

鲁棒的无线传感器网络空间 范围聚集查询处理算法

刘 亮, 秦小麟, 李博涵, 刘宇雷

(南京航空航天大学 信息科学与技术学院, 江苏南京 210016)

摘 要: 现有无线传感器网络环境下的空间范围聚集查询处理算法没有综合考虑节点移动、节点失效和通信链路失效等因素对查询处理的影响, 导致查询成功率低和能量消耗大. 本文提出了利用节点冗余保证查询处理过程鲁棒性的空间范围聚集查询处理算法 RSA (Robust Spatial window aggregation query processing Algorithm), 它将查询区域划分为若干个网格, 沿一条路线依次收集各网格内所有节点的数据. RSA 算法给出了一种自适应的网格大小设置方法和基于网格面积的网格代表节点选择策略, 以减少算法的能量消耗和提高查询成功率; 给出了基于目标矩形的方法以绕过不存在节点的区域, 从而避免查询处理过程出现中断. 仿真实验结果表明, RSA 算法优于现有的 IWQE 算法.

关键词: 无线传感器网络; 查询处理; 空间范围查询; 鲁棒性; 网格

中图分类号: TP393 **文献标识码:** A **文章编号:** 0372-2112 (2012) 07-1465-06

电子学报 URL: <http://www.ejournal.org.cn> **DOI:** 10.3969/j.issn.0372-2112.2012.07.029

Robust Spatial Window Aggregation Query Processing Algorithm in Wireless Sensor Networks

LIU Liang, QIN Xiao-lin, LI Bo-han, LIU Yu-lei

(College of Information Science & Technology, Nanjing University of Aeronautics and Astronautics, Nanjing, Jiangsu 210016, China)

Abstract: The current state-of-the-art spatial window aggregation query processing algorithms in wireless sensor networks do not take node movement, link failures and node failures into consideration in an integrated way, which results in low query success rate and large energy consumption. We propose a spatial window query processing algorithm called RSA which takes advantage of node redundancy to ensure the robustness of query processing. It divides the query region into several grids and collects the data of the nodes within each grid along a well-designed itinerary. An adaptive grid size setting method and a representative node selection strategy for the grids is put forward to reduce energy consumption and increase query success rate. We propose an approach based on target rectangle to bypass the region without nodes in order to avoid interruptions in the process of query processing. Experimental results show that RSA outperforms the existing IWQE algorithm.

Key words: wireless sensor network; query processing; spatial window query; robustness; grid

1 引言

无线传感器网络是以数据为中心的网络, 用户使用它的主要目的是查询其感知的数据. 在各种查询中, 典型的查询是空间范围聚集查询, 如获得当前时刻某大桥的最大压力值, 森林某区域的最大温度等. 考虑到传感器节点的能量由电池供应且通常无法更换, 能量十分有限. 因此, 能量高效的空间范围聚集查询处理算法是目前亟需研究的问题.

近年来, 学术界对传感器网络环境下的空间范围查询处理技术进行了一定的研究, 现有的算法可以分为两类: 基于全局树的算法^[1,2]、基于局部树的算法^[3~5]和基于路线的算法^[6]. 基于全局树的算法先广播查询消息至查询区域内的所有节点, 建立一棵以查询发起节点为根的树. 树中的非叶子节点接收其孩子节点的感知数据, 对这些数据和本地的感知数据进行聚集运算, 生成部分查询结果并将其发送到父节点, 如此继续, 直至到达查询发起节点, 生成最终的查询结果. 基于局部树的

收稿日期: 2010-01-23; 修回日期: 2012-02-21

基金项目: 国家自然科学基金 (No. 60673127); 国家高技术研究发展计划 (863 计划) 基金 (No. 2007AA01Z404); 高等学校博士学科点专项科研基金 (No. 20103218110017); 江苏省科技支撑计划基金 (No. BE2008135); 工信部电子信息产业发展基金; 中国博士后科学基金 (No. 20100481133); 江苏省博士后基金 (No. 1001005B); 江苏省“物联网与控制技术”优势学科建设项目

