

基于时延优化的蜂窝D2D通信联合用户关联及内容部署算法

柴蓉* 王令 陈明龙 陈前斌

(重庆邮电大学通信与信息工程学院 重庆 400065)

摘要: 针对蜂窝网络传输性能及基站(BS)缓存能力受限,多用户内容请求难以满足用户服务质量(QoS)需求等问题,该文提出一种蜂窝终端直通(D2D)通信联合用户关联及内容部署算法。考虑到位于特定区域的多用户可能对于相同内容存在内容请求,该文引入成簇思想,提出一种成簇及内容部署机制,通过为各簇头推送热点内容,而簇成员基于D2D通信模式关联簇头获取所需内容,可实现高效内容获取。综合考虑成簇数量、用户关联簇头、簇头缓存容量及传输速率等限制条件,建立基于用户总业务时延最小化的联合成簇及内容部署优化模型。该优化问题是一个非凸的混合整数优化问题,该文运用拉格朗日部分松弛法,将原优化问题等价转换为3个凸优化的子问题,并基于迭代算法及Kuhn-Munkres算法联合求解各子问题,从而得到联合成簇及内容部署优化策略。最后通过MATLAB仿真验证所提算法的有效性。

关键词: 蜂窝网络; D2D通信; 用户关联; 内容部署; 业务时延

中图分类号: TN915

文献标识码: A

文章编号: 1009-5896(2019)11-2565-06

DOI: 10.11999/JEIT180408

Joint Clustering and Content Deployment Algorithm for Cellular D2D Communication Based on Delay Optimization

CHAI Rong WANG Ling CHEN Minglong CHEN Qianbin

(School of Communication and Information Engineering, Chongqing University of Posts and Telecommunications, Chongqing 400065, China)

Abstract: Due to the limited transmission performance of cellular network and the buffering capabilities of the Base Station (BS), it is very difficult to achieve the Quality of Service (QoS) requirements of multi-user content requests. In this paper, a joint user association and content deployment algorithm is proposed for cellular Device-to-Device (D2D) communication network. Assuming that multiple users located in a specific area may have content requests for the same content, a clustering and content deployment mechanism is presented in order to achieve efficient content acquisition. A joint clustering and content deployment optimization model is formulated to minimize total user service delay, which can be solved by Lagrange partial relaxation, iterative algorithm and Kuhn-Munkres algorithm, and the joint clustering and content deployment optimization strategies can be obtained. Finally, the effectiveness of the proposed algorithm is verified by MATLAB simulation.

Key words: Cellular network; Device-to-Device (D2D) communication; User association; Content deployment; Service delay

1 引言

随着移动通信技术的快速发展,蜂窝网络面临前所未有的挑战,亟待研究新型通信技术以满足爆

炸性增长的业务需求^[1]。终端直通(Device-to-Device, D2D)技术可支持邻近终端用户设备之间直接进行通信,而无需经基站中继,可实现更高传输速率及终端能耗的显著降低^[2,3],有效提升频谱效率及用户服务质量(Quality of Service, QoS)^[4]。

文献^[5]将联合模式选择、无线资源分配和功率协调问题建模为端到端总速率最大化问题,并通过基站调度算法进行求解,从而有效提高D2D用户的传输速率。在满足传输速率要求的条件下,文献^[6]通过建模联合模式选择和功率控制优化问题,来实

收稿日期: 2018-05-02; 改回日期: 2019-05-21; 网络出版: 2019-07-19

*通信作者: 柴蓉 chairong@cqupt.edu.cn

基金项目: 国家自然科学基金(61571073), 国家科技重大专项(2016ZX03001010-004)

Foundation Items: The National Science Foundation of China (61571073), The National Science and Technology Specific Project of China (2016ZX03001010-004)

现系统中蜂窝用户和D2D用户能效最大化。针对资源分配问题,文献[7]建模能耗最小化问题,并提出一种能量最优模式选择策略进行求解,从而提高系统能效。文献[6,7]均没有考虑资源利用率问题,在资源受限情况下,文献[8]将模式选择问题建模为资源利用率最大化问题,并提出一种基于信道状态的模式抉择机制进行求解,从而获得更高的资源利用率和D2D增益。

通过在接入网基站处或终端用户处缓存热点内容,可有效降低用户业务获取时延,提升业务体验。针对多用户的单小区蜂窝网络D2D通信场景,文献[9]提出一种以信息为中心的无线虚拟化和D2D通信的新型网络架构,实现虚拟移动运营商总收益的最大化。在考虑用户内容偏好和共享概率的前提下,文献[10]提出基于网络卸载增益最大化的用户内容推送策略。

