

多 Sink 群智感知网络数据收集方法*

王继良^{1,2}, 黄丽嫦², 唐晖²

¹(长沙环境保护职业技术学院 环境信息系, 湖南 长沙 410004)

²(湖南大学 信息科学与工程学院, 湖南 长沙 410082)

通讯作者: 王继良, E-mail: 673303216@qq.com



摘要: 随着 Android 和 iPhones 等移动设备的广泛普及, 群智感知网络成为研究热点, 人们携带这些移动设备在日常生活中收集环境感知数据, 人们的移动具有社会性, 其移动轨迹难以预测, 如何设计一种有效的数据收集算法是一个值得研究的问题。针对多 Sink 群智感知网络, 提出一种基于地点的感知数据收集方法。首先采用多目标决策的层次分析法, 以移动设备与 Sink 节点之间的距离、连接时间和相遇概率为性能指标, 提出最优 Sink 的选择机制, 确定感知数据的目标传输节点; 受 PeopleRank 启发, 提出基于地点的数据转发方法, 以优化下一跳选择策略。最后, 通过一系列实验对该方法的可行性和有效性进行验证, 实验结果表明, 该方法不仅大大提高了感知数据的传输成功率, 而且转发开销和延迟有了明显的降低。

关键词: 群智感知; 数据收集; 数据转发; 机会网络

中文引用格式: 王继良, 黄丽嫦, 唐晖. 多 Sink 群智感知网络数据收集方法. 软件学报, 2016, 27(Suppl. (1)): 102-112. <http://www.jos.org.cn/1000-9825/16011.htm>

英文引用格式: Wang JL, Huang LC, Tang H. Data gathering scheme for multi-sink crowd sensing networks. Ruan Jian Xue Bao/ Journal of Software, 2016, 27(Suppl. (1)): 102-112 (in Chinese). <http://www.jos.org.cn/1000-9825/16011.htm>

Data Gathering Scheme for Multi-Sink Crowd Sensing Networks

WANG Ji-Liang^{1,2}, HUANG Li-Chang², TANG Hui²

¹(Environmental Information Department, Changsha Environmental Protection College, Changsha 410004, China)

²(College of Computer Science and Electronic Engineering, Hu'nan University, Changsha 410082, China)

Abstract: With the widespread popularity of sensor-rich mobile devices, such as Android phones and iPhones, crowdsensing networks have attracted enormous attentions from researchers recently. Mobile devices, carried by people in their daily life, move randomly in monitoring areas and acquire environmental sensing data. However, it is an arduous task to design an efficient data gathering approach for crowdsensing since the movements of mobile devices are difficult to predict. In this paper, a location based data gathering scheme for multi-sink crowdsensing networks is proposed. First, a selection method based on multi-objective decision-making is presented to find the optimal sink according to the distances, connection durations and encounter probabilities between mobile devices and sinks. Then, inspired by PeopleRank algorithm, a location based sensing data forwarding algorithm is presented to optimize the data forwarding strategy. Furthermore, the feasibility and effectiveness of the proposed scheme are validated through a series of experiments employing opportunistic network environment simulator. The results show that the proposed scheme not only improves the transmission rate of perception data but also has lower cost and forwarding delay than the existing methods.

Key words: crowdsensing; data gathering; data forwarding; opportunistic network

目前, 内嵌多种传感器的智能手机等移动设备呈爆炸式普及, 人们已经变成了一个“传感器”。人们随身携

* 基金项目: 湖南省科技计划(2013FJ4041); 湖南省自然科学基金(14JJ2051)

Foundation item: The Planned Science and Technology Project of Hunan Province (2013FJ4041); Natural Science Foundation of Hu'nan Provincial (14JJ2051)

收稿时间: 2016-05-31; 采用时间: 2016-09-29

带移动设备,随时随地进行感知,通过互联网进行协作,实现感知数据收集,形成群智感知网络^[1,2]。收集数据是群智感知网络最重要的任务之一。群智感知的一个重要特点就是“众人参与”,这个特点为群智感知应用的实用化带来了前所未有的机会,具体体现在:

(1) 城市中有大量的移动设备且移动设备基本都具备蓝牙或 Wi-Fi 等短距离无线通信技术,无须专门部署,感知成本低。

(2) 人的流动带来移动设备的移动,促进了感知覆盖与感知数据的传输。

群智感知网络数据收集最直接的方法是移动设备借助蜂窝网络(如 GSM、3G/4G)直接与互联网相连接,将感知数据上传到数据中心。然而,对需要收集大量感知数据的群智感知应用而言,一方面,会消耗用户较多的网络流量和电量,收集成本高导致用户参与热情不高;另一方面,大量用户参与群智感知会给蜂窝网络带来额外的负载。

机会传输是群智感知网络数据收集的另一种研究思路。机会传输的思路来源于机会网络^[3,4]。其中,移动设备借助人们相遇的机会,基于蓝牙、Wi-Fi 等短距离无线通信技术实现数据交换。机会传感不依赖蜂窝网络,节省了用户的通信流量。文献[5]提出 MV Routing 数据路由算法,利用节点访问某个区域的历史记录来计算相遇概率,并以该概率为依据有选择地转发数据分组。文献[6]提出 PROPHET 路由算法,基于概率策略,预先估计到达目的节点每条链路的传送概率,并依此来进行消息转发。然而,人们相遇的时间短暂,彼此的设备很难建立稳定可靠的连接,数据传输的成功率低。

