

Received April 2, 2018, accepted April 28, 2018, date of publication May 2, 2018, date of current version June 5, 2018.

Digital Object Identifier 10.1109/ACCESS.2018.2832442

Towards Energy Efficient Device-to-Device Content Dissemination in Cellular Networks

YANLI XU¹, (Member, IEEE), SHENGMING JIANG¹, (Senior Member, IEEE),
AND JINSONG WU², (Senior Member, IEEE)

¹College of Information Engineering, Shanghai Maritime University, Shanghai 200135, China

²Department of Electrical Engineering, Universidad de Chile, Santiago 1058, Chile

Corresponding author: Yanli Xu (ylxu@shmtu.edu.cn)

This work was supported in part by the National Natural Science Foundation of China under Grant 61601283, Grant 61472237, and Grant 61271283, and in part by the Chile CONICYT FONDECYT Project under Grant 1181809.

ABSTRACT Considering that the future cellular networks are more content-centric, and content sharing among user equipments (UEs) is more universal, device-to-device (D2D) communication techniques become attractive for their applications to content delivery due to the characteristic of proximal transmissions without needs of high-rate backhaul. However, D2D content delivery may not be energy efficient due to shortage of useful contents for neighbor UEs when content caching and delivery are not well designed. In this paper, we propose an architecture using D2D multicast for content delivery in cellular networks and a practical methodology, which can indicate the time for delivery, the content to be disseminated, and the UEs to receive the delivery, to achieve more energy efficient content delivery for future green cellular networks.

INDEX TERMS D2D communications, content delivery, energy efficiency, green cellular networks.

I. INTRODUCTION

Recently, data traffic grows rapidly due to the development of mobile devices and wireless communication technologies. Mobile data traffic is expected to increase by a factor of 40 over the next five years [1]. Unlike the previous growth of mobile traffic, which mainly depended on the demands of connection-centric communications, such as phone calls and text messages, the recent and future growth would come from the explosion of content-centric communications such as video streaming and content sharing. However, content delivery aggravates the burden of the core network and is not energy efficient due to frequent content retrieval from base stations (BSs).

Towards green fifth-generation (5G) cellular networks, some researchers have studied the utilization of small cell BSs with large storage units (such as hard-disks, solid-state drives) for content delivery to reduce the resource consumption by bringing content closer to users [2]–[5]. For example, system performance metrics such as outage probability and content delivery rate were analyzed in [2]. Caching strategies about how to store data at small cell BSs were studied in [3]–[5]. However, such small-cell architecture may cause bottlenecks of content delivery due to the requirement of high-rate backhaul links as indicated in [6]. In device-to-device (D2D) communications, data are directly transmitted

between user equipments (UEs) without the traverse of BSs and the transmission distance is usually short [7], [8], which will improve the resource efficiency. In addition, there is a large amount of content reuse among UEs, especially for popular contents. Thereby, D2D communications for content delivery have attracted great attention recently [9]–[11].

In [9], D2D pairs with common interests construct a cluster for D2D content sharing, where the number of requests from D2D UEs that could be served in a cluster and the optimized global density of served requests were analyzed. The work [10] analyzed the optimal cluster size for random and fixed caching via assuming that users cache popular video files and send them after receiving requests from other users in the same cluster. A tractable closed-form equation to find when redundant caching should be used in order to minimize the expected energy consumption was derived in [11]. However, these research works were based on a simplistic system model with the assumptions which did not accurately consider the effects of D2D interference or fading. Thus, some works analyzed the D2D content delivery problem based on stochastic geometry theory [12] with the Poisson Point Process (PPP) UE distribution [13]–[15]. For example, caching policy was studied for D2D communications in millimeter-wave (mmWave) networks and the offloading gain of the proposed policy was analytically derived based on

stochastic geometry theory in [13]. The work [14] optimized the D2D caching to maximize the total success probability of content delivery. The work [15] analyzed sharing content failure rate in a certain range and proposed an algorithm to determine the probability that content is cached in a server to minimize the failure probability. Particularly in [16], based on social relationships among mobile users and a heterogeneous but overlapping interests of users model, a graph based social-aware algorithm was proposed, in which cellular links and D2D links were established according to social ties and social contributions of users for efficient multi-file delivery. Through integrating network coding with D2D communication, [17] demonstrated the benefits of network coding for D2D content delivery and that the performance can be improved even by simple network coding.

