

断续连接移动自组网络中的数据复制*

石柯¹⁺, 董燕²

¹(华中科技大学 计算机科学与技术学院 计算机科学系,湖北 武汉 430074)

²(华中科技大学 电子与信息工程系,湖北 武汉 430074)

Replication in Intermittently Connected Mobile Ad Hoc Networks

SHI Ke¹⁺, DONG Yan²

¹(Department of Computer Science, College of Computer Science and Technology, Huazhong University of Science and Technology, Wuhan 430074, China)

²(Department of Electronics and Information Engineering, Huazhong University of Science and Technology, Wuhan 430074, China)

+ Corresponding author: E-mail: keshi@mail.hust.edu.cn, <http://cs.hust.edu.cn>

Shi K, Dong Y. Replication in intermittently connected mobile ad hoc networks. *Journal of Software*, 2010,21(10):2677–2689. <http://www.jos.org.cn/1000-9825/3642.htm>

Abstract: Accessing data in a intermittently connected mobile ad hoc networks is a challenging problem that is caused by frequent network partitions due to node mobility and to impairments of wireless communications. This partitioning pattern is studied by examining the statistics of network partitions for a number of mobility models. This paper then establishes the relationship between the network partitioning pattern and the effectiveness of the data replication scheme, and then gives an upper bound of data availability, achieved by an ideal replication scheme under standard operational conditions. The data availability of a totally random scheme can achieve is also given. Based on these results, a novel replication scheme, RICMAN (replication in intermittently connected mobile ad hoc networks), which takes into account the fact that the network is often partitioned into smaller and uses only intermittent connectivity, thanks to mobile nodes traveling across partition, is proposed. In RICMAN, data items are replicated with the nodes using rather stable neighboring topology and with enough resources. A semi-probabilistic data disseminating protocol is employed to distribute the replicas and propagate the updates and can identify the potential mobile nodes traveling across partitions to maximize data delivery. To maintain replica consistency, a weak, eventual consistency model is utilized to ensure all updates eventually propagate to all replicas in a finite delay. Simulation results demonstrate that RICMAN scheme can achieve high data availability with low overhead. With optimized parameters, the data availability is just 10%~15% lower than the upper bound under certain network conditions.

Key words: data replication; intermittently connected mobile ad hoc networks; network partition; data availability

* Supported by the National Natural Science Foundation of China under Grant No.50805058 (国家自然科学基金); the National High-Tech Research and Development Plan of China under Grant No.2007AA04Z186 (国家高技术研究发展计划(863)); the Joint Funds of NSFC-Microsoft Research Asia under Grant No.60933012 (NSFC-微软亚洲研究院联合资助项目)

Received 2007-10-31; Revised 2008-04-02; Accepted 2009-05-05

摘要: 在移动自组网络中,节点的移动或是无线连接的中断会引起频繁的网络分割.因此,访问节点并获取相应的数据是相当困难的.通过理论和统计分析得到特定运动模型对应的网络分割模式,建立了网络分割模式与数据复制有效性之间的联系,推导出了理想复制方法在特定网络环境下能够获得的数据可用性的上限,也指出纯随机复制方法可提高数据可用性.基于上述分析,提出了一种新的数据复制方法 RICMAN(replication in intermittently connected mobile ad hoc networks)来提高断续性连接移动自组网络的数据可用性.该方法将所需数据以副本的形式复制到一系列拓扑结构相对稳定和资源充足的特定节点上,为处于同一分区的节点提供数据服务.副本的分布和更新基于半概率性数据分发协议实现.此协议能够识别可能的跨越多个网络分区的运动节点,由这些节点传播数据及其更新,从而在断续性连接网络中最大化数据传输.为了保持副本的一致性,该方法使用一种弱一致性模型——最终一致性模型,以确保所有的更新最终在有限的延迟内传送到所有的副本处.仿真结果显示,RICMAN 方法能够以较小的开销获取较高的数据可用性,经过优化后,数据可用性仅比理想上限低 10%~15%.

关键词: 数据复制;断续连接的移动自组网络;网络分割;数据可用性

中图法分类号: TP393 **文献标识码:** A

在移动自组网络中,由于距离的变化、障碍物的出现、节点进入节能状态等原因,节点间的连接经常中断.这可能会造成任意两个节点之间不存在端到端的路径连接,网络被分割成多个相互间断开的网络分区.我们称这样的网络为断续性连接的移动自组网络.远程数据访问是固定网络和移动网络的基础应用之一,但在断续性连接移动自组网络中,访问分布式数据要比在全连通网络中困难得多.节点无法访问处于另外一个网络分区内的节点上的数据,相应数据访问的成功率和可用性比全连通网络的要低得多.

将数据复制到若干移动节点上是解决移动自组网络中数据可用性低的一种有效策略.如图 1 中的例子所示,在节点 Q 移动造成网络分割之前,将数据 D_1 和 D_2 的副本分别复制到节点 C 和 T 上(这两个节点在网络分割后处在不同的网络分区中).这样,尽管发生了网络分割,仍然能够使网络中所有的移动节点都可以访问相应数据.假设每个数据有 k 个副本,分别存储在网络中的不同节点中.一个节点为了获得相应的数据,只需访问存储该数据副本的 k 个节点中的一个即可.在全连通网络中,网络的连通性为 100%,我们可以假定,所有节点的可用性是独立的,并且符合相同的随机分布(i.i.d). k 的值可以根据数据可用性的期望值 ε 和节点平均可用性 α 来确定.

$$k = \frac{\log(1-\varepsilon)}{\log(1-\alpha)} \quad (1)$$

基于等式(1),如果网络中节点的可用性高(例如,节点失效不频繁),则通常只需建立较少的副本就能达到较高的数据可用性.然而在移动自组网络中,节点的可用性不再是独立分布的,而是依赖于节点的位置.假设一个节点只有一个邻接节点,当它不能与其邻接节点建立连接时,它就无法访问系统中的任意一个副本.因此,等式(1)在移动自组网络中并不适用.所以,我们需要为移动自组网络设计符合其网络环境的复制方法.该方法必须尽可能地保证在每个网络分区中存在一个有效的数据副本.

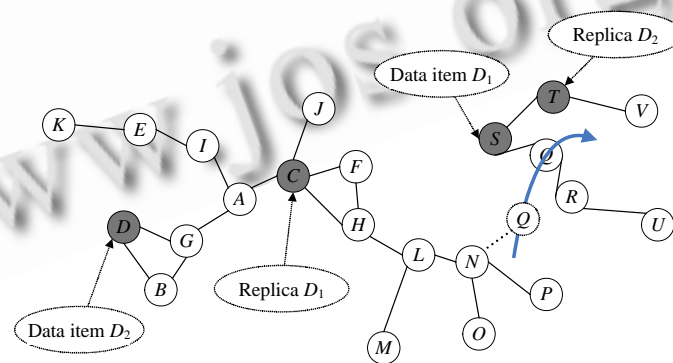


Fig.1 Data replication can improve data accessibility

图 1 数据复制方法能够提高数据的可用性

即使所有分区中都有对应的数据副本,我们仍须考虑数据更新问题.一旦数据更新了,这个更新需要尽快地传输到位于其他网络分区的副本处;否则,处在其他网络分区中的节点只能访问旧的数据.因此,移动自组网络的网络分割模式对于数据复制的性能影响极大.首先,有限的副本应该被尽可能地分配到较大的网络分区中;其次,副本必须最大限度地利用断续的连接机会进行更新,如图 1 所示中数据 D_1 更新后,必须利用 Q 节点的移动将两个被分割开的网络分区再次连接起来的机会更新副本.