近年来,借助于地点来降低连接机会的不确定性成为研究热点。人们在地点处有足够长的时间对感知数据进行转发和选路。Link^[7]等人假设携带移动终端的人群会在一些固定的地点随机分布,预测出节点接触目的节点的概率,从而决定该节点是否适宜作为中继节点进行数据传输;但是此种路由方案过于依赖移动节点移动模型的精度,所以不能够准确表示节点相遇带来的转发机会。Vaidya 等人^[8]提出将数据发送到某个地点,节点会记录与其他节点的接触情况,同时要记录其他节点访问的地点,并以此为基础决定数据如何转发。但节点学习其他节点的行为涉及隐私问题。Daly 等人^[9]提出,社会网络结构中获取节点与目的地点的节点之间的社会相似性,将高相似性的节点作为传输数据的中继节点,但其需要先验知识,进行一系列的数据挖掘操作造成运算复杂。Lu 等人^[10]提出了 LOOP 算法,每个节点记录本身的移动轨迹,每个节点同时学习自己移动轨迹中隐含的规律性和固定模式。在机会性接触期间,节点依据学习的移动规律计算自身发送邻居节点数据到其目的地的概率并通知邻居节点,邻居节点根据所收到的概率选择相应的转发节点。

实际生活中,城市中有很多地方部署了 AP(access point),可以免费接入 Internet。携带智能设备的人们在指定的监测区域内收集感知数据,每一部智能手机内置的物理传感器可以获取监测区域内的经度、纬度、光线强度、温度、噪音等监测数据。当人们与 AP 相遇时,就将感知数据上传到数据中心,从而可以避免产生过多的流量和电量,同时也可减轻蜂窝网络的负载压力。人类的移动与地点息息相关^[11],本文研究的是一种基于地点的群智感知网络数据收集方法。与传统机会网络不同的是:(1) 移动设备并不感知数据的目标节点,本文将 AP 比拟成 Sink,感知数据只要传输到任意一个能免费接入互联网的 Sink 就可以将数据上传到数据中心;(2) 移动设备不需要实时地将感知数据上传,而是选择适当时机转发。因此本文首先确定感知数据的目标节点;其次考虑移动设备之间应基于什么样的路由方法来转发感知数据,其中选择合适的中继转发节点来转发感知数据显得尤为重要。

鉴于此,本文提出一种基于地点的多 Sink 群智感知网络数据收集方法。首先,针对多个 Sink 的群智感知网络的感知数据收集,基于多层次分析法的 Sink 选择机制,选择出一种相对最优的 Sink 作为目标 Sink;其次,利用人类的移动规律作为是否转发数据的参考依据,提出了一种基于地点的感知数据转发方法。最后将选择最优 Sink 机制和基于地点的转发方法在模拟器 ONE^[12]中实现,对湖南大学群智感知网络场景进行了多次模拟,并与现有的路由算法进行了比较。结果表明,本文提出的方法不仅提高了传输成功率,而且消息转发开销和转发延迟有了明显的降低。

1 基于地点的感知数据收集方法

本节首先给出问题的定义和假设,再对移动传感节点与 Sink 空间、时间、相遇概率属性的度量指标加以分析,引入多目标决策的层次分析法^[13],求解出最优的 Sink,即确定感知数据的目标节点.然后分析人类的社会属性,受 PeopleRank^[14]启发,在此基础上提出基于地点的感知数据路由方法.

1.1 最优Sink选择方法

1.1.1 问题定义和假设

在本文中,携带移动设备的人类被抽象成“传感器”节点.城市监测区域内分布多个访问点(AP),我们将 AP 比拟成传感器网络中的 Sink.对于移动的“传感器”节点而言,需要选择一个最优 Sink 作为感知数据的目标节点.我们将对移动的“传感器”节点和 Sink 之间空间和时间属性以及相遇概率进行分析,计算移动的“传感器”节点与 Sink 之间相关属性的指标值,利用层次分析法求解出一个最优 Sink.

针对此背景,我们给出如下假设:

(1) 网络由 m 个移动的“传感器”节点 $S(S_1, S_2, \dots, S_i, \dots, S_m)$ 和 n 个 Sink($Sink_1, Sink_2, \dots, Sink_n$) 组成,其中 m 远大于 n ;

(2) S_i 可以任意移动,在本文中以人类作为载体,且具备短距离通信能力;

(3) Sink 是静止的,在本文中就是部署了 Wi-Fi 热点的 AP 地点;

(4) S_i 的能量和缓存大小有限,然而 Sink 是 AP 接入点,能量和缓存不受限制;

(5) S_i 可以知晓自己的位置信息,并可以记录下自己的位置;

(6) 同一个消息的所有拷贝具有相同的消息 ID.

1.1.2 基于空间属性的度量指标及数据分析

在移动群智感知网络中,由于移动传感节点 S_i 的位置在不停地变动,导致 S_i 与 $Sink_j$ 之间的网络拓扑结构也随之变化.本文以“间距”作为研究空间属性的度量指标.

群智感知网络中分布着 m 个移动的“传感”节点 S_i 与 n 个 $Sink_j$,彼此之间的间距不同,用 $d(S_i, Sink_j)$ 表示 S_i 与 $Sink_j$ 之间的路径间距. r 表示 $Sink_j$ 的通信半径,若携带移动设备的人们(“传感器”节点)进入 $Sink_j$ 的通信范围,则可由 $Sink_j$ 收集其感知数据.一般情况下, S_i 与 $Sink_j$ 的间距远远大于其通信半径 r .

