

动态无线传感网中数据查询技术的研究进展

梁俊斌 马方强 蒋 婵

(广西大学计算机与电子信息学院 南宁 530004)

(广西大学广西多媒体通信与网络技术重点实验室 南宁 530004)

摘要 无线传感器网络(Wireless Sensor Networks, WSN)是由大量通信、计算及存储能力有限的传感器节点通过自组织方式构成的网络,可以部署在恶劣环境中执行长期的监测任务。数据查询是 WSN 中最基础的操作之一,主要指用户通过特定节点向网络分发查询请求,而网络中满足要求的节点再将数据返回给用户。在查询过程中,由于网络存在动态性(如:节点由于受到外力破坏而损毁、移动或睡眠,导致网络拓扑及连通性发生变化;通信链路不可靠等),使得数据传输出现延迟大、能耗高甚至数据丢失等情况,导致查询成功率低。目前有很多学者对该问题进行了研究,并取得了一定的进展,但实际应用中依然有很多问题仍待解决。为了进一步推进动态无线传感网中数据查询技术的深入研究,对近年来的典型工作进行了分析、归纳和总结,并对比了它们的优缺点;然后讨论了该领域亟需解决的关键问题,并指出了下一步的研究方向。

关键词 无线传感器网络,数据查询,动态性,数据传输,查询成功率

中图分类号 TP393 **文献标识码** A **DOI** 10.11896/jsjcx.181202258

Research Progress on Data Query Technology in Dynamic Wireless Sensor Networks

LIANG Jun-bin MA Fang-qiang JIANG Chan

(School of Computer and Electronics Information, Guangxi University, Nanning 530004, China)

(Guangxi Key Laboratory of Multimedia Communications and Network Technology, Guangxi University, Nanning 530004, China)

Abstract Wireless sensor networks (WSN) are a self-organizing network composed of a large number of sensor nodes with limited communication, computing and storage capabilities, which can be deployed to perform long-term monitoring tasks in harsh environments. Data query processing is one of the most basic operations for WSN to obtain monitoring data. It mainly refers to the user distributing query requests to the network through a specific node, and then the node that satisfies the requirements returns the data to the user. In the process of query, because of the dynamic nature of the network (e. g. , the destruction of nodes by external forces, the movement or sleep, resulting in changes in network topology and connectivity, and unreliable communication links, etc.), the data transmission has large delay, high energy consumption and even data loss, resulting in low success rate of query. At present, many scholars study this problem and make some progress, but there are still many problems to be solved in practical application. In order to further promote the in-depth study of data query technology in dynamic wireless sensor networks, this paper analysed and summarized the typical work in recent years, and compared their advantages and disadvantages. Then, this paper discussed the key issues that need to be solved in this field, and finally pointed out the next research directions.

Keywords Wireless sensor network, Data query, Dynamism, Data transmission, Success rate of query

1 引言

随着微电子技术和无线通信技术的兴起,无线传感器网络被应用于诸多领域,如工业、军事、环保及医疗卫生等^[1-3]。WSN 是由数量庞大且通信范围有限的传感器节点采用多跳的传输方式自组织而成的网络,节点具有电池能量、存储空间、传输带宽等资源有限的特点,通常被部署在人类无法进入

的恶劣环境中来执行长期的事件监测和数据收集任务^[3-6]。

然而,在恶劣的监测环境中,为了避免节点随机分布出现网络空洞或覆盖范围不全等问题,网络采用一定的策略让部分节点通过移动来修复网络空洞或扩大覆盖范围,导致网络拓扑结构不断发生改变^[7-9]。此外,节点的意外损毁、异步占空比的工作模式(即每个传感器节点的睡眠/苏醒时刻及工作周期不同)及数据传输链路的不可靠性,也会导致网络拓扑或

收稿日期:2018-12-04 返修日期:2019-03-27 本文受国家自然科学基金资助项目(61562005,61762010),广西自然科学基金资助项目(2018GXNSFBA281169),广西高等学校千名中青年骨干教师培育计划基金资助项目(桂教人(2017)49)资助。

梁俊斌(1979-),男,博士,教授,硕士生导师,主要研究方向为无线传感器网络,E-mail:liangjb2002@163.com;马方强(1993-),男,硕士生,主要研究方向为无线传感器网络,E-mail:1964195239@qq.com(通信作者);蒋婵(1980-),女,硕士,讲师,主要研究方向为无线传感器网络。

连通性不断发生改变^[10-12]。这种网络拓扑或连通性动态改变的网络被称为动态 WSN。

数据查询是 WSN 中最重要和最基础的操作之一,主要分为数据查询分发和结果回传两个阶段。前者是用户将需求以数据包的形式注入到网络中网络通过一定的执行策略将查询请求发送至部分或所有节点;后者将符合查询条件的目标数据返回给用户,它的主要目标是在网络节点能耗尽可能少的情况下将需求数据及时、可靠地回传给用户,从而让用户对监测事件信息进行及时的处理和分析^[12-15]。

但是,由于以下原因,在动态 WSN 中,数据的快速、可靠查询是一个具有挑战性的问题。

1)在恶劣环境中,节点的损毁是随机发生的,或者存在节点异步地进入睡眠状态等情况,使得网络拓扑无规律地变化,甚至断裂,从而导致数据查询失败或延迟过大^[11-12,16];

2)节点的移动或时钟偏移,使得网络中的节点需要经常相互通信来发现新的邻居或进行时钟同步,导致通信开销和延时增大^[12,15];

3)节点的意外损毁可能会导致部分感知数据丢失而影响查询成功率^[16]。

目前,虽然已经存在许多有关数据查询处理的研究工作,也取得了一定的成果,但是,在实际的应用场景中,网络拓扑或传输路径往往受多种因素(节点损毁、节点移动、链路不可靠等)综合影响,数据查询算法的实用性和可靠性较差,平均通信延迟和能耗过大。为了促进动态 WSN 中数据查询技术的深入研究,本文对近年来的经典算法进行了归纳和总结,分析了它们的优缺点,明确当前研究存在的问题,并指出了下一步的研究方向。

本文第 2 节对查询分发、数据回传路由方式及查询算法进行了分类;第 3 节对比分析了目前基于非结构化路由的经典查询算法;第 4 节对比分析了目前基于结构化路由的经典查询算法;第 5 节对两大类算法进行了对比分析;第 6 节提出了亟需解决的问题和下一步的研究方向;最后总结全文。

