

Research Article

Energy Efficiency for Data Offloading in D2D Cooperative Caching Networks

Weiguang Wang , **Hui Li**, **Wenjie Zhang**, and **Shanlin Wei**

School of Electronics and Information, Northwestern Polytechnical University, Xi'an 710129, China

Correspondence should be addressed to Weiguang Wang; laughing_wwg@163.com

Received 6 December 2019; Revised 21 May 2020; Accepted 11 June 2020; Published 27 June 2020

Academic Editor: Manuel Fernandez-Veiga

Copyright © 2020 Weiguang Wang et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

D2D communication improves the cellular network performance by using proximity-based services between adjacent devices, which is considered an effective way to solve the problem of spectrum scarcity caused by tremendous mobile data traffic. If the cache-enabled users are willing to send the cached file to the requesters, the content delivery traffic can be offloaded through the D2D link. In this paper, we strive to find the maximum energy efficiency of the D2D caching network through the joint optimization of cache policy and content transmit power. Specifically, based on stochastic geometry-aided modeling of the network, we derive the data offloading rate in closed form, which jointly considers the effects of success sensing probability and success transmission probability. According to the data offloading rate, we formulate a joint optimization problem integrating cache policy and transmit power to maximize the system energy efficiency. To solve this problem, we propose two optimization algorithms: the cache policy optimization algorithm based on gradient update and the joint optimization algorithm. The simulation results demonstrate that the joint optimization has twice the superiority in improving the energy efficiency of the D2D caching network compared with other schemes.

1. Introduction

With the popularity of intelligent electronic devices and the emergence of various applications, the explosive growth of mobile data traffic in recent years will reach the limit of current cellular network infrastructure. According to Cisco's latest Visual Networking Index (VNI) report, mobile video data traffic will account for 82% of the global mobile data traffic to 2021 [1], and the core network will be overloaded by 2025 [2]. The tremendous increase in video traffic has brought severe challenges to the 5th generation mobile networks. However, compared to the massive data contents, the cache spaces of nodes are limited. Therefore, the caching placement problem is the key to optimizing the performance of edge caching [3]. In Device-to-Device (D2D) communications, mobile devices also play a role in content delivery and enable direct communication links between users. The D2D communications can increase cellular network throughput, reduce energy consumption, and improve spectrum utilization, which can also promote the research of the 5th genera-

tion (5G) mobile communication technology [4, 5]. Cache-enabled D2D communications have been shown to achieve significant offloading gains in networks, and there are higher chances to retrieve the desired data pieces right from the content-related users [6]. Generally, cache-enabled D2D communications included three basic processes, namely, cache placement, content sensing, and content delivery. In the cache placement process, contents are cached by the cache-enabled users based on the cache strategy. When a user device acts as a content requester, the process of finding the requested content in the surrounding cached users is called content sensing. Finally, when the file in the cache user is requested and the request information can be sensed, the content transfer process will be triggered immediately.

In the above three processes, the content placement process is an essential prerequisite of content perception and content transmission, which determines whether the requesting user can find the required content in the nearby cache-enabled users. Therefore, the cache policy must be designed carefully, which is very important to improve the

offload gain of the D2D cache network [7, 8]. Due to the mobility and spatial randomness of mobile users, the authors use random geometric theory to study user association and caching strategies in D2D caching networks in [9]. In [10], they optimize the caching policy with the knowledge of user preference and activity level to maximize the offloading probability for cache-enabled device-to-device communications. Rao et al. [11] optimize the cache strategy to improve the probability that the requester can find relevant content within the D2D communication range. However, they ignore the influence of channel fading and network interference on the content delivery process. As a supplement, the authors consider the channel changes among users when optimizing the cache policy to improve the probability of successful content transmission in [12–14]. Most existing works focus on the optimization of cache policy for different users with very limited cache to improve the proactive cache gain. Due to the limited battery life, some users are always selfish and unwilling to waste their energy to help other users [13]. No matter what purpose some cache users provide request content to requesters, their real purpose is to maximize the energy efficiency of the cache-enabled D2D network [15]. Although D2D caching has a large potential in terms of offloading data traffic, it also brings many urgent problems to be studied. For example, how can we offload more data traffic through the D2D cache? What is more, how can we maximize network energy efficiency? This study is not only a vital part of the global green energy conservation and environmental protection plan but also conducive to the healthy development of network technology.

In the previous research on cache-enabled D2D communication [12–14], the energy cost of cache-enabled users is ignored. They assume that the battery energy is infinite, the optimal cache policy may not achieve the maximum energy efficiency. In [16], a distributively implementable algorithm is proposed to cluster users. It assumes that only users in the cluster can establish D2D links, but the proposed optimal cache policy cannot maximize the data offloading, and the energy cost is very high. This is because when only users in the cluster can establish D2D links, it cannot obtain files from the nearest clusters with low transmit power. In [17], in order to reduce the energy cost in the process of content transmission, Chen et al. independently optimized the cache policy and transmit power, then analyzed the trade-off between traffic offloading and energy cost. However, there is a balance between cache policy and transmission power to maximize energy efficiency for cache-enabled D2D networks.

In this paper, we jointly optimize the cache policy and transmit the power to maximize the energy efficiency of the cache-enabled D2D network. Specifically, the main work of this paper is summarized as follows:

Firstly, based on stochastic geometry theory, we derive the data offloading rate in closed form, which jointly considers the effects of success sensing probability and success transmission probability. In the probability of successful data transmission, we take the data transmission rate and the D2D establishment constraint into account.

Secondly, since energy efficiency is an NP-hard problem, we propose an iterative optimization algorithm based on

independent optimization (optimize cache policy or transmit power, respectively). These two subproblems are proved to converge to the optimal solution, respectively. The joint optimization algorithm can achieve the maximum energy efficiency.

Thirdly, numerical results indicate that the proposed joint optimization scheme is more than twice as much as any independent optimization scheme in the D2D caching network energy efficiency.

The remainder of this paper is organized as follows. The system model is described in Section 2. The problem formulation is described in Section 3. Joint optimization cache policy and transmit power scheme to maximize the energy efficiency is described in Section 4. Simulation results are demonstrated in Section 5. Finally, the conclusion for this paper is in Section 6.

2. System Model

We considered a cache-enabled D2D network where the cache-enabled users and the requesters are, respectively, modeled as two Homogeneous Poisson Point Process (PPP) with density λ_p and λ_r , as shown in Figure 1. Each cache-enabled single antenna user has a local cache to store files and the same transmit power P_t . The BS can detect user's channel status and location information and also coordinate cache-enabled D2D communication [18]. The signal will be affected by various fading factors in the transmission process. For the large scale fading, we assume that the signal from the user will be attenuated by $r^{-\alpha}$, where α is the path loss exponent and $\alpha > 2$ is satisfied. For small-scale fading, we assumed that any pair of transmitters and receivers are Rayleigh fading channels and obeys a zero-mean complex Gaussian distribution [19]. Rayleigh fading is proved to be an effective statistical model to evaluate the fluctuation of multipath channels [20].

