"十一五"国家重点图书·军事信息技术学术丛书

无线传感器网络中
高效传输技术

彭绍亮 彭宇行 李姗姗 等编著

国防科技大学出版社

国家重点基础项目研究资助
国防科技大学学术专著专项经费

“十一五”国家重点图书·军事信息技术学术丛书

无线传感器网络中
高效传输技术

彭绍亮　彭宇行　李姗姗
陈松政　车永刚　朱培栋　编著
鄢来斌

国防科技大学出版社
湖南·长沙

内容简介

无线传感器网络中的信息传输是用户分发控制信息和获取感知数据的根本途径，而高效的信息传输技术则是一种提高网络能源使用效率的重要手段，其目的是在源节点(Source)和汇聚节点(Sink)之间建立信息传输的双向通道，用尽可能少的传输开销和延迟在两者间实现信息交互。然而，传感器节点资源受限、网络带宽受限、网络规模大等问题，给高效传输的研究带来了巨大的挑战。本书针对高效传输技术在信息分发和收集时的通信开销、实时性、可扩展性等问题，对多播、稀疏采样、簇内聚合、非结构化的存储和查询等关键问题提出了有效的解决方案，希望对于推进无线传感器网络中高效传输技术的研究和实用化能有一定的理论意义和应用价值。

图书在版编目(CIP)数据

无线传感器网络中高效传输技术/彭绍亮，彭宇行，李姗姗等编著.—长沙：国防科技大学出版社，2010.2
(军事信息技术学术丛书)
ISBN 978-7-81099-621-1

Ⅰ.无… Ⅱ.①彭… ②彭… ③李… Ⅲ.无线电通信—传感器—数据传输 Ⅳ.TP212

中国版本图书馆CIP数据核字(2009)第210402号

国防科技大学出版社出版发行
电话:(0731)84572640 邮政编码:410073
http://www.gfkdcbs.com
责任编辑:黄 煌 责任校对:王 嘉
新华书店总店北京发行所经销
国防科技大学印刷厂印装
*
开本:850×1168 1/32 印张:6.5 字数:169千
2010年2月第1版第1次印刷 印数:1-500册
ISBN 978-7-81099-621-1
定价:25.00元

前　言

无线传感器网络集信息获取、处理和传输为一体，在工业、农业、军事、环境监控、生物医疗、城市管理和抢险救灾等领域有着非常广泛的应用前景。传感器网络中信息传输是用户分发控制信息和获取感知数据的根本途径，而高效的信息传输技术则是一种提高网络能源使用效率的重要手段，其目的是在源节点（Source）和汇聚节点（Sink）之间建立信息传输的双向通道，用尽可能少的传输开销和延迟在两者间实现信息交互。然而，传感器节点资源受限、网络带宽受限、网络规模大等问题，给高效传输的研究带来了巨大的挑战。

无线传感器网络中的高效传输技术主要包含以下两个过程：Sink 到 Source 的高效控制信息分发；Source 到 Sink 的高效数据收集。现有的工作主要停留在使用一些诸如单播、广播、聚合之类的传统技术方法，这些方法普遍存在通信开销大、实时性差、可扩展性不强等特点。本书针对无线传感器网络的诸多特点和现有研究的不足，以最小化总的传输开销为目标，从控制信息分发和数据收集两个方向入手，系统地研究了高效传输技术涉及的四个关键性问题：（1）Sink 到 Source 的多播问题；（2）Sink 到 Source 的采样问题；（3）Source 到 Sink 的紧急报文聚合问题；（4）Sink 与 Source 之间的双向协同问题，即数据存储和查询问题。

本书首先分析了由Sink向Source进行控制信息分发的多播技术，并针对现有对等通信模式下多播研究通信开销大、效率不高、可扩展性不强等问题，提出了一种基于基站模式的新型多播协议SenCast，减少了控制信息分发的传输开销，提高了信息分发的实时性。SenCast首先利用了基站较强的计算和存储能力，用组件化的思想计算一棵全局近似最优多播树，并基于该树进行信息分发，最大限度地减少了多播传输开销。然后，证明了SenCast在解决MNN（最小非叶节点树）Steiner树这一NP难问题时，获得了小于 $\ln|R|$ 的近似率（R 为目标节点数），该近似率也是MNN问题理论上的最优近似率。最后，还针对路由时路径长、分支多的特点，设计了SRL和HLB两种可扩展性机制，使得SenCast可以解决大规模多播问题。实验结果表明，SenCast适用于大规模群组通信，是一种能量有效的、可扩展性强的新型多播协议。