本文研究特定运动模型对应的网络分割模式,建立了网络分割模式与数据复制有效性之间的联系,推导出了理想复制方法在特定网络环境下能够获得的数据可用性的上限,也指出纯随机复制方法可提高数据可用性.基于上述分析,本文提出了一种新的数据复制方法 RICMAN(replication in intermittently connected mobile ad hoc networks)来应对网络分割频繁发生的断续性连接移动自组网络.在此方法中,部分具有相对稳定拓扑结构和足够资源的节点容纳数据副本,副本的分布和更新基于半概率性数据分发协议实现,此协议能够识别出可能跨越多个网络分区的运动节点,由这些节点传播数据及其更新.为了保持副本的一致性,本方法使用最终一致性模型以确保所有的更新最终在有限的延迟内传送到所有的副本处.

1 相关研究

数据复制是分布式系统中以增加存储空间和通信开销为代价来提高数据可用性和缩短数据读/写操作时间的一种常用方法,很多技术文献都提出过该方法的实现方案,分别用于广域网络^[1]、内容分发网络^[2]和对等网络^[3].但是,这些研究都是关于有线网络的.在有线网络中极少发生网络分割,并且节点的可用性和相应的网络连接通常被认为是相互独立的.文献[3]提出的用于对等网络的数据复制策略虽然考虑了节点的动态加入和退出,但这些不会影响底层网络连接,也不会造成其他节点间连接的中断.

在移动计算领域,文献[4,5]曾经提出过一些考虑节点断续性连接的数据复制方法.这些方法旨在降低无线网络连接的开销,补偿无线通信造成的可靠性下降,通过对数据可用性、数据一致性和通信开销进行权衡找到较好的解决方案.但在这些研究中,无线通信都是单跳的,无法用在多跳的自组网络环境中.

Hara 曾提出过一系列在自组网络中进行数据复制的方法^[6,7].为了在网络分割的情况下保证数据的可用性,这些研究通过定时的副本再分配优化副本在网络中的位置.这些方法都假设节点数据访问频率已知且固定,并将其作为副本再分配的依据.然而,数据访问频率在实际中是经常变化的,而且副本再分配的间隔也难以确定,间隔太长,难以提高数据可用性,间隔太短,频繁的副本拷贝又会带来极大的网络开销.

类似于 Hara 的研究,Wang^[8]和 Huang^[9]的研究也考虑了副本的分配问题.根据特定的网络分割探测方法,当且仅当网络拓扑结构发生变化可能导致网络分割时,数据才进行复制.但在他们的研究中,网络分割的探测取决于所使用的运动模型.他们的方法假设所有移动节点的位置、移动方向和速度都是已知的.

上述这些方法都未研究特定网络条件下的网络分割模式对数据复制性能的影响,不知道数据可用性的上限,因此无法对算法的有效性和开销做出合适的评价.

2 网络分割模式及其对数据复制性能的影响

2.1 定义

移动自组网络被定义为随机几何图 $G(n, r, A)$,其中, n 个节点均匀分布在有限域 A 中,构成节点集合 V ,若任意两个节点之间的距离小于 r ,它们之间就存在一条直接连接.网络分区 NP 被定义为满足下列要求的 V 的子集, $NP \subseteq V$:(1) 对于任意 $n_i, n_j \in NP$,都存在一条连通路程,无论该路径是直接的还是多跳的;(2) 对于任意 $n_i \in NP, n_k \in V - NP$,不存在连通路程.

节点 v 的度 $d(v)$,代表与节点 v 存在直接连接的节点数量.图 G 的度定义为节点度的平均值.

$$d = \frac{1}{n} \sum_{v \in G} d(v) = (n-1) \times \frac{\pi r^2}{A} \quad (2)$$

2.2 网络分割模式

根据对随机几何图的研究^[10],随机几何图的度确定了随机几何图的连通性.对于任意随机几何图,存在一个阈值 d_c ,若图的度 d 大于此阈值 $d_c, d \geq d_c$,则该随机几何图存在一个无穷大的分区,它是连通的.目前,只能通过数值方法确定 d_c 的值,对于无穷二维平面区域 $d_c \approx 4.5$,对于有界区域 d_c 约为 6.因此,当网络的平均邻居节点数目小于 d_c 时,网络中一定存在多个网络分区,其连接是断续性的.

本文关注的重点就是平均邻居节点数目小于 d_c 的网络,在这样的网络中到底存在多少个网络分区,每个分区的大小(包含多少个节点),这些分区对网络节点间的连接有什么影响.为了回答上述问题,本文用网络分区大小的分布和节点连接断续时间的分布来描述网络分割模式.网络分区的大小是离散值,因此,我们用概率群分布函数(probability mass function,简称 PMF)来描述其分布, $p(k)=P(\|PN\|=k)$ 定义大小为 k 的分区出现的概率.给定 PMF,可以得到网络分区的平均大小 $S_{NP}^a (S_{NP}^a = \|PN\|)$.

$$S_{NP}^a = \sum_{k=1}^n p(k) \times k \tag{3}$$

节点连接断续时间(inter-connection time,简称 ICT)描述网络中任意两节点连接中断的周期长短,这是一个连续值,我们用累积概率密度函数(cumulative distribution function,简称 CDF)来描述其分布.采用计算机数值求解方法获取了这两种分布.

图 2 显示了网络分区的大小分布,当节点的平均度 d 处于 d_c 附近时,分区大小的分布为无尺度的幂指数分布, $p(k)=P(S_{NP}=k)=b \times k^{-a}+c$,网络中存在几个较大的分区包含了大部分节点.随着 d 的减小,网络变得越来越稀疏,分区大小的分布可以近似为几何分布,如公式(4)所示.

$$p(k)=P(S_{NP}=k)=p(1-p)^{k-1} \tag{4}$$

其中, $p = \sqrt{e^{-\lambda}}$, $\lambda=(n-1) \times q, q=\pi r^2 / A$.

当 d 超过 d_c 时,网络趋向全连通,出现一个大的网络分区包含了几乎所有节点.分区大小分布的参数取决于节点的分布,可以通过数值方法求得.运动模型只要不改变节点的分布,相应的网络就有相同的分区大小分布.因此,只要节点的平均度相同,使用任何 Random Trip 模型^[11](包括随机行走的模型、改进的随机 waypoint 模型、基于 Community 的运动模型等),只要能够到达一个稳定的状态,网络分区大小遵循同样的分布.

