

基于多点合作的分段式内容分发技术研究*

马帅帅¹ 张一治¹ 张柏林²

(1. 航天科工防御技术研究试验中心 北京 100854)(2. 北京百度网讯科技有限公司 北京 100000)

摘 要 D2D 通信技术存在着通信范围受限、网络拓扑不稳定等问题。针对移动网络中缓存资源非结构化且分布不均的缺陷,研究了移动 D2D 网络资源规划问题,提出了区域协作缓存资源规划机制;同时考虑移动 D2D 网络单节点缓存容量受限、利用率低,单区域对用户移动路径覆盖不足的特点,提出了基于区域协作缓存的移动内容分发机制。论文提出了一种基于区域协作缓存的移动内容分发技术,以期提升移动 D2D 网络环境下的内容分发效率,通过仿真试验验证了该技术的有效性。

关键词 移动内容分发; 区域协作缓存; 移动趋势; 终端直通通信
中图分类号 TP393 **DOI:**10. 3969/j. issn. 1672-9722. 2019. 01. 021

Research on Mobile Content Delivery Technology Based on Muti-points and Area Cooperation

MA Shuaishuai¹ ZHANG Yizhi¹ ZHANG Bolin²

(1. Aerospace Science and Industry Defense Technology Research Testing Center, Beijing 100854)
(2. Beijing Baidu Network Technology Co., Ltd., Beijing 100000)

Abstract D2D communication technology has some problems, such as the limited communication range and the instable network topology. Against the uneven distribution of cache resources in the mobile D2D network, a cache resource planning mechanism based on area cooperation is proposed. As to the problems of limited buffer capacity of single node, the user's insufficient mobile path coverage in single area, etc, this paper puts forward a mobile content distribution mechanism based on area cooperative caching. This paper proposes a mobile content distribution technology based on area cooperative caching (ACC), in order to improve the efficiency of content distribution in mobile D2D network environment. It is proved to be efficient through the simulation.

Key Words mobile content distribution, area cooperative caching, movement trends, D2D
Class Number TP393

1 引言

针对移动蜂窝通信系统中频带资源利用效率低,信道带宽不足的问题,学术界已有广泛的研究^[1]。D2D 技术提供了一种新的通信渠道与数据传输方式,其优势有:信道带宽更大、工作能耗更低;通信距离较近,可以实现频谱资源的复用;数据传输率更高、时延更低,实现多终端间的自组网;具有很强的扩展性,网络的覆盖范围大,边缘用户的数据传输率得到加强^[2]。为此,3GPP 提出在下一代的 IMT-A 系统中引入终端直通通信技术——D2D (Device-to-Device),对移动蜂窝网络进行增强^[3]。

针对 D2D 通信的移动中断问题,有大量学者开展了基于多节点协作分配缓存资源的协作缓存机制的研究^[4]。COCA^[5]是一种基于移动蜂窝网络和终端直通技术的协作缓存机制。在此机制中,节点可以通过点对点通信从邻居节点处获取所需的缓存数据,将这种从邻居节点处获取缓存资源的机制称为 GlobalCache。网络中的基站会定期向处于其通信范围内的节点发送访问率较高的资源,以提升 GlobalCache 的缓存质量^[6]。Groco^[7]是一种基于分组的协作缓存机制,它在 COCA 的基础上进行了扩展,根据节点的移动属性和对资源的不同兴趣将网络中的节点划分至不同的组中,并通过节点分组

* 收稿日期:2018 年 7 月 20 日,修回日期:2018 年 8 月 19 日
作者简介:马帅帅,男,硕士研究生,助理工程师,研究方向:元器件可靠性。

来统筹管理本地和全局的缓存资源分布^[8]。

本文分别提出了面向移动内容分发的区域协作缓存资源规划机制和基于区域协作缓存的移动内容分发机制。在某个节点发起请求时,首先利用资源规划机制确定缓存区域的数量、位置及缓存任务,再利用内容分发机制完成缓存区域向请求节点的内容分发。利用 omnet++ 搭建了系统仿真模型,将 ACC 技术与基于地理位置的随机转发路由技术^[9](GeRaF, Geograpic Random Forwarding)进行对比,得出 ACC 技术在不同网络环境下均有更高的内容分发效率与可靠性,并且在节点移动较快、请求距离较远的情况下优势更为明显。

