

机会网络中基于有权社团结构图的路由协议研究

马学彬,白 婧,郑田玉

(内蒙古大学计算机学院,内蒙古呼和浩特 010021)

摘 要: 基于社团检测的机会网络路由算法大多采用无权重网络拓扑划分社团,仅将节点间的关系抽象为一条简单的无权重的边,忽略了节点关系的强弱程度.本文通过引入权重策略改进了QCA社团更新算法,提出了一种基于有权社团结构的路由算法,该算法解决了社团关系定量化单一的问题,更能真实反映出社团成员之间的关系.算法中,节点间的交互信息转化为权重,根据不同的网络环境选择不同的权重转化方案——归一化权重(normalized weight)和非归一化权重(non-normalized weight).路由算法在检测到周围网络环境变化时自动切换权重计算方案以适应网络环境的变化.通过在仿真环境和真实数据集上测试和分析,该算法能够将网络中的节点划分出合理的社团结构,并在保证较高的传输成功率的情况下降低网络开销.

关键词: 机会网络; 社团划分; 路由算法; 有权拓扑

中图分类号: TN92 **文献标识码:** A **文章编号:** 0372-2112 (2016)10-2449-10

电子学报 URL: <http://www.ejournal.org.cn> **DOI:** 10.3969/j.issn.0372-2112.2016.10.024

A Routing Algorithm Based on Weighted Community Detection for Opportunistic Networks

MA Xue-bin, BAI Jing, ZHENG Tian-yu

(School of Computer, Inner Mongolia University, Hohhot, Inner Mongolia 010021, China)

Abstract: Most of the opportunistic networks routing algorithms based on community detection use an un-weighted network which ignores the degree of intensity of relations between nodes. This paper proposes a routing algorithm based on community detection of weighted network. We improve quick community adaptation (QCA) and make it adapt to the opportunistic networks by using weighted networks. The algorithm calculates link weights by the connection information between nodes in the network. According to the different network environments, we present two weight calculation strategies: normalized weight strategy and non-normalized weight strategy. The algorithm detects the environment around the current node, and then chooses the right strategy. To illustrate the performance of our algorithm, we test the algorithm by using a simulation environment and a real dataset. The results demonstrate that our algorithm gets a reasonable community structure and reduces the overhead ratio and keeps a higher delivery probability.

Key words: opportunistic networks; community detection; routing algorithm; weighted topology

1 引言

在机会网络^[1]中,人们随身携带的智能移动设备构成了网络中的节点,并且随着人的移动而移动.由于人作为社会群体中的一员体现出社会性特征,所以这些移动设备也具有一定的社会性特征.社团划分就是基于网络中节点间的相互关系将它们划分到各个社团,并且同一社团内成员间的交互概率会明显高于社团间成员的交互概率^[2-4].利用节点间的这个特点可以

为选择转发节点提供极为重要的参考依据,因此划分网络的社团结构用于数据转发,对提高机会网络中的路由性能具有重大意义.

本文提出了一种基于社团划分方法的路由算法.在机会网络中,大部分基于社团划分的路由算法只是将网络中两个节点的连接简单映射为网络拓扑中一条无权重的边,忽视了节点间的连接所反映的节点间关系强弱的信息.根据节点间的连接信息可以分析出它们间的关系强度以及节点移动模式的共同特征^[5],利

用这些信息可以划分出高质量的社团结构并因此提高路由性能^[6].

2 相关工作

机会网络中,有关社团划分的路由算法多是从复杂网络的社团划分算法演变而来的,比较经典的算法包括 K 派系过滤算法^[7-9]、GN 算法^[10]等等.在文献[7]中,作者使用 K-clique 算法通过添加一些社会信息来检测社会网络,并形成具有自报告功能的社会网络.研究发现通过使用自报告的社会网络信息虽然不能明显的提升消息传输成功率,但是能有效地降低开销.在文献[8]中,作者改进了文献[7]中的想法,在两个不同的场景下使用 K-clique 算法,分别是室内场景和室外场景.作者发现能够影响机会网络解决方案和算法的关键是社会信息和移动行为信息,网络性能能够通过增加社会网络参与者的社会联系来增加.同时,通过组合这两种信息,可以提高传输成功率,并且几乎没有影响交付延迟.在文献[9]中,Fu 提出了一个基于联合发现结构的改进的 k-clique 发现算法,它缩短了社团发现的时间,其划分社团的时间复杂度接近线性.Newman 于 2004 年提出的模块度 (Modularity)^[10,11]更是成为衡量社团结构划分质量的一个重要的指标.由于 Newman 提出的快速算法时间复杂度较高,改进的算法引入了模拟退火算法、禁忌搜索算法、退火算法等来加快计算速度.

目前,对于社团划分算法方面的研究,如何快速高效、高精度地划分社团仍是研究热点之一.此外,研究者从多个角度去分析社团结构,进而提高划分精度.模块度、节点的社会属性,节点的私人属性等方面已经成为研究社团结构需要考虑的因素.Nam 等人^[12]提出了一种在动态社交网络中自适应更新社团结构的算法 Quick Community Adaptation(QCA).通过推导简化了模块度计算的复杂度,形成了一个新的社团更新算法.QCA 根据网络变化实时更新节点的所属社团,最终得到高质量的社团划分结果.该算法在运行时间上有显著的提高,适合划分大型的、快速变化的网络上的社团结构,如在线社交网络、人际关系网等.Costa 等人^[13]提出了一种在机会网络中使用发布/订阅机制的路由算法——SocialCast 路由算法.在该路由算法中,作者建立了社会网络模型和移动模型,将社会网络中节点的社会性和移动性以量化的方式转化为协同定位和连接度的改变.该路由算法只将消息发布给订阅它的节点,而具有相同兴趣的节点会聚集在一起,这能够有效地控制网络中的副本数量,并通过社交矩阵预测并选择最佳的消息载体.该路由算法一共分为兴趣分发、中继选择、消息发布三个阶段.该算法适合具有小世界特性的