现有基于缓存的蜂窝网络D2D通信的相关研究,较少联合考虑内容部署及多用户内容共享问题,难以保障网络中所有用户的业务需求体验[11]。此外,现有相关研究较少考虑用户业务传输时延特性,难以保障传输能及时敏感用户的需求体验。基于此,本文综合考虑邻近用户内容共享以及不同传输模式业务特性,引入成簇机制,提出一种基于时延优化的蜂窝D2D通信联合用户关联及内容部署算法。

2 系统模型

本文考虑蜂窝D2D通信上行传输系统场景,包括1个基站及 M 个内容请求用户(Request User, RU)及服务用户(Serving User, SU)。假设RU以一定概率请求内容,记RU所需内容文件总数量为 K ,令 $p_{i,k}$ 为用户 i 对内容 k 的偏好概率,则 $p_{i,k} \in [0, 1]$, $0 \leq \sum_{k=1}^K p_{i,k} \leq 1$, $1 \leq i \leq M, 1 \leq k \leq K$ 。为提升用户内容请求服务性能,降低内容获取时延,本文假设基站或SU处可缓存部分热点内容,各RU可通过蜂窝通信模式,即连接蜂窝网络基站获取所需内容或基于D2D通信模式与某一SU交互实现内容获取。简便起见,采用蜂窝通信模式以及D2D传输模式的RU均基于正交频分多址接入(Orthogonal Frequency Division Multiple Access, OFDMA)方式,占用不同正交子载波进行通信,因而信息传输无干扰。图1所示为本文考虑的蜂窝D2D通信系统。给定网络拓扑及用户文件需求,本文将研究联合优化设计RU关联模式及SU内容部署策略。

3 本文所提成簇机制

在蜂窝D2D通信系统中,由于部分RU可能对

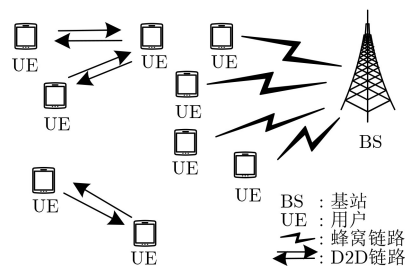


图1 蜂窝D2D通信系统场景图

相同内容文件提出访问需求,通过将部分热点内容缓存至某些SU,可支持多个RUs的同时内容访问服务。针对各种业务类型的差异性及其内容共享难以满足所有用户QoS体验,本文基于成簇的思想,提出一种成簇机制。根据网络状态及用户特性,将网络中用户动态划分为多个簇,各簇由1个簇头用户与多个簇成员用户(简称簇成员)组成,其中,簇头用户(为SU)可与蜂窝基站直接相连,并由基站为其推送热点内容;而簇成员(为RU)通过与簇头关联可获取所需内容。

假设簇头具有内容缓存功能,即为SU,基站可向簇头推送热点内容,簇成员(为RU)采用D2D传输模式与簇头关联,可获取热点内容,从而实现簇内高效内容共享。假设系统内最大簇头数目为 N_1 ,每个簇头可关联多个簇成员,令最大可关联簇成员数目为 N_2 ,簇头 j 的最大缓存容量为 C_j^{\max} , $1 \leq j \leq M$ 。另假设系统内的每个簇成员最多可关联1个簇头。

图2给出了基于用户成簇机制的蜂窝D2D系统场景图,选择特定SU作为簇头,部署相关热点内容,簇成员可采用D2D链路关联簇头获取所需内容,或采用蜂窝通信模式与基站关联进行内容请求。考虑网络中信道特性、簇成员内容需求及簇头缓存内容能力的差异性,不同簇头选择、簇头与簇成员关联以及簇头内容部署策略会导致不同用户业务体验及网络性能,因而需联合优化设计成簇及内容部署策略。