2 数据查询路由及查询算法的分类

由于节点移动、睡眠、损毁等原因,在数据查询过程中,寻找节能路由方式进行查询请求分发和数据回传是动态 WSN 中一个谨慎的问题。

2.1 查询请求分发路由方式

根据网络结构,可将目前已存在的查询请求分发路由方法分为平面路由和分层路由两种协议^[8,12,15,17],如图 1 所示。

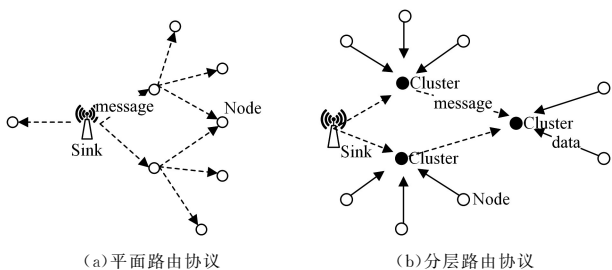


图 1 查询请求分发路由方式的分类

Fig. 1 Classification of routing methods of query request dissemination

如图 1(a)所示,平面路由协议^[17]是将网络中所有节点赋予相同的角色来执行感知任务。Sink 节点采用广播或多播的方式将查询请求发送给邻居节点,邻居节点迭代分发,直到不符合分发条件时停止查询分发。在动态 WSN 中,该协议只需要最小的通信开销来维护节点之间的连通性,但是查询分发的消息包数量和通信冲突通常较大。

如图 1(b)所示,分层路由协议^[17]是为同构网络强加一个结构来实现查询分发的低能耗和低延迟,即将网络划分为多个簇,其中一个节点(具有较高的剩余能量和较大的存储空间)成为簇头(CH),CH 汇聚成员节点数据并在不同集群之间传输,Sink 节点分发(广播、多播或单播)查询消息至附近的 CH,CH 重复分发,直到不符合分发条件时停止查询分发。在动态 WSN 中,所有集群需要动态更新成员节点和维护 CH 之间的连通性。

2.2 数据回传路由方式

根据文献^[18-20],将已存在的数据回传路由方式分为基于位置的数据回传路由方法和基于路由表的数据回传路由方法,如图 2 所示。

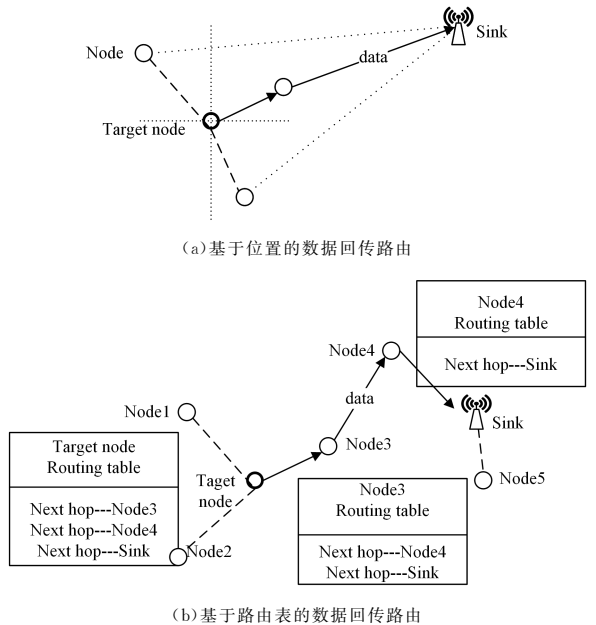


图 2 数据回传路由方式的分类

Fig. 2 Classification of routing methods for data return

如图 2(a)所示,目标节点/转发节点选取 Sink 节点与其所有邻居节点之间距离最小的一个邻居作为新转发节点,重复上述操作,直到数据传输到 Sink 节点。该方法适用于动态 WSN 中的数据回传,但是需要满足以下条件:

- 1)网络中每个节点需要获取自身和它的邻居节点的位置信息;
- 2)目标节点已获取 Sink 位置信息并将其传输给下一跳转发节点。

如图 2(b)所示,目标节点/转发节点按照查询分发时构建的路由列表将目标数据回传到 Sink 节点。该方法适用于节点定位自身位置信息的静态 WSN,其虽然简单且易实现,但是实用性较差。

根据查询请求分发和数据回传所采用的路由方式,可将目前已有的数据查询工作分为两大类:基于非结构化路由的

查询算法和基于结构化路由的查询算法。前者根据是否访问所有节点,可被继续细分为随机分发查询算法和泛洪查询算法;后者根据查询路由所采用的拓扑结构,可被继续划分为基于链拓扑、基于集群拓扑、基于集群树拓扑这 3 类查询算法。查询算法的详细分类如图 3 所示。

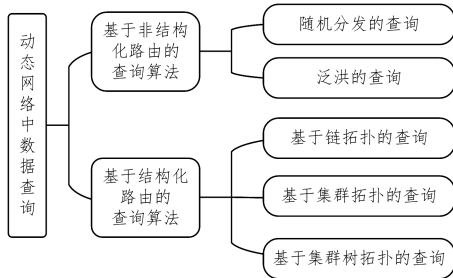


图 3 查询算法的分类

Fig. 3 Classification of query algorithms

3 基于非结构化路由的查询算法

基于非结构化路由的查询算法是在无法预知目标节点所在范围的情况下,查询节点根据一定规则将查询数据包发送给其他邻居,迭代执行上述操作,直到搜索到目标数据时将其返回给用户。该算法的操作简单,且无须预先构建固定的查询路由。

3.1 随机分发查询

随机分发查询^[21-23]是指查询节点按照一定的规则将查询消息随机分发给部分邻居节点,重复上述过程,直到查询到目标数据时将其返回给用户。虽然早期已经有研究者提出了 Gossing^[24], DD^[25], SPIN^[26], GBR^[27] 等查询路由协议,但是这些协议的通信能量消耗和通信延迟等性能较差。