Suppose a limited content library $\mathcal{F} \triangleq \{1, 2, \dots, F\}$, where the 1-st file represents the most popular content and the F -th file represents the least popular content. For simplicity, we assume that all files have an identical size. It has been found that only a small number of files are frequently accessed by most users. More specifically, each user can request files independently from the content library, and the f -ranked file is requested with a probability of $p_f = f^{-\epsilon} / \sum_{i=1}^F i^{-\epsilon}$, where ϵ represents the shape parameter of the file popularity. Each cache-enabled D2D user caches the file independently according to the specific cache policy $\mathbf{q} = [q_1, \dots, q_f, \dots, q_F]$ where q_f is the probability of the f -th file cached by cache-enabled D2D user. Given the limited storage of the wireless terminals, the file cache probability must be satisfied $\sum_{f=1}^F q_f \leq 1$ [17]. According to the feature of the Poisson Process, the distribution of cache-enabled users having the f -th file cached follow HPPP with an density of $q_f \lambda_p$.

3. Problem Formulation

3.1. Success Sensing Probability. Specifically, the success sensing probability that the requester successfully perceives

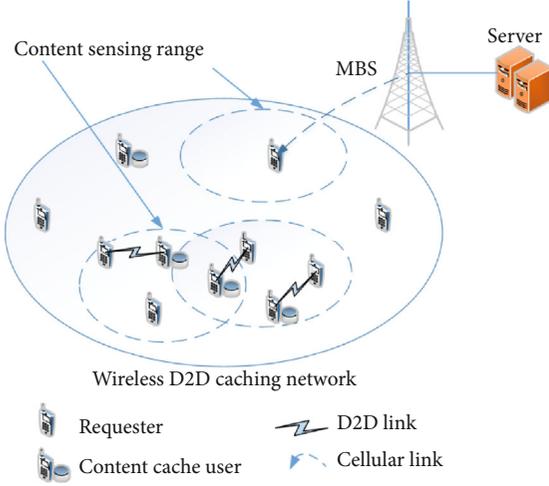


FIGURE 1: Cache-enabled D2D network model.

the required content within the communication range. In the D2D cache network, each requester can randomly and independently request f -th file from the nearest cache-enabled user, $f \in \mathcal{F}$. The location distribution of cache-enabled users (termed cp_0) who have cached f -th file follows a HPPP with density $q_f \lambda_p$. Thus, the probability density function of the association distance r between the requester and the cache-enabled user with f -th file cached is denoted as

$$f_{0f}(r) = 2\pi\lambda_p q_f r \exp\left(-\pi\lambda_p q_f r^2\right) \quad (1)$$

Suppose z is the sensing radius of D2D communication, the probability that the requesters can successfully sensing the file f within the communication range z is formulated as (2).

$$S_f = \int_0^z f(r) dr = 1 - e^{-\pi\lambda_p q_f z^2} \quad (2)$$

The success sensing probability of all files is denoted as

$$S = \sum_{f=1}^F p_f S_f = \sum_{f=1}^F p_f \left(1 - e^{-\pi\lambda_p q_f z^2}\right) \quad (3)$$

3.2. Success Transmission Probability. In practice, users are randomly distributed and the association distance between users will be very long. On the one hand, the cache-enabled users must increase the transmit power to complete long-distance transmission, which not only increases the network interference but also loses the power gain. On the other hand, if they do not change the transmit power, the probability that the requester obtains the file is relatively low and the transmit power consumption is also very wasteful. So we should strictly control the establishment conditions of D2D communication. When the average long-term receiving power of the requester is higher than θ and the data rate is higher than R_0 , the content transmission can be completed successfully. Let θ indicate the D2D establishment threshold.

The probability of successful transmission is expressed as $D_f = \Pr(P_t r^{-\alpha} \geq \theta, R_f \geq R_0)$, where $P_t r^{-\alpha} \geq \theta$ can be rewritten as $0 < r \leq (P_t/\theta)^{1/\alpha} = z$. Obviously, when θ is given, the D2D communication radius is determined by the transmit power. According to Shannon's theorem, the data transmission rate is expressed as $R_f = \omega \log_2(1 + \gamma_f)$. File can be transferred successfully only when $R_f \geq R_0$, which is further equivalent to $\gamma_f \geq \gamma_0$, where $\gamma_0 = 2^{R_0/\omega} - 1$ denotes the SINR threshold for file transmission. Therefore, the equivalent form of D_f can be further expressed as $D_f = \Pr(P_t r^{-\alpha} \geq \theta, \gamma_f \geq \gamma_0)$.

Specifically, the SINR of the requesting user within the communication range is expressed as $\gamma_f = P_t g_{0f} r^{-\alpha} / I_1 + N_0 \omega$, where I_1 is the interference caused by the other cache-enabled users in the communication range. g_{0f} represents the channel gain between the requester and the cache-enabled user with file f cached. The distribution of other cache-enabled users are modeled as HPPP Φ_1 with density $\lambda_r S$. $N_0 \omega$ represents the noise power of the receiver. Further, the success transmission probability can be described as below.

$$\begin{aligned} \Pr\left(P_t r^{-\alpha} \geq \theta, \gamma_f \geq \gamma_0\right) &= \Pr\left(0 < r \leq z, \gamma_f \geq \gamma_0\right) \\ &= \int_0^z \Pr\left(\gamma_f \geq \gamma_0\right) f(r) d(r) \\ &\stackrel{(a)}{\approx} \int_0^z \Pr\left(g_{0f} \geq P_t^{-1} r^{-\alpha} I_1 \gamma_0\right) f(r) d(r) \\ &\stackrel{(b)}{\equiv} \int_0^z \mathbf{E}_{I_1}\left(e^{-P_t^{-1} r^{-\alpha} I_1 \gamma_0}\right) f(r) d(r) \\ &\stackrel{(c)}{\equiv} \int_0^z \mathcal{L}_{I_1}\left(P_t^{-1} r^{-\alpha} \gamma_0\right) f(r) d(r), \end{aligned} \quad (4)$$

where step (a) is due to the consideration of an interference limited region. Step (b) represents a small-scale fading model $g_{0f} \sim \exp(1)$. Step (c) represents the Laplace transform of the random variable I_1 . Specifically, $I_1 = \sum_{i \in \Phi_1 \setminus cp_0} P_t g_i r_i^{-\alpha}$, where g_i indicates the interference channel gain of the i -th cache-enabled user and the requesters, whose distance indicates r_i . The Laplace Transform of I_1 is as shown in (5).

$$\begin{aligned} \mathcal{L}_{I_1}(s) &= \mathbf{E}_{\Phi_1, g_i} \left[e^{-s \sum_{i \in \Phi_1 \setminus cp_0} P_t g_i r_i^{-\alpha}} \right] \\ &= \mathbf{E}_{\Phi_1} \left[\prod_{i \in \Phi_1 \setminus CP_0} (1 + s P_t r_i^{-\alpha})^{-1} \right] \\ &\stackrel{(a)}{\equiv} \exp \left\{ -2\pi\lambda_r \int_0^\infty \left(1 - \frac{1}{1 + s P_t x^{-\alpha}}\right) x dx \right\} \\ &= \exp \left\{ -2\pi\lambda_r \int_0^\infty \left(\frac{y^{2/\alpha-1}}{1 + (s P_t)^{-1} y} \right) \frac{1}{\alpha} dy \right\} \\ &\stackrel{(b)}{\equiv} \exp \left\{ -2\pi^2 \lambda_r (s P_t)^{2/\alpha} \csc(2\pi/\alpha) \alpha^{-1} \right\}, \end{aligned} \quad (5)$$

where step (a) follows from the probability generating functional of the PPP. Step (b) is obtained from ([21], eq. (3.194.4)). By substituting $s = P_t^{-1} r^\alpha \gamma_1$ into (5), we can obtain (6).

