

C-Cast: A Location-Free Data Dissemination and Discovery Approach in Sensor Networks *

LI Zhigang^{*}, WU Junqing

(Institute of Command Automation, PLA Univ. of Sci. and Tech., Nanjing 210007, China)

Abstract: Peer-to-peer data dissemination and discovery is critical for ad-hoc wireless sensor networks. Most existing research depends on location information that is not always obtained easily, efficiently and accurately. We propose the concept of C-cast, a location-free data dissemination and discovery approach for large-scale wireless sensor networks. One important property of C-cast is that it does not depend on physical position or accurate localization services. The other advantage is that each node needs not to maintain too much global topology information. We prove that under the ideal scenario, C-cast can achieve 100% data retrieval success ratio, while under the random scenario, it can achieve 80% in reality. We evaluate C-cast thoroughly using metrics including data retrieval success ratio, storage cost, and load balance. Evaluation results show that C-cast can reach comparable functionalities and performance as the other approaches with physical location information.

Key words: sensor networks; data-centric storage and discovery; data retrieval; C-cast

EEACC: 7230

doi: 10.3969/j.issn.1004-1699.2012.07.023

C-cast: 传感器网络中无位置数据分发和发现技术 *

李志刚^{*}, 吴君青

(解放军理工大学指挥自动化学院, 南京 210007)

摘要: 在无线传感器网络中, 基于对等节点的数据分发和发现技术是一项关键的工作和研究问题。现有的研究技术和协议设计多数依赖于节点的位置信息, 但是在传感器网络中节点获取位置信息需要精确的定位算法和大量的计算、通信开销。为满足在没有位置信息的情况下, 节点依然能够有效地进行数据分发和发现工作, 本文提出了 C-cast 协议。C-cast 协议不仅不依赖于精确的物理位置信息和大功耗的定位服务, 而且节点不需要存储维护太多的全局拓扑信息。本文证明了在稠密网络的理想模型下, C-cast 协议能够达到 100% 的数据查询成功率; 在随机模型下, 通过选择跳步距离适当的两个信标节点, C-cast 协议能够达到 80% 以上的实际数据查询成功率。本文测试了 C-cast 协议的三种性能, 包括数据查询成功率、存储代价和负载均衡。测试结果显示 C-cast 协议达到甚至超过基于位置信息的协议性能。

关键词: 无线传感器网络; 数据存储发现; 数据查询; C-cast

中图分类号: TP301

文献标识码: A

文章编号: 1004-1699(2012)07-0985-08

近年来, 大规模无线传感器网络在很多关系到国计民生的应用中得到广泛的研究, 比如科学数据收集, 环境监测, 以及军事侦察等领域^[1-3]。一般来说, 传感器网络中节点的电池容量、计算能力和存储能力都是有限的, 传统网络协议由于需要较多的存储和较高的计算能力等原因所以一般不适合传感器网络的应用^[1]。设计轻量级的协议对传感器网络的实用性至关重要。在协议设计过程中, 目前的研究大多基于传感器网络符合服务器-客户端网络模式, 即在网络中部署一个静态的基站节点, 其他节点通过多跳传输, 周期性地向基站节点传送感知数

据^[3]。不过, 在一些应用中, 传感器节点之间需要进行对等的点对点数据传输和分发工作^[4]。

关于传感器网络中对等节点的数据分发和发现问题已经有一些相关的研究。现有的工作可以分成两类: 基于位置的策略和无位置策略。相关工作 TTDD^[5], GHT^[6], Double Rulings^[7] 和振荡轨迹^[8], 都属基于位置的策略, 如果每个节点都能获知其位置信息的话, 这些策略都可以取得良好的性能。在基于位置的策略中, 感知到数据的节点可以将数据存储在预先定义的位置附近某个节点 R 上(可以通过 GHT 获得这个位置信息), 同时其他节点可以通

项目来源: 江苏省自然科学基金项目(BK2011119)

收稿日期: 2012-01-21 修改日期: 2012-05-03

过访问该节点 R 来发现感知数据。但是定位算法^[9-10]的局限性限制了基于位置策略的应用。而在无位置策略中,假设每个节点不能获知其位置信息,也不能获知其他节点的位置信息,在没有全局信息的假设下,就不能预先知道哪一个节点需要查询数据,也不能确定将数据存储到什么位置。因此为满足有效的数据发现,需要在每个节点上存储维护大量的全局信息。但是维护全局信息,需要很多能量、计算开销和存储资源。如果不维护全局信息,而采用简单的泛洪策略和 rumor 路由策略^[11],同样会产生大量的冗余开销。本文的目标是设计一套轻量级的策略,允许在没有位置和不需要太多拓扑信息的条件下,进行数据的分发和发现。