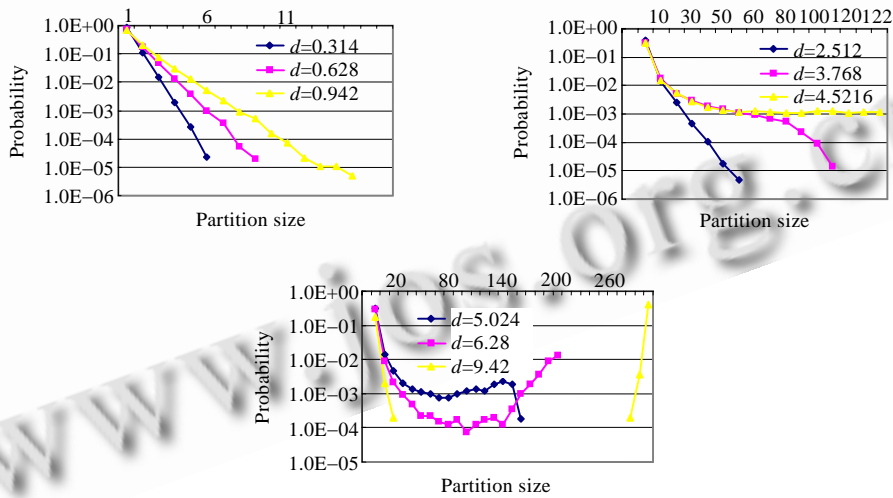


Fig.2 Distribution of partition size for Random Trip models with different average node degrees

图 2 针对任何 Random Trip 模型,不同节点平均度下的网络分区大小分布

另一个重要的网络分割模式的特性是其对具体节点连接的影响,始终存在的这些分区会不会造成某些节

点间始终不存在网络连接.由 ICT 的仿真和统计分析显示,ICT 的分布可以近似为指数分布, $f(t_{ICT})=a \times e^{-at_{ICT}}$,如图 3 所示,右侧的图采用对数轴,CDF 呈直线状.这说明网络分割造成两个节点之间的连接长时间中断的概率是非常小的.虽然网络中存在许多分区,分区的大小相对固定,但随着节点的运动,分区包含的节点不断发生变化,为节点间的连接提供了机会.节点间的连接是断续的,且中断的间隔是有界的,节点间的连接中断一段时间后,重新恢复连接的概率非常高.具体的 ICT 参数与节点的分布和节点间的相对速度(由运动模型决定)相关,可以通过仿真来确定.

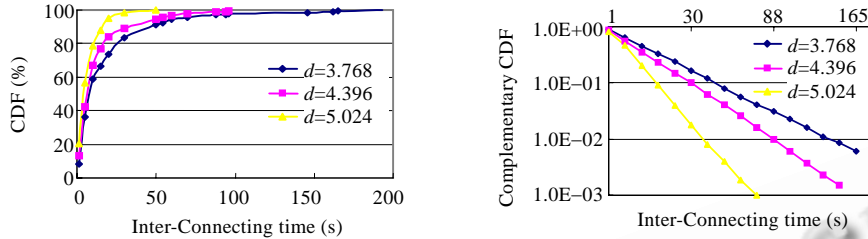


Fig.3 CDF of the inter-connecting time of two nodes

图 3 网络分割下的 ICT 分布

2.3 网络分割模式对数据复制性能的影响分析

网络分区大小的分布对于复制算法的性能有着重要影响,将数据副本分配到较大的网络分区中可以极大地提高数据可用性.数据可用性可以定义为任何一个节点发出的数据访问请求被满足的概率 P_{da} ,如下式所示:

$$P_{da} = P_r + (1 - P_r) \times P_{NP} \tag{5}$$

即为发出数据请求的节点自身或其所在网络分区内的任一节点有所需的数据副本的概率.其中: P_r 是发出数据请求的节点自身有所需的数据副本的概率, $P_r = n_r/n$; n_r 为复制的数据副本个数, $n_r \leq n-1$; P_{NP} 是所在网络分区内的任一节点有所需的数据副本的概率,如公式(6)所示:

$$P_{NP} = \sum_{i=1}^{N_{NP}} P(a, NP_i) \times P(r, NP_i) \tag{6}$$

其中: $P(a, NP_i)$ 是发出数据访问请求的节点属于某一特定分区的概率, $P(a, NP_i) = S_{NP_i}/n$; $P(r, NP_i)$ 是该分区中有存储数据副本的节点存在的概率.

理想化的数据复制算法能够获取所有需要的信息,将 n_r 个数据副本分配到前 n_r 个最大的网络分区内,从而获得最高的数据可用性.因此,前 n_r 个最大的网络分区 $P(r, NP_i) = 1$,其余的 $P(r, NP_i) = 0$,数据可用性如公式(7)所示.这可以看作是数据复制算法所能达到的性能上限.

$$P_{da} = \frac{n_r}{n} + \left(1 - \frac{n_r}{n}\right) \times \sum_{i=1}^{n_r} \left(\frac{S_{NP_i}}{n}\right), \forall i = 1, \dots, n_r, j = n_r + 1, \dots, n, S_{NP_i} \geq S_{NP_j} \tag{7}$$

若采用纯随机算法,副本随机分布在 n_r 个节点上. $P(r, NP_i)$ 可以如公式(8)所示来计算:

$$P(r, NP_i) = 1 - \frac{\binom{n - n_r - 1}{S_{NP_i} - 1}}{\binom{n - 1}{S_{NP_i} - 1}} \tag{8}$$

最终的数据可用性如公式(9)所示.公式(9)给出了数据复制算法的性能基线.给出网络的平均度信息,根据网络分区的分布信息可以得出 N_{NP} 和 S_{NP_i} .给定 n_r 就能知道可获取的性能上限和基线,从而能够有效地评价算法的性能和开销.

$$P_{da} = \frac{n_r}{n} + \left(1 - \frac{n_r}{n}\right) \times \sum_{i=1}^{N_{NF}} \left(\frac{S_{NF_i}}{n} \times \left(1 - \frac{\binom{n-n_r-1}{S_{NF_i}-1}}{\binom{n-1}{S_{NF_i}-1}} \right) \right) \quad (9)$$

上述公式的推导假定只要能够访问到数据,数据访问就是成功的,因此只适用于只读环境.若数据进行更新,还必须考虑访问到的副本是否是最新的.若不是最新的,则数据访问就不能认为是成功的.因此,数据可用性会下降,下降的程度取决于节点的连接断续时间以及数据更新和访问的模式.如图4所示,当数据*i*的所有者对数据进行更新后,拥有数据*i*副本的节点经过 t_{udi} 的延迟后才能收到 $Update_i$,在此期间发生的数据请求 $Data\ access_i$ 到 $Data\ access_{i+1}$ 无法获取最新的数据.对非数据所有者的节点来说,数据可用性的下降程度由更新分发时延、数据访问请求时延和数据更新时延确定.若假定数据更新和访问的分布都是均匀的,则受影响的数据访问请求占整个数据访问请求的比例 $f_U = t_{ud}/t_u, t_u$ 为更新间隔时间, t_{ud} 可定义为ICT的平均值 $E(ICT)$.这一比例可以看作是相应数据副本节点所能提供的数据可用性的下降比例,使用该比例可以得出存在更新时的数据可用性.