$$\mathcal{L}_{I_1}(P_t^{-1} r^\alpha \gamma_1) = \exp \left\{ -2\pi^2 \lambda_p \gamma_1^{2/\alpha} \csc(2\pi\alpha^{-1}) \alpha^{-1} r^2 \right\} \quad (6)$$

Therefore, the probability that the file f is successfully transmitted can be denoted as (7).

$$\begin{aligned} D_f &= \Pr \left(0 < r \leq z, \gamma_f \geq \gamma_0 \right) \\ &= \int_0^z \mathcal{L}_{I_1}(P_t^{-1} r^\alpha \gamma_0) f(r) d(r) \\ &\approx 2\pi \lambda_p q_f \int_0^z \exp \left\{ -2\pi^2 \lambda_p \gamma_1^{2/\alpha} \csc(2\pi\alpha^{-1}) \alpha^{-1} r_{0f}^2 \right\} \\ &\quad \cdot \exp \left(-\pi \lambda_p q_f r^2 \right) r d(r) \\ &= 2\pi \lambda_p q_f \int_0^z e^{-\varphi_f r^2} r dr \stackrel{(a)}{=} \frac{\pi q_f \lambda_p}{S_f} \left(\frac{1 - e^{-\varphi_f z^2}}{\varphi_f} \right), \end{aligned} \quad (7)$$

where step (a) is obtained from ([21], eq.(3.326.2)) and $\varphi_f = \pi q_f \lambda_p + 2\pi^2 \lambda_p \gamma_1^{2/\alpha} \cdot S \cdot \csc(2\pi\alpha^{-1}) \alpha^{-1}$.

Finally, we can get the approximate expression of data offloading rate by combining equations (2) and (7), as shown (8).

$$\rho(q_f, z) = \sum_{f=1}^F p_f S_f D_f = \sum_{f=1}^F p_f \pi q_f \lambda_p \left(\frac{1 - e^{-\varphi_f z^2}}{\varphi_f} \right) \quad (8)$$

Energy efficiency is defined as the ratio of the average number of successfully transmitted content bits per unit time to the total power consume required, expressed as

$$\eta_{EE} = \frac{\vartheta}{E} \quad (9)$$

where E represents the total power consume of successfully transmitted content, ϑ is the average number of successfully transmitted content bits per unit time, $\vartheta = R_t \cdot \rho(q_f, z)$, R_t indicates the transfer rate of content.

$$\eta_{EE} = R_t \cdot \rho(q_f, z) / \left(\frac{P_t}{\eta} + P_c \right) \quad (10)$$

Where P_c is the circuit power consumed at the cache-enabled users, and η is the power amplifier efficiency.

Then, combining (8) with (10), we can derive the expression of energy efficiency as

$$\eta_{EE} = \sum_{f=1}^F p_f \pi q_f \lambda_p R_t \left(\frac{1 - e^{-\varphi_f z^2}}{\varphi_f} \right) \left(\frac{\eta}{P_t + \eta P_c} \right) \quad (11)$$

Obviously, the impact of cache policy \mathbf{q} on energy efficiency is mainly reflected in the data offloading rate. Both the offloading ratio $\rho(q_f, z)$ and the total power consume E increase with the transmit power P_t . We also find that P_t and \mathbf{q} are closely coupled in the exponential term which is introduced by the D2D range limit. Each transmit power will correspond to an optimal cache policy, and each cache policy will correspond to an optimal transmit power. Therefore, it implies that there is a trade-off between the offload rate and cache policy to maximize energy efficiency. In the following section, we strive to study the joint impact of the cache policy and the transmit power of cache-enabled users on energy efficiency in detail and derive the joint solution.

4. Joint Optimization for Maximize Energy Efficiency

4.1. Optimal Caching Policy. In this section, our purpose is to obtain the optimal cache policy \mathbf{q} , which maximizes the data offloading ratio. The optimal cache policy problem can be formulated as

$$\begin{aligned} \mathbf{P1} : \max_{\mathbf{q}} \rho &= \sum_{f=1}^F p_f \pi q_f \lambda_p \left(\frac{1 - e^{-\varphi_f z^2}}{\varphi_f} \right) \\ \text{s.t.} \quad \sum_{f=1}^F q_f &\leq 1, q_f \geq 0 \end{aligned} \quad (12)$$

Proposition 1. *The objective function ρ is a concave function about q_f on a convex set, and the problem P1 is a convex programming problem.*

Proof: see the appendix.

The optimization problem P1 can be solved by the Lagrange algorithm. With the Lagrange multiplier of μ , we can get the Lagrange function of the problem P1; the expression is as follows

$$L(\mathbf{q}, \mu) = \sum_{f=1}^F p_f \pi q_f \lambda_p \left(\frac{1 - e^{-\varphi_f z^2}}{\varphi_f} \right) + \mu \left(1 - \sum_{f=1}^F q_f \right) \quad (13)$$

Where μ is the nonnegative Lagrange multiplier associated with constraint (12). We can obtain the optimal cache policy \mathbf{q} by solving the equation (14).

$$\frac{\partial L(\mathbf{q}, \mu)}{\partial q_f} = \frac{\partial \rho}{\partial q_f} - \mu = 0 \quad (14)$$

```

1: Initialize  $t:=1$ , Lagrangian variable  $\mu$ , step size  $\delta$ , error factor  $\varepsilon$ ; Calculate initial cache policy  $\mathbf{q}[1]$  based on (14) and (15)
2: while  $|\sum_{f=1}^F q_f[t] - 1| \leq \varepsilon$  do
3:   if  $\sum_{f=1}^F q_f[t] > 1$  then
4:      $\mu^{t+1} = \mu^t + \delta^t$ 
5:   else
6:      $\mu^{t+1} = \mu^t - \delta^t$ 
7:   end if
8:   Updating cache policy  $\mathbf{q}[t+1]$  based on (14) and (15)
9:   if  $\sum_{f=1}^F q_f[t+1] \cdot \sum_{f=1}^F q_f[t] < 0$  then
10:     $\delta^{t+1} = \delta^t/2$ 
11:  end if
12:  Updating cache policy  $\mathbf{q}[1] = \mathbf{q}[t+1]$ 
13:  Set  $t = t + 1$ 
14: end while
15: The final  $\mathbf{q}[1]$  is obtained.

```

ALGORITHM 1: The optimal caching policy algorithm.

Specifically, the first derivative of ρ on q_f can be expressed as

$$\frac{\partial \rho}{\partial q_f} = \frac{p_f \pi \lambda_p}{\varphi_f^2} \left[\varphi_f \left(1 - e^{-\varphi_f z^2} \right) + q_f \left(\pi \lambda_p + p_f 2\pi^3 \lambda_r \lambda_p \gamma_1^{2/\alpha} \csc(2\pi\alpha^{-1}) \alpha^{-1} z^2 e^{-\pi \lambda_p q_f z^2} \right) \right] \cdot \left((z^2 \varphi_f + 1) e^{-\varphi_f z^2} - 1 \right) \quad (15)$$

Obviously, the equation (15) is very complex, so it is difficult to find the analytical expression of q_f from the equation (14). It has been proved in the appendix that $\partial \rho / \partial q_f$ is a continuous monotonic function. Therefore, when the Lagrange multiplier μ is given, the solution of equation (14) can be found mathematically, such as the bisection method [22]. Hence, we can obtain the optimal caching policy from (16).

$$q_f^* = \left[q_f(\mu) \right]_0^1 \quad (16)$$

Where $q_f(\mu)$ means q_f is a function of μ and $[\cdot]_0^1 = \min(\max(\cdot, 0), 1)$. We can update μ to get the optimal cache policy \mathbf{q}^* . The details are described in Algorithm 1.

