移动社会网络中基于多维上下文匹配的数据转发算法

徐 方[1,2] 邓 敏[1] 熊曾刚[1] 叶从欢[1] 徐 宁[2]

(湖北工程学院计算机与信息科学学院 湖北 孝感 432000)[1] (武汉大学计算机学院 武汉 430072)[2]

摘 要 通过研究移动社会网络中的多种上下文信息对节点移动模式的影响,提出了基于多维上下文认知的数据转发算法MCMF。该算法综合考虑物理邻接性、社会相似性以及社会交互性3个维度的上下文信息来进行动态数据转发决策。首先消息携带者节点通过物理邻接匹配获得邻居节点集合;然后通过社会相似性匹配在邻居节点集合中选出候选节点子集,并基于社会网络的社群特征,采用马尔可夫预测方法在候选节点子集中选出最优中继节点;最后设计高效的数据转发算法。仿真实验表明,相比于其他3种著名算法,该算法在交付比率和开销比率方面具有较好的性能。

关键词 移动社会网络,多维,上下文匹配,预测模型,转发算法

中图法分类号 TN915 **文献标识码** A **DOI** 10.11896/j.issn.1002-137X.2019.02.013

Data Forwarding Algorithm Based on Multidimensional Context Matching in Mobile Social Networks

XU Fang[1,2] DENG Min[1] XIONG Zeng-gang[1] YE Cong-huan[1] XU Ning[2]

(School of Computer and Information Science, Hubei Engineering University, Xiaogan, Hubei 432000, China)[1]

(School of Computer Science, Wuhan University, Wuhan 430072, China)[2]

Abstract Through studying the effect of a variety of context information on the mobility patterns in mobile social networks, this paper proposed a multidimensional context matching forwarding (MCMF) algorithm. In this novel algorithm, three dimension contexts, which are physical adjacency, social similarity and social interactivity, are used to make routing decisions dynamically. Firstly, message carrier obtains its neighbor node sets by using physical adjacency matching at the present moment. Then social similarity matching is used to search relay candidate subset of neighbor node sets, and the discrete-time semi-Markov prediction model is used to determine the best relay node. At last, the efficient data forwarding algorithm is designed. Simulation experiments based on real traces show that the proposed MCMF algorithm is more efficient in terms of maximizing the delivery ratio and minimizing the overhead ratio than other three state-of-the-art algorithms.

Keywords Mobile social networks, Multi dimensional, Context matching, Prediction model, Forwarding algorithm

1 引言

在过去的十年中,智能移动设备的数量迅速增长,这为移动社会网络的发展提供了物质基础。智能移动设备具有友好的用户接口,用户可以非常方便地在智能移动设备上创建各类应用,大量的应用需要利用网络与其他用户进行通信,移动社会网络的出现为这些应用提供了一种方便、廉价的通信服务[1]。由于移动设备通常由人类、车辆或动物所携带,因此研究者开始考虑携带者的特征和移动规律对网络通信的影响。移动社会网络的间歇性连接导致节点之间缺乏稳定的端到端通路,因而使得设计高效的路由和转发策略成为当前研究的热点问题。从另一个角度看,网络节点随着人类(车辆)等携带者一起移动,人类的活动和社会关系显著影响着节点运行的模式[2]。在研究移动社会网络中的数据转发时,充分考虑这些上下文信息能获得较好的效果。

上下文认知是实现网络智能化的关键技术,在网络数据转发中引入上下文认知,使得数据转发算法和路由协议能根据网络自身和外部的状态智能地选择路径,从而最大限度地提高网络的服务质量和节约网络资源。PRoPHET[3]是一种概率路由算法,它利用收集到的节点的历史接触信息计算一个节点到目的节点的投递预测值。投递预测值就是一个节点未来遇到目的节点的概率。当一个节点频繁遇到目的节点时,概率就会增加,反之则可以利用老化函数来减小概率。CAR路由算法[4]利用卡尔曼滤波模型评估多种上下文信息,

到稿日期:2018-01-07 返修日期:2018-03-30 本文受教育部人文社会科学研究青年基金项目(17YJCZH203),湖北省教育厅科研计划重点项目(D20182702)资助。

徐 方(1981-),男,博士,副教授,硕士生导师,CCF会员,主要研究方向为无线移动网络、智能感知计算;邓 敏(1983-),女,硕士,讲师,主要研究方向为计算机网络,E-mail:dm2010@whu.edu.cn(通信作者);熊曾刚(1974-),男,博士,教授,硕士生导师,主要研究方向为计算机网络、云计算;叶从欢(1980-),男,博士,副教授,硕士生导师,主要研究方向为社交多媒体;徐 宁(1989-),男,博士生,主要研究方向为无线移动网络、智能感知计算。

通过对未来上下文信息的预测来进行路由决策。

研究者提出了一些基于社群的路由协议。具有社会联系、共同兴趣和相似性的人们通常会组成一个群体，相比于群体成员间的社会交往，群体内部成员的社会交往会更加频繁。较知名的基于社群的路由协议有 LABEL[5]，SGBR[6]，CAOR[7]，LocalCom[8] 和 Friendship Based Routing[9] 等。基于社会上下文认知的算法不仅需要利用节点移动性相关的上下文信息，而且还要考虑将节点社会方面的属性作为一个重要参数来参与路由决策。实际上，在大多数情况下，节点的移动特征由该节点的携带者的行为所决定，这些携带者可能是人、动物或者车辆。3 种著名的基于社会上下文的路由算法是 HiBOp[10]，Rubble Rap[11] 和 dLife[12]。