网络和很难获取全局信息的网络.Pan Hui 等^[14]提出一种基于社会属性的路由算法 Bubble Rap.该路由算法根据目的地社团和节点中心度的排名来选择转发节点.在该算法中,每个节点至少属于一个社区,每个节点有两个中心性,一个是本地中心性,一个是全局中心性.整个路由过程分为两个阶段,全局转发和局部转发.在全局转发阶段,该算法尽最大努力使消息到达目的社团;在局部转发阶段,该算法将把消息转发到目的节点.通过这两个阶段的相互配合使消息快速高效地到达目的节点.但是该算法在所有节点的全局中心性都很低时可能失效.Bulut 等^[15]提出了基于朋友关系的路由,在相遇次数、相遇持续时间、相遇规律这三个维度定义朋友的特性,并且基于朋友关系的质量来选择转发节点,通过将消息交付给高质量的朋友来增加消息的传输成功率.Mei 等提出了一种社会性敏感的无状态路由方法 SANE (Social-Aware and Stateless Routing)^[16].在这种路由算法中一个人的兴趣由 k 维的兴趣向量表示,两个用户 u 和 v 之间的兴趣相似性用向量间的 Cosine 相似性来表示.在路由过程中,消息被转发给与目的地兴趣最相似的节点.该算法维护和更新社会度量相对简单,占用空间小,时间复杂度和空间复杂度都很低.在文献[17]中,Xuebin M 提出了一个基于节点间耦合程度的路由算法 CR,每个节点根据它与周围相遇节点的相似性确定它们属于哪个社团,并根据消息的目的节点选择与其最属性接近的社团进行转发,最终将消息发送到目的社团并进一步转发到目的节点.该算法通过记录的接触历史信息计算关系强度及其节点间的耦合程度,使得该算法的计算复杂度和时间复杂度都较低.实验结果表明,该算法减少了转发的次数,提高了传输成功率.在文献[18]中,Theus Hossmann 研究了四个数据集和三个先进的合成数据的社团结构、图谱、社团内部以及社团间的权重分布.在此基础上作者提出一个框架用于分析基于 SNA 的 DTN 路由方案的性能.在文献[19]中提出了一个新的路由算法 NBR,该算法使用一个加入节点回归因素的模型.在社区内采用混合路由算法,并加入了正反馈思想计算节点活跃度来寻找目的节点;在社区间采用查询路由表和判断节点回归相结合的方法进行消息转发.该算法与其他算法相比,不仅提高了传输成功率,也降低了开销.

上面所述路由算法虽然从不同角度解决了社团划分和机会网络中基于社团结构的路由算法问题,但是他们大多只考虑了影响社团结构的一个或几个因素,并没有综合考虑各种因素.本文通过给出一个通用的社团划分框架,可以把各种因素转换为边之间的权重,通过划分社团结构来提高机会网络的路由性能.该路由算法还考虑到不同拓扑结构的特点,自主选择不同

的权重计算方法,使得划分的社团更准确、更高效.

3 基于权重的社团更新算法

本文基于 Nam P Nguyen^[12]等人的工作,对其提出的 QCA 算法进行改进,将其应用于机会网络并引入权重概念以提升算法性能.对于单个节点来说,在无权重网络拓扑中,它涉及的网络拓扑的变动有两种,分别是边的添加和删除.而对于有权网络拓扑,边的添加和删除相对应的转变为边权重的增加和减少,因此 QCA 社团更新算法^[12]中的计算公式不再适合本文中的情况,需要重新推导.接下来会用到的一些概念和符号如下.

有权网络拓扑表示为 $G = (V, E, w)$, 这里 V 是网络拓扑中的节点集合, $E = \{(u, v) | u \in V, v \in V\}$ 是网络拓扑 G 的边集合, $w(u, v)$ 是节点 u 和 v 之间的边的权重.

本文采用模块度公式 $Q = \sum (\frac{m_c}{M} - \frac{d_c^2}{4M^2})$ ^[11], 其中 m_c 是社团 C 内部所有的边权重的总和, d_c 是 C 中所有顶点的度的总和(即 C 中每个顶点的每条边的权重相加的和), M 是网络拓扑中所有的边权重的总和. e_c^u 是节点 u 与 C 中所有节点连接边的权重的总和. 在 t 时刻, 网络拓扑为 G_t , $t+1$ 时刻, 网络拓扑经过 ΔG 的变化变为 G_{t+1} , 即 $G_t \xrightarrow{\Delta G} G_{t+1}$. ΔG 为边的权重的增加或减少和节点加入或离开的集合.

假设整个网络拓扑中所有的边权重都是大于 0 的, 且整个网络中均匀分布多个社团(四个以上), 即 $w_{ij} \geq 0, M > 0, d_c > 0, M$ 大于任意 d_c , 两个节点间的边的权重由 w_1 变为 w_2 , 变化量为 $|\Delta w|$. 经过推导, 可以得到以下几条结论:

(1) 社团内的边权重的增加能使该社团对全局模块度的贡献增加.

证明

$$\text{权重变化之前的模块度 } Q' = (\frac{m_c}{M} - \frac{d_c^2}{4M^2}) + A$$

$$\text{权重变化后的模块度 } Q = (\frac{m_c + \Delta w}{M + \Delta w} - \frac{(d_c + 2\Delta w)^2}{4(M + \Delta w)^2}) + A$$

$$\begin{aligned} Q' - Q &= (\frac{m_c + \Delta w}{M + \Delta w} - \frac{(d_c + 2\Delta w)^2}{4(M + \Delta w)^2}) - (\frac{m_c}{M} - \frac{d_c^2}{4M^2}) \\ &= \frac{4M^3 \Delta w - 4M^2 d_c \Delta w - 4M^2 m_c \Delta w + 2Md_c^2 \Delta w}{4M^2 (M + \Delta w)^2} \\ &\quad - \frac{4Mm_c \Delta w^2 - d_c^2 \Delta w^2}{4M^2 (M + \Delta w)^2} \\ &\geq \frac{4M^3 \Delta w - 6M^2 d_c \Delta w + 2Md_c^2 \Delta w - 2Md_c \Delta w^2 + d_c^2 \Delta w^2}{4M^2 (M + \Delta w)^2} \end{aligned}$$

$$\times (m_c < \frac{d_c}{2})$$

$$= \Delta w \frac{(4M^3 - 6M^2 d_c + 2Md_c^2 - 2Md_c \Delta w + d_c^2 \Delta w)}{4M^2 (M + \Delta w)^2}$$

$$= \Delta w \frac{(2M^2 - 2Md_c - d_c \Delta w)(2M - d_c)}{4M^2 (M + \Delta w)^2}$$