Note that the architecture of D2D content delivery was based on request-and-response pattern in most existing proposals. As advocated in [18], the multicast gain can play an important role in proactive caching for the 5G cellular networks. Realizing the advantage of multicast, UE multicast were used for content recovery of UEs in the same multicast group as an assistant method of cellular multicast in [19]–[21]. For content delivery based on D2D multicast, a model based on PPP was proposed and the system performance was analyzed in [22]. Considering that proximal UEs have clustering distribution and there may be multiple potential transmitters in a cluster, a multicast model based on Poisson cluster process was investigated for D2D cache and delivery in [23]. To distinguish appropriate conditions of employing unicast and multicast, performance metrics under different transmission environment parameters are investigated in [24]. In addition to the content delivery pattern, the energy efficiency of D2D content delivery needs to be considered for the sake of green 5G [25]. Towards energy efficient content delivery, centralized and distributed Medium Access Control (MAC) strategies are studied in [26] and [27] to achieve balance between conserving energy and content delivery. In [28], social awareness was exploited in order to improve the energy efficiency of D2D cooperative communication. Different from previous works, we target to design energy-efficient D2D multicast from the perspective of network control in this paper, such as multicast range, multicast strategy and mode selection. Firstly, we propose a multicast content delivery architecture for content-centric cellular networks to improve the energy efficiency of video transmissions, based on which a transmission zone is defined and optimized to minimize the expected energy consumption for a required file per UE. In addition, a new caching and delivery strategy is proposed to maximize the number of UEs that can receive their required files. Furthermore, an evaluation metric is proposed to determine when to use D2D delivery. The contributions of this work are listed as follows.

- A novel architecture is proposed for D2D content delivery, by which a D2D UE multicasts cached files following an optimized distribution and other UEs are scheduled to receive it.

- A transmission zone is defined for the content delivery, which gives a constraint to UEs receiving the delivery. This zone is optimized to minimize the expected energy consumption of a required file per UE.
- A new caching and delivery strategy is proposed to maximize the number of UEs that can receive their required files. It can be used for determining which file is cached or disseminated.
- The energy efficiency problem is analyzed for D2D content delivery in comparison with that of cellular content delivery. Furthermore, an evaluation metric is proposed to determine which content delivery mode is better.

The remaining part of the paper is organized as follows. The system model is described and the D2D delivery problem is formulated in Section II. In Section III, the transmission energy of content delivery is analyzed and optimized, and the delivery distribution is modeled as an optimal problem and the solution is provided. Then the energy efficiency problem of D2D delivery is discussed and compared with that of BS content delivery in Section IV. Simulation results and related discussion are presented in Section V. Finally, Section VI concludes this paper.

II. SYSTEM MODEL

A. NETWORK AND CHANNEL MODELS

For the sake of clarity, Table 1 lists main symbols used in the paper.

1) NETWORK MODEL

We consider a cellular network where UEs are randomly located. UEs transmit data under either cellular mode or D2D mode. The transmission mode of a UE is selected by the BS based on a given mode selection criterion. Two UEs are neighbors if the Euclidean distance between them is smaller than a distance threshold and may communicate in the D2D mode. This maximally allowable distance for D2D communication is determined by the transmission power and the used power control scheme of transmitters. There are two ways for the in-band coexistence of different modes, i.e., D2D UEs reusing resource of cellular UEs (underlying coexistence) and using the orthogonal resource of cellular UEs (overlying coexistence), respectively. Both of coexistent ways have been widely studied in prior works [29]–[31]. Compared to orthogonal resource allocation, a resource reuse scheme may improve spectrum efficiency while leads to severer interference and higher implementation complexity. In this paper, we focus on the efficiency of orthogonal resource allocation case and readers can refer to prior works on the spectrum efficiency of different coexistence ways such as [32]–[34]. For simultaneous D2D transmitters, uplink resources are spatially reused to save bandwidth since interference among them can be well controlled due to the short transmission range of these communication links. Similar to [22], stochastic geometry is used for the modeling of UE distributions for D2D multicast. Particularly, the positions of multicast D2D

TABLE 1. Definition of used symbols.