4.2. Optimal Transmit Power. In this section, we are working to optimize the power of any caching policy to maximize energy efficiency. We consider a large-scale fading effect and set the path loss factor to 2, which is also set in the simulation in Section 5. We can do some variable substitutions for (11) by setting $\psi_f = p_f \pi q_f \lambda_p \eta R_t$, $t_f = \varphi_f \eta P_c$, and $y = \varphi_f P_t / \theta$. For any given cache policy, the optimal transmit power problem can be described as

$$\mathbf{P2} : I(y) = \max_{P_t} \sum_{f=1}^F \psi_f \cdot \left(\frac{1 - e^{-y}}{y\theta + t} \right) \quad (17)$$

s.t. $0 < P_t \leq P_{\max}$.

Then, we taking the first-order derivative of $I(y)$ with regard to y can be obtained as

$$I'(y) = \sum_{f=1}^F \psi_f \cdot \frac{\varphi_f}{\theta} \cdot \frac{y\theta e^{-y} + t e^{-y} - \theta + \theta e^{-y}}{(y\theta + t)^2}. \quad (18)$$

Set $g(y) = y\theta e^{-y} + t e^{-y} - \theta + \theta e^{-y}$ and taking the first-order derivative of $g(y)$ with regard to y , we can get $g'(y) = -y' y \theta e^{-y} - t y' e^{-y} \leq 0$. So $g(y)$ is a monotone minus function. When $y \rightarrow 0$, $g(y) = t > 0$. When $y \rightarrow \infty$, $g(y) = -\theta < 0$. Therefore, $I(y)$ increases first and decreases with the increase of the transmit power P_t , and the maximum value is obtained at $g(y) = 0$. Thus, we can find the optimal transmit power P_t^* by searching from $\theta(e^y - y) = t + \theta$, which is obtained from $g(y) = 0$. We can get the first-order derivative of $\theta(e^y - y) = t + \theta$ with regard to P_t is not less than 0. So $\theta(e^y - y) = t + \theta$ is a continuous and monotonic function, and we can find the optimal transmit power P_t^* by bisection search.

4.3. Joint Optimization Algorithm of Transmit Power and Cache Policy. We discuss the optimal transmit power and the optimal caching policy algorithm to maximize the energy efficiency, as shown in Algorithm 2. T^{\max} is the maximum number of iterations. In the initialization phase of Algorithm 2, the D2D users' caching decisions $\mathbf{q}[1]$ is made based on the historical profiles of user requests and file attributes [23, 24], and $\eta_{EE}(P_t, \mathbf{q})$ is calculated by (16). Then, we use a fixed transmit power $P_t[t]$ to solve

1:**Initialize:** $T^{\max}, P_t[1], \mathbf{q}[1]$, calculate $\eta_{EE}(P_t, \mathbf{q})[1]$ based on (16), the maximum tolerance ε
2:**Repeat**{Loop}:
(a) Through $P_t[t]$ solve P1 and find the optimal cache policy $\mathbf{q}[t+1]$ with Algorithm 1.
(b) Through $\mathbf{q}[t+1]$ solve P2 and find the optimal transmit power $P_t[t+1]$ with bisection search.
(c) Update $\eta_{EE}(P_t, \mathbf{q})[t+1]$.
(d) Set $t=t+1$.
3:**Until** $|\eta_{EE}(P_t, \mathbf{q})[t] - \eta_{EE}(P_t, \mathbf{q})[t+1]| \leq \varepsilon$ or $t = T^{\max}$
4:Output η_{EE}, P_t and \mathbf{q}

ALGORITHM 2: Joint optimization algorithm.

the optimal cache strategy $\mathbf{q}[t+1]$ using Algorithm 1. Next, we use the optimal cache strategy $\mathbf{q}[t+1]$ to find the optimal transmit power $P_t[t+1]$ by bisection search. Note that the iterative process will continue until convergence P_t and \mathbf{q} or reach the maximum number of iterations t . The algorithm is summarized in Algorithm 2.

Computational complexity: The computational complexity of Algorithm 2 consists of two parts: the complexity from the transmission optimization and the cache optimization. The optimal cache strategy can also be obtained by Algorithm 2, with computation complexity $O(F \cdot t)$, where F is the number of files to be calculated per iteration and the number of iteration times for caching scheme is t . In each iteration, the optimal transmit power is obtained by a bisection search that converges to a solution with a certain error tolerance σ with computation complexity [23] $O(\log_2(P_{\max}/P_t))$, where P_{\max} is the maximum transmit power of D2D users. Let T_{\max} be the maximum number of iterations needed for iteration loops. Accordingly, the total computational complexity [24, 25] can be approximated as $T_{\max} \cdot (O(F \cdot t) + O(\log_2(P_{\max}/P_t)))$.

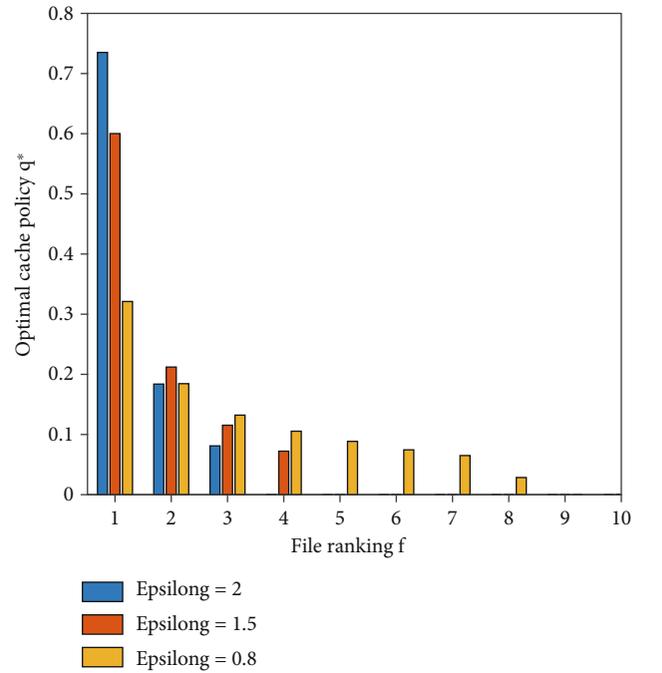
5. Performance Evaluation

In this section, we study the energy efficiency of the joint optimization scheme, compared with the traditional independent optimization scheme proposed in [13]. Then present some results to evaluate the performance of our proposed algorithm. Unless otherwise specified, the detailed simulation parameters are listed in Table 1.

Firstly, we tested the optimal cache strategy with different Zipf parameters at a given transmit power $P_t = -10\text{dBW}$ as shown in Figure 2, i.e., an ε value from 2 to 1.5 and 0.8. It can be seen from Figure 2 that when ε is certain, the cache probability of the file is proportional to the popularity ranking of the file. With the decrease of file popularity, the probability of file being cached tends to decrease, and the cache probability of the last few files ranking in popularity is less than one percent. For example, when $\varepsilon = 0.8$, almost all files are cached. When $\varepsilon = 2$, only the top three files are cached. This is because caching the most popular files can improve the data offloading rate. This is consistent with the fact that only a few files are repeatedly requested in a certain period of time in the actual network [15]. For the same reason, due to the limited storage capacity of cache-enabled users, the last few files ranked in popularity may not be cached.