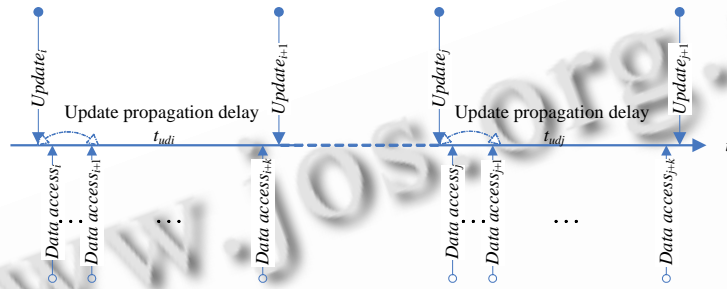


Fig.4 Impact of data update on the performance of replication

图4 数据更新对复制性能的影响

3 RICMAN 方法

根据上面对网络分割模式的讨论,好的数据复制算法应该能够将数据副本放置在较大的分区内,并利用经常在分区间运动的节点传递更新.然而,在实际的网络环境中,获取具体网络分区的信息十分困难,节点通过无限制的广播来获取网络信息进而确定其所在分区的大小和构成会带来极大的网络开销,而且不能保证信息的准确性和及时性.RICMAN 方法的基本思想是,利用可获取的网络局部信息识别两类节点:一类节点在其所在的网络分区中具有相对稳定的拓扑结构和足够的资源(如能源和存储空间等),能够存储数据副本;另一类节点是可能跨越多个网络分区的运动节点,RICMAN 方法使用半概率性的数据分发协议来利用这些节点传播数据,进行副本的分发和更新.此外,RICMAN 方法中假设系统中的所有节点都能相互协作,分享数据,包括保存副本和传输数据.

3.1 半概率性数据分发协议

在半概率性数据分发协议中,每个移动节点处在以其自身为中心、半径为 h 跳的数据分发区域内,并维护该分发区域内节点之间的连接信息.如果接收节点在发送节点的数据分发区域内,则按照分发区域信息在两节点间同步传输数据(确定性传输).如果不能同步传输数据,则将数据发送给一系列有较高概率成功完成传输的节点(概率性传输).节点成功完成传输的概率将根据可用的局部区域信息综合得出.

节点需要收集局部信息来优化数据传输过程.为了使用确定性传输,每个节点维护一张路由表,该表记录了本节点到位于其数据分发区域内其他节点的连接路径.为了使用概率性传输,每个节点还必须维护另一张表来记录其传输指定数据的传输概率.这些信息只在相邻 h 跳的节点间传输.

每个节点定期地根据其本地信息来计算传输指定数据的传输概率.节点 i 传输数据 j 的传输概率可用下面

的公式计算:

$$P_{ij} = \begin{cases} 1, & \text{if node } i \text{ is the receiver of data } j \\ P_{ij}^d, & \text{otherwise} \end{cases} \quad (10)$$

如果节点 i 是数据 j 的接收节点,节点 i 的数据传输概率应等于 1.如果节点 i 不是数据的接收节点,它同样可以在该数据的分发过程中作为一个数据转发节点或者一个数据携带节点, P_{ij}^d 代表它作为数据转发节点或者数据携带节点将数据 j 传输给相应接收节点的能力.我们将 P_{ij}^d 定义为

$$P_{ij}^d = w_{cdc} U_{cdc,i} + w_{col} U_{col,ij} \quad (11)$$

$U_{cdc,i}$ 表示节点 i 局部拓扑结构的变化程度.例如,在最近 T 秒内该节点连接建立和断开连接的次数.较高的值意味着节点 i 有可能和许多不同的节点进行连接. $U_{col,ij}$ 描述了节点 i 和数据 j 的接收节点能够进行直接同步通信的可能性.高 $U_{col,ij}$ 值意味着节点 i 有较高几率与数据 j 的一个或更多的接收节点处于同一数据分发区域,从而进行直接通信.

设 $n(t)$ 为时间 t 时节点 i 的同域节点(处于同一数据分发区域的节点)的集合. t 时刻的 $U_{cdc,i}$ 值通过比较 $n(t)$ 中的成员和上一个周期开始时刻的成员来计算得出:

$$U_{cdc,i} = \frac{|n(t-T) \cup n(t)| - |n(t-T) \cap n(t)|}{|n(t-T) \cup n(t)|} \quad (12)$$

如上式所示, $U_{cdc,i}$ 等于在时间间隔 $[t-T, t]$ 内该节点新增同域节点的数量除以在同一时间内该节点所有的同域节点数量. $U_{col,ij}$ 的计算要容易得多:当数据 j 的接收节点在节点 i 的数据分发区域中时, $U_{col,ij}=1$, 否则 $U_{col,ij}=0$. 这种计算方法只需要我们知道节点之间的逻辑连接关系,不需要知道节点的位置、运动速度和方向等数据,因而不需要 GPS 之类的定位设备或定位方法.

具体的数据分发过程如下:(1) 确定性数据分发:节点 i 总是将接收到的数据 j 通过同步路由机制发送给所有处于其数据分发区域内数据 j 的接收节点;(2) 概率性数据分发包括两个部分:数据的转发和数据的携带.转发过程使数据快速地在同一个网络分区向该数据的接收节点传递.数据 j 会被发送到那些处于节点 i 的数据分发区域中,并且其传输概率不小于其转发阈值 σ 的节点.携带过程使用存储-携带-转发的方法来利用断续性网络连接在不同的网络分区中传输数据.如果节点 i 的传输概率 P_{ij} 不小于它的携带阈值 τ ,在数据 j 被节点 i 转发给其他节点后,数据 j 仍将存储在它的缓冲存储空间中.当携带信息的节点在移动过程中遇到该数据的接收节点时,则将该数据传输给相应的接收节点.基于合适的 τ 值,数据将有较高的可能性在不同网络分区中,在规定的延期内,通过合理的开销(例如,缓冲存储空间、传输次数)进行传输.

3.2 复制节点的选举

在 RICMAN 方法中,我们使用分布式的复制节点选举方法来找寻在各个网络分区中具有相对稳定拓扑结构和具有充足资源(如能源和存储空间)的节点作为复制节点.复制节点的选举在网络构造的初始化阶段完成.如果一个节点在其 h 跳内的邻接节点中是最稳定和処理能力最强的节点,则选举它为复制节点.每个节点都要计算一个能力值来表明它成为复制节点的可能性.该能力值通过如下公式来计算:

$$f = \frac{C}{M} = \frac{w_s S + w_p P + w_e E}{U_{cdc}} \quad (13)$$

上式中: C 是指节点的数据处理能力,该数据处理能力为标准化后的存储空间、剩余能量和处理器能力值的加权平均; M 是指节点的移动性,该移动性定义为节点局部拓扑结构的改变程度.节点的 C 值越大, M 值越小,则其 f 值越大.