当两个移动的“传感器”节点 S_1, S_2 相遇时,假设两者的位置坐标为 (S_{1x}, S_{1y}) 和 (S_{2x}, S_{2y}) ,则计算出两者之间的间距为 $\sqrt{(S_{1x} - S_{2x})^2 + (S_{1y} - S_{2y})^2}$.一般来说,该间距与 Sink 的间距相比非常小,可将 S_1 和 S_2 看成一个点,与 $Sink_1, Sink_2, \dots, Sink_n$ 的距离分别为 $d(S_i, Sink_1), d(S_i, Sink_2), \dots, d(S_i, Sink_n)$.我们用 (S_{ix}, S_{iy}) 表示 S_i 的坐标, $(Sink_{jx}, Sink_{jy})$ 表示 $Sink_j$ 的坐标,间距计算如式(1)所示.

$$d(S_i, Sink_j) = \sqrt{(S_{ix} - Sink_{jx})^2 + (S_{iy} - Sink_{jy})^2} \quad (1)$$

显然,若只考虑间距这一个指标,收集数据的最佳 Sink 则为距离 S_i 最近的 Sink,如式(2)所示.

$$Sink = \{Sink_m \mid d(S_i, Sink_m) < d(S_i, Sink_j), \forall m \neq j\} \quad (2)$$

例如,图 1 中有 3 个 Sink 节点,计算出 $d(S_i, Sink_1) < d(S_i, Sink_2) < d(S_i, Sink_3)$,则优先考虑间距最小的 $Sink_1$ 节点作为目标 Sink.当 S_1 和 S_2 相遇时,以 $Sink_1$ 为感知数据的目标节点,分析 S_1 和 S_2 的下一跳,若 S_1 的下一跳地点为 $Sink_1$, S_2 就将感知数据转发给 S_1 .

1.1.3 基于时间属性的度量指标及数据分析

在以人为主体的移动群智感知网络中,移动的“传感器”节点 S_i 通信范围有限, S_i 与 $Sink_j$ 的通信链路频繁断接,使得感知数据在传输的过程中丢失,所以移动的“传感器”节点 S_i 与 $Sink_j$ 之间建立连接的持续时间可以反映

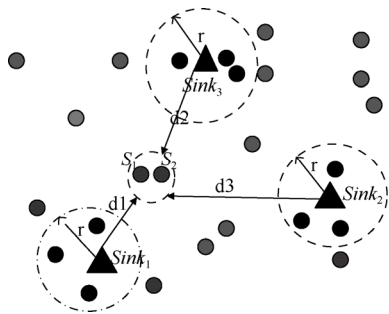


Fig.1 S_1 meets S_2

图 1 S_1 与 S_2 相遇

出该连接的连接强度.同理, S_i 与 $Sink_j$ 之间断开的持续时间可以反映出该两者之间没有连接的强度.期望传输延迟表示移动的“传感器”节点 S_i 与 $Sink_j$ 直接相遇需等待的时间,本文以期望传输延迟,即“传感器”节点 S_i 与 $Sink_j$ 直接相遇需等待的时间作为研究时间属性的度量指标.

假设在某一段时间 t 内,移动的“传感器”节点 S_i 与 $Sink_j$ 的链路状态如图2所示,设 $DC_1, DC_2, \dots, DC_i$ 表示时间 t 内移动的“传感器”节点 $S_i(1 \leq i \leq m)$ 与 $Sink_j$ 断开连接的持续时间, $W_{ij}(t)$ 表示移动的“传感器”节点 S_i 与 $Sink_j$ 的传输等待时间,用 $L_{S_i}(Sink_j)$ 表示 S_i 与 $Sink_j$ 的期望传输延迟,则 $L_{S_i}(Sink_j)$ 的计算公式如式(3)所示,其中, N 表示时间 t 内链路断开的次数.

$$L_{S_i}(Sink_j) = \int_0^t \frac{1}{t} W_{is}(t) dt = \frac{1}{t} (W_{is}(t)) = \frac{\sum_{k=1}^N \frac{1}{2} DC_k^2}{t} = \frac{1}{2t} \sum_{k=1}^N DC_k^2 \quad (3)$$

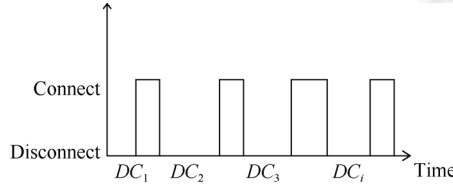


Fig.2 Link state

图2 链路状态

显然,若只考虑期望传输延迟这一个指标,收集数据的最佳 Sink 为期望传输延迟最小的 Sink,如式(4):

$$Sink = \{Sink_m | L_{S_i}(Sink_m) < L_{S_i}(Sink_j), \forall m \neq j\} \quad (4)$$

1.1.4 基于相遇概率的度量指标及数据分析

在移动群智感知网络中,移动的“传感器”节点 S_i 随机移动.除了间距和期望传输延迟,在选择最优 Sink 时也应考虑 $Sink_j$ 与所有移动的“传感器”节点 S_i 之间的相遇概率,依据历史相遇信息,加以计算出相遇概率.相遇概率定义为不同的 $Sink_j$ 与 S_i 相遇的概率大小,用 $P_{S_i}(Sink_j)$ 表示.相遇概率的计算公式如式(5)所示,其中, N_j 表示时间 t 内移动传感节点 S_i 与 $Sink_j$ 的相遇次数, n 表示 Sink 的个数.

$$P_{S_i}(Sink_j) = \frac{N_j}{\sum_{i=1}^n N_i} \quad (5)$$

显然,若只考虑相遇概率这一个指标,收集数据的最佳 Sink 则为相遇概率最大的节点,如式(6)所示.

$$Sink = \{Sink_m | P_{S_i}(Sink_m) < P_{S_i}(Sink_j), \forall m \neq j\} \quad (6)$$

1.1.5 最优 Sink 选择机制

在前面小节中,我们已经对移动的“传感器”节点 S_i 与 $Sink_j$ 之间的空间属性、时间属性、相遇概率属性进行了分析.显然,若单独考虑这3个属性很容易选择出最优 Sink,但是联合考虑就很困难.因为彼此的度量单位不一致,需要将其化为无量纲的数量标度,并且按统一的规则进行归纳与综合,才使得不同的目标对应的结果值具有可比性的数量关系.