TABLE 1: Simulation parameters.

Parameters	Value
Intensity of cache-enabled users λ_p	500 nodes/km ²
Intensity of requesters λ_r	500 nodes/km ²
D2D bandwidth W	20MHz
Path loss exponent α	3
SINR threshold γ_0	-20dB
D2D establishment threshold θ	0dBm
Circuit power consumption P_c	10mW
Power amplifier efficiency η	0.8
Zipf exponent δ	2
The number of the contents F	10 files

FIGURE 2: The impact of different Zipf parameters ε .

In Figure 3, we show the impact of different transmit power on the optimal cache policy under given Zipf parameters $\varepsilon = 2$. It can be seen from Figure 3 that when the transmit power P_t is given, the file popularity ranking is still the main

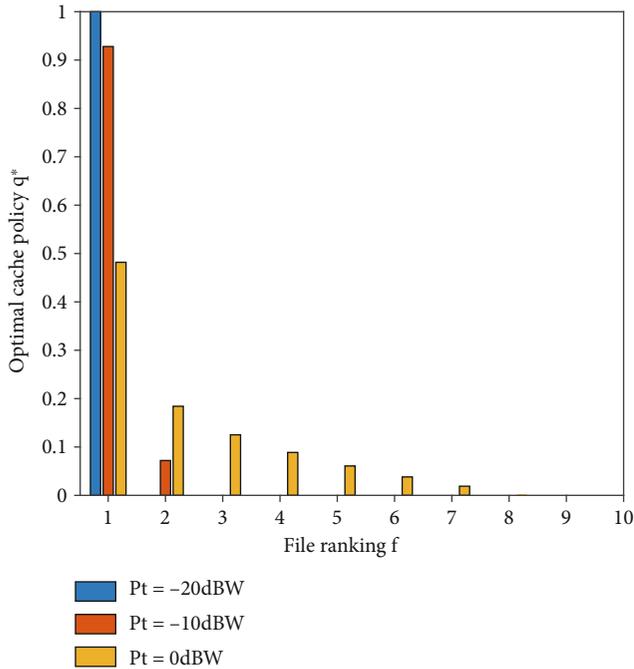


FIGURE 3: The impact of different transmit power P_t .

factor affecting the optimal cache strategy. With the increase of transmit power, the number of cache files increases gradually. When transmitting power $P_t = -20\text{dBW}$, only the files with the highest popularity are cached. When transferring power $P_t = 0\text{dBW}$, almost all files are cached in the library. The main reason can be analyzed from the equation $z = (P_t/\theta)^{1/\alpha}$, when the threshold θ is established for D2D communication, the communication radius z is proportional to the transmit power P_t . Therefore, when the transmit power is small, the correlation distance r_{of} between users must also be reduced to satisfy the D2D establishment condition. We must increase the location probability density $q_f \lambda_p$ of the cache-enabled users to satisfy the small correlation distance. Thus, we need to increase the cache probability of file f .

In Figure 4, we show the simulated offloading rate for the different collaboration distance. We compared the proposed cache policy with uniform caching policy (i.e., all users select a file from the content library uniformly, with legend “Uniform-baseline”) as a caching baseline and popularity based on caching policy (legend “Optimal-Caching-Popularity”). Clearly, the offloading rate of the three schemes has been remarkably improved upon increasing of the collaboration distance. This is because as the association distance increases, each user’s neighbors will gradually increase. If the association distance is large enough, all files in the cached content library can be found in adjacent users. By contrast, the optimized caching policy can offload more traffic than the others as expected. However, with the increase of the association distance, the transmit power consumption will also increase. Thus, we can maximize the network energy efficiency by controlling the association distance r , where r should satisfy $0 < r \leq (P_t/\theta)^{1/\alpha}$.

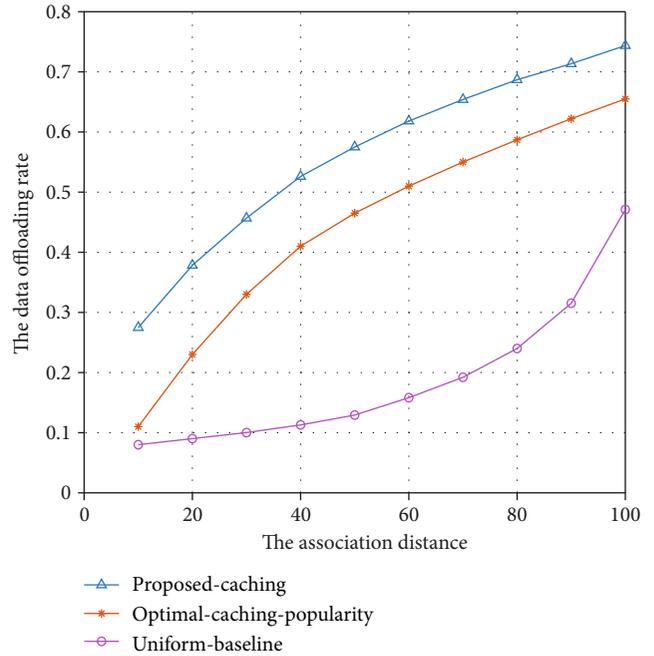


FIGURE 4: Offloading opportunity versus the association distance.

In Figure 5, we predetermine transmit power $P_t = -10\text{dBW}$ and Zipf parameters $\epsilon = 2$ and compare the energy efficiency of the proposed cache policy with the traditional cache policy proposed in [13]. Clearly, the two schemes tend to be stable rapidly with the increase of the number of iterations. It is obvious that the energy efficiency of our proposed optimal caching policy is better than the energy efficiency of the caching policy proposed in [13]. The reason could be summed up in two aspects. On the one hand, we derive the data offloading ratio in closed form based on consider the probability of successful sensing and transmission. However, the most current research is based on success sensing as data offloading. On the other hand, we jointly analyze the impact of the SINR threshold γ_0 and D2D communication threshold θ on the network energy efficiency, while only the SINR threshold γ_0 was considered in [13].

We present the simulated results of the energy efficiency η_{EE} with the given cache policy q_2 (uniform caching policy) in Figure 6, versus its analytical value with respect to P_t^* . The red line shows the process of using bisection search to find the optimal transmit power P_t^* to achieve the maximum energy efficiency η_{EE} . It is observed through numerical results that the approximation on P_t^* given by bisection search is valid. Specifically, the cache policy q_2 obtains the maximum energy efficiency when transmitting power $P_t^* = -16.53\text{dBW}$.

We use different Zipf parameters and the number of files to evaluate the performance of the proposed energy efficiency joint optimal scheme in Figure 7. To make the comparison fair, the overall parameters are set as the same. Specifically, the initial transmit power and cache policy is set to $P_t = -10\text{dBW}$ and q_2 , respectively. From the figure, it is obvious that Algorithm 2 needs at most

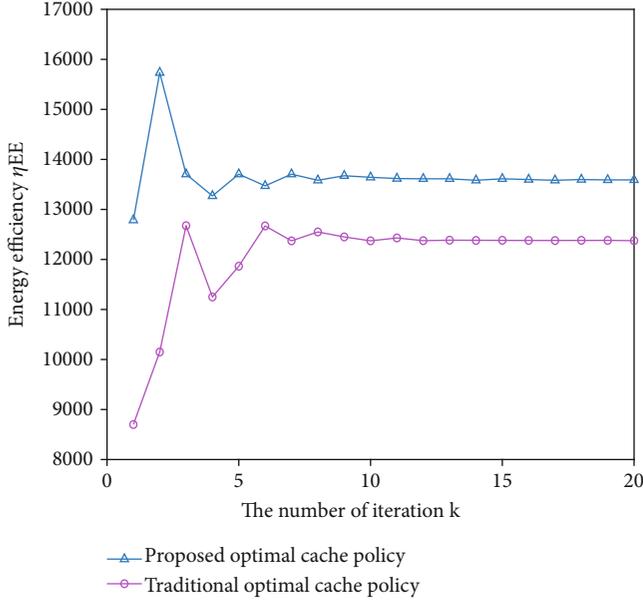


FIGURE 5: The energy efficiency performance.