网络初始化构造的最初 TA 时间称为复制节点的选举期,在此期间,所有的节点都在其 h 跳内的邻接节点中广播它们的能力值.当 TA 时间过后,每个节点将自己的能力值与其 h 跳内所有邻接节点的能力值进行比较.如果该节点的值最大,它就成为复制节点.所有节点都有一张记录已选出的复制节点的信息表,并且定期地相互交换该表的信息.在 RICMAN 方法中,半概率性数据分发协议将该表中的复制节点作为副本或更新的接收节点.这种

分布式的复制节点选举过程能够保证每个网络分区中至少有一个复制节点存在.在大网络分区中,每隔 h 跳可能会有多个复制节点存在.

因此, h 的取值同时确定了网络中复制节点的个数和节点可获取信息的局部区域的大小,对 RICMAN 方法的性能有重要的影响. h 值太大,会带来相当大的网络开销,当 h 取无穷大时,RICMAN 方法退化为试图以无限制的广播来获取网络信息. h 值太小,会限制算法识别合适的复制节点和数据分发节点的能力,同时使较大的网络分区内存在多个复制节点,带来不必要的存储空间和网络传输消耗.RICMAN 方法采取的策略是根据节点的度和网络规模确定 h 的取值.从第 2 节的分析可知,知道了节点的度可以确定网络分区大小的分布,知道了网络规模就可以知道分区的平均大小 S_{NP}^a . h 值的设置应该保证其所划分区域包含节点的平均数量 n_c 大于 S_{NP}^a . n_c 的值可以通过路由表来统计.

3.3 副本分发

副本分发负责将副本分发给已选举出的复制节点.数据的拥有者根据半概率性数据分发协议将数据传输给其复制节点列表中所有的复制节点,以完成副本分发过程的初始化.在副本分发过程中,较小的 σ 和 τ 值将会加快在所有复制节点间分发副本的速度,但会带来较大的开销,因此需要选择合适的 σ 和 τ 值.

RICMAN 方法采用的是无状态的复制策略,数据的拥有者不需要通过各种形式的数据结构(如表)来维护它的副本状态,因为在移动自组网中,维护副本状态信息的巨大开销是难以接受的,并且由于频繁的网络分割,这些信息很可能是不准确的.图 5 显示了一个在被分割的移动自组网中使用半概率性数据分发协议分发副本的例子.节点 A, L 和 T 是复制节点,它们的数据传输概率都为 1.0.节点 E 是数据的拥有者.其他节点的数据传输概率根据第 3.1 节中给出的公式来计算,计算的结果已在图中标记出来.网络中所有节点的数据分发区域的半径都是 2 跳,节点的 σ 值设定为 0.65, τ 值设定为 0.8.由于节点 A 在节点 E 的数据分发区域内,可以使用确定性方法传输数据,将该数据通过节点 T 传输到节点 A .此时,节点 A 成为该数据的数据分发节点.由于节点 C 的数据传输概率为 0.7,大于 σ 值,所以节点 A 接着将数据传输给节点 C .由于节点 L 位于节点 C 的数据分发区域内,所以节点 C 将数据通过节点 H 传输给节点 L .因为节点 N 和 Q 两者的数据传输概率都大于 σ 值,于是,节点 L 把数据分别传输给节点 N 和 Q .但是,节点 N 只能将同一个数据传输一次,所以节点 N 不能再将该数据传输给节点 Q .而节点 Q 的数据传输概率大于 τ 值,所以节点 Q 携带信息.当节点 Q 移动到另一个网络分区中时,它通过节点 S 将数据传输给节点 T .

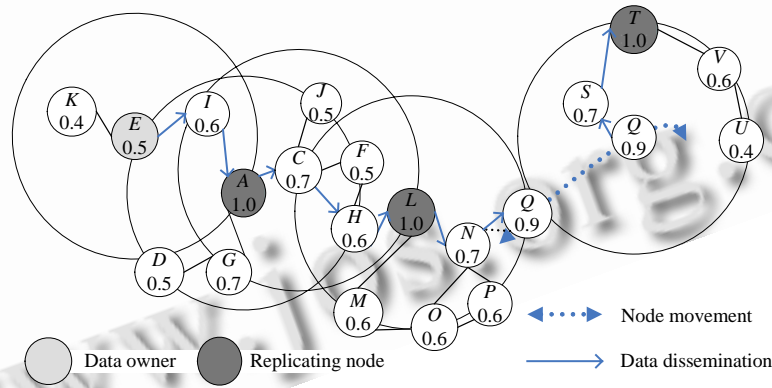


Fig.5 An example of partitioned MANET where semi-probabilistic data dissemination protocol is exploited to distribute replicas
图 5 在被分割网络中使用半概率性数据分发协议分发副本的例子

3.4 更新和数据一致性

由于无线连接的中断和频繁的网络分割,在各复制节点间保持较强的数据一致性是不切实际的.于是,在我

们的复制方法中,采用了一种弱的一致性模型——最终一致性模型,用它来约束所有的副本,使它们最终达到一致的状态.为了保证最终的数据一致性,必须做到:(1) 每个更新必须最终到达所有副本处,并且(2) 所有副本必须以相同的次序提交更新,以保证所有副本提交的最后更新相同.

在 RICMAN 方法中,我们假定只有最初的数据拥有者可以更新该数据.该节点使用半概率性数据分布协议将更新传播给其复制节点表中的所有复制节点.我们可以根据更新的紧急程度和网络运行的状态来调整 τ 和 σ 的值.由于半概率性数据分发协议并不能保证数据传输的结果,所以更新到达各个复制节点的次序可能会有所不同.为了解决更新次序的问题,最初的数据拥有者利用一个不断增加的版本号来记录该数据的更新次序,版本号的初始值为 0 并且每更新一次加 1.复制节点按照更新的版本号的升序次序提交更新.由于更新分发过程中的延迟可能造成版本号小的更新晚于版本号大的更新到达目的节点,因此,目的节点需要暂时缓存先到的、版本号大的更新,等待版本号小的更新到达后再提交更新,以消除乱序现象.

对于数据的请求节点,当其最终成功获得数据时,它将该数据的版本号记录下来.如果当前获得的数据版本号小于该记录值,它就能确定接收到的是过期数据.

3.5 副本检索

副本检索负责使节点能够有效地在复制节点中找到其请求的数据.当一个节点想要访问某一特定数据时,它首先检查其数据分发区域中有没有复制节点.如果在其数据分发区域中存在复制节点,则该节点就可以直接通过同步路由机制获得请求的数据.

如果在数据请求节点的数据分发区域中没有复制节点,则该节点使用半概率性数据分发协议初始化一个数据查询过程来分发它的数据请求.相关的 τ 值设定为一个大于 1 的值,使网络中的任何节点都不会携带这种数据请求信息.这样,该数据请求只能在请求节点所在的网络分区中传播.这是因为,我们的复制方法能够在大多数的网络分区中保证复制节点的可用性,而在不同网络分区中,传播数据请求信息可能会造成不必要的开销.数据请求节点的 σ 值应与其数据传输概率相同.

