输累加数据仅传输融合后的多个包；

H(Heterogeneous): 表示网络节点是否为异构的,Y表示是异构的,N表示是同构的。

Hn(Hop-number): 表示结构化网络中转结点之间数据转发方式，1表示单跳网络，直接由中转节点转发给汇聚结点，m表示多跳网络，数据由多个中转节点一跳一跳地传输给汇聚节点；

Tp(topology): 表示网络的拓扑结构，其中p(plane)表示平面网络结构，l(chain) : 表示链状网络结构，c(cluster)表示簇状网络结构，t(tree): 表示树状网络结构

T(Lifetime): 表示网络生命周期，1表示以第一个结点死亡持续时间定义网络生命周期，x表示特定数量结点的存活或死亡持续时间定义网络生命周期，n%表示以某种比例结点的存活或死亡持续时间定义网络生命周期，n表示以最后一个结点死亡或全部结点死亡持续时间定义网络生命周期，o(others): 其他用于定义网络生命的方式或衡量协议或算法性能的方式，比如：1 汇聚结点在一段时间内接受的数据总量，2 发送一定量的数据后的存活结点数等。

### 2.5.1 平面网络与结构化网络

无线传感器网络的耗能优化与数据处理策略中，在数据采集模式和数据路由方面，传统的平面网络，如Directed Diffusion等协议，由于距离汇聚节点远的区域的“热点”问题，使得网络过早地瘫痪，而且网络通常不支持节点的异构性。结构化的网络，如LEACH, SCA, UCR, HEED, UCS, PEGSIS, PEDAP等，通常支持和能够有

效利用节点的异构性。对于链状协议，如PEGSIS等和树状的网络协议，如PEDAP等，由于采用集中式的算法来计算网络的拓扑和数据的路由信息，不适合网络规模的可扩展性，而分簇的网络与传统的平面网络和链状、树状的结构化网络相比，分簇后的网络，有这几个方面的优势：

表 1 经典算法协议总结

p	ins	Inf	Intf	H	Hn	T	Tp
Directed Diffusion	—	—	—	N	1	o	p
LEACH	1	1	—	N	1	o	c
UCS	1	1	1+,x+	N	2	1, n%	c
HEED	1	1	1	N	m	1, n, o	c
UCR	1	1	1	N	m	1, n	c
SCA	1	n	x+	N	m	1, n%	c
PEGSIS	1	1	1	N	1, m	1, n%	l
PEDAP	1	1	1,x	N	m	1, n	t

1) 由于局部数据处理的代价远远低于网络数据传输的代价，簇内局部的数据处理可以大大减少所需传输的数据量，从而大大减少整个网络的数据负载，大大减少网络总能量消耗；

2) 而且簇内数据融合可以大大减少基站处理的数据量，增强数据信号和减少不相干的噪音从而增强数据的可靠性；

3) 而且分簇后的网络可以减少与基站进行直接通信的节点数目，将整个网络结构化，成员节点之间向相应的簇头进行汇报，簇头只负责收集簇内成员节点的采集数据，这样就降低数据路由的复杂性，便于数据处理；

此外，局部的簇节点随机轮转协调和控制建簇使得动态网络具有可扩展性和稳健性。

### 2.5.2 单跳与多跳数据传输

在网络数据出和汇聚模式方面，由于传感器节点具有通信距离的局限性，单跳网络协议如LEACH等，通常只适合小规模的网络，对于多跳的网络协议，如SCA，UCR，HEED，UCS等，虽然适合大规模网络，但离汇聚节点较近的区域依然存在能耗不均的“热点”问题。

### 2.5.3 网络数据融合

在网络的数据融合方面，在平面网络结构中，目前大多数协议，如Directed

---

Diffusion等协议，只是简单地论述消除冗余数据，而结构化网络中的协议，如 LEACH, SCA, UCR, HEED, UCS, PEGSIS, PEDAP等，通常采取将中转节点采集到的数据融合为一个数据包或者对整个网络采用统一的融合比进行融合的方式。

#### 2.5.4 结论

在数据聚集方面，数据采集模式和数据路由的研究中，进行比较和总结并发现，在结构化的无线传感器网络数据方法中，尤其是大规模的结构化多跳网络中，发现现存大多数的算法或协议虽然能够在一定程度上延长网络的生命周期，但对于解决大规模同构网络下的能耗问题存在一定局限，数据在多跳累积导致靠近汇聚节点区域的部分节点过早死亡，网络过早瘫痪、网络节点能量利用率低下的问题依然存在。此外，现存大多数协议或算法，当全网采用统一的融合比时，没有全面地考虑不同的数据融合比对无线传感器网络拓扑、网络能耗均衡及网络能量利用率的影响。同时发现，相比其它无线传感器网络能耗优化协议策略，分簇机制和多跳机制是一种更为有效的网络组织、扩展网络规模、数据融合以节省网络能耗、延长网络生命周期的技术，作为本文的基本数据处理与能耗优化技术而引入。

### 第三章 无线传感器网络数据处理与能耗优化策略建模分析

本章首先定义了大规模无线传感器网络的能耗问题，结合对网络拓扑、簇头数目、数据融合等影响网络数据处理与能耗的因素进行建模优化，主要做了以下几个方面的工作：首先，在给定网络面积的前提下，寻找最优的网络跳数以保证在网络生命周期的一轮里面全网内所有簇头总耗能最小、提高簇头节点的能量利用率；其次，在量化多跳转发过程中不同跳数的同时，求解各个子网内分担转发数据的簇头数以保证全网所有总簇头耗能最小的同时任意两个相邻子网间的簇头在一轮里面耗能均衡、从而延长整个网络的生命周期。最后，对同时满足使所有簇头总能耗最小和任意相邻子网内簇间平均能耗负载均衡时，离汇聚节点远近不同跳数或子网内簇尺寸分布所应满足的规律以及不同的数据融合比对不同子网内簇头数分布以及网络能耗均衡的影响进行进一步定量分析。

#### 3.1 系统模型

##### 3.1.1 参数说明

表 2 模型参数说明

参数名称	具体说明	条件限制
$d$	传感器节点数据传输距离	$d > 0$
$d_0$	传感器节点传输距离阈值	$d_0 > 0$
$d_i$	第 $i$ 、 $i-1$ 跳两相邻子网内簇头之间的平均传输距离	$d_i > 0$
$\alpha$	传感器节点传输距离指数	$\alpha = 2$ 或 $\alpha = 4$
$\epsilon d^\alpha = \begin{cases} \epsilon_{fs} d^\alpha & d < d_0 \\ \epsilon_{mp} d^\alpha & d \geq d_0 \end{cases}$	传感器节点将 1 bit 的数据传输距离 $d$ 时 传感器节点放大器的能耗级别	$\epsilon d^\alpha > 0$
$q_0$	任意传感器节点传输每次一个数据包或控制消息包的长度	$q_0 \in \mathbb{Z}^+$
$q$	传感器节点之间传输长度为 $q$ 比特的数据或控制消息包	$q \in \mathbb{Z}^+$

表 3 模型参数说明 (续)

$E_{elec}$	传感器节点发送端或接受端电路消耗能量	$E_{elec} > 0$
$E_{Tx-elec}(q)$	传感器节点传输长度为 $q$ 比特的数据或控制消息包的电子能耗	$E_{Tx-elec}(q) > 0$
$E_{Tx-amp}(q, d)$	传感器节点传输长度为 $q$ 比特的数据或控制消息包的放大器能耗	$E_{Tx-amp}(q, d) > 0$
$E_{Tx}(q, d)$	传感器节点将长度为 $q$ -bit 的数据包发送距离长度 $d$ 时耗能	
$E_{Rx-elec}(q)$	传感器节点接受长度为 $q$ 比特的数据或控制消息包的电子能耗	
$E_{Rx}(q)$	传感器节点接收长度为 $q$ -bit 的数据包时耗能	
$E_{DA}$	传感器节点融合每个比特数据所耗能量	$E_{DA} > 0$
$E_{CH}(q, d)$	传感器簇头节点经过距离 $d$ 传输长度为 $q$ -bit 的数据包时的总耗能	
$S$	网络部署面积	$S > 0$
$N$	全网传感器节点总数	$N \in Z^+$
$\rho$	网络节点部署密度	$\rho > 0$
$m$	簇间跳数或网络子网划分总数	$m \in Z^+$
$m_u$	全网子网跳数或子网区域划分数的上界	$m_u \in Z^+$
$m_{opt}$	全网实际最优跳数或子网划分数	$m_{opt} \in Z^+$
$l$	正方形子网的边长	$l > 0$
$l_i$	第 $i$ 个子网的边长	$l_i > 0$
$a$	全网统一数据融合比	$0 < a \leq 1$
$E_{Rx}(a, i)$	第 $i$ 跳子网内每个簇头接收数据平均能耗	
$E_{agg}(i)$	第 $i$ 跳子网内每个簇头数据融合平均能耗	
$E_{Tx}(a, i)$	第 $i$ 跳子网内每个簇头数据转发平均能耗	
$E_{CHi}(a)$	全网数据融合比为 $a$ 、第 $i$ 跳子网内每个簇头所耗平均能耗	通常 $E_{CHi} \leq E_{CHi-1}$

表 4 模型参数说明 (续)

$c_i$	第 $i$ 跳子网内的簇头数	$c_i \leq n_i$
$n_i$	第 $i$ 跳子网内的节点数	$c_i, n_i \in Z^+$
$E_{CHtotal}(a, m)$	全网数据融合比为 $a$ 、子网跳数为 $m$ 的全网总能耗	
$\Delta E_{CH}(a, c_{i-1}, c_i)$	第 $i$ 、 $i-1$ 跳两相邻子网内簇头数分别为 $c_i$ 、 $c_{i-1}$ 且全网数据融合比为 $a$ 时簇头平均能耗负载差值	$\Delta E_{CH}(a, c_{i-1}, c_i) \geq 0$

### 3.1.2 无线传感器网络模型和传感器节点模型

我们研究的无线传感器网络部署区域模型采用学术界广泛使用的矩形区域抽象部署模型，而且我们考察的矩形部署区域假设具有如下属性：

- 1) 部署的矩形区域具有较大面积，通常在几万至十几万平方米以上；
- 2) 在网络部署区域中，所有的传感器节点随机、均匀部署，节点的部署密度能够保证整个目标区域节点之间能够良好通信，而且一旦部署后不再受人工干预，静止不动，因此不能再充电；
- 3) 传感器节点之间的数据传输链路是对称的，即任意两个节点之间能够采用相同的能耗级别通信；
- 4) 所有传感器节点都是同构的：具有相同的初始能量，最大的传输距离，具有固定的传输能耗级别，并能根据不同通信距离调控相应能耗级别，相同的数据处理能力（包括相同的 CPU 和存储器）。节点不具地址感知功能，不配备 GPS 设备；
- 5) 网络只有一个汇聚节点，位于传感器网络部署区域的一端，负责汇聚整个网络的数据。

整个网络部署场景如图 12 所示：

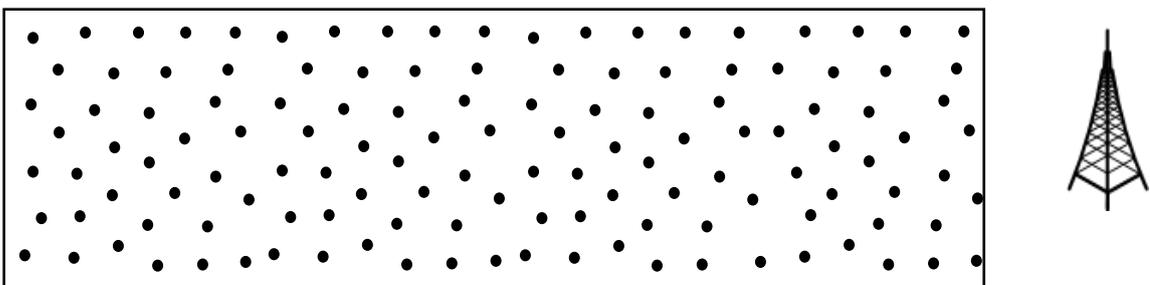


图 12 大规模无线传感器网络矩形部署区域

### 3.1.3 传感器节点传感数据处理能耗模型

#### 1、传感器节点基本能耗模型

对于无线传感器网络节点的能耗模型，我们采用国内外文献（如 LEACH<sup>[31]</sup>，UCS<sup>[33]</sup>，HEED<sup>[35]</sup>，UCR<sup>[13]</sup>，SCA<sup>[14]</sup>，PEGSIS<sup>[36]</sup>，PEDAP<sup>[37]</sup>等）广泛引用的基本能耗模型，如下图 13 所示。

1) 传感器节点发送端能耗：传感器节点将长度为  $q$ -bit 的数据包发送距离长度  $d$  时耗能如下：

$$\begin{aligned} E_{Tx}(q, d) &= E_{Tx-elec}(q) + E_{Tx-amp}(q, d) \\ &= q(E_{elec} + \epsilon d^\alpha) \\ &= \begin{cases} q(E_{elec} + \epsilon_{fs}d^2) & \alpha = 2 \quad d < d_0 \\ q(E_{elec} + \epsilon_{mp}d^4) & \alpha = 4 \quad d \geq d_0 \end{cases} \end{aligned} \quad (1)$$

2) 传感器节点接收端能耗：传感器节点接收长度为  $q$ -bit 的数据包时耗能

$$E_{Rx}(q) = E_{Rx-elec}(q) = qE_{elec} \quad (2)$$

$E_{elec}$ 为传感器节点发送端或接受端电路消耗能量，是与传感器节点硬件本身相关的常量； $\alpha$  为节点电路信号衰减系数，当数据发送或传输距离小于特定阀指  $d_0$  ( $d_0$ 为传感器硬件固定系统常数)时取 2，当数据发送或传输距离大于等于阀指  $d_0$  时取 4； $\epsilon_{fs}d^2$  或  $\epsilon_{mp}d^4$  为放大器将信号放大到一定级别时节点所耗能量。在本文中，我们假定所有簇内节点之间的通信采用  $d^2$  能耗模型，簇间通信采用  $d^4$  的通信模型。

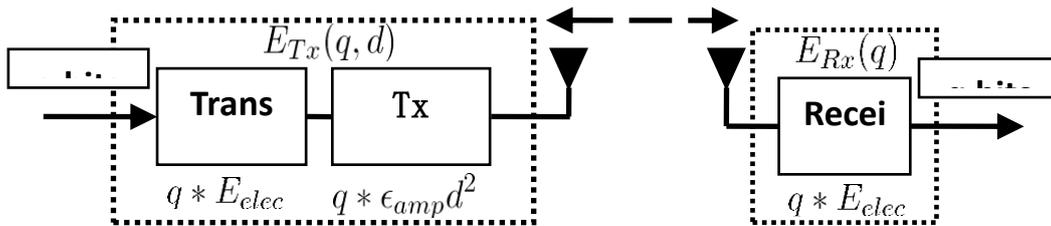


图 13 传感器节点能耗模型

#### 2、多跳分簇网络簇头节点能耗模型

在多跳分簇无线传感器网络协议中，如 UCS<sup>[33]</sup>，HEED<sup>[35]</sup>，UCR<sup>[13]</sup>，SCA<sup>[13]</sup>等中的能耗模型，对传感器网络节点传感数据处理的基本能耗模型进行了扩展，其中，非簇头的普通节点采用上述基本能耗模型，而对簇头传感数据处理能耗模型，包括以下几个方面：

1) 传感器簇头接收端能耗：在本文中，传感器节点接收上一跳的长度为  $q_{last-hop}$ -bit 的数据包和本跳内部采集的长度为  $q_{intra}$ -bit 的数据包时耗能如下：

$$\begin{aligned} E_{CH-Rx}(q_{next-hop}) &= E_{CH-Rx}(q_{last-hop}) + E_{CH-Rx}(q_{intra}) \\ &= (q_{intra} + q_{last-hop})E_{elec} \end{aligned} \quad (3)$$

这里,  $q_{next-hop} = q_{intra} + q_{last-hop}$ 。

2) 数据融合能耗: 簇头将本簇内成员节点采集的长度为  $q_{intra}$ -bit 的数据包融合成长度为  $q'_{intra}$ -bit 的数据包耗能为:

$$E_{agg} = q_{intra}E_{DA} \quad (4)$$

这里  $E_{DA}$  为网内压缩单位比特数据时所耗的能量。

3) 传感器簇头转发端能耗: 在多跳分簇网络中, 传感器簇头节点的转发能耗包括用于转发本簇内融合后的数据能耗和簇间累加转发的数据能耗。中继簇头节点将本簇内部融合后的数据与上一跳簇头转发而来的数据一起经过距离  $d$  发送到下一跳时耗能为:

$$\begin{aligned} E_{CH-Tx}(q'_{next-hop}, d) &= E_{CH-Tx}(q'_{intra}, d) + E_{CH-Rx}(q_{last-hop}) \\ &= (q'_{intra} + q_{last-hop})(E_{elec} + \epsilon d^\alpha) \end{aligned} \quad (5)$$

这里,  $q'_{next-hop} = q'_{intra} + q_{last-hop}$ 。

因此, 每一个中转簇头节点传输长度为  $q'_{next-hop}$ -bit 的数据包时的总耗能为:

$$\begin{aligned} &E_{CH}(q'_{next-hop}, d) \\ &= E_{CH-Rx}(q_{next-hop}) + E_{CH-Tx}(q'_{next-hop}, d) + q_{intra}E_{DA} \\ &= (q_{intra} + q_{last-hop})E_{elec} + (q'_{intra} + q_{last-hop})(E_{elec} + \epsilon d^\alpha) + q_{intra}E_{DA} \end{aligned} \quad (6)$$

### 3.1.4 网络生命周期定义

无线传感器网络通常以节点的能耗情况作为衡量和定义网络生命周期的主要依据。对于不同的应用场景, 无线传感器网络具有不同的生命周期定义方式。对于这些不同的应用场景, 通常有如下几种网络生命周期的定义方式<sup>[13]</sup>:

#### 1、基于单个节点死亡的定义

这种无线传感器网络生命周期定义方式又分为网络中第一个传感器节点死亡和最后一个节点死亡时无线传感器网络所能持续工作的时间长度两种方式。在本文中, 我们采取的单个传感器节点死亡定义下的网络生命周期是指网络正常工作, 直到其中的一个簇头不能继续将它自身的数据转发到下一跳(下一个转发簇头或汇聚节点)或成员节点不能将其自身的传感数据汇报给相应的簇头节点而导致该节点与网络其它节点中断通信所持续的时间。

目前, 有许多学者在无线传感器网络的研究过程中采用了这种定义方式。其中, 对于第一个传感器节点死亡的网络生命周期定义方式在很多实际应用场景下不太合理,

---

尤其对于大规模的无线传感器网络，这种定义方式会导致网络资源的大量浪费。在大规模的无线传感器网络的部署场景中，当其中仅有一个节点死亡时，网络其他大量的节点依然存活，能通过其他存活节点继续工作来实现网络主要功能，以此继续保证用户所需基本功能或服务。由于本文研究的网络模型是大规模的无线传感器网络，所以，在本文中的网络生命周期除了采用这种网络生命周期的定义方式外，还考虑了以下另外一种更加符合实际应用的定义方式。