Symbol	Meaning
λ_t	UE transmitter density
λ_b	BS density
λ_r	Receiver density
E_i	Transmission power of transmitter i
h_{ij}	Coefficient of channel between nodes i and j
n_0	Noise at receiver
d_{ij}	Distance between nodes i and j
H_{ij}	Fast fading power of the channel between nodes i and j
N_0	Noise power
α	Path loss factor
I_j	Interference at the receiver j
γ_j	SINR at the receiver j
γ_{th}	SINR threshold for correctly decoding data
$f_s(n)$	Transmission probability for file n at a transmitter
$f_r(n)$	Required probability for file n at a UE
β	Zipf exponent
N	Number of files
\mathcal{T}_i	Targeted sharing zone (TSZ) of transmitter i
d_0	Radius of TSZ
k	Path loss compensatory exponent
η_0	Targeted receiving power
$M_j(n)$	Conditional probability that receiver j requires file n when the transmitter transmits n
ω_j	Probability that receiver j correctly receives its required file
$f_c(n)$	Caching probability for file n at a UE
d_σ	Infinite small area
N_r	Number of receivers that correctly receive the required file in a TSZ
$N_r^{u.b.}$	The upper bound of N_r
ρ	Transmission efficiency coefficient
E_i^*	Optimal transmission power of UE i
E_{max}	Maximal transmission power of a UE
d_0^*	Optimal radius of TSZ
ζ	Caching and delivery efficiency coefficient
δ	Lagrange multiplier
$\xi_D(j)$	The probability that a UE j receives a required file under the D2D delivery mode
ϕ	Empty set
ξ_D	The expectation of a UE receiving a required file under the D2D delivery mode
$\xi_C(j)$	The probability that a UE j receives a required file under the BS delivery mode
$f(r)$	Probability distribution function (pdf) of distance between UE and its served BS
ξ_C	The expectation of a UE receiving a required file under the BS delivery mode
ε_e	D2D energy efficiency coefficient

transmitters form an independent PPP Ψ_t with density λ_t . For each D2D transmitter, the positions of its intended receivers are independently uniformly scattered around the transmitter in the multicast range which is centered at the transmitter with radius d_0 . Thereby, the receiver density of a multicast transmitter is a constant independent to d_0 , which is denoted as λ_r here.

2) CHANNEL MODEL

We assume that UEs are half-duplex constraint. All UEs are synchronized with BSs. For a communication link $i \rightarrow j$,

the signal received at the node j can be expressed as

$$y_j = \sqrt{E_i} h_{ij} x_i + n_0, \tag{1}$$

where E_i denotes the transmission power of node i , and n_0 is the thermal noise. h_{ij} denotes the channel coefficient between nodes i and j , which satisfies $|h_{ij}|^2 = |h'_{ij}|^2 d_{ij}^{-\alpha}$, h'_{ij} characterizes the rapid fading, d_{ij} is the distance between i and j , and α is the path loss exponent. It is assumed that the channel undergoes the flat Rayleigh fading and h'_{ij} follows zero-mean circular-symmetric complex Gaussian distribution $\mathcal{CN}(0, 1/\mu_{ij})$. Signal-interference-ratio (SIR) distribution is independent of μ_{ij} in the interference limited system while μ_{ij} will affect the system performance when the noise is considered. Here we let $\mu_{ij} = 1$ for simplification when noise is ignored.

In this work, we focus on the performance of multicast delivery, which is analyzed by randomly selecting a multicast transmitter i and its corresponding receivers. Considering that the transmitter is the same for these receivers, with a slight abuse of notation, we omit i when we analyze the multicast performance of a receiver of i in the left part of this work, e.g., receiving signal strength, SINR and the probability that the multicast file meeting the requirement of a receiver. For a receiver j , the received signal to noise and interference ratio (SINR) γ_j can be expressed as

$$\gamma_j = \frac{E_i d_{ij}^{-\alpha} H_{ij}}{I_j + N_0}, \tag{2}$$

where I_j is the interference at the receiver j and $H_{ij} = |h'_{ij}|^2$.

Then, a successful transmission of the link $i \rightarrow j$ needs to satisfy

$$\gamma_j \geq \gamma_{th}, \tag{3}$$

where γ_{th} is the required SINR for an effective communication.

B. CONTENT DELIVERY MODELS

Each UE has a cache with size \tilde{N} , which is populated with some video files with the same size. We assume that the probability that a UE caches a file n is $f_c(n)$, and the distribution of cached files is independent identically distributed (iid) at each UE. Not all of these cached contents are required by the UE itself. However, it may provide useful contents to its neighbors by D2D multicast. Similarly, this UE can also benefit from the content cached by other UEs. Hence, this caching policy is mutually beneficial. Assuming that all contents have the same size, UEs disseminate these cached contents to nearby UEs in their transmission resources to bring popular contents closer to users. Here two ways for content delivery can be utilized according to a pre-defined mode selection strategy, i.e., D2D delivery and cellular delivery.