若权重变化后模块度增加, 则需满足条件 $Q' - Q > 0$

因 $\Delta w > 0$, 可知若要 $Q' - Q > 0$, 需要 $(2M^2 - 2Md_c - d_c \Delta w)(2M - d_c) > 0$,

$$\text{即 } \begin{cases} 2M^2 - 2Md_c - d_c \Delta w > 0 \\ 2M - d_c > 0 \end{cases},$$

$$\text{或 } \begin{cases} 2M^2 - 2Md_c - d_c \Delta w < 0 \\ 2M - d_c < 0 \end{cases}$$

$$\text{① } \begin{cases} 2M^2 - 2Md_c - d_c \Delta w > 0 \\ 2M - d_c > 0 \\ \Delta w > 0 \end{cases}$$

$$\Rightarrow \begin{cases} 0 < \Delta w < 2M(\frac{M}{d_c} - 1) \\ d_c < 2M \end{cases}$$

$$\Rightarrow \begin{cases} 0 < \Delta w < 2M(\frac{M}{d_c} - 1) \\ 2M(\frac{M}{d_c} - 1) > 0 \\ d_c < 2M \end{cases} \Rightarrow \begin{cases} 0 < \Delta w < 2M(\frac{M}{d_c} - 1) \\ d_c < M \end{cases}$$

$$\text{② } \begin{cases} 2M^2 - 2Md_c - d_c \Delta w < 0 \\ 2M - d_c < 0 \\ \Delta w > 0 \end{cases} \Rightarrow \begin{cases} \Delta w > 2M(\frac{M}{d_c} - 1) \\ \Delta w > 0 \\ d_c > 2M \end{cases}$$

因为 $2M$ 为整个网络中所有节点的度的和, 所以一个社团内的节点的度的和不可能大于 $2M$, 因此 $d_c > 2M$ 不成立, 不存在上面这一种情况.

由上面的推导①可以看出, 只要网络中均匀分布两个以上的社团, 且权重的增加不是特别大的话, 社团内部的边权重的增加必然能够使该社团的模块度增大. 由此可知, 一般情况下, 社团内的边权重的增加能使该社团模块度的增加.

通过上面的推导可知, 一般情况下社团内的边权重的减少能够使该社团对全局模块度的贡献减少.

(2) 两个社团间的边权重减少时能够提高两个社团对全局模块度的贡献.

证明 由结论(1)可反向推导得出.

(3) 当社团内部两个子社团间的连接符合 $\frac{d_1 d_2}{2M} < e_{12}$

$< \Delta w + \frac{d_1 d_2 + d_c \Delta w + \Delta w^2}{2(M + \Delta w)}$ 时, 该社团应该被分裂为两个子社团.

证明 假设社团 C 内部边权重的减少会导致该社团分裂为 C_1 和 C_2 , 那么权重变化前模块度应该符合下面公式

$$Q_1 + Q_2 < Q$$

$$\frac{m_1}{M} - \frac{d_1^2}{4M^2} + \frac{m_2}{M} - \frac{d_2^2}{4M^2} < \frac{m_1 + m_2 + e_{12}}{M} - \frac{(d_1 + d_2)^2}{4M^2}$$

$$e_{12} > \frac{d_1 d_2}{2M}$$

权重减少后模块度应该符合下面公式

$$Q'_1 + Q'_2 > Q'$$

$$e_{12} < \Delta w + \frac{d_1 d_2 + d_c \Delta w + \Delta w^2}{2(M + \Delta w)}$$

因此, 当社团内部两个子社团间的连接符合 $\frac{d_1 d_2}{2M} < e_{12} < \Delta w + \frac{d_1 d_2 + d_c \Delta w + \Delta w^2}{2(M + \Delta w)}$ 时, 该社团应该被分裂为两个子社团.

(4) 当社团内的一条边权重减少, 并且这条边的两端节点中至少有一个节点只有一条边与其相连, 则这条边的变化不会分裂两个社团.

证明

若社团内的一条边 (u, v) 的权重减少 $|\Delta w|$, 且这条边两端的两个节点 u 和 v 中至少有一个的度为 1, 若这条边的权重变化会导致本地社团 C 的分裂, 则应该满足以下两个条件权重变化前的模块度

$$Q_1 + Q_2 < Q$$

$$\frac{m_1}{M} - \frac{d_1^2}{4M^2} + \frac{m_2}{M} - \frac{d_2^2}{4M^2} < \frac{m_1 + m_2 + e_{12}}{M} - \frac{(d_1 + d_2)^2}{4M^2}$$

$$e_{12} > \frac{d_1 d_2}{2M}$$

权重变化后的模块度

$$Q'_1 + Q'_2 > Q'$$

$$e_{12} < \frac{d_1(d_2 + \Delta w)}{2(M + \Delta w)} = \frac{d_1 d_2 + d_1 \Delta w}{2(M + \Delta w)}$$

即

$$\frac{d_1 d_2}{2M} < e_{12} < \frac{d_1(d_2 + \Delta w)}{2(M + \Delta w)} = \frac{d_1 d_2 + d_1 \Delta w}{2(M + \Delta w)}$$

因为

$$\frac{d_1 d_2 + d_1 \Delta w}{M + \Delta w} - \frac{d_1 d_2}{M} = \frac{d_1 \Delta w (M - d_2)}{M(M + \Delta w)} < 0 (\Delta w < 0)$$

$$\text{所以 } \frac{d_1 d_2 + d_1 \Delta w}{M + \Delta w} < \frac{d_1 d_2}{M}, \text{ 则 } \frac{d_1 d_2}{2M} < e_{12} < \frac{d_1(d_2 + \Delta w)}{2(M + \Delta w)}$$

不成立.

由上面推导可知, 当社团内一条权重减少的边的两端的节点只要其中一个的度为 1, 则这条边的变化不会分裂两个社团.

(5) 当两个社团间一条边的权重增长时, $F_{\text{out}}^D(u)$ 会增强, $F_{\text{in}}^C(u)$ 会减弱, 则这种情况可能导致节点 u 所属社团的变化.

对于处于社团 C 中的节点 u , 若它与社团 D 间的一条边的权重增加了 $\Delta w (\Delta w > 0)$, 很明显可以看出节点 u 连接社团 C 和社团 D 间的边权重的变化影响了社团 C 对节点 u 的 F_{in}^C 和社团 D 对节点 u 的 F_{out}^D .