针对现有方法在获得网络全局信息时冗余传输太多，从而导致网络带宽、节点资源浪费、传输延迟长的问题，本书提出了一种基于多播的“释放/捕获”采样技术FLAKE，有效减少了获取全局信息时的通信开销，提高了实时性。FLAKE是一种快速稀疏采样技术，其基本思想是：首先在网内均匀、快速地“释放” m 条消息种子，然后对消息种子进行采样，如果 m' 个消息种子被 n' 个节点收到，则目前网内的工作节点总数 n 估算为 mn'/m'。更重要的是，本书基于逆采样理论开展研究，有效解决了估算精度和采样尺寸之间的折中问题，即在保证一定误差精度的情况下，给出了最优的消息种子数 m 和逆采样的数量 m'。FLAKE还

包含了一个分布式的稀疏采样算法，不需要进行任何中央式的控制，就可以在一定精度范围内，以很少的开销和延迟估算出网络的全局状态信息。理论分析和实验模拟验证了 FLAKE 在通信开销、实时性、可扩展性等方面都获得了较好的性能。

大规模传感器网络中发生大量突发事件时会产生大量的紧急报文，已有方法不能快速、实时地把紧急报文传回 Sink，从而引发网络拥塞。本书提出了一个基于分簇网络延迟敏感的控制反馈模型 DSFC 和簇间紧急报文信道预留机制 EPCR，有效提高了紧急报文端到端传输的实时性，并减少了传输开销。DSFC 模型采用自适应闭环控制反馈的方法，每一轮自适应地只聚合一部分紧急报文，在保证一定数据聚合精度的基础上，尽可能地减少簇内聚合开销和延迟。EPCR 机制优先调度传输紧急报文，并为紧急报文预留信道，从而有效缓解了紧急报文和普通报文间的信道竞争和路径拥塞问题。实验表明，在保证相同数据精度的情况下，紧急报文的聚合等待延迟、端到端的总传输延迟、聚合能量开销以及平均丢包率等方面都获得了较好的性能。

针对目前结构化的方法在进行数据复本放置和查询分发时易出现传输开销大、实时性差、热点区域瓶颈等问题，提出了一种非结构化的数据存储和查询策略，有效减少了关键数据复本分发和查询时总的传输开销，在保证查询成功率的前提下，提高了查询的实时性。该策略在保证一定查询成功率的基础上，以最小化总能耗为目标，建立了 MESQ 优化问题模型，并在存储受限和不受限两种情况下，分别给出了复本和查询个数（d，q），证明了其最优性；基于 MESQ 模型，设计了一个实用的分布式实时信息

分发算法 BubbleGeocast，该算法以（d，q）为参数，采用自适应分支扩散的虚拟多播技术进行数据分发，加速了复本和查询的分发速度。分析和实验表明，在相同查询成功率时，BubbleGeocast 有效降低了能耗，减少了复本放置和查询分发的延迟。

综上所述，本书针对高效传输技术在信息分发和收集时的通信开销、实时性、可扩展性等问题，对多播、稀疏采样、簇内聚合、非结构化的存储和查询等关键问题提出了有效的解决方案，对于推进无线传感器网络高效传输的研究和实用化，具有一定的理论意义和应用价值。

PREFACE

Wireless sensor networks (WSNs) are integrated networks which can perform information gathering, processing and delivering. There are wide applications for WSNs in industry, agriculture, military affairs, environment monitoring, biomedicine, city managing and disaster succoring. As a basic issue in WSNs, data delivery determines how control packets are disseminated to sources and how data packets are gathered to sink through multi-hop routing. Moreover, high efficient data delivery determines the performance and energy efficiency in WSNs. The goal of high efficient data delivery is to bridge source and sink, so that they can exchange valid information with the least transmission cost and delay. However, WSNs are resource-constrained, bandwidth-constrained and of large scale, which bring us great challenges to achieve high efficient data delivery.

High efficient data delivery in WSNs includes two main parts: high efficient data dissemination from sink to source and high efficient data gathering from source to sink. Recent works mainly focus on utilizing traditional techniques, such as unicast, broadcast and aggregation. Those works almost all suffer high transmission cost, high delay and poor scalability. Aiming at the inherent characteristics of WSN and the limitation of current works, this thesis studies four key problems on data dissemination and gathering comprehensively to minimize the total transmission cost: (1) multicast from sink to source; (2) sampling from sink to source; (3) aggregation of emergent packets from source to sink; (4) two directional data coordinations between sink and source, data storage and query.