## 2、基于子区域部分节点死亡的定义

在分簇结构化的多跳网络中，网络能耗不均，离汇聚节点较近的“热点区域”内才传感器节点能耗负载较重，较其它区域的节点较早死亡，当网络正常工作，直到该区域内存活的节点的剩余能量小于转发本跳内的数据所需的能量时，一定数量的节点死亡所持续的时间。

在本文中，这种网络生命周期定义方式的具体含义是指网络正常工作，直到离汇聚节点最近的那个子区域内的存活节点数少于由建模得到的用于转发本跳数据所需的理论簇头数而导致该子区域与网络其它区域断开所持续的时间。

## 3.2 大规模无线传感器数据处理与网络能耗问题

### 3.2.1 大规模无线传感器网络的能耗问题

由前文的相关工作分析可知，相对于其它几种无线传感器网络数据处理方式，基于分簇的数据处理方式被证明是一种有效的网络组织方式。为了处理大规模无线传感器网络能耗问题，支撑无线传感器网络大规模的应用部署和数据融合技术，需要建立有效的网络拓扑结构，本文将引入基于分簇的数据处理方式作为基本的网络组织方式。

首先，为了节省网络能量、延长网络生命周期，无线传感器网络所有节点自组织成簇结构，成员节点负责传感数据，然后汇报给相应簇头进行数据融合以除去冗余数据、减少能耗，如图 14 所示。其中，每个 Vronoi 区域<sup>[34]</sup>表示一个簇结构。一旦簇结构形成、簇头选定之后，所有的成员节点传感数据，然后便按照恒定的速率将传感数据直接发送给相应的簇头。当簇头接受完簇内所有成员节点汇报的数据后，再将这些成员节点采集到的数据按照一定的数据融合算法进行处理。这里，尽管簇头能够进行簇内和簇间数据融合，簇内数据压缩，但是在无线传感器网络的实际应用中，尤其是大规模网络，不同的簇头通信时簇间的数据空间、时间冗余度非常小，因此，在本文

中，我们采用的数据融合策略中仅仅考虑簇内数据融合，簇间数据不融合，直接进行转发；而且，在本文中，我们只关心数据融合处理后的数据量而不考虑具体的数据融合算法，如数据压缩、网络编码等，假设簇内网络数据相似、冗余度较高，全网采用统一均匀融合方式，融合后的数据量与簇内成员节点数成正比。

其次，为了实现网络的可扩展性，网络内的所有簇头按照簇间路由算法经过多个簇头由远及近地经过多跳，最后传输到汇聚节点。

网络在建簇过程中，必须满足以下几个方面的条件<sup>[35]</sup>：1) 成簇过程完全是分布式的，每个节点仅仅基于其局部信息做出判断；2) 成簇过程在有限时间内终止，与网络规模无关；3) 在建簇过程结束后，每个节点要么是簇头，要么属于而且仅属于一个簇结构的成员节点。根据网络在建簇过程，网络的生命周期，如同[13]等文献中定义的一样，通常按照轮数来度量，每一轮又包含这样四个子过程：1) 全网内不同跳数子网区域进行簇头选取；2) 网络不同跳数子网区域内部簇结构建立；3) 网络不同跳数子网区域内部各个簇内成员节点向选取的簇头进行数据汇报的过程；4) 网络不同跳数子网区域簇头之间进行数据转发过程。

由前面分析可知，在网络生命周期每一轮中，网络中所有节点的耗能分别如下。在每个簇结构中，对于簇头节点，其能耗主要包括以下几个方面：

- 1) 簇头自身传感数据感知能耗；
- 2) 生成和传输控制信息能耗：这部分能耗又包括簇头选取过程中簇头竞选能耗和簇结构建立过程给成员节点反馈信息的能耗；
- 3) 传感数据处理能耗：这部分能耗主要包括三个方面：簇内数据和簇间数据接收能耗，数据融合能耗以及簇头之间或簇头与汇聚节点间的数据转发能耗（距离汇聚节点最远端区域内的簇头不需要接受和转发前面一跳簇头转发的数据）。传感数据处理能耗是簇头的主要能耗<sup>[9]</sup>。

对于每个成员节点，能耗主要包括传感数据感知能耗和传感数据发送能耗。本文主要考察传感器网络传感器节点的数据处理能耗，主要研究内容分为以下几点：

第一，相比于成员节点，簇头是网络中能耗最多的节点，本文的主要工作之一就是提高簇头节点的能量利用率。在这种多跳数据传输模式中，离汇聚节点越近的区域内的簇头节点要转发的平均数据负载量就越大，使得这部分区域的簇头节点能耗加剧，比离汇聚节点远的区域的节点过早地耗能量、导致这些区域内的节点过早死亡，形成所谓的“热点”区域出现常见的“热点区域(Hot-Spot)”问题，使得整个网络瘫痪，无法满

---

足用户的需求。

这个问题首先在文献 [33] 中提出来，但是他们仅仅对两跳的网络进行了研究、而且没有考虑网络规模的可扩展性。对于这个问题所采取的最直观的策略就是在离汇聚节点越近的区域部署越多的簇头节点来分担上面多跳转发而来的数据、从而缓解能耗不均。但是，在这样大规模的无线传感器网络的不同区域究竟要部署多少节点才能保证不同区域的簇头能耗负载保持均衡？

第二，要解决上述问题，需要首先解决另外一个问题：在图 14 所示的大规模网络中，离汇聚节点远端区域的簇头所汇聚的数据在整个多跳网络传输过程中需要多少跳才能传输到汇聚节点同时保证进行多跳转发的所有簇头总能耗最小？当跳数过多时，传输等量的数据需要更多的簇头来接受和转发才能传输到汇聚节点，从而消耗簇头更多的接受能量，从而浪费更多的能量；当跳数过少时，由于负责转发的簇头数量偏少，能耗负载过大，从而导致网络能耗更加不均衡。

第三，全网采用统一均匀融合方式，由于融合后的数据量与簇内成员节点数成正比，因此对于不同的数据融合比，网络的数据融合比愈大，簇头需要处理的数据愈多，全网所有节点能耗负载，尤其是簇头能耗负载愈大，怎样通过不同的数据融合比来定量分析和确定相应的簇头分布规律？

第四，由于在同构的无线传感器网络中，由于簇头能耗负载远远大于成员节点能耗，从而，在网络生命周期的每一轮里面，簇头工作的时间愈长，簇头能耗愈多大，网络能耗愈不均，怎样才能进一步促进包括簇头和成员节点在内的所有节点能耗均衡一致、延长网络的生命周期？

总之，本文的主要工作就是针对如图 14 所示的大规模网络，通过对网络拓扑、簇头数目及尺寸分布、数据融合等影响网络能耗的因素进行建模分析，针对以上几个问题，提出有效的数据处理与能耗优化策略来提高全网所有节点的能量利用效率、延长网络的生命周期，促进无线传感器网络大规模的部署和应用。

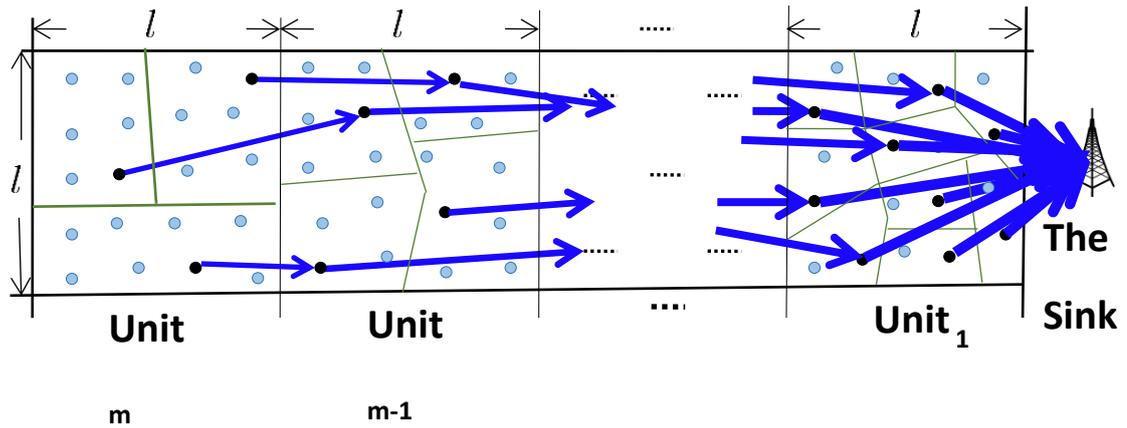


图 14 多跳分簇结构化无线传感器网络部署场景

### 3.2.2 大规模无线传感器网络数据处理与能耗优化策略

首先，我们对如图 14 所示的大规模网络，求解簇间最优的跳数以保证全网内所有簇头节点能耗最小。为了便于问题的分析，我们假设全网由一系列虚构的小规模的正方形子区域或子网构成（对于该子网模型假设的合理性及基于矩形网络模型的其它种类的子网拓扑方式放在第六章的讨论部分专门进行讨论）。为了保持各个子网内簇头能耗均衡，在同一子网内的各个簇头尺寸保持一致，只负责接收成员节点采集的数据和转发上一跳的数据而不接收同一子网内其它簇头转发的数据，而且每个簇头离汇聚节点具有相同的跳数，即经过相同的跳数将数据转发传输至汇聚节点，相邻子网之间进行单跳转发。这样，问题等价于对于如图 14 所示的一个大规模网络，在给定网络面积  $S$  的前提下，寻找最优的子网的数目  $m$ （也就是寻找最优的网络跳数），同时得到了每个子网相应的边长  $l$  及划分整个网络的子网区域大小，以达到在一轮里面全网内所有簇头总耗能最小、从而延长整个网络的生命周期。

其次，在求得最优的跳数保证全网内所有簇头节点能耗最小的同时，由于离汇聚越近的子区域内的簇头由于平均能耗负载越大耗能越多，为了缓解不同子网内能耗不均，应当在离汇聚节点越近的区域部署越多的簇头节点来。因此，对于如图 14 所示的大规模多跳传感器网络，我们需要量化在多跳转发过程中不同跳数时，各个子网内分担转发数据的簇头数、寻找全网不同子网间最优的簇头分布以保证全网所有总簇头耗能最小的同时任意两个相邻子网间的簇头在一轮里面耗能均衡、从而延长整个网络的生命周期。

这样，求解出上述两个问题，本质上就确定了在一个大规模多跳分簇的无线传感器网络中，在网络生命周期的一轮里，离汇聚节点远近不同区域的簇尺寸如何分布以

保证全网内所有簇头的总能耗最小同时保持均衡。

再次，由于网络数据融合后的数据量与簇内成员节点数成正比，我们将考察不同的数据融合比对簇头数量的分布以及网络能耗的影响。

最后，本文将依据簇头工作时间的长短、节点剩余能量状态以及簇间传输距离等因素，结合已有的被证明为有效的簇头轮转、动态建簇以及在簇间动态能量感知路由等策略，来设计模型相关的仿真协议，进一步促进整个网络所有节点的能耗均匀。

### 3.3 无线传感器网络簇头节点总能耗最小策略

在本节中，我们首先对如图 14所示的大规模无线传感器网络中，簇头的数据及能耗负载进行分析，接着对决定簇头负载主要因素的多跳传输平均距离进行估算，然后在此基础上，实现我们的第一个研究目标，即对如图 14所示的大规模网络，求解簇间最优的跳数以达到在一轮里面全网内所有簇头节点总耗能最小（即求解最优的子网数目和每个子网相应的边长及划分整个网络的子网区域大小），从而提高网络所有簇头能耗利用率、延长网络的生命周期。

#### 3.3.1 子网内簇头数据及能耗负载分析

本小节主要对子网内簇头的传感数据处理能耗负载进行分析。

在大规模多跳分簇网络中，簇头是网络中能耗最多的节点，其传感数据处理能耗主要包括以下三个方面：1) 簇内和簇间数据接收能耗，包括用于接收从之前子区域簇头转发的多跳累加的数据和簇内所有成员节点的数据的能耗，2) 数据融合能耗 和 3) 向下一子区域或汇聚节点发送簇间转发的数据的能耗。

在分析簇头能耗之前，我们假设在如图 14 所示子网区域  $i$  内的传感器节点数为  $n_i$ ，簇头数为  $c_i$ ；在每一轮里，每个成员节点向相应的簇头发送一个长度  $q_0$  比特的数据包。如果记子网区域  $i$  内每个簇头用于接收数据的平均能耗为  $E_{Rx}(a, i)$ ， $a$  为全网数据融合比率 ( $0 < a \leq 1$ )，根据方程 (3)，有：

$$E_{Rx}(a, i) = \frac{\sum_{k=i+1}^m an_k q_0 + n_i q_0}{c_i} E_{elec} \quad (7)$$

每个簇头用于簇内数据融合的平均能耗，根据 (4)，有：

$$E_{agg}(i) = \frac{n_i q_0}{c_i} E_{DA} \quad (8)$$

这里  $E_{DA}$  表示融合每个比特的数据所耗的能量。

每个簇头节点用于转发本簇内融合后的数据能耗和簇间累加转发的数据的平均

能耗，根据方程 (5)，有：

$$E_{Tx}(a, i) = \frac{\sum_{k=i}^m a n_k q_0}{c_i} (E_{elec} + \epsilon_{mp} d_i^A) \quad (9)$$

这里， $1 \leq i \leq m$ 。

由于每个子网区域为正方形区域，网络中所有节点均匀分布在这些子区域中因此有：

$$n_1 = n_2 = \cdots n_m = n \quad (10)$$

$$m = \lfloor \frac{S}{l^2} \rfloor. \quad (11)$$

这里， $N$  为整个网络的总节点数， $n = \lfloor \frac{N}{m} \rfloor$ 。

根据方程 (6)~(10)，可以得到第  $i$  个子区域内簇头的平均能耗为：

$$\begin{aligned} E_{CH_i}(a) &= E_{Rx}(a, i) + E_{Tx}(a, i) + E_{agg}(i) \\ &= \frac{\sum_{k=i+1}^m n_k a + n_i q_0}{c_i} E_{elec} + \frac{\sum_{k=i}^m n_k a}{c_i} (E_{elec} + \epsilon d_i^A) + \frac{n_i}{c_i} q_0 E_{DA} \end{aligned} \quad (12)$$

在网络簇结构尚未形成的初始阶段，所有节点扮演同样的角色、拥有同样的被当选为簇头的概率。由上述表达式可知，如果要求每个簇头的平均能耗负载并减少和均衡网络能耗，我们要求解上述任意两个相邻的子网之间的任意两个节点的平均传输距离以及相应网络的子网数目。

### 3.3.2 大规模无线传感器网络多跳传输距离期望

在大规模的无线传感网络中，传感器节点通常都是随机部署的。根据我们调研得知，如何评估和求解节点之间传输距离，却是被大多数学者和研究人员所忽略或由于问题本身的难度，目前传感器网络解决能耗优化的各种协议中，如 UCS<sup>[33]</sup>，HEED<sup>[35]</sup>，UCR<sup>[13]</sup>，SCA<sup>[13]</sup>等协议，大部分未对随机部署的传感器节点之间的平均传输距离做出科学合理的估算而被简化。而无线传感器网络节点之间传输的距离，是决定簇头节点能耗的最关键因素，对簇头之间的平均距离的准确估算直接决定协议性能的优劣。

由上小节对簇头的数据能耗负载进行分析可知，我们要求解对于类似如图 14 所示的矩形网络拓扑结构，我们要求解任何两个相邻方形子网内任意两个簇头节点之间的平均距离。由于网络节点随机均匀部署，那么，如果两个节点集合分别随机、均匀分布在任意两个的相邻的方形区域内，如图 15 所示，分别位于这两个节点集合内的任意两个节点之间的平均距离可以看作为分别位于这两个子区域内的任意两个簇头之间的平均传输距离。为了解决该问题，我们进行了广泛的调研。根据调研得知，这个问题其实最终归结于求两个分别部署在相邻的两个单位正方形的点之间的欧氏距离的期望值，

理论表述如下：

**定理 1：**两个分别随机、均匀部署在相邻的两个单位正方形的点之间的欧氏距离的期望值为一个常数，其值为 $\lambda = 1.088$ <sup>[43] [44] [44] [45] [46]</sup>（具体证明过程请参考附录 B）。

由**定理 1**可以得出如下推论：

**推论 1：**假设  $d$  为任意两个分别随机、均匀部署在相邻两个正方形区域的簇头之间的平均距离期望值， $l$  为正方形边长，那么有：

$$d = \lambda l, \lambda = 1.088 \quad (13)$$

**证明：**假设如图 14 所示的大规模无线传感器网络中，相邻的两个子网区域为如图 15 所示的两个正方形相邻区域，两个正方形子网内分别随机、均匀部署相同数目的传感器节点，其中 $D_1 = \{(x, y) | 0 \leq x \leq l, 0 \leq y \leq l\}$ ， $D_2 = \{(x, y) | l \leq x \leq 2l, 0 \leq y \leq l\}$ 。

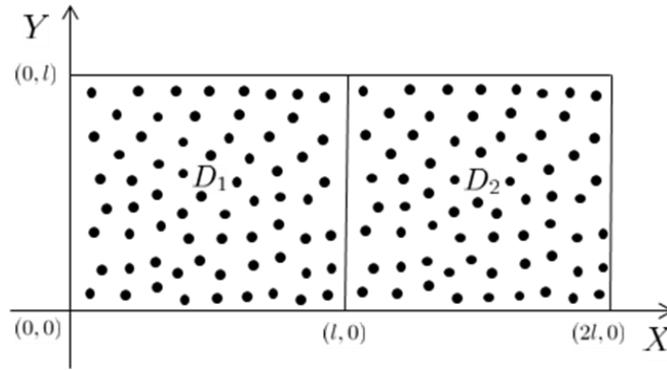


图 15 同构无线传感器网络传感器节点部署场景

我们首先考虑每个子网中只有一个节点随机均匀部署的情形。假设 $X, Y$ 表示这两个点位置的随机变量， $d(X, Y)$ 表示它们之间的欧氏距离， $E(d(X, Y))$ 表示该欧氏距离的期望值。根据**定理 1**，可以得到

$$E(d(X, Y)) = \lambda l.$$

该结论很容易推广到每个子网部署  $n$  个节点的情形。假设  $S$  表示分别位于这两个相邻子网内所有的点对之间的欧氏距离的总和，那么有：

$$S = \sum_{i=1}^n \sum_{j=1}^n d(X_i, Y_j)$$

这里， $\{X_i \in D_1, i = 1, \dots, n\}$ ， $\{Y_j \in D_2, j = 1, \dots, n\}$ 。

由于所有的节点都是随机、均匀分布在每个子网中，而且所有点对之间的欧氏距离变量是相互独立的，因此有：

$$E(S) = E\left(\sum_{i=1}^n \sum_{j=1}^n d(X_i, Y_j)\right) = n^2 E(d(X, Y)).$$

如果假设分别位于这两个相邻子网内任意点对间的距离期望值为  $d$ ，那么有：

$$d = \frac{E(S)}{n^2} = \lambda l, \lambda = 1.088.$$

**推论 1 成立。**

如果假设  $d_i$  表示如图 14 所示网络模型中任意两个相邻两跳子网区域  $i$  和  $i-1$  ( $i = m-1, \dots, 2$ ) 内任意两簇头之间的平均传输距离，由于子网拓扑是对称分布的，因此有：

$$d = d_{m-1} = \dots = d_2 = \lambda l, \lambda = 1.088 \quad (14)$$

### 3.3.3 网络簇头总能耗最小及最优总簇间跳数

在以上两小节中，我们对簇头的传输数据处理能耗进行了分析，并对多跳无线传感器网络簇间传输距离的期望值进行了精确的估算。在此基础上，求解簇间最优的跳数（即求解最优的子网数目）以保证在一轮里面全网内所有簇头节点总耗能最小、延长网络的生命周期。