为此,彭绍亮等^[28]提出了一种新的实时、高效的数据查询方式,即采用 Birthday paradox 的思想,让感知数据与查询消息随机复制转发,直到两种报文相遇。该过程具体可分 3 个阶段:1)以通信能耗最小为目标函数、高的成功查询率为约束条件建立优化模型,计算最优查询副本和数据副本;2)设计一种分布式分发算法 BubbleGeocast,用于保证数据副本和查询副本随机且均匀地分发并精准地控制副本的分发数量;3)匹配成功的节点将目标数据按照查询路由回传给用户。该算法采用源数据多路分发,导致节点能耗过大。

为了解决上述问题,Chen 等^[29]提出了一种 SAoD 查询算法,它采用时分协作技术,让同一事件附近的多个节点记录存储数据。源节点再利用 Bloom Filter(BF)压缩源数据块,并使用与 BubbleGeocast 类似的方法多路分发数据,直到压缩数据与查询数据匹配成功。与 BubbleGeocast 算法不同,SAoD 算法利用移动 Sink 节点对查询到的目标区域源数据依次进行收集,从而极大地降低了通信能耗和延时。但是,BF 存在假阳性误判概率和节点意外损毁的情况,会降低查询成功率。

针对 SAoD 算法的缺点,Kolcun 等^[30]提出了一种基于分布式的静态属性表的查询算法(DSAT)。其基本思想是:将全网节点数据的属性列表划分为 P 个容量相同的子表,并将其余 $P-1$ 个子表广播给邻居节点,查询消息最多只需要查找存储 P 个子表的节点即可找到目标节点。具体步骤如下:

1)源节点广播子表,邻居节点存储转发该子表,直到 $P-1$ 个子表被全部存储;2)查询消息根据 Routing Table(RT)快速查询子表,从而获得目标数据。此外,邻居节点之间通过定期更新 RT 来解决节点失效导致查询失败的问题。该算法虽然加快了查询速度,但是对节点的能耗和存储空间的要求较高,并不适合在大规模动态 WSN 中进行数据查询。

表 1 直观地列出了几种随机分发查询算法的平均能耗、平均延迟、查询成功率及它们的优缺点。可以看到,Bubble-Geocast 和 DSAT 算法主要是通过存储冗余数据来提高查询成功的概率,往往会导致网络寿命缩短而过早死亡;SAoD 主要通过减小数据体积或传输总跳数来缩短平均通信延迟。

表 1 随机分发查询算法的性能对比

Table 1 Performance comparison of random dissemination query algorithms

算法	平均能耗	平均延迟	查询成功率	优点	缺点
Bubble Geocast	较高	较高	较高	网络容错性高	网络均匀分布
SAoD	较低	较低	较低	减少数据容量和存储	网络容错性差
DSAT	较高	较低	较高	扩展性强且易实现	存储空间消耗大

3.2 泛洪查询

在 WSN 中,泛洪查询^[22,31]是指查询消息被广播给所有邻居节点,重复执行上述操作,直到查询消息到达目标节点,并将符合查询条件的数据返回给用户。这类查询容易造成传输能量消耗过大、传输冲突较高及查询成功率较低的问题,已经很少应用在动态 WSN 中。

Cao 等^[32]对简单泛洪算法进行了改进,提出了一种占空比网络中完全分布式的低功率监听泛洪算法(LPL),其利用满足并发广播的严格信号时间和强度要求来减少争用和冲突至其他节点。具体步骤如下:1)节点随机调整前导码的间隔,并将其广播;2)接收器基于轻量级信号模式识别的方法识别前导码,确定苏醒邻居节点;3)苏醒节点接收查询消息,并将其与存储数据进行匹配。该算法可能造成冗余的苏醒时隙,从而导致能耗过高,而且数据查询的实时性较差。

为了解决 LPL 算法存在的缺点,Chen 等^[33]提出了一种延迟有界的多播新算法(MAADM)。其主要思想是:在保证任意有限的传输时延的条件下,使节点增加的苏醒时隙最小化。由于它是一个 NP 难问题,当目的节点数量 $k=2$ 时,以预设的延迟截止时间为界,源节点采用启发式方式进行传输数据;当目的节点数量 $k>2$ 时,采用 $2(k-1)$ -可近似算法求解最小扩充时隙,再进行数据传输。该算法通过增加节点苏醒时隙来有效地控制通信时延和能耗。

Cao 和 Chen 等提出的算法主要针对链路可靠的占空比传感网(DC-WSN)。Guo 等^[34]针对节点异步协作工作的不可靠通信的低占空比传感网(LDC-WSN),提出了一种新的机会泛洪算法(OF)。其基本思想是:从能量优化结构树开始,每一个发送节点基于下一跳接收器的延迟分布做出概率性的转发决策,以节省节点传输能耗;此外,采用基于链路质量的退避方法解决多个节点同时转发的冲突,从而降低通信延迟。

针对 OF 算法无法保证数据查询实时性的问题,Cheng

等^[35]提出了一种分布式最小延迟能量有效泛洪树算法(MDET)。该算法首先基于预期传输计数(ETX)的方法构建一棵最短路径树;然后在数据分发过程中,发送节点根据通信时延和传输能耗最小原则动态地局部调整泛洪树,选取最佳传输链路质量的父节点进行数据广播,直至查询到目标节点。

上述4种查询算法都一定程度地降低了查询的通信能耗和时延,但均未考虑节点损毁而导致查询成功率降低等情况。此外,OF和MDET算法针对的是节点异步协作工作的链路不可靠的LDC-WSN,而LPL和MAADM算法针对的是链路可靠的DC-WSN,其中LPL算法考虑的是节点同步协作工作的网络。几种泛洪查询算法的性能对比如表2所列。

表2 泛洪查询算法的性能对比

Table 2 Performance comparison of flooding query algorithms

算法	平均能耗	平均延迟	网络模型	优点	缺点
LPL	较高	较高	DC-WSN	扩展性强且易实现	传输链路不可靠且实用性差
MAADM	较低	较低	DC-WSN	算法适用范围更广	算法复杂度高且传输链路不可靠
OF	较高	较低	LDC-WSN	数据传输成功率高	前期能耗大且实时性较差
MDEF	较高	较低	LDC-WSN	扩展性和数据传输成功率高	前期能耗大且存在节点损毁

4 基于结构化路由的查询算法

基于结构化路由的查询算法是指查询节点按照预先构造的网络拓扑路由将查询消息传递给固定路径的路由节点,网络再将符合查询条件的数据返回给用户。如图3所示,我们将分别对基于链拓扑、基于集群拓扑及基于集群树拓扑这3种路由查询算法进行对比分析。