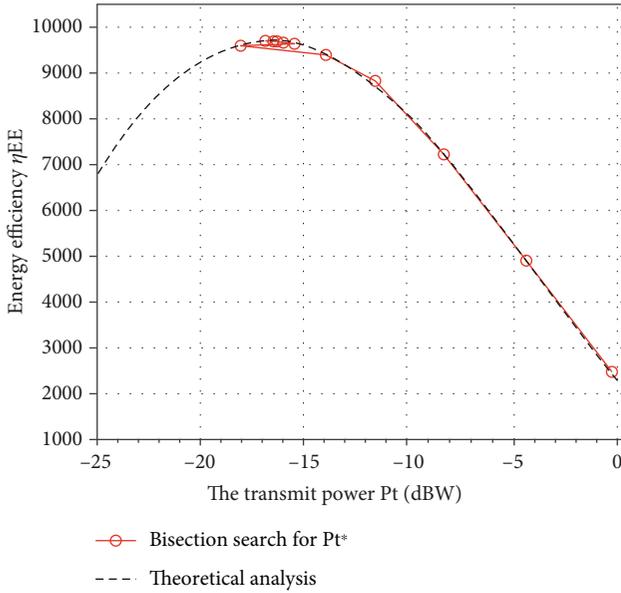


FIGURE 6: The optimal transmit power with the bi-section search method.

three iterations to achieve the maximum energy efficiency. When the Zipf parameters ε increases, the optimal energy efficiency η_{EE} also increases. However, when the number of files increases, the optimal energy efficiency will decrease. The reason could be, (1) when the smaller transmit power and the larger Zipf parameters, the cache probability of files with high popularity will be larger to obtain the maximum throughput, and (2) when $F = 50$, the probability of the first ranked file being requested is lower than $F = 10$.

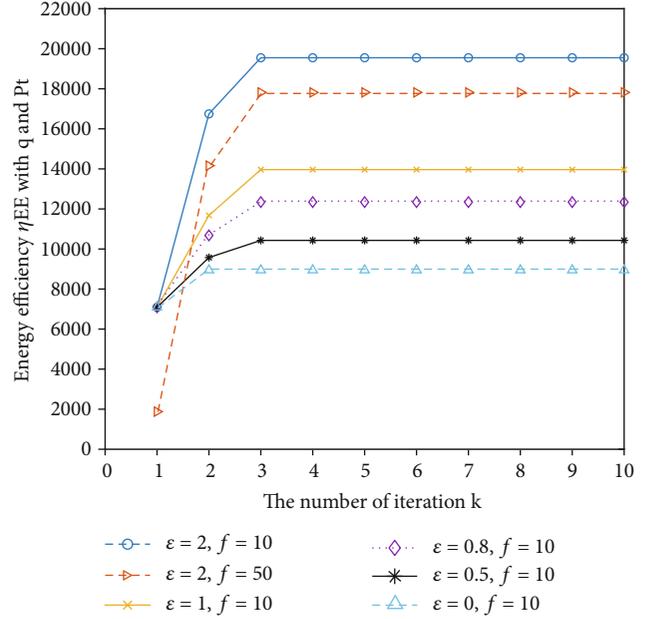


FIGURE 7: The energy efficiency performance for joint algorithm.

TABLE 2: Comparison of the energy efficiency with different schemes.

Problem	Energy efficiency (bits/s/W)
P1	$\eta_{EE}(\mathbf{q}^*, -10\text{dBW}) = 9561.93$
P2	$\eta_{EE}(\mathbf{q}_2, P_t^*) = 8172.8$
Joint optimization of P1 and P2	$\eta_{EE}(\mathbf{q}^*, P_t^*) = 19550.3$

In Table 2, we give the exact energy efficiency values for different problems to further illustrate the effectiveness of the joint optimization scheme. To make the comparison fair, $\varepsilon = 2$ and $F = 10$ are set as the same. The problems P1 and P2 are independent optimization schemes, respectively. It can be clearly seen from Table 2 that our proposed joint optimization scheme is better than the independent optimization scheme. Whether the cache policy or the transmission power is optimized separately, the joint optimization scheme in energy efficiency is more than twice of them.

Finally, we compared the proposed energy efficiency alternating optimization scheme with other joint optimization schemes in Figure 8. With legend “Two-step joint optimization scheme” [16] (i.e., optimize the active cache strategy and the transmit power of each user separately, and then combine the cache strategy and transmit power to control the energy consumption of content transmission) and “Maximum power transmission scheme” (i.e., maximize the hit rate to optimize the cache strategy, each user transmits content at the maximum transmit power). As can be seen from Figure 8, with the increase of θ ,

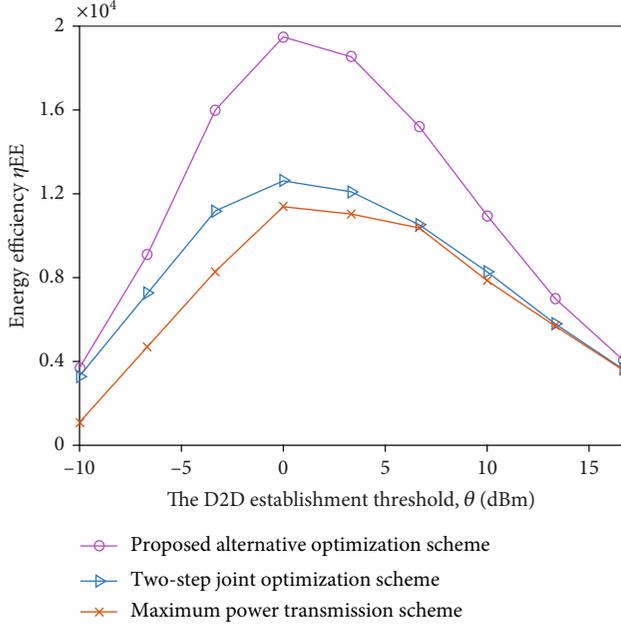


FIGURE 8: Comparison of energy efficiency under three strategies.

the energy efficiency joint optimization scheme proposed by us is superior to the other two joint optimization schemes. In fact, the influence of θ is mainly reflected in the D2D communication range, $z = (P_t/\theta)^{1/\alpha}$, and it acts as a bridge between the cache strategy and transmit power. Therefore, the proposed joint optimization scheme can alternately optimize the cache strategy and transmit power to maximize energy efficiency. Obviously when $\theta = 0$ dBm, there is no D2D range limit anymore. When θ is large, D2D communications establishment is difficult. When θ is relatively small, the effect of z is diminishing. Therefore, the threshold is assumed as preset parameter in this study.

