

## 基于社团结构的容迟传感器网络多副本数据传输策略研究

付彬<sup>1</sup>, 李仁发<sup>1</sup>, 肖雄仁<sup>1</sup>, 黄亚薇<sup>2</sup>

(1. 湖南大学 湖南省嵌入式与网络计算重点实验室, 湖南 长沙 410082; 2. 重庆理工大学 计算机科学与工程学院, 重庆 400054)

**摘要:** 在社团结构网络模型下, 建立社团间多副本数据转发策略的传输概率模型, 提出一种基于社团结构的多副本数据传输(CMDD)策略。CMDD 策略动态计算社团间移动节点的平均相遇间隔时间, 并根据传输概率模型估算借助该社团分发副本可获得的成功传输概率。模拟实验表明, CMDD 策略能以较低的传输延迟和较低的传输开销获得较高的数据传输成功率。

**关键词:** 移动传感器网络; 容迟网络; 数据传输; 社团结构; 多副本转发

中图分类号: TP393

文献标识码: A

文章编号: 1000-436X(2013)07-0174-10

## Community based multi-copied data delivery scheme for delay tolerant sensor network

FU Bin<sup>1</sup>, LI Ren-fa<sup>1</sup>, XIAO Xiong-ren<sup>1</sup>, HUANG Ya-wei<sup>2</sup>

(1. Key Lab for Embedded and Network Computing of Hunan Province, Hunan University, Changsha 410082, China;

2. School of Computer Science and Engineering, Chongqing University of Technology, Chongqing 400054, China)

**Abstract:** The characteristics of community structure have effect on the data transmission procedure of delay tolerant mobile sensor network. Under the community network model, data delivery probabilities of multi-copied delivery were deduced. A community based multi-copy data delivery (CMDD) scheme was also proposed. By dynamically computing the average inter-meeting time between the communities and using the delivery probability formulations, the scheme can estimate the delivery probability when data are delivered through such community. Simulation results show that the proposed CMDD scheme achieves a relatively high data delivery ratio and a lower message delivery latency and transmission overhead.

**Key words:** mobile sensor network; delay/disruption tolerant network; data delivery; community structure; multi-copied delivery

### 1 引言

在无线传感器网络的一类应用中, 需要把传感器节点附着在运动物体(人、动物或车辆等载体)上, 如野生动物生活习性数据收集、流感病毒传播跟踪、高危人群健康监测等。物体的运动导致传感器节点之间的链路形成一个间歇连通的移动传感器网络。该类网络具有拓扑变化频繁、节点稀

疏、间断连通、延迟容忍等容迟网络(DTN, delay/disruption tolerant network)的特点<sup>[1]</sup>, 称之为容迟移动传感器网络(DTMSN, delay/disruption tolerant mobile sensor network)<sup>[2]</sup>。

在 DTMSN 中, 由于难以维护稳定的传输路径, 节点采用“存储—携带—转发”(store-carry-forward)的机会传输模式, 在遇到目标节点或者找到更合适的中转节点前, 当前节点需要保存中转的数据分

收稿日期: 2012-07-12; 修回日期: 2013-04-20

基金项目: 国家自然科学基金资助项目(61173036, 60873074); 国家高技术研究发展计划(“863”计划)基金资助项目(2012AA01A301-01); 中央高校基本科研业务基金资助项目

**Foundation Items:** The National Natural Science Foundation of China(61173036,60873074); The National High Technology Research and Development Program of China (863 Program) (2012AA01A301-01); The Fundamental Research Funds for the Central Universities

组。在这种架构中，传统网络中的路由发现和路由维护机制被弱化，传输协议不仅需要设计高效的转发策略，利用有限的转发机会提高传输性能，还需要考虑如何降低“存储—携带—转发”模式导致的各种额外开销，例如存储开销。

经典路由协议利用节点移动规律进行预测或判定来提高转发效率，避免随机转发带来的开销。例如 RED<sup>[2]</sup>、PROPHET<sup>[3]</sup>、MaxProp<sup>[4]</sup>、RAPID<sup>[5]</sup>，均基于相遇历史预测未来相遇概率，依据和目标节点的相遇概率来选择中转节点或管理存储空间。该方法从时间维度体现了移动特性对提升协议性能的价值，但监测时间范围的选择是此类协议的一大难题<sup>[6]</sup>，并且在社团类型网络中可能失效<sup>[7]</sup>。现实中，人类的移动必然受到人的社会性影响，节点的社会性特征可以用于协助传感器节点的数据收集，例如，相遇次数决定了数据转发的机会、接触时间长短决定了传输数据量的大小。近年来，学者们通过挖掘人类的社会行为特征用于研究 DTN 的数据转发<sup>[7-10]</sup>。这些策略主要强调个体移动性及社会属性，如节点的中心性、流行度、朋友关系等，未考虑社团结构对网络数据传输带来的影响。文献[11]分析了社会网络对容迟网络数据传输的影响，但该文只对网络中存在 2 个社团的情况进行了分析，未深入分析多个社团相互作用对数据传输性能的影响。因此，如何充分利用社团结构来提升 DTN 数据传输的性能是一个待解决的重要问题。

本文首先分析了社团结构网络的特点，指出节点的交互并不是完全随机和同构的，而在一定程度上是同构和异构行为的混合。接着，在社团结构网络模型下详细分析副本通过不同的社团进行转发时可获得的成功传输概率，以此作为副本分发社团选择的依据。最后，提出一种基于社团结构的多副本数据传输 (CMDD, community based multi-copy data delivery) 策略。由于节点只需要记录社团的相遇信息，而不必记录每个节点的相遇信息，在一定程度上节省了存储开销。仿真实验表明，本文提出的 CMDD 策略与直接传递、洪泛策略、SNW 策略相比，能以较低的传输延迟和较低的传输开销获得较高的数据传输成功率。

## 2 相关工作

根据转发策略的不同，容迟网络数据传输策略

主要有基于冗余的数据传输策略、基于效用的数据传输策略和冗余效用混合等几类<sup>[12]</sup>。

病毒式洪泛<sup>[13]</sup>是最简单的冗余数据传输策略，节点把携带的数据转发给所有相遇的节点。在不考虑资源限制的情况下，该方法能够最快实现数据的传输，但不适用于节点资源有限场景。与洪泛策略相反，直接传递策略<sup>[14]</sup>不选择中转节点，而是直到与目标节点相遇时才转发数据。这种方式在最小化传输开销的同时，大大增加了传输延迟。SWIM 系统<sup>[15]</sup>用于收集鲸的生物信息，假设各节点和汇聚点相遇概率相同，源节点传输消息的有限副本给相遇节点，从而实现一定有效率的数据传输。SNW<sup>[16]</sup>是一种副本数受限的复制策略，在第 1 阶段中，报文的源节点仅向网络中复制固定数量的副本，即喷洒报文；在第 2 阶段，持有该报文副本的节点仅仅是等待，直至遇到报文传输的目标节点时，才将报文转发给该节点以完成传输过程。该文假设节点处于随机运动模型下，节点之间相遇的概率相等，因此在喷洒阶段，源节点将副本分发给最初遇到的节点。