本文提出的 C-cast 协议,只需要选择两个节点作为信标节点^[12],利用它们在网络中构建两个独立的轮廓覆盖网络(contour overlay network),在此基础上进行数据分发和发现工作。C-cast 可以达到基于位置信息的 Double rulings 策略的性能和效果。同时,协议的存储开销只是需要每个节点存储一个跳步计数对,而不用存储太多的信息。C-cast 需要的只是节点链接信息而不需要测距和定位。同时 C-cast 路由可以达到负载均衡,避免在网络中出现空洞^[13]。

1 轮廓覆盖网构建

1.1 网络模型

本文假设的网络是大规模稠密网络,节点按照泊松分布或者均匀分布部署在一个正方形的区域内,且在传感器网络所在的区域内任意画一条曲线 l ,到 l 距离小于某一阈值 d 的所有节点集合满足连通性。全网节点可以分为三类:数据消费者节点、数据生产者节点和中间节点。数据生产者节点为感知到数据的节点^[5-8]。数据消费者节点为在无线传感器网络中,在某一时间间隔 $[T_1, T_2]$ 中,发起数据查询操作的节点。中间节点是指用于数据消费者节点与生产者节点进行多跳数据转发和存储数据的节点。本文假设任何节点都有可能成为数据消费者节点或者数据生产者节点。数据消费者节点和数据生产者节点构成对等网络^[4-8,11]。

定义 1 数据查询成功率:数据消费者节点发出的查询消息能够发现数据生产者节点产生的数据或者数据摘要,则称为一次数据查询成功。否则称为一次数据查询失败。数据查询成功的次数占所有数据查询次数的比例称为数据查询成功率。

1.2 轮廓覆盖网

本文的工作,主要是利用两个信标节点来抽取

整个网络的拓扑和链接信息。为方便下文的论述,称其中一个信标为蓝色信标,另一个为红色信标。两个信标都向网络中其他节点广播跳步计数消息。其他节点在收到跳步计数消息以后,分别计算到两个信标的最短跳步数。本节主要描述如何构建和初始化两个轮廓覆盖网,暂时不考虑信标的选择过程,假设两个信标已经选定。关于如何选择信标在下面的章节会有描述。信标节点只在网络初始阶段起作用,在轮廓覆盖网构建以后,其能量消耗过程和其他节点是相似的。

首先为每个节点赋值一个跳步计数对 (b, r) , b 代表该节点到蓝色信标的跳步数,称为 *bluehop* 数; r 代表该节点到红色信标的跳步数,称为 *redhop* 数。 (b, r) 初始值为 (b_{\max}, r_{\max}) , 比如 b_{\max} 和 r_{\max} 都可以设置为 127。跳步计数消息, M 包括两部分,第一部分为颜色标识,用 1 个比特位来表示,比如说“1”代表蓝色消息,“0”代表红色消息;另一部分就是跳步计数器。当一个节点收到消息 M 的时候,首先检查 M 的颜色标识。然后比较当前跳步计数对中相同颜色标识的跳步数与 M 的计数器的大小。如果节点的跳步数大于等于 M . Hop+2, 节点的跳步数则修改为 M . Hop+1。在一个较短的时间内,节点有可能收到多个具有相同颜色标识的消息,那么节点只需要选择具有最小跳步数的消息即可。如果节点修改了某种颜色的跳步数,则向其邻居广播一个消息,通知其邻居节点它的当前跳步数。在几次迭代之后,所有的节点都会计算出到相应的信标的最短跳步数。

算法 1 轮廓覆盖网的建立

- 1 对每个节点初始化 $\text{bluehop} = 127, \text{redhop} = 127$;
- 2 (a) 蓝色信标广播消息 BM 给所有的邻居节点,其中 $\text{BM.hop} = 0, \text{BM.flag} = \text{blue}$; 蓝色信标的 $\text{bluehop} = 0$;
(b) 红色信标广播消息 RM 给所有的邻居节点,其中 $\text{RM.hop} = 0, \text{RM.flag} = \text{red}$; 红色信标的 $\text{redhop} = 0$;
- 3 (a) 当节点 u 接收到一个蓝色消息 BM, 如果 $(\text{BM.hop} + 1 < u.\text{bluehop})$ 则 $u.\text{bluehop} = \text{BM.hop} + 1$; 否则丢掉 BM。
(b) 当节点 u 接收到一个红色消息 RM, 如果 $(\text{RM.hop} + 1 < u.\text{redhop})$ 则 $u.\text{redhop} = \text{RM.hop} + 1$; 否则丢掉 RM。
- 4 (a) 节点 u 将 $\text{BM.hop} = \text{BM.hop} + 1$; 并且将新的 BM 发送给邻居节点。
(b) 节点 u 将 $\text{RM.hop} = \text{RM.hop} + 1$; 并且将新的 RM 发送给邻居节点。

以红色信标为例,所有的 redhop 为 k 的节点形成一个轮廓(contour), 记为 CR_k (对应于蓝色信标的轮廓记为 CB_k)。所有根据红色信标形成的轮廓称