在设计移动社会网络数据转发策略时，充分考虑网络节点携带者的上下文信息可以提高数据转发效率，而综合利用多维度上下文信息对节点的影响可以进一步提升网络的性能。本文提出了基于多维度上下文匹配的数据转发(Multidimensional Context Matching Forwarding，MCMF)算法。该算法首先进行物理邻接匹配，找到与转发节点有通信连接的邻居节点；然后在邻居节点中通过社会相似性匹配找到与目的节点具有相似社会特征的中继候选者集合；最后应用社会交互性匹配找到最佳的信息中继节点。该算法利用了 3 个维度的上下文信息，采用马尔可夫预测模型对上下文信息进行推理和预测，最终综合上述信息选择最适合的中继节点，从而提高数据转发的效率。

2 系统模型

2.1 问题描述

本文以校园中的移动社会网络为例来说明网络中的数据转发问题。网络中的节点主要在若干地理社区停留或社区之间移动，此处假设处于同一地理社区中的节点之间可以利用自身所携带的短距离无线接口直接通信，即任何两个同时处于同一地理社区的节点可以建立连接并交换数据。处于不同地理社区的节点不能直接通信，地理社区由唯一的标识(ID)进行区分。网络中的节点由校园中的老师和学生所携带的智能移动设备组成，在具有社会属性的网络中，节点的移动具有一定的规律。图 1 是一个基于校园环境的移动社会网络，网络中的节点按一定的日常规律在不同的地理社区停留和移动，这些地理社区为教室(Classroom)、实验室(Lab)、食堂(Canteen)、图书馆(Library)和体育场(Stadium)。

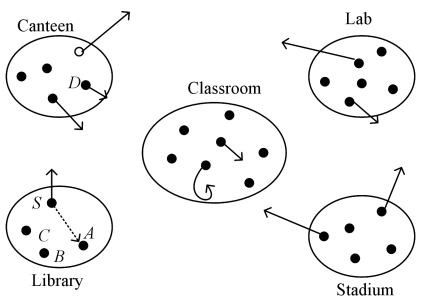


图 1 基于校园环境的移动社会网络

Fig. 1 Mobile social networks based on campus

图 1 描绘了基于校园环境的移动社会网络的数据转发过程，节点 S 需要将数据转发给目的节点 D，节点 A，B 和 C 与节点 S 处于同一地理社区，基于社会上下文信息和历史移动信息，节点 S 预测图书馆社区(Library)中的节点 A 为数据转发的最优中继节点，于是节点 S 将数据转发给节点 A，经过一段时间，节点 A 离开图书馆社区(Library)，在教室社区(Classroom)内与目的节点 D 相遇，并将数据转发给节点 D。此处仅仅是以人们熟悉的大学校园为场景来描述相关问题，该算法的实际应用场景为不存在蜂窝移动通信基础设施的情况、因各种灾害导致基站损毁时的应急通信、不愿意使用公用网络传输数据的场景以及蜂窝移动通信费用较高的场景。另一方面，该算法主要用于中小规模网络节点的自组织通信。

在数据转发过程中，如何利用上下文信息来预测节点未来的移动情况，以及如何利用这些预测信息进行高效的数据转发，是当前研究的热点问题。本文通过多维上下文匹配的方法，在候选节点中搜索最优中继节点，并设计合理的数据转发策略为网络提供数据转发服务。利用这些方法可提高数据成功投递的比率，同时减少投递时延。

2.2 社会相似性建模

在以前提出的大多数路由协议的设计中，数据转发决策只是简单地考虑了节点的社会属性和邻近节点[13-14]。研究表明[15]，具有相同兴趣倾向和相似社会特征的人们容易相聚在一起。基于这一规律，本文为每个节点设计了一个社会属性表，用于比较节点之间的社会相似性。通过社会相似性比较，我们可以找到更适合的中继节点，利用这些中继节点可以把消息传递到目的节点。

(1)社会属性表

为了计算某个节点与目的节点的社会相似性，本文首先定义了网络节点共同的社会属性集合，该集合包含一组社会属性以及属性的权重，权重值表示这个属性在社会网络中的重要性，每个节点建立一个社会属性表(Social Profiles)。社会属性表由一组成对的属性 e_i、值 v_i，以及哈希值 $H(e_i,v_i)$ 组成。

本文使用社会属性表描述节点携带者的社会信息，这些信息主要包含设备携带者的信息(姓名、住址、工作单位、爱好等)，由社会属性与对应的取值组成。考虑到便于比较和安全方面的因素，采用哈希函数对属性和值进行处理，并将函数生成哈希值存入对应的社会属性表中。

社会属性表的例子如表 1 所列，每个节点的社会属性表具有相同的属性集合，并且属性排列的顺序是一致的。每个社会属性会被分配一个权重用于体现其重要程度，属性权重的分配与实际应用相关。