节点之间的相遇历史记录反映了节点的移动特性，为预测将来的相遇情况提供了决策依据。PROPHET<sup>[3]</sup>定义了传输预测值来描述节点间成功传输的概率。每当节点相遇时，只有在对方节点到目标节点的传输预测值大于自身到目标节点的传输预测值时，才向对方节点转发报文。由于预测的准确度会随着时间的推移降低，因此该算法适用于相遇频繁的小规模移动网络。MaxProp<sup>[4]</sup>采用增量平均化的方法来估算节点间成功传输报文的概率，并要求节点根据彼此相遇概率计算单跳成本以及计算全局最优的传输路径。全局算法增加了计算和存储开销。RAPID<sup>[5]</sup>将转发决策转化为一个效用驱动的资源分配问题，根据延迟为每一个数据分组计算效用值，效用值的计算需要全局相遇信息的支持。RDAD<sup>[17]</sup>采用传感器节点到汇聚点的相对距离来计算节点传输概率的大小，作为消息传输时选择下一跳的依据。该策略需要汇聚点周期性地以大发射功率等级发出广播消息。

以上数据传输策略未充分考虑移动节点的社会性及其对数据传输策略性能的影响。SimBet<sup>[7]</sup>利用社会网络分析技术，提出了通过局部信息计算节点中心性 (betweenness) 和社会相似度 (similarity) 的方法，并通过这 2 个属性值计算转发的效用函数。BUBBLE RAP<sup>[8]</sup>是一种分布式路由算法，利用节点

流行度 (popularity) 和归属团体 (social community) 进行决策, 通过选择更高流行度和目标节点所在团体成员作为中转节点, 提高报文转发的效率。文献 [9] 提出了一个衡量朋友关系的度量指标, 为节点定义其直接及间接的朋友社团, 利用朋友社团进行数据转发。由于在真实移动网络中, 人们常常以组或社团的形式活动, 文献 [11] 分析了社团网络结构对 DTN 路由的影响, 定义了社团内部和社团之间的相遇间隔时间, 讨论了 2 个社团情况下的多副本传输方法的传输成功率和时间延迟, 多个社团间的相互作用对 DTN 路由的影响未做深入分析。

### 3 社团结构网络模型

社团结构是人类社会活动中形成的一种重要结构。从大量人类移动轨迹数据分析中发现, 节点移动在空间或社会形式上体现社团结构<sup>[8,18]</sup>。一般社团内部个体接触时间长, 接触次数多; 不同社团之间的个体接触时间短, 接触次数少。社团结构下的节点移动特征反映了节点的交互并不是完全随机和同构的, 而在一定程度上是同构和异构行为的混合<sup>[11,19]</sup>。在现实中, 虽然人们接触的次数和时间不可预测, 然而, 从平均值看, 社团内部以及社团之间的平均相遇间隔时间是较稳定的, 该值可以通过节点相遇的历史记录得到。

文献 [11,19,20] 对社团结构模型进行了定义, 本文遵循该通用模型, 采用的模型定义如下。假设整个网络由  $m$  个社团  $C_i (i=1, \dots, m)$  构成, 社团  $C_i$  中的节点数以  $N_i$  表示。  $B_{ij}$  表示  $C_i$  中的节点与  $C_j$  中的节点的平均相遇间隔时间, 即平均每  $B_{ij}$  个时间单元 2 个社团中的节点相遇一次, 有一次交换数据的机会; 社团  $C_i$  内部节点的平均相遇间隔时间用  $B_i$  表示, 即平均每  $B_i$  个时间单元社团  $C_i$  内部的节点相遇一次。在同一社团内部, 节点之间的平均相遇间隔时间相同; 在不同社团之间, 根据社团之间联系的紧密程度, 具有不同的平均相遇间隔时间。

图 1 为由 4 个社团构成的社团结构网络模型。在社团  $C_1$  内部, 社团成员节点的平均相遇间隔时间为  $B_1$ ; 在社团之间, 社团  $C_1$  内的节点以  $B_{13}$  的平均相遇间隔时间与社团  $C_3$  中的节点接触, 而以  $B_{14}$  的平均相遇间隔时间与社团  $C_4$  中的节点接触。由于社团内部节点互相接触更频繁, 因此社团  $C_1$  内的节点的平均相遇间隔时间  $B_1$  小于  $B_{13}$ 、 $B_{14}$ 。该模型描述了社团结构网络的一般行为, 表达了社团内部的同

构特性和不同社团之间的异构特性。

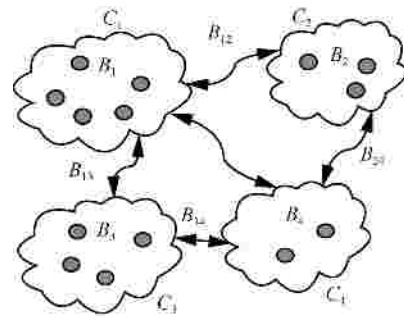


图 1 社团结构网络模型

### 4 副本分发过程分析与建模

基于复制的冗余数据传输是一种通过将原始信息复制多个副本, 然后将多个副本在网络中同时传输的策略。只要其中一个副本到达目的地, 就实现了信息的成功传输。副本的数量及副本的分发方式决定了数据传输的成功率、延迟及开销等性能。

#### 4.1 副本分发过程描述

如图 2 所示, 以 3 个社团为例,  $C_s$  为数据源节点  $s$  所在社团,  $C_d$  为目的节点  $d$  所在社团,  $C_e$  为其他社团, 假设各个社团中的节点数都为  $n$ 。限定网络中副本数为  $L$ , 并且只有源节点可以复制副本, 其中,  $L_{in}$  个副本在  $C_s$  内部分发,  $L_{out}$  个副本向其他社团分发。假设社团内部节点的平均相遇间隔时间为  $G$  个时间单元, 即  $B_s=B_d=B_e=G$ ; 在不同社团之间, 节点的平均相遇间隔时间为:  $B_{sd}=k_1G$ 、 $B_{se}=k_2G$ 、 $B_{ed}=k_3G$ 。由于社团内部节点之间的接触比不同社团之间的节点接触更频繁, 因此, 系数  $k_1$ 、 $k_2$ 、 $k_3$  取值大于 1。