证明

在边权重变化之前, 节点 u 的 $F_{\text{in}}^C(u)$ 为 $e_c^u - \frac{d_u(d_c - d_u)}{2M}$, $F_{\text{out}}^D(u)$ 为 $e_d^u - \frac{d_u d_{\text{out}D}}{2M}$. 边权重变化之后, 节

点 u 的 $F_{\text{in}}^C(u)$ 变为 $e_c^u - \frac{(d_u + \Delta w)(d_c - d_u)}{2(M + \Delta w)}$.

$$F_{\text{out}}^D(u) \text{ 变为 } (e_d^u + \Delta w) - \frac{(d_u + \Delta w)(d_{\text{out}D} + \Delta w)}{2(M + \Delta w)}.$$

$$F_{\text{out}}^D(u)_{\text{new}} - F_{\text{out}}^D(u)_{\text{old}} = \frac{2M\Delta w + d_u d_{\text{out}D}}{2M} - \frac{d_u d_{\text{out}D} + d_{\text{out}D}\Delta w + d_u \Delta w + \Delta w^2}{2(M + \Delta w)}$$

$$\geq \frac{2M\Delta w + d_u d_{\text{out}D}}{2(M + \Delta w)} - \frac{d_u d_{\text{out}D} + d_{\text{out}D}\Delta w + d_u \Delta w + \Delta w^2}{2(M + \Delta w)}$$

$$= \Delta w \frac{2M - d_{\text{out}D} - d_u - \Delta w}{2(M + \Delta w)}$$

已知 $M > 0, d_{\text{out}D} > 0$, 且 $2M = \sum d_i$, 可推出 $2M > d_{\text{out}D} + d_u + \Delta w$ 必然存在. 且 $\Delta w > 0$, 则 $F_{\text{out}}^D(u)_{\text{new}} - F_{\text{out}}^D(u)_{\text{old}} > 0$. 由此可以看出, 社团间的边权重增加后, 与之相连的节点 u 的 $F_{\text{out}}^D(u)$ 变大.

$$F_{\text{in}}^C(u)_{\text{new}} - F_{\text{in}}^C(u)_{\text{old}} = \Delta w (d_c - d_u) \frac{d_u - M}{2M(M + \Delta w)}$$

已知 $\Delta w > 0, M > 0, d_c - d_u > 0$, 且 M 为网络中所有边的总和必然大于 d_u , 可推导出 $d_u - M$ 必然小于 0, 则 $F_{\text{in}}^C(u)_{\text{new}} - F_{\text{in}}^C(u)_{\text{old}} < 0$. 由此可以看出, 社团间的边权重增加后, 与之相连的节点 u 的 $F_{\text{in}}^C(u)$ 会变小.

(6) 当社团间的一条边权重增加时, 若这条边所连接的一端的节点符合公式

$$4(M + \Delta w)(e_d^u - e_c^u + \Delta w) + (d_u + \Delta w)(d_c - d_u - d_d) - \Delta w^2 > 0,$$

则该节点应该加入到另一端的社团. 若两端的节点都符合该公式, 则选择值较大的一段加入.

证明

当两个社团间一条边的权重增长时, 节点若留在原来的社团 C 中时, 模块度为

$$Q_C + Q_D$$

$$= \left(\frac{m_c}{M + \Delta w} - \frac{(d_c + \Delta w)^2}{4(M + \Delta w)^2} \right) + \left(\frac{m_d}{M + \Delta w} - \frac{(d_d + \Delta w)^2}{4(M + \Delta w)^2} \right)$$

若节点 u 离开社团 C , 加入社团 D , 模块度变为

$$Q_{C-u} + Q_{D+u} \\ = \left(\frac{m_c - e_c^u}{M + \Delta w} - \frac{(d_c - d_u)^2}{4(M + \Delta w)^2} \right) \\ + \left(\frac{m_d + e_d^u + \Delta w}{M + \Delta w} - \frac{(d_d + d_u + 2\Delta w)^2}{4(M + \Delta w)^2} \right)$$

若要证明节点 u 加入社团 D 后模块度会增加, 只需证明

$$Q_{C-u} + Q_{D+u} > Q_C + Q_D$$

即

$$4(M + \Delta w)(e_d^u - e_c^u + \Delta w) + (d_u + \Delta w)(d_c - d_u - d_d) - \Delta w^2 > 0$$

本文所提出的基于权重的社团更新算法如算法 1 所示. 其中, $\text{increaseWeight}(u, v, \Delta w)$ 算法如算法 2 所示, 算法中使用的 $\text{decreaseWeight}(u, v, \Delta w)$ 、 $\text{newNode}(u)$ 、 $\text{removeNode}(u)$ 三种方法的内部实现方式与 QCA 算法中类似, 主要区别在于使用的推导公式不同, 因此这里不做赘述.

算法 1 WQCA

Input: $G_t \xrightarrow{\Delta G} G_{t+1}, \Delta G = \{\xi_1, \xi_2, \dots, \xi_n\}$

Output: Community which node u belongs to.

Begin

For ΔG do

If $\xi_i = \text{changeWeight}(v)$ then

If $\Delta w > 0$ then

increaseWeight($u, v, \Delta w$)

Else If $\Delta w < 0$ then

decreaseWeight($u, v, \Delta w$)

End If

Else If $\xi_i = \text{newNode}(v)$ then

newNode(v)

Else If $\xi_i = \text{removeNode}(v)$ then

removeNode(v)

End If

End For

End

算法 2 increaseWeight($u, v, \Delta w$)

Begin

If $C(u) \neq C(v)$ then

$\Delta q = \max \{ \Delta q_{u, C(u), C(v)}, \Delta q_{v, C(u), C(v)} \}$

$w = \arg \max \{ \Delta q_{u, C(u), C(v)}, \Delta q_{v, C(u), C(v)} \}$

If $\Delta q > 0$ then

w join into another community

End If

End If

End

4 基于权重的社团划分路由协议

基于权重的社团划分路由协议分为信息采集、权重转化和路由转发三部分. 其中信息采集和权重转化为社团划分的前期工作, 它们将机会网络中的节点接触信息转化为边权重, 并形成有权网络拓扑图. 社团划分算法就是在这两部分工作的基础上运行的. 消息的路由转发则是根据社团划分算法得到的结果和权重等信息判断转发节点. 社团划分算法流程图如图 1 所示.