表 1 社会属性表的例子

Table 1 Example of social profiles

Evidence name	Value	Hashed values
Name	Jack	$H(Name,Jack)$
Profession	researcher	$H(Profession,researcher)$
Workplace	WHU	$H(Workplace,WHU)$
Residence	Wuhan	$H(Residence\ place,Wuhan)$
Interest	Football	$H(Interest,Football)$

(2)社会相似性

当两个节点在同一社群相遇时,它们可以互相交换信息,并且假设网络中的节点都是无私的,愿意为其他节点转发数据。通过匹配节点的属性表可以计算两个节点的社会相似性。节点 N 的社会属性表中的属性(evidence)和对应值(value)的集合记为 $N(e,v)$;目的节点 D 的属性表中成对的属性和值集合记为 $D(e,v)$;节点 N 和目的节点 D 的属性值的交集记为 $M(e,v)$,则有:

$$M(e,v)=N(e,v)\bigcap D(e,v) \tag{1}$$

定义 1(社会相似性度量)　移动社会网络中一个节点 N 与目的节点 D 的社会属性的相似程度。

网络中每个节点内存在一个含有属性和对应取值的社会属性表,节点 N 和目的节点 D 之间的社会相似性计算可以通过匹配两个节点的社会属性表来实现。利用表 1 中的哈希值可以很方便地计算社会相似性,哈希函数的输出是固定的长度,并且按照属性的顺序对哈希值进行对比。属性的权重值表明该属性在信息成功投递中的重要性,每个属性均设置有一个权重。设 W 为 $M(e,v)$ 中属性的权值 W_m 的集合,W_D 为 $D(e,v)$ 中属性权值 W_d 的集合,则节点 N 和目的节点 D 之间的社会相似性 $Sim_{[N,D]}$ 的计算如式(2)所示。

$$Sim_{[N,D]}=\frac{\sum\limits_{m\in W}W_m}{\sum\limits_{d\in W_D}W_d} \tag{2}$$

2.3　社会网络建模

社会网络可以用一个无向图 $G(V,E,W_{\text{society}})$ 表示。由图 G 可知,社会网络中存在 N 个节点、M 条边。图中的每个节点表示一个用户携带的移动设备,并且所有设备的无线信息的传输半径相同。图中的每条边代表两个节点之间存在社会关系,边的权值 W_{society} 代表两个节点社会关系的亲密程度。每个节点可能有多条边,表示每个节点可能与其他多个节点存在社会关系。但是一条边只能表示两个节点之间的关系。

定义 2(基于地理社区的社会网络)　移动节点以地理社区的形式组成连通域,某一时刻一个节点只属于其中的某个连通域,具有此特征的社会网络为基于地理社区的社会网络。

从 1 到 Y 对社区进行编号,则基于地理社区的社会网络可表示为 $C(t)$,如式(3)所示:

$$\begin{aligned}&C(t)=\{C(t)_1,C(t)_2,\cdots,C(t)_Y\}\\&C(t)_k\in C(t),k\in[1,2,\cdots,Y]\end{aligned} \tag{3}$$

其中,t 为时间变量。

社会网络图 G 可以用 $C(t)$ 表示,它拥有相应的图特征,包含若干个图 G 的子集。$C(t)_k$ 是图 G 中的一个子集,也是其中的一个地理社区。图 G 中所有的节点都属于 $C(t)$ 中一个确定的地理社区。

此外,$C(t)$ 依赖时间槽 t 捕获节点移动的动态特征,如式(4)所示:

$$G=\bigcup_{i=1,\cdots,k}C(t)_i \tag{4}$$

在这个图模型中,每个节点的连通性随着移动性和位置的变化而变化。因此,社区内节点的状态随时间而变化,这样就形成了网络图。这个网络图明确地表明了在 t 时刻一个网络节点可以与其他哪些节点直接通信。随着时间从 t 迁移到 $t+1$,节点从一个社区移动到另一个社区,因此网络图也在发生变化。

2.4　预测模型

根据社会学的研究成果,社会网络表示节点间社会交往(传递性)的程度。如果两个人有一些共同的特征,那么在将来的某个时间段,他们在一起交往和相聚的概率非常大。本文采用半马尔可夫过程建立社会移动模式,节点从一个地理社区到另一个地理社区的移动性具有马尔可夫的记忆无关属性,那么节点从一个地理社区到另一个地理社区移动的概率不依赖于以前的概率。因此,我们可以利用马可夫过程建立节点移动的预测模型。

节点 m 的状态集合表示为 $X^m=\{1,2,3,\cdots,Y\}$,用于表示节点 m 所处的地理社区编号,Y 表示图中总地理社区数量。节点 m 的第 k 个状态表示为 X_k^m,节点 m 进入第 k 个状态的时刻表示为 T_k^m,那么对于节点 m 来说,形成了一个时间和状态离散的马尔可夫链 X_k^m,则有节点 m 的第 $k+1$ 个状态 X_{k+1}^m 只与其第 k 个状态有关,而与之前的状态无关。此时对于节点 m,形成了状态离散的半马尔可夫模型 (X_k^m,T_k^m),该模型描述了节点 m 随着时间的变化而转移状态的过程。节点 m 在第 k 个状态的停留时间表示为 R_k^m,则 R_k^m 可由式(5)表示:

$$R_k^m=T_{k+1}^m-T_k^m \tag{5}$$

半马尔可夫模型 (X_k^m,T_k^m) 的核心表达式如式(6)所示,此处 T_k^m 为一般分布,重点分析状态 X_k^m 的随机过程。模型可以表示节点 m 在 t 时间内从状态 i 转移到状态 j 的概率,X_{k+1}^m 的状态取决于当前状态 X_k^m,而与之前的 X_{k-1}^m 等状态无关。

$$\begin{aligned}Q_{ij}^m&=P(X_{k+1}^m=j\mid X_k^m,X_{k-1}^m,\cdots,X_0^n;t_k,t_{k-1},\cdots,t_0)\\&=P(X_{k+1}^m=j\mid X_k^m=i,t_k),i,j,k\in\{1,2,\cdots,Y\}\end{aligned} \tag{6}$$