根据方程 (10), (11), (12) 和 (14)，可知，在一轮里，子网  $i$  中所有簇头的平均耗能为：

$$\begin{aligned} E_{CH_i}(a) &= E_{Rx}(a, i) + E_{Tx}(a, i) + E_{agg}(i) \\ &= \frac{[(m-i)a+1]nq_0}{c_i} E_{elec} + \frac{(m-i+1)naq_0}{c_i} [E_{elec} + \epsilon_{mp}(\frac{\lambda^2 S}{m})^2] + \frac{nq_0}{c_i} E_{elec} \end{aligned} \quad (15)$$

进而由此方程可以得到全网所有簇头的总耗能为：

$$\begin{aligned} E_{CHtotal}(a, m) &= \sum_{i=1}^m c_i E_{CH_i} \\ &= \left\{ \left( \frac{m-1}{2} a + 1 \right) E_{elec} + \frac{m+1}{2} a [E_{elec} + \epsilon_{mp}(\frac{\lambda^2 S}{m})^2] + E_{DA} \right\} Nq_0 \end{aligned} \quad (16)$$

对于给定的不同的网络面积  $S$ ，对上述方程进行数值分析可知，存在这样一个最佳的子网划分数目使得全网簇头的总耗能最小。因为全网的总簇头数  $N$  和数据包的长度  $q_0$  是已知的，我们假设：

$$g(a, m) = \left( \frac{m-1}{2} a + 1 \right) E_{elec} + \frac{m+1}{2} a [E_{elec} + \epsilon_{mp}(\frac{\lambda^2 S}{m})^2] + E_{DA}$$

为了得到使全网簇头总能耗最小的最佳的簇间跳数，我们对上式关于簇间跳数  $m$  进行求偏导可得：

$$\frac{\partial g(a, m)}{\partial m} = \frac{1}{2}aE_{elec} + \frac{1}{2}a(E_{elec} + \frac{\epsilon_{mp}\lambda^4 S^2}{m^2}) - \frac{(m+1)a\epsilon_{mp}\lambda^4 S^2}{m^3} \quad (17)$$

令  $\frac{\partial g(a, m)}{\partial m} = 0$ ，我们发现有一个合理的正实数根满足该等式（其它两个根均为虚根没有实际含义，故舍去），该实根的表达式为：

$$m_{root} = \frac{[6\epsilon_{mp}\lambda^4 S^2(18 + \sqrt{\frac{324E_{elec}-6\epsilon_{mp}\lambda^4 S^2}{E_{elec}}})E_{elec}^2]^{\frac{1}{3}}}{6E_{elec}} + \frac{\epsilon_{mp}\lambda^4 S^2}{[6\epsilon_{mp}\lambda^4 S^2(18 + \sqrt{\frac{324E_{elec}-6\epsilon_{mp}\lambda^4 S^2}{E_{elec}}})E_{elec}^2]^{\frac{1}{3}}} \quad (18)$$

在求得使全网簇头总能耗最小的最佳的簇间跳数时，还需考虑到无线传感器能耗模型对网络拓扑及簇间跳数的影响。为此，我们有以下推论：

**推论 2：**对于基于矩形拓扑的大规模多跳分簇无线传感器网络，网络中所有的节点均匀、随机部署，网络划分成多个正方形子区域，子区域边长均为  $l$ ，如图 14 所示，子网区域间簇头数据转发的平均距离大于或等于  $d_0$ ，如果网络部署面积为  $S$ ，簇间传输的平均距离为  $d$ ，则将大规模无线传感器网络划分能够保证任意两相邻子网间的簇头能够通信的簇间跳数或子网区域数的上界  $m_u$  为：

$$m_u = \begin{cases} 1, & 0 \leq S < \frac{2d^2}{\lambda^2} \\ \lfloor \frac{S\lambda^2}{d_0^2} \rfloor, & others. \end{cases} \quad (19)$$

**证明：**如果我们将一个基于矩形拓扑的大规模多跳分簇无线传感器的网络划分成  $m$  个相同的子区域，如图 14 所示，则有  $S = ml^2$ 。由于所有的传感器节点服从均匀分布在网络部署区域，根据推论 1，我们有  $d = \lambda l$ ,  $\lambda = 1.088$ ，因此：

$$m = \lfloor \frac{S\lambda^2}{d^2} \rfloor.$$

此外，当子网区域簇间数据转发时能耗模型放大器能耗衰减系数为  $\epsilon_{mp}d^4$ ，转发的平均距离必须满足限制条件，即大于或等于  $d_0$ ； $d \geq d_0, m \in Z^+$ ，因此有，

$$m \leq \lfloor \frac{S\lambda^2}{d_0^2} \rfloor.$$

所以， $\lfloor \frac{S\lambda^2}{d_0^2} \rfloor$  为能够保证任意两相邻子网间的簇头能够通信的子网区域划分数的上界。

特别，当  $\lfloor \frac{S\lambda^2}{d^2} \rfloor \leq 2$  时，即  $0 < S < \lfloor \frac{2d^2}{\lambda^2} \rfloor$  有  $m_u = 1$ 。因此，推论 2 成立。

**证明完毕。**

对于不同的网络规模，根据推论 1 和公式 (11)， $m_{root}$  也可以表示为  $\lfloor \frac{S\lambda^2}{d^2} \rfloor$ 。同时，根据推论 2，可以得到能保证如图 14 所示的大规模多跳分簇无线传感器网络，所有簇

头总能耗最小而且任意两相邻子网间的簇头能够通信的子网区域实际的最优的簇间跳数或子网区域数:

$$m_{opt} = \min\{\lfloor m_{root} \rfloor, m_u\} = \lfloor m_{root} \rfloor$$

因此, 我们有以下定理:

**定理 2:** 对于基于矩形拓扑的大规模多跳分簇无线传感器网络, 网络中所有的节点均匀、随机部署, 网络划分成多个正方形子区域, 子区域边长均为  $l$ , 如图 14 所示, 子网区域间簇头数据转发的平均距离大于或等于  $d_0$ , 如果网络部署面积为  $S$ , 簇间传输的平均距离为  $d$ , 则为保证网络生命周期一轮中, 所有簇头节点总能耗最小的实际的最优跳数或子网划分数  $m_{opt}$  可以由如下表达式计算:

$$m_{opt} = \begin{cases} 1, & 0 \leq S < \lceil \frac{2d^2}{\lambda^2} \rceil \\ \lfloor m_{root} \rfloor, & others. \end{cases} \quad (20)$$

这里  $m_{opt}$  与全网的统一数据融合比无关,  $m_{opt} \in Z^+$ 。

对于  $0 < S < \lceil \frac{2d^2}{\lambda^2} \rceil$  的情形, 网络的最优子网数是 1, 即为单跳网络, 这包括和论证了能量优化协议 LEACH<sup>[31]</sup>所采用网络模型中, 从簇头到汇聚节点只有一跳的传输策略。对于其它情形, 不同网络规模的实际的最优的簇间跳数或子网数可由公式 (20) 计算得到。

### 3.4 大规模多跳无线传感器网络簇头能耗均衡策略

在上节中, 我们求解到了对于如图 14 所示的大规模多跳分簇无线传感器网络, 能保证所有簇头总能耗最小而且任意两相邻子网间的簇头能够通信的子网区域实际的最优的簇间跳数或子网区域数。在此基础上, 在本节中, 主要工作如下:

首先, 我们对如图 14 所示网络中任意两个相邻子网内簇头能耗均衡能耗的条件进行分析:

其次, 我们通过建立模型, 着力于量化不同跳数时每个子网区域内的簇头数、寻找全网不同子网间最优的簇头分布以保证任意两个相邻子网间的簇头能耗均衡。这样, 就确定了在一个大规模多跳分簇的无线传感器网络中, 在网络生命周期的一轮里, 离汇聚节点远近不同区域的簇尺寸如何分布以保证全网内所有簇头的总能耗最小同时保持均衡、从而延长整个网络的生命周期。

最后, 我们在求得的最佳的子网间簇头尺寸的分布基础上, 定量分析不同的数据融合比对不同子网内簇头数分布以及网络能耗均衡的影响。

### 3.4.1 网络簇头能耗均衡影响因素分析

在本小节中，我们对如图 14 所示网络中任意两个相邻子网内簇头能耗均衡能耗的条件进行分析。为了避免整个网络中出现“热点区域”、达到能耗均衡，网络中所有的节点在每一轮里能量负载应该保持均衡。由于簇头是网络中主要的耗能节点要使整个网络能耗均衡，首先应该保证全网内所有簇头的能耗均衡。为便于分析，我们先求解任意两个相邻子网内簇头能耗均衡能耗的条件，从而保证全网内所有簇头能耗均衡。如果记子网划分数目为  $m$ ，对于不同的数据融合比  $a$ ，则要保证任意两个相邻子网内的簇头的平均能耗均衡能耗，有：

$$E_{CH_i}(a) = E_{CH_{i-1}}(a), \forall i \in \{2, \dots, m\} \wedge i \in Z^+. \quad (21)$$

这里， $E_{CH_i}(a)$ 和  $E_{CH_{i-1}}(a)$  分别为每一轮里子区域  $i$  和子区域  $i-1$  的簇头平均能耗负载。

当然，由于在每一轮里，簇头的能耗负载远远大于成员节点的能耗负载，仅仅使得簇头之间的能耗负载均衡只能从较大程度上缓解网络的能耗不均、而不能完全保证全网所有节点能耗均匀。

### 3.4.2 网络簇头数与能耗均衡

本小节主要考察对于如图 14 所示的大规模多跳分簇无线传感器网络，在网络生命周期的一轮里，为了达到全网内所有簇头能耗均衡，不同跳数或子网内簇头数分布所应满足的规律。通过求解网络的跳数和不同跳数或子网内的簇头数来确定网络簇头尺寸的分布后，定量分析不同的数据融合比对不同子网内簇头数分布以及网络能耗均衡的影响。

根据上小节分析可知，为了使整个网络不同跳数时簇头能耗均衡，对于不同子网内每个簇头的平均能耗负载相同，对于不同的数据融合比  $a$ ，根据方程 (21)，即有：

$$E_{CH_1}(a) = E_{CH_2}(a) = \dots = E_{CH_{m-1}}(a) = E_{CH_m}(a)$$

对如图 14 所示的大规模多跳传感器网络，我们通过求得不同跳数时每个子网内的最优簇头数分布来保证不同子网区域的簇头能耗均衡，即使得  $E_{CH_{i-1}}(a) - E_{CH_i}(a)$  最小。通过方程 (15)，可以得到相邻子两跳或子网区域簇头之间的平均能量负载差值为：

$$\Delta E_{CH}(a, c_{i-1}, c_i) = E_{CH_{i-1}}(a) - E_{CH_i}(a) = f(a, c_{i-1}, c_i) \frac{N}{m} q_0 \quad (22)$$

这里

$$\begin{aligned}
& f(a, c_{i-1}, c_i) \\
&= \left( \frac{m-i+1}{c_{i-1}} - \frac{m-i}{c_i} \right) a \left[ 2E_{elec} + \epsilon_{mp} \left( \frac{\lambda^2 S}{m} \right)^2 \right] \\
&+ \left( \frac{1}{c_{i-1}} - \frac{1}{c_i} \right) \left[ a \left( E_{elec} + \epsilon_{mp} \left( \frac{\lambda^2 S}{m} \right)^2 \right) + E_{elec} + E_{DA} \right]
\end{aligned} \tag{23}$$

通常，由于随着簇头通过多跳至离汇聚节点越近的子网区域时，簇头的平均能耗负载越大，因此  $\Delta E_{CH}(a, c_{i-1}, c_i) \geq 0$ ，从而  $f(a, c_{i-1}, c_i) \geq 0$ 。根据上面分析可知，要想使不同跳数或子网区域内簇间平均能耗负载差值最小，即有：

$$\min\{\Delta E_{CH}(a, c_{i-1}, c_i)\} = 0, f(c_{i-1}, c_i) = 0 \tag{24}$$

因此，当离汇聚节点最近或最远的子网区域内的簇头作为前提条件给定时，其它子网内的簇头节点数可以根据等式方程 (23) 和 (24) 采用迭代的方法计算得知。

在本文中，我们采用 LEACH 中计算最优簇头数的方法来计算出离汇聚节点最远的子网区域内的初值。由于离汇聚节点最远的子网区域  $m$  内不需要转发上一跳的数据，根据 LEACH<sup>[31]</sup> 中公式 (19)，通过对该区域内簇头节点的总耗能关于该区域内的簇头数  $c_m$  求导，得到：

$$c_m = \sqrt{\frac{N}{2\pi m}} \sqrt{\frac{\epsilon_{fs}}{\epsilon_{mp}}} \frac{l}{d^2} = \frac{1}{\lambda^2} \sqrt{\frac{N}{2\pi S}} \sqrt{\frac{\epsilon_{fs}}{\epsilon_{mp}}} \tag{25}$$

### 3.4.3 网络簇尺寸分布、数据融合比与能耗均衡

由上小节的分析，我们可以求得使所有簇头总能耗最小的最优簇间跳数或子网区域数的前提下，进一步得到了为使任意相邻两跳子网内簇间平均能耗负载均衡时任一子网内的簇头数目的表达式。在此基础上，我们进一步定量分析，同时满足使所有簇头总能耗最小和任意相邻子网内簇间平均能耗负载均衡时，离汇聚节点远近不同跳数或子网内簇尺寸分布所应满足的规律以及不同的数据融合比对不同子网内簇头数分布以及网络能耗均衡的影响。我们得到以下理论：

**定理 3：**对于基于矩形拓扑的大规模多跳分簇无线传感器网络，网络划分成若干个相同的正方形子区域，所有的节点随机、均匀部署在整个网络中，网络中所有传感器采集的数据流沿着一个方向经过多跳最终汇聚到网络部署区域外的一个汇聚节点，如图 14 所示。在网络生命周期的每一轮里，不同跳数的子网区域内所有的簇头将成员节点采集到的数据按照统一的数据融合比进行处理，融合后的数据量与该子区域内的节点数成正比。那么，我们有：

1) 为了避免“热点区域”、达到全网所有簇头的能耗均衡，不同跳数的子网区域的

簇头随着离汇聚节点的距离递减而其数目严格地线性递增；

2) 对于不同的数据融合比，规律 1) 仍然成立，而且数据融合比越大，为了维持不同跳数子网间簇头能耗均衡所需簇头数目越多。

**证明：**对于如图 14 所示网络部署场景，在网络生命周期的每一轮，为了避免“热点区域”、达到全网所有簇头的能耗均衡，任意两个相邻的跳数的子网内的簇头数应满足等式(24)。假设

$$E_1 = 2E_{elec} + \epsilon_{mp} \left( \frac{\lambda^2 S}{m} \right)^2 \quad (26)$$

$$E_2 = E_{elec} + \epsilon_{mp} \left( \frac{\lambda^2 S}{m} \right)^2 \quad (27)$$

$$E_3 = E_{elec} + E_{DA} \quad (28)$$

将 (26)，(27)和(28)代入方程 (24)得到：

$$c_{i-1}(a) = \left[ 1 + \frac{a(E_1 + E_2) + E_3}{a(m-i)E_2 + E_3} \right] c_i(a). \quad (29)$$

从而，由等式 (25) 和 (29) 我们可以得到，为使不同跳数或子网区域内簇间平均能耗负载差值最小时任一子网内的簇头数目：

$$c_i(a) = \left[ 1 + \frac{a(m-i)E_1}{aE_2 + E_3} \right] c_m, \forall i \in \{1, \dots, m\}. \quad (30)$$

由表达式 (30) 可知，为使不同跳数或子网区域内簇间平均能耗负载差值最小，任一子网内的簇头数目取决于数据融合比及该簇头到汇聚节点之间的跳数或距离（即与该簇头所在的子网区域 ID 有关）。记：

$$H(a, \mu) = 1 + \frac{a(m-\mu)E_1}{aE_2 + E_3}. \quad (31)$$

这里， $a$  (这里， $0 < a \leq 1$ )是全网统一的数据融合比， $\mu$ 是不同的跳数或子网 ID 且  $\forall \mu \in \{1, \dots, m\} \wedge \mu \in Z^+, \forall a \in (0, 1]$ 。

首先，我们分析对于给定的不同的数据融合比，为达到全网所有簇头的能耗均衡，全网不同跳数的子网区域内的簇头数所应满足的函数关系，在此基础上得到全网簇尺寸分布规律。对  $H(a, \mu)$  求关于  $\mu$  的偏导数可得：

$$\frac{\partial H(a, \mu)}{\partial \mu} = -\frac{aE_1}{aE_2 + E_3}, a \in (0, 1], \forall \mu \in \{1, \dots, m\}. \quad (32)$$

由于 $E_1, E_2, E_3$  全是正整数，因此， $\frac{\partial H(a, \mu)}{\partial \mu} < 0$ 。在给定不同的数据融合比时， $h(a, \mu)$  在其区间内对于不同的跳数或子网 ID，是一个严格单调递减函数。因此，我们可以得知，

为达到全网所有簇头的能耗均衡时，不同跳数的子网内的簇头数与子网 ID 之间呈严格的线性关系。而且，为了避免“热点区域”的出现和全网的能量均衡，簇头跳数或子网区域离汇聚节点愈近（在方程 (30) 中意味着子网 ID  $\mu$  从  $m$  递减至 1），所需的簇头数目越多，而且呈严格的线性递增趋势。而且由此可以得出，文献[8]SCA 协议只是 COCA 协议当数据融合比为  $\alpha = 0.1$  时的一个特例。