The pre-defined delivery strategy is that UE or BS transmits content following a certain distribution f_s , which is iid at each UE and BS. Based on this distribution, the transmission

probability of a content in a cache is determined. For example, multicast transmitters transmit file n with probability $f_s(n)$ in allocated bandwidth during a multicast period. All neighbor UEs of a multicast transmitter will be potential receivers. These potential receivers which have no data to transmit will be scheduled by BS to listen the muticast files of their neighbor transmitters. When a UE be a potential receiver of multiple multicast transmitters, BS randomly selects a multicast transmitter for it. The neighbor relationship among UEs can be obtained by BS during the neighbor discovery period which is an essential procedure before D2D communications standardized by the 3rd Generation Partnership Project (3GPP) [35].

We use popularity to describe the probability of a file being required by a UE. That is, the required statistics for video files are modeled by a Zipf distribution, which has been shown to fit well with video requirement and widely used for the analysis of D2D caching [36]–[38]. Hence, the probability $f_r(n)$ that file n is required by UE can be expressed by

$$f_r(n) = \frac{1}{n^\beta} \left[\sum_{l=1}^N \frac{1}{l^\beta} \right]^{-1}, \quad (4)$$

where β is the Zipf exponent, n is the n th most popular video and N is the number of files ($N \leq \tilde{N}$).

III. D2D CONTENT DELIVERY STRATEGIES

A. DELIVERY ENERGY AND COVERAGE

We call a UE, which can benefit from the content delivery by receiving its required file, as a benefiting UE (BUE). For D2D delivery, the number of BUEs is one key evaluation metric to determine the efficiency of delivery. Since the number of BUEs depends on the coverage of a transmitter, we first give the following definition.

Definition 1: A targeted sharing zone (TSZ) is the area that a transmitting UE targets to cover for sharing its cached files.

For a transmitter i , its TSZ \mathcal{T}_i can be modeled as a circle with center i and radius d_0 . We assume that the path loss is compensated by power control, i.e.,

$$E_i = \eta_0 d_0^{k\alpha}, \quad (5)$$

where k is the path loss compensatory exponent, and η_0 is the targeted receiving power at the receiver.

Since the goal of D2D caching and delivery is to meet file requirement of UEs in the network through proximal transmission, a UE having no data to transmit is scheduled to receive files as long as it locates in the TSZ of a transmitter. Several transmitters may cover a receiver simultaneously and the receiver randomly selects a transmitter. Since simultaneous transmitters spatially reuse transmission resources, the signals from different transmitters may be interfered with each other at the receiver. For a receiver j and its transmitter i , the SINR of the received signal γ_j at j can be

expressed as

$$\begin{aligned} \gamma_j &= \frac{E_i H_{ij} d_{ij}^{-\alpha}}{I_j + N_0} \\ &= \frac{\eta_0 d_0^{k\alpha} H_{ij} d_{ij}^{-\alpha}}{\sum_{l \in \Psi_i \setminus i} \eta_0 d_0^{k\alpha} H_{lj} d_{lj}^{-\alpha} + N_0} \\ &= \frac{H_{ij} d_{ij}^{-\alpha}}{\sum_{l \in \Psi_i \setminus i} H_{lj} d_{lj}^{-\alpha} + \frac{N_0}{\eta_0 d_0^{k\alpha}}}. \end{aligned} \quad (6)$$

Denote the file sent by transmitter i as S_i and the file required by receiver j as R_j . When i transmits file n , the conditional probability $M_j(n)$ that n meets the requirement of j is

$$M_j(n) = \Pr \{ \gamma_j \geq \gamma_{th} \} \Pr \{ R_j = n | S_j = n \}, \quad (7)$$

where $\Pr \{ R_j = n | S_j = n \}$ is the conditional probability that UE j requires file n when UE i sends this file.

We assume that H_{ij} is exponentially distributed and noise is negligible. The first part in (7) can be calculated by (8), i.e.,

$$\begin{aligned} \Pr \{ \gamma_j \geq \gamma_{th} \} &= \Pr \left\{ H_{ij} \geq \gamma_{th} d_{ij}^\alpha \sum_{l \in \Psi_i} d_{lj}^{-\alpha} H_{lj} \right\} \\ &= \mathbb{E} \left[\exp \left(-\gamma_{th} d_{ij}^\alpha I_j' \right) \right] \\ &= \mathcal{L}_{I_j'} \left(\gamma_{th} d_{ij}^\alpha \right) \\ &= \exp \left(-\pi \lambda_t m \Gamma(m) (1-m) \gamma_{th}^m d_{ij}^2 \right) \\ &= \exp \left(-\kappa d_{ij}^2 \right), \end{aligned} \quad (8)$$