4.1 基于链拓扑的查询

基于链拓扑的查询^[36-37]指选择部分传感器节点作为中继节点构成链路径,查询节点沿着链路径传输查询消息,直到查找到目标数据并回传给用户。该类算法的主要缺点是可能存在较长链路上的通信延迟较大的情况。

因此,Sun等^[38]提出了一种组播查询的数据分发(Multi-cast-Query-based Data Dissemination, MQDD)协议。其基本思想是:基于地理路由算法,使源节点在水平方向双向传输源数据的摘要信息,Sink节点选取周围节点组播查询消息以搜索目标数据。当感知数据与查询消息匹配成功时,目标节点将以Sink节点位置为方向,选取距离Sink节点最近的邻居节点快速地返回查询数据。与之前提出的GHT^[39],LB-DD^[40],RailRoad^[41]等协议相比,MQDD协议无需获取全网拓扑结构信息即可执行低能耗、低延迟的数据查询。但是,网络拓扑的动态变化会导致水平方向的链拓扑断裂而降低查询成功率。

为了解决这个问题,Boukerche等^[42]提出了一种新的节能、稳健的串行查询处理算法,它是基于GBT(Greedy & Boundary Traversal)的一种分布式查询方法。具体分为4步:1)每个节点将感知数据存储在部分邻居节点上进行备份存储;2)查询节点根据贪婪算法,选取位置最远的邻居节点作为下一跳节点进行查询,并将节点位置存储到路由表(RT)中;3)如果形成查询回路,当前节点切换到边界遍历方式,继

续寻找未被访问过的邻居节点;4)重复地交替执行两种算法,直到查询到目标数据,并将查询结果返回给查询节点。GBT算法具有良好的容错能力,保证了查询的可靠性,而且解决了查询时的网络‘漏洞’问题。但是,构建查询路径的过程比较复杂,使得通信延迟较大。

为了克服GBT算法的缺点,Merzoug等^[43]提出了一种新的串行处理算法GSS(Geometric Serial Search)。其基本思想是:首先,用户选取网内任何一个节点注入查询请求,选取半径为 $r/2$ 的rolling-ball几何图形进行遍历;然后,每个当前节点广播查询请求,修改接收到查询请求的节点的访问位;最后,发送节点选取rolling-ball内部均为访问过节点的邻居节点作为下一跳节点,重复执行查询操作。如果查询到所需要的数据,采用与MQDD算法相似的方法,目标节点以查询节点的地理位置为方向回传目标数据。GSS算法解决了传输冲突和网络拓扑改变导致查询失败的问题,但是网络的平均能量消耗过大。

如表3所列,在动态WSN中,虽然GBT算法和GSS算法的查询成功率较高,但是GBT算法以较高的通信延迟为代价,而GSS算法则需要发送大量的数据副本进行查询。MQDD协议虽然具有较低的平均通信能耗和延迟,但是与另两种算法相比,网络扩展性和容错性较差,因此查询成功率不高。

表3 基于链拓扑的查询算法的性能对比

Table 3 Performance comparison of query algorithms based on chain topology

算法	平均能耗	平均延迟	查询成功率	优点	缺点
MQDD	较低	较低	较低	简单易实现	扩展性和容错性差
GBT	较低	较高	较高	扩展性和容错性强	操作过程复杂
GSS	较高	较低	较高	扩展性强且易实现	广播数据量大

4.2 基于集群拓扑的查询

基于集群拓扑的查询^[36,44]是指查询消息只需要对簇头节点中的数据进行检索,并将满足查询条件的数据返回给用户的方式。其中,集群内的通信采用单跳方式(即CH与成员节点直接连接),集群间的通信采用多跳方式(即CH之间通过其他节点间接相连)。

首先,于秦等^[45]针对通信不可靠的网络提出了一种以最优搜索代价和最大探测成功概率为目标的查询算法(OSFP)。它对已经被分簇的传感网络建立最优搜索概率模型 $P(f) = \sum_{j \in J} p(j) [1 - (1 - a_j)^{f(j)}]$ (其中, $p(j)$ 为目标节点在簇 j 中的概率, a_j 为访问一次目标簇 j 查找成功的概率, $f(j)$ 为对目标簇 j 的搜索次数),以确定最优搜索方案。在此基础上,Sink节点再分发查询消息给相应的簇头节点进行查询。该算法虽然缩短了查询时间,但是需要预先获取完全CH位置信息,且未考虑CH失效等情况。

针对OSFP算法的缺点,Zhao等^[46]提出了一种基于流行度的缓存机制的周期性查询算法(PCMPQ)。在OSFP基础上,每个CH节点统计最近一段时间被查询成功的次数(即流行度),如果流行度大于其预先设置的阈值,则对应CH节点去冗余数据尽可能多地缓存在网络中的其他CH节点上,从而

加快了查询速度,提高了查询成功概率。然而,该算法需要 CH 节点进行大量的数据传输,从而导致通信能耗较大;此外,它还存在网络负载不均衡等问题。

针对上述算法的缺点,Liu 等^[47]和 Martina 等^[48]分别提出了 RSA 算法和 EeFTSQ 算法。前者根据能耗公式定期对每个单元网格大小和 CH 节点进行优化,并通过 CH 计算成员节点的感知数据,将符合查询条件的部分查询结果发送给下一网格中的 CH,依次执行,直到遍历完查询区域内的所有网格单元。后者在 CH 节点失效时,选取距离其最近的节点作为新的汇集节点,并利用容错路由协议将查询消息转发至目标区域,再将目标数据回传给用户。这两种算法虽然通过去除冗余数据降低了通信能耗,而且解决了单点失效导致的查询成功率降低的问题。但是,两种算法的通信延迟均较高。

针对上述算法存在的缺点,Cheng 等^[49]提出了一种曲线查询处理算法(CQP)。该算法可分为以下几个步骤:1)将离散的感知数据连续化,并采用感知曲线推导算法获取监测环境的感知曲线;2)CH 对来自邻居的多个感知曲线进行汇聚;3)CH 根据用户给定的查询错误边界,将符合查询请求的感知曲线返回给 Sink 节点。该算法不仅降低了网络通信开销,缩短了延迟,而且提高了查询可靠性。但是,它的计算复杂度较高,且未考虑网络存在传输链路不可靠等问题。