2 移动 D2D 网络模型

移动 D2D 网络环境如图 1,在区域范围内,有一座基站 B,其余均为移动节点 M。基站可与其覆盖范围内的节点通信,而节点之间也可通过 D2D 技术通信^[10]。

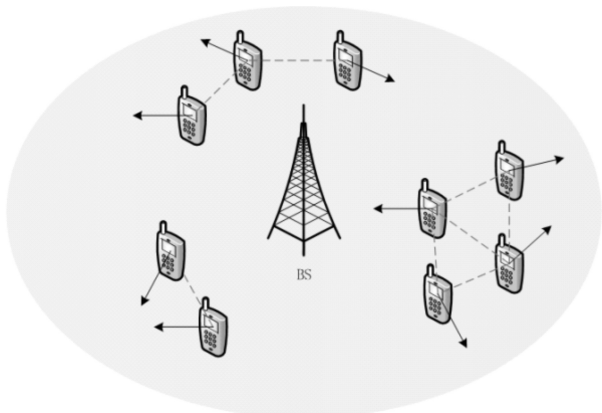


图 1 移动 D2D 网络模型

在移动 D2D 网络中,同时存在蜂窝通信和 D2D 通信两种方式,因此在数据源节点 S 向请求节点 D 分发内容时^[11],可能存在三种数据传输方式:基站直传方式^[12]、基站协传方式^[13]和节点协传方式^[14]。本文重点研究节点协传方式可靠性差的问题^[15]。

3 面向移动内容分发的区域协作机制研究

3.1 面向移动内容分发的区域协作缓存资源规划机制

在 D2D 网络中,移动节点通过缓存数据获取资源,会出现请求与数据源节点远距离多跳传输的情况。合理部署 D2D 网络中的资源,使得源更多、易获取,会提升移动内容分发的效率^[14]。移动 D2D 网络的资源规划问题即选择何种资源部署到哪个节点上的问题^[16]。

本文提出的区域协作缓存资源规划机制,如图 2,根据请求的数据信息,将所需的资源规划到请求节点未来移动路径上的多个缓存区域内,每个区域保存其中一部分资源,并通过协作完成全部内容的部署。

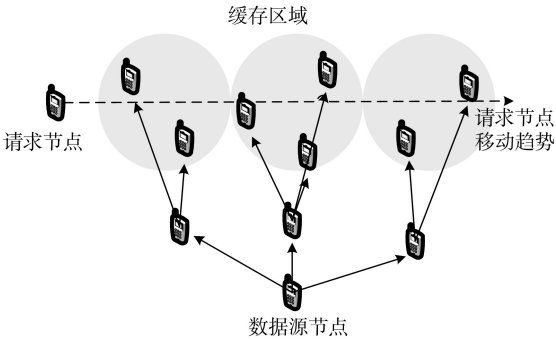


图 2 区域协作缓存资源规划机制示意图

区域协作缓存资源规划机制的基本工作流程如下:

- 1) 获取全网节点地理位置信息并预测请求节点移动趋势;
- 2) 根据 D2D 通信带宽与请求数据量大小计算出所需的缓存区域个数;
- 3) 根据请求节点移动趋势和数据源节点地理位置计算出缓存区域的潜在选择范围;
- 4) 确定各缓存区域的具体位置;
- 5) 根据缓存区域数量及请求数据信息为各区域分配缓存任务。

3.2 基于停留时间分析的缓存区域数量规划策略

在本文所述机制中,根据请求数据包中数据量的不同,规划不同数量的缓存区域。在确定缓存区域的数量时,要保证请求节点在区域内的停留时间大于数据包传输所需时间。记数据包所需传输总时间为 t ,请求节点在每个区域中的期望停留时间为 t_i ,则 $\sum_{i=1}^{n_a} t_i \geq t$,其中 n_a 为所需区域总数。在满足上述条件时,可尽量减少区域的冗余面积,使用较少的缓存节点完成同样的任务,记每个区域内期望缓存节点数为 m_i ,则 $\text{Min} \sum_{i=1}^{n_a} m_i$ 可使得缓存节点最少。同时还要保证区域内至少有两个缓存节点,即 m_i 应有一个最小值 $m_{i\min}$ 。综上可得约束条件:

$$\begin{cases} \sum_{i=1}^{n_a} t_i \geq t, t > 0 \\ m_i \geq m_{i\min} > 2, 0 \leq i \leq n_a \\ \text{Min} \sum_{i=1}^{n_a} m_i \end{cases} \quad (1)$$