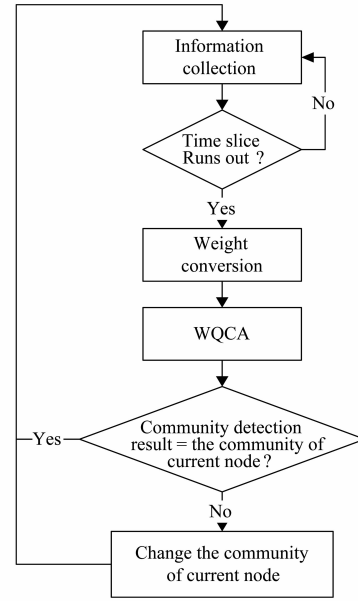


图1 社团划分算法流程图

4.1 信息采集

节点采集信息的质量十分重要, 会直接影响到之后的社团划分和路由性能. 为了保留节点间的强连接以及剔除一些无意义的弱连接(例如两个节点间只存在一次连接时间极短的连接等情况), 需要采用合适大小的时间窗口和合适的策略对连接进行筛选. 如果选择不恰当, 会造成网络拓扑越来越稠密, 经过一段时间后网络拓扑不再能明确的反映出节点间的关系, 甚至网络拓扑有可能变为连通图. 本路由协议提出两种方式, 在规定的窗口内以最新的最近连接(Most Recent Connection, MRC)和最新的最频繁连接(Most Frequent Connection, MFC)两种方式筛选部分连接^[20]. 最新的最近连接(MRC)是根据表中记录的节点最近连接时间选择距离当前时间最远的一部分记录删除. 最新最频繁连接(MFC)则是选择当前时间片内连接次数最少的一部分记录删除.

在本路由协议中, 网络中的每个节点都要维护一个

记录自身与其他节点的连接信息的网络拓扑信息记录表。每个节点以其全局唯一的 ID 的形式存放在网络拓扑信息记录表中。网络拓扑信息记录表中的每一项都记录着另一节点与本节点间的连接次数、连接时长、两节点间的权重等信息,在社团划分时根据表中记录的连接信息计算两个节点之间的边的权重并记录在表中相应的字段。本文所用到的社团更新算法是根据网络拓扑的变动更新社团,但是考虑到若是每一次网络拓扑变动都更新社团结构会造成大量的计算以及节点间的社团记录无法同步等问题,因此本文采用时间片的形式更新社团结构,每当时间片结束时调用社团更新算法更新网络中的社团结构。在网络拓扑信息记录表存在一个名为标志位的字段,该字段用来记录网络拓扑在时间片内的变动情况,具体来说标志位包含四个值 new、delete、change、old。下面为标志位变动的几种情况:

算法 3 信息采集算法

```

Begin
  Before a time slice running
  For The Connection Information List do
    If conItem.flag = "delete" then
      remove conItem from the list
    Else If conItem.flag = "new" or "change" then
      conItem.flag = "old"
    End If
  End For
  Time slice running
  For < ConnectionNode > con do
    If con doesn't exist in the connection information list then
      add con into the list and conItem.flag = "new"
    Else
      conItem.flag = "change"
    End If
  End For
  Time slice runs out
  choose remove conItem in the connection information by MRC
  or MFC
  removeItem.flag = "delete"
End

```

在网络拓扑记录表中的节点表示当前节点与表中记录的其它节点间存在边,这条边的权重根据所采集的连接信息计算。另外,每个节点保存着一个黑名单列表(BlackList),用来记录拒绝接收转发消息的节点。信息采集阶段的具体算法如算法 3 所示。

4.2 权重转化

在不同的网络环境下,权重计算公式应该相应的改变以维持社团更新算法的最优性能。对于节点密度比较高接触时间长而且频繁的网络环境,节点间的边

的权重值会比较高。而非归一化权重在这种环境下能够保持较高的分辨率,划分出合理的社团结构,不会将这些高权重的节点都划分到一个社团。而在一般的情况下,归一化权重能够更敏感的察觉网络拓扑的变化,得到合理的社团划分结果。

对于网络节点间的边权重则是由节点间的接触信息转化。节点间的接触信息包括节点间的总连接时长、连接次数、黑名单事件、发送失败等等。本文对边权重的定义包含两部正权重和负权重,正权重和负权重相加得到的和为两个节点间的边权重,即 $w = w_1 + w_2$ 。 w_1 为正权重,由节点间的平均连接时间计算,其中非归一化的权重计

算公式为 $w_{u,v} = \frac{t_D}{n_c}$, 归一化的权重计算公式为 $w_{u,v,normalized}$

$= \frac{w_{u,v}}{t_{totalD}}$ 。 t_D 为节点 u, v 之间总的持续连接时间, n_c 为节点

u, v 之间总的连接次数。 $w_{u,v}$ 为节点 u 和节点 v 之间的非归一化正权重,即节点 u, v 之间的平均持续连接时间。

$w_{u,v,normalized}$ 为节点 u 和节点 v 之间的归一化边权重, t_{totalD}

为所有与节点 u 连接的节点的平均持续连接时间的总和。

节点根据与周围节点的连接密度选择相应的计算公式。 w_2 为负权重,根据发生的黑名单、信息发送失败等事件

衰减两节点间的权重。在实际生活中,节点之间可能会发生一些主动拒绝接受或转发来自特定节点的消息,这种情况被称为黑名单事件。

为了模拟这种情况,在路由算法中加入了一个黑名单队列,当发生黑名单事件时将拒绝来往的节点加入到此队列中。

然后,在之后每次与其他节点发生连接时都首先检查该节点是否存在于黑名单队列中。

如果存在,则两个节点间不转发任何消息断开连接,若不存在则继续进行后面的操作。

当发生黑名单事件时, w_2 让边权重 w 直接降为 0。发生信息发送失败等事件

的发生次数则是与这两个节点间的链路稳定性成反比,这类事件发生的越多说明这两个节点间的链路稳定性越差,同时也说明虽然这两个节点间的行为模式比较接近,但是实际情况中这两个节点的关系并不十分紧密。因此当发生这类事件时,负权重则按照事件发生的次数以一定比例衰减边权重 w 。这里,负权重是一个衰减的变量。当 w 衰减至 0 时,说明两个节点间的边是极度不稳定的或者两节点间的关系是极差的,应该将这条边从网络拓扑记录表中剔除掉。

根据上面的公式计算节点间的边权重,将前面采集到的有关连接信息的数据充分提取利用,以进一步体现节点间关系的紧密程度,并在此基础上划分出更加合理的社团结构。

4.3 路由转发

在消息转发阶段,节点在与其他节点相遇时,首先检查消息目的地是否存在于相遇的网络拓扑信息记录

表中. 因为网络拓扑信息记录表的筛选策略, 留在表中的大部分是与当前节点经常相遇或频繁相遇的节点, 所以先检查表中记录的节点有助于更快找到合适的转发节点. 若消息目的地存在于相遇节点的表中, 则选择权重大的节点转发消息.