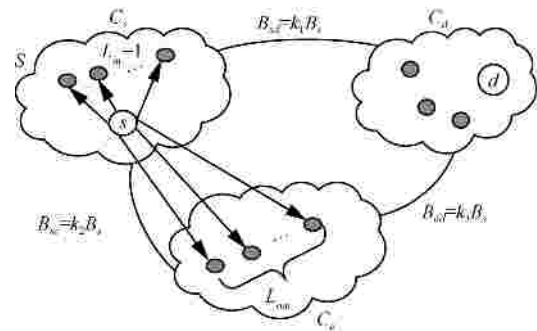


图 2 3 个社团副本分发过程

本文采用“喷洒” (spray) 和“等待” (wait) 2 个阶段组成数据传输策略。“喷洒”阶段完成副本分发, 由本地喷洒 (local spraying) 和全局喷洒 (global

spraying) 2 个独立的过程组成。在本地喷洒过程, 源节点将  $L_{in}-1$  个副本分发给自己在社团中的节点 (源节点持有 1 个副本); 在全局喷洒过程, 源节点将  $L_{out}$  个副本分发至其他社团。因此, 整个数据传输过程可分为 3 个阶段: 第 1 阶段 (all spraying), 本地喷洒和全局喷洒同时进行; 第 2 阶段 (mix phase), 本地喷洒或全局喷洒中的一个已经结束, 进入等待阶段, 另一个仍在继续; 第 3 阶段 (all waiting), 所有副本分发过程都结束, 进入等待阶段。在全局喷洒过程, 由于社团之间节点相遇间隔时间不同, 通过不同的社团转发副本的传输效率不同。若源节点与目的节点所在社团  $C_d$  的相遇概率大, 那么直接向  $C_d$  社团分发副本传输成功概率大; 若源节点与  $C_d$  相遇概率很小, 而与  $C_e$  社团相遇概率大,  $C_e$  社团与  $C_d$  社团又联系紧密, 那么, 通过  $C_e$  社团转发副本效率更高。因此, 需要讨论向目标节点所在社团分发副本和通过其他社团节点分发副本 2 种情况的传输成功概率。

#### 4.2 通过目标节点社团分发副本的成功传输概率

数据副本在源节点所在社团和目标节点所在社团分发。 $P^d(t)$  表示任意时刻  $t$  的数据传输成功概率;  $P_i^d$  表示第  $i$  个阶段可获得的数据传输成功概率;  $C_i^d$  表示在第  $i$  阶段数据未能成功传输至目标节点的概率;  $D_i^d(j)$  表示在第  $i$  阶段的第  $j$  个时间单元以前数据未能成功到达目标节点的概率。

1) 当分发的副本数满足  $L_{in} \leq Lk_1/(k_1+1) + 1/(k_1+1)$  时, 本地副本喷洒过程比全局副本喷洒过程先结束, 各个阶段的数据传输成功概率如下。

第 1 阶段 (all spraying), 在第  $i$  个时间单元, 源节点  $s$  最多向  $C_s$  中的节点分发  $i/G$  个副本 (源节点  $s$  自己携带一个副本);  $C_s$  中的每个副本与目的节点  $d$  相遇的概率为  $1/(k_1nG)$ ; 同时, 源节点最多向  $C_d$  中的节点分发  $i/(k_1G)$  个副本,  $C_d$  中的每个副本与目的节点  $d$  相遇的概率为  $1/(nG)$ 。因此, 在第 1 阶段的第  $i$  个时间单元, 至少有一个副本成功发送至目标节点  $d$  的概率  $P_1^d(i)$  为

$$P_1^d(i) = 1 - \left(1 - \frac{1}{k_1nt}\right)^{i+1} \left(1 - \frac{1}{nt}\right)^{\frac{i}{k_1t}}$$

由此可得, 第 1 阶段可获得的成功发送概率  $P_1^d$  以各个时间单元成功发送概率的和表示如下

$$P_1^d = \sum_{i=1}^{(L_{in}-1)t} D_1^d(i) P_1^d(i) \quad (1)$$

其中,  $D_1^d(i) = \prod_{j=1}^{i-1} (1 - P_1^d(j))$ 。 $D_1^d(i)$  表示在第 1 阶段的第  $i$  个时间单元以前数据未能成功到达目标节点的概率。

第 2 阶段 (mix phase), 本地喷洒过程已经结束, 在第  $i$  个时间单元,  $C_s$  中的节点拥有  $L_{in}$  个副本,  $C_d$  中的节点拥有  $i/(k_1G)$  个副本, 至少有一个副本成功发送至目标节点  $d$  的概率  $P_2^d(i)$  为

$$P_2^d(i) = 1 - \left(1 - \frac{1}{k_1nt}\right)^{L_{in}} \left(1 - \frac{1}{nt}\right)^{\frac{i}{k_1t}}$$

因此, 第 2 阶段可获得的成功传输概率  $P_2^d$  为

$$P_2^d = \sum_{i=(L_{in}-1)t+1}^{(L-L_{in})k_1t} C_1^d D_2^d(i) P_2^d(i) \quad (2)$$

其中,  $C_1^d = \prod_{s=1}^{(L_{in}-1)t} (1 - P_1^d(s))$ ,  $D_2^d(i) = \prod_{j=(L_{in}-1)t+1}^{i-1} (1 - P_2^d(j))$ 。

第 3 阶段 (all waiting), 所有副本分发过程都结束,  $C_s$  中的节点拥有  $L_{in}$  个副本,  $C_d$  中的节点拥有  $L - L_{in}$  个副本, 携带副本的节点等待与目标节点相遇。在第  $i$  个时间单元, 至少有一个副本成功发送至目标节点  $d$  的概率  $P_3^d(i)$  为

$$P_3^d(i) = 1 - \left(1 - \frac{1}{k_1nt}\right)^{L_{in}} \left(1 - \frac{1}{nt}\right)^{L-L_{in}}$$

因此, 第 3 阶段可获得的成功传输概率  $P_3^d$  为

$$P_3^d = \sum_{i=(L-L_{in})k_1t+1}^{\infty} C_1^d C_2^d D_3^d(i) P_3^d(i) \quad (3)$$

其中,  $C_2^d = \prod_{s=(L_{in}-1)t+1}^{(L-L_{in})k_1t} (1 - P_2^d(s))$ ,  $D_3^d(i) = \prod_{j=(L-L_{in})k_1t+1}^{i-1} (1 - P_3^d(j))$ 。