算法先路由查询消息至查询区域内的一个节点 h , 然后将查询区域内的节点组织成一棵以 h 为根的树, 利用基于全局树的算法相同的方法, 收集并聚集查询区域内节点的感知数据至节点 h , 最后 h 将最终的查询结果返回至查询发起节点. 相比基于全局树的算法, 基于局部树的算法仅将查询消息发送到查询区域内而无需在整个网络内广播, 减少了算法能量的能耗. 基于路线的算法 IWQE^[6] (Itinerary-based Window Query Execution) 的核心思想是沿一条或多条查询路线收集查询区域内节点的感知数据并进行聚集操作. 查询路线上的节点 (查询节点) 首先将查询消息广播至其邻居节点 (数据节点), 然后收集数据节点的感知数据, 最后对收集到的数据、本地的感知数据和从上一个查询节点接收到的部分查询结果进行聚集运算 (如计算最大值等), 计算出部分查询结果并将其发送至下一个查询节点, 如此继续. 收集到查询区域内所有节点的数据后, 利用位置路由协议^[7] 将最终的查询结果返回至查询发起节点. 另外, 文献[8,9]研究了连续近似聚集查询. 考虑到在许多现实应用中, 传感器感知的数据具有不确定性且用户能够接受查询结果有误差. 它利用节点感知数据具有时间相关性, 提出了一种自适应的设置节点误差配额的算法 (setting error budgets), 以在满足用户查询精度的前提下, 减少感知数据的发送.

基于路线的算法 IWQE 的查询路线是在查询处理过程中形成的, 因而相比前两类算法, 能有效地消除通信链路失效和节点移动的影响. 但 IWQE 算法还存在以下问题: (1) 节点失效问题. IWQE 的性能瓶颈在于查询节点. 当数据节点失效时, 则仅丢失了该节点的数据. 而当查询节点失效时, 则前面收集的查询结果均丢失且查询处理过程无法继续进行. (2) IWQE 要求节点维护其邻居节点信息且保证其正确性. 当查询节点维护的邻居节点信息不正确时, 查询节点 n_1 依据该信息选择的下一跳查询节点 n_2 可能不在 n_1 的其通信半径内, 导致基于路线的数据收集过程中断, 查询处理过程无法完成.

为了避免这些问题, 本文提出了一种利用节点冗余保证查询处理过程鲁棒性的空间范围聚集查询处理算法 RSA (Robust Spatial window aggregation query processing Algorithm). 它将查询区域划分为若干个网格, 每个网格中有一个代表节点用于收集所在网格其他节点的感知数据, 并将其发送至下一个网格, 直至收集到所有网格中的数据. 给出了一种自适应的网格大小设置方法和基于网格面积的网格代表节点选择策略, 以减少算法的能量消耗和提高查询成功率. 给出了基于目标矩形的方法以绕过不存在节点的区域, 从而避免查询处理过程出现中断. 仿真实验结果表明, RSA 算法在能

量消耗和查询成功率方面均优于现有的 IWQE 算法.

2 鲁棒的空间范围聚集查询处理算法

2.1 算法思想

如图 1 所示, RSA 将查询区域划分为若干个网格, 沿一条路线依次收集各网格内所有节点的数据. 查询处理过程如下: (1) 查询发起节点 S 以位于查询区域左上角的网格 $ABCD$ 的中心为目的位置, 利用无状态位置路由协议^[10,11] (即无需维护邻居节点信息的位置路由协议) 将查询消息发送至网格 $ABCD$ 中的所有节点. 其中, 网格 $ABCD$ 满足条件: 网格 $ABCD$ 中的节点均在彼此的通信半径内. 在位置路由由查询消息的过程中, 各中间转发节点判断矩形 $ABCD$ 内的所有节点是否都能收到该节点广播的消息. 若否, 则继续路由. 若是, 则停止路由并将查询消息广播至矩形 $ABCD$ 中的节点. 在图 1 中, 查询消息到达节点 R_2 后, R_2 将查询消息广播至网格 $ABCD$ 内的节点; (2) 网格 $ABCD$ 内的节点收到查询消息后, 选出一个代表节点用于接收该网格内非代表节点 (数据节点) 的数据, 并计算出到当前网格为止的部分查询结果. 在图 1 中, 网格 $ABCD$ 的代表节点为 Q_1 , 节点 a 、 b 、 c 、 d 为数据节点; (3) 为了避免无线通信的消息碰撞, 网格内的数据节点按一定的次序 (如顺时针方向) 依次广播其感知数据. 此时, 代表节点根据所在网格数据节点广播的感知数据和本地的感知数据, 能计算出到所在网格为止的部分查询结果, 而数据节点能计算出除代表节点外的部分查询结果. (4) 代表节点计算其下一个网格并将到所在网格为止的部分查询结果和查询消息广播至下一个网格. 代表节点的下一个网格是满足以下约束条件的网格: ① 下一个网格内的节点均在彼此的通信半径内; ② 下一个网格中的所有节点均在代表节点的通信半径内. 在图 1 中, Q_1 的下一个网格是网格 $CDEF$ (网格 $CDEF$ 中的节点均在 Q_1 的通信半径内且网格 $CDEF$ 中的节点均能互相通信). 网格 $CDEF$ 中的节点收到部分查询结果和查询消息后, 进行以下步骤: (1) 在网格 $CDEF$ 中选出一个代表节点; (2) 网格 $CDEF$ 中的数据节点依次广播其感知数据. 网格 $CDEF$ 的代表节点根据所在网格内数据节点广播的数据、本地的感知数据以及从上一个网格发送来的部分查询结果, 能计算出到所在网格为止的部分查询结果. 而网格 $CDEF$ 中的数据节点能计算出到所在网格为止、除代表节点外的部分查询结果; (3) 网格 $CDEF$ 的代表节点计算其下一个网格并将到所在网格为止的部分查询结果和查询消息发送至该网格. 如此继续, 沿图 1 所示的路线到达查询区域内最后一个网格后, 该网格的代表节点利用无状态位置路由协议将最终的查询结果返回至查询发起节点.