当目的地节点不存在于任何一个相遇节点的网络拓扑信息记录表中时, 路由转发根据社团内/外转发以及路由调度策略转发消息. 为了便于选择转发节点, 节点先统计自身与各个社团间的总权重并建立一张社团权重查询表, 如图 2 所示. 节点与各个社团间的总权重为网络拓扑信息记录表中属于同一社团的节点成员与当前节点间的边权重的总和. 根据消息目的地的所属社团判断该消息目前是处于社团内部转发还是社团外部转发, 采用相应的转发策略. 下面分别详细介绍这两部分以及相应的路由调度策略. 社团查询的时间复杂度是 $O(m+n)$, 其中 m 是拓扑信息数目, n 是社团表大小.

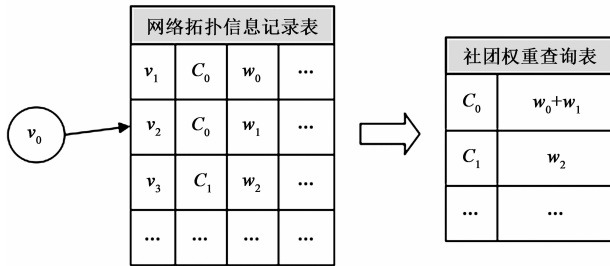


图2 社团权重查询表

算法 4 WQCA 路由算法

```

Begin
  For < ConnectionNode > con do
    For < Message > msg do
      If msg.destination exists in The Connection Information List then
        add msg to ForwardList
      Else If msg exists in destination community then
        If con.CommWeight > CurCommWeight then
          add msg to ForwardList
        End If
      Else
        If con.Community = msg.DesComm then
          add msg to ForwardList
        Else If con.CommNum > curCommNum then
          add msg to ForwardList
        End If
      End If
    End For
  End For
  sort(ForwardList)
  forwardMessage(ForwardList)
End For
End

```

若消息已经处于目的地社团, 则采取社团内转发的方式转发消息. 消息处于社团内转发时, 消息不再转发向非目的地社团成员的节点, 只在目的地社团内部传播. 比较相遇节点与目的地社团的权重, 选择权重较大的节点转发. 处于社团中心的节点同其他社团成员的接触比较频繁, 相应的它与其所属社团的权重也会更大一些. 消息逐渐向与目的地社团权重大的节点转移, 意味着它将逐渐接近社团中心, 进而遇见更多的目的地社团成员, 有更大的概率遇到目的地节点.

若消息未传递到目的地社团, 则采取社团外转发的方式. 根据社团权重查询表, 找到节点与目的地社团的总权重, 选择与目的地社团权重最大的节点转发. 若没有与目的地社团连接的相遇节点, 则选择连接邻接社团数量最多的节点转发. 只有离开当前这个非目的地社团才能够有更多的机会遇见目的地社团, 因此期望还未传递到目的地社团的消息能够向社团边缘传递. 边缘节点能够有更多的机会接触其他社团的成员将消息扩散出去.

路由转发阶段的具体算法实现如算法 4 所示. 路由转发的时间复杂度是 $O(m * n)$, m 是连接节点数目, n 是消息数目.

5 仿真实验

5.1 社团结构

在 ONE^[20] 平台上使用 Pittsburgh Working Day Model^[21] 分别运行基于模块度的社团划分算法、QCA 以及 WQCA 算法三种算法. 通过三者的社团划分结果比较算法的性能. 仿真时长为 2 周, 每天进行一次社团划分, 并统计当前网络中的社团总数和全局模块度.

图 3、图 4 分别为基于最新最频繁连接和最新的最近连接信息采集两种方式在每次进行社团划分时划分出的社团数量. 从中可以看出, 基于模块度的社团划分算法趋向于将整个网络中的节点划分到少数几个社团内, 整个网络中的社团数量是随着仿真时间的增长变得越来越少. 在仿真中, 整个网络中节点的移动模式没有太过剧烈的改变. 随着仿真时间的增长, 可以看出 WQCA 算法划分出的社团结构比 QCA 算法更快的趋于稳定, 而且 WQCA 算法划分出的社团结构相较另两种算法相比波动相对较小.

并且, 从图 3 和图 4 中可以看出 MFC 方式划分出的社团结构要比 MRC 方式划分出的社团结构更稳定一些. 虽然对于 QCA 和 WQCA 这两种根据网络拓扑变化更新社团的算法来说, 连接产生的时间离社团更新的时间越近, 它对节点所属社团产生的影响就越大, 并且整个网络中的节点移动轨迹变动的越剧烈越随机, 这种影响越明显, 但是在本仿真环境中, 节点的移动轨迹

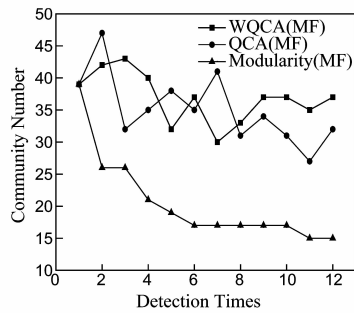


图3 基于MFC方式划分的社团数量

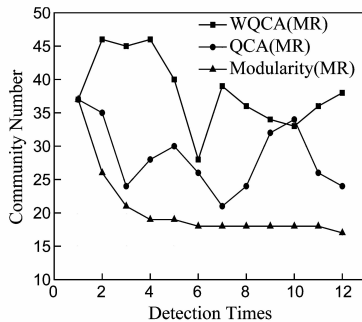


图4 基于MRC方式划分的社团数量

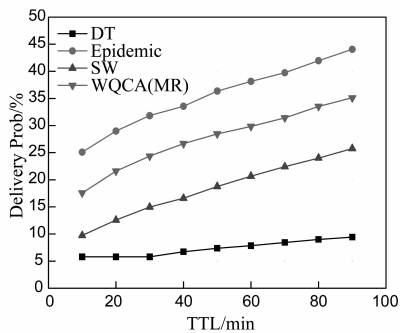


图5 消息传输成功率(Inforcom06)

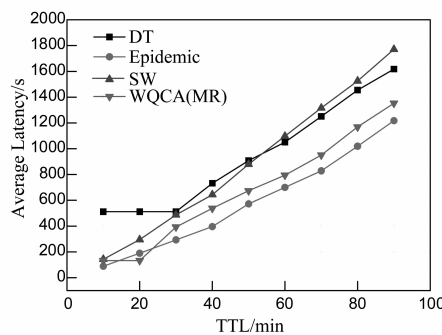


图6 平均时延(Inforcom06)