其次，我们定量分析不同的数据融合比对不同子网内簇头数分布以及网络能耗均衡的影响。图 16 和图 17 分别显示了当全网总节点数  $N = 400$  和  $N = 800$  时，不同数据融合比下为达到全网所有簇头的能耗均衡时全网不同跳数的子网区域内的簇头数的分布规律。从图中可以看出，对于不同的数据融合比，随着跳数或子网与汇聚节点之间距离的递减，所需的簇头数严格地线性递增，而且对于不同网络规模和不同的数据融合比，增长规律保持一致。而且数据融合比越小，所需的簇头数就越少。当  $\alpha = 0$  时，簇头的变化曾一条直线分布，这是因为：

$$\lim_{\alpha \rightarrow 0^+} H(\alpha, \mu) = 1.$$

即在理想状态下，当数据融合比为 0 时，簇间没有累加转发数据，不同跳数子网内的每个簇头能耗负载相同，因此不同跳数子网内的簇头数目保持不变就能保证能耗均衡。

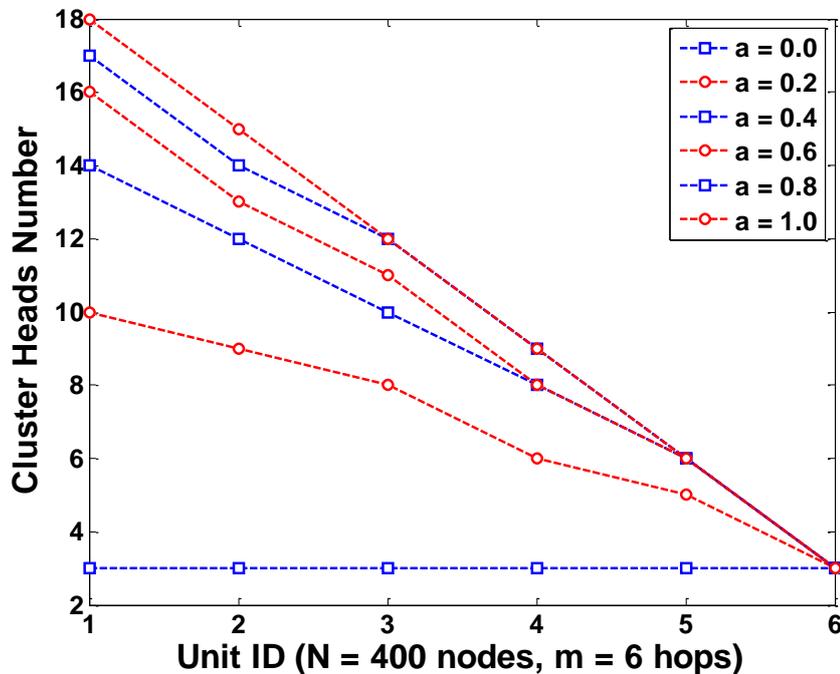


图 16 不同数据融合比时簇头分布规律

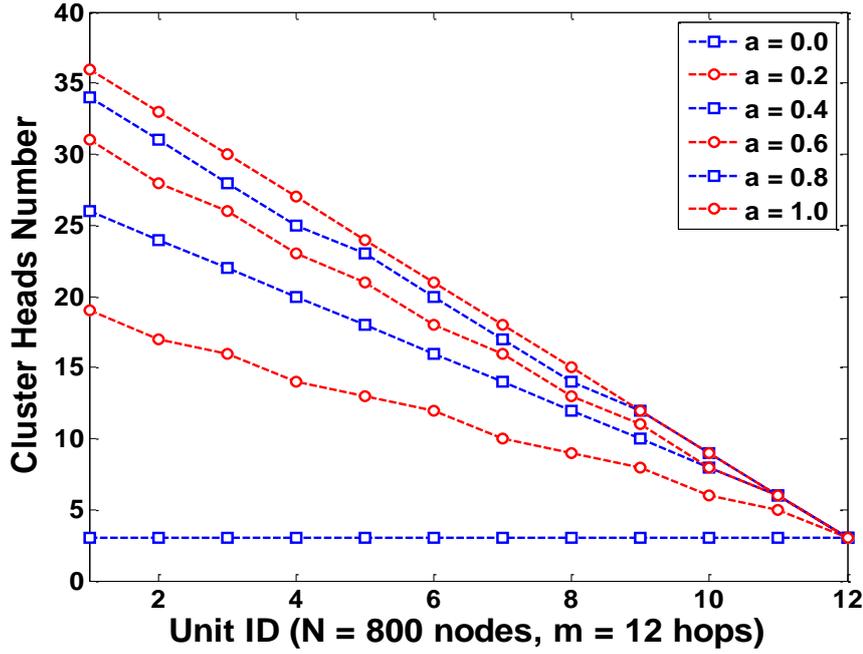


图 17 不同数据融合比时簇头分布规律

当然，在实际应用中，数据压缩比不可能取 0，总存在一个下界。在本文中，我们采取如下的方式定义数据融合比的下界：数据融合比至少应该保证离汇聚节点最远端的那个网内的每个簇头通过数据融合后至少分别有一个数据包可以通过多跳转发成功到达汇聚节点：

$$a * n \geq c_m(a)$$

通过方程 (30) 和  $n = \lfloor \frac{N}{m} \rfloor$ ，我们有：

$$NE_2a^2 + [NE_3 - mc_mE_2 - mc_m(m-i)E_1]a - mc_mE_3 \geq 0$$

我们将满足上式的数据融合比定义为数据融合比的下界  $\min a$ ，那么在实际应用中的合理的数据融合比值变化范围为： $\min a \leq a \leq 1$ 。

接着，我们研究不同的数据融合比对于每跳单个网内簇头增量的影响。对  $H(a, \mu)$  关于数据融合比  $a$  求偏导数可得：

$$\frac{\partial H(a, \mu)}{\partial a} = \frac{(m - \mu)E_1E_3}{(aE_2 + E_3)^2}, \forall a \in (0, 1), \mu \in \{1, \dots, m\}.$$

从该函数关系可以看出，满足使所有簇头总能耗最小和任意相邻网内簇间平均能耗负载均衡时，每跳单个网内的簇头的增量随着数据融合比的增大而增加，但这个增幅的变化随着数据融合比的增大而逐渐减缓。

证明完毕。

如图 18 和图 19 分布描述了当全网总节点数  $N = 400$  和  $N = 800$  时，不同的数

据融合比对离汇聚节点最近一跳那个子网内簇头数量增量的影响。可以看出，当数据融合比较小小时，数据融合比对单个子网内簇头增幅变化影响比较大，当数据融合比较大时，增幅变化减缓。而且当数据融合比在一定小范围内变化时，簇头的增量为一个常量，簇头数量增量随着数据融合比的变化整体呈阶梯状递增趋势，且对较大规模网络（ $N=800$ ）增幅的变化更为剧烈。这是因为，较大规模的网络需要累加传输的数据量越多，对全网数据融合比的变化更为敏感。

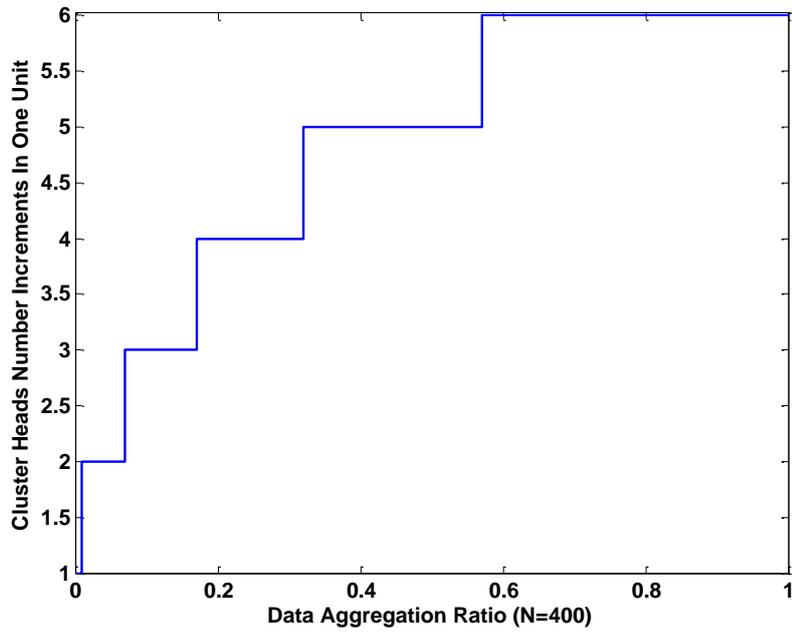
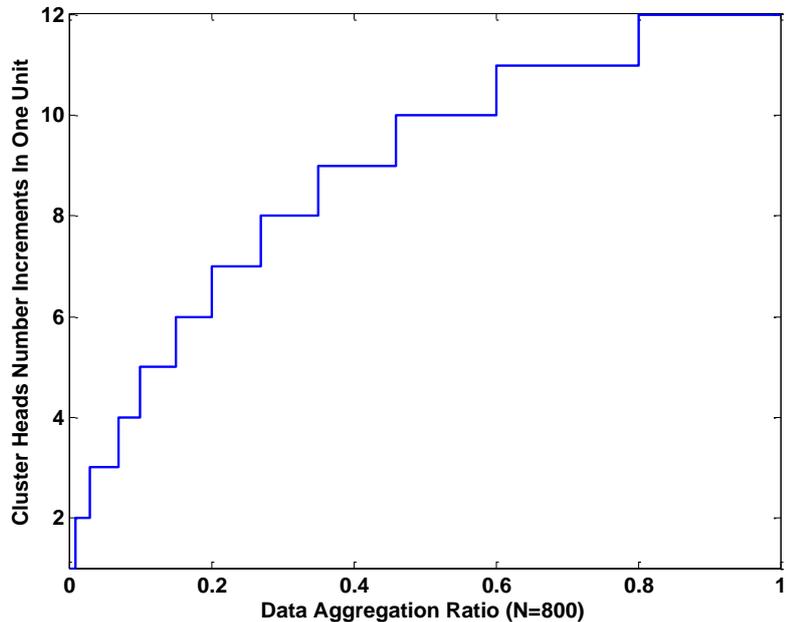


图 18 数据融合比对单个子网内簇头数量增量的影响（ $N=400$ ）



---

图 19 数据融合比对单个子网内簇头数量增量的影响 (N=800)

### 3.5 本章小结

本章针对大规模无线传感器网络的能耗问题，通过对网络拓扑、簇头数目、数据融合等影响网络能耗的因素进行建模分析，寻找到了最优的网络跳数以保证在网络生命周期的一轮里面全网内所有簇头总耗能最小；求得了各个子网内分担转发数据的簇头数以保证全网所有总簇头耗能最小的同时任意两个相邻子网间的簇头在一轮里面耗能均衡。最后，进一步定量分析了同时满足使所有簇头总能耗最小和任意相邻子网内簇间平均能耗负载均衡时，离汇聚节点远近不同跳数或子网内簇尺寸分布所应满足的规律以及不同的数据融合比对不同子网内簇头数分布以及网络能耗均衡的影响。

## 第四章 模型仿真及性能分析

本章主要对第三章所建立的无线传感器数据处理与能耗优化理论模型进行仿真及其对仿真结果进行对比和分析。其中，4.1 节介绍了仿真平台及系统相关的参数设置以及本文设计策略相应的仿真协议 COCA(Constructing Unequal Clustering Architecture)协议及其对比协议 UCR 协议参数设置，4.2 节介绍了模型相关的仿真协议的设计，4.3 节测试和选取了网络簇头最佳轮转周期以及对 COCA 协议与 UCR 协议在不同大规模网络下网络的生命周期的性能进行了对比与分析，4.4 节考察了 COCA 与 UCR 协议下网络节点的能量利用率以及不同的数据融合比对网络节点能量利用率的影响。

### 4.1 仿真平台和系统参数设置

#### 4.1.1 仿真平台

本文的仿真系统基于 Debian GNU/Linux 系统，采用经典的 NS2 仿真模拟器，用 TCL (Tool Command Language) 脚本语言开发<sup>[41]</sup>。仿真结果采用 MatLab 进行统计分析。

#### 4.1.2 系统参数设置

仿真系统的参数设置主要包括以下三个方面：

- 1) 传感器节点能量模型系统参数设置：包括传感器节点发送端电路能量，放大器能量，距离阈值，节点初始能量及数据融合比；
- 2) 无线传感器网络模型参数设置：网络生命周期，包括网络生命周期种类的定义，衡量网络生命周期的一轮的的定义，簇头最佳轮转周期的定义；网络拓扑，包括网络面积，传感器节点的部署位置，部署密度；网络通信，包括传感器节点数据包和控制消息通信包的大小的设置；
- 3) 协议相关参数设置：包括 COCA 协议和 UCR 协议相关的参数设置。

其中，仿真系统前两部分参数的设置分别见表 5 和表 6。

#### 4.1.3 COCA 协议参数设置

我们的策略相应的仿真 COCA 协议，其中的能量模型和网络模拟相关设置同 4.1.2 小节相同。其它协议相关参数，根据第三章所建立的无线传感器数据处理与能耗优化理论模型得到的理论分析值，包括不同网络规模对应的最优的簇间跳数或区域划分数目及

相应的正方形子网区域边长，不同跳数内子网簇头数的分布，分别设置如表 7。

表 5 系统模型相关参数设置

参数名称	参数值
传感器节点初始能量	2.0 J
传输距离阈值 $d_0$	87.0 m
传感器节点发送端或接受端电路消耗能量 $E_{elec}$	50 nJ/bit
电容放大器能耗 $\epsilon_{fs}d^2$ 或 $\epsilon_{mp}d^4$	$\epsilon_{fs}$ 10pJ/bit/m <sup>2</sup> , $\epsilon_{mp}$ 0.0013pJ/bit/m <sup>4</sup>
数据融合能耗 $E_{DA}$	50 nJ/bit/signal
数据融合比 $\alpha$	0.2, 0.4, 0.6, 0.8

表 6 无线传感器网络模型参数设置

网络生命周期种类定义	网络第一个节点死亡和部分节点死亡定义
网络生命周期的一轮 round	包括选簇或重选簇及簇头轮转，建簇，簇内数据汇报，簇间路由及数据转发阶段
簇头最佳轮转周期	依据前期实验确定
单个数据包或控制消息包 Data 字节长度	500Bytes
单个数据包头或控制消息包 HDR 字节长度	25Bytes
网络部署面积 $S$	$S = 3.0 * 10^4 m^2 \sim S = 1.2 * 10^5 m^2$
节点部署密度 $\rho$	1 node/100m <sup>2</sup>
传感器节点的部署位置	采用均匀分布随机函数模拟

表 7 COCA 协议参数设置

网络规模（全网总结点个数）	簇头跳数或子网个数	正方形子网区域边长	每跳子网簇头数（距离汇聚节点依次由远及近）
300	4	86.602540	3, 4,5,6
400	6	81.649658	3,4,5,6,7,8
500	7	84.515426	3,4,5,6,7,8,9
600	9	81.649658	3,4,5,6,7,8,9,10,11
700	10	83.666003	3,4,5,6,7,8,9,10,11,12
800	12	81.649658	3,4,5,6,7,8,9,10,11,12,13,14
900	13	83.205029	3,4,5,6,7,8,9,10,11,12,13,14,15
1000	15	81.649658	3,4,5,6,7,8,9,10,11,12,13,14,15,16,17
1100	16	82.915620	3,4,5,6,7,8,9,10,11,12,13,14,15,16,17,18
1200	18	81.649658	3,4,5,6,7,8,9,10,11,12,13,14,15,16,17,18,19,20

#### 4.1.4 UCR 协议参数设置

由于大量的仿真结果显示，最新的多跳分簇无线传感器网络协议 UCR 协议<sup>[13]</sup>比经

典的 HEED 协议<sup>[35]</sup>能更进一步提高能耗利用效率、提高网络生命周期。因此，本文将选取 UCR 协议作为我们的对比协议，来进行仿真结果的对比和分析。这里，不像文献 UCR 协议中簇头将所接受的数据全部融合成一个包，我们模拟的作为本文对比的 UCR 协议中，所有的簇头将其采集到簇头的的数据均匀地按照融合比  $\alpha$  进行融合。UCR 协议

的能量模型和网络模拟相关设置同 4.1.2 小节相同，协议其它参数设置如表 8。

表 8 UCR 协议参数设置

参数名称	参数值
普通节点成为候选簇头概率 T	0.2
簇头竞争半径差异比率 c	0.3
候选簇头最大竞争半径 $R_0$	1) 保证产生的簇头数与 COCA 一样多 2) 在区间 $[\frac{l}{2}, l]$ 之间
最佳簇间候选簇头数 $k$	2
与汇聚节点直接通信半径阈值 TD MAX	200m

## 4.2 模型与协议仿真设计

在这节中，我们采用类似文献 [14] 中 SCA 协议中的路由设计原理来设计我们的路由。在大规模多跳无线传感器分簇的同构网络中，簇头能耗负载远远大于成员节点能耗，是能耗最多的关键节点，在第三章关于数据处理与能耗优化策略的理论建模分析中，我们主要对簇头的能耗进行分析。本节通过考察簇头剩余能量状态以及簇间传输距离等因素，在不同跳数的子网内所有节点之间采用动态建簇及在不同子网间建立动态路由，从而保证网络的传感数据经过处理后能够传输至汇聚节点的同时，使各个子网内和子网间的能耗更加均衡。

### 4.2.1 网络模型初始化建立

这个阶段主要是汇聚节点在全网内对网络进行初始化设置，建立多跳虚拟的子网模型。

在初始阶段，汇聚节点首先向全网广播网络生命周期定义类型，由第三章建模分析得到的为保证全网簇头总能耗最小的网络跳数  $m_{opt}$  和相应的虚构子网划分模型以及为保证全网簇头能耗均衡时不同跳数子网内簇头数的分布  $\vec{C} = \{c_m, \dots, c_1\}$ 。然后，在全网不同跳数的子网区域进行簇头选取。在网络簇结构尚未形成的初始阶段，所有簇头节点剩余能量是随机的，簇头的选取是随机而且每个节点具有同等的概率被当选为簇头的概率。当网络运行一定阶段后，网络中所有节点的剩余能力发生了变化，簇头选取将

---

依据节点的剩余能量最高、簇内传感数据传输距离最短的节点重新作为簇头，其中簇头最佳重选轮转周期将在下一节进行讨论。确定后的簇头最佳轮转周期由汇聚节点在网络的初始化阶段向全网广播。当传感器节点接收到汇聚节点广播的消息之后，建立了多跳虚拟的子网模型。

#### 4.2.2 子网内部簇头选取、簇结构形成及簇内数据处理

网络初始化之后，各个子网模型建立，然后各个子网内所有节点之间根据节点的剩余能量、传感器节点之间的数据传输距离竞选簇头，根据选定的簇头建立簇结构以及簇内数据汇报及融合处理。