由于本机制根据请求来决定缓存区域,计算量

不宜过大,可令每个区域拥有相同的期望缓存节点数 m_p ,期望传输时间也相同,记为 t_{se} 。同时,当各区域期望传输时间之和刚好等于数据包所需传输时间 t 时,缓存节点的冗余最小,则式(1)可简化为

$$\begin{cases} t_{se} * n_a = t, t > 0 \\ m_p \geq m_{p\min} > 2 \\ \text{Min } m_p * n_a \end{cases} \quad (2)$$

式(2)的求解如下:当有节点请求数据时,节点所需数据传输总时间 $t^{[11]}$,则

$$t = \frac{m}{B} \quad (3)$$

上式中 m 为请求数据包大小, B 为通信带宽。根据网络中的节点密度 p (每平方米中的节点数量)可算出圆形缓存区域半径 r : 由于不同缓存区域的大小相同且面积固定不变,而网络中各处的节点密度相同,因此每个缓存区域中所停留的期望节点数 m_p 相同,可据此来计算缓存区域的半径,即

$$r = \sqrt{\frac{m_p}{\pi p}} \quad (4)$$

根据缓存区域半径 r 与请求节点移动速度 v_d 可计算出请求节点在区域内的最大停留时间 t_{sm} ,即请求节点由缓存区域直径穿越区域时,所需要的时间:

$$t_{sm} = \frac{2r}{v_d} \quad (5)$$

由于请求节点可能不是沿直径穿越区域,但在大多数情况下穿越区域的路程不会小于直径的 $\frac{1}{\sqrt{2}}$,因此期望停留时间不会小于最大停留时间的 $\frac{1}{\sqrt{2}}$,取 $t_{se} = \frac{t_{sm}}{\sqrt{2}}$ 作为请求节点在缓存区域中的期望停留时间,即

$$t_{se} = \frac{t_{sm}}{\sqrt{2}} = \frac{\sqrt{2}r}{v_d} \quad (6)$$

根据传输总时间与请求节点在每个区域的期望停留时间可计算出所需区域总数,即

$$\begin{aligned} n_a &= \left\lceil \frac{t}{t_{se}} \right\rceil = \left\lceil \frac{m}{B t_{se}} \right\rceil = \left\lceil \frac{m}{B \frac{\sqrt{2}r}{v_d}} \right\rceil = \left\lceil \frac{\sqrt{2}m}{B \frac{2r}{v_d}} \right\rceil \\ &= \left\lceil \frac{mv_d}{\sqrt{2}B \sqrt{\frac{m_p}{\pi p}}} \right\rceil = \left\lceil \frac{mv_d}{B} \cdot \sqrt{\frac{\pi p}{2m_p}} \right\rceil \end{aligned} \quad (7)$$

由式(2)和(7)可知:

$$m_p * n_a = m_p * \left\lceil \frac{mv_d}{B} \cdot \sqrt{\frac{\pi p}{2m_p}} \right\rceil \quad (8)$$

式(8)与 $\sqrt{m_p}$ 成正比,要使其最小,则要使 m_p 取最小值。由式(2)可知, m_p 最小可取 3,因此某次请求缓存区域的数量应不小于 $\left\lceil \frac{mv_d}{B} \cdot \sqrt{\frac{\pi p}{6}} \right\rceil$ 。若网络环境对于资源利用率的要求不高,而对内容分发可靠性的要求较高,可适当提升 m_p 的取值。

在确定缓存区域的数量后,可以进一步确定每个区域大致所需缓存的数据量。由于每个缓存区域的大小相同,因此可给每个区域规划同等数量的缓存,将总数据量按照区域数量等分可计算出每个区域所需部署的缓存数据量 m_n ,即

$$m_n = \frac{m}{n_a} = \frac{m}{\left\lceil \frac{mv_d}{B} \cdot \sqrt{\frac{\pi p}{2m_p}} \right\rceil} \quad (9)$$