4 优化问题建模

本文提出一种蜂窝D2D通信联合用户关联及内容部署算法,以下对优化问题建模进行详细描述。

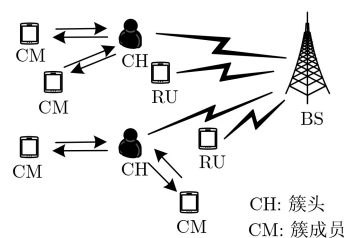


图2 基于用户成簇机制的蜂窝D2D系统场景图

4.1 用户总业务时延建模

系统中用户总业务时延定义为簇内D2D通信时延、簇头内容推送时延及蜂窝通信模式时延之和，本文用 D 表示用户总业务时延， D 可建模为

$$D = D^{\text{CM}} + D^{\text{CH}} + D^{\text{B}} \quad (1)$$

其中， D^{CM} 表示簇内D2D通信时延， D^{CH} 表示簇头内容推送时延， D^{B} 表示蜂窝通信模式时延。式(1)中，簇内D2D通信时延 D^{CM} 可表示为

$$D^{\text{CM}} = \sum_{i=1}^M \sum_{j=1}^M \sum_{k=1}^K \delta_{i,j} \beta_{j,k} p_{i,k} D_{i,j,k}^{\text{d}}, j \neq i \quad (2)$$

其中， $\delta_{i,j} \in \{0,1\}$ 表示用户 i 与簇头 j 的关联因子，若 $\delta_{i,j} = 1$ ，表示用户 i 关联簇头 j ；反之， $\delta_{i,j} = 0$ 。 $D_{i,j,k}^{\text{d}}$ 表示用户 i 关联簇头 j 获取内容 k 时对应的业务时延， $D_{i,j,k}^{\text{d}}$ 可表述为

$$D_{i,j,k}^{\text{d}} = \frac{C_k}{R_{i,j}^{\text{d}}} \quad (3)$$

其中， $R_{i,j}^{\text{d}}$ 表示D2D通信模式下簇成员 i 关联簇头 j 获取内容时对应的传输数据速率。

式(1)中簇头内容推送时延 D^{CH} 可建模为

$$D^{\text{CH}} = \sum_{j=1}^M \sum_{k=1}^K \delta_{j,j} \beta_{j,k} D_{j,k}^{\text{c}} \quad (4)$$

其中， $D_{j,k}^{\text{c}}$ 表示基站向簇头 j 推送内容 k 时对应的业务时延， $D_{j,k}^{\text{c}}$ 可表述为

$$D_{j,k}^{\text{c}} = \frac{C_k}{R_j^{\text{c}}} \quad (5)$$

其中， C_k 表示内容 k 的大小， R_j^{c} 表示基站向簇头 j 推送内容时对应的传输数据速率。

式(1)中蜂窝通信模式时延 D^{B} 可建模为

$$D^{\text{B}} = \sum_{i=1}^M \sum_{k=1}^K \left(1 - \sum_{j=1}^M \delta_{i,j} \beta_{j,k} \right) p_{i,k} D_{i,k}^{\text{b}}, j \neq i \quad (6)$$

其中， $D_{i,k}^{\text{b}}$ 为用户 i 关联基站获取内容 k 时对应的业务时延， $D_{i,k}^{\text{b}}$ 可表述为

$$D_{i,k}^{\text{b}} = D_{i,k}^{\text{t}} + D_{i,k}^{\text{w}} \quad (7)$$

其中， $D_{i,k}^{\text{t}}$ 表示用户 i 关联基站获取内容 k 时对应的数据传输时延， $D_{i,k}^{\text{w}}$ 表示用户 i 关联基站获取内容 k 时对应的排队时延。

4.2 优化模型建立

在满足用户成簇、簇头缓存容量及RU最低传输数据速率需求等限制条件下，本文联合优化设计

用户成簇及内容部署策略，建模蜂窝D2D通信联合成簇及内容部署问题为

$$\left. \begin{array}{l} \min_{\delta_{i,j}, \beta_{j,k}} D \\ \text{s.t. C1: } \delta_{i,j} \in \{0,1\}, \forall i, j \\ \text{C2: } \beta_{j,k} \in \{0,1\}, \forall j, k \\ \text{C3: } \sum_{j=1}^M \delta_{j,j} \leq N_1 \\ \text{C4: } \sum_{i=1}^M \delta_{i,j} \leq N_2, \forall j \neq i \\ \text{C5: } \sum_{j=1}^M \delta_{i,j} \leq 1, \forall i \neq j \\ \text{C6: } \sum_{k=1}^K \beta_{j,k} C_k \leq C_j^{\text{max}}, \forall j \\ \text{C7: } R_i \geq R_i^{\text{min}}, \forall i \end{array} \right\} \quad (8)$$