6. Conclusions

Due to the emergence of large-scale video services and low-latency transmission requirement of users, the enable-D2D caching technology will play a central role in the future wireless networks. According to the different energy efficiency optimization algorithm, a novel energy efficiency optimization scheme has been proposed through a joint optimization cache policy and transmit power. Based on the stochastic geometry, we derive the data offloading rate in closed form, which jointly considers the effects of success sensing probability and success transmission probability. In the success transmission probability, we take the data transmission rate and the D2D establishment constraint into account. Then, we formulate an iterative optimization algorithm based on independent optimization (for a fixed transmit power, optimize cache policy or for a fixed cache policy, optimize transmit power). We prove that two sub-problems can converge to the optimal solution. The optimi-

zation of energy efficiency has been developed by leveraging the joint optimization algorithm. Numerical results show that the energy efficiency of our proposed joint optimization scheme is more than twice that of any independent optimization scheme (cache policy or transmit power). Compared with other joint optimization algorithms, it also shows obvious advantages. At present, most researches are based on the same cache capacity of terminal devices, but the remaining cache capacity of each device is different, which is worth for us to conduct further study.

Appendix

A.1. Proof of Proposition

We set $A = \varphi_f/q_f = \pi\lambda_p + 2\pi^2\lambda_r\lambda_p\gamma_1^{2/\alpha} \csc(2\pi\alpha^{-1})\alpha^{-1}((1 - e^{-\pi\lambda_p q_f z^2})/q_f)$, the equation (9) can be expressed as

$$\rho = \sum_{f=1}^F p_f \pi \lambda_p \left(\frac{1 - e^{-\varphi_f z^2}}{A} \right) \quad (\text{A.1})$$

We can get the first derivative of ρ with respect to q_f ($0 \leq q_f \leq 1$), as shown in (A.2).

$$\rho' = \frac{\partial \rho(\mathbf{q})}{\partial q_f} = p_f \pi \lambda_p \frac{z^2 A \varphi_f' e^{-\varphi_f z^2} - A' (1 - e^{-\varphi_f z^2})}{A^2} \quad (\text{A.2})$$

Where $\varphi_f' = \pi\lambda_p + 2\pi^3\lambda_r\lambda_p\gamma_1^{2/\alpha} p_f \csc(2\pi\alpha^{-1})\alpha^{-1} z^2 e^{-\pi\lambda_p q_f z^2}$, A'_{q_f} is expressed as the first derivative of A with respect to q_f .

$$\begin{aligned} A'_{q_f} &= \frac{dA(q_f)}{dq_f} \\ &= 2\pi^2\lambda_r\gamma_1^{2/\alpha} p_f \csc(2\pi\alpha^{-1})\alpha^{-1} \frac{(\pi\lambda_p q_f z^2 + 1)e^{-\pi\lambda_p q_f z^2} - 1}{q_f^2} \end{aligned} \quad (\text{A.3})$$

For simplicity, we set $v(q_f) = (\pi\lambda_p q_f z^2 + 1)e^{-\pi\lambda_p q_f z^2} - 1$. By deriving the first-order derivative of $v(q_f)$, we find that $v'(q_f) = -\pi^2\lambda_p^2 q_f z^4 e^{-\pi\lambda_p q_f z^2} \leq 0$, which indicates that $v(q_f)$ is an decreasing function. As a consequence, we can obtain $v(q_f) \leq v(0) = 0$. It is clear that the first derivative of A with respect to q_f is not greater than zero, $A'_{q_f} = (dA(q_f)/dq_f) \leq 0$. Because of $z^2 A \varphi_f' e^{-\varphi_f z^2} \geq 0$ and $1 - e^{-\varphi_f z^2} \geq 0$, we can get the first derivative of the objective problem, $\rho' \geq 0$.

Then, we can further obtain the second derivative of ρ with respect to q_f , as shown in (A.4).

$$\begin{aligned}
\rho'' &= \frac{\partial^2 \rho}{\partial q_f^2} = \frac{\partial \rho'}{\partial q_f} \\
&= \frac{p_f \pi \lambda_p}{A^4} \left[\begin{aligned} & z^2 A^2 A'_{q_f} \phi'_f e^{-\phi_f z^2} + z^2 A^3 \phi''_f e^{-\phi_f z^2} - z^4 A^3 \phi'^2_f e^{-\phi_f z^2} - A''_{q_f} A^2 (1 - e^{-\phi_f z^2}) \\ & - z^2 A'_{q_f} A^2 \phi'_f e^{-\phi_f z^2} - 2AA'_{q_f} (z^2 A \phi'_f e^{-\phi_f z^2} - A'_{q_f} (1 - e^{-\phi_f z^2})) \end{aligned} \right] \quad (\text{A.4}) \\
&= \frac{p_f \pi \lambda_p}{A^3} \left[z^2 A^2 \phi''_f e^{-\phi_f z^2} - z^4 A^2 \phi'^2_f e^{-\phi_f z^2} - A''_{q_f} A (1 - e^{-\phi_f z^2}) - 2A'_{q_f} z^2 A \phi'_f e^{-\phi_f z^2} + 2A'^2_{q_f} (1 - e^{-\phi_f z^2}) \right] \\
&= \frac{p_f \pi \lambda_p}{A^3 e^{\phi_f z^2}} \left[(-z^2 A \phi'_f - 2A'_{q_f}) z^2 \phi'_f A + \phi''_f z^2 A^2 + (2A'^2_{q_f} - A''_{q_f} A) (e^{\phi_f z^2} - 1) \right],
\end{aligned}$$

where $A''_{q_f} = 2\pi^2 \lambda_r \gamma_1^{2/\alpha} p_f \csc(2\pi\alpha^{-1}) \alpha^{-1} ((-\pi\lambda_p z^2 q_f - 2)\pi\lambda_p q_f z^2 e^{-\pi\lambda_p q_f z^2} + 2(1 - e^{-\pi\lambda_p q_f z^2})/q_f^3)$ expressed as the second-order derivative A with respect to q_f . We set $h_1(q_f) = (-\pi\lambda_p z^2 q_f - 2)\pi\lambda_p q_f z^2 + 2(e^{\pi\lambda_p q_f z^2} - 1)$, the first derivative $h_1'(q_f)$ is readily follows from $h_1(q_f)$.

$$\begin{aligned}
h_1'(q_f) &= -\pi\lambda_p z^2 (\pi\lambda_p q_f z^2) - \pi\lambda_p z^2 (\pi\lambda_p q_f z^2 + 2) \\
&\quad + 2\pi\lambda_p z^2 e^{\pi\lambda_p q_f z^2} \\
&\geq -\pi\lambda_p z^2 (\pi\lambda_p q_f z^2) - \pi\lambda_p z^2 (\pi\lambda_p q_f z^2 + 2) \\
&\quad + 2\pi\lambda_p z^2 (1 + \pi\lambda_p q_f z^2) = 0. \quad (\text{A.5})
\end{aligned}$$