将数据包按此数据量分块,每块数据任务对应其中一个缓存区域。

3.3 基于移动趋势分析的缓存区域位置规划策略

在请求节点移动路线上,最早能接收到全部缓存资源的位置,将所述位置作为圆心的缓存区域记为“最快到达缓存区域”。每次确定缓存区域位置时,都要首先计算“最快到达缓存区域”的位置。

在无阻挡的地理环境下,网络中节点按照直线移动,此时“最快到达缓存区域”位置的计算方法如下:

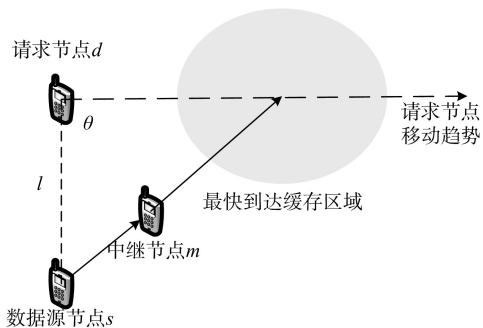


图3 无阻挡环境下计算最快到达缓存区域位置

由图3,在 T_0 时刻请求节点 d 与数据源节点 s 之间距离为 l 。请求节点 d 有一近似直线传输方向,在 t 时刻后与最快到达的中继节点相遇,此相遇位置为 d_i 。请求节点 d 有一平均速度 v_i ,因此 d 与 d_i 之间距离为 $v_i t$ 。请求节点 s 通过一次或多次中继节点携带数据,将数据包从 s 位置传输到 d_i 位置。假设各中继节点 m 有相同的移动方向,均向 d_i 直线移动,此时携带缓存的中继节点将会以最快的速度与请求节点相遇,因此 d_i 是请求节点最早可以获得数据的地点。在 d_i 之前的位置,请求节点不可能与带有缓存的中继节点相遇,在 d_i 之后的位置均有可能

相遇。 ds 与 dd 之间夹角为 θ ,根据余弦定理:

$$\cos \theta = \frac{l^2 + (v_1 t)^2 - (v_2 t)^2}{2lv_1 t}$$

(10)

由上式可解得:

$$t = \frac{lv_1 \cos \theta \pm l \sqrt{v_1^2 (\cos \theta)^2 - v_1^2 + v_2^2}}{v_1^2 - v_2^2}$$

(11)

取上式中的正值作为 t 的取值。由此可确定请求节点在 t 时刻后的潜在位置 d_t 。以 d_t 为圆心,请求节点通信范围为半径做圆,则该圆形区域为“最快到达缓存区域”。

3.4 基于区域协作的缓存资源分配策略

在本文所述的区域协作缓存资源规划机制中,缓存区域的数量很可能不止一个,因此需要给不同的区域分配不同的缓存数据。基本思路就是按照区域的数量,将原有数据包分为区域数量对应的块数,给每个区域都分配独有的缓存资源。如果用户请求的资源是流媒体,则需要按照时间序列将数据按顺序分配给离请求节点最近到最远的区域。同时,需要在分配缓存资源时,提前考虑可能出现的问题,设置一部分区域间的冗余缓存,兼顾可靠性。由于只有在请求节点到达某个区域时,才能判断其在前一区域是否获得了足够的数据,因此可在每个区域设置一部分前一区域的缓存任务,以弥补可能出现的缓存故障。

资源分配策略如下:

- 1)判断请求节点与数据源节点的相对位置与移动趋势,如果两者位置很近,不仅已经处于D2D通信范围内,而且在断开连接前可以传输全部的数据,则无需对数据包进行处理,直接发送即可;
- 2)判断缓存区域的数量,若区域个数仅为1个,则该资源数据量较小,无需对其进行分块处理,直接寻找中继节点发往缓存区域即可;
- 3)若缓存区域的个数大于1,则将其分为区域个数对应的块数。其中,如果该数据包是流媒体资源,则对其按照时间序列的顺序进行分块,保证每一块数据都是可以连续的,可以直接观看其内容;
- 4)在未分配的数据块中挑选时间序列最靠前的,分配给没有数据块任务的区域中离请求节点最近的区域;
- 5)重复步骤4),直至所有区域都分配好缓存数据任务;
- 6)如图4,在基本缓存任务分配好后,为增强可靠性,增设前向协作缓存,内容为每个区域所承担数据任务在时间序列上的前向延伸,具体数据量