5 优化问题求解

由于优化问题式(8)是一个非凸的混合整数优化问题，很难用常规的凸优化算法进行求解。本文通过运用麦考密克凸松弛法^[12]和拉格朗日部分松弛法^[13]，将原优化问题等价转换为3个凸优化子问题，并通过改进的Kuhn-Munkres(K-M)算法^[14]来求解各子问题，从而得到对应最优用户总业务时延的联合成簇及内容部署策略。

5.1 原优化问题等价转换

优化问题式(8)中包含 δ, β 等多个布尔变量，属于非线性整数优化问题，且变量之间存在耦合关系，难以直接求解，为实现问题简化，引入解耦变量 $\alpha_{i,j,k} = \delta_{i,j} \beta_{j,k}$ ，将原优化问题转换为

$$\left. \begin{array}{l} \min_{\delta_{i,j}, \beta_{j,k}, \alpha_{i,j,k}} \sum_{i=1}^M \sum_{j=1}^M \sum_{k=1}^K \alpha_{i,j,k} D_{j,k}^{\text{c}} \\ + \sum_{i=1}^M \sum_{j=1, j \neq i}^M \sum_{k=1}^K \alpha_{i,j,k} p_{i,k} D_{i,j,k}^{\text{d}} \\ + \sum_{i=1}^M \sum_{k=1}^K \left(1 - \sum_{j=1, j \neq i}^M \alpha_{i,j,k} \right) p_{i,k} D_{i,k}^{\text{b}} \\ \text{s.t. C1} \sim \text{C7} \\ \text{C8: } \alpha_{i,j,k} = \delta_{i,j} \beta_{j,k} \end{array} \right\} \quad (9)$$

其中，C8为非凸优化限制条件，通过运用麦考密克凸松弛方法，可等效转化为凸优化的限制条件C9~C12。

优化问题式(9)可转换为

$$\left. \begin{aligned} \min_{\delta_{i,j}, \beta_{j,k}, \alpha_{i,j,k}} & \sum_{i=1}^M \sum_{j=1}^M \sum_{k=1}^K \alpha_{i,j,k} D_{j,k}^c \\ & + \sum_{i=1}^M \sum_{j=1, j \neq i}^M \sum_{k=1}^K \alpha_{i,j,k} p_{i,k} D_{i,j,k}^d \\ & + \sum_{i=1}^M \sum_{k=1}^K \left(1 - \sum_{j=1, j \neq i}^M \alpha_{i,j,k} \right) p_{i,k} D_{i,k}^b \\ \text{s.t.} & \text{C1} \sim \text{C7} \end{aligned} \right\} \quad (10)$$

为求解式(10), 利用拉格朗日局部松弛法。

拉格朗日对偶问题可表示为

$$\left. \begin{aligned} \max_{\eta_{i,j,k}, \varphi_{i,j,k}, \theta_{i,j,k}} & \min_{\delta_{i,j}, \beta_{j,k}, \alpha_{i,j,k}} \\ & L(\delta_{i,j}, \beta_{j,k}, \alpha_{i,j,k}, \eta_{i,j,k}, \varphi_{i,j,k}, \theta_{i,j,k}) \\ \text{s.t.} & \text{C1} \sim \text{C7} \end{aligned} \right\} \quad (11)$$

对于给定的拉格朗日因子 $\eta_{i,j,k}$, $\varphi_{i,j,k}$, $\theta_{i,j,k}$, 拉格朗日函数可表示为

$$\begin{aligned} L(\delta_{i,j}, \beta_{j,k}, \alpha_{i,j,k}, \eta_{i,j,k}, \varphi_{i,j,k}, \theta_{i,j,k}) \\ = f_1(\delta_{i,j}) + f_2(\beta_{j,k}) + f_3(\alpha_{i,j,k}) \end{aligned} \quad (12)$$

由于式(12)右边的3个函数中的变量之间无耦合关系, 且对偶函数分为3个独立区域, 故原对偶问题可转换为3个子问题分别进行求解, 即用户关联子问题SP1、内容部署子问题SP2以及联合优化子问题SP3。

5.2 基于迭代算法的子问题求解

由于所建模3个子问题中优化变量和拉格朗日因子之间具有耦合关系, 为获得各子问题的最优解, 应联合求解优化变量和拉格朗日因子, 因此, 本节提出一种迭代算法联合求解优化变量和拉格朗日因子。