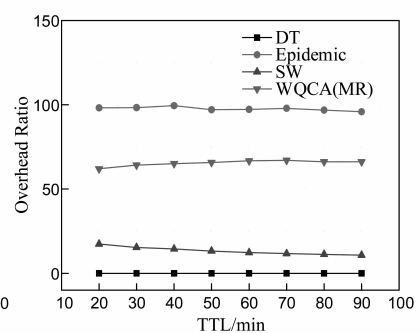


图7 网络开销比率(Inforcom06)

图5、6和7分别为DT、Epidemic、SW和WQCA这四种路由算法随TTL时间变化的消息传输成功率、平均时延和网络开销率变化曲线。可以看出,随着TTL的增加这四种路由算法的传输成功率和平均时延都保持着持续增长。这是因为原本一些在TTL内传递不到目的地的消息副本,由于TTL延长而找到了目的地节点传输成功。WQCA路由的传输成功率和平均时延仅次于Epidemic路由。WQCA的传输成功率只比Epidemic路由低了7%至8%,平均时延比Epidemic高了100s左右,而网络开销则比Epidemic低了将近35%。再看WQCA与DT的比较,虽然在网络开销方面WQCA比DT高很多,这是因为在DT当中始终只保持一个副本,而在传输成功率方面WQCA比DT高12%至25%,在平均时延方面,WQCA比DT低180s至390s。可以看出WQCA在网络开销增加不多的情况下在传输成功率和

具有一定的规律性,节点间的接触也比较频繁,因此使用MFC的信息采集方式不仅不会剔除掉很多重要的接触信息,反而会过滤掉大部分干扰社团划分的无意义接触信息,所以在本仿真环境中使用MFC信息采集方式更加合理。

5.2 路由性能

为了得到WQCA算法在真实环境下的性能表现,在Inforcom06数据集上对Direct Transmission (DT)^[22]、Epidemic^[23]、SW^[24]和WQCA这四种路由算法进行了比较。机会网络路由算法中,最基本的是DT、Epidemic、SW这三种路由算法。Epidemic是理想中成功率最高的路由算法,这也是冗余度最高的路由算法。DT路由算法冗余最低,成功率也是最低的。SW也是一个经典的算法,它是DT、Epidemic算法的折中,冗余和成功率也在这两者之间。如果拿这些算法比,不具有代表意义。由于Inforcom06数据集中的节点数量相对较少,节点间接触比较稀疏,而且真实环境中人的移动轨迹变化更具有随机性,所以这里WQCA算法采用MRC方式收集接触信息,以避免将对社团划分影响比较大的接触信息筛选掉。

的优势很明显。最后比较WQCA和SW算法。由于SW算法是Epidemic和DT算法的折中,所以SW在传输成功率方面要优于DT算法,而在网络开销方面要优于Epidemic算法。在传输成功率方面WQCA比SW算法高7%至10%,在平均时延方面,WQCA比SW少30s至450s,在网络开销方面WQCA要比SW高,这是因为,SW的网络副本数量是固定的,而WQCA是可变的,但是从整个机会网络的环境来看,增加的网络副本数是可接受的。

为了比较WQCA算法与其他采用社团划分的路由算法,本文选取了经典的Bubble_rap^[25]算法的改进算法dLife^[26]与本算法进行比较。移动模型选取了工作日模型(WDM)^[27],实验场景地图是赫尔辛基地图,这是因为场景中的节点数量相对较少,节点间接触比较稀疏,而节点之间的社团结构明显,适合采用基于社团结

构的路由算法. 图 8、9、10 为 WQCA 和 dLife 这两种路由算法随 TTL 时间变化的消息传输成功率、平均时延和网络开销率变化曲线. 可以看出, WQCA 的传输成功率明显高于 dLife, 平均时延也比 dLife 低. 这是因为两方面的原因: 首先, dLife 采用的是 K-clique 社团划分方法, K-clique 社团划分方法的对于社团划分不如 WQCA 社团划分方法稳定; 其次 dLife 在转发时, 如果目的节点不在源节点社团内, 就选择权重高的节点作为转发

节点, 而 WQCA 算法选取的是社团的边缘节点, 这使得下一跳节点选取更为合理. 在网络开销方面, WQCA 在 TTL 为 600 分钟以内时与 dLife 算法差别不大, 在 600 分钟以上时明显高于 dLife 算法, 这是因为在本算法中如果在很长时间无法寻找到目的社团将通过增加副本数来提高传输成功率. 如果网络环境对网络开销要求较高, 可以根据情况把 TTL 设置为 600 分钟以内.

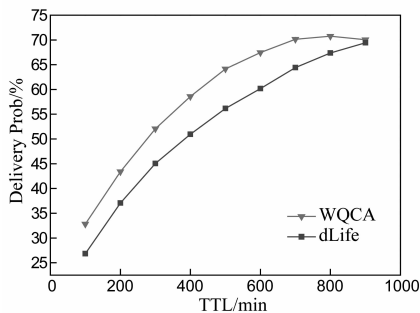


图8 消息传输成功率(WDM)

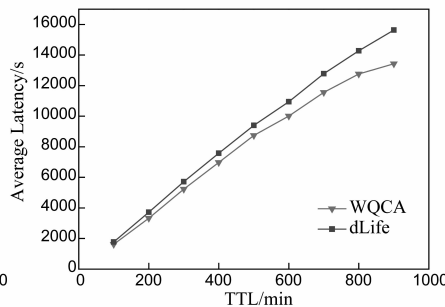


图9 平均时延(WDM)

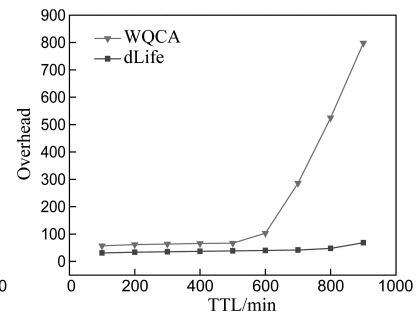


图10 网络开销(WDM)

6 结论

本文提出了一种基于有权社团结构的路由算法, 该路由算法可以选择两种不同的方式 (MRC 或 MFC) 采集信息, 根据采集的节点间接触信息进行权重转化, 在此基础上划分社团并判断转发节点. 同时, 本文根据网络环境的变化提出了两种适应不同环境的权重转化方案, 让路由算法在检测到周围网络环境变化时能够自主切换权重计算方式. 通过与其他路由算法的比较, 本文提出的路由算法能够在保证较高的传输成功率的情况下保持较低的平均时延, 并有效降低网络开销. 该路由算法性能稳定, 能够自主适应较大的网络拓扑变化. 目前该路由算法还有很大的提升空间, 未来工作将针对进一步提升路由性能以及重叠社团等方面作进一步研究.