假设簇头重新选取的周期是预先知道的（簇头的重选周期是指各个网络每个一定的时间段便重新选取簇头，我们将在下节讨论重选簇头的原因及如何获取和设置簇头重选周期值），一旦建簇过程开始，每个节点在相应的子网内向其它节点广播一个簇头竞选消息 Adv\_Comp\_CH，发送它的剩余能量状态值，子网内的每个节点从其他  $n_i - 1$  个节点接收到 Adv\_Comp\_CH。对于任意一跳子网  $i$ ，子网内的簇头数  $c_i$  是预先知道的，子网内的每个节点能够监听其它节点，因此剩余能量最高的  $c_i$  个候选簇头能够感知其它  $n_i - c_i$  个成员节点的位置，每个存活的候选簇头向其它  $n_i - 1$  节点中的存活节点广播消息 Adv\_CH，通知自己当选成为新的簇头。其他  $n_i - c_i$  个成员节点接收到  $c_i$  个簇头中存活的簇头广播的当选簇头的消息 Adv\_CH。这样，每个子网最终选出了簇头。

当  $c_i$  个簇头选定后，每个子网内， $n_i - c_i$  个节点中存活的成员节点分别向新选出的  $c_i$  个簇头中存活的簇头广播加入簇头的成簇消息 Adv\_Join\_CH。成员节点加入簇头的衡量标准是它们接收到的信号强度和与簇头之间的距离。每个节点将选择信号强度最大而且距离自己最近的簇头加入。每个簇头的成员数目是  $Mem_{num} = n_i/c_i$ ，对于剩余的  $n_i \% c_i$  个节点依次从这  $c_i$  个簇头中依次选择能量最高的分别加入其中。每个簇头等待一定的时间，以保证所有的节点都不再发送成簇消息。最后，当个簇头全部接收到成员节点的成簇消息后，再向全部成员节点发送一个确认消息 Confirm\_Mmb。这样，在每个子网内最终建立了簇结构。由于每个子网只有有限个节点，因此建簇过程能在有限的时间内完成，不难得知，每一轮簇结构形成阶段全网消息的复杂度为  $O(N)$ ， $N$  为网络节点总数。

#### 4.2.3 子网簇间数据处理动态能量感知路由

为了网络的可扩展性，本文采取了多跳数据转发的模型。用于转发数据的下一跳候

选簇头从离汇聚节点较近的相邻跳数内的子网内选取。根据前面的讨论可知，离汇聚节点越近，子网内的簇头数就越多。因此，在簇间路由选取阶段，每个簇头可以从离汇聚节点更近的相邻子网内按照如下原则选取候选簇头：

$$w(s_i, s_{i-1}) = E_{rem}(s_{i-1}) * (E_{rem}(s_{i-1}) - E_{rem}(d(s_i, s_{i-1})))$$

$$\text{或 } w(s_i, s_{i-1}) = E_{rem}(s_{i-1})/d(s_i, s_{i-1})$$

其中  $s_i$  分别为第  $i$  跳子网内的数据转发簇头， $s_{i-1}$  为下一跳即第  $i-1$  跳内的候选转发簇头， $E_{rem}$  为传感器节点的剩余能量， $w(s_i, s_{i-1})$  为簇头  $s_i$  和  $s_{i-1}$  数据传输路由权值。则每次选取不同的路由权值  $w(s_i, s_{i-1})$  最大的簇头作为其下一跳的候选簇头。

在一个簇头重新选取周期内的每一轮里，每个簇头通过与相应的  $r$  个候选簇头进行通信后，将选取其中剩余能量最高的候选簇头作为其下一跳的转发节点。根据等式 (30)，易知，在每一轮里，子网间簇间路由选取阶段的消息复杂度是  $O(1)$ 。

#### 4.2.4 子网簇间数据多跳转发

一旦簇间路由建立后，除了距离汇聚节点最远端区域内的簇头只需要转发本簇内融合后的成员节点的数据外，其它不同跳数子网内的簇头便将从之前子区域簇头转发的多跳累加的数据及自身簇内所有成员节点的汇报数据经过融合后的数据一起转发给下一跳子网区域内由路由算法确定的相应簇头或汇聚节点。

从簇结构形成阶段和子网簇间路由建立阶段的整个过程可知，路由形成阶段均是基于子网内部的局部信息，因此，是分布式算法，从而对大规模网络是可扩展的。

### 4.3 大规模多跳网络生命周期

#### 4.3.1 簇头最佳轮转周期选取

由于簇头工作的时间愈长，簇头能耗愈多大，网络能耗愈不均，容易迅速死亡而导致网络生命周期大大缩短。因此，若经过一定周期后，将不同跳数子网内部的簇头在所有节点间进行轮转，选取节点的剩余能量最高、簇内传感数据传输距离最短的节点作为簇头，将进一步减少节点能耗和促进节点能耗均衡、延长网络生命周期。如果我们将所有的簇头节点将所有的传感器节点采集的数据进行处理后完成一次向汇聚节点的汇报定义为一轮，我们需要确定不同跳数子网内部每隔多少轮进行一次重选簇头作为簇头的最佳轮转周期。然而，簇头最佳轮转周期的选取仍旧是一个尚待解决的问题<sup>[47]</sup>。在无线传感器网络的数据处理能量感知路由的形成和簇头轮转过程中，一方面，如果簇头的轮转周期过短，网络将花费大量的能量用于竞选簇头和簇结构的形成的通

信能耗方面，不利于提高网络传感的数据处理的能耗利用率；另一方面，如果簇头的轮转周期过长，簇头能耗愈多大，网络能耗愈不均，容易迅速死亡而导致网络生命周期大大缩短。因此，本小节的主要内容在于确定簇头的最佳轮转周期，将不同跳数子网内部的簇头在所有节点间进行轮转，选取节点的剩余能量最高、簇内传感数据传输距离最短的节点作为簇头，以此来进一步减少节点能耗和促进节点能耗均衡、延长网络生命周期。但是，由于簇头最佳轮转周期的选取仍旧是一个尚待解决的问题<sup>[47]</sup>，本文采用类似文献 [14] 中 SCA 协议所采取的簇头轮转策略，通过下面的仿真模拟实验来确定簇头的最佳轮转周期，以此来缓解或消除最佳簇头轮转周期的选取对网络生命周期衡量的影响。

为了得到相对稳定和准确的簇头的最佳轮转周期，在我们的仿真测试方案设计中，我们尽可能选取较多的典型的测试用例。在本次仿真模拟试验中，我们对 10 个不同网络规模（从  $S = 3.0 \times 10^4 m^2$  到  $S = 1.2 \times 10^5 m^2$ ）分别进行了测试，每个网络的部署密度均相同。为了描述清晰，我们只描述其中五个规模的结果。图 20 和图 21 分别描述了在第一个节点死亡定义和部分节点死亡定义下网络的生命周期和最佳簇头轮转数之间的关系。

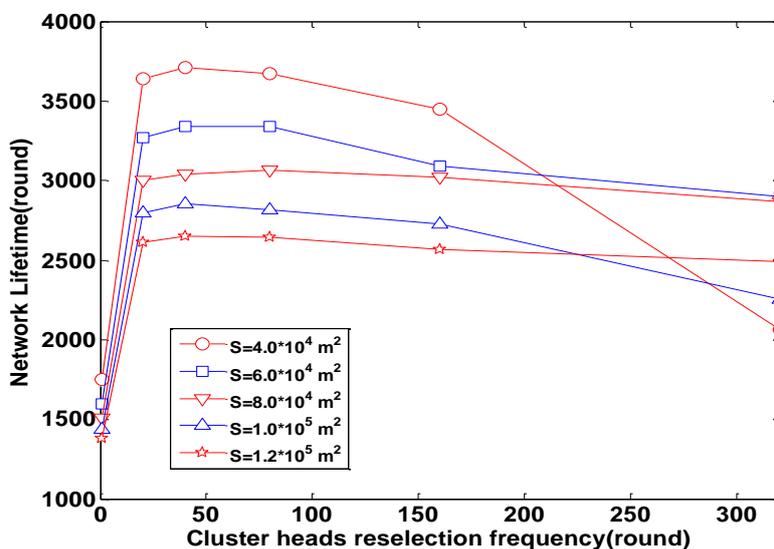


图 20 第一个节点死亡定义下的簇头轮转周期与网络生命周期之间的关系

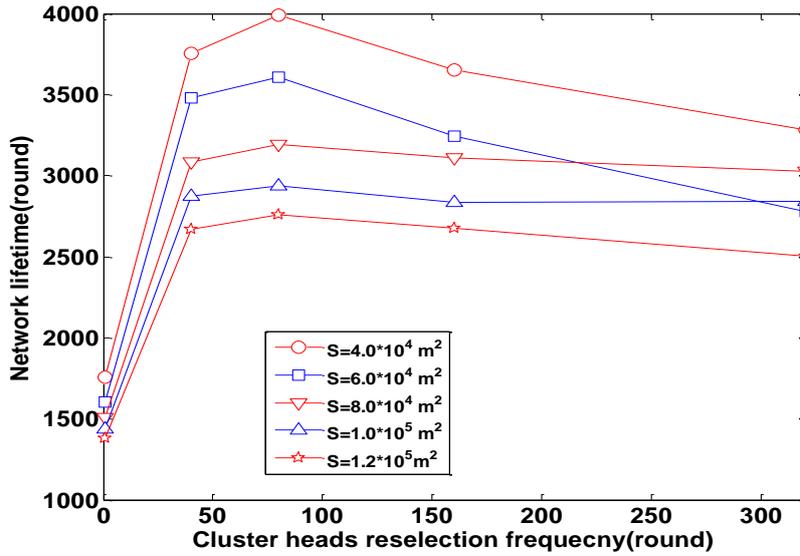


图 21 部分节点死亡定义下的簇头轮转周期与网络生命周期之间的关系

从图 20 和图 21 我们可以看出,通常情况下,网络的生命周期首先随着簇头轮转周期的增大而延长。这是因为重选簇头的轮转周期过短会导致大量的能量浪费在节点的成簇阶段。当轮转周期增长到一定程度后,网络的生命周期会接近一个最大值,然后又逐渐开始缩短。因为当选簇头的时间过长,簇头用于数据转发的能耗更大,每一轮内节点的能耗更不均匀,加速节点的死亡、缩短网络的生命周期。显然,在第一个节点死亡和部分节点死亡的定义簇头的最佳轮转周期分别为 40 轮和 80 轮。

值得注意的是,对于不同的网络规模,尽管它们的网络生命周期不同,但它们的最优的簇头轮转周期几乎一致。因此,尽管网络规模能够影响网络的生命周期,对不同的网络规模我们接下来仍然可以利用这里测得的簇头的最优轮转周期。

#### 4.3.2 大规模多跳网络生命周期

在这节中,我们通过通过不同的大规模多跳无线传感器网络规模来模拟验证我们提出的数据处理和能耗优化策略对于提高网络能耗、延长网络生命周期的影响。这里,我们主要对比我们提出的 COCA 协议和 UCR 不等分簇路由协议下网络的生命周期。

对于影响 UCR<sup>[13]</sup>网络生命周期的最主要的参数之一,候选簇头最大竞争半径  $R_0$ ,对于 10 个不同的网络规模,我们测试了以下两组相应的不同的值:

首先,对于 10 个不同的网络规模,我们对 UCR 测试了一组相应的最大竞争半径,以保证协议执行的每一轮里,对于不同的网络规模,分别产生与 COCA 协议基本相同的总簇头个数。我们发现,为了产生与 COCA 相同的簇头个数,最大竞争半径的选取随着网络规模的增大而减小,而且当网络规模增至  $S = 1.2 * 10^5 m^2$ 时,最大竞争半径减至可

能达到的最小的距离，即网络节点部署的半径，因此，不适应更大规模的模拟。

其次，我们对 UCR 测试了另一组最大竞争半径下，网络的生命周期。我们使最大竞争半径  $R_0$  在区间  $[\frac{l}{2}, l]$  之间变化为了寻找网络最佳的生命周期，这里  $\frac{l}{2}$  和  $l$  分别为 COCA 可能的保证全网能够被覆盖到的最小和最大的簇头最大竞争半径。我们发现，当簇头最大的竞争半径  $R_0$  从  $\frac{l}{2}$  递增至  $l$  时，网络的生命周期逐渐缩短，当  $R_0$  递增至  $l$  时，UCR 甚至网络生命周期不到一轮的时间。因此，我们仅仅只描述了当簇头的最大竞争半径  $R_0$  等于  $\frac{l}{2}$  时 UCR 协议的最佳网络生命周期。此外，我们发现 UCR 协议网络路由在成员节点汇报完数据后、在进行簇间路由通信之前必须及时跟新以达到最佳的网络生命周期。

图 22 显示了，在第一个节点死亡的定义下，对于不同的网络规模，对于不同的簇头最大竞争半径，COCA 与 UCR 相比，能够将网络的生命周期大约延长 14%至 133%左右。如图 23 所示在部分节点死亡的定义下，对于不同的簇头最大竞争半径，COCA 与 UCR 相比，能够将网络的生命周期大约延长 53%至 172%左右。

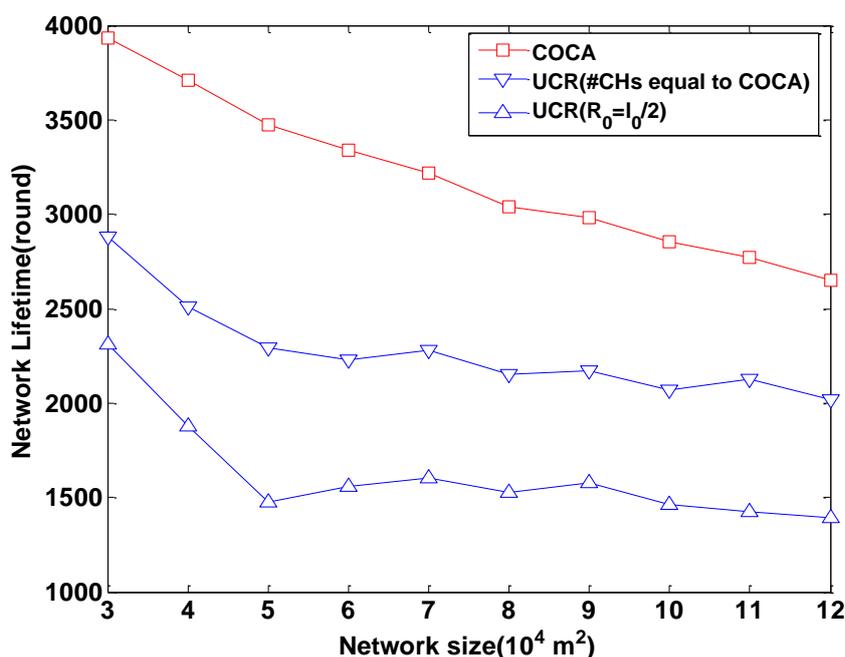


图 22 第一个节点死亡定义下网络生命周期的比较

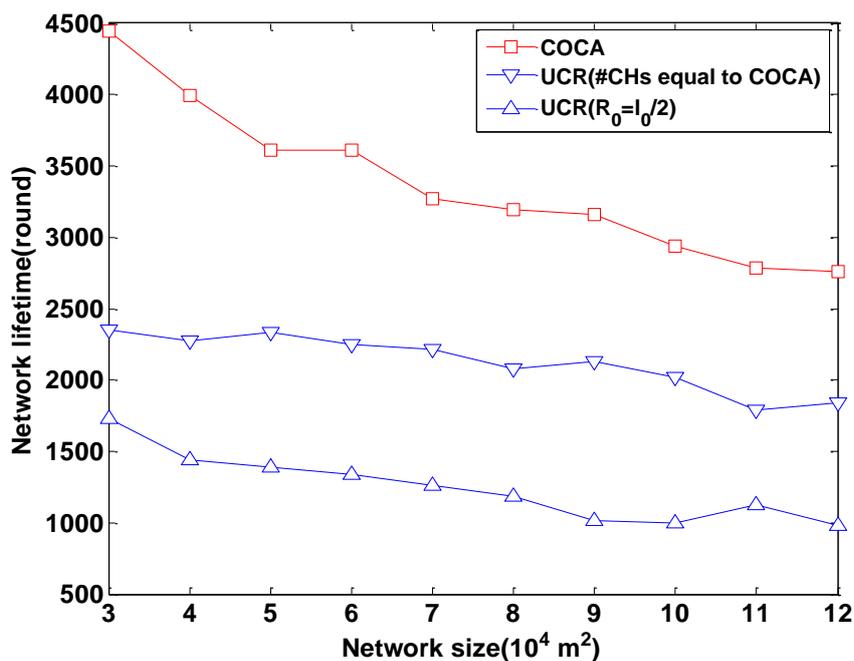


图 23 部分节点死亡定义下网络生命周期的比较

此外，应当指出的是，当网络中某个节点死亡时，整个网络仍然能够运转，直到网络中死亡的节点数过多以至网络的部分“热点区域”被断裂开来。因此，网络生命周期部分节点死亡的定义比第一个节点死亡的定义更具有实际意义。从图 22 和图 23 可以看出，与 UCR 协议相比，在部分节点死亡定义下，COCA 能在更大程度上延长网络的生命周期，更有利于网络大规模的扩展。

#### 4.3.3 协议稳定性分析

在这部分内容中，我们主要讨论 COCA 和 UCR 协议对于不同网络和第一个节点死亡和部分节点死亡两种不同的网络生命周期定义下的协议的稳定性，分别如图 24 和图 25 所示。

对于 COCA 协议，由图 24 可知，对于 10 个不同的网络规模，部分节点死亡定义下的网络生命周期与第一个节点死亡定义下的生命周期都要长。但是，随着网络规模的不断增大，尽管两种不同定义下的协议运转的最佳簇头轮转周期不同，对于特定的网络规模，两种不同定义下的网络生命周期之间的差值越来越小。这说明，在这两种网络生命周期定义下，COCA 协议能够很好地保证全网内能耗均匀而且具有很好的稳定性，有利于网络规模的扩展，尤其适合大规模的网络。