为了解决 CQP 算法存在的问题,Chen 等^[50]提出了一种自适应容错的 QoS 控制查询算法(AFTQC)。其基本思想是:采用源数据冗余存储和在 CH 之间构建冗余传输路径这两种机制来保证网络链路的服务质量,延长网络寿命。具体操作如下:首先,每个集群中的 CH 通过聚合成员节点的数据来响应查询请求;其次,在集群内采用轮换机制定期更换 CH,而在集群间构建多条通路来进行相互通信。但是,该算法在前期的数据冗余存储和构建冗余连通路时的通信能耗较大。

表 4 从算法的基本思想、优缺点等方面对基于集群拓扑的查询算法进行了直观对比和分析。从中可以看出,在动态 WSN 中,大部分算法通过访问全网中的所有 CH 来提高查询成功概率。目前,综合考虑低能耗、低延迟及高成功查询概率等多因素的实用性研究还较少。

表 4 基于集群拓扑的查询算法的对比

Table 4 Comparison of query algorithms based on cluster topology

算法	主要思想	优点	缺点
OSFP	广播查询所有 CH	操作简单且易实现	成功查询的概率较低
PCMPQ	多次访问的 CH 将数据缓存在其他 CH 中	加快查询速度且查询成功率较高	传输能耗较高
RSA	线性查询所有 CH	操作简单且扩展性强	通信延迟较高
EeFTSQ	多路分发查询消息	查询成功率较高且能耗较低	通信延迟较高
CQP	基于感知曲线的聚合操作去除冗余数据	通信能耗和延迟较低	算法复杂度较高且查询成功率较低
AFTQC	源数据冗余存储在 CH 间构建多条传输路径	查询成功率较高且通信延迟较低	网络通信能耗较高

4.3 基于集群树拓扑的查询

基于集群树拓扑的查询^[36,51-52]是指将网络中所有集群中的 CH 节点按照一定的规则构建成一棵树,在查询消息时按照这棵树进行分发查询,并将查找到的目标数据回传给用户。在动态 WSN 中,CH 节点失效、网络负载不均衡及网络拓扑改变成为该查询方法的一个难点问题。

Cui 等^[53]提出了一种基于最小生成树(Minimum Spanning Tree, MST)的查询协议(QPMST),它以传输能耗最小为原则将所有节点构建成一棵 MST, Sink 节点根据 MST 传输查询请求,叶子节点将满足查询条件的数据传输到父节点上,父节点对这些数据进行聚合操作,去除了冗余数据,依次迭代执行,直到将目标数据传输给 Sink 节点。该算法虽然避免了冗余数据的传输,但是需要维护所有节点构成 MST 的连通性。

针对 QPMST 算法存在的缺点,Sarode 等^[54]提出了一种自适应动态剪枝数据聚合查询算法(APDA)。该算法可分为 3 个步骤:1)每个子区域节点传输感知数据到 CH,CH 对数据进行归一化处理;2)采用与 QPMST 算法相似的方法,将所有 CH 构建成一棵 MST;3)查询节点按照 MST 转发查询消息,在此过程中将不符合查询条件的 CH 节点剪切掉,以降低网络通信能耗。此外,为了解决网络拓扑动态变化导致查询成功率降低的问题,该算法会定期更换 CH 节点并重构最小生成树(MST)。该算法只需要通过维护部分节点链的连通性,即可获得目标数据。

在此基础上,Gao 等^[55]提出了一种基于数据属性的空间索引查询算法(SIQ)。SIQ 可分为 4 个部分来实现低能耗、快速的数据查询目标:1)采用分簇存储模式将感知相似数据的传感器节点划分在同一地理位置;2)CH 存储成员节点的感知数据和感知区域的矩形边界;3)将所有 CH 作为叶子节点插入到 R-tree 中,依次迭代,构造一个空间索引 R-tree;4)查询请求根据空间索引树结构快速地查找目标节点所在的区域,并将目标数据返回给用户。此外,为了解决 CH 失效的问题,每个区域根据能量均衡原则定期更新 CH。与 APDA 算法相比,SIQ 的查询速度更快,且需要维护索引树的能耗更少。但是,它会导致网络中的部分存储空间不足。

针对上述算法的不足,Chen 等^[56]提出了一种异构网络中基于压缩感知的自适应查询算法(CSA),其中普通传感器节点感知数据,存储节点(SN)负责汇聚、存储和处理感知数据。其基本思想为:首先,网络以 Sink 为根节点,将所有 SN 构建成一棵 MST;然后,在数据感知过程中,普通节点利用压缩感知(CS)技术采集样本数据并将其存储在本地的 SN 上;最后,在数据查询过程中, Sink 节点发送请求至 SN 上,SN 根据等级从压缩域导出所需数据并将查询结果返回给 Sink。该算法根据查询者发起的等级动态地改变压缩数据,节省了网络的通信开销。但是,该算法未考虑 SN 失效导致查询失败等情况。

表 5 直观地从算法的主要思想和优缺点几个方面对上述基于集群树拓扑的查询算法进行了对比。从中可以看到,大部分集群树拓扑路由查询算法需要获取全网拓扑信息,构建

树形结构来保证查询的实时性。但是,网络拓扑的动态变化性,使得网络需要定期更新集群的 CH 节点和维护树拓扑结构来提高成功查询的概率,导致网络通信能耗过高。

表 5 基于集群树拓扑的查询算法的对比

Table 5 Comparison of query algorithms based on cluster tree topology

算法	主要思想	优点	缺点
QPMST	所有节点构成 MST,父节点聚合子节点数据,迭代执行	去除冗余传输,加快查询速度	维护 MST 的连通性使得能耗大
APDA	所有 CH 构成 MST,查询消息剪切不符合查询条件的子区域	通信延迟较低且查询成功率高	定期更新 CH 节点和 MST 使得能耗大
SIQ	利用 CH 构造空间索引 R-tree,查询消息根据索引树进行查找	构造 R-tree 的能耗低且查询简单	节点负载(存储空间、能量)不均衡
CSA	节点压缩采样传输至 SN 上,根据 SN 构建的 MST 进行 top-k 查询	节省节点能耗和存储空间	两层网络,查询成功率较低