This paper studies the multicast from sink to source to disseminate control information. Aiming at the high communication cost, energy inefficiency and poor scalability, we first propose a base-station model-based multicast protocol, SenCast, to decrease the transmission cost and delay. SenCast can compute an approximately multicast tree globally to route the multicast packets because sink owns more powerful CPU and larger storage. By introducing an MNN (Minimum Nonleaf Nodes) Steiner tree problem which is NP-hard, SenCast builds an MNN multicast tree using a global approximation algorithm. Theoretical analysis shows that SenCast is able to approximate the MNN problem to a ratio of $\ln|R|$ (R is the set of all destinations), the best known lowest bound. Consequently, the multicast traffic can be decreased significantly. We further design two scalable schemes, SRL and HLB which compress the multicast tree information and deliver the multicast messages without information loss. Experimental results demonstrate that SenCast is a scalable and energy efficient multicast protocol when the scale of WSN or the number of destination nodes is large.

Efficient estimation of global information is a common requirement for many WSN applications. Aiming at the high communication cost and delay of recent works, we develop a novel protocol called FLAKE that can efficiently and accurately estimate the global information of large-scale sensor networks based on the fast release/capture sampling. As an example, we outline the basic idea of estimating the number of nodes alive in a large sensor network. We first uniformly disseminate m messages called seeds into the network, which is referred to as seed release. Then a small number of nodes (say m) are queried about whether they have received a seed, which is referred to as seed capture. Suppose the number of seeds received by the n' nodes is m'. The total number of alive nodes in the network can be estimated as mn'/m'. The similar idea can be easily applied to capture other global information of a network. Moreover, to reduce the number of nodes to query in seed capture, FLAKE is based on the inverse sampling theory. Specifically, a query is injected into the network to count the number of seeds. The query process stops immediately after the total of number of seeds found reaches a given threshold. Our analysis show that, by controlling the number of seeds to release and capture, a desirable trade-off between the accuracy of information and the communication overhead can be achieved. To implement the aforementioned idea, FLAKE employs a distributed sparse sampling algorithm that adaptively expands the branches of seed dissemination based on neighborhood sizes. The algorithm can effectively lower the delay of seed release and achieve the global informaticn within the accuracy bound. Our theoretical analysis and simulations show that FLAKE significantly outperforms several existing schemes on message overhead, delay and scalability.

Works in WSNs recently seldom use the differentiated services, priority based packets classification and scheduling method, which causes large overhead and poor delay of packets transmission when bursty events happen. The cluster-based data aggregation has emerged as a basic and general technique in WSNs. However, there is a trade-off issue between aggregation waiting (AW) delay and aggregation accuracy. Therefore, we propose a DSFC model to solve the trade-off issue. Moreover, we present a new inter-cluster congestion avoidance mechanism, EPCR to accelerate the transmission of emergent packets in the inter-cluster routing. We try to find the optimal point for the trade-off between AW delay and accuracy. That is to decrease AW delay as much as possible within the accuracy bound. In this paper, a distinguished feedback control model is proposed to adaptively aggregate partial data in a delay sensitive manner while not damage the gathering accuracy. EPCR differentiates packets with different priority and adopts a method of channel reservation to avoid inter-cluster congestion and accelerate the transmission of emergent packets. EPCR builds a fast end-to-end path for emergent packets which alleviates the congestion between emergent packets and regular packets. Simulation results verify the performance of DSFC and EPCR, and demonstrate their attractions in accuracy, real-time, energy efficiency and ratio of packets loss.

Storage and query are also important problems in WSNs. There are unstructured and structured methods to solve the problem. While the structured methods need an index or hash function to get the locations of replicas beforehand. There are some problems on structured methods, such as high delay, hot spot on popular data, and long stretch. Contrarily, unstructured methods need not know where the replicas are, or need not any storage structure and index. It is more simple and flexible. But a few unstructured methods mainly focus on theory and are not energy efficient. In this work, we focus on unstructured random data storage and query. Since the energy is one of the most precious resources, we formulate an MESQ optimization problem whose aim is to select the optimum number of replicas and queries that minimize the total energy cost, (d, q), subject to unrestrained or restrained storage. In order to make our works more practical, we also design a localized data dissemination algorithm, BubbleGeocast. BubbleGeocast uses (d, q) as parameters and branch adaptively to diffuse data replicas and queries as soon as possible based on virtual multicast tree. We show by theoretical analysis and simulations that our BubbleGeocast achieves the distinguished performance on energy efficiency and delay, within the bound of successful query.

In summary, aiming at scalability, less transmission cost and delay on high efficient data dissemination and gathering, our works present solutions to several key problems of multicast, sparse sampling, aggregation in clustered networks and unstructured storage and query, which have academic and practical value for advancing the theory and practicability of high efficient data delivery in WSNs.

目　录

前　言

PREFACE

第一章　绪　论

第二章　相关研究

第三章　基站模式下的可扩展多播技术

第四章　实时的"释放/捕获"采样技术

第五章　层次化网络结构下的紧急报文实时传输

第六章 非结构化的实时数据存储和查询