为红色轮廓覆盖网。在稠密网络中,假设存在 k' ,使得所有位于轮廓 $CR_{k'}$ 的节点都是连通的,其中 $k < k'$, k 称为 CR_k 的半径;对于所有 $k > k'$ 的轮廓 CR_k 有可能分成几段,在这种情况下,可以通过边界上的点辅助进行路由,而且在正方形网络中 k' 与 k 差别很小。所以在下面我们都假设位于同一轮廓上的节点都是连通的。对 redhop 为 k ($k > 0$)的个体节点来说,如果 k 不是具有最大的 redhop 的节点,则其邻居可以分为三种。第一种是 k -hop 邻居,具有相同的 redhop 数,并同样位于相同的轮廓 CR_k 上;第二种为 $(k-1)$ -hop 邻居,为当前节点的上一跳节点;第三种为 $(k+1)$ -hop 邻居,为当前节点的下一跳节点。轮廓覆盖网建立以后,具有 (b, r) 节点简称为节点 (b, r) 。节点 (b, r) 属于两条轮廓,一条为蓝色轮廓 CB_b ,一条为红色轮廓 CR_r 。节点不需要存储轮廓上其他所有节点的信息。实际上可以将 (b, r) 看作一种虚拟坐标。不过相比欧式距离空间的位置坐标,该坐标不需要精确测量节点之间的距离,而是完全基于节点之间的链接关系,能够更准确的反映节点的连接和拓扑信息。节点可以利用该坐标来指导数据的分发存储和发现过程。算法1描述了轮廓覆盖网的建立过程,图1中表示随机选择的信标建立轮廓覆盖网以后的情况。事实上两个覆盖网是重叠的。假设所有的属于同一轮廓的节点都是连通的,可以定义下面一种新的数据转发协议。



图1 轮廓线的建立(信标节点及跳步数为偶数的节点用粗点表示)

定义2 C-cast:在轮廓覆盖网中,节点 (b, r) 将数据传送到轮廓 CB_b 或者 CR_r 的所有节点上,称为一次C-cast路由。C-cast路由可以用来分发数据副本和查询。

2 理想 C-cast 模型

在本章中,假设传感器网络部署在一个正方形的区域内,节点稠密且均匀分布。文献[12]的方法可以用来选择两个信标节点,即首先随机选择一个节点A,然后选择距离节点A最远的节点B作为红色信

标;选择到节点B最远的节点C,然后选择距离节点B和C之和最大的节点D为蓝色节点,文献[12]指出,在稠密均匀的正方形区域,该过程选择的两个信标节点位于正方形同一侧的顶点附近。在某些情况下和应用中,甚至可以直接指定两个节点作为信标。本节假设蓝色信标位于正方形区域的左上角,红色信标位于正方形区域的左下角,称具有这样性质的网络为理想C-cast模型,见图2。本章主要研究理想C-cast模型下的数据副本分发和查询过程。

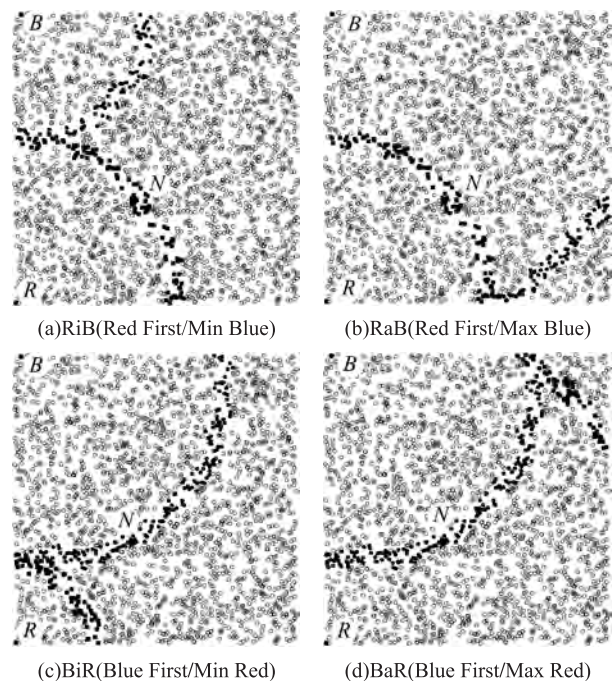


图2 数据副本的四种分发策略(在基本查询过程中也用到这四种策略)