X_k^n 表示节点 m 在时刻 t 停留在社区 k 的状态。Q_{ij}^m 是节点 m 从社区 i 到社区 j 的转移概率。假设每个节点只在离散时隙内移动,并且相遇节点在每个时隙内更新信息。最终,每个节点保存一个转移概率矩阵。下式中,Y 表示网络中社区的数量。

$$A_m=\{Q_{ij}^m;i,j=1,2,\cdots,Y\} \tag{7}$$

$$A_m=\begin{pmatrix}Q_{11}^m & Q_{12}^m & \cdots & Q_{1Y}^m\\ Q_{21}^m & \cdots & \cdots & Q_{1Y}^m\\ \cdots & \cdots & \cdots & \cdots\\ Q_{Y1}^m & \cdots & \cdots & Q_{YY}^m\end{pmatrix} \tag{8}$$

式(7)和式(8)中,A_m 是离散马尔可夫链的转移概率矩阵,它包含了每个节点的所有过渡状态。在一个时间段内,如果跟踪记录节点 m 所访问的社区,那么节点 m 从社区 i 到社区 j 的转移概率 Q_{ij}^m 的定义如式(9)所示:

$$Q_{ij}^m=\frac{Num_{ij}}{Num_i} \tag{9}$$

其中,Num_{ij} 表示某段时间内从社区 i 移动到社区 j 的节点的数量,Num_i 为在该时间段内从社区 i 转移出的所有节点的数量。利用上述马尔可夫模型和相关的转移概率矩阵,可以对移动社会网络节点的移动规律进行形式化描述。在上述模型中,网络中的每个地理社区有一个唯一的编号,在节点移动的

每个时刻，节点都能感知自己所处的社区编号。同时，假设网络由紧紧相邻的社区组成，节点从一个社区到达另一个社区的时间可以忽略不计，那么节点在任一时刻总是处于某一社区之中。

因此，我们可以针对具有社区特性的移动社会网络建立相应的马尔可夫预测模型，然后利用该模型预测网络节点在社区间的移动情况和停留时间。

3 MCMF 数据转发策略

在间歇性连通的移动社会网络中，数据转发利用网络节点之间的相遇机会逐跳将数据传递到目的节点，那么与目的节点相遇概率较大的节点将容易被选作中继节点。当一个节点需要发送一个数据包时，它就会触发转发算法，在自己的邻近节点中寻找最佳的中继节点，数据包被一次次转发给最佳的中继节点，直到它达到目的节点为至。这是容迟网络单拷贝路由算法中较为常见的策略，本文提出的路由算法的特别之处在于，利用社会上下文信息来评估和预测相遇概率，并选择最优的中继节点。

基于以上分析，本文提出了一种适合于间歇性连通的移动社会网络数据转发算法——基于多维度上下文匹配的数据转发(MCMF)算法。MCMF 算法是一种分布式算法，该算法应用 3 个上下文信息匹配过程来求解最优中继节点。这 3 个匹配过程分别为物理邻接匹配、社会相似性匹配和社会交互性匹配，它们构成了整个分布式算法的核心部分。MCMF 算法的伪代码如算法 1 所示。

算法 1 MCMF Algorithm

```
(Suppose node m attempts to send a packet to node d)
Input: NP, C(t), A_m, d
1. FH=Max_hops
2. repeat
3.   R=Ø // R is relay candidates set
4.   Execute PAM (m, C(t))
5.     return L_m
6.   If d∈L_m then
7.     {relay_node=d
8.     Forward_Packet (relay_node)}
9.   Else {R=R∪L_m
10.          Execute SRM(L_m, NP)
11.          return relay_candidates_set=R
12.          Execute SAM(R, A_m)
13.          return relay_node=m'}
14.  Endif
15.  Forward_Packet (relay_node)
16.    {Send the packet to relay_node
17.     Delete the packet in m's buffer
18.    FH=FH-1}
19.  m=relay_node
20. until {FH=0 or relay_node=d}
```