本小节研究如何将3个属性联合考虑,求解出最优 Sink.将多个 Sink 看成是多个目标,即将选择最优的 Sink 问题可以看成多目标最优化求解问题.我们从以下4个步骤来阐述其求解过程.

(1) 建立层次结构模型

首先我们根据实际应用建立起基于多 Sink 的层次结构模型,包括3层结构,即目标层、准则层和方案层.目标层所要达到的目标是在基于多 Sink 的移动群智感知网络场景中针对移动的“传感器”节点 S_i 在 $Sink_j(1 \leq j \leq n)$ 中选择出最优的 Sink.为了这个目标,我们评价方案时所要考虑到的准则是移动的“传感器”节点 S_i 与 Sink 的间距、期望传输延迟和相遇概率.最后在方案层中,给出解决方案:选择 $Sink_i(1 \leq i \leq n)$ 中的一个 Sink 作为最优目标

Sink.

(2) 构造判断矩阵

其次通过相互比较确定准则层中各准则对于目标的权重,来构造判断矩阵.判断矩阵标度取值和含义见表1.准则层有3个准则 C1:间距,C2:期望传输延迟,C3:相遇概率.构造判断矩阵 $A_{3 \times 3}$,其中判断矩阵中的 A_{ij} 表示两个要素的相对重要性的数量尺度, A 称作判断尺度矩阵.

Table 1 Matrix judgment

表 1 判断矩阵标度

标度(A_{ij})	含义
1	两两对比,同等重要
3	两两对比,前者稍微重要
5	两两对比,前者明显重要
7	两两对比,前者强烈重要
9	两两对比,前者极端重要
2,4,6,8	两者的中间值

$$A = \begin{matrix} & \begin{matrix} C_1 & C_2 & C_3 \end{matrix} \\ \begin{matrix} C_1 \\ C_2 \\ C_3 \end{matrix} & \begin{pmatrix} A_{11} & A_{12} & A_{13} \\ A_{21} & A_{22} & A_{23} \\ A_{31} & A_{32} & A_{33} \end{pmatrix} \end{matrix}$$

两两比较打分,在 n 个 Sink 中选择目标 Sink 时,设定好判断尺度矩阵.分别根据3个准则针对不同目标建立起相关矩阵. B_1 是相对间距的尺度矩阵,对间距而言,间距越小越好.

$$B_1 = \begin{matrix} & \begin{matrix} Sink_1 & Sink_2 & Sink_3 & \dots & Sink_n \end{matrix} \\ \begin{matrix} Sink_1 \\ Sink_2 \\ Sink_3 \\ \dots \\ Sink_n \end{matrix} & \begin{pmatrix} B_{11} & B_{12} & B_{13} & \dots & B_{1n} \\ B_{21} & B_{22} & B_{23} & \dots & B_{2n} \\ B_{31} & B_{32} & B_{33} & \dots & B_{3n} \\ \dots & \dots & \dots & \dots & \dots \\ B_{n1} & B_{n2} & B_{n3} & \dots & B_{nn} \end{pmatrix} \end{matrix}$$

矩阵 B_1 中的元素 B_{jk} 计算如式(7)所示,其中, $d_{S_i, Sink_k}$ 表示节点 S_i 和 $Sink_k$ 之间的距离. B_2 是相对期望传输延迟的尺度矩阵.对相遇传输延迟而言,取值越小越好. B_3 是相对于相遇概率的尺度矩阵,对相遇概率而言,取值越大越好. B_2 和 B_3 中元素值的计算类似于 B_1 ,这里不再赘述.

$$B_{jk} = \frac{d_{S_i, Sink_k}}{d_{S_i, Sink_j}} \quad (7)$$

(3) 层次排序

对于 $n-1$ 层的某因素,第 n 层的各因素按照重要性进行排序.具体计算是:对于判断矩阵 B_i ,计算需满足 $B_i w_i^{(3)} = \lambda_{\max} w_i^{(3)}$;对于判断矩阵 A ,计算需满足 $A w^{(2)} = \lambda_{\max} w^{(2)}$.这里, $\lambda_{\max}$ 取矩阵特征值的最大值, $w^{(2)}$ 或 $w^{(3)}$ 是对应于 $\lambda_{\max}$ 的正规化的特征向量,它们的分量即是相应元素单排序的权值.

对判断矩阵 A ,对应于 $\lambda_{\max}$ 的正规化的特征向量为 $w^{(2)}$;相对于间距的判断矩阵 B_1 而言,对应的特征向量为 $w_1^{(3)}$,同理,相对于期望传输延迟和相遇概率的判断矩阵 B_2 和 B_3 而言,所对应的特征向量分别为 $w_2^{(3)}$ 和 $w_3^{(3)}$.以 $w_1^{(3)}$ 、 $w_2^{(3)}$ 和 $w_3^{(3)}$ 为列向量构成矩阵 $w^{(3)}$.