视网络情况而定(若请求节点移动方向改变的几率较大,则设置较多的前向协作缓存;若请求节点移动方向改变几率较小,则设定较少的前向协作缓存)。

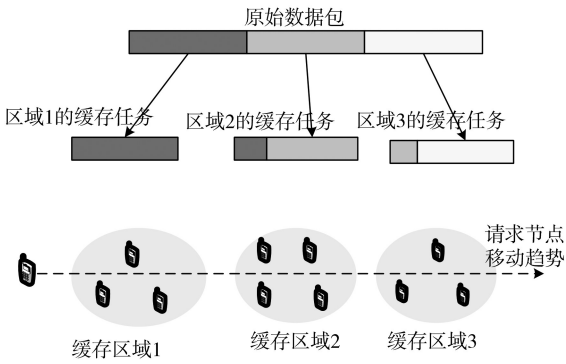


图4 缓存区域所承担的缓存任务

3.5 区域协作内容分发策略

由于各区域所分配的缓存任务各不相同,因此需要各区域协作将内容分发给请求节点。若某个区域中的缓存节点与请求节点的通信时间不足以完成该区域的既定内容分发任务,则由与其相邻的区域完成剩余数据的分发。而当某个区域内容分发任务完成后,请求节点仍未离开,则由当前区域继续向请求节点分发下一区域的内容。具体实现步骤如下:

- 1)在相邻两区域缓存资源充足,而请求节点尚未到达时,由后一区域的节点将该区域数据传输给前一区域的节点,此部分数据记为后向协作缓存数据。两相邻区域是紧密相联的,它们之间必然有节点处于通信范围内,可以在请求节点到达前持续传输数据,数据量最多不超过后一区域既定任务的1/3;
- 2)在请求节点到达某个缓存区域后,判断其在前一区域是否完全获取了既定内容数据。若完全,则由本区域节点直接向其分发既定的任务数据;若不完全,则本区域节点先向请求节点补充发送前一区域内容,完成后再分发本区域既定的任务内容;
- 3)若请求节点在某区域时,本区域任务提前完成,则由本区域内的节点继续向请求节点发送后向协作数据,直至离开;
- 4)重复步骤3)、4),直至请求节点经过所有缓存区域,若收集到足够数据则内容分发完成。

4 性能仿真与分析

4.1 实验环境和参数设定

本次性能仿真实验使用omnet++平台来搭建仿真系统。仿真实验在500m*500m的方形区域范围内进行,区域正中有一基站,其不直接参与内容

分发,其余节点均为移动节点并随机分布在实验区域内。

表 1 仿真实验参数

参数	取值
仿真区域范围	500m*500m
请求时刻间隔时间	6s
单次请求的内容分发时限	30s
D2D通信带宽	10Mb/s
默认D2D通信范围	175m
默认数据包大小	100Mb
默认节点移动速度	10m/s
默认节点规模	40个
默认每跳选择中继节点数	2个

4.2 针对内容分发效率的仿真实验

实验内容为,将本文的 ACC(Area Cooperative Caching)技术与前述的 GeRaF 技术,在不同的网络环境下进行对比,评价标准为限时内容分发成功率。限时内容分发成功率的定义为单次请求内容分发时间小于某一阈值的次数与总请求次数的比值。

4.2.1 不同网络环境下两种技术的性能对比

图 5 对比了在不同节点移动速度的情况下,ACC 技术与 GeRaF 技术的限时内容分发成功率。由图中知:在节点不移动时,两种技术的限时内容分发成功率相近;而在节点移动的情况下,不论速度如何变化,本文技术的限时内容分发成功率均大于 GeRaF 技术。

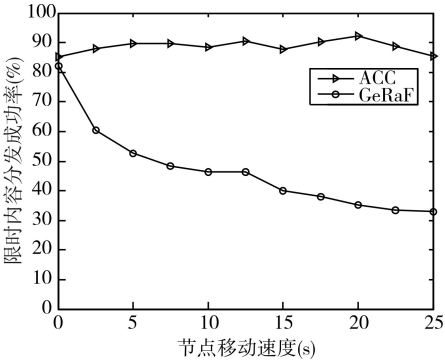


图 5 限时内容分发成功率与节点移动速度的关系 (D2D 通信范围 175m, 节点规模 40 个)