给定最大簇头数目 N_1 , 本文考虑不同的簇头选择可能性, 假设可得到 L 种簇头组合策略。本节首先分别针对各种簇头组合策略, 基于给定的拉格朗日因子 $\eta_{i,j,k}$, $\varphi_{i,j,k}$, $\theta_{i,j,k}$, 迭代求解各子问题, 进而选择 L 种簇头组合策略中对应总业务时延最小的联合成簇及内容部署策略作为全局最优策略。

5.2.1 基于K-M算法的子问题求解

由于针对第 l 种簇头组合策略, $1 \leq l \leq L$, 给定拉格朗日因子 $\eta_{i,j,k}$, $\varphi_{i,j,k}$, $\theta_{i,j,k}$, 所建模的各子问题均为包含2元变量的整数优化问题, 可等价于二分图中的匹配问题, 因此, 针对用户成簇子问题、内容部署子问题和联合优化子问题, 在加权二分图 G 中分别建立各个子问题对应的二分图集合, 并采用改进的K-M算法求解各子问题, 进而得到用户成簇因子 $\delta_{i,j}^{(l)*}$, 内容部署变量 $\beta_{j,k}^{(l)*}$ 及联合优化变量 $\alpha_{i,j,k}^{(l)*}$ 。

5.2.2 拉格朗日因子更新

根据局部最优解 $\delta_{i,j}^{(l)*}$, $\beta_{j,k}^{(l)*}$, $\alpha_{i,j,k}^{(l)*}$, 利用梯度迭代算法可更新拉格朗日因子

$$\eta_{i,j,k}(t+1) = [\eta_{i,j,k}(t) - \omega_1(\alpha_{i,j,k}(t) + 1 - \delta_{i,j}(t) - \beta_{j,k}(t))]^+ \quad (13)$$

$$\varphi_{i,j,k}(t+1) = [\varphi_{i,j,k}(t) - \omega_2(\delta_{i,j}(t) - \alpha_{i,j,k}(t))]^+ \quad (14)$$

$$\theta_{i,j,k}(t+1) = [\theta_{i,j,k}(t) - \omega_3(\beta_{j,k}(t) - \alpha_{i,j,k}(t))]^+ \quad (15)$$

其中, 迭代指数 ω_x , $x \in \{1, 2, 3\}$ 是步长。

针对第 l 种簇头组合, 通过迭代重复上述步骤, 可实现算法收敛, 得到第 l 种簇头组合下的最优用户关联及内容部署策略, 进而得到最优用户业务时延, 最后, 通过比较 L 种簇头组合对应的用户业务时延, 可确定最优用户关联及内容部署优化策略。本文所提具体算法如表1所示。

表1 联合用户关联及内容部署算法

(1) 确定 L 种簇头组合策略;
(2) for $l = 1$, 针对第 l 种簇头组合策略;
(3) 设置最大迭代次数 T^{\max} 和最大容忍值 ε ;
(4) 初始化拉格朗日因子 $\eta_{i,j,k}$, $\varphi_{i,j,k}$, $\theta_{i,j,k}$;
(5) 重复主程序;
(6) 求解用户关联子问题得到局部变量 $\delta_{i,j}$;
求解内容部署子问题得到局部变量 $\beta_{j,k}$;
求解联合优化子问题得到局部变量 $\alpha_{i,j,k}$;
(7) 更新拉格朗日因子;
$\eta_{i,j,k}(t+1) = [\eta_{i,j,k}(t) - \omega_1(\alpha_{i,j,k}(t) + 1 - \delta_{i,j}(t) - \beta_{j,k}(t))]^+$,
$\varphi_{i,j,k}(t+1) = [\varphi_{i,j,k}(t) - \omega_2(\delta_{i,j}(t) - \alpha_{i,j,k}(t))]^+$,
$\theta_{i,j,k}(t+1) = [\theta_{i,j,k}(t) - \omega_3(\beta_{j,k}(t) - \alpha_{i,j,k}(t))]^+$;
(8) 若 $\sum_{i=1}^M \sum_{j=1}^M \sum_{k=1}^K [\eta_{i,j,k}(t+1) - \eta_{i,j,k}(t) + \varphi_{i,j,k}(t+1) - \varphi_{i,j,k}(t) + \theta_{i,j,k}(t+1) - \theta_{i,j,k}(t)] \leq \varepsilon$;
(9) 算法收敛;
返回 $\delta_{i,j}^{(l)*} = \delta_{i,j}$, $\beta_{j,k}^{(l)*} = \beta_{j,k}$, $\alpha_{i,j,k}^{(l)*} = \alpha_{i,j,k}$;
(10) 否则 $t = t + 1$;
(11) 重复步骤(6)—步骤(10), 直到算法收敛或 $t = T^{\max}$;
(12) $l = l + 1$, 重复步骤(5)—步骤(11), 得到 $\delta_{i,j}^{(l)*}$, $\beta_{j,k}^{(l)*}$, $\alpha_{i,j,k}^{(l)*}$ 及 $D^{(l)*}$, 直至 $l = L$;
(13) 比较 L 种簇头组合下的最优业务时延, 选择最优用户关联及内容部署优化策略, 即 $\{\delta_{i,j}^{(l)*}, \beta_{j,k}^{(l)*}, \alpha_{i,j,k}^{(l)*}\} = \arg \min D^{(l)*}$ 。