根据以上对各个阶段数据成功传输率的分析以及式(1)~式(3)可得, 到任意时刻  $t$  的数据传输成功率  $P^d(t)$  为:

$$P^d(t) = \begin{cases} \sum_{i=1}^t D_1^d(i) P_1^d(i), & t \leq (L_{in}-1)t \\ P_1^d + \sum_{i=(L_{in}-1)t+1}^t C_1^d D_2^d(i) P_2^d(i), & (L_{in}-1)t < t \leq (L-L_{in})k_1t \\ P_1^d + P_2^d + \sum_{i=(L-L_{in})k_1t+1}^t C_1^d C_2^d D_3^d(i) P_3^d(i), & t > (L-L_{in})k_1t \end{cases} \quad (4)$$

2) 当分发的副本数满足： $L_{in} > Lk_1 / (k_1+1) + 1 / (k_1+1)$  时，全局副本喷洒过程比本地副本喷洒过程先结束，各个阶段的数据传输成功概率如下

$$P_1^d = \sum_{i=1}^{(L-L_{in})k_1t} D_1^d(i)P_1^d(i) \quad (5)$$

$$P_2^d = \sum_{i=(L-L_{in})k_1t+1}^{(L_{in}-1)t} C_1^d D_2^d(i)P_2^d(i) \quad (6)$$

$$P_3^d = \sum_{i=(L_{in}-1)t+1}^{\infty} C_1^d C_2^d D_3^d(i)P_3^d(i) \quad (7)$$

其中， $P_1^d(i)$ 、 $P_2^d(i)$ 、 $P_3^d(i)$ 、 $D_1^d(i)$ 、 $D_2^d(i)$ 、 $D_3^d(i)$ 、 $C_1^d$ 、 $C_2^d$  定义如下

$$P_1^d(i) = 1 - \left(1 - \frac{1}{k_1nt}\right)^i \left(1 - \frac{1}{nt}\right)^{L-L_{in}}$$

$$D_2^d(i) = \prod_{j=(L-L_{in})k_1t+1}^{i-1} (1 - P_2^d(j))$$

$$D_3^d(i) = \prod_{j=(L_{in}-1)t+1}^{i-1} (1 - P_3^d(j))$$

$$C_1^d = \prod_{s=1}^{(L-L_{in})k_1t} (1 - P_1^d(s))$$

$$C_2^d = \prod_{s=(L-L_{in})k_1t+1}^{(L_{in}-1)t} (1 - P_2^d(s))$$

该情况下，根据式(5)~式(7)，可得到任意时刻  $t$  的数据成功传输率  $P^d(t)$  为

$$\begin{cases} \sum_{i=1}^t D_1^d(i)P_1^d(i), & t \leq (L-L_{in})k_1t \\ P_1^d + \sum_{i=(L-L_{in})k_1t+1}^t C_1^d D_2^d(i)P_2^d(i), & (L-L_{in})k_1t < t \leq (L_{in}-1)t \\ P_1^d + P_2^d + \sum_{i=(L_{in}-1)t+1}^t C_1^d C_2^d D_3^d(i)P_3^d(i), & t > (L_{in}-1)t \end{cases} \quad (8)$$

### 4.3 通过中间社团分发副本的成功传输概率

假设通过其他社团  $C_e$  来分发副本， $P^e(t)$  表示在任意时刻  $t$  的数据传输成功概率， $P_i^e$  表示第  $i$  个过程可获得的数据传输成功概率， $C_i^e$  表示在第  $i$  阶段数据未能成功传输至目标节点的概率， $D_i^e(j)$  表示在第  $i$  阶段的第  $j$  个时间单元以前数据未能成功到达目标节点的概率。

1) 当满足条件  $L_{in} > Lk_2 / (k_2+1) + 1 / (k_2+1)$  时，本地副本分发过程比全局副本分发过程先结束。此

时，各个阶段的数据传输成功概率如下

$$P_1^e = \sum_{i=1}^{(L_{in}-1)t} D_1^e(i)P_1^e(i) \quad (9)$$

$$P_2^e = \sum_{i=(L_{in}-1)t+1}^{(L-L_{in})k_2t} C_1^e D_2^e(i)P_2^e(i) \quad (10)$$

$$P_3^e = \sum_{i=(L-L_{in})k_2t+1}^{\infty} C_1^e C_2^e D_3^e(i)P_3^e(i) \quad (11)$$

其中， $P_1^e(i)$ 、 $P_2^e(i)$ 、 $P_3^e(i)$ 、 $D_1^e(i)$ 、 $D_2^e(i)$ 、 $D_3^e(i)$ 、 $C_1^e$ 、 $C_2^e$  定义如下

$$P_1^e(i) = 1 - \left(1 - \frac{1}{k_1nt}\right)^i \left(1 - \frac{1}{k_3nt}\right)^{\frac{i}{k_2t}}$$

$$P_2^e(i) = 1 - \left(1 - \frac{1}{k_1nt}\right)^{L_{in}} \left(1 - \frac{1}{k_3nt}\right)^{\frac{i}{k_2t}}$$

$$P_3^e(i) = 1 - \left(1 - \frac{1}{k_1nt}\right)^{L_{in}} \left(1 - \frac{1}{k_3nt}\right)^{L-L_{in}}$$

$$D_1^e(i) = \prod_{j=1}^{i-1} (1 - P_1^e(j))$$

$$D_2^e(i) = \prod_{j=(L_{in}-1)t+1}^{i-1} (1 - P_2^e(j))$$

$$D_3^e(i) = \prod_{j=(L-L_{in})k_2t+1}^{i-1} (1 - P_3^e(j))$$

$$C_1^e = \prod_{s=1}^{(L_{in}-1)t} (1 - P_1^e(s))$$

$$C_2^e = \prod_{s=(L-L_{in})k_2t+1}^{(L-L_{in})k_2t} (1 - P_2^e(s))$$

该情况下，根据式(9)~式(11)，到任意时刻  $t$  的数据成功传输率  $P^e(t)$  为

$$\begin{cases} \sum_{i=1}^t D_1^e(i)P_1^e(i), & t \leq (L_{in}-1)t \\ P_1^e + \sum_{i=(L_{in}-1)t+1}^t C_1^e D_2^e(i)P_2^e(i), & (L_{in}-1)t < t \leq (L-L_{in})k_2t \\ P_1^e + P_2^e + \sum_{i=(L-L_{in})k_2t+1}^t C_1^e C_2^e D_3^e(i)P_3^e(i), & t > (L-L_{in})k_2t \end{cases} \quad (12)$$

2) 当满足条件  $L_{in} > Lk_2/(k_2+1)+1/(k_2+1)$  时，全局副本分发过程比本地副本分发过程先结束，各个阶段的数据传输成功概率如下

$$P_1^e = \sum_{i=1}^{(L-L_{in})k_2t} D_1^e(i)P_1^e(i) \quad (13)$$

$$P_2^e = \sum_{i=(L-L_{in})k_2t+1}^{(L_{in}-1)t} C_1^e D_2^e(i)P_2^e(i) \quad (14)$$

$$P_3^e = \sum_{i=(L_{in}-1)t+1}^{\infty} C_1^e C_2^e D_3^e(i)P_3^e(i) \quad (15)$$