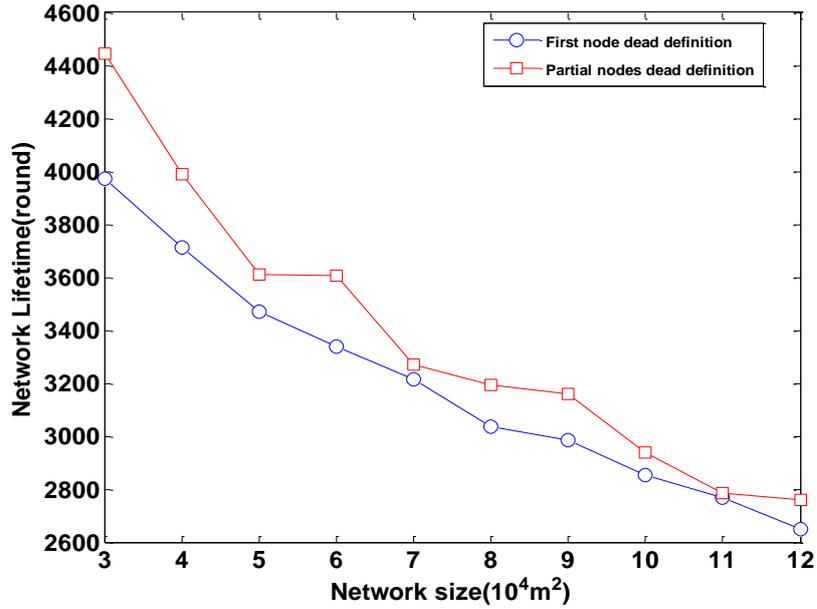


图 24 COCA 两种不同定义下的生命周期的比较

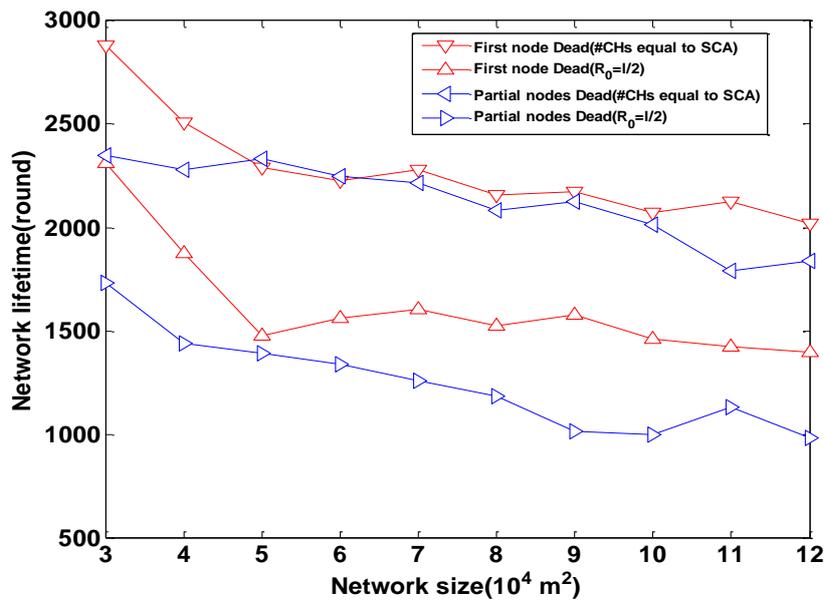


图 25 UCR 两种不同定义下的生命周期的比较。

然而，由于 UCR 的选簇协议——EEUC 是一个基于概率的算法，一方面，它的网络生命周期随着网络协议的参数，比如簇头最大竞争半径  $R_0$  和簇头竞争半径差异比率  $c$  等的变化而波动很大，如图 24 和图 25 所示，而且不适合大规模网络的扩展。另一方面，对于在不同的最佳簇头轮转周期下的两种不同的网络生命周期的定义，对于某些网络规模，部分节点定义下的网络生命周期甚至比第一个节点死亡定义下的网络生命周期还要短，如图 25 所示。因此，对于不同的网络规模，与 UCR 相比，COCA 具有更加良好的稳定性，更适合于大规模网络的扩展。

## 4.4 网络节点能量利用率

### 4.4.1 不同网络规模的节点剩余能量

在本小节中，我们比较了 COCA 和 UCR 协议不同网络规模在不同网络生命周期定义下所有网络节点的剩余能量。由于空间限制，我们仅仅描绘了四个不同网络规模的节点的剩余能量： $S = 3.0 * 10^5 m^2$ ， $S = 5.0 * 10^5 m^2$ ， $S = 7.0 * 10^5 m^2$ ， $S = 9.0 * 10^5 m^2$ ，相应的网络节点总数分别为 300, 500, 700 和 900 个。相应的模拟结果显示在图 26 中。从这四幅图中可以看到：

首先，对于不同的网络生命周期定义，相对于 UCR，COCA 协议下网络节点的剩余能量比较小，很大程度上提高网络中传感器节点的能量利用率、延长了网络的生命周期。

其次，对于一个特定的网络生命周期的定义，相对于 UCR，COCA 协议下每跳子网内节点的几乎具有相同的剩余能量，而且对于不同的网络规模子网之间的剩余能量水平差值非常小。这表明 COCA 协议能够在很大程度上保证网络能耗均衡、从而延长网络的生命周期。

最后，对于一个给定的网络规模，相对于 UCR，COCA 协议下网络中大多数存活节点的剩余能量在部分节点死亡网络生命周期定义下比头结点死亡网络生命周期定义下要少。因此，相对于 UCR，COCA 协议具有更好的实际部署性能。但是，对于不同的网络规模，COCA 协议每个存活的节点在第一个节点死亡和部分节点死亡两种不同的生命周期定义下几乎保持了相同的剩余能量。这说明，COCA 协议对于大规模网络，在全网内保持能耗均衡具有良好的稳定性能。

总之，COCA 协议能够在不同的网络规模和不同的网络生命周期定义下，能在很大程度上保持网络节点能耗均衡的同时提高网络能量利用效率。因此，COCA 对不同网络规模具有良好的适用性。

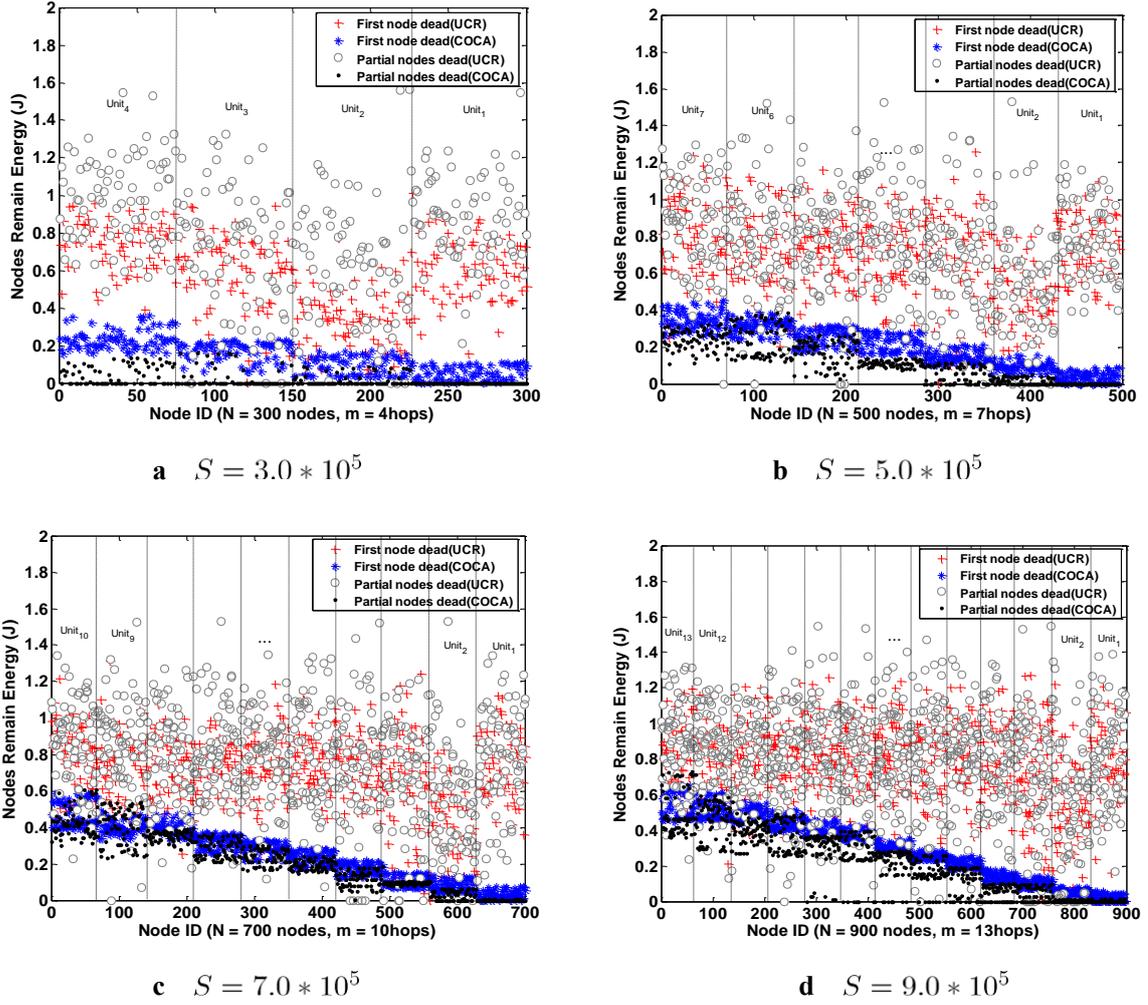


图 26 COCA 和 UCR 在头结点死亡和部分节点死亡情形下的网络剩余能量

#### 4.4.2 数据融合比与网络节点剩余能量

在本小节中，我们主要通过仿真试验验证 COCA 协议在不同网络规模下，不同的数据融合比，在部分节点死亡生命周期定义下所有网络节点的剩余能量，以此得出数据融合对网络能耗均衡和能量利用率的影响。由于空间限制，我们仅仅描绘了两个不同网络规模的节点的剩余能量： $S = 4.0 \times 10^5 m^2$  和  $S = 8.0 \times 10^5 m^2$ ，分别拥有 400 和 800 个网络节点。相应的模拟结果显示在图 27 中。

在图 27 中，我们测试了对于不同的网络规模在不同的数据融合比在部分节点死亡生命周期定义下，网络节点的剩余能量。从图中不难看出，对于特定的网络规模，对于特定的数据融合比，每跳子网内节点的几乎具有相同的剩余能量，而且对于不同的网络规模子网之间的剩余能量水平差值非常小。此外，从图中可以看出，当数据融合比愈小时，网络能耗愈均衡。这说明，我们的协议比较适合于数据冗余度比较高的大规模的网络。但是，随着数据融合比越大，不同子网间的剩余能量差值也越大。这是因为，当全

网采用统一的均匀数据压缩方式时，网络传输的数据量越大，相同的时间内，不同子网间簇头负载的能耗就愈不均匀，而且网络节点剩余能量越大，节点能量利用率越低。这同时也说明，仅仅依靠非均匀建簇、簇头轮转和数据融合的方式，无法避免彻底消除能耗不均，而必须采用其他方式，比如增加新的节点、给网络补给新能量等方式。

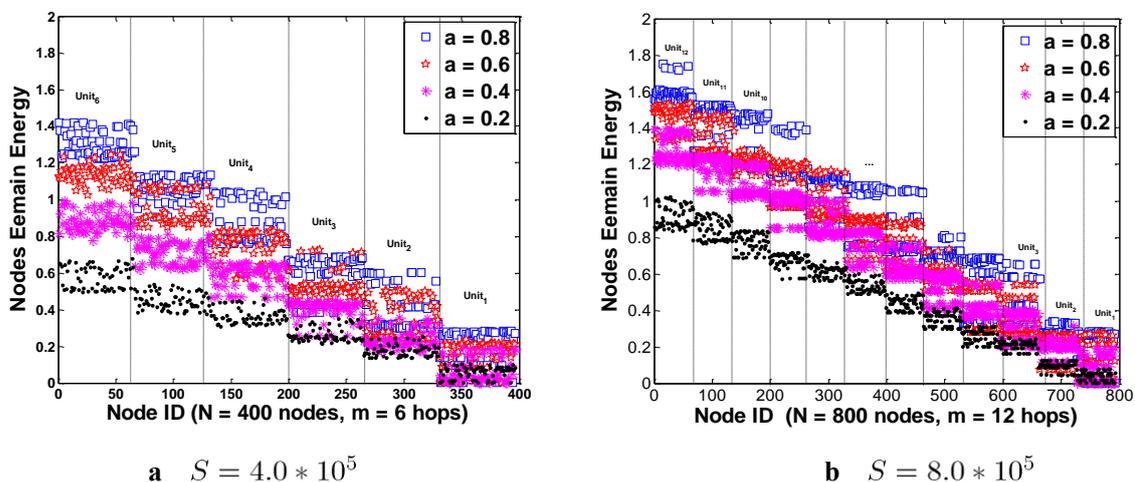


图 27 COCA 在不同数据融合比时网络剩余能量

## 4.5 本章小结

本章对本文设计的无线传感器数据处理与能耗优化理论模型进行了仿真，结合已有的被证明为有效的动态建簇、动态路由以及簇头轮转的策略设计了仿真协议，进一步促进了全网所有节点的能耗均衡，进一步延长网络的生命周期。通过对 COCA 协议与 UCR 进行的仿真结果进行对比和分析，论证了 COCA 协议相对于 UCR 协议，具有良好的稳定性能及实际应用部署能力，能够在不同的网络规模和不同的网络生命周期定义下，能在很大程度上保持网络节点能耗均衡的同时提高网络能量利用效率。同时发现，全网的数据融合比对网络能量利用率有重大影响。

---

## 第五章 模型与协议扩展及实际部署应用

本章首先对第三章无线传感器网络数据处理与能耗优化策略建模分析过程中。所采纳的基于矩形的子网拓扑及基于矩形其它种类的子网拓扑原型方式的合理性进行论证；其次，针对本文提出的无线传感器网络数据处理与能耗优化策略，提出了如何在任意大规模实际应用部署场景下，如何进行扩展以及组织部署的方案。

### 5.1 不同子网拓扑分析

本节针对基于矩形的多跳无线传感器网络，整个网络被划分成各种不同类型的矩形子网拓扑原型的合理性论证。在如图 14 所示的大规模多跳无线传感器矩形拓扑网络，基于矩形的各个子网的变化方式分为以下三种可能的基本拓扑原型。

#### 5.1.1 递增子网模型

对于如图 14 所示的网络，被划分成如下一类的子网区域：被划分成的矩形子区域的面积沿着汇聚节点不断增长，也就是说  $l_{i+1} < l_i, \forall i \in \{m-1, \dots, 1\}$ ，如图 28 所示。根据节点的能量模型，子网内节点之间的通信能量模型采取  $d^2$  的形式能量模型，节点之间的通信距离必须小于阈值  $d_0$  和节点之间的最大传输距离  $t_x$ 。当子网的尺寸沿着汇聚节点方向逐渐增大，增大到一定程度后，尤其是处于子网对角线位置的节点之间的通信距离将不再能够通信，以至于沿着汇聚节点之的子网被分裂开来。因此，这种方式子网增长方式是不可行的，也是不可扩展的。

#### 5.1.2 递减子网模型

对于如图 14 的网络，被划分成如下一类的子网区域：背离着汇聚节点方向的矩形子区域的面积不断增长，也就是说  $l_{i+1} > l_i, \forall i \in \{m-1, \dots, 1\}$ ，如图 29 所示。当背离汇聚节点的子网区域的面积增大到一定尺寸后，节点之间的通信距离必须大于阈值  $d_0$  或节点之间的最大传输距离后，这种方式也是不合理的，原因类似于情形 1。此外，由于离汇聚节点越近，每个子网的尺寸反而越小，意味着这些子网内的节点数越少，子网内的每个簇头能耗负载更大，以致不可能达到能耗均衡。

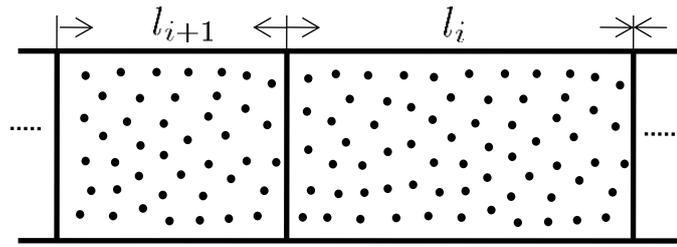


图 28 递增子网模型

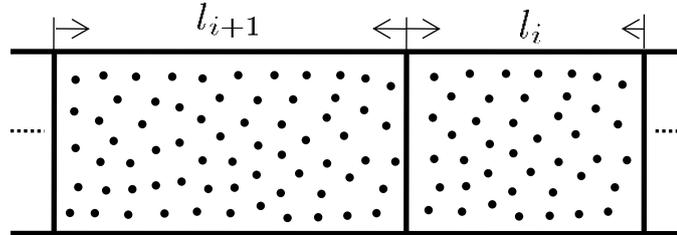


图 29 递减子网模型

### 5.1.3 均匀子网模型

在此类网络原型划分方式中，如图 14 整个网络被划分如下一类网络：子网面积均匀分布： $l_{i+1} = l_i = l, \forall i \in \{m-1, \dots, 1\}$ 。这种类型能进一步分为三种子类型：一种情形是子网区域拥有相同的宽度和高度，如图 30 所示，这种情形也就是我们研究过程中所采用的网络拓扑原型。在这种网络模型中，子网的尺寸必须保证所有的节点的通信距离小于阈值  $d_0$  和节点的最大传输距离  $t_x$ 。对于边长为  $l$  的正方形，节点之间的最大通信距离  $R_{max}$  是  $\sqrt{2}l$ 。因此，如果  $R_{max} < \min\{d_0, t_x\} = d_0$  ( $d_0$  通常远远小于  $t_x$ )，即  $l < \frac{d_0}{\sqrt{2}}$ ，就能保证所有节点的通信距离小于  $d_0$ 。然而，根据前面理论模型的数值分析结果可知，这个子网尺寸下界过于严格，并不适于大规模网络的实际应用部署。

这里，我们找到了一种更加合理的易于实际部署的子网尺寸分布区间。假设  $X, Y$  是记录两个传感器节点位置的随机变量， $d(X, Y)$  是他们之间的欧氏距离。根据文献 [48] - [51] 和附录 A 给出的两个随机、均匀部署在同一单位正方形内的节点之间的欧式距离的概率分布函数可知，两个节点之间的距离小于正方形边长的概率为  $P(d(X, Y) \leq l) = 0.9749259333$ 。也就是说，如果正方形的边长只需要小于阈值  $d_0$ ，它就能在很大程度上保证子网内所有的节点的通信距离均小于阈值  $d_0$ 。而且，由于子网内所有的节点组织成簇，而且每个节点均采用最短距离的方式加入距离自己最近的簇头，实际部署时节点之间的距离将远远小于阈值  $d_0$ 。因此，如果以正方形的边长  $l$  作为子网节点的最大的传输距离，我们得到如下结论：

推论 3: 对于一个大规模的多跳分簇无线传感器网络, 网络被划分成相同的边长为  $l$  的正方形子区域, 子网内簇头与成员节点之间的通信局限于子网内部, 如果正方形子网边长  $l$  小于阈值  $d_0$ , 那么就能在很大概率上保证子网内所有节点与簇头之间的通信距离小于阈值  $d_0$ .