算法 1 描述了 MCMF 算法的执行过程。算法的输入：NP，为网络中节点的社会属性；$C(t)$，表示在 t 时刻，网络节点分布于各社区的状态；A_m，表示马尔可夫转移矩阵；d，表示数据包的目的节点。首先，设置数据包在网络中传输的最大转发跳数 FH 为一个常数 Max_hops，Max_hops 的取值通过仿真实验取得，在第 4 节的性能评估部分将有详细的描述。算法 1 的第 2－20 行是一个循环体，其目的是找到从源节点到目的节点的一系列适当的中继节点并转发数据。首先，定义一个中继节点集合 R，并将 R 赋值为空集。然后执行物理邻接匹配(PAM)子程序，该子程序返回转发节点 m 的邻接节点集合 L_m。如果 L_m 中包含数据包的目的节点 d，则转发节点 m 把相关数据发送给目的节点 d，算法结束运行；如果 L_m 中不包含目的节点 d，则将邻接节点集合 L_m 并入中继节点集合，然后执行社会相似性匹配(SRM)子程序。SRM 子程序利用节点社会属性的相似性做出选择，从中继节点集合中选出与目的节点的社会属性相似性较高的节点，进一步减少中继节点集合中元素的数量。接下来，运行社会交互性匹配(SAM)子程序，利用马尔可夫转移矩阵的运算，从中继节点集合中找出到目的节点社区转移概率最高的节点 m'。此时节点 m 选择 m' 作为最优中继节点，并将数据转发给 m'。以上过程循环进行，以找到从源节点到目的节点路径上的一系列中继节点，并转发数据，直到跳数计数器为 0 或者数据已经到达目的节点。

(1)物理邻接匹配

物理邻接匹配(Physical Adjoin Matching，PAM)子程序的伪代码如算法 2 所示。在某一时刻，某些移动网络的节点处于彼此无线射频的通信范围内并能够互相交换数据，称这些节点是物理邻接的。节点 m 的物理邻接节点集合也可以称为节点 m 的一跳邻居集合。节点 m 通过与其周围的邻居节点发送 hello 报文来感知其物理邻接节点。

算法 2 Physical Adjoin Matching (PAM)

```
Input: m, C(t)
Output: L_m
1. n∈C(t), m∈C(t)
2. m broadcast Hello message
3. n reply Hello message replying package at time T
4. L_m={n|n is a neighbor of m at time T}
5. return L_m
```

(2)社会相似性匹配

社会相似性匹配(Social Similarity Matching，SSM)子程序的伪代码如算法 3 所示。该子程序的主要目的是在发送节点的邻接节点集 L_m 中查找出与目的节点 d 有相似社会属性特征的节点。

算法 3 Social Similarity Matching (SSM)

```
Input: L_m, NP
Output: R
1. for all node k∈L_m do
2.   search for k such that
3.   S={k|Sim_[k,d]≥λ; ∀k, k∈L_m}
4.   R=R∩S
5. end for
6. return R
```

节点之间的社会相似性匹配主要用于评估相关节点与目的节点在社会特征上的相似程度，这在 2.2 节中有详细的描

述。相关研究表明,若两个节点具有相似社会属性,则它们之间将会存在较高的相遇概率。节点之间的社会属性匹配程度越高,它们之间相遇的概率就越大。因此,在移动社会网络中,数据包的成功投递率与节点之间的社会相似性密切相关。在算法 3 中,程序的输入有:转发节点 m 的邻接集合 L_m 和网络节点的社会属性 NP。算法 3 第 1—5 行是一个循环体,其目的是把邻接集合 L_m 中的每个节点与目的节点进行社会相似性匹配,如果社会相似性取值大于或等于某一阈值 λ,那么就把集合 L_m 中对应的节点加入到中继节点集 R 中。当循环体结束时,就可以在邻接集合 L_m 中找出与目的节点有较高相似性的节点集合。本子程序输出的节点集合将用于下文的社会交互性匹配。

(3)社会交互性匹配

社会交互性匹配(Social Activity Matching,SAM)子程序的伪代码如算法 4 所示。本模块利用离散的半马尔可夫模型来预测在未来某个确定的时刻两个节点的接触情况。SAM 模块的输入为:中继节点集合 R 以及马尔可夫转移概率矩阵 A_m。A_m 中保存了网络中节点访问不同地理社区的历史信息,我们利用这些信息进行社会交互性匹配。在算法 4 中,第 1—4 行是一个循环体,该循环体的主要目的是计算出中继节点集合 R 中每个节点到目的节点的转移概率,然后在其中选择出转移概率最高的节点 m'。最后节点 m'作为子程序的返回值输出,即 m'是最终选定的数据包的中继节点。

算法 4　Social Activity Matching (SAM)

Input:R,A_m

Output:m'

1. for all node $m' \in R$ do

2.　search for m' such that

3.　$m' = \underset{m'}{\operatorname{argmax}}\{ \sum_{i=C(t)_k}^{j=Y} Q_{ij}^{m'} * \sum_{\substack{i'=1 \\ \forall i' \neq i, \forall j' = j}}^{i'=Y} Q_{i'j'}^{d} \}$

4. end for

5. return m'

4　仿真与分析

本文针对 MCMF 的性能进行了大量模拟实验。为了进一步证明本文所提算法的优点,我们把 MCMF 与其他相关的经典路由转发算法进行了比较分析,这些路由转发算法分别为基于洪泛的经典算法 Epidemic 算法[16]、没有利用社区的路由 ProPHET 算法以及基于社区的 BubbleRap 算法。本文的仿真实验采用机会网络和移动社会网络中广泛使用的仿真工具 ONE[17],实际数据集为 MIT Reality Data[18]。本文提及移动社会网络是一种分布式自组织网络,网络节点由较高性能的智能移动设备组成,用于承担 MCMF 算法的运算和存储。