其中， $P_1^e(i)$ 、 $P_3^e(i)$ 、 $D_1^e(i)$  定义同上  $P_2^e(i)$ 、 $D_2^e(i)$ 、 $D_3^e(i)$ 、 $C_1^e$ 、 $C_2^e$  定义如下

$$P_2^e(i) = 1 - \left(1 - \frac{1}{k_1 nt}\right)^{i+1} \left(1 - \frac{1}{k_3 nt}\right)^{L-L_{in}}$$

$$D_2^e(i) = \prod_{j=(L-L_{in})k_2t+1}^{i-1} (1 - P_2^e(j))$$

$$D_3^e(i) = \prod_{j=(L_{in}-1)t+1}^{i-1} (1 - P_3^e(j))$$

$$C_1^e = \prod_{s=1}^{(L-L_{in})k_2t} (1 - P_1^e(s))$$

$$C_2^e = \prod_{s=(L-L_{in})k_2t+1}^{(L_{in}-1)t} (1 - P_2^e(s))$$

该情况下，根据式(13)~式(15)，到任意时刻  $t$  的数据成功传输率  $P^e(t)$  为

$$\begin{cases} \sum_{i=1}^t D_1^e(i)P_1^e(i), & t \leq (L-L_{in})k_2t \\ P_1^e + \sum_{i=(L-L_{in})k_2t+1}^t C_1^e D_2^e(i)P_2^e(i), & (L-L_{in})k_2t < t \leq (L_{in}-1)t \\ P_1^e + P_2^e + \sum_{i=(L_{in}-1)t+1}^t C_1^e C_2^e D_3^e(i)P_3^e(i), & t > (L_{in}-1)t \end{cases} \quad (16)$$

#### 4.4 2种副本分发方式比较

社团之间平均相遇间隔时间的参数  $k_1$ 、 $k_2$  和  $k_3$  的大小描述了社团之间的亲疏程度。对该参数不同取值条件下的数据传输成功概率进行比较。实验各参数取值如表 1 所示，副本通过目标节点  $d$  所在社团  $C_d$  转发以及通过其他社团  $C_e$  转发获得的传输成功概率如图 3 所示。

表 1 模型中参数取值

参数	取值
$n$	10
$G$	1
$L$	10
$L_{in}$	4
$k_1$	6
$k_2, k_3$	$k_2=4, k_3=5$ 或 $k_2=2, k_3=2$

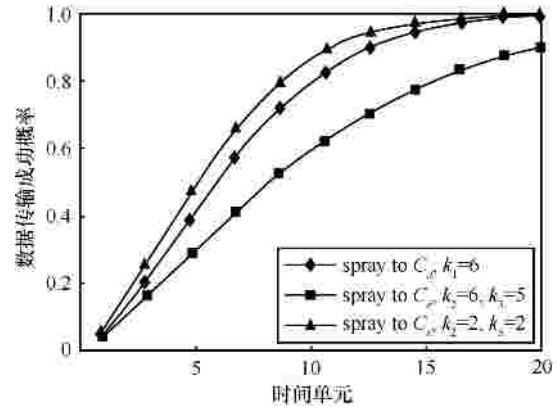


图 3 数据传输成功概率对比

从图 3 可以看出，选择不同的社团来分发副本对数据传输成功概率有较大的影响。当源节点与目标节点所在社团的平均相遇间隔时间较大 ( $k_1=6$ ) 时，而中间社团与源节点所在社团和目的节点所在社团都关系较密切 ( $k_2=2, k_3=2$ ) 时，选择通过该中间社团来分发副本可以获得更高的成功传输概率；而当中间社团与源节点所在社团和目的节点所在社团都关系较疏远 ( $k_2=4, k_3=5$ ) 时，直接向目的节点所在社团分发副本的传输成功概率更高。

由此可见，合理地选择副本分发社团有利于提高数据传输成功率。

#### 5 面向社团结构的多副本数据传输策略

假设节点已通过社团发现算法获得自己所属的社团信息，社团发现算法如文献[7,8]。为便于讨论，将节点的移动区域划分为等面积的网格区域，每个社团随机选择一个网格作为其主要活动区域，并以该网格所在位置坐标标识该社团。节点通过交换社团相遇时的相关信息来确定彼此的移动特征，并确定最佳副本分发策略。面向社团结构的多副本数据传输策略主要包括建立社团相遇矩阵、选择副本分发社团、快速分发副本 3 个步骤。

##### 1) 建立社团平均相遇间隔时间矩阵

假设，从时刻  $T_{start}$  开始统计社团间的相遇信息。

每个节点记录其所在社团与其他社团的相遇次数，以及平均相遇间隔时间。例如，节点  $i$  与节点  $j$  相遇(2 个节点所在社团分别为社团  $i$  和社团  $j$ )，节点  $i$  计算  $ContactTimes_{ij}=ContactTimes_{ij}+1$ ，并计算  $B_{ij}=(CurrentMeetTime_{ij}-T_{start})/ContactTimes_{ij}$ ，然后更新记录中对应的各项值。从记录形式看，节点不需要记录移动轨迹的具体历史，也不需要记录与每个节点之间相遇的信息，而是记录与社团相遇的次数及平均相遇间隔时间，存储开销较小。统计社团间相遇信息的时间周期可以根据具体应用灵活设定，例如，可以按照人的活动周期以小时、天、周为单位进行统计。

### 2) 选择副本分发社团

假设数据源节点  $s$  需要发送数据至目标节点  $d$ ，源节点  $s$  根据自身和目标节点  $d$  所在社团的移动特征选择最佳的副本分发社团。副本分发社团的选择，主要依据第 3 节计算的数据传输成功概率。源节点根据自己记录的社团平均相遇间隔时间去估算通过哪个社团分发副本的成功率更高，具体过程如下。首先，源节点计算通过目标节点  $d$  所在社团  $C_d$  分发副本的成功传输概率  $P_1^d$  (根据式(1)和式(5))；然后，若源节点社团与其他社团的相遇间隔时间小于与目标节点所在社团的相遇间隔时间，那么源节点计算通过中间社团分发副本的成功概率  $P_1^e$  (根据式(9)和式(13))；若  $P_1^e > P_1^d$ ，则选择通过社团  $e$  分发副本，否则选择通过社团  $d$  分发副本。在此，由于一般社团特性相对较稳定，通过估算第 1 阶段传输成功概率大小来选择副本分发社团。若第 1 阶段传输成功率高，也就可以判断该社团是和目的社团联系更紧密的社团，从而是更合适地协助数据分发的社团。