Therefore, $h_1(q_f)$ is an increasing function of q_f , when $q_f \rightarrow 0$, $h_1(q_f) \rightarrow 0$. Further, we can get $A''_{q_f} \geq 0$. So the second-order derivative of ρ with respect to q_f , i.e., ρ'' , we have

$$\begin{aligned}
\rho'' &= \frac{p_f \pi \lambda_p}{A^3 e^{\phi_f z^2}} \left[(-z^2 A \phi'_f - 2A'_{q_f}) z^2 \phi'_f A + \phi''_f z^2 A^2 \right. \\
&\quad \left. + (2A'^2_{q_f} - A''_{q_f} A) (e^{\phi_f z^2} - 1) \right] \\
&\stackrel{(a)}{\leq} \frac{p_f \pi \lambda_p}{A^3 e^{\phi_f z^2}} \left[(-z^2 A \phi'_f - 2A'_{q_f}) z^2 \phi'_f A + \phi''_f z^2 A^2 \right. \\
&\quad \left. + 2A'^2_{q_f} (e^{\phi_f z^2} - 1) \right] \quad (\text{A.6}) \\
&\stackrel{(b)}{\approx} \frac{p_f \pi \lambda_p}{A^3 e^{\phi_f z^2}} \left[-z^4 \phi'^2_f A^2 + \phi''_f z^2 A^2 \right] \leq 0,
\end{aligned}$$

where step (a) is obtained from (A.5), and step (b) is obtained from A'_{q_f} approximated with ([21], eq. (1.211.1)). Because $\phi''_f = -2\pi^4 \lambda_r \lambda_p^2 \gamma_1^{2/\alpha} p_f \csc(2\pi\alpha^{-1}) \alpha^{-1} z^4 e^{-\pi\lambda_p q_f z^2} \leq 0$, we know that $\rho'' \leq 0$ always holds. Therefore, the objective function ρ is a concave function about q_f on a convex set;

then, the problem is a convex programming problem. Thus, Proposition 1 is proved.

Data Availability

The data used to support the findings of this study are available from the corresponding author upon request (Email: Laughing_wwg@163.com). The lemma or theorem cited in this article can be publicly obtained according to the reference.

Conflicts of Interest

The authors declare that there are no conflicts of interest regarding the publication of this paper.

Acknowledgments

This study was funded by the National Natural Science Foundation of China under Grant no. 61571364.

References

- [1] Cisco, *Cisco visual networking index: global mobile data traffic forecast update, 2016-2021*, Cisco Syst. Inc., San Jose, CA, USA, 2017, collateral/service-provider/visual-networking-index-vni/mobile-white-paper-c11-520862.pdf.
- [2] M. Abana, M. Peng, Z. Zhao, and L. A. Olawoyin, "Coverage and rate analysis in heterogeneous cloud radio access networks with device-to-device communication," *IEEE Access*, vol. 4, pp. 2357-2370, 2016.
- [3] T. Zhang, X. Fang, Y. Liu, G. Y. Li, and W. Xu, "D2D-enabled mobile user edge caching: a multi-winner auction approach," *IEEE Transactions on Vehicular Technology*, vol. 68, no. 12, pp. 12314-12328, 2019.
- [4] P. Gandotra and R. K. Jha, "Device-to-device communication in cellular networks: a survey," *Journal of Network & Computer Applications*, vol. 71, pp. 99-117, 2016.
- [5] P. Liu, Y. Li, W. Cheng, W. Zhang, and X. Gao, "Energy-efficient power allocation for millimeter wave beamspace MIMO-NOMA systems," *IEEE Access*, vol. 7, pp. 114582-114592, 2019.

- [6] L. Wang, H. Wu, and Z. Han, "Wireless distributed storage in socially enabled D2D communications," *IEEE Access*, vol. 4, pp. 1971–1984, 2016.
- [7] J. Jiang, S. Zhang, B. Li, and B. Li, "Maximized cellular traffic offloading via device-to-device content sharing," *IEEE Journal on Selected Areas in Communications*, vol. 34, no. 1, pp. 82–91, 2016.
- [8] D. Malak, M. Shalash, and J. G. Andrews, "Optimizing content caching to maximize the density of successful receptions in device-to-device networking," *IEEE Transactions on Communications*, vol. 64, no. 10, pp. 4365–4380, 2016.
- [9] C. Jarray and A. Giovanidis, "Successful file transmission in mobile D2D networks with caches," *Computer Networks*, vol. 147, no. 24, pp. 162–179, 2018.
- [10] B. Chen and C. Yang, "Caching policy for cache-enabled D2D communications by learning user preference," *IEEE Transactions on Communications*, vol. 66, no. 12, pp. 6586–6601, 2018.
- [11] J. Rao, F. Hao, C. Yang, Z. Chen, and B. Xia, "Optimal caching placement for D2D assisted wireless caching networks," in *2016 IEEE International Conference on Communications (ICC)*, Kuala Lumpur, Malaysia, May 2016.
- [12] A. Mete and S. Moharir, "Caching policies for D2D-assisted content delivery systems," in *Proceedings of the 2018 on Technologies for the Wireless Edge Workshop - WirelessEdge '18*, pp. 3–7, New Delhi, India, 2018.
- [13] B. Chen, C. Yang, and Z. Xiong, "Optimal caching and scheduling for cache-enabled D2D communications," *IEEE Communications Letters*, vol. 21, no. 5, pp. 1155–1158, 2017.
- [14] Y. Cui, D. Jiang, and Y. Wu, "Analysis and optimization of caching and multicasting in large-scale cache-enabled heterogeneous wireless networks," *IEEE Transactions on Wireless Communications*, vol. 16, no. 1, pp. 250–264, 2017.
- [15] Y. Long, D. Wu, Y. Cai, and J. Qu, "Joint cache policy and transmit power for cache-enabled D2D networks," *IET Communications*, vol. 11, no. 16, pp. 2498–2506, 2017.
- [16] A. C. Kazez and T. Girici, "Clustering-based device-to-device cache placement," *Ad Hoc Networks*, vol. 84, pp. 170–177, 2019.
- [17] B. Chen, C. Yang, and A. F. Molisch, "Cache-enabled device-to-device communications: offloading gain and energy cost," *IEEE Transactions on Wireless Communications*, vol. 16, no. 7, pp. 4519–4536, 2017.
- [18] Y. Y. Chen, L. Wang, R. Q. Ma, B. L. Jiao, and L. Hanzo, "Cooperative full duplex content sensing and delivery improves the offloading probability of D2D caching," *IEEE Access*, vol. 7, no. 7, pp. 29076–29084, 2019.
- [19] Z. Chen, P. Nikolaos, and K. Marios, "Probabilistic caching in wireless D2D networks: cache hit optimal versus throughput optimal," *IEEE Communications Letters*, vol. 21, no. 3, pp. 584–587, 2017.
- [20] F. Baccelli and B. Blaszczyszyn, "Stochastic geometry and wireless networks: volume I theory," *Foundations and Trends® in Networking*, vol. 3, no. 3-4, pp. 249–449, 2009.
- [21] I. S. Gradshteyn and I. M. Ryzhik, *Table of Integrals, Series, and Products*, Academic, New York, NY, USA, 7th edition, 2007.
- [22] L. Yanshan, W. U. Dan, C. Yueming, M. Wang, and J. Guo, "Cache placement optimization scheme in D2D networks with heterogeneous cache capacity," *Journal of Computer Applications*, vol. 38, no. 5, pp. 1453–1457, 2018.
- [23] G. R. Wood, "The bisection method in higher dimensions," *Mathematical Programming*, vol. 55, no. 1-3, pp. 319–337, 1992.
- [24] L. Xiang, D. W. K. Ng, R. Schober, and V. W. S. Wong, "Cache-enabled physical layer security for video streaming in backhaul-limited cellular networks," *IEEE Transactions on Wireless Communications*, vol. 17, no. 2, pp. 736–751, 2018.
- [25] L. Xiang, D. W. K. Ng, R. Schober, and V. W. S. Wong, "Secure video streaming in heterogeneous small cell networks with untrusted cache helpers," *IEEE Transactions on Wireless Communications*, vol. 17, no. 4, pp. 2645–2661, 2018.