(4) 最优 Sink 求解

最后,进行层次总排序,求出不同的结果,选择最优的 Sink 作为目标 Sink,各个方案优先程度的排序向量计算如式(8)所示.

$$w = w^{(3)} w^{(2)} \quad (8)$$

若有 n 个 Sink, w 是一个 n 行 1 列的矩阵,其中 $w_1, w_2, \dots, w_n$ 分别为 n 个 Sink 的结果,若第 j 行的值最大,则取

$Sink_j$ 为目标 Sink,公式如式(9)所示.

$$Sink = Sink_j \{j \in \max w_j (1 \leq j \leq n)\} \quad (9)$$

1.2 基于偏好地点的感知数据转发方法

1.2.1 基于地点的群智感知网络构建模型

人类在他们的日常社会生活中,不断地在不同的地点之间相遇或移动,发现他们会访问不同的地点,但却会花费大量的时间在一些他们喜欢去的地点,这为数据传输的高可靠性奠定了基础.例如,校园中的学生 A、B、C 有共同的偏好地点 P 食堂,学生 A、B、C 随身携带着智能手机并根据监测任务采集感知数据,在时刻 t ,他们在 P 食堂相遇,并停留较长的时间.在此期间,学生 A、B、C 彼此建立稳定的连接,能在食堂内将感知数据进行交换.以学生 A 为例,根据学生 B 和 C 的移动规律决定将自身携带的感知数据转发给 B 还是 C.

把人类日常运动中的共有且稳定的特征提取出来,能发现出人类的移动与地点息息相关,人类活动具有与地点密切相关的规律性和重复性,每个人都会有自己固定的工作地点、生活地点、休闲娱乐地点.因此,分析节点的社会关系特征并利用节点的移动规律作为是否转发数据的参考依据.

正如我们所知,人类携带个人移动设备(如智能手机)在多个地点运动,移动设备具备存储-转发能力.其中,移动设备收集环境感知数据,且这些感知数据最终要上传到服务器,所以可以将部署了 Wi-Fi 热点的 AP 地点作为 Sink,即作为感知数据的“目标地点”.如果移动设备移动到某个部署了 Wi-Fi 热点的地点,则携带的感知数据可以直接由 Wi-Fi 上传到服务器.但是事实上只有少量的地点部署了 Wi-Fi 热点的 AP,则在遇到 AP 接入前,移动节点在地点停留较长时间并可以建立稳定连接,所以本文假定移动节点只在地点处彼此通过机会路由转发数据.

1.2.2 基于地点感知数据转发方法

本文的数据转发方法是建立在多 Sink 机会移动传感网络中,我们选择一个最优的 Sink 作为感知数据的最终接收者.在移动的“传感器”节点与 Sink 未相遇之前,将感知数据的转发和地点结合起来,在长时间停留的地点进行数据的充分交换,才能高效地收集感知数据.

在能够形成稳定拓扑结构的地点内,移动的“传感器”节点彼此会建立稳定的连接,建立连接的目的是为了将感知数据转发出去.最简单的方法是将感知数据转发给相遇的每一个节点,会提高感知数据的传输成功率,但却造成消息冗余,且开销很大.在此背景下,我们应该优化下一跳选择策略.本节对移动的“传感器”节点 S_i 如何选择中继节点来转发感知数据进行完整描述.

假设移动的“传感器”节点 S_i 与 k 个移动的“传感器”节点($S_1, S_2, \dots, S_k$)相遇,我们需要考虑将 S_i 的感知数据转发给哪个节点.在本文中,人作为社会的一份子总会在某些地点频繁出现,这也使得人随身携带的移动设备会稳定地偏向于访问某些地点,我们将这些地点定义为偏好地点 L ,偏好地点的受欢迎程度表示人们对偏好地点的偏好程度,用 $Pop(L_i)$ 来表示偏好地点 L_i 的受欢迎程度.其中,一个地点的受欢迎程度,可以用很多指标来衡量.在文献[15]中提出地点的受欢迎程度由两个要素所决定:出度和人口量;地点的受欢迎程度排名=出度排名 $\times$ 人流量排名;即地点 L_i 的受欢迎程度值计算如式(10)所示.

$$Pop(L_i) = L_{cvg}(L_i) \times P_{opu}(L_i) \quad (10)$$

人倾向于访问这些地点,就有越多的转发候选者可以选择,感知数据转发的机会越高,在本文我们只考虑地点的人流量,如式(11)所示.

$$Pop(L_i) \approx P_{opu}(L_i) \quad (11)$$

在这里,我们对多个偏好地点进行排序,即考虑到移动设备下一跳偏好地点信息,根据偏好地点的受欢迎程度给不同的移动节点赋予不同的优先级,将感知数据转发给优先级高的移动节点.

如表 2 所示,在 $T1$ 时,传感节点 S_i 运动到地点 L_0 处,并与 S_j, S_k 节点相遇,考虑移动节点 S_i 要选择哪个移动节点将感知数据转发:考虑与其相遇的 S_j, S_k 下一跳要访问的偏好地点,看是否为最优 Sink.

Table 2 Node i meets node j, k **表 2** 节点 i 与 j, k 相遇

Node	Time	Location	Next location
i	T_1	L_0	NULL
j	T_1	L_0	L_B
k	T_1	L_0	L_C

如果是,转发给它;如果都没有,就考虑下一跳地点的受欢迎程度,看谁的下一跳地点的受欢迎程度高,就将感知数据转发给它.式(12)~式(14)为转发节点 $Fnode$ 的计算公式:

$$Fnode = S_j, L_B \in BestSink \quad (12)$$

$$Fnode = S_k, L_c \in BestSink \quad (13)$$

$$Fnode = \begin{cases} \text{Sink}_j, & \text{Pop}(L_B) > \text{Pop}(L_c) \\ \text{Sink}_k, & \text{Pop}(L_c) > \text{Pop}(L_B) \end{cases} \quad (14)$$

受 PeopleRank 启发,我们提出移动传感节点执行的数据转发算法,具体描述如下所示.

Location Based Data Forwarding Algorithm.