4.1　实验设置

在仿真实验中,为了使得仿真尽量接近真实环境,我们采用 MIT Reality Data 实际数据集作为外部数据源来对算法进行实验分析。在计算社会相似性时,需要用到社会属性表,不同的数据集包含的社会属性不同。在 MIT Reality Data 数据集中,社会属性表的属性主要为 Name,Position,Neihborhood 和 Hangouts,根据属性的重要性,这些属性对应的权值分别为:0.1,0.2,0.4,0.3。社群信息来源于数据集中的Celltower_oid 字段。节点的物理邻接信息可以从数据集中节点的接触记录中获得,节点的社会属性和社交联系也可以很容易地从数据集中取得。

实验主要从交付比率、开销比率、平均时延和平均跳数 4 个方面来对算法进行评估和分析。首先对算法中的最大转发跳数(Max_hops)的设置进行分析,因为最大转发跳数设置得过大或过小都会对算法性能产生不利影响。然后,评估和分析 MCMF 中的 λ 参数的取值对算法性能的影响,基于分析结果选取最佳的 λ 参数取值,使得 MCMF 的性能最优。最后,将 MCMF 算法的性能与 Epidemic,ProPHET 和 BubbeRap 进行对比分析。

每个仿真实验运行 30 次,然后分别收集统计数据,并取平均值。有关 Epidemic 和 ProPHET 中的参数设置,仿真中采用 ONE 中相关协议的默认设置。仿真实验中的主要参数设置如表 2 所列。在仿真实验中,模拟实验运行 1h 相当于实际生活场景的 3h,因此表 2 中的仿真时间 40h 为考察了 5 天的实际场景的运行数据。整个网络的消息发送频率为每秒 500～600 个。MCMF 算法的最大转发跳数设置为 6。

表 2　参数设置

Table 2　Setting of parameters

参数	取值
仿真时间/h	40
预热时间/s	5000
运动模型	External Movement
节点数量	97,76
数据集	MIT Reality Data
无线传输类型	Bluetooth
传输速率/(kB/s)	250
传输范围/m	10
消息发送频率/s	500～600
消息大小/kB	256～512
节点缓存容量/MB	5
消息生存时间/min	600

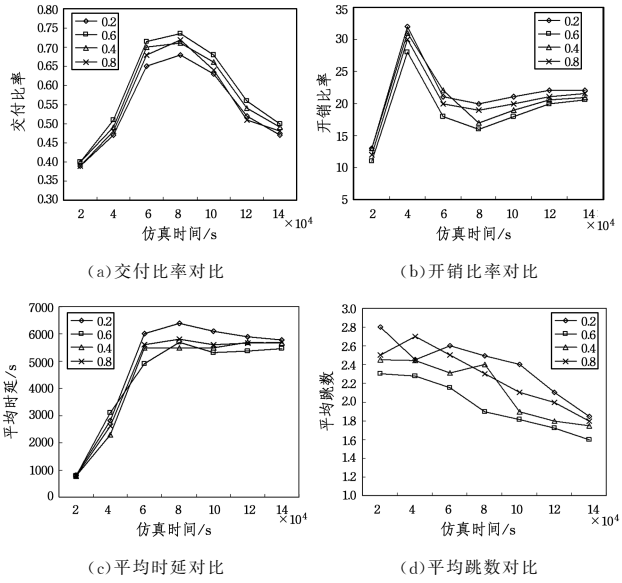
4.2　参数对 MCMF 的影响分析

在计算社会相似性匹配时,参数 λ 的取值会对 MCMF 算法的性能产生较大的影响。基于这一考虑,本文首先进行了一组仿真实验,用于比较和分析参数 λ 对算法的影响,并获得使算法性能最优的参数 λ 的取值。

MCMF 通过多维度的上下文信息匹配的方式为转发节点寻找最适合的中继节点,在进行社会相似性匹配时,利用参数 λ 来调节中继节点侯选集中节点的数量。当参数 λ 的取值较小时,算法把社会相似度较低的节点都纳入侯选节点,增加了侯选节点的数量和网络资源的消耗,不利于提高消息转发的效率。当参数 λ 的取值较大时,只有与目的节点社会相似度很高的邻居节点才能被选为侯选中继节点,中继节点侯选集中的节点数会较少,甚至容易出现空集现象。由于社会相似性匹配只是影响转发成功的因素之一,因此参数 λ 的取值较大时会影响到消息转发的交付比率。

接下来,本文采用 MIT Reality Data 实际数据集来评估参数 λ 对算法性能的影响。在评估中,分别设置参数 λ 的值

为 0.2,0.4,0.6,0.8。图 2 展示了 MIT Reality Data 数据集下 MCMF 算法的仿真结果,分别对比了在参数 λ 取不同值的情况下,算法的交付比率、开销比率、平均时延和平均跳数。从图中可以看出,当参数 λ 取值为 0.4 时,MCMF 算法获得了最高的性能。因此,在下文与其他路由和转发算法进行比较时,MCMF 算法中的参数 λ 的取值为 0.4 较为合适。