2.1 数据副本分发过程

当节点 (b, r) 感知到数据以后,则触发数据副本分发过程。在该过程中,可以定义四种策略。分别为(a) RiB (Red First/Min Blue), (b) RaB (Red First/Max Blue), (c) BiR (Blue First/Min Red), and (d) BaR (Blue First/Max Red)。图2显示了这四种策略。每个节点采用每种策略的概率都为25%。不失一般性,用RiB策略作为例子说明数据副本分发的过程。在策略RiB中,节点 (b, r) 需要做两个决定。第一,决定首先沿红色轮廓 CR_r C-cast数据。第二,在数据中加入通知所有位于红色轮廓 CR_r 上的节点寻找具有极大bluehop的节点的消息。如果节点 $M(m, r)$ 发现自己的bluehop比所有邻居的bluehop都要小,则M发起沿蓝色轮廓 CB_m 的C-cast。数据在经过的节点上留下一个副本。该过程产生的红色轮廓和蓝色轮廓,由节点M连接在一起,形成一条路径,称为联合C-cast路径,简称为C-cast路径。设计四种策略的原

因是为了满足负载均衡。

2.2 数据查询过程

如果某个节点成为数据消费者节点,则需要启动查询过程。如果数据消费者节点采用类似于数据副本分发过程一样的策略,则称之为基本查询过程。除此之外,也可以利用其他的一些策略进行数据查询工作。

基本查询过程的思想是,当数据生产者 P 将数据副本利用以上四种策略中的任意一种分发到相应的 C-cast 路径上以后,另一个节点 Q 作为消费者节点需要查询该数据。节点 P 的 C-cast 存储路径包括一段红色轮廓 CR_r 和一段蓝色轮廓 CB_b 。基本查询过程利用以上四种策略中的任意一种分发查询信息。可以证明 C-cast 查询路径和 C-cast 存储路径是相交的,即 C-cast 查询路径和 C-cast 存储路径,至少共享一个节点。如果 C-cast 路径为连续曲线,则可以很容易证明任意两条路径都相交于一点。在实际情况下,C-cast 路径是离散的,而不是连续的。下面的定理 1 保证了在离散的情况下任意两条 C-cast 路径仍然存在相交的节点。

定理 1 在理想 C-cast 模型中,任意两条 C-cast 路径至少相交于一个节点。(证明略)

定理 1 说明,如果数据消费者节点需要发现期望数据,只需要检索所有在相同轮廓上的邻居即可,而不需要考虑所有的邻居,从而避免了向 $(k-1)$ -hop 邻居和 $(k+1)$ -hop 邻居转发数据。

除了基本查询过程,针对某些应用设计了双 C-cast 查询策略。在双 C-cast 查询中,节点 (b, r) ,在轮廓 CB_b 和 CR_r 上同时分发查询信息。

3 随机 C-cast 模型

出于网络实际情况的考虑,信标节点不一定位于网络中的特殊位置,而是允许出现在任何位置,这种情况称为随机 C-cast 模型。

3.1 信标选择

在随机 C-cast 模型中,依然假设网络区域为正方形区域。下面首先讨论信标的随机选择过程。

3.1.1 红色信标选择

首先,在网络中随机选择一个节点作为红色信标。然后红色信标向网络中的其他节点广播跳步计数消息。如果中间节点发现不能再向其他节点发送计数消息,则逆向往红色信标发送 ACK 消息。通过 ACK 消息,红色信标可以计算网络中最大的 redhop 数 K 。在没有全局调度的情况下,可能有几个节点同时发起红色信标竞争的消息。

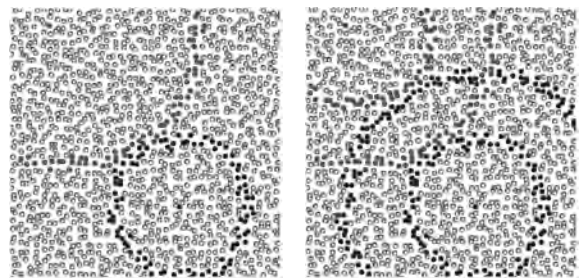
在这种情况下,这些节点可以按照时间戳或者节点 ID 进行竞争。

3.1.2 蓝色信标选择

蓝色信标的选择过程需要通过红色信标的辅助。当红色信标计算出最大 redhop 数 K 以后,随机按梯度路由选择一条路径发送关于 K 和 K' 的消息,其中 K' 为红色信标选取的不超过 K 的整数,当消息抵达某个 redhop = K' 的节点上,消息停止,并且该 redhop = K' 的节点选为蓝色信标。蓝色信标同样向网络中其他节点广播计数消息。同时利用 piggyback 消息将 K 和 K' 传送给其他节点。我们下面重点研究如何确定 K' 。

3.2 相切轮廓策略 TC

在随机 C-cast 模型中,采用相切轮廓 (tangent contours, TC) 数据副本分发策略 (如图 3 所示),TC 类似于理想模型中的 RiB 和 BiR。在 TC 策略中,数据生产者节点首先比较自己的 bluehop 和 redhop 数的大小。如果 bluehop 小于 redhop,数据生产者采用 BiR 策略。如果 redhop 小于 bluehop,数据生产者采用 RiB 策略。不过在随机模型中,数据查询过程采用同样的 TC 策略,并不能保证 100% 的数据查询成功率。数据查询成功率主要受两个信标之间的跳步距离 (即 K') 的影响。下面通过研究 TC 策略的最差情况,来分析信标之间的跳步距离与数据查询成功率的关系,并给出确定信标跳步距离的启发式算法。