Input: S_i meets S_j and S_K

Calculate:

Next Location(S_j)= L_B

Next Location(S_K)= L_C

Calculate $\text{Pop}(L_B)$;

Calculate $\text{Pop}(L_C)$;

If L_B is Sink $_j$

then $Fnode=S_j$ send data to S_j

Else If L_C is Sink $_j$

then $Fnode=S_K$ send data to S_K

Else If $\text{Pop}(L_B) > \text{Pop}(L_C)$

then $Fnode=S_j$ send data to S_j

Else If $\text{Pop}(L_C) > \text{Pop}(L_B)$

then $Fnode=S_K$ send data to S_K

End;

Output: $Fnode$.

2 模拟实验与分析

本节基于 THE ONE 模拟器进行模拟实验,评估算法的性能,并与 EpidemicRouter、FirstContactRouter、SprayAndWaitRouter、ProphetRouter、MaxPropRouter 经典机会路由算法进行比较.在实验中,我们绘制真实的湖南大学地图并导入到 THE ONE 模拟器中,同时设置 11 个地点作为用户相遇的偏好地点,其中 3 个位置部署了 AP.每个模拟场景重复 5 次取平均.

2.1 湖南大学群智感知网络模拟仿真

首先使用开源软件 OpenJump,分层绘制湖南大学南校区地图,主要绘制 road 层与 Community 层.其中,地图数据以 WKT 文件的方式,文件中 LINestring 格式内的所有坐标形成一条完整的路径曲线,ONE 模拟平台从 WKT 文件读取 LINestring 形式的坐标.

在实验中,将在设置的偏好地点的范围中心内部设置一个固定不动的真实节点,该节点的消息收发能力等于节点内所有移动节点收发能力之和,并为这一系列替代偏好地点的真实节点进行单独编组,使用基于地点的路由转发协议.如图 3 所示,矩形表示的是移动传感节点的 11 个偏好地点($L_0, L_1, \dots, L_{10}$).我们在偏好地点区域内分别选取一个内存无限大、位置固定的节点代替这个偏好地点,黑色三角形表示的是选取的 3 个目标 Sink,即感知数据被 3 个目标 Sink 中的任意一个所接收就可认为感知数据已成功传输.节点之间的通信接口是蓝牙,通信距离为 15m.每个模拟场景重复 5 次.

实验模型基本参数设置见表 3.在实验中,根据 default_settings.txt 的这个配置文件,节点基于感兴趣地点移

动,并且节点只在固定的偏好地点处转发感知数据(在道路上不转发数据),当移动节点到达某个地点时,就将自己所携带的感知数据托管给地点,本身不考虑转发.在实验中,我们用 c12.wkt、c13.wkt、c14.wkt 代替这 3 个 Sink,将消息的接受方设置为这 3 个 Sink 的其中一个.

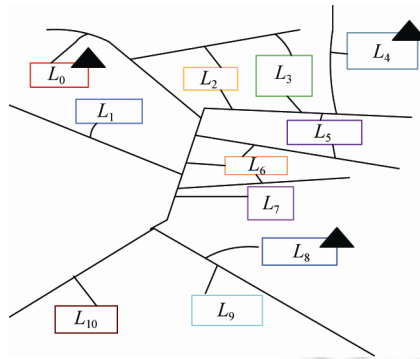


Fig.3 The distribution of interested locations and Sinks

图 3 偏好地点和目的 Sink 分布

Table 3 Simulation setting

表 3 模拟设置

参数	数值
Scenario time	12 000 S (12h)
Scenario.updateInterval	0.1
Group.bufferSize	10M
Transmission interface	Bluetooth (15m)
Number of preference location	11
Group.waitTime	100, 200

2.2 网络性能评估

本节通过模拟实验对基于地点的群智感知网络数据收集方法的各项性能进行评估.为定量比较,以下 3 个指标作为评价路由算法的度量依据,分别是传输成功率、消息转发开销率和消息转发延迟.

2.2.1 传输成功率

传输成功率用成功到达 Sink 的消息个数与网络中生成消息总数的比值表示.传输成功率是衡量一个路由算法好坏的重要指标.在基于最优 Sink 的地点路由协议的情况下,其中感知数据只在指定的场所内进行数据的转发,且在转发的时候进行基于下一跳地点排序选择中继转发节点.虚线表示在机会网络情况下以往几个路由协议的传输成功率,不对转发地点和最优 Sink 作限制要求;横坐标 node number 代表移动节点个数,从 5~55 逐渐增加;纵坐标 delivery-prob 代表传输成功率;图 4 表示与各路由协议,在不同节点个数情况下的传输成功率的对比.虚线表示传统的机会议由协议的实验结果,实线表示本文算法的实验结果.

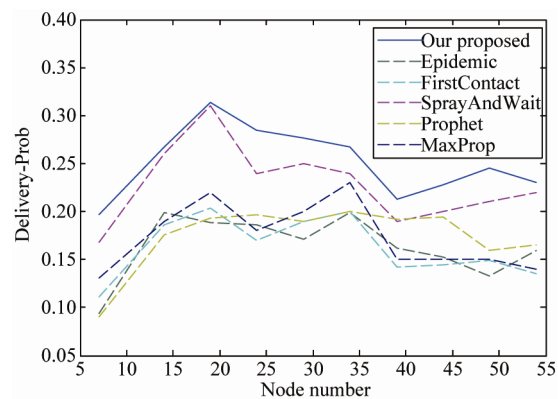


Fig.4 The comparison of the transmission success rates

图 4 不同节点个数情况下的传输成功率对比

从图 4 可以看出,除了 SprayAndWaitRouter 这种严格控制消息副本数量防止冗余发生的路由算法能够达到 20%以上的交付率以外,其他传统的机会议由路由算法在包含虚拟节点的移动模型中交付率均很低.分析其原

因是,移动节点大量高密度地存在于某个 Sink 节点附近时,而源节点和 Sink 由于不存在一条有效的链路甚至没有办法进行通信,无意义的相互转发消息占用了大量内存和时间,影响了移动节点和 Sink 节点间的消息转发效率.