(a)交付比率对比 (b)开销比率对比

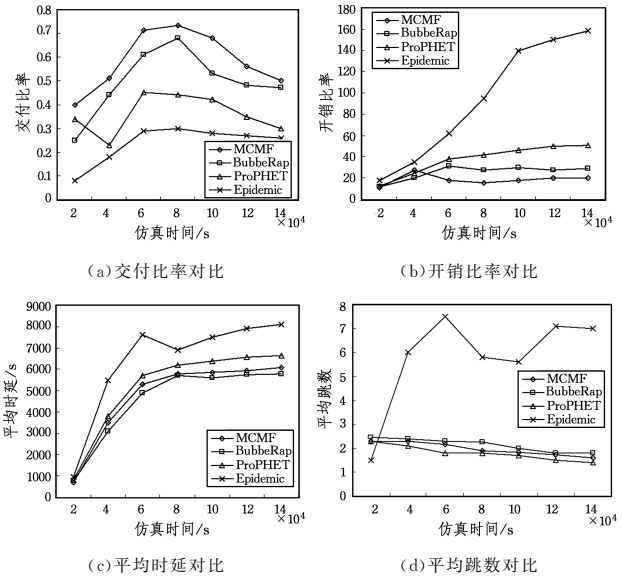
(c)平均时延对比 (d)平均跳数对比

图 2 参数 λ 对 MCMF 的影响

Fig. 2 Infulence of parameter λ on MCMF

4.3 相关算法的对比实验与分析

图 3 描述了在相同仿真时间情景下,各算法在交付比率、开销比率、平均时延和平均跳数 4 个评价指标方面的性能对比情况。从图中可以看出,MCMF 获得了优于 Epidemic, Prophet 和 BubbeRap 3 种算法的性能,具有更高的网络数据交付比率以及更少的开销比率。



(a)交付比率对比 (b)开销比率对比

(c)平均时延对比 (d)平均跳数对比

图 3 各算法的性能比较

Fig. 3 Performance comparison of each algorithm

图 3(a)描述了 4 种算法的交付比率随仿真时间变化的情况。从图中可以看出,Epidemic 保持着最低的交付率,这是由于 Epidemic 采用最简单的方式在网络中洪泛数据包,该方法会大量占用通信带宽和节点的缓存空间,因此容易导致网络拥塞和缓存溢出等现象,从而影响数据包的交付比率。这种方式在数据转发时不对中继节点进行筛选,会造成同一数据的转发次数远高于单拷贝转发方式。因此,这种方式对于资源受限的移动社会网络极为不利,在实际应用中也很难实行。

MCMF 由于利用了多维上下文信息来寻找最优中继节点,特别是利用了节点的社会上下文信息,因此在交付比率上的性能最优,图 3(a)所示,在 80000 s 处,MCMF 的交付比率为 73.5%,比 BubbleRap, ProPHET 和 Epidemic 分别高出 5.5%,29.5%和 43.5%。MCMF 有如此优越的表现主要有如下两个原因:首先,MCMF 利用社会属性相似性匹配来选择与目的节点紧密相关的中继节点集合。相似性匹配的主要思想为:当节点之间存在相似的社会属性和兴趣爱好时,它们会有更高的相遇概率。这样避免了 BubbleRap 和 ProPHET 算法存在的一些问题,由于 BubbleRap 和 ProPHET 算法只选择少数中心节点作为中继节点,容易导致这些中间节点的缓冲区溢出和能量过快耗尽,从而出现数据包丢失的现象。其次 MCMF 使用了上下文信息预测模型,该模型采用半马尔可夫过程来表示节点在不同社区之间的转移概率分布。我们通过了解节点相遇的概率分布可以提高网络的投递率。

图 3(b)给出了随着仿真时间的变化,Epidemic, PorPHET, BubbleRap 和 MCMF 4 种算法开销比率的变化情况。从总体上看,MCMF 算法的开销比率最小。显然,由于 Epidemic 采用洪泛的方式转发数据包,因此其投递开销最大。其他 3 种基于上下文匹配的算法采用选择中继节点的模式,MCMF 算法的开销比率与 BubbleRap 较为接近。在 80000 s 处,MCMF 算法的开销比率为 16,PorPHET 算法的开销比率为 43,而 Epidemic 的开销比率则高达 95。本文提出的 MCMF 算法在开销比率上的表现最好,主要有两个原因:首先,在 MCMF 算法进行数据转发时引入了合理的最大转发跳数,通过合理地控制转发次数来减少投递开销;其次,BubbleRap 和 PorPHET 在中继节点的选择上没有采用目的节点的相关信息(目的节点无关性),而是采用了节点间的相遇概率和中心性度量进行数据转发,这样在网络中可能造成瓶颈问题。在本文提出的方案中,MCMF 首先进行转发节点的邻接匹配,然后进行社会关系匹配,最后利用时间半马尔可夫模型来预测两个节点在指定时间的相遇情况。

图 3(c)描述了 4 种路由算法的平均时延情况,MCMF 的总体表现不如 BubbleRap 算法,Epidemic 表现最差。在 80000 s 处,MCMF 算法的平均时延为 5800 s,介于 BubbleRap(5710 s)和 PorPHET(6214 s)之间。在移动社会网络中,数据包的投递时延通常较大,并且投递时延抖动也很大。图 3(d)描述了 4 种路由算法的平均跳数情况,在 80000 s 处,MCMF 的平均跳数表现不错,为 1.9,低于 BubbleRap(2.26),略高于 PorPHET(1.8)。Epidemic 的平均跳数最高为 5.8。