接收到该数据请求的节点先检查其所在数据分发区域中有没有复制节点,一旦找到一个复制节点,就把一个答复信息(包括它到复制节点的路由信息)传回给数据请求节点以标示到复制节点的路由路径.检索进程可能返回多条答复.请求节点只利用其中的第 1 个答复来建立到复制节点的连接以缩短延时.通过对路由信息的叠加,则可以建立起数据请求节点和复制节点之间的数据传输路径.这里的副本检索与标准的泛洪协议是不同的,它同时使用了确定性和概率性方法来传播数据请求,对该区域的网络拓扑结构的了解能够提高数据请求的传播效率,从而增加请求节点找到其目的地节点的可能性.在如图 2 所示的例子中,所有节点都可以使用确定性的数据传输方法直接获得所需数据.

3.6 副本维护

在完成副本选举和分发的初始化后,副本维护过程负责保证网络中副本个数和分布的稳定,特别是在复制节点故障或节点运动模式发生显著变化的情况下,可以防止数据访问性能的剧烈下降.当一个复制节点意识到它将要离开网络(如关机)时,会自动地把共享的数据传输给 h 跳内 f 值最高的节点,并且使用半概率性数据分发协议将改变后的复制节点表传输给其他节点.但复制节点不是总能提前预见到它将要离开网络.所以,当某一节点发现在一段时间内它访问不了任何复制节点时,就将在其 h 跳的邻接节点内重新开始复制节点的选举过程.

4 RICMAN 方法的性能评估

4.1 仿真设置

我们使用 OPNET 软件通过仿真来评估 RICMAN 方法的性能,并采用基于社区的运动模型^[12]方式.节点分成两种:一种(被称为普通节点)在特定区域内运动,模仿社区内的交互;另一种特殊的节点(称作旅行者)在区域间运动,模仿跨区域的交互.具体的网络分布在 $1000\text{m} \times 1000\text{m}$ 的平面区域内,分为 4×4 的方格,每个方格对应于一个特定的区域.普通节点在特定的方格内做随机行走,节点运动速度 v 在 $0 \sim 3\text{m/s}$ 间选择,周期为 30s.旅行者节

点在整个平面区域内做随机行走,运动速度在 0~4m/s 的范围内选取,周期为 5s.每个移动节点的通信范围是以 r 为半径的圆, r 值约为 100m.网络的 MAC 协议为 802.11b 协议.节点的个数可变,以对应不同的度,其中,20%为旅行者节点.

在仿真实验中,10%的节点被随机选为原始数据的拥有者.数据的大小为 10KB.这些节点负责分发数据并把相应的更新传输给复制节点.每个复制节点都有足够的空闲空间存储所有相关数据.数据更新平均间隔为 120s,服从均匀分布.每个节点产生的数据请求平均间隔为 20s,也服从均匀分布.

间隔 T 决定了节点交换信息和计算数据传输概率的频率,它的设置值应与节点的移动速度成反比关系.这样,运动特点不同的网络会有同样合理的信息准确度.这里的 T 设定为 6s.相对于上面给出的节点通信范围和移动速度的值,该值的大小足以保证在时间间隔 T 内网络拓扑结构发生变化的概率很小.用于计算数据传输概率的加权系数为 $w_{cdc}=w_{cot}=0.5$.半概率性数据分发协议中各参数的默认值为 $\sigma=0.4$, $\tau=0.6$, $h=2$.仿真的时间设定为 1 200s.在仿真的前 600s 内不进行数据的采集,因为节点的分布状态达到稳态需要时间.复制节点的初始化工作在该时间段内完成.

4.2 仿真结果

4.2.1 性能比较

我们分别选取节点总数为 40,80,120,160,200 的情况,得出了采用理想算法获得的性能上限、采用纯随机算法获得的性能基线和采用 RICMAN 获得的性能,如图 6 所示.其中,

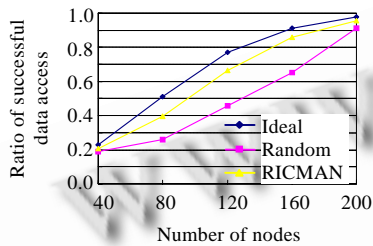


Fig.6 Performance comparison of replication schemes

图 6 复制方法性能的比较

数据成功访问率是指网络上所有节点成功访问所请求的数据与所有的请求之间的比率,用来描述数据可用性.对 RICMAN 方法的参数进行了优化,以获取较高的性能指标.对应的理想算法和纯随机算法中副本的个数与 RICMAN 方法中的复制节点个数相同.当节点数量为 200 时,节点的平均度为 6.28,网络几乎是连通的.因此,理想算法、纯随机算法和 RICMAN 方法均能获取较高的数据可用性,其差别较小, RICMAN 方法相比纯随机算法的性能提高了约 5%.当节点数量为 40 时,节点的平均度为 1.256,网络分割成许多小的分区,限制了理想算法、纯随机算法和 RICMAN 方法对数据可用性的提高.当节点的数量为 80,120 和 160 时,网络中存在较大的分区,因此,相对于纯随机算法,理想算法能够大幅度地提高数据可用性.相应地, RICMAN 方法也能获取较高的数据可用性,比纯随机算法高了 30%,能够达到理想算法 85%~90%左右的性能.

相应地, RICMAN 方法也能获取较高的数据可用性,比纯随机算法高了 30%,能够达到理想算法 85%~90%左右的性能.

4.2.2 半概率性数据分发协议的性能评价

我们比较了使用半概率性数据分发协议与使用同步路由和纯概率性路由方法进行副本分发和更新的性能差别,如图 7 和图 8 所示(节点数量为 160).由于同步路由方法不能穿过不同的网络分区进行传输数据,因此,在我们仿真的结果中,它的数据成功访问率最低.当 $\sigma=0.4$, $\tau=0.6$ 时, RICMAN 方法的数据成功访问率比前者提高了几乎 100%,而只额外增加了约 20%的网络负荷.当使用同步路由方法时,网络负荷主要来自于在节点间找寻端到端的路径,由于经常无法找到这样的路径,不能进行成功的数据访问,因此网络负荷低. RICMAN 方法增加的网络负荷主要来自于数据的成功传输.

图 7 和图 8 也显示了使用纯概率路由协议进行副本分发和更新的性能结果.在该方法中,每个节点都以一定概率(为了比较的方便,通常设为 $1-\sigma$ 或者 $1-\tau$)转发和保存数据,当节点遇到新的邻接节点时,它会把数据重新传输给新邻接节点.在大多数情况下,半概率性数据分发协议比纯概率路由方法的性能更好.虽然使用纯概率路由的数据复制方法能够获得较高的数据成功访问率,但却极大地增加了网络的负荷.与半概率性数据分发方法相比,该方法以两倍的网络负荷为代价获得了与 RICMAN 方法同样的数据成功访问率.这是因为纯概率性路由方法实际上是一种可控的泛洪方法.它对于信息转发节点和携带节点的选择完全是随机的,并没有考虑局部网络拓扑结构的信息和连接历史的信息,因此造成信息的盲目携带和转发,从而加大了网络的负荷.