6 仿真结果及分析

本节应用MATLAB仿真软件评估分析本文所提算法性能。仿真场景考虑存在单个基站、多个RUs及多个SUs。网络内所有基站和用户均分布在 100×100 m的区域内, 基站的坐标位置为(100 m,

100 m), 用户的位置随机分布。仿真中选择的用户数量为8个, 基站的发送功率为26 dBm, 用户最低速率需求为2 Mbit/s, 噪声功率谱密度分别为-174 dBm/Hz, -160 dBm/Hz和-150 dBm/Hz。

图3给出了采用本文所提算法所得用户总业务时延与迭代次数关系图, 分别考虑了簇头数目为2及4, 以及不同子信道带宽对应的系统传输性能。可以看出, 总业务时延在较少迭代次数内可趋于收敛, 表明本文算法的有效性。比较不同子信道带宽对应的总业务传输时延, 可得总业务时延随着子信道带宽的增加而减少。此外, 对比不同簇头数目对应的业务时延性能, 可以看出, 当簇头数目增加时, 对应用户总业务时延增加, 这是因为簇头数目增加导致簇头与基站之间传输链路负荷增加, 而通常情况下, 簇头与基站之间链路性能不及簇内链路性能, 进而导致用户总业务时延增加。

图4给出了采用本文所提算法所得用户总业务时延与基站业务到达速率关系图, 可以看出, 用户总业务时延随着业务到达速率的增加而增加, 总业务时延随着平均服务速率的增加而减少。此外, 最大用户关联簇头数目增加导致用户总业务时延减少, 这是因为, 增加最大用户关联簇头数这一限制条件, 可实现用户簇头关联选择的更大灵活性, 进而导致用户总业务时延性能提升。

图5给出了不同最低传输速率限制条件下, 用户总业务时延与业务到达率关系图。基站平均服务速率设定为125 包/s, 噪声功率为-174 dBm/Hz。

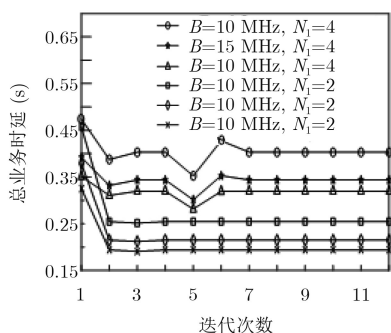


图 3 总业务时延与迭代次数关系图

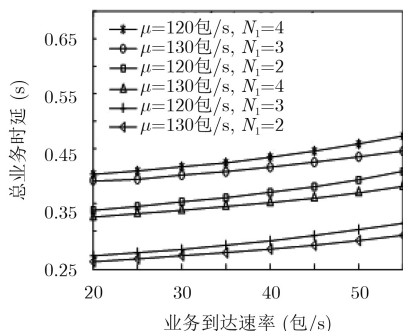


图 4 总业务时延与业务到达速率关系图(不同服务速率)

可以看出, 若最低传输速率限制较低, 用户总业务时延较低, 这是因为大量链路满足传输限制, 因而可根据系统性能及其他限制条件实现联合用户关联及内容部署策略优化; 而若最低传输速率要求较高, 将导致用户总业务时延较高, 原因是部分链路可能无法满足传输速率要求, 因而限制了簇头选择、簇成员关联及内容部署策略的灵活性, 导致系统传输性能受限。

图6给出了在不同噪声功率谱密度的情况下, 用户总业务时延与子信道带宽关系图。可以看出, 用户总业务时延随着噪声功率谱密度的增加而增加, 这是因为更高噪声功率谱密度对应更低传输速率及更高业务传输时延。而总业务时延随着子信道带宽的增加而减少, 因为更高子信道带宽对应更高传输速率及更低传输时延。