中的所有节点均能互相通信.为了减小 E_c, h_4 取最大值 $\min(d(P, S), d(P, U))$. 若以 Q_4 为圆心、 R 为半径的圆与线段 PT 不相交或相交于点 P 时,或者当 $d(Q_4, P) > R$ 时, Q_4 的下一个网格不存在.此时,采用基于目标矩形的方法,将当前的部分查询结果和查询消息发送至下一个存在节点的目标矩形.

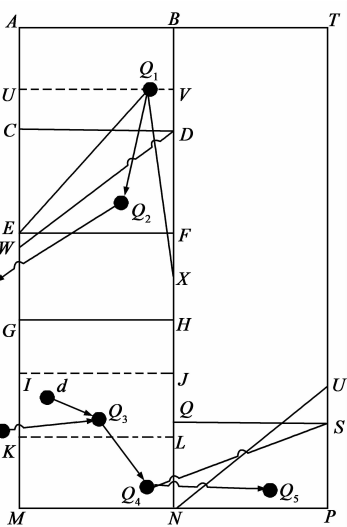


图3 计算下一个网格的大小

(3) 基于网格面积的网格代表节点选择策略

当前网格内的节点收到上一个网格发送的查询消息和部分查询结果后,分别广播其位置信息.此时,当前网格内的节点均能知道彼此的位置信息.根据该信息,当前网格内的任意一个节点都能计算出所在网格内其他节点下一个网格的面积.为了减少算法消耗的能量,选择当前网格内下一个网格面积最大的节点作为网格代表节点.值得注意的是,若节点维护其邻居节点位置信息且能保证其正确性,则节点无须广播其位置信息,直接根据维护的邻居节点位置信息判断是否成为所在网格的代表节点.

(4) 绕过不存在节点的网格

当节点密度较小或节点分布不均匀时,网格代表节点的下一个网格中可能不存在节点.以图 3 为例, Q_2 为网格 $CDEF$ 的代表节点, Q_2 计算出的下一个网格是矩形 $EFGH$.由于 $EFGH$ 中不存在节点,此时查询处理过程无法继续进行.为了避免该问题, Q_2 利用位置路由协议绕过不存在节点的下一个网格.即 Q_2 以目标矩形 $GHIJ$ (其高度 $d(G, I) = (R^2 - gw^2)^{1/2}$) 的中心为目的位置,利用位置路由协议将部分查询结果和查询消息发送至该目标矩形内的所有节点.消息被路由至节点 b 后,节点 b 发现无法将消息发送至目标矩形 $GHIJ$.此时,节点 b 计算新的目标矩形 $IJKL$ (其高度 $d(I, K) = (R^2 - gw^2)^{1/2}$).节点 b 利用位置路由协议将消息发送至该目标矩形内的节点.部分查询结果和查询消息到达节点 c 后,目标矩形 $IJKL$ 中的所有节点均能收到 c 广播的消息. c 将部分查询结果和查询消息发送至该矩形中的节点,继续查询处理过程.