5 算法的对比分析总结

由第 3 节和第 4 节可知,在动态 WSN 中,基于非结构化和结构化的路由查询算法都是通过数据冗余存储和查询消息的大量分发来提高查询成功的概率。它们的区别是,前者采用随机复制转发的方式,将查询消息或数据多路随机传输至其他节点中进行查询,以提高查询成功率;但是,在查询过程中需要大量的数据传输,使得节点的平均能量损耗和平均通信时延较大。后者采用固定路由结构分发查询消息,降低了数据查询的平均能耗和通信延迟;但是,网络自身的动态特性,容易导致查询路由断裂而查询失败,因此该算法通常需要预先构建多条冗余的查询路由或定期构建新的查询路由,这增大了节点的平均能耗。具体算法的分析对比如表 6 所列。

表 6 基于非结构化和结构化的路由查询算法的对比

Table 6 Comparison of route query algorithms based on unstructured and structured

算法	基本思想	优点	缺点
基于非结构化路由查询	采用随机多路转发数据报文	查询成功率较高且简单易实现	数据转发数量较多和传输冲突较短
基于结构化路由查询	采用固定路由结构转发数据报文	数据转发数量较少且查询时间较短	为保证查询成功率,需要定期构建新路由

6 待解决的问题及未来的研究方向

在动态 WSN 中,针对数据查询的研究虽然已经取得了一些进展,但是仍然存在一些亟待解决的问题。下面将对这些问题进行分析总结,并展望未来的研究方向。

(1) 如何保证高概率的查询成功率

在动态 WSN 中,影响查询成功率的因素主要可分为两个方面。1) 节点的意外损毁或睡眠。由于网络通常将感知数据存储在本地区或其他节点上,如果这些节点损毁或处于睡眠状态,就意味着查询请求无法访问到这部分节点的感知数据,

使得查询成功的可能性降低。2) 网络链路的不可靠性。查询节点一般采用多跳的传输方式将查询消息分发至网络中。如果在传输过程中存在数据丢失的情况,则将导致数据查询失败。因此,需要研究一种新的数据查询算法来保证查询的可靠性。

未来可从以下 3 个方向进行研究:1) 基于容错的数据查询方式,研究如何使数据在最小备份数量的条件下,确保查询请求能够搜索到目标数据;2) 基于数据冗余编码的多路分发机制,结合数据传输的丢包率,研究如何对数据进行最小代价的冗余编码和分组划分,从而多路分发给邻居节点以快速到达 Sink 节点,Sink 节点可根据获得的部分编码数据恢复完整数据;3) 针对 LDC-WSN,采用何种数据分发方式,使得睡眠的 CH 接收查询消息,并将符合查询结果的数据返回给用户。

(2) 如何平衡传输能耗和时延

在拓扑结构动态变化的传感网中,通常采用以下两种方式:数据的冗余随机分发和基于固定拓扑路由的数据分发。这两种方式虽然都缩短了数据查询的平均通信时间,但是前者需要分发大量不必要的查询消息,后者需要动态更新固定拓扑结构,从而使得网络通信能耗过高。因此,在动态 WSN 中,如何在保证高成功查询率的同时平衡网络通信能耗和时延是另一个研究重点。

针对上述问题,未来可以研究的方向如下:1) 在拓扑路由查询方面,需要结合实际监测环境和网络拓扑的动态性,低能耗地选取下一时刻能量最大的邻居节点,重新构建新的拓扑路由,实现低能耗、低延迟的数据查询;2) 在数据处理查询方面,在对感知数据进行压缩去冗的同时,使数据传输的丢失率最低且通信时延最短。

结束语 本文综述了近年来动态 WSN 中数据查询的研究进展,相比静态 WSN,动态 WSN 因节点自身特性(移动性、占空比工作等)或外部因素(节点损毁、传输链路不可靠)的干扰,网络拓扑结构不断改变,为有效的数据查询应用带来了新的难点问题。本文从网络结构化拓扑路由和非结构化拓扑路由这两个方面,对数据查询算法进行了归纳和总结,分析和对比了当前具有代表性的研究成果。可以看到,现有的工作虽然已经取得了一定的进展,但是在实际应用方面还存在诸多不足。为了推进数据查询算法的研究,提出了动态 WSN 中数据查询技术研究的两个待解决的科学问题:如何保证高概率的查询成功率以及如何平衡网络的通信能耗和时延。同时,针对每个问题存在的原因进行了分析,并且展望了下一步的研究方向。

参考文献

- [1] TANG J,ZHANG B,ZHOU Y,et al. An Energy-Aware Spatial Index Tree for Multi-Region Attribute Query Aggregation Processing in Wireless Sensor Networks[J]. IEEE Access,2017,5: 2080-2095.
- [2] REN H,LI H,DAI Y,et al. Querying in Internet of Things with Privacy Preserving:Challenges,Solutions and Opportunities[J]. IEEE Network,2018(99):1-8.