(a)单C-cast路径

(b)双C-cast路径

图 3 随机 C-cast 模型的数据副本分发

3.3 TC 最差情况

3.3.1 网络区域与纺锤体区域

假设两个信标之间的跳步数记为 h ,理论上如果网络覆盖的区域无限大,所有半径小于 h 的轮廓会形成一个环,而且两种颜色的轮廓在两个信标之间的区域相互重叠形成一个纺锤体区域。在如图 4 中的纺锤体区域内,如果某个红色轮廓和蓝色轮廓相切,那么两个轮廓共享的节点满足 bluehop + redhop = h 。红色轮廓和蓝色轮廓合称为一个 BR 轮廓对。在图 4(a)中,任意两个 BR 轮廓对在纺锤体区域相交于两点,即假设所有的节点都落在该纺锤体之内,并且信

标节点位于两个圆心点上, 则可以证明在这种模型下数据查询成功率也可以达到 100%。不过真实的网络区域不是无限的, 而且也不一定是纺锤体的部署形状, 所以真实的网络区域和纺锤体区域不一定是重叠的(图 4(b))。下面论述纺锤体区域和网络区域的重叠程度决定了数据查询成功率。

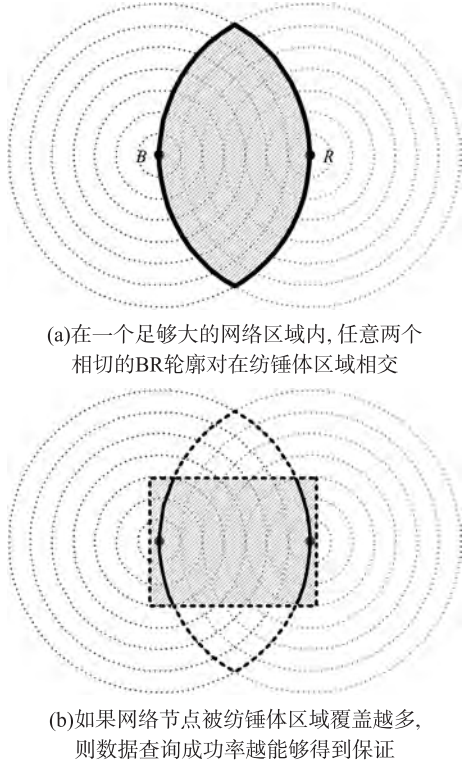


图 4

3.3.2 最差情况下的 K' 值选取

因为红色信标是随机选取的, 所以 K 值不是固定的, 不过可以证明 $K \in [K_d/2, K_d]$, K_d 为网络的直径。本节研究在 K 值未知的情况下如何确定 K' 值。

定义 3 信标扫视区域。分别过两个信标点作垂直于两个信标连线的直线。这两条直线将整个网络区域分成三个部分。其中位于两条直线之间的区域内的节点被两个信标连线分成两个子区域。在这两个子区域中, 其中覆盖节点面积较大的一个区域称为信标扫视区域。其面积记为 SA 。图 5(a) 中的四边形 $ABEF$ 为 AB 的信标扫视区域。

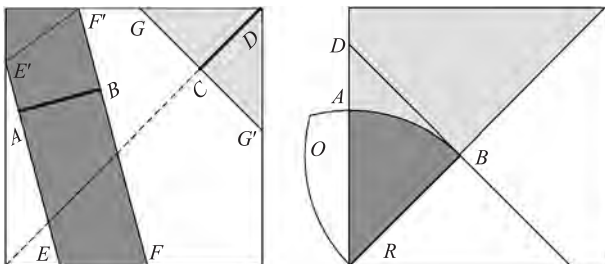


图 5 相切轮廓策略最差情况证明示意图

定义 4 信标扫视半纺锤体区域。两个信标生成的纺锤体区域被两个信标连线等分为两部分, 这两部分都称为信标扫视半纺锤体区域。图 5(b) 中的半个纺锤体 ORB 即为 RB 的信标扫视半纺锤体区域。

定理 2 假设两个信标之间的跳步距离为 h , 且 h 小于网络的直径 K_d 。当两个信标位于正方形网络区域的对角线上, 且其中一个位于正方形区域的顶点位置, 那么对于距离 h , 此时信标扫视区域的面积最小。

证明: 如图 5(a) 所示, 对任意线段 AB , 如果 $|AB| = |CD|$, 那么 $|EE'| + |FF'| \geq |GG'| \Rightarrow S_{EE'FF'} \geq S_{CC'DD'} = S_{GG'DD'}/2 \leq S_{EE'FF'}/2 \leq S_{AB}$ 。