对于其他两种高度和宽度不等的矩形拓扑的情形, 如图 31 和图 32 所示, 可以采取本文类似的方式扩展到这两种类型的拓扑原型。在本文中, 这两种情形的讨论不再给出。

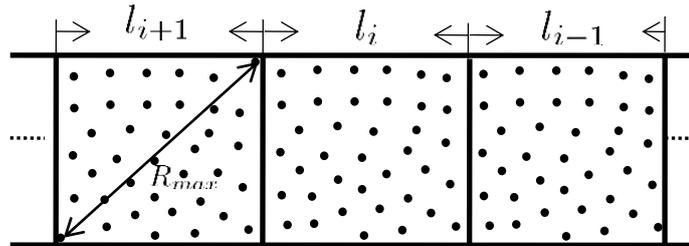


图 30 正方形均匀子网模型

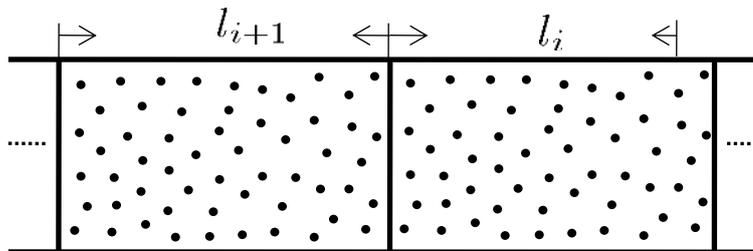


图 31 短高与长宽不等的均匀子网模型

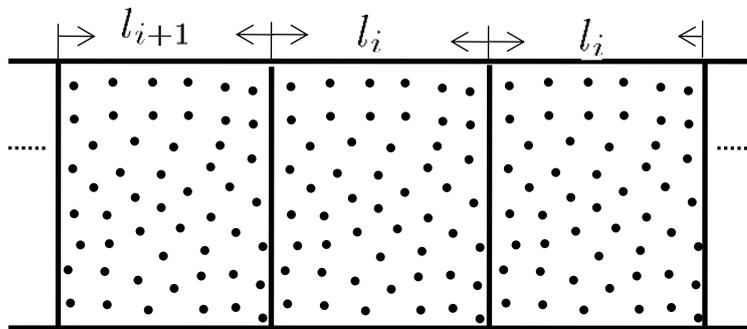


图 32 长高与短宽不等的均匀子网模型

## 5.2 大规模网络实际部署建议方案

对于任意大规模矩形区域的无线传感器网络应用部署场景, 如图 33 所示, 本文所设计的多跳数据, 模型协议依然适用。可以采用如下方式来组织部署: 首先, 将如图所示的大规模的网络合理划分成一定数量的小规模的大网络。每个划分后的网络部署一个汇聚节点; 然后, 这些网络再根据我们设计的模型和协议划分成不同数量的正方形小子网; 最后网络中所有的传感数据由划分后的网络经过汇聚节点收集后最终汇聚到数据中

心。采取这种分层的组织方式，我们的模型和协议能够扩展运用到任意大规模的不同的实际应用场景中。

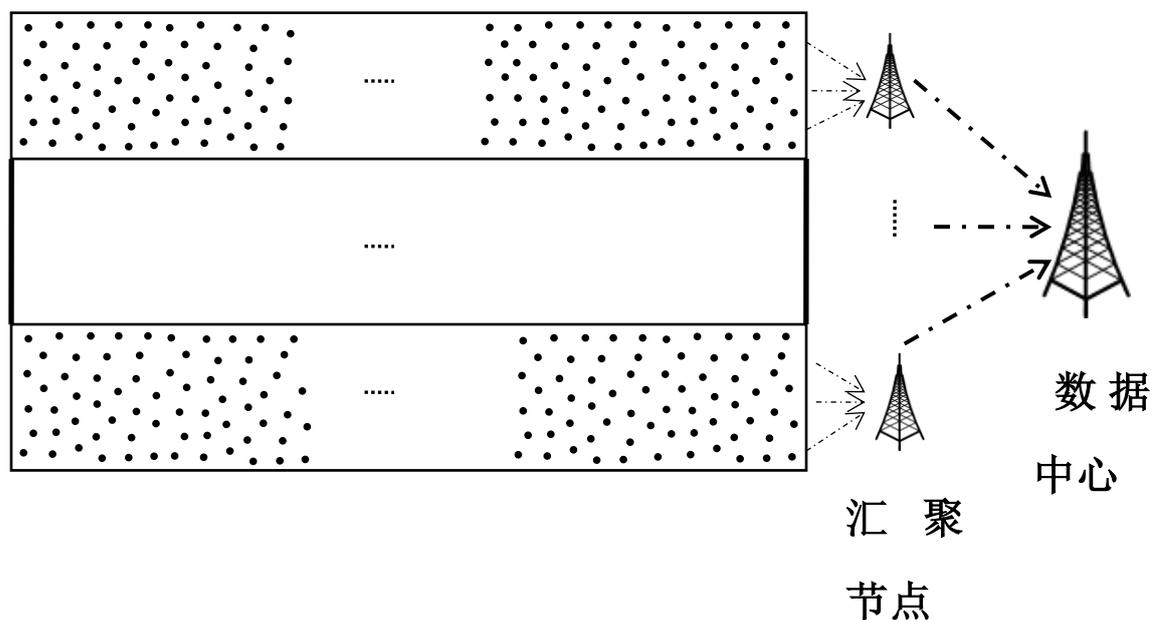


图 33 大规模无线传感器网络实际应用部署场景

### 5.3 本章小结

本章讨论并证明了数据处理及能耗优化策略建模分析过程中涉及基于矩形的各种子网拓扑原型方式的合理性；其次，提出了在任意大规模实际应用部署场景下对本文提出的数据处理及能耗优化策略进行扩展以及组织部署的方案。

---

## 总结与展望

### 本文主要工作总结

本文针对大规模的多跳分簇无线传感器网络,对网络拓扑结构、多跳簇间传输距离、全网的簇头及簇尺寸分布、数据融合、簇头工作时间的长短等影响网络能耗的因素进行了分析,提出有效的数据处理与能耗优化策略,来提高全网所有节点的能量利用效率、延长网络的生命周期。本论文的主要工作分为以下几个方面:

首先,由于簇头是主要的耗能节点,我们对簇头的能耗进行了分析,提高簇头节点的能量利用率,求解出了网络簇头间最优的跳数或子网数以保证全网内所有簇头节点总耗能最小。本文将大规模网络划分成由一系列合理的小规模的正方形子区域或子网,在给定网络面积的前提下,保证各个子网内簇头能耗均衡的同时,求解出划分整个网络的最优的子区域大小以保证整个网络所有簇头的总耗能最小以达到在一轮里面全网内所有簇头总耗能最小,从而延长整个网络的生命周期。

其次,在求得最优的跳数保证全网内所有簇头节点能耗最小的同时,由于离汇聚越近的子区域内的簇头由于平均能耗负载越大耗能越多,为了缓解不同子网内能耗不均,我们量化了在多跳转发过程中不同跳数时、各个子网内分担转发数据的簇头数、求解出了全网不同子网间最优的簇头分布以保证任意两个相邻子网间的簇头在一轮里面耗能均衡、从而延长整个网络的生命周期。找到了最优的子网数及每个子网内簇头数的分布,实际上也就确定了全网不同区域内为了达到全网簇头总耗能最小以及不同子网间簇头能耗均衡的最优的簇尺寸分布规律。

再次,由于在一定的网络周期内,对于簇头数据融合比愈大,需要处理的数据愈多,全网所有簇头节点耗能愈大,我们分析了不同的数据融合比下簇头数量的分布及其对网络能耗的影响。

最后,由于在同构的无线传感器网络中,由于簇头能耗负载远远大于成员节点能耗,从而,簇头工作的时间愈长,簇头能耗愈多大,网络能耗愈不均,我们依据簇头工作时间的长短、节点剩余能量状态以及簇间传输距离等因素,通过结合已有的被证明为有效的簇头轮转、动态建簇的策略设计模型相关的仿真协议来进一步促进了整个网络所有节点的能耗均匀。通过实验仿真得到的结论可知,我们针对大规模多跳分簇无线传感器网络提出的数据处理和能耗优化策略,与之前同类协议相比,显著地提高了网络能量利用效率及网络规模的可扩展性,具有良好的可适性,非常易于实际部署应用。本文最后还

给出了基于矩形的各种子网拓扑原型方式的合理性论证及在任意大规模实际应用部署场景下的扩展和组织部署本文提出的策略的实际方案。

本论文的主要创新点在于针对大规模的无线传感器网络，基于分簇及多跳数据处理策略，综合网络拓扑、多跳簇间距离、全网的簇头及簇尺寸分布、数据融合、簇头轮转等数据处理与能耗优化策略进行建模分析，设计了一个有效的数据处理策略来提高大规模无线传感器网络的能量利用率。同时，根据理论建模的结果结合簇头的工作时间、剩余能量状态以及簇间传输距离等因素，结合已有的被证明为有效的动态建簇、簇间能量感知路由以及簇头轮转等策略来进一步减少和均衡全网的节点能耗、延长网络生命周期。

## 后续工作展望

本文针对大规模无线传感器网络能耗问题设计的数据处理与能耗优化策略，在很大程度上提高了网络能耗利用率，促进了网络的可扩展性，并具备良好的实际部署性能。但是，并没能完全避免网络中的能耗不均问题，网络能量利用率依然有待进一步的改善和提升。具体，可以从以下几个方面着手考虑：

1、簇头节点部署密度与簇头数量分布规律之间的关系：如果给定整个网络的网络规模、部署密度以及数据融合模式和方式，簇头在不同的子区域内均匀分布，如何寻找最优的网络的划分方式和相应的网络部署密度使得不同的子区域的簇头能耗均衡、全网簇头的总耗能最小是后续可行的一个研究方案。

2、数据融合规律和簇头数量分布规律之间的关系：虽然，在大规模的多跳分簇无线传感器网络中，簇头通常将传感数据融合成一个数据包或者按照一定的数据融合比来处理数据，结合本文提出的数据处理和能耗优化策略，提出一个性能更加优异的数据融合模型，是我们接下来的一个研究方向。

3、结构化分簇网络的异构节点、投入成本与网络能耗之间的关系：如果给定整个网络的网络规模、网络的划分方式和部署密度、子网的数量以及数据融合模式和方式，如果簇头节点与成员节点是异质的，即簇头节点具有较成员节点更高的储备能量的处理能力、经济花费更高，且簇头固定，即不采用轮转策略来重新选取簇头，可以在寻找相应的簇头数量分布规律使得不同子区域内的簇头能耗均衡和最小而且簇头的经济代价最小。

## 参考文献

- [1] Cyber-physical system[EB/OL], <http://www.nsf.gov/pubs/2008/nsf08611/nsf08611.htm>.
- [2] Internet of Things[EB/OL], <http://www.internetofthings.net.cn/html/index.php/Index/d/id/35>.
- [3] Werner-Allen G., Swieskowski P., Welsh M.: Real-time volcanic earthquake localization[C].SenSys, 2006, 357-358.
- [4] Tiwari A., Ballal P., and Lewis F. L.. Energy-efficient wireless sensor network design and implementation for conditionbased maintenance[J]. ACM Transactions on Sensor Networks (TOSN), 2007,V3(1).
- [5] Akyildiz I., Su W., Sankarasubramaniam Y., and Cayirci E.. A Survey on Sensor Networks [J]. IEEE Communication Magazine,2002,V40(8):102-114.
- [6] 传感器网络标准工作组[EB/OL], <http://www.wgsn.org/getIndex.req?action=quary&req=modulenvpromote&id=635&type=0&moduleId=680&sid=51>.
- [7] 传感器行业协会[EB/OL], <http://www.chinasensor.com.cn/gl/inf/list.asp?id=238>.
- [8] 崔莉. 传感器网络研究概况简介[J]. 信息技术快报,2008,6(3):1-4.
- [9] Kimura N. and Latifi S..A Survey on Data Compression in Wireless Sensor Networks[C]. Proceedings of the International Conference on Information Technology: Coding and Computing, (ITCC),2005: 8-13.
- [10] Akkaya K. and Younis M..A survey on routing protocols for wireless sensor networks[J].Ad Hoc Networks,2005,V3(3):325-349.
- [11] Olariu S. and Stojmenovic I..Design Guidelines for Maximizing Lifetime and Avoiding Energy Holes in Sensor Networks with Uniform Distribution and Uniform Reporting[C].Proc. of the 25<sup>th</sup> IEEE International Conference on Computer Communication (INFOCOM), 2006:1-12.
- [12] Li J. and Mohapatra P..Analytical Modeling and Mitigation Techniques for the Energy Hole Problem in Sensor Networks[J].Pervasive and Mobile Computing, 2007,V3(3):233-254.
- [13] Chen G., Li C., Ye M. and Wu J.. An Unequal Cluster-Based Routing Strategy in Wireless Sensor Networks [J]. Wireless Networks, 2009,V15(2): 193-207.
- [14] Li H., Cao J., and Xiong J.W..Constructing Optimal Clustering Architecture for Maximizing Lifetime in Large-scale Wireless Sensor Networks [C]. Proc. of the 15th Int'l Conf. on Parallel and Distributed Systems (ICPADS), 2009: 182-189.
- [15] Wu X., Chen G., and Das S. K..Avoiding Energy Holes in Wireless Sensor Networks with Nonuniform Node Distribution[J].IEEE Transactions on Parallel and Distributed Systems, 2008,V19(5):710-720.
- [16] Gupta G. and Younis M..Load-balanced clustering of wireless sensor networks[C].IEEE International Conference on Communications, 2003:1848-1852.
- [17] Gupta G. and Younis M..Performance evaluation of load-balanced clustering of wireless sensor networks[C].Proceedings of IEEE International conference on communications, 2003:1577-1581.

- 
- [18] Efthymiou C., Nikiletsas S., and Rolim J..Energy Balanced Data Propagation in Wireless Sensor Networks[J].Wireless Networks,2006,V12(6): 691-707.
- [19] Lenone P., Nikiletsas S., and Rolim J..An Adaptive Blind Algorithm for Energy Balanced Data Propagation in Wireless Sensors Networks[J].Distributed Computing in SensorSystems,2005, V3560: 35-48.
- [20] Zhang H., Shen H., and Tan Y..Optimal Energy Balanced Data Gathering in Wireless Sensor Networks[C].Proc. of the 21st IEEE Int'l Parallel and Distributed Processing Symposium(IPDPS), 2007:1-10.
- [21] Lian J., Nail K., and Agnew G..Data Capacity Improvement of Wireless Sensor Networks Using Non-Uniform Sensor Distribution[J]. Int'l Journal of Distributed Sensor Networks,2005,V2: 121-145.
- [22] Mhatre V. and Rosenberg C..Design guidelines for wireless sensor networks: communication, clustering and aggregation[J].Ad Hoc Networks, 2004,V2(1):45-63.
- [23] Abbasi A. A. and Younis M..A Survey on Clustering Algorithms for Wireless Sensor Networks[J]. Computer Communications, 2007,V30(14-15): 2826-2841.
- [24] Rajagopalan R. and Varshney P. K..Data Aggregation Techniques in Sensor Networks:A survey[J]. Comm. Surveys & Tutorials,2006,V8:48-63.
- [25] Krishnamachari B., Estrin D. and Wicker S..The Impact of Data Aggregation in Wireless Sensor Networks[C].Proc. of the 22nd Intl Conf. on Distributed Computing Systems Workshops (ICDCSW), 2002:575-578.
- [26] Yoon S. and Shahabi C..The clustered aggregation (cag) technique leveraging spatial and temporal correlations in wireless sensor networks[J].ACM Transactions on Sensor Networks(TOSN), 2007,V3(1):3-es.
- [27] Baek S. J., Veciana G. D., and Su X..Minimizing energy consumption in large-scale sensor networks through distributed data compression and hierarchical aggregation[J]. IEEE Journal on Selected Areas in Communications, 2004,V22(6):1130-1140.
- [28] Heinzelman W., Kulik J., and Balakrishnan H..Adaptive Protocols for Information Dissemination in Wireless Sensor Networks[C].Proc. 5th ACM/IEEE Mobicom Conference, Seattle, WA,1999: 174-185.
- [29] Kulik J., Heinzelman W.R. and Balakrishnan H..Negotiation-based Protocols for Disseminating Information in Wireless Sensor Networks[J].Wireless Networks, 2002,V8(2-3):169-185.
- [30] Intanagonwiwat C., Govindan R. and Estrin D..Directed Diffusion: A Scalable and Robust Communication Paradigm for Sensor Networks[C]. Proceedings of the 6th annual international conference on Mobile computing and networking(MobiCOM), 2000:56-67.
- [31] Heinzelman W. , Chandrakasan A., and Balakrishnan H..Energy-Efficient Communication Protocol for Wireless Microsensor Networks[C].Proceedings of the 33rd Hawaii International Conference on System Sciences,2000: 3005-3014.
- [32] Heinzelman W., Chandrakasan A. and Balakrishnan H..An Application-Specific Protocol

- Architecture for Micro-Sensor Networks[J].IEEE Transactions on Wireless Communications,2002, V1(4):660-670.
- [33] Manjeshwar A. and Agrawal P..TEEN: A Protocol for Enhanced Efficiency in Wireless Sensor Networks[C].Proc. of the 15th Parallel and Distributed Processing Symposium,2001:2009-2015.
- [34] Soro S. and Heinzelman W..Prolonging the Lifetime of Wireless Sensor Networks via Unequal Clustering[C].Proceedings of the 19th IEEE International Parallel and Distributed Processing Symposium (IPDPS),2005.
- [35] Younis O. and Fahmy S.. HEED: A Hybrid, Energy-Efficient, Distributed Clustering Approach for Ad-hoc Sensor Networks[J]. IEEE Transaction on Mobile Computing, 2004,V3(4): 660-669.
- [36] Lindsey S., Raghavendra C., and Sivalingam K.M..Data Gathering Algorithms in Sensor Networks Using Energy Metrics[J].IEEE Transactions on Parallel and Distributed Systems, 2002,V13(9): 924-935.
- [37] Tan H. O. and Korpeoglu I..Power Efficient Data Gathering and Aggregation in Wireless sensor Networks[J].SIGMOD Record, 2003,V32(4):66-71.
- [38] Haiyun Luo, Fan Ye, Jerry Cheng, et al.TTDD: Two-Tier Data Dissemination in Large-Scale Wireless Sensor Networks[J]. Wireless Networks, 2005,V11(1-2): 161-175.
- [39] Zhang Z., Ma M., and Yang Y..Energy-Efficient Multi-Hop Polling in Clusters of Two-Layered Heterogeneous Sensor Networks[J]. IEEE Transactions on Computers, 2008,V57(2):231-245.
- [40] Fasolo E., Rossi M., Widmer J., et al.In-Network Aggregation Techniques for Wireless Sensor Networks: A Survey[J].IEEE Wireless Communication, 2007,V14(2):784-798.
- [41] The Network Simulator, Version 2 [EB/OL]. <http://www.isi.edu/nsnam/ns/>.
- [42] Ghosh B..Random distances within a rectangle and between two rectangles[J].Bulletin of the Calcutta Mathematical Society,1951,V43:17-24.
- [43] Daley D. J..Problems in applied mathematics: selections from SIAM review[J/OL].Society for Industrial Mathematics, 1990:77.
- [44] Vaughan R. J..Problems in applied mathematics: selections from SIAM review[J/OL].Society for Industrial Mathematics, 1990:78-79.
- [45] Alagar V. S..The Distribution of the Distance between Random Points[J]. Journal of Applied Probability, 1976:V13(6):558-566.
- [46] Oser H. J..Problems in applied mathematics: selections from SIAM review[J/OL]. Philadelphia, PA: Society for Industrial Mathematics, 1990:76.
- [47] Younis O.,Krunz M.and Ramasubramanian S..Node Clustering in Wireless Sensor Networks:Recent Developments and Deployment Challenges[J].IEEE Network,2006,V20(3):20–25.
- [48] Gaboun B. and Laporte G. and Soumis F..Expected Distances between Two Uniformly Distributed Random Points in Rectangles and Rectangular Parallepipeds[J].The Journal of the Operational Research Society,1993,V44(5):513–519.
- [49] Lazoff D., Lazoff D. M. and Sherman A. T..An Exact Formula for the Expected Wire Length Between Two Randomly Chosen Terminals[R].University of Maryland Baltimore County report

- CS-94-08, 1994.
- [50] Mathai A., Moschopoulos P. and Pederzoli G.. Random Points Associated with Rectangles[J]. Rendiconti del Circolo Matematico di Palermo, 1999, V48(1):163-190.
- [51] Miller L. E.. Distribution of Link Distances in a Wireless Network[J]. Journal of Research of the National Institute of Standards and Technology, 2001, V106(2):401-412.