图 6 对比了在不同 D2D 通信范围下,ACC 技术与 GeRaF 技术的限时内容分发成功率。由图可知:随着通信范围的增大,两种技术的限时内容分发成功率都随之增大。

图 7 对比了在不同节点规模下,ACC 技术与 GeRaF 的限时内容分发成功率。由图中数据可知:随着节点规模的增大,两种技术的限时内容分发成功率都有所提高,但本文技术的限时内容分发成功

率始终高于 GeRaF。

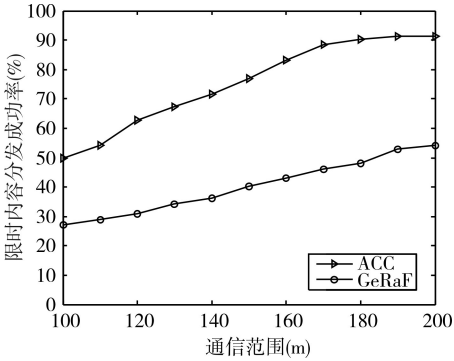


图 6 限时内容分发成功率与 D2D 通信范围的关系 (节点移动速度为 10m/s, 节点规模为 40 个)

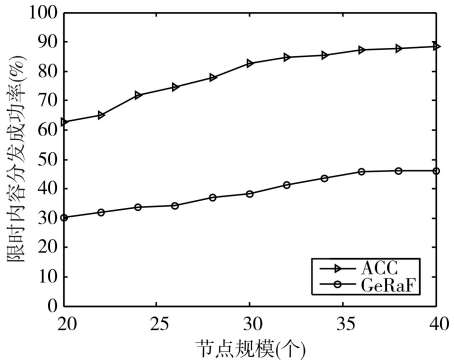


图 7 限时内容分发成功率与节点规模的关系 (节点移动速度为 10m/s, D2D 通信范围为 175m)

综合可见:不论节点规模、节点移动速度、D2D 通信范围如何变化,ACC 技术的限时内容分发成功率对比 GeRaF 均有较大优势,且在网络中节点移动速度更快时,ACC 技术的内容分发效率优势也更大。

4.2.2 不同请求距离下两种技术的性能对比

图 8 对比了在不同请求节点发起请求时刻,其与数据源节点的相对距离处于不同范围时,ACC 技术与 GeRaF 技术的限时内容分发成功率。由图可知:随着请求距离的增大,二者的限时内容分发成功率均呈下降趋势,但 ACC 的下降幅度更小,且成功率始终大于 GeRaF。由此可见,ACC 在请求距离更远时,有更高的内容分发效率与可靠性。

图 9 对比了在不同请求节点发起请求时刻,其与数据源节点的相对距离处于不同范围时,ACC 技术与 GeRaF 技术的平均成功内容分发时间。其中,平均成功内容分发时间的定义为成功完成一次请求内容分发所需的平均时间。由图可知:随着请求距离的增大,二者的平均成功内容分发时间均呈增大趋势,但 ACC 的增长幅度相比 GeRaF 要更小。同时,在请求距离较小时,GeRaF 的平均成功内容分发时间较少;而请求距离较远时,ACC 的平均成