3 实验分析

为了对算法的性能进行比较,我们用 JAVA 语言实现了一个用于仿真传感器网络查询处理算法的模拟器 NHSensorSim.在 NHSensorSim 上实现了 IWQE 和 RSA 算法.实验的硬件环境为 P4(3.0GHz) CPU, 512M 内存;软件环境为 Windows XP 操作系统、Eclipse 开发工具.实验的参数选择如下:根据文献[12],无线通信电路发送和接收 1byte 的能量消耗公式为 $E_t = \alpha + \gamma \times d^n$, $E_r = \beta$.采用文献[13]的参数: $\gamma = 10\text{pJ/bit/m}^2$, $\alpha = 45\text{nJ/bit}$, $\beta = 135\text{nJ/bit}$, $n = 2$.节点失效概率记为 p_f .算法的查询成功率用 p 表示, RSA 和 IWQE 的查询成功率分别记为 $p(\text{RSA})$ 和 $p(\text{IWQE})$.算法 RSA 和 IWQE 处理查询消耗的能量分别记为 $E_{\text{total}}(\text{RSA})$ 和 $E_{\text{total}}(\text{IWQE})$.其他参数如表 1 所示.

表 1 实验参数表

参数名	参数值
网络覆盖区域	$450 \times 450\text{m}^2$
网络生成方法	随机均匀分布
节点通信半径	52m
节点数	800
感知数据消息大小	40bytes
查询消息大小	20bytes
查询区域占网络覆盖区域的百分比	36%
节点失效概率 p_f	4%
网格宽度占通信半径的比率	$\sqrt{3}/2$
IWQE 算法路线宽度占通信半径的比率	$\sqrt{3}/2$

(1) 节点密度对查询成功率的影响.如图 4 所示,随着节点数目的增加, IWQE 算法的查询成功率的提高较小.而 RSA 的查询成功率有显著的提高.可见,当网络节点密度变大时,网格中的节点数目也随之增大.这些节点作为网格代表节点的备份节点,能有效地降低节点失效对查询处理的影响.

(2) 网格代表节点的选择对查询成功率的影响.在 RSA 算法的基础上,实现了 RSARandom 算法.与 RSA 算法不同, RSARandom 算法随机地在下一个网格中选择一个节点作为下一个网格的代表节点.如图 5 所示,相比

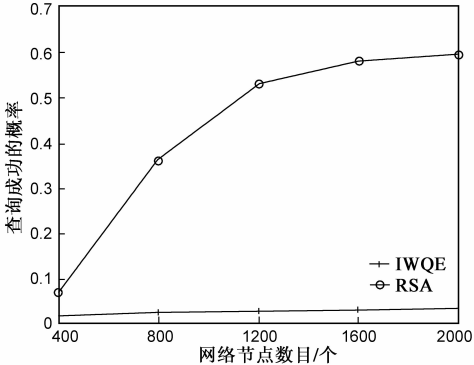


图4 $p_f=0.04$ 时节点密度对 p 的影响

RSARandom,本文提出的网格代表节点选择策略选择一个网格面积最大的节点作为网格代表节点,使得网格代表节点的备份节点数目增多,提高了查询的成功率.

(3)网络节点密度对能量消耗的影响.图 6 显示了节点不失效的情况.随着节点密度的增大,需要收集更多节点的感知数据,因而两种算法的能量消耗随之变大.相比 IWQE,RSA 算法采用了自适应的网格代表节点选择策略,查询区域内的节点只需发送一次感知数据消息或部分查询结果消息,并将 IWQE 算法中的发送查询消息至数据节点和发送部分查询结果与查询消息发送至下一个查询节点合并为一个过程,减少了数据通信量,因而,当 $p_f=0$ 时,RSA略优于IWQE.图7显示了