下面的公式描述网络区域与纺锤体区域的重叠程度,

$$r = \frac{S_c}{S_N} \cdot \frac{S_c}{S_o} \quad (2)$$

其中 S_c 为信标扫视半纺锤体区域覆盖的信标扫视区域面积, S_N 为信标扫视区域的面积。 S_o 为信标扫视半纺锤体区域的面积。定理 2 描述了针对 h , 两个信标位置的最差情况。根据式(2)得出在最差情况下, 网络区域与纺锤体区域的重叠程度为,

$$r(h) = \begin{cases} \frac{\pi}{8} h^2 \cdot \frac{\pi}{8} & 0 \leq h \leq L \\ \frac{L^2}{2} \cdot \frac{\pi - \sqrt{3}}{3 - 4} & L < h \leq \sqrt{2}L \end{cases} \quad (3)$$

其中 L 近似为正方形区域边长的跳步距离。当 $0 \leq h \leq L$, $r(h)$ 的最大值约为 50%; 当 $L < h \leq \sqrt{2}L$, $r(h)$ 的最大值约为 52%, 其中 $h \approx 1.1L$ 。这意味着最差情况的下界当 $h \approx 1.1L$ 达到最大。因为红色信标是随机选取的, 所以 $K \in [K_d/2, K_d]$ 。 K 的期望值为,

$$E(K) = 4L \int_{\frac{1}{2}}^1 \int_{\frac{1}{2}}^1 \sqrt{x^2 + y^2} dx dy \approx 1.07L \quad (4)$$

当 $r(E(K)) \approx 51%$, 为期望最差情况的网络区域与纺锤体区域的重叠程度。

最差情况网络区域与纺锤体区域的重叠程度反映了任意一对 BR 轮廓的相交率的下界。在实际情况中相交率可以达到 80%。因为在上面的假设和公式中, 如果节点没有被信标扫视半纺锤体区域覆盖, 则认为不能发现其他数据也不能被发现。在实际情况中即使节点位于信标扫视半纺锤体区域以外, 仍然有可能经过绕路转发抵达信标扫视半纺锤

体区域之内。即使在轮廓被分成两个段的情况下,同样可以利用边界点进行绕路,以便达到数据发现的目的。即使在理想模型下,利用上面的公式得到的相交率也仅为 60%,所以说公式 3 的结果是非常保守的。

根据以上结果,提出如何确定 K' 值的启发算法。首先红色信标得到 K 值,并估计出 L 的值。如果 K 值小于 $1.1L$,则 K' 设为 K ;如果 K 值大于 $1.1L$,则 K' 设为 $1.1L$ 。下面给出如何估计 L 值的方法。

3.4 L 值估计算法

假设 L 值在网络部署以前没有分配给每个节点,根据式(4)红色信标节点需要估计 L 值的大小。如果网络是正方形区域且稠密的,则红色信标节点在收到的所有 ACK 中可以提取到 4 个 K 值,分别为 $K_1 < K_2 < K_3 < K_4$ 。如图 6 所示,下面的方程可以求解 L 的近似值。设 $x_1 < x_2 < x_3 < x_4$,则满足

$$\begin{aligned} x_1^2 + x_2^2 &= K_1^2 \\ x_1^2 + x_3^2 &= K_2^2 \\ x_2^2 + x_4^2 &= K_3^2 \\ x_3^2 + x_4^2 &= K_4^2 \end{aligned} \tag{5}$$

方程 5 可以解出 x_1, x_2, x_3, x_4 , 易得 $x_1 + x_4 = x_2 + x_3 = L$ 。

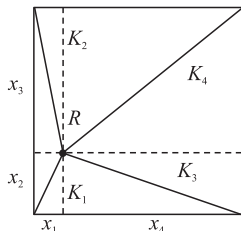


图 6 L 值估计

4 性能测试

本节讨论 C-cast 的模拟测试结果。因为理想 C-cast 模型在理论上能够保证数据查询成功率为 100%,所以只针对随机模型进行了模拟。在模拟中,考虑了三种性能指标:数据查询成功率、存储开销和负载均衡。虽然随机 C-cast 模型可以用于其它形状的网络中,但为了方便起见,只研究和模拟了在正方形区域网络的情况。使用扰动网格拓扑来部署传感器节点。在扰动网格拓扑网络中,每个传感器节点按照一定的扰动因子偏离原来的网格点,扰动因子的幅度为 1 m ~ 10 m。对于 C-cast,假设节点没有位置信息。为了模拟 GHT 和 Double rulings,假设节点能够通过定位算法来确定位置,定位的误差是随机产生的。节点的传输半径为 15 m。