### 3) 快速分发副本

确定副本分发社团后，便启动“喷洒”和“等待”的两阶段数据传输机制。在喷洒阶段，源节点将  $L_{in}$  个副本分发给自己在所在社团中的节点，将  $L_{out}$  个副本分发至选定的其他转发副本社团；喷洒阶段采用二元策略快速分发副本，即每当节点遇到选定的可以协助分发副本的节点时，将其所携带的  $c$  份副本的  $c/2$  转交给相遇节点，直到只携带一份冗余则停止分发。携带冗余  $c > 1$  的节点如遇到目标节点，则停止分发并将数据传送给目标节点。在等待阶段，持有该报文副本的节点等待与目标节点相遇，再把数据转发给目标节点。

算法伪代码如下。

**CMDD 算法 节点  $i$ , pseudo-code of node  $i$**   
 /\*假定节点  $i$  在社团  $C_i$  中,节点  $j$  在社团  $C_j$  中,目的节点  $d$  在社团  $C_d$  中\*/

```

begin
  if (node  $i$  contact node  $j$ ) then
     $ContactTimes_{ij} = ContactTimes_{ij} + 1$ 
     $LastMeetTime_{ij} = CurrentTime - T_{start}$ 
     $B_{ij} = CurrentMeetTime_{ij} / ContactTimes_{ij}$ 
  end if
  if (node  $i$  has data to send to node  $d$ ) then
     $P_1^d = CalculateDeliveryProb(C_d)$ 
     $P_1^e = 1$ 
    for (each community  $z$  that  $i$  meets) do
      if ( $B_{sz} < B_{sd}$ ) then
         $P_1^z = CalculateDeliveryProb(C_z)$ 
        if ( $P_1^z < P_1^d$  &&  $P_1^z < P_1^e$ ) then
           $P_1^e = P_1^z$ 
           $ChosedCommunityId = C_z$ 
        else
           $ChosedCommunityId = C_d$ 
        end if
      end if
    end for
     $Deliver-Copy(ChosedCommunityId,$ 
 $L_{out})$ 
     $DeliveryCopy(C_s, L_{in})$ 
  end if
end

```

## 6 模拟实验

模拟实验平台为 ONE (opportunistic network environment)<sup>[21]</sup>，实验比较了直接传递、洪泛策略、SNW、CMDD 策略的平均传输成功率、平均传输延迟和平均冗余传输 3 个性能指标。其中，传输成功率是指在给定时间内成功传输到目标节点的报文与源节点发送报文的比值；传输延迟指报文从源节点传输到目标节点所需的平均时间；定义成功接收单个数据分组所引起的平均冗余传输为传输开销，平均冗余传输=(转发的数据总量-成功接收的数据)/成功接收的数据。

### 6.1 参数设置

在仿真平台下实现了第 2 节提到的社团结构网络模型，整个运动区域为 1 000 m×1 000 m 的正方

形，每个网格大小为 50 m×50 m，节点移动遵循的主要规则如下。

1) 初始时，每个社团随机选择一个网格作为其主要活动区域，该社团中的节点随机选择其社团所在网格中的一个位置作为其初始位置。

2) 社团  $C_i$  中的节点平均每  $B_i$  个时间单元随机选择自己社团所在区域的一个目的位置向其移动。由于人们在社团内部活动一般为步行方式，设置节点移动速度在 [1,3] m/s 内随机选择。节点到达目的位置后，在该位置暂停一段时间后再移动，暂停时间服从 [60,600] s 内的均匀分布。

3) 社团  $C_i$  中的节点平均每  $B_{ij}$  个时间单元随机选择  $C_j$  社团所在区域的一个目的位置向其移动。人们在社团间移动，根据社团距离的远近，可能采取步行或车行等方式。因此，节点移动速度在 [1,11] m/s 内随机选择，在目的位置的暂停时间服从 [60,600] s 内的均匀分布。

传感器节点的通信范围为 10 m，网络带宽为 1 Mbit/s。仿真开始的前 500 s 为节点的热身时间，源节点社团从 600 s 开始，每间隔 10 s 产生一个大小为 1 KB 的数据分组，源节点共产生 100 个数据分组，节点缓存大小 50 KB。为了减小仿真的随机误差，所示结果均为 30 次仿真的平均值。

### 6.2 4 种策略性能对比

在本组实验中网络中存在的社团数在 [2,6] 内随机产生，每个社团中的节点数在 [10,20] 内随机产生，社团内部节点的相遇间隔时间服从均值为 300 s 的均匀分布，表示各个社团之间联系紧密程度的系数参数  $k$  在 [2,8] 内随机产生，社团之间节点的相遇间隔时间服从均值为  $k \times 300$  s 的均匀分布，SNW 和 CMDD 的副本数在 [5,10] 内随机产生。

图 5 为仿真不同时刻各个策略的平均传输成功率。从图中可以看到，随仿真时间的增加，直接传递、SNW、CMDD 策略的平均传输成功率都在增加，洪泛策略的平均传输成功率达到 0.7 后不再增加，这是因为大量的洪泛消息使传感器节点的缓冲队列迅速被填满，无法再接受新的消息。在仿真 35 min 之前，洪泛策略的平均传输成功率比其他 3 个策略要略高；35 min 之后，CMDD 策略的平均传输成功率高于其他策略，这是因为 CMDD 策略考虑了社团之间的相遇间隔时间，把副本转交给与目标节点所在社团相遇更频繁的社团，与 SNW 相比提高了报文传输成功率。直接

传递策略只有遇到目标节点时才发送数据，在源节点社团与目标节点社团相遇不频繁的情况下传输成功率非常低。

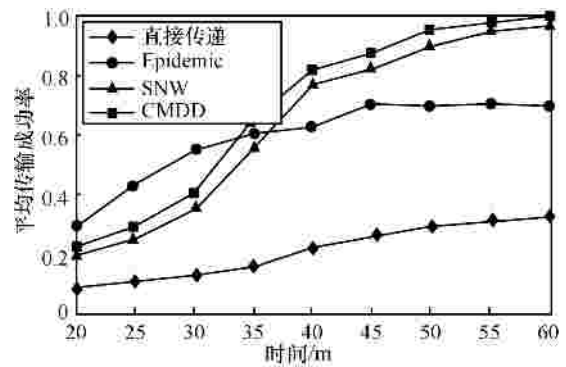


图 5 平均传输成功率

图 6 为在相同的传输成功率下各个策略的平均传输延迟。从图中可以看到，随着平均传输成功率的增加，报文的平均传输延迟都在增大。在保证一定传输成功率下，4 个策略中直接传递策略的平均传输延迟最大，其余 3 个策略由于是基于复制策略的数据传输，利用冗余传输技术提高了传输成功率，平均传输延迟较低，并且，CMDD 策略的平均传输延迟略小于 SNW 策略。这是因为 SNW 策略把副本分发给最初遇到的节点，未考虑社团之间的关系，主要通过源节点所在社团转发副本，在源节点所在社团与目标节点所在社团相遇间隔时间较大的情况下，副本成功到达目标社团的效率低，而 CMDD 策略选择了更合适的社团进行副本分发。