2.2.2 消息转发开销率

本文用消息平均中转数表示消息转发开销,我们定义消息平均中转数如式(15)所示.

$$\text{消息平均中转数} = \frac{\text{网络中消息传递总次数} - \text{成功传达目标的消息个数}}{\text{成功传达目标的消息个数}} \quad (15)$$

消息平均中转数表示所有成功传达目标的消息,需要中转的平均次数,该值能在一定程度上反映消息传输的效率和消息转发开销.该值越小,表明消息成功传达目标需要的消息中转次数越少,消息转发开销越小,因而传输的效率越高.图 5 表示各路由协议,在不同节点个数情况下的开销率对比.虚线表示传统的几率路由协议的实验结果,实线表示本文算法的实验结果.

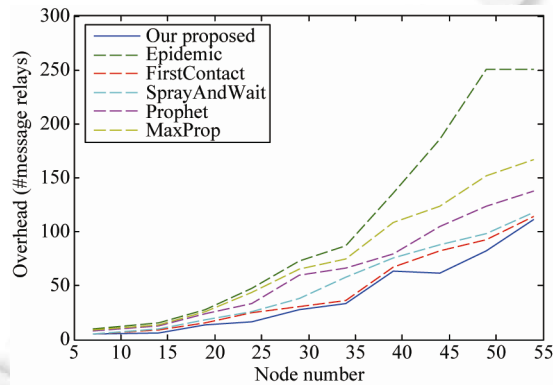


Fig.5 The comparison of overhead rates

图 5 开销率对比

从图 5 可以看出,其他传统的路由算法的开销率均很高,特别是 Epidemic.其原因在于,Epidemic 中移动节点只要遇到其他移动节点就将消息转发,消息转发开销率很高.同时,占用过多的消息缓存抑制了其他消息的存储和携带.在我们提出的方法中,当移动传感节点到达偏好地点处时,对比相遇的移动传感节点的下一跳偏好地点,将消息转发给前往 Sink 可能性最高的节点,如果都不前往,则选择最受欢迎的下一跳偏好地点,因此有较低的转发开销率.

2.2.3 消息转发延迟

消息转发延迟为数据包从发送端传输至 Sink 接收端所需花费的时间,该值越小,表明消息传输成功所需的时间越短,占用系统资源越少.图 6 表示与各路由协议,在不同节点个数情况下的传输延迟率对比.虚线表示传统的几率路由协议的实验结果,实线表示本文算法的路由实验结果.

从图 6 可以看出,其他传统的路由算法的转发延迟均很高,本文提出的算法分析节点的下一跳使消息转发有明确的目的性,消息转发后遇到 Sink 的概率更高,从而有较低的转发延迟.

2.2.4 产生不同数量的消息的路由性能对比

在模拟实验中,我们设置缓冲区的内存大小为 10M 字节,目的是分析不同数量的消息对路由性能的影响.我们将消息的产生数量从 50 逐渐增长到 400.图 7~图 9 表明了在不同消息个数情况下与 EpidemicRouter,First-ContactRouter,SprayAndWaiter,Prophet,MaxProp 路由协议的传输成功率、消息转发开销率、消息转发延迟的对比情况.