图7给出了在不同业务到达速率的情况下, 用户总业务时延与子信道带宽关系图。可以看出, 在相同的业务到达速率条件下, 用户总业务时延随着子信道带宽的增加而减少, 这是因为更高子信道带宽对应更高传输速率及更低传输时延。给定子信道带宽, 总业务时延随着到达速率的增加而增加, 因为基站处理业务的排队时延随着到达速率的增加而增加, 导致用户总业务时延随之增大。同时, 本文算法与文献[15]所提出算法相比较, 可以看出本文算法性能优于文献[15]所提算法性能, 验证了蜂窝D2D通信联合用户关联及内容部署算法可实现更低业务时延。

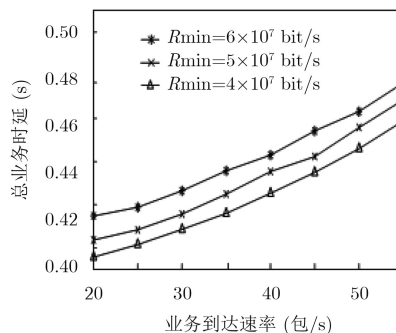


图 5 总业务时延与业务到达速率关系图(不同最低传输速率)

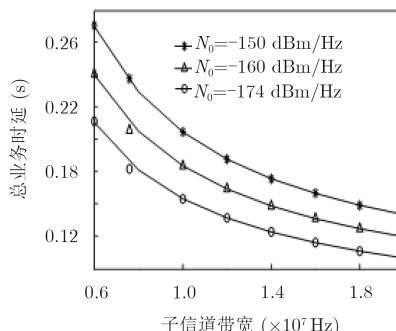


图 6 总业务时延与子信道带宽关系图(不同噪声功率谱密度)

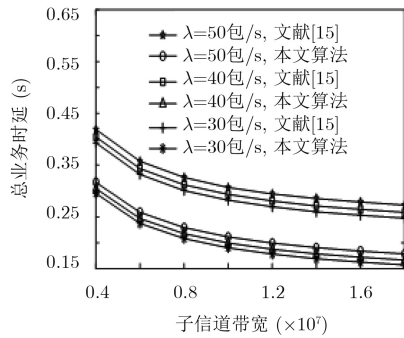


图7 总业务时延与子信道带宽关系图(不同业务到达速率)

7 结束语

针对由多个RU及多个SU构成的单基站蜂窝D2D通信系统场景, 本文提出了一种基于时延优化的蜂窝D2D通信联合用户关联及内容部署算法。为实现多数用户获取性能, 提出一种成簇机制, 支持SU作为簇头, 支持RU的共享获取内容。综合考虑成簇数量、用户关联簇头、簇头缓存容量及传输速率等限制条件, 建立基于用户总业务时延最小化的联合用户关联及内容部署优化模型。本文运用拉格朗日部分松弛法, 将原优化问题等价转换为3个凸优化的子问题, 并提出一种迭代算法联合求解各子问题, 从而得到联合成簇及内容部署优化策略。最后通过MATLAB仿真验证所提算法可实现业务传输时延优化。