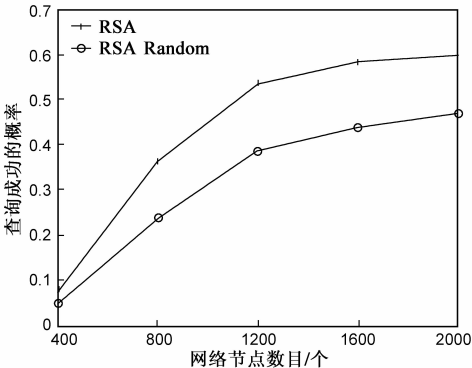


图5 $p_f=0.04$ 时网格代表节点的选择对 p 的影响

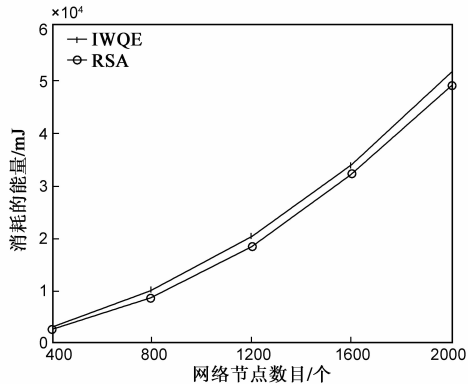


图6 $p_f=0$ 时网络节点密度对能量消耗的影响

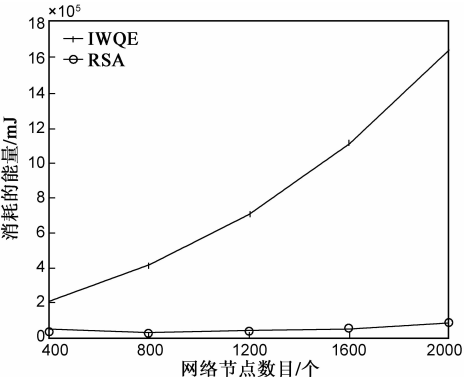


图7 $p_f=0.04$ 时网络节点密度对能量消耗的影响

$p_f=0.04$ 的情况,因为 $p(RSA) > p(IWQE)$ 且随着节点密度的增大, $p(RSA)$ 显著提高,而 $p(IWQE)$ 提高较小,所以RSA更节省能量.

(4)网格代表节点的选择对能量消耗的影响.图 8 显示了节点不失效的情况.可见,RSA 算法消耗的能量消耗略小于 RSARandom 算法.图 9 显示了 $p_f=0.04$ 的情况.随着节点数目的增多,需要收集更多节点的数据,算法的能量消耗均变大.但由于当 $N=400$ 时,算法的查询成功率较小,因而其能量消耗大于 $N=800$ 时消耗的能量.另外,由于 $p(RSA) > p(RSARandom)$,故RSA算法消耗的能量更少.

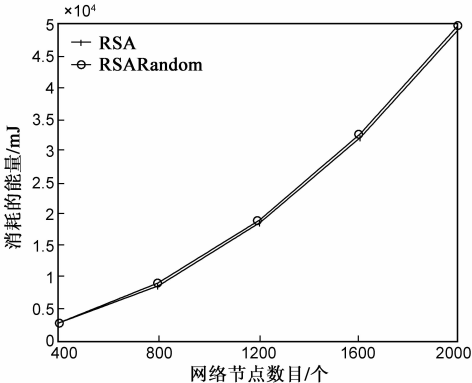


图8 $p_f=0$ 时网格代表节点对能量消耗的影响

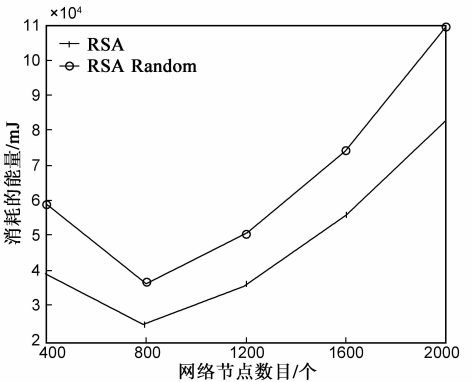


图9 $p_f=0.04$ 时网格代表节点对能量消耗的影响

4 结论

当节点不失效时,由于RSA算法采用了自适应的网格大小设置方法和基于网格面积的网格代表节点选择策略,并将IWQE算法中的发送查询消息至数据节点和发送部分查询结果与查询消息至下一个查询节点合并为一个过程,且查询区域内的节点只需发送一次感知数据消息或部分查询结果消息,减少了数据通信量,因而,RSA算法的能量消耗少于IWQE算法;当节点失效时,RSA算法利用冗余节点恢复查询处理过程,因而其查询成功率高于IWQE算法,且随着网络节点密度的增大,冗余节点数目增大,使得RSA的查询成功率显著提高,而IWQE的查询成功率提高较小.当查询区域变

大时, IWQE 的能量消耗急剧增大, 而 RSA 的能量消耗增幅较小. 当节点失效概率较大时, RSA 消耗能量远小于 IWQE 算法. 因此, RSA 算法优于 IWQE 算法.