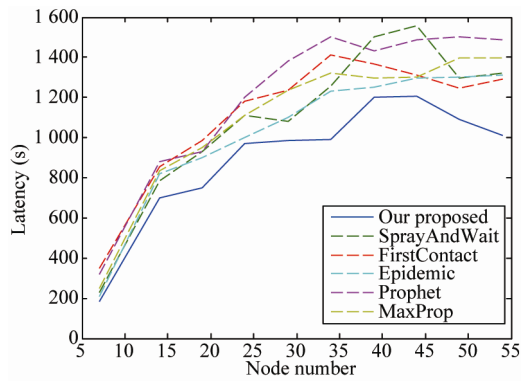


Fig.6 Delays with different number of nodes

图 6 不同节点个数情况下的延迟

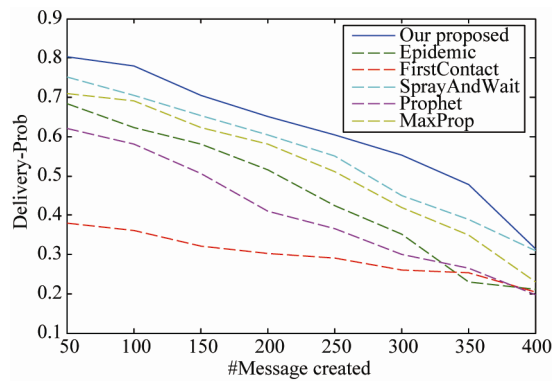


Fig.7 Transmission rates with different number of messages

图 7 不同消息个数情况下的传输成功率

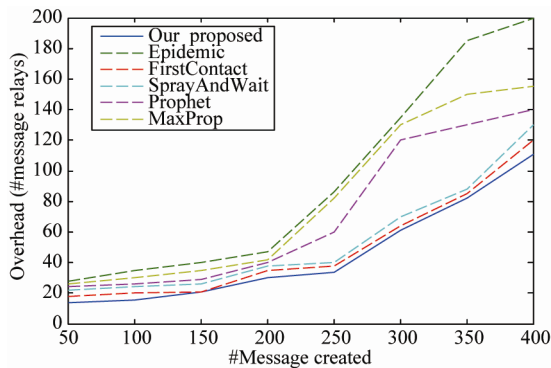


Fig.8 Overhead rates with different number of nodes

图 8 不同消息个数情况下的开销率

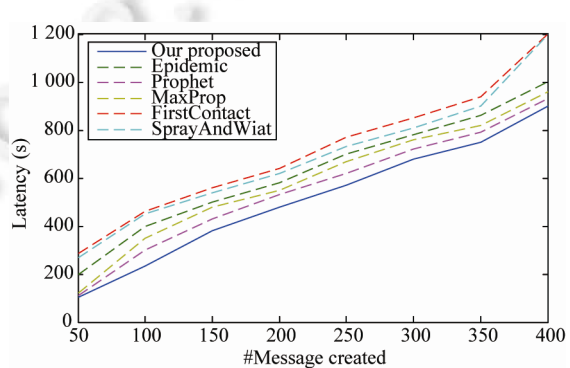


Fig.9 Forward delays with different number of messages

图 9 不同消息个数情况下的延迟

从图 7~图 9 可以看出,当网络中的消息数量逐渐增加时,即交通负载越来越大时,各路由协议的消息交付率都逐渐下降,我们提出的算法的效果仍好于传统路由协议。

3 总结与展望

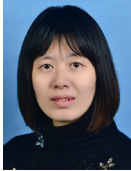
本文针对实际应用中的群智感知网络,提出了一种考虑网络中存在多个 Sink 的感知数据收集方案.首先不同于以往消息有特定的接收方,感知数据的目标节点可以是任意一个 Sink,考虑到实际环境中,很多地点部署了 Wi-Fi 的 AP,将 AP 比拟为 Sink,引入多目标层次分析法,选择出一个最优的 Sink 作为感知数据最终的接收者;其次人们所携带的智能设备只在建立稳定连接的地点内交换感知数据,不同于以往只要相遇就交换数据,所以利用地点的不变性,可形成稳定的网络拓扑,最终实现基于地点感知数据路由方法。

但是在未来的日子里,还有很多方面值得改进:(1) 如部署 Sink 的 AP 要设置在哪些场所,数据收集的效率最高;(2) 优化多目标分析法,更快地找到最优 Sink,从而提高数据传输效率;(3) 能基于机器学习来优化数据转发的路由算法,例如 SVM 算法,能够更智能地选择中继节点。

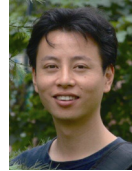
References:

- [1] Wang L, Zhang D, Xiong H, Gibson JP, Chen C, Xie B. ecoSense: Minimize participants' total 3G data cost in mobile crowd sensing using opportunistic relays. IEEE Trans. on Systems, Man, and Cybernetics: Systems, 2016, 1-14.
- [2] 刘云浩.群智感知计算.中国计算机学会通讯,2012,8(10):38-41.

- [3] Chakchouk N. A survey on opportunistic routing in wireless communication networks. IEEE Communications Surveys & Tutorials, 2015,17(4):2214–2241.
- [4] Wang S, Wang X, Huang J, Bie R. The potential of mobile opportunistic networks for data disseminations. IEEE Trans. on Vehicular Technology, 2016,65(2):912–922.
- [5] Weinschrott H, Durr F, Rothermel K. StreamShaper coordination algorithms for participatory mobile urban sensing. In: Proc. of the IEEE 7th Int'l Conf. on Mobile Adhoc and Sensor Systems. San Francisco, 2010. 195–204.
- [6] Kerancn A, Ott J, Karkkainen T. The ONE simulator for DTN protocol evaluation. In: Proc. of the 2nd Int'l Conf. on Simulation Tools and Techniques. Begijnhoflaan: Institute for Computer Sciences, Social Informatics and Telecommunications Engineering (ICST), 2009. 1–10.
- [7] Burns B, Brock O, Levine BN. MORA routing and capacity building in disruption tolerant networks. In: Proc. of the IEEE 24th Annual Joint Conf. of the IEEE Computer and Communications Societies. 2005. 398–408.
- [8] Lindgren A, Doria A, Schelen O. Probabilistic routing in intermittently connected networks. ACM SIGMOBILE Mobile Computing and Communications Review, 2003,7(3):19–20.
- [9] Link JAB, Schmitz D, Wehrle K. GeoDTN: Geographic routing in disruption tolerant networks. In: Proc. of the IEEE Global Telecommunications Conf. Houston, 2011. 1–5.
- [10] Ko YB, Vaidya NH. Geocasting in mobile ad hoc networks: Location-Based multicast algorithms. In: Proc. of the IEEE Workshop on Mobile Computing Systems and Applications. New Orleans, 1999. 101–110.
- [11] Daly EM, Haahr M. Social network analysis for information flow in disconnected delay-tolerant MANETS. IEEE Trans. on Mobile Computing, 2009,8(5):606–621.
- [12] Lu S, Liu Y, Liu Y, Kumar M. LOOP: A location based routing scheme for opportunistic networks. In: Proc. of the IEEE 9th Int'l Conf. on Mobile Adhoc and Sensor Systems. Las Vegas, 2012. 118–126.
- [13] Chen M, Wang H. A multi-objective routing decision-making model for opportunistic network. In: Proc. of the IEEE Int'l Conf. on Cloud Computing and Intelligence Systems. 2011. 316–320.
- [14] Mtibaa A, May M, Diot C. PeopleRank: Social opportunistic forwarding. In: Proc. of the IEEE INFOCOM. 2010. 1–5.
- [15] Liu Y, Lu S, Liu Y. Routing design and analysis of place based opportunistic networks. In: Proc. of the IEEE 16th Int'l Conf. on Computational Science and Engineering (CSE). 2013. 252–259.



王继良(1973—),女,湖南长沙人,副教授,主要研究领域为传感器网络,教育教学改革.



唐晖(1972—),男,副教授,主要研究领域为计算机网络.



黄丽嫦(1993—),女,硕士生,主要研究领域为传感器网络.