4.1 数据查询成功率

随机 C-cast 模型中,首先我们通过模拟来验证

数据查询成功率受两个信标之间的跳步距离影响。在测试环境中部署了 1 600 个节点,并随机生成 50 个数据生产者和 50 个数据消费者节点。然后随机选择一个节点作为红色信标节点,在不同的红色轮廓上分别随机选择一个节点作为蓝色信标节点。对每组红色和蓝色信标,50 个数据生产者根据 C-cast 路径分发数据,50 个数据消费者利用基本查询过程检索数据。网络的直径为 39 跳。当两个信标的距离介于 20 到 30 之间时,数据查询成功率的平均值大于 75% (图 7)。

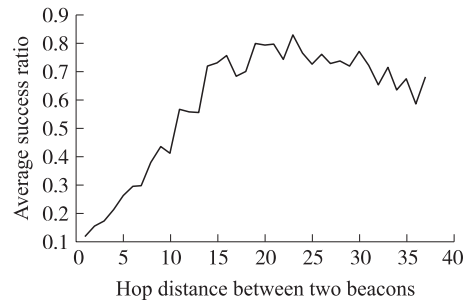


图 7 蓝色信标和红色信标的不同跳步距离影响了平均数据查询成功率

通过 3.3 节中的算法选择两个信标之间的跳步距离,并利用边界辅助查询策略与 GHT 和 rumor 路由进行比较。在 GHT 中,考虑了定位的误差。通过模拟发现(图 8),定位误差会严重的影响 GHT 数据查询成功率。这是因为在 GHT 中使用的是地理贪婪路由协议,即使在网络中不存在空洞,由于定位误差的影响,网络中也会产生一些虚拟的空洞,造成局部极小值的发生。我们将误差从 1m 逐步提高到 9m,发现 GHT 的平均数据查询率随着误差增大而降低。而 C-cast 和 rumor 路由不受定位误差的影响。rumor 路由的数据查询成功率平均为 67%,而 C-cast 的平均数据查询成功率可以达到 80%。因为 Double rulings 的机制和 GHT 是一样的,其数据查询成功率也会受到定位误差的影响。

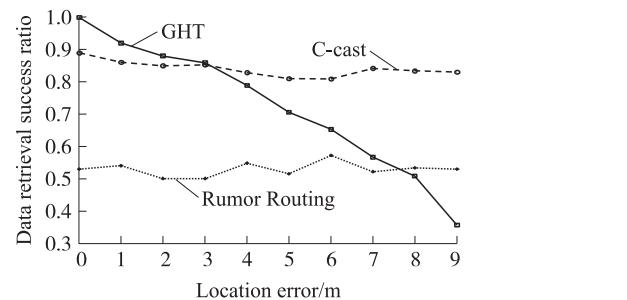


图 8 GHT, rumor 路由和 C-cast 在 ($K_d = 30, K = 25$) 时的数据查询成功率

4.2 存储开销

另一个值得关注的性能指标为存储开销(图 9)。

比较 C-cast 与 Double rulings 的存储开销。模拟不同规模的网络中存储开销。对每种规模,进行了 1 000 次测试。通过比较,C-cast 的存储开销要稍微大于 Double rulings,比如说在图 9 中,当网络为 3000 个节点的时候,C-cast 需要的存储节点个数为 170,而 Double rulings 为 140。不过在理论上 C-cast 和 Double rulings 所需要的存储节点的个数是同一规模的,即假设网络的节点规模为 N ,则一条存储路径的规模为 $O(N^{1/2})$ 。因此说 C-cast 在存储开销上和 Double rulings 是同一水平。

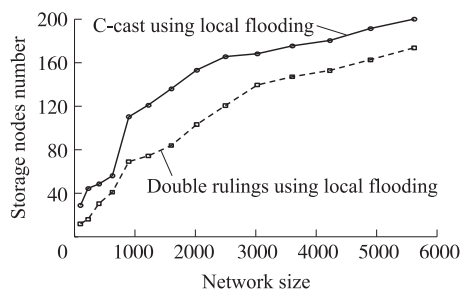
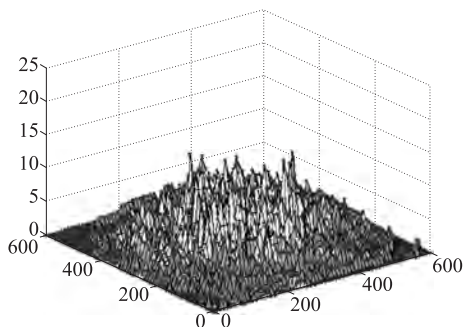