根据仿真实验结果, σ 和 τ 的取值对分发协议的性能有着极大的影响. 如图 7 所示, 过大的 σ 值可能会导致找不到进行数据转发的节点, 数据成功访问率和网络负荷将会很小. 如果 σ 值过小, 每个节点都会转发数据, 这将增加不必要的传输次数而有助于提高数据成功访问率. 因此, 需要对数据成功访问率和网络负荷这两项指标进行权衡以确定合适 σ 的值. 在仿真过程中, σ 值约为 0.4, 可以达到较好的均衡效果. 观察图 4 可以发现, 对于携带阈值 τ 也存在相同的现象. 较小的 τ 值会使更多的节点参与携带和转发数据, 会加大网络的负荷, 节点也需要更多的空闲缓冲存储空间来存储这些数据. 当 τ 取较大的值时, 网络中只会有很少的节点携带和转发数据, 这会降低数据成功访问率. 本仿真中的 τ 值约为 0.6, 可以达到较好的均衡效果.

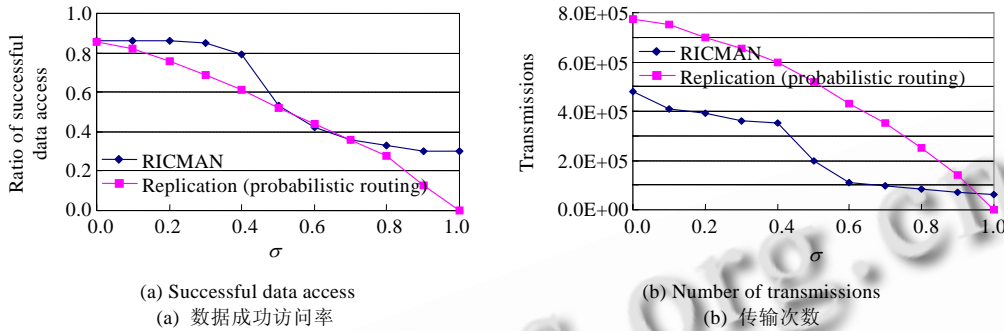


Fig.7 Successful data access ratio and network traffic vs. forwarding threshold σ ($\tau=0.6$)

图 7 数据成功访问率和网络负荷与转发阈值 σ ($\tau=0.6$) 的关系

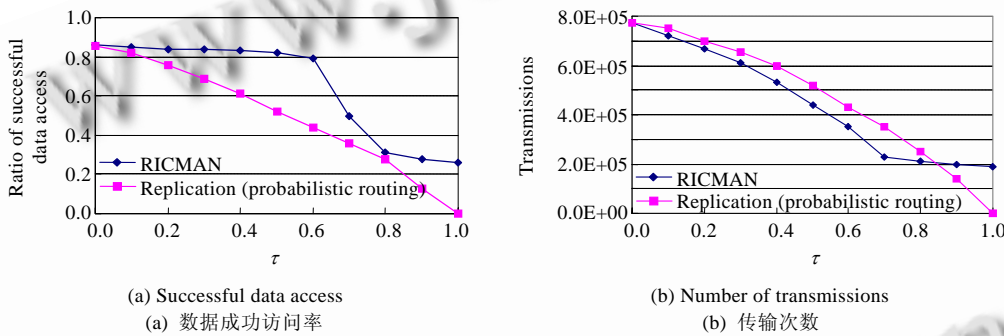


Fig.8 Successful data access ratio and network traffic vs. carrying threshold τ ($\sigma=0.4$)

图 8 数据成功访问率和网络负荷与携带阈值 τ ($\sigma=0.4$) 的关系

4.2.3 数据分发区域半径的影响

参数 h 决定了节点数据分发区域的大小并且直接影响确定性信息维护的开销. 当 h 值较小时, 节点只有有限的确定性信息指导其传输数据或更新, 这将增加传输的次数. 同时, 网络中会出现较多的复制节点, 进一步加大副本分发和更新带来的网络负荷. 当 h 值较大时, 会有更多的确定性信息来加速数据的传输. 然而, 维护数据分发区域信息会加重网络负荷, 可能使更新的数据不能准时地到达每个复制节点, 这会使数据分发区域的信息不准确从而降低传输性能. 此时, 网络中的复制节点数量又会减少, 影响数据的可用性.

从图 9 (节点数量为 160) 可以看出, 当仿真中 h 值设定为 2 时, 增加的确定性信息能够减少盲目的转发, 却没有带来网络开销的加重; 同时, 复制节点的数量也比较适中.

一般来说, 移动性低意味着网络的拓扑变化较小、较慢, 节点的平均度高意味着分区较大, h 可以选取较大的值; 移动性较高, 网络的拓扑变化较大、较快, 确定性信息的维护需要大量的开销, 节点的平均度低, 分区较小, h 应选取较小的值.

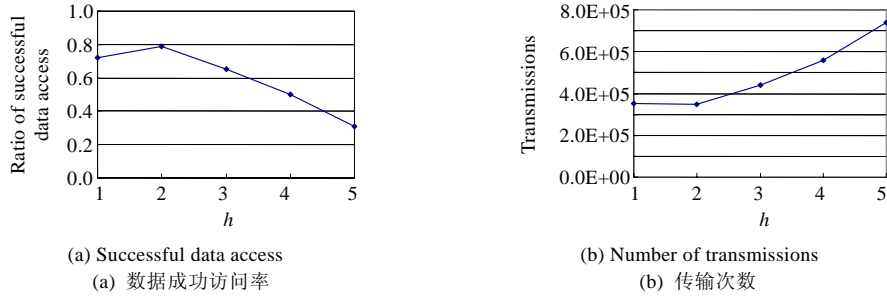


Fig.9 Successful data access ratio and network traffic vs. disseminating radius h

图9 数据成功访问率和网络负荷与数据分发区域半径 h 的关系

4.2.4 运动特性的影响

RICMAN 方法利用所选出的数据携带节点的跨网络分区运动来实现在相互间断开的网络分区中传输数据.换句话说,只有可能成为数据携带者的节点才能存储、携带和转发数据.对于均匀分布的随机行走模型,没有旅行者节点,所有节点的拓扑结构变化程度相同,这会造成难以确定节点能否成为数据转发者和携带者,降低数据成功访问率.在 τ 取一个较大值时,我们有可能找不到任何信息携带节点,这将导致如图 10(a)所示的非常低的数据成功访问率.在这种情况下,首选较低的 σ 和 τ 值.