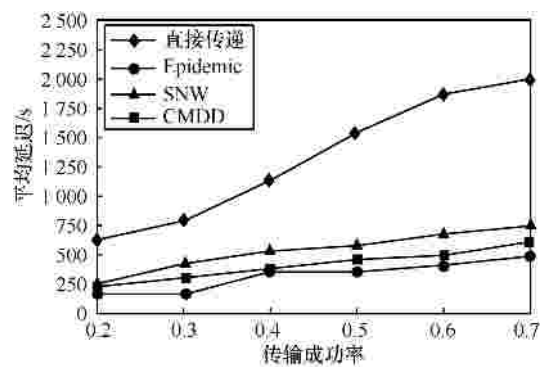


图 6 平均传输延迟

图 7 为不同节点数下，各个策略的平均冗余传输值。洪泛策略的平均冗余传输值大于其他 3 个策略，并随节点数线性增加；直接传递策略的平均冗余传输值为 0，因为该策略直到遇到目标节点才转发

数据；CMDD 和 SNW 策略的平均冗余传输值基本相同，等于该策略中限定的副本数量，由于本组实验副本数量在[5,10]内随机产生，因此该值约为 7.5。

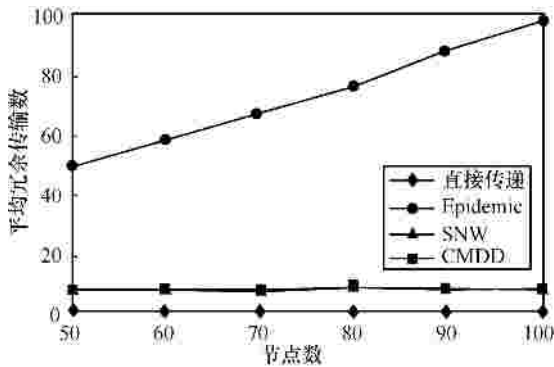


图 7 平均冗余传输值

通过本组实验发现，与洪泛、直接传递、SNW 策略相比，CMDD 策略在获得较高的传输成功率下，传输延迟较低、传输开销较小。

### 6.3 副本数对策略性能的影响

SNW 和 CMDD 策略都是基于复制的数据传输策略，本组实验考察副本数对策略性能的影响。本组实验设置网络中的社团数为 3，每个社团中的节点数为[10,20]内的随机值，社团内部节点的相遇间隔时间服从均值为 300 s 的均匀分布，社团之间节点的相遇间隔时间服从均值为  $k_i \times 300$  s 的均匀分布，其中源节点所在社团和目的节点所在社团系数  $k_1=8$ ，中间社团和源节点所在社团以及目的节点所在社团系数  $k_2=k_3=2$ 。

图 8 表明增加副本的数量可以提高 SNW 和 CMDD 策略的传输成功率，在同样的副本数下，CMDD 策略的平均传输成功率明显高于 SNW 策略。这是因为，在本组实验中，源节点所在社团和目的节点所在社团相遇的平均间隔时间大 ( $k_1=8$ )，SNW 策略主要将副本分发给了源节点社团中的节点，这些节点与目标节点相遇需要花费较长的时间；而 CMDD 策略选择通过中间社团进行副本分发，中间社团与源节点所在社团和目的节点所在社团相遇较频繁 ( $k_2=k_3=2$ )，通过中间社团分发提高了传输效率。因此，在同样的时间内，CMDD 策略的传输成功率高于 SNW 策略。

图 9 为在同样的传输成功率为 0.5 的情况下，CMDD 和 SNW 策略的平均传输延迟随副本数变化的结果。从图中可以看到，随着副本数的增加，2 个策略的平均传输延迟都减小，并且在源节点社团

与目标节点社团关系疏远的情况下，CMDD 策略的平均传输延迟大大低于 SNW 策略。

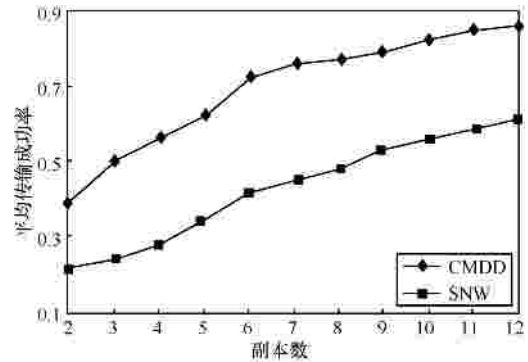


图 8 不同副本数下的平均传输成功率

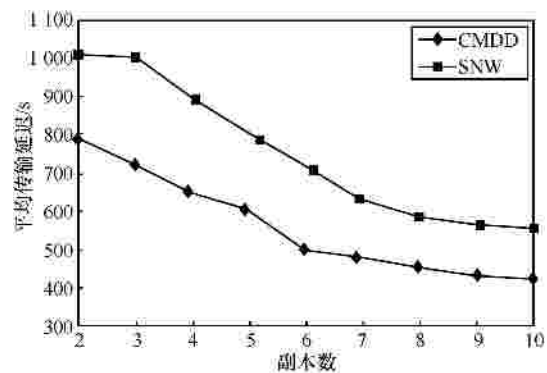


图 9 不同副本数下的平均传输延迟

图 10 为 CMDD 策略的平均传输延迟随传输成功率和副本数变化的情况。图 10 表明，在同样的副本数下，要获得更高的传输成功率都会增加传输延迟。在获得同样的传输成功率下，副本数越多传输延迟越小。因此，副本数量对基于复制的冗余传输策略性能有重要影响。