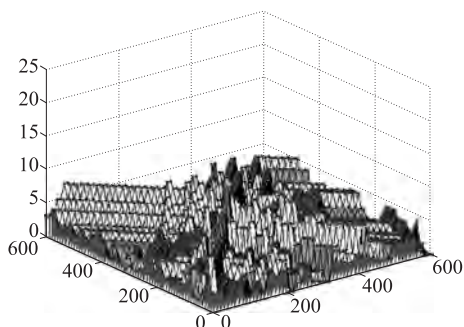
图9 一个数据生产者的存储节点的个数

4.3 负载均衡

利用网格网络拓扑来测试 Double rulings, C-cast 的负载均衡。网络的规模为 3600 个节点。随机产生 50 个数据生产者,然后分别利用 Double rulings 和 C-cast 来分发数据副本。图 10(a)为 Double rulings 的负载状况。Double rulings 中每个节点的存储小于 10,网络的中心位置的负载远远大于网络边缘的节点负载。图 10(b)说明了 C-cast 的负载状



(a) Double rulings 的负载



(b) C-cast 负载

图10 负载均衡比较

况,C-cast 中每个节点的负载也小于 10,但是在 C-cast 中整个网络的负载更加平均一些。

综上所述,通过模拟发现 C-cast 在没有位置的情况下,最关键的数据查询成功率是最好的,其余指标能够达到基于位置的策略相当的水平。

5 结论

针对大规模无线传感器中对等节点的数据发现和传输问题,本文提出了 C-cast 协议,一种无位置的数据存储和发现策略。该协议通过两个信标节点构建轮廓覆盖网,数据沿着轮廓根据不同的策略进行数据转发。该协议的优势是不依赖于节点的位置信息。本文提出了两种 C-cast 网络模型,理想 C-cast 模型和随机 C-cast 模型。C-cast 可以保证数据查询成功率,同时能够兼顾与基于位置策略等同的存储开销和负载均衡。未来的工作包括对 C-cast 中多数数据进行网络编码和在稀疏网络中的应用。

参考文献:

- [1] 孙利民,李建中,陈渝,等. 无线传感器网络[M]. 清华大学出版社,2005.
- [2] 卢汉良,李德骏,杨灿军,等. 深海海底观测网络信息采集监测系统设计与实现[J]. 传感技术学报,2011,24(3):407-411.
- [3] 樊镭,龚闻天,施晓秋. 大规模无线传感器网络数据收集与处理系统设计[J]. 传感技术学报,2011,24(11):1611-1616.
- [4] 陶孜谨,酆苏丹,徐金义,等. 大规模无线传感器网络中面向 ANY 型查询的能量高效数据分发算法[J]. 国防科技大学学报,2009,31(1):64-69.
- [5] Luo H, Ye F, Cheng J, et al. TTDD: Two-Tier Data Dissemination in Large-Scale Wireless Sensor Networks[J]. Wireless Networks, 2005, 11: 161-175.
- [6] Karp B, Ratnasamy S, Yin L, et al. GHT: A Geographic Hash Table for DataCentric Storage [C]//Proceedings of the First ACM international Workshop on Wireless Sensor Networks and Applications (WSNA) [C]//Atlanta, Georgia, USA, 2002.
- [7] Sarkar R, Zhu X, Gao J. Double Rulings for Information Brokerage in Sensor Networks [C]//Proceedings of the 12th ACM Annual International Conference on Mobile Computing and Networking (MobiCom) [C]//Los Angeles, CA, USA, 2006.
- [8] 李志刚,肖依,褚福勇. 大规模无线传感器网络中基于振荡轨迹的数据存储与发现机制[J]. 计算机研究与发展,2010,47(11):1911-1918.
- [9] Li M, Liu Y. Rendered Path: Range-Free Localization in Anisotropic Sensor Networks with Holes [C]//Proceedings of the 13th ACM Annual International Conference on Mobile Computing and Networking (MobiCom) [C]//2007.
- [10] Yang Z, Liu Y. Quality of Trilateration: Confidence-Based Iterative Localization, presented at IEEE ICDCS [C]//Beijing, China, 2008.

- [11] Braginsky D, Estrin D. Rumor Routing Algorithm for Sensor Networks [C]//Proceedings of the 8th ACM Annual International Conference on Mobile Computing and Networking (MobiCom) [C]//Atlanta, Georgia, USA, 2002.
- [12] Chessa S, Caruso A, De S, et al. GPS Free Coordinate Assignment and Routing in Wireless Sensor Networks [C]//Proceedings of the 24th IEEE Conference on Computer Communications (InfoCom) [C]//Miami, FL, USA, 2005.
- [13] Wu X B, Chen G, Sajal K Das. Avoiding Energy Holes in Wireless Sensor Networks with Nonuniform Node Distribution [C]//IEEE Transactions on Parallel and Distributed Systems [J]. 2008, 19 (5): 710-720.



李志刚(1980-),男,山东人,博士,解放军理工大学讲师,研究方向为无线通信与传感器网络,lzgnudt@gmail.com;



吴君青(1979-),男,浙江人,解放军理工大学讲师,研究方向为计算机网络。