- [3] YOUNIS M, SENTURK I F, AKKAYA K, et al. Topology Management Techniques for Tolerating Node Failures in Wireless Sensor Networks: A survey[J]. *Computer Networks*, 2014, 58:254-283.
- [4] KHAN A W, ABDULLAH A H, RAZZAQUE M A, et al. VG-DRA: A Virtual Grid-Based Dynamic Routes Adjustment Scheme for Mobile Sink-Based Wireless Sensor Networks[J]. *IEEE Sensors Journal*, 2015, 15(1):526-534.
- [5] CHEN Y B, NEVAT I, ZHANG P, et al. Query-Based Sensors Selection for Collaborative Wireless Sensor Networks with Stochastic Energy Harvesting[J]. *IEEE Internet of Things Journal*, 2018, 10:1109-1122.
- [6] MAHMOOD M A, SEAH W K G, WELCH I. Reliability in Wireless Sensor Networks: A Survey and Challenges ahead[J]. *Computer Networks*, 2015, 79:166-187.
- [7] GU Y, REN F, JI Y, et al. The Evolution of Sink Mobility Management in Wireless Sensor Networks: A Survey[J]. *IEEE Communications Surveys & Tutorials*, 2017, 18(1):507-524.
- [8] CHEN X, YU P. Research on Hierarchical Mobile Wireless Sensor Network Architecture with Mobile Sensor Nodes[C]//2010 3rd International Conference on Biomedical Engineering and Informatics. Yantai:IEEE, 2010, 7:2863-2867.
- [9] MAHBOUBI H, AGHDAM A G. Distributed Deployment Algorithms for Coverage Improvement in a Network of Wireless Mobile Sensors:Relocation by Virtual Force[J]. *IEEE Transactions on Control of Network Systems*, 2017, 4(4):736-748.
- [10] ANASTASI G, CONTI M, FRANCESCO M D. A Comprehensive Analysis of The MAC Unreliability Problem in IEEE 802.15.4 wireless sensor networks[J]. *IEEE Transactions on Industrial Informatics*, 2011, 7(1):52-65.
- [11] KAFI M A, OTHMAN J B, BADACHE N. A Survey on Reliability Protocols in Wireless Sensor Networks[J]. *ACM Computing Surveys*, 2017, 50(2):31.
- [12] HAO J, ZHANG B, MOUFTAH H T. Routing Protocols for Duty Cycled Wireless Sensor Networks: A survey[J]. *Communications Magazine IEEE*, 2012, 50(12):116-123.
- [13] SILVA R D D, MACEDO D F, NOGUEIRA J M S. Spatial Query Processing in Wireless Sensor Networks-A survey[J]. *Information Fusion*, 2014, 15(1):32-43.
- [14] PANTAZIS N A, NIKOLIDAKIS S A, VERGADOS D D. Energy-Efficient Routing Protocols in Wireless Sensor Networks: A survey[J]. *IEEE Communications Surveys & Tutorials*, 2013, 15(2):551-591.
- [15] LIU Z, LI Z, LI M, et al. Path Reconstruction in Dynamic Wireless Sensor Networks Using Compressive Sensing[J]. *IEEE/ACM Transactions on Networking*, 2016, 24(4):1948-1960.
- [16] TANG B, JAGGI N, TAKAHASHI M. Achieving Data k-Availability in Intermittently Connected Sensor Networks[C]//2014 23rd International Conference on Computer Communication and Networks. Shanghai:IEEE, 2014:1-8.
- [17] WARRIER M M, KUMAR A. Energy Efficient Routing in Wireless Sensor Networks: A survey[C]//2016 International Conference on Wireless Communications, Signal Processing and Networking. Chennai:IEEE, 2016:1987-1992.
- [18] GOYAL D, TRIPATHY M R. Routing Protocols in Wireless Sensor Networks: A survey[C]//2012 Second International Conference on Advanced Computing & Communication Technologies. Rohtak:IEEE, 2012:474-480.
- [19] LIU X. Atypical Hierarchical Routing Protocols for Wireless Sensor Networks: A Review[J]. *IEEE Sensors Journal*, 2015, 15(10):5372-5383.
- [20] CHATAP A, SIRSIKAR S. Review on Various Routing Protocols for Heterogeneous Wireless Sensor Network[C]//I-SMAC (IoT in Social, Mobile, Analytics and Cloud), 2017 International Conference on. Palladam:IEEE, 2017:440-444.
- [21] HEDETNIEMI B S M, HEDETNIEMI S T, LIESTMAN A. A Survey of Gossiping and Broadcasting in Communication Networks[J]. *Networks*, 2010, 18(4):319-349.
- [22] AHMAD N, HUSSAIN S Z. Broadcast Expenses Controlling Techniques in Mobile Ad-Hoc Networks: a survey[J]. *Journal of King Saud University-Computer and Information Sciences*, 2016, 28(3):248-261.
- [23] HASAN M Z, AL-RIZZO H, AL-TURJMAN F. A Survey on Multipath Routing Protocols for QoS Assurances in Real-Time Wireless Multimedia Sensor Networks[J]. *IEEE Communications Surveys & Tutorials*, 2017, 19(3):1424-1456.
- [24] DIMAKIS A G, SARWATE A D, WAINWRIGHT M J. Geographic Gossip: Efficient Aggregation for Sensor Networks[C]//Proceedings of the 5th International Conference on Information Processing in Sensor Networks. New York:ACM, 2006:69-76.
- [25] INTANAGONWIWAT C, GOVINDAN R, 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. New York:ACM, 2000:56-67.
- [26] HEINZELMAN W R, KULIK J, BALAKRISHNAN H. Adaptive Protocols for Information Dissemination in Wireless Sensor Networks[C]//Proceedings of the 5th Annual ACM/IEEE International Conference on Mobile Computing and Networking. New York:ACM, 1999:174-185.
- [27] SCHURGERS C, SRIVASTAVA M B. Energy Efficient Routing in Wireless Sensor Networks[C]//2001 MILCOM Proceedings Communications for Network-Centric Operations: Creating the Information Force. McLean:IEEE, 2001, 1:357-361.
- [28] PENG S L, LI S S, PENG Y X, et al. A Real-Time and Efficient Data Storage and Query Method in Wireless Sensor Networks [J]. *Journal of Communications*, 2008, 29(11):128-138. (in Chinese)
彭绍亮,李姗姗,彭宇行,等.无线传感器网络中一种实时高效的数据存储和查询方法[J].*通信学报*, 2008, 29(11):128-138.
- [29] CHEN H, JIN H, GUO L. Sink-Free Audio-on-Demand over Wireless Sensor Networks[J]. *IEEE Transactions on Computers*, 2016, 65(5):1606-1618.
- [30] KOLCUN R, BOYLE D E, MCCANN J A. Efficient Distributed Query Processing[J]. *IEEE Transactions on Automation Science & Engineering*, 2016, 13(3):1230-1246.
- [31] CAO Z, LIU D, WANG J, et al. Chase:Taming Concurrent Broadcast for Flooding in Asynchronous Duty Cycle Networks