参考文献

- [1] TEHRANI M N, UYSAL M, and YANIKOMEROGLU H. Device-to-device communication in 5G cellular networks: Challenges, solutions, and future directions[J]. *IEEE Communications Magazine*, 2014, 52(5): 86–92. doi: [10.1109/MCOM.2014.6815897](https://doi.org/10.1109/MCOM.2014.6815897).
- [2] ASADI A, WANG Qing, and MANCUSO V. A survey on device-to-device communication in cellular networks[J]. *IEEE Communications Surveys & Tutorials*, 2014, 16(4): 1801–1819. doi: [10.1109/COMST.2014.2319555](https://doi.org/10.1109/COMST.2014.2319555).
- [3] FODOR G, DAHLMAN E, MILDH G, *et al.* Design aspects of network assisted device-to-device communications[J]. *IEEE Communications Magazine*, 2012, 50(3): 170–177. doi: [10.1109/MCOM.2012.6163598](https://doi.org/10.1109/MCOM.2012.6163598).
- [4] ZHU Huiling. Radio resource allocation for OFDMA systems in high speed environments[J]. *IEEE Journal on Selected Areas in Communications*, 2012, 30(4): 748–759. doi: [10.1109/JSAC.2012.1205099](https://doi.org/10.1109/JSAC.2012.1205099).
- [5] MA Ruofei, XIA Nian, CHEN H H, *et al.* Mode selection, radio resource allocation, and power coordination in D2D communications[J]. *IEEE Wireless Communications*, 2017, 24(3): 112–121. doi: [10.1109/MWC.2017.1500385WC](https://doi.org/10.1109/MWC.2017.1500385WC).
- [6] WEN Dingzhu, YU Guanding, and XU Lukai. Energy-efficient mode selection and power control for device-to-device communications[C]. Proceedings of 2016 IEEE Wireless Communications and Networking Conference, Doha, Qatar, 2016: 1–7. doi: [10.1109/WCNC.2016.7565099](https://doi.org/10.1109/WCNC.2016.7565099).
- [7] PENDA D D, FU Liqun, and JOHANSSON M. Mode selection for energy efficient D2D communications in dynamic TDD systems[C]. Proceedings of 2015 IEEE International Conference on Communications, London, UK, 2015: 5404–5409. doi: [10.1109/ICC.2015.7249183](https://doi.org/10.1109/ICC.2015.7249183).
- [8] KLUGEL M and KELLERER W. Leveraging the D2D-gain: Resource efficiency based mode selection for device-to-device communication[C]. Proceedings of 2016 IEEE Global Communications Conference, Washington, USA, 2016: 1–7. doi: [10.1109/GLOCOM.2016.7841953](https://doi.org/10.1109/GLOCOM.2016.7841953).
- [9] WANG Kan, YU F R, and LI Hongyan. Information-centric virtualized cellular networks with device-to-device communications[J]. *IEEE Transactions on Vehicular Technology*, 2016, 65(11): 9319–9329. doi: [10.1109/TVT.2016.2518658](https://doi.org/10.1109/TVT.2016.2518658).
- [10] PAN Yijin, PAN Cunhua, ZHU Huiling, *et al.* On consideration of content preference and sharing willingness in D2D assisted offloading[J]. *IEEE Journal on Selected Areas in Communications*, 2017, 35(4): 978–993. doi: [10.1109/JSAC.2017.2680938](https://doi.org/10.1109/JSAC.2017.2680938).
- [11] LI Xiaoshuai, MA Lin, SHANKARAN R, *et al.* Joint mode selection and proportional fair scheduling for D2D communication[C]. Proceedings of the 2017 IEEE 28th Annual International Symposium on Personal, Indoor, and Mobile Radio Communications, Montreal, Canada, 2017: 1–6. doi: [10.1109/PIMRC.2017.8292254](https://doi.org/10.1109/PIMRC.2017.8292254).
- [12] LIBERTI L and PANTELIDES C C. An exact reformulation algorithm for large nonconvex NLPs involving bilinear terms[J]. *Journal of Global Optimization*, 2006, 36(2): 161–189. doi: [10.1007/s10898-006-9005-4](https://doi.org/10.1007/s10898-006-9005-4).
- [13] LEMARÉCHAL C. S. Boyd, VANDENBERGHE L. Convex optimization[J]. *European Journal of Operational Research*, 2006, 170(1): 326–327. doi: [10.1016/j.ejor.2005.02.002](https://doi.org/10.1016/j.ejor.2005.02.002).
- [14] HUANG Yifei, NASIR A A, DURRANI S, *et al.* Mode selection, resource allocation, and power control for D2D-enabled two-tier cellular network[J]. *IEEE Transactions on Communications*, 2016, 64(8): 3534–3547. doi: [10.1109/TCOMM.2016.2580153](https://doi.org/10.1109/TCOMM.2016.2580153).
- [15] JIANG Wei, FENG Gang, and QIN Shuang. Optimal cooperative content caching and delivery policy for heterogeneous cellular networks[J]. *IEEE Transactions on Mobile Computing*, 2017, 16(5): 1382–1393. doi: [10.1109/TMC.2016.2597851](https://doi.org/10.1109/TMC.2016.2597851).

柴蓉: 女, 1974年生, 教授, 研究方向为通信网络架构及关键技术、无线资源管理及移动性管理技术等。

王令: 女, 1993年生, 硕士生, 研究方向为无线通信、无线资源管理及网络虚拟化等。

陈明龙: 男, 1993年生, 硕士生, 研究方向为无线资源管理、网络虚拟化及内容缓存等。

陈前斌: 男, 1967年生, 教授, 研究方向为无线通信、通信网络理论、软件定义网络、多媒体技术等。