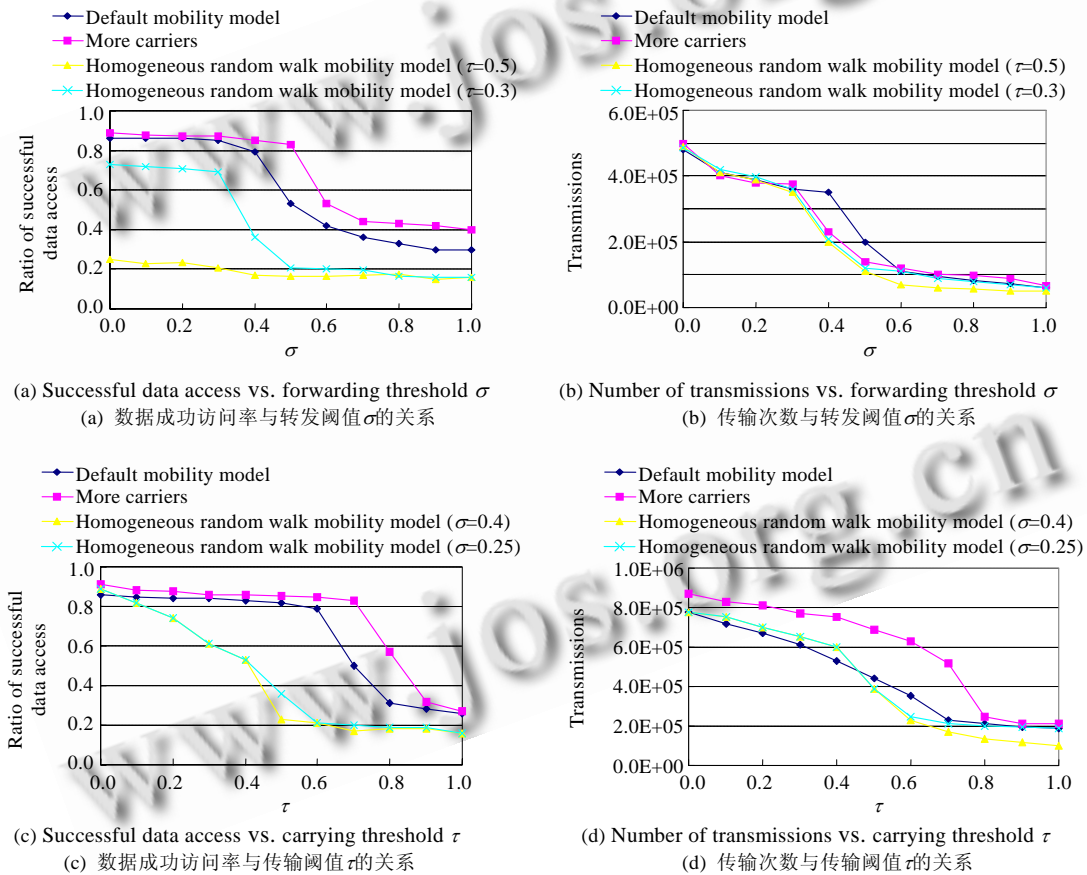


Fig.10 Impact of mobility

图10 运动特性的影响

旅行者节点是在网络的远程分区间进行数据传输的关键,而通过 RICMAN 方法能够识别这些特殊的节点.如图 10 所示,增加旅行者节点的数目可以增加数据成功访问率.与此同时,由于网络中有更多的携带者节点,网络负荷也相应地增加了一些.在这种情况下,由于更多的旅行者节点增加了节点的拓扑结构改变程度,所以首选较高的 σ 和 τ 值.根据仿真结果,节点具有相同的移动性需要选择较低的 σ 和 τ 值,而节点具有异质的移动性可以选择较高的 σ 和 τ 值.

5 结 论

本文分析了网络分割模式与数据复制有效性之间的联系,推导出了理想复制方法在特定网络环境下能够获得的数据可用性的上限,也指出纯随机复制方法可提高数据可用性,并提出了一种新的数据复制方法——RICMAN 来提高断续性连接移动自组网络的数据可用性.仿真结果显示,RICMAN 方法比现有的方法(包括纯随机方法)性能更好,经过优化后数据可用性仅仅比理想上限低 10%~15%.未来的研究包括丰富本地信息的内容.例如,在计算传输概率时加入节点剩余能量的信息,利用相应的预测方法预测节点与复制节点相遇的概率.如何根据网络分割模式和应用需求自适应地决定参数 h , σ 和 τ 的取值也是一个值得深入研究的问题.

References:

- [1] Mondal A, Kitsuregawa M. Effective dynamic replication in wide-area network environments: A perspective. In: Tjoa AM, ed. Proc. of the IWDEXA 2005. Copenhagen: IEEE Press, 2005. 287–291.
- [2] Li YL, Liu MT. Optimization of performance gain in content distribution networks with server replicas. In: Helal S, ed. Proc. of the SAINT 2003. Washington: IEEE Press, 2003. 182–189.
- [3] Cohen E, Shenker S. Replication strategies in unstructured peer to peer networks. In: Steenkiste P, ed. Proc. of the SIGCOMM 2002. Pittsburgh: ACM Press, 2002. 177–190.
- [4] Huang Y, Sistla P, Wolfson O. Data replication for mobile computer. In: Snodgrass RT, ed. Proc. of the ACM SIGMOD'94. Minneapolis: ACM Press, 1994. 13–24.
- [5] Ratner D, Reiher PL, Popek GJ, Kuenning GH. Replication requirements in mobile environments. *Mobile Networks and Applications*, 2001,6(6):525–533. [doi: 10.1023/A:1011862121702]
- [6] Hara T, Madria S. Data replication for improving data accessibility in ad hoc networks. *IEEE Trans. on Mobile Computing*, 2006, 5(11):1515–1532. [doi: 10.1109/TMC.2006.165]
- [7] Hara T, Madria SK. Dynamic data replication using aperiodic updates in mobile ad hoc networks. In: Lee Y, ed. Proc. of the DASFAA 2004. Jeju Island: Springer-Verlag, 2004. 869–881.
- [8] Wang KH, Li B. Efficient and guaranteed service coverage in partitionable mobile ad-hoc networks. In: Kermani P, ed. Proc. of the INFOCOM 2002. New York: IEEE Press, 2002. 1089–1098.
- [9] Huang JL, Chen MS, Peng WC. Exploring group mobility for replica data allocation in a mobile environment. In: Carbonell J, ed. Proc. of the CIKM 2003. New Orleans: ACM Press, 2003. 161–168.
- [10] Goel A, Rai S, Krishnamachari B. Sharp thresholds for monotone properties in random geometric graphs. In: Babai L, ed. Proc. of the TC 2004. Chicago: ACM Press, 2004. 580–586.
- [11] Le Boudec J-Y, Vojnovic M. Perfect simulation and stationarity of a class of mobility models. In: Makki K, ed. Proc. of the INFOCOM 2005. Miami: IEEE Press, 2005. 13–17.
- [12] Musolesi M, Mascolo C. A community based mobility model for ad hoc network research. In: Conti M, ed. Proc. of the REALMAN 2006. Florence: ACM Press, 2006. 31–38.



石柯(1973—),男,河南社旗人,博士,教授,CCF 会员,主要研究领域为无线自组织与传感器网络,嵌入式系统,智能控制,并行分布式计算.



董燕(1972—),女,博士,副教授,主要研究领域为移动通信系统与无线网络.