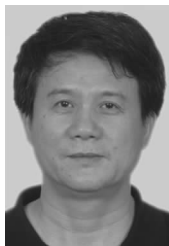
参考文献

- [1] S Madden, M J Franklin, J M Hellerstein, et al. Tag: A tiny aggregation service for ad-hoc sensor networks[A]. Proc of 5th Symposium on Operating Systems Design and Implementation [C]. New York: ACM Press, 2002. 131 – 146.
- [2] D Goldin, M Song, A Kutlu, et al. Georouting and delta-gathering: Efficient data propagation techniques for geosensor networks[A]. Proc of First GeoSensor Networks Workshop [C]. Portland, Maine, 2003. 73 – 95.
- [3] A Coman, M A Nascimento, J Sander. A framework for spatio-temporal query processing over wireless sensor networks[A]. Proc of the 1st Int'l Workshop on Data Management for Sensor Networks in Conjunction with VLDB 2004 [C]. Washington: IEEE Computer Society, 2004. 104 – 110.
- [4] A Coman, M A Nascimento, J Sander. Exploiting redundancy in sensor networks for energy efficient processing of spatiotemporal region queries[A]. Proc of the 14th ACM Conf. Information and Knowledge Management [C]. New York: ACM Press, 2005. 187 – 194.
- [5] 刘亮, 秦小麟, 戴华等. 能量高效的无线传感器网络时空查询处理算法[J]. 电子学报, 2010, 38(1): 54 – 59.
Liu Liang, Qin Xiao-Lin, Dai Hua, et al. Energy-efficient spatio-temporal query processing algorithm in wireless sensor networks[J]. Acta Electronica Sinica, 2010, 38(1): 54 – 59. (in Chinese)
- [6] Yingqi Xu, Wang-Chien Lee, Jianliang Xu, et al. Processing window queries in wireless sensor networks[A]. Proc of the 22nd International Conference on Data Engineering [C]. Washington: IEEE Computer Society, 2006. 70 – 80.
- [7] Karp B, Kung HT. GPSR: Greedy perimeter stateless routing for wireless networks[A]. Proc of the 6th Annual International Conference on Mobile Computing and Networking [C]. New York: ACM Press, 2000. 243 – 254.
- [8] N Jain, P Yalagandula, M Dahlin, et al. Self-tuning, bandwidth-aware monitoring for dynamic data streams[A]. Proc of the 22nd International Conference on Data Engineering [C]. Washington: IEEE Computer Society, 2009. 114 – 125.
- [9] N Jain, M Dahlin, Y Zhang, et al. STAR: Self-tuning aggregation for scalable monitoring[A]. Proc of the 33rd International Conference on Very Large Data Bases [C]. Washington: IEEE Computer Society, 2007. 962 – 973.
- [10] Yingqi Xu, Wang-Chien Lee, Jianliang Xu, et al. PSGR: Priority-based stateless geo-routing in wireless sensor networks [A]. Proc of the 2nd IEEE Int'l Conference on Mobile Ad-hoc and Sensor Systems [C]. Washington: IEEE Computer Society, 2005. 120 – 125.
- [11] M Zorzi, R R Rao. Geographic random forwarding (GeRaF) for ad hoc and sensor networks: Energy and latency performance[J]. IEEE Transactions on Mobile Computing, 2003, 2(4): 349 – 365.
- [12] T Rappaport. Wireless Communications: Principles and Practice[M]. New Jersey: Prentice-Hall Press, 1996.
- [13] M Bhardwaj. Power-Aware Systems [DB/OL]. http://www.mtl.mit.edu/researchgroups/icsystems/pubs/theses/manishb_sm_2001.pdf. 2001.

作者简介



刘 亮 男, 1985 年出生于江西省景德镇市, 2005 年在西北工业大学获得工学学士学位, 现为南京航空航天大学博士研究生. 研究方向为传感器网络数据库、流数据库的研究.
E-mail: liangliu@nuaa.edu.cn



秦小麟 男, 1953 年出生于江苏省南京市, 现为南京航空航天大学信息科学与技术学院教授, 博士生导师. 研究方向为分布式环境数据管理与容灾、安全数据库、时空数据库等.
E-mail: qinxcs@nuaa.edu.cn

李博涵 男, 1979 年出生于黑龙江省哈尔滨市, 南京航空航天大学计算机科学与技术流动站博士后. 主要研究方向为空间数据库的索引及查询优化, 传感器网络的多范畴时空查询等.

刘宇雷 男, 1974 年出生于江苏海门, 现为南京航空航天大学博士研究生, 主要从事传感器网络等方面的研究工作.