- [J]. *IEEE/ACM Transactions on Networking*, 2017, 25(5): 2872-2885.
- [32] CAO Z, HE Y, MA Q, et al. Lazy Forwarding in Low-Duty-Cycle Wireless Sensor Network[J]. *IEEE/ACM Transactions on Networking*, 2015, 23(3): 922-930.
- [33] CHEN Q, GAO H, CHENG S, et al. Centralized and Distributed Delay-Bounded Scheduling Algorithms for Multicast in Duty-Cycled Wireless Sensor Networks[J]. *IEEE/ACM Transactions on Networking*, 2017, 25(6): 3573-3586.
- [34] GUO S, HE L, GU Y, et al. Opportunistic Flooding in Low-Duty-Cycle Wireless Sensor Networks with Unreliable Links[J]. *IEEE Transactions on Computers*, 2014, 63(11): 2787-2802.
- [35] CHENG L, NIU J, LUO C, et al. Towards Minimum-Delay and Energy-Efficient Flooding in Low-Duty-Cycle Wireless Sensor Networks[J]. *Computer Networks*, 2018, 134: 66-77.
- [36] VELMANI R, KAARTHICK B. An Efficient Cluster-Tree Based Data Collection Scheme for Large Mobile Wireless Sensor Networks[J]. *IEEE Sensors Journal*, 2015, 15(4): 2377-2390.
- [37] LUO H, YE F, CHENG J, et al. TTDD: Two-Tier Data Dissemination in Large-Scale Wireless Sensor Networks[J]. *Wireless networks*, 2005, 11(1/2): 161-175.
- [38] SUN L, LONG C, CHEN C, et al. A Multicast-Query-Based Data Dissemination Protocol for Wireless Sensor Networks with Multiple Mobile Sinks[J]. *Intelligent Control & Automation World Congress on*, 2010, 20(1): 1240-1244.
- [39] RATNASAMY S, KARP B, YIN L, et al. GHT: A Geographic Hash Table for Data-Centric Storage[C]// *Proceedings of the 1st ACM International Workshop on Wireless Sensor Networks and Applications*. New York: ACM, 2002: 78-87.
- [40] HAMIDA E B, CHELIUS G. Strategies for Data Dissemination to Mobile Sinks in Wireless Sensor Networks[J]. *Wireless Communications IEEE*, 2008, 15(6): 31-37.
- [41] SHIN J H, KIM J, PARK K, et al. Railroad: Virtual Infrastructure for Data Dissemination in Wireless Sensor Networks[C]// *Proceedings of the 2nd ACM International Workshop on Performance Evaluation of Wireless Ad Hoc, Sensor, and Ubiquitous Networks*. New York: ACM, 2005: 168-174.
- [42] BOUKERCHE A, MOSTEFAOUI A, MELKEMI M. Efficient and Robust Serial Query Processing Approach for Large-Scale Wireless Sensor Networks[J]. *Ad Hoc Networks*, 2016, 47: 82-98.
- [43] MERZOUG M A, BOUKERCHE A, MOSTEFAOUI A. Serial In-Network Processing for Large Stationary Wireless Sensor Networks[C]// *Proceedings of the 20th ACM International Conference on Modelling, Analysis and Simulation of Wireless and Mobile Systems*. New York: ACM, 2017: 153-160.
- [44] CHOUDHARYL R, KUMAR S, DEEPAK A, et al. Data aggregation in Wireless Sensor Network: An integer linear programming formulation for energy optimization[C]// *2016 Thirteenth International Conference on Wireless and Optical Communications Networks*. Hyderabad: IEEE, 2016: 1-6.
- [45] YU Q, WANG W D, QIN Z G, et al. Research on Optimal Search Flood Protocol for Wireless Sensor Networks[J]. *Computer Science*, 2010, 37(2): 56-60. (in Chinese)
于秦, 王伟东, 秦志光, 等. 无线传感器网络最优搜索泛洪协议研究[J]. *计算机科学*, 2010, 37(2): 56-60.
- [46] ZHAO D, ZHOU Z, NING K, et al. Cache-Based Periodic Query Optimization for Wireless Sensor Networks[C]// *2014 International Conference on Identification, Information and Knowledge in the Internet of Things*. Chennai: IEEE, 2014: 216-219.
- [47] LIU L, QIN X, LI B, et al. Energy-Efficient and Robust Spatial Window Aggregation Query Processing Algorithm in Wireless Sensor Networks[C]// *International Conference on Distributed Computing Systems Workshops*. Macau: IEEE, 2012: 169-177.
- [48] MARTINA M P P, FELIX E V S. Energy-Efficient and Fault Tolerant Spatial Query Processing in Wireless Sensor Networks[C]// *International Conference on Advanced Communication Control and Computing Technologies*. Ramanathapuram: IEEE, 2015: 790-794.
- [49] CHENG S, CAI Z, LI J. Curve Query Processing in Wireless Sensor Networks[J]. *IEEE Transactions on Vehicular Technology*, 2015, 64(11): 5198-5209.
- [50] CHEN I R, SPEER A P, ELTOWEISSY M. Adaptive Fault-Tolerant QoS Control Algorithms for Maximizing System Lifetime of Query-Based Wireless Sensor Networks[J]. *IEEE Transactions on Dependable and Secure Computing*, 2011, 8(2): 161-176.
- [51] MOH'D ALIA O. Dynamic Relocation of Mobile Base Station in Wireless Sensor Networks Using a Cluster-Based Harmony Search Algorithm[J]. *Information Sciences*, 2017, 385: 76-95.
- [52] YAN H, AL-HOQANI N, YANG S H. In-network Multi-Sensors Query Aggregation Algorithm for Wireless Sensor Networks Database[C]// *2018 IEEE 15th International Conference on Networking, Sensing and Control*. Zhuhai: IEEE, 2018: 1-8.
- [53] CUI Y, QIN H. Data Query Protocol Based on Minimum Spanning Tree for Wireless Sensor Network[C]// *2010 Fourth International Conference on Genetic and Evolutionary Computing*. Shenzhen: IEEE, 2010: 798-801.
- [54] SARODE P, NANDHINI R. APDA: Adaptive Pruning & Data Aggregation Algorithms for Query Based Wireless Sensor Networks[C]// *2016 International Conference on Global Trends in Signal Processing, Information Computing and Communication*. Jalgaon: IEEE, 2016: 219-224.
- [55] GAO Z, ZHENG J. Queries Based on Data Attribute Spatial Index in Wireless Sensor Networks[C]// *2018 International Conference on Intelligent Transportation, Big Data & Smart City*. Xiamen: IEEE, 2018: 183-186.
- [56] CHEN Y S, TSOU Y T. Compressive Sensing-Based Adaptive Top-k Query over Compression Domain in Wireless Sensor Networks[C]// *Wireless Communications and Networking Conference*. San Francisco: IEEE, 2017: 1-6.