在 MCMF 算法中,网络被分成不同的社区,当两个节点

处于同一社区时,它们之间很容易转发数据。如果两个节点处于不同的社区,MCMF 采用离散的半马尔可夫链来建模节点的移动模式并进行社会相似性匹配,因而,在进行下一跳中继节点的选择时其比 BubbleRap 和 PorPHET 更加准确。最终 MCMF 提高了数据交付比率和开销比率。根据以上仿真结果可以看出,本文提出的路由算法 MCMF 利用移动节点的多维上下文信息提高了移动社会网络的性能。MCMF 的基本原理是利用这些匹配方法寻找与目的节点相遇概率更高的中继节点,最终在每一跳中找到最佳中继节点,有效提高了数据转发的效率。通过以上仿真实验验证了 MCMF 算法的有效性和优越性。

结束语　本文首先以校园中的移动社会网络为例来说明网络中的数据转发问题,分析了移动社会网络中各种上下文信息对节点移动模式的影响。然后提出了系统模型,包括社会相似性建模、社会网络建模和用户移动建模,为数据转发算法的设计提供理论依据。接下来详细描述了基于多维度上下文匹配的数据转发算法,该算法通过使用物理邻接匹配、社会相似性匹配和社会交互性匹配 3 个过程,在每一跳的数据转发过程中寻找最优的中继节点,并设计高效转发策略。最后利用 ONE 仿真平台进行了大量的仿真实验,通过与其他著名算法进行比较,验证了本文提出的数据转发策略的优越性。在下一步的工作中,我们将采用其他真实数据集和移动模型生成的轨迹数据进一步验证本文提出的算法,并对该算法进一步完善。

参 考 文 献

[1] LU Z,WEN Y,ZHANG W,et al. Towards information diffusion in mobile social networks[J]. IEEE Transactions on Mobile Computing,2016,15(5):1292-1304.

[2] ZHANG Y,SONG L,JIANG C,et al. A Social-Aware Framework for Efficient Information Dissemination in Wireless Ad Hoc Networks[J]. IEEE Communications Magazine,2017,55(1):174-179.

[3] LEE H J,NAM J C,SEO W K,et al. Enhanced PRoPHET routing protocol that considers contact duration in DTNs[C]// 2015 International Conference on Information Networking (ICOIN). IEEE,2015:523-524.

[4] MUSOLESI M,MASCOLO C. CAR:Context-aware adaptive routing for delay-tolerant mobile networks[J]. IEEE Transactions on Mobile Computing,2009,8(2):246-260.

[5] HUI P,CROWCROFT J. How small labels create big improvements[C]// Fifth Annual IEEE International Conference on Pervasive Computing and Communications Workshops. IEEE,2007:65-70.

[6] ABDELKADER T,NAIK K,NAYAK A,et al. SGBR:A routing protocol for delay tolerant networks using social grouping[J]. IEEE Transactions on Parallel and Distributed Systems,2013,24(12):2472-2481.

[7] XIAO M,WU J,HUANG L. Community-aware opportunistic routing in mobile social networks[J]. IEEE Transactions on Computers,2014,63(7):1682-1695.

[8] LI F,WU J. LocalCom:a community-based epidemic forwarding scheme in disruption-tolerant networks[C]// 6th Annual IEEE Communications Society Conference on Sensor,Mesh and Ad Hoc Communications and Networks. IEEE,2009:1-9.

[9] BULUT E,SZYMANSKI B K. Friendship based routing in delay tolerant mobile social networks[C]// Global Telecommunications Conference (GLOBECOM 2010). IEEE,2010:1-5.

[10] BOLDRINI C,CONTI M,JACOPINI J,et al. Hibop:a history based routing protocol for opportunistic networks[C]// IEEE International Symposium on World of Wireless,Mobile and Multimedia Networks. IEEE,2007:1-12.

[11] 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.

[12] MOREIRA W,MENDES P,SARGENTO S. Opportunistic routing based on daily routines[C]// 2012 IEEE International Symposium on World of Wireless,Mobile and Multimedia Networks (WoWMoM). IEEE,2012:1-6.

[13] XU F,ZHANG H,DENG M,et al. Social-aware data forwarding in smartphone-based Delay-Tolerant Networks[C]// 2016 IEEE International Conference on Computational Electromagnetics (ICCEM). IEEE,2016:84-86.

[14] LIU G,LI Y. Social-aware data dissemination service in mobile social network with controlled overhead[J]. Pervasive and Mobile Computing,2016,33:127-139.

[15] HOM J,GOOD L,YANG S. A survey of social-based routing protocols in Delay Tolerant Networks[C]// 2017 International Conference on Computing,Networking and Communications (ICNC). IEEE,2017:788-792.

[16] TIAN D,ZHOU J,WANG Y,et al. An adaptive vehicular epidemic routing method based on attractor selection model[J]. Ad Hoc Networks,2016,36(P2):465-481.

[17] KERÄNEN A,JÖRG O,TEEMU K,et al. The ONE simulator for DTN protocol evaluation[C]// Proceedings of the 2nd International Conference on Simulation Tools and Techniques. Brussels,Belgium,2009:81-91.

[18] EAGLE N,PENTLAND A. Reality mining:sensing complex social systems[J]. Personal and Ubiquitous Computing,2009,10(4):255-268.