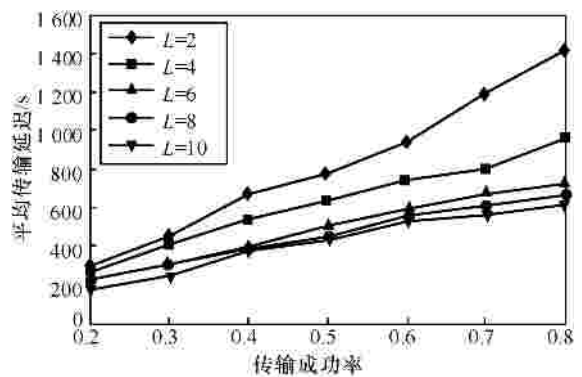


图 10 副本数对平均传输延迟的影响

## 7 结束语

在人携带传感器节点这一类容迟移动传感器

网络的应用中，人类社会的各种属性如移动性、社会网络结构及社会特征会对传感器节点的数据传输性能造成影响。

本文研究了社团网络结构对多副本数据传输的影响。在社团结构网络模型下，通过建立社团间多副本数据分发过程的数学模型，得到副本通过不同的社团进行转发时可获得的成功传输概率。提出了一种基于社团结构的多副本数据传输策略 CMDD，与直接传递、洪泛策略、SNW 策略相比，能以较低的传输延迟获得较高的数据传输成功率，并且传输开销较低。

在基于复制的冗余传输策略中，合理设置副本数量可以在传输性能和传输开销两方面取得较好的平衡。进一步工作将分析如何在各个社团中分配副本，以及如何充分利用人类的社会属性以提高 DTMSN 传输策略的性能。

#### 参考文献：

- [1] FALL K. A delay-tolerant network architecture for challenged internets[A]. Proceedings of SIGCOMM'03[C]. Karlsruhe, Germany, 2003.27-34.
- [2] WANG Y, WU H Y. Replication-based efficient data delivery scheme (RED) for delay/fault-tolerant mobile sensor network (DFT-MSN)[A]. Proceedings of the 4th Annual IEEE International Conference on Pervasive Computing and Communications Workshops[C]. Pisa, Italy, 2006.
- [3] LINDGREN A, DORIA A, SCHELEN O. Probabilistic routing in intermittently connected networks[J]. ACM SIGMOBILE Mobile Computing and Communications Review, 2003, 7(3):19-20.
- [4] BURGESS J, GALLAGHER B, JENSEN D, *et al.* MaxProp: routing for vehicle-based disruption-tolerant networks[A]. Proceedings of IEEE Infocom[C]. Barcelona, Catalunya, Spain, 2006.
- [5] BALASUBRAMANIAN A, LEVINE B N, VENKATARAMANI A. DTN routing as a resource allocation problem[A]. Proceedings of ACM SIGCOMM 07[C]. Kyoto, Japan, 2007.
- [6] KARVO J, OTT J. Time scales in delay-tolerant routing protocols[A]. Proceedings of the Third ACM Workshop on Challenged Networks[C]. San Francisco, California, USA, 2008.
- [7] DALY E, HAAHR M. Social network analysis for routing in disconnected delay-tolerant MANETs[A]. Proceedings of the ACM MobiHoc07[C]. Montreal, Quebec, Canada, 2007.
- [8] PAN H, CROWCROFT J, YONEKI E. BUBBLE rap: social-based forwarding in delay tolerant networks[J]. IEEE Transactions on Mobile Computing, 2011,10(11):1576-1589.
- [9] BULUT E, SZYMANSKI B K. Friendship based routing in delay tolerant mobile social networks[A]. Proceedings of IEEE Globecom 2010[C]. Miami, Florida, USA, 2010.
- [10] ABDELKADER T, NAIK K E. A socially-based routing protocol for delay tolerant networks[A]. Proceedings of IEEE Globecom 2010[C]. Miami, Florida, USA, 2010.
- [11] BULUT E, WANG Z J, BOLESZAW K S. Impact of social networks in delay tolerant routing[A]. Proceedings of IEEE GLOBECOM 2009[C]. Honolulu, USA, 2009.
- [12] 熊永平, 孙利民, 牛建伟等. 机会网络[J]. 软件学报, 2009, 20(1): 124-137.  
XIONG Y P, SUN L M, NIU J W, *et al.* Opportunistic networks[J]. Journal of Software, 2009, 20(1): 124-137.
- [13] VAHDAT A, BECKER D. Epidemic Routing for Partially Connected Ad Hoc Networks[R]. Technical Report, VoLCS-200006, Durham, Duke University, 2000.
- [14] WANG Y, WU H. Delay/fault-tolerant mobile sensor network (DFT-MSN): a new paradigm for pervasive information gathering[J]. IEEE Trans on Mobile Computing, 2006, 6(8):1021-1034.
- [15] SMALL T, HAAS Z J. The shared wireless info station model—a new ad hoc networking paradigm[A]. Proceedings of MobiHoc 2003[C]. San Diego, CA, USA, 2003.
- [16] SPYROPOULOS T, PSOUNIS K, RAGHAVENDRA C. Efficient routing in intermittently connected mobile networks: the multiple-copy case[J]. IEEE/ACM Transactions on Networking, 2008, 16(1):77-90.
- [17] 许富龙, 刘明, 龚海刚等. 延迟容忍传感器网络基于相对距离的数据传输[J]. 软件学报, 2010,21(3):490-504.  
XU F L, LIU M, GONG H G, *et al.* Relative distance-aware data delivery scheme for delay tolerant mobile sensor networks[J]. Journal of Software, 2010,21(3):490-504.
- [18] PIORKOWSKI M, SARAFIJANOVIC-DJUKIC N, GROSSGLAUSER M. A parsimonious model of mobile partitioned networks with clustering[A]. Proceedings of COMSNETS 2009[C]. Bangalore, India, 2009.
- [19] LEE C H, DO Y E. Exploiting heterogeneity in mobile opportunistic networks: an analytic approach[A]. Proceedings of 7th Annual IEEE Communications Society Conference on Sensor Mesh and Ad Hoc Communications and Networks (SECON)[C]. Boston, Massachusetts, USA, 2010.
- [20] LEE C H, EUNT D Y. Heterogeneity in contact dynamics: helpful or harmful to forwarding algorithms in DTN[A]. Proceedings of the 7th International Conference on Modeling and Optimization in Mobile, Ad Hoc and Wireless Networks[C]. Seoul, Korea, 2009.
- [21] KERANEN A, KARKKAINEN T, OTT J. Simulating mobility and DTN with the ONE[J]. Journal of Communications, 2010, 5(2):92-105.

#### 作者简介：



付彬 (1978-)，女，广东英德人，湖南大学博士生、讲师，主要研究方向为无线网络、无线传感器网络。

李仁发 (1957-)，男，湖南郴州人，博士，湖南大学教授、博士生导师，主要研究方向为嵌入式计算、无线网络、CPS、虚拟与仿真。

肖雄仁 (1978-)，男，湖南长沙人，湖南大学博士生、讲师，主要研究方向为嵌入式系统、计算机网络。

黄亚薇 (1991-)，女，江西南昌人，重庆理工大学本科生，主要研究方向为计算机网络。