参考文献

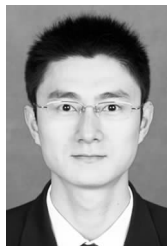
- [1] Y P Xiong, L M Sun, J W Niu, Y Liu. Opportunistic networks[J]. Journal of Software, 2009, 20(1): 124 - 137.
- [2] Newman M E J, Girvan M. Finding and evaluating community structure in networks[J]. Physical Review E, 2004, 69(2): 026113.
- [3] Traag V A, Bruggeman J. Community detection in networks with positive and negative links[J]. Physical Review E, 2009, 80(3): 036115.
- [4] E M Daly, M Haahr. Social network analysis for information flow in disconnected delay-tolerant MANETs [J]. IEEE Transactions on Mobile Computing, 2009, 8(5): 606

- 621.

- [5] Onnela J P, Saramäki J, Hyvönen J, et al. Structure and tie strengths in mobile communication networks[J]. Proceedings of the National Academy of Sciences, 2007, 104(18): 7332 - 7336.
- [6] Xiang R, Neville J, Rogati M. Modeling relationship strength in online social networks[A]. Proceedings of the 19th International Conference on World Wide Web[C]. New York: ACM, 2010. 981 - 990.
- [7] Bigwood Greg, et al. Exploiting self-reported social networks for routing in ubiquitous computing environments [A]. IEEE International Conference on Wireless and Mobile Computing [C]. Avignon: IEEE, 2008. 484 - 489.
- [8] Ciobanu R I, et al. Social aspects for opportunistic communication[A]. 11th International Symposium on Parallel and Distributed Computing (ISPDC) [C]. Bavaria: IEEE, 2012. 251 - 258.
- [9] F Cai, et al. K-clique community detection based on union-find[A]. International Conference on Computer, Information and Telecommunication Systems (CITS) [C]. Jeju: IEEE, 2014. 1 - 5.
- [10] Newman M E J. Assortative mixing in networks [J]. Physical Review Letters, 2002, 89(20): 208701.
- [11] Newman M E J. Analysis of weighted networks[J]. Physical Review E, 2004, 70(5): 056131.
- [12] Nguyen N P, et al. Adaptive algorithms for detecting community structure in dynamic social networks [A]. Proceedings of IEEE INFOCOM [C]. Shanghai: IEEE, 2011. 2282 - 2290.
- [13] Daly E M, Haahr M. Social network analysis for routing in

- disconnected delay-tolerant MANETs [A]. Proceedings of the 8th ACM International Symposium on Mobile Ad Hoc Networking and Computing [C]. New York: ACM, 2007. 32 – 40.
- [14] Hui P, Crowcroft J, Yoneki E. Bubble rap: Social-based forwarding in delay-tolerant networks [J]. IEEE Transactions on Mobile Computing, 2011, 10(11): 1576 – 1589.
- [15] E Bulut, B K Szymanski. Exploiting friendship relations for efficient routing in mobile social networks [J]. IEEE Transactions on Parallel and Distributed Systems, 2012, 23(12): 2254 – 2265.
- [16] A Mei, et al. Social-aware stateless forwarding in pocket switched networks [A]. Proceedings IEEE INFOCOM [C]. Shanghai: IEEE, 2011. 251 – 255.
- [17] Xuebin Ma, et al. A community-based routing algorithm for opportunistic networks [A]. 2013 Fifth International Conference on Ubiquitous and Future Networks (ICUFN) [C]. Da Nang: IEEE, 2013. 701 – 706.
- [18] Hossmann T, Spyropoulos T, Legendre F. Social network analysis of human mobility and implications for DTN performance analysis and mobility modeling [R]. Switzerland: Computer Engineering and Networks Laboratory ETH Zurich, 2010. 323.
- [19] Ma H, Du Q. Research on opportunistic network routing based on community structure [J]. Electronic Science and Technology, 2013, 2013(5): 037.
- [20] A Keränen, J Ott, T Kärkkäinen. The ONE simulator for DTN protocol evaluation [A]. The 2nd International Conference on Simulation Tools and Techniques [C]. Brussels: ICST, 2009. 55.
- [21] Ekman Frans, Ari Keränen, Jouni Karvo, Jörg Ott. Working day movement model [A]. Proceedings of the 1st ACM SIGMOBILE Workshop on Mobility Models [C]. New York: ACM, 2008. 33 – 40.
- [22] Spyropoulos, et al. Single-copy routing in intermittently connected mobile networks [A]. Proceedings of Sensor and Ad Hoc Communications and Networks SECON [C]. Los Angeles: IEEE, 2004. 235 – 244.
- [23] Vahdat A, Becker D. Epidemic routing for partially connected ad hoc networks [R]. Technical Report CS-200006, Duke University, 2000.
- [24] Spyropoulos T, Psounis K, Raghavendra C S. Spray and wait: an efficient routing scheme for intermittently connected mobile networks [A]. Proceedings of the 2005 ACM SIGCOMM Workshop on Delay-tolerant Networking [C]. New York: ACM, 2005. 252 – 259.
- [25] Hui P, Crowcroft J, Yoneki E. Bubble rap: Social-based forwarding in delay – tolerant networks [J]. IEEE Transactions on Mobile Computing, 2011, 10(11): 1576 – 1589.
- [26] Moreira W, Mendes P, Sargento S. Opportunistic routing based on daily routines [A]. 2012 IEEE International Symposium on World of Wireless, Mobile and Multimedia Networks (WoWMoM) [C]. San Francisco: IEEE, 2012. 1 – 6.
- [27] Ekman F, Keränen A, Karvo J, et al. Working day movement model [A]. Proceedings of the 1st ACM SIGMOBILE Workshop on Mobility Models [C]. New York: ACM, 2008. 33 – 40.

作者简介



马学彬 男, 1981 年生, 内蒙古大学副教授, 从事计算网络技术方面的有关研究.

E-mail: csmaxuebin@imu.edu.cn



白婧 女, 1989 年 12 月出生, 内蒙古大学硕士研究生. 研究方向为计算机网络.

郑田玉 男, 1993 年 12 月出生, 内蒙古大学硕士研究生. 研究方向为计算机网络.