功内容分发时间较少。由此可见,ACC在请求距离更远时的内容分发效率更高,而在请求距离较近时的内容分发效率不如GeRaF。

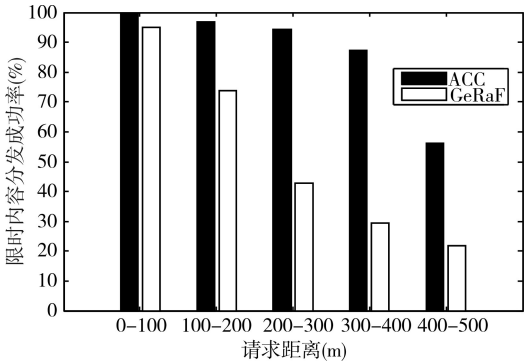


图8 限时内容分发成功率与请求距离的关系

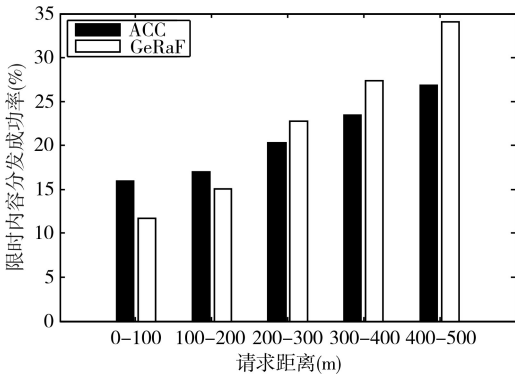


图9 平均成功内容分发时间与请求距离的关系

5 结语

本文提出了一种更适合移动 D2D 网络环境的基于区域协作缓存的移动内容分发技术。从移动 D2D 网络环境中节点移动频繁、通信范围受限等特点出发,提出了区域协作缓存资源规划机制;在此基础上又提出了一种移动 D2D 网络环境下基于区域协作缓存的内容分发技术,通过仿真证明,该技术比现有 GeRaF 技术在不同的网络环境下均对移动内容分发效率有较大提升。

参考文献

[1] 王俊义, 巩志帅, 符杰林, 等. D2D 通信技术综述[J]. 桂林电子科技大学学报, 2014, 34(2): 114-119.
WANG Junyi, GONG Zhishuai, FU Jieli, et al. A survey on device-to-device communication [J]. Journal of Guilin University of Electronic Technology, 2014, 34(2): 114-119.

[2] Wang Rong, Luo Xinlong, Lin Jiaru. Cooperative relay selection scheme based on geographical information [C]// 2011 IEEE International Conference on Cloud Computing and Intelligence Systems (CCIS), 2011:387-390.

[3] Chow C Y, Hong V L, Chan A T S. Group-based coopera-

tive cache management for mobile clients in a mobile environment[J]. Iccp, 2004:83-90.

[4] Chow C Y, Leong H V, Chan A T S. Distributed group-based cooperative caching in a mobile broadcast environment [C]// International Conference on Mobile Data Management. ACM, 2005:97-106.

[5] 4G AMERICAS. 4G Americas' Recommendations on 5G Requirements and Solutions, White Paper [EB/OL], 2014.

[6] Zorzi M, Rao R R. Geographic Random Forwarding (GeRaF) for Ad Hoc and Sensor Networks: Multihop Performance [J]. IEEE Transactions on Mobile Computing, 2003, 2(4):349-365.

[7] 顾闻. 基于 D2D 的中继选择算法研究[D]. 南京:南京邮电大学, 2013.

GU Wen. A Study of Relay Selection Algorithms Based On D2D Communication [D]. Nanjing: Nanjing University of Posts and Telecommunications, 2013.

[8] Hara T, Madria S K. Data Replication for Improving Data Accessibility in Ad Hoc Networks [J]. IEEE Transactions on Mobile Computing, 2006, 5(11): 1515-1532.

[9] Cichon J, Cichon K, Kobylanski P. Node Evaluation in the Chord P2P Systems [C]// Fourth International Conference on Dependability of Computer Systems. IEEE Computer Society, 2009:168-175.

[10] Zhuo X, Li Q, Cao G, et al. Social-Based Cooperative Caching in DTNs: A Contact Duration Aware Approach [C]//IEEE Eighth International Conference on Mobile Ad-hoc & Sensor Systems, 2011:92-101.

[11] Ming Z, Xu M, Wang D. Age-based cooperative caching in Information-Centric Networks [J]. Proceedings - IEEE Infocom, 2012:268-273.

[12] Li W, Chan E, Chen D. Energy-Efficient Cache Replacement Policies for Cooperative Caching in Mobile Ad Hoc Network [C]//Wireless Communication & Networking Conference, 2007:3347-3352.

[13] Fu A, Sadeghi P. r-Regular P2P Broadcast Networks: Optimal Delay and Throughput Using Network Coding [C]//IEEE International Conference on Communications. IEEE, 2011:1-5.

[14] Meng J, Pan X. The optimization study of network coding for coding-aware wireless MESH networks [C]// International Conference on Future Computer and Communication. IEEE, 2010: V2-537-V2-540.

[15] Luby M, Watson M, Gasiba T, et al. Raptor codes for reliable download delivery in wireless broadcast systems [C]//Consumer Communications and NETWORKING Conference, 2006. Ccnc. 2006:73-78.

[16] Chow C Y, Leong H V, Chan A. Peer-to-peer cooperative caching in mobile environments [C]//IEEE Computer Society, 2004.