## 附录 A

Ghosh<sup>[42]</sup> 似乎最先给出了随机、均匀部署在两个相邻单位正方形（就像棋盘上的白子和黑子一样）内的两点之间的欧式距离期望值，而且该值为 1.088<sup>[44]</sup><sup>[43]</sup><sup>[44]</sup>。在 Ghosh<sup>[42]</sup> 中所求得的该距离  $R$  的密度分布函数为：

$$f(R) = \frac{2R\phi(R)}{a^4}$$

这里，

$$\phi(R) = \begin{cases} aR - \frac{1}{2}R^2 & \text{for } 0 \leq R \leq a \\ 2a^2 \cos^{-1}\left(\frac{a}{R}\right) - 2a\sqrt{R^2 - a^2} - 2aR + R^2 + \frac{3}{2}a^2 & \text{for } a \leq R \leq 2\sqrt{a} \\ 2a^2 \sin^{-1}\left(\frac{a}{R}\right) + 2a\sqrt{R^2 - a^2} - 2aR - \frac{1}{2}a^2 & \text{for } \sqrt{2a} \leq R \leq 2a \\ 2a^2[\sin^{-1}\left(\frac{a}{R}\right) - \cos^{-1}\left(\frac{2a}{R}\right)] + 2a\sqrt{R^2 - a^2} + a\sqrt{R^2 - 4a^2} & \text{for } 2a \leq R \leq \sqrt{5a} \\ -\frac{1}{2}R^2 - \frac{5}{2}a^2 & \text{for } 2a \leq R \leq \sqrt{5a} \end{cases}$$

那么，所求距离期望的精确值即为随机变量  $R$  关于原点的一阶矩，即对  $Rf(R)$  在其整个定义域内的积分值，其近似估计值为 1.088 $a$ 。详细的推导过程请参照 Ghosh<sup>[42]</sup> 附录。

为求两个分别部署在相邻的两个单位正方形的点之间的欧氏距离的期望值，Oser H.J.<sup>[44]</sup> 通过定义了以下一个四重积分来对其进行精确计算：

$$F = \int_0^1 \int_{-1}^0 \int_{-\frac{1}{2}}^{\frac{1}{2}} \int_{-\frac{1}{2}}^{\frac{1}{2}} \{(x_1 - x_2)^2 + (y_1 - y_2)^2\}^{\frac{1}{2}} dx_1 dx_2 dy_1 dy_2$$

Daley D.J.<sup>[43]</sup> 用以下方法给出了对上面的这个四重积分的求解方法。定义  $D(1, a)$  为分别随机均匀分布在长与宽分别为 1 和  $a$  的矩形内的两点之间的距离的期望值，表示如下，所求距离期望的精确值即一阶矩，

$$\begin{aligned} D(1, a) &= 4a^{-2} \int_0^1 (1-v) dv \int_0^a (a-u)(u^2 + v^2)^{\frac{1}{2}} du \\ &= \frac{1 + a^5 + [a^2 - (a^2 - 1)^2(a^2 + 1)^{\frac{1}{2}}]}{15a^2} + \frac{a^3 \log\left[\frac{1+(a^2+1)^{\frac{1}{2}}}{a}\right] + \log[a + r(a^2 + 1)^{\frac{1}{2}}]}{6a} \end{aligned}$$

这样就可以通过如下表达式把求  $F$  这个四重积分的值转化为求  $D(1, a)$  这个更加通用和简洁的二重积分表示形式的值：

$$D(1, 2) = \frac{1}{2}[D(1, 1) + F]$$

将  $a = 1$  和  $a = 2$  分别代入上述两个表达式，即可求得  $F$  的精确值和近似值：

$$F = \frac{29 - 2\sqrt{2} - 5\sqrt{5}}{30} + \frac{8\log(\frac{1+\sqrt{5}}{2}) + \log(2 + \sqrt{5}) - 2\log(1 + \sqrt{2})}{6}$$

$$\approx 1.08814$$

Vaughan R. J.<sup>[44]</sup> 也通过首先建立概率分布函数的方式给出两个分别部署在相邻的两个单位正方形的点之间的欧氏距离的期望值的估计值为  $\frac{4}{3} * \frac{1}{1.213} \approx 1.10$

基于 Crofton 的期望理论, Alagar V.S.<sup>[45]</sup> 给出了两个分别部署在相邻的两个单位正方形的点之间的欧氏距离的概率分布函数如下:

$$F(1) = \begin{cases} 2\frac{x^3}{3} - \frac{x^4}{4} & \text{for } 0 \leq x \leq 1 \\ -\frac{1}{4} + 2\frac{x^2}{2} - 4\frac{x^3}{3} + \frac{x^4}{2} - 2\frac{(1+2x^2)(x^2-1)^{\frac{1}{2}}}{3} + 2x^2\arccos(\frac{1}{x}) & \text{for } 1 \leq x \leq \sqrt{2} \\ -\frac{11}{12} - \frac{x^2}{2} - 4\frac{x^3}{3} + 2\frac{(1+2x^2)(x^2-1)^{\frac{1}{2}}}{3} + 2x^2\arcsin(\frac{1}{x}) & \text{for } \sqrt{2} \leq x \leq 2 \\ -\frac{75}{12} - 9\frac{x^2}{2} + 5\frac{x^4}{12} + 2\frac{(1+2x^2)(x^2-1)^{\frac{1}{2}}}{3} + 2\frac{(2+2x^2)(x^2-4)^{\frac{1}{2}}}{3} + 2x^2\arcsin(\frac{1}{x}) \\ + 2x^2\arccos(\frac{2}{x})[-1 + \frac{(5-x^2)}{(x^2-4)^{\frac{1}{2}}}] & \text{for } 2 \leq x \leq \sqrt{5} \\ 1 & \text{for } x \geq \sqrt{5} \end{cases}$$

这里,  $F(t) = P\{d \leq x | t \geq 0\}$ ,  $d$  是独立随机分布在边长为  $t$  和 1 的相邻的两个矩形的两点之间的随机距离。那么, 独立随机分布在两个相邻的边长分别为  $t$  和 1 的两点之间的距离期望的精确值即关于原点的一阶矩为如下表达式:

$$t^2 F(t) = \frac{1}{30} \{30 - t^5 + 2(t^2 + 1)^{\frac{1}{2}}(t^4 - 3t^2 + 1) - (t^2 + 4)^{\frac{1}{2}}(t^4 - 12t^2 + 16) \\ - 40t\log(2) - 5t\log[t + (t^2 + 1)^{\frac{1}{2}}] - 5t^4\log[1 + (t^2 + 1)^{\frac{1}{2}}] \\ 40t\log[t + (t^2 + 4)^{\frac{1}{2}}] + 5t^4\log[2 + (t^2 + 1)^{\frac{1}{2}}]\}$$

因此可得两个分别均匀部署在相邻的两个单位正方形的点之间的欧氏距离的期望值为:

$$\frac{29 - 2\sqrt{2} - 5\sqrt{5} + 5\log((2 + \sqrt{5})(-1 + \sqrt{2})^2) + 40\log((1 + \sqrt{5})/2)}{30} \approx 1.08814$$

因此, 综上所述, 依据 Ghosh<sup>[42]</sup>, Daley D.J.<sup>[43]</sup>, Vaughan R. J.<sup>[44]</sup> 和 Alagar V.S.<sup>[45]</sup> 等人的解法, 我们可以求出该欧氏距离。

## 附录 B

Ghosh在文献 [42] 中对在如何求随机、均匀部署在“一个矩形”和分别部署在“两个相邻矩形”的两点之间的欧氏距离的期望值这类问题进行了严格的数学定义，并对此类问题提供了一种通用的一般解法。利用这种通用的解法，可以得到一些特殊的实例的解，如随机、均匀部署在一个单位正方形内和分别部署在相同的相邻正方形内的两点之间的欧氏距离的期望值。Ghosh<sup>[42]</sup>似乎最先给出了随机、均匀部署在一个单位正方形内的两点之间的距离期望值。Ghosh<sup>[42]</sup>给出了随机、均匀在一个长与宽分别为  $a$  和  $b$  的矩形内的两点欧式距离  $R$  的密度分布函数：

$$f(R) = \frac{4R\phi(R)}{a^2b^2} \quad (\text{A1})$$

这里

$$\phi(R) = \frac{1}{2}\pi ab - aR - bR + \frac{1}{2}R^2, 0 \leq R \leq b$$

显然，对这个一般分布密度函数取  $a = b$  便得到了正方形情形下的特例。那么，两点之间的欧氏距离的期望值或关于原点的一阶矩  $\alpha_1$  就是对  $\{Rf(R)\}$  在  $R$  的整个值域上进行积分。详细的推导过程请参照文献 [42] 附录。

Gaboun B.<sup>[48]</sup>等人给出了随机、均匀分布在一个长与宽分别为  $a$  和  $b$  的矩形内的两点欧式距离  $L$  的期望值：

$$E(L) = \frac{1}{6} \left[ \frac{a^2}{b} \cosh^{-1} \left( \frac{(a^2 + b^2)^{\frac{1}{2}}}{a} \right) + \frac{b^2}{a} \cosh^{-1} \left( \frac{(a^2 + b^2)^{\frac{1}{2}}}{b} \right) \right] + \frac{1}{15} \left( \frac{b^3}{a^2} + \frac{a^3}{b^2} \right) - \frac{(a^2 + b^2)^{\frac{1}{2}}}{15} \left( \frac{b^2}{a^2} + \frac{a^2}{b^2} - 3 \right) \quad (\text{A2})$$

Lazoff D.<sup>[49]</sup>等人给出了随机、均匀分布在长与宽分别为  $a$  和  $b$  的任意矩形内的两点欧式距离  $D$  的期望值的计算公式：

$$\frac{a^5 + b^5 - (a^4 - 3a^2b^2 + b^4)\sqrt{a^2 + b^2}}{15a^2b^2} + \frac{a^2}{6b} \ln \left( \frac{b + \sqrt{a^2 + b^2}}{a} \right) + \frac{b^2}{6a} \ln \left( \frac{a + \sqrt{a^2 + b^2}}{b} \right) \quad (\text{A3})$$

在公式(A4)中，通过替换  $b = a$  便可得到随机、均匀部署在一个单位正方形内的两点之间的距离期望值。

Mathai A.<sup>[50]</sup>等人对如何求解随机、均匀部署在一个单位正方形内的两点之间的距离期望值的密度分布函数进行了调研。给出了随机、均匀分布在长与宽分别为  $a$  和  $b$  的任意矩形内的两点欧式距离为  $x$  的密度函数：

$$f(x) = \frac{4x}{a^2b^2}\phi(x) \quad (\text{A4})$$

这里,

$$\phi(x) = \begin{cases} \frac{ab\pi}{2} - (a+b)x + \frac{x^2}{2}, & \text{for } 0 \leq x < a \\ ab\sin^{-1}\left(\frac{a}{x}\right) - \frac{a^2}{2} - b(x^2 - a^2)^{\frac{1}{2}}, & \text{for } a \leq x < b \\ ab\left[\sin^{-1}\left(\frac{a}{x}\right) - \sin^{-1}\sqrt{1 - \frac{b^2}{x^2}}\right] - \frac{a^2+b^2}{2} - \frac{x^2}{2} \\ \quad + a\sqrt{x^2 - b^2} + b\sqrt{x^2 - a^2}, & \text{for } b \leq x \leq \sqrt{a^2 + b^2} \\ 0 & \text{elsewhere} \end{cases}$$

而且欧氏距离的期望值表达式为

$$E(x) = \frac{1}{15}\left[\frac{a^3}{b^2} + \frac{b^3}{a^2} + \frac{5}{2}\left(\frac{a^2}{b}\ln\left(\frac{a + (a^2 + b^2)^{\frac{1}{2}}}{b}\right) + \frac{b^2}{a}\ln\left(\frac{b + (a^2 + b^2)^{\frac{1}{2}}}{a}\right)\right) + (a^2 + b^2)^{\frac{1}{2}}\left(3 - \frac{b^2}{a^2} - \frac{a^2}{b^2}\right)\right] \quad (\text{A5})$$

Miller L. E.<sup>[51]</sup>给出了随机、均匀分布在单位正方形内的两点欧式距离的累积分布函数:

$$F_d(\gamma = \xi D) = \begin{cases} 0 & \text{for } \xi \leq 0 \\ \xi^2\left(\frac{1}{2}\xi^2 - \frac{8}{3}\xi + \pi\right) & \text{for } 0 \leq \xi < 1 \\ \frac{4}{3}\sqrt{\xi^2 - 1}(2\xi^2 + 1) - \left(\frac{1}{2}\xi^4 + 2\xi^2 - \frac{1}{2}\right) \\ \quad + 2\xi^2\left[\sin^{-1}\left(\frac{1}{\xi}\right) - \cos^{-1}\left(\frac{1}{\xi}\right)\right] & \text{for } 1 \leq \xi < \sqrt{2} \\ 1 & \text{for } \sqrt{2} < \xi \end{cases} \quad (\text{A6})$$

这里,  $D$  是正方形的边长 (同本文的正方形子网区域边长  $l$  含义相同),  $\xi$  是定义为  $\xi \triangleq \frac{\gamma}{D}$  的一个标准化变量,  $\gamma$  是所求欧式距离。

因此, 基于Ghosh<sup>[42]</sup>, Gaboun B.<sup>[48]</sup>, Lazoff D.<sup>[49]</sup>和Mathai A.<sup>[50]</sup>等学者提出的方法, 均可以得到两个随机、均匀分布在单位正方形内的两点欧式距离的密度函数。特别, 如果假设  $X, Y$  表示两个随机、均匀分布在边长为  $l$  正方形内的两点位置的随机变量,  $d(X, Y)$  表示它们的欧氏距离, 那么可以依据公式(A1) (取  $a = b$ ), (A4) (取  $a = b$ ) 或 (A6) (取  $\xi = 1$ ) 中的任意一个分布函数推导求解得出, 这两点的欧氏距离小于该正方形边长的概率为:

$$P(d(X, Y) \leq l) = 0.9749259333.$$

例如, 考虑公式(A6), 取  $\xi = 1$ , 两点的欧氏距离小于该正方形边长的概率为:

$$P(d(X, Y) \leq D) = F_d(\gamma = D) = F_d(\gamma = l) = \left(\frac{1}{2} - \frac{8}{3} + \pi\right) = 0.9749259333.$$

这里的  $D$  为正方形的边长, 与本文的正方形子网区域边长  $l$  含义相同。

## 攻读硕士学位期间取得的学术成果

### 硕士期间发表的论文

[1] Li H., Cao J., and **Xiong J.W.** Constructing Optimal Clustering Architecture for Maximizing Lifetime in Large-scale Wireless Sensor Networks [C]. Proc. of the 15<sup>th</sup> Int'l Conf. on Parallel and Distributed Systems (ICPADS'09), on Pages: 182-189, Dec. 8–11, 2009, Shenzhen, China(**EI** 索引)

### 正在修订论文

Li Huan, Xiong Junwu, and Jia Weijia, "COCA: Constructing Optimal Unequal Clustering Architecture for Maximizing Lifetime in Large-Scale Wireless Sensor Networks", (working in progress)

## 致 谢

首先,我对我的导师李欢副教授表示最深挚的感谢。自从 2008 年 3 月被北航录取成为您的学生、到您推荐我到香港城大深圳研究院工作、并在整个研究生阶段两年多来由您亲自悉心的指导,是我研究生阶段最大的幸运。您对学术研究的巨大热情和执着、极其严谨而开放的治学作风、遇到任何困难挫折绝不妥协止步和勇于创新的态度,在我最困难的时候给了我最大的精神支柱和坚定不移的信念,让学生领略到了一个真正的学者应有的风范和品格,是值得学生一辈子效仿的楷模。您对我学术上极其耐心细致的教导和开诚布公的讨论与交流,平时学习工作处事作风的严格要求以及生活上热心的关照帮助,不仅使得我的学位论文得以顺利开展和完成、我的整个研究生阶段变得更加紧凑、充实和完整,更使得我尽力去发掘一个更加刚健完善的自我,不断积累更多的勇气与智慧去开拓以后的人生。李老师,感谢您这两年多来对我所有的谆谆教诲和为我所付的全部心血,学生永远铭刻于心!祝福您身体健康,家庭幸福,硕果累累、桃李满天下!

特别感谢曹杰瑞师兄、邢航、李延凯和梁栋同学。你们对我整个研究生研究过程耐心热情的帮助,给我留下了非常深刻的印象,每次和你们的研究讨论过程中都大大地开阔了我的思路。感谢你们在我论文的选题、实现和论文撰写的过程中给我的帮助与支持。感谢王芳、孙哲峰、谢丽慧、刘彦磊同学,能与你们一起共同讨论问题、朝夕相处共渡的一点一滴,给我硕士期间留下最美好的回忆。祝福你们学业有成。

衷心感谢网格组的郑海滨、周德辉、李岑、蒋昌绘、高崎、衣敏、侯毅和刘超同学,有幸和你们共渡研究生生活,非常快乐和幸运。感谢钟亮博士和崔文超师兄、耿剑平师兄在我研究生学习和生活中给予的指点和帮助,祝你们身体健康,事业有成。感谢 ACT 实验室的所有兄弟姐妹,祝大家学业有成。

衷心感谢 ACT 实验室的怀进鹏老师、刘旭东老师、韩军老师、李先贤、胡春明老师,感谢你们忘我的工作给全体 ACT 学生带来了如此良好的学术科研环境,让我度过了充实的研究生阶段。特别感谢网络计算组李建欣老师、沃天宇老师,在我的研究生阶段,给予了积极的指导和帮助。

感谢各位评审老师在百忙之中抽出时间评阅我的论文。

最后,谨以此文献给我最深爱的父母。在我整个求学生涯中,你们默默为我付出无尽的血汗,儿子永铭于心,祝愿爸爸妈妈身体永远健康快乐!感谢你们的支持与爱,我将自强不息,勇往直前!