where $\mathcal{L}_{I_j'}$ is the Laplace transform of $I_j' = \sum_{l \in \Psi_i} d_{lj}^{-\alpha} H_{lj}$, $\kappa = \pi \lambda_t m \Gamma(m) (1-m) \gamma_{th}^m$, $m = \frac{2}{\alpha}$ and $\Gamma(\cdot)$ is the Gamma function. The impact of interference from simultaneous transmitters is implicitly embodied by κ based on stochastic geometry analyses (Appendix A in [39]), which depends on the density of simultaneous transmitters (λ_t). That is, κ increases with interference which is led by the increase of the density of simultaneous transmitters. From (8), we see that the average successful transmission is affected by path loss (determined by link distance) and the density of simultaneous transmitters (determined by scheduling).

The second part in (7) can be derived by

$$\begin{aligned} \Pr \{ R_j = n | S_j = n \} &= \frac{\Pr \{ R_j = n, S_j = n \}}{\Pr \{ S_j = n \}} \\ &= \Pr \{ R_j = n \} \\ &= \frac{1}{n^\beta} \left[\sum_{l=1}^N \frac{1}{l^\beta} \right]^{-1}. \end{aligned} \quad (9)$$

Including (8) and (9) into (7) yields the conditional probability that file n meets the requirement at j as follows:

$$M_j(n) = \frac{\exp\left(-\kappa d_{ij}^2\right) \frac{1}{n^\beta}}{\sum_{l=1}^N \frac{1}{l^\beta}}. \quad (10)$$

For the receiver j , the probability ω_j that its receiving file meets its requirement when it is in the TSZ of transmitter i can be given by

$$\omega_j = \sum_{n=1}^N \Pr\{S_j = n\} M_j(n), \quad (11)$$

where $\Pr\{S_j = n\}$ is the probability that transmitter i caches file n and sends it, which can be written as

$$\Pr\{S_j = n\} = f_c(n) f_s(n), \quad (12)$$

where $f_c(n)$ and $f_s(n)$ are probabilities that the transmitter caches file n and sends it, respectively.

Including (10) and (12) into (11), we have

$$\begin{aligned} \omega_j &= \sum_{n=1}^N \Pr\{S_j = n\} M_j(n) \\ &= \sum_{n=1}^N f_c(n) f_s(n) \frac{\exp\left(-\kappa d_{ij}^2\right) \frac{1}{n^\beta}}{\sum_{l=1}^N \frac{1}{l^\beta}} \\ &= \frac{\exp\left(-\kappa d_{ij}^2\right)}{\sum_{l=1}^N \frac{1}{l^\beta}} \sum_{n=1}^N f_c(n) f_s(n) \frac{1}{n^\beta}. \end{aligned} \quad (13)$$

From (13), we observe that larger transmission distance d_{ij} leads to a decrease of meeting probability ω_j . From this perspective, the coverage of i , TSZ \mathcal{T}_i , does not need to be too large. However, larger \mathcal{T}_i may bring benefits to more receivers who need this file for a fixed receiver density. For an optimal TSZ of a transmitter in a general view, we then analyze the total number of receivers that successfully receive their required files, i.e., the BUE number N_r in \mathcal{T}_i . Let d_σ be an infinitely small area of the TSZ on interval $[\tau, \tau + d\tau]$. Since the probability that there are more than two UEs in d_σ can be neglected, the number of receivers that obtain required files in d_σ can be expressed as $\lambda_r \omega_{d_\sigma}$, where ω_{d_σ} is the receiving probability at d_σ and can be obtained from (13) through using τ instead of d_{ij} , i.e.,

$$\omega_{d_\sigma} = \frac{\exp\left(-\kappa \tau^2\right)}{\sum_{l=1}^N \frac{1}{l^\beta}} \sum_{n=1}^N f_c(n) f_s(n) \frac{1}{n^\beta}. \quad (14)$$

With the number of BUEs at location d_σ , N_r in the TSZ can be calculated by traversing all locations in the TSZ as follows.

$$\begin{aligned} N_r &= \int_{\mathcal{T}_i} \lambda_r \omega_{d_\sigma} d\sigma \\ &= \lambda_r \int_0^{2\pi} \int_0^{d_0} \omega_{d_\sigma} \tau d\tau \\ &= \frac{\pi \lambda_r}{\kappa \sum_{l=1}^N \frac{1}{l^\beta}} \left[1 - \exp\left(-\kappa d_0^2\right)\right] \sum_{n=1}^N f_c(n) f_s(n) \frac{1}{n^\beta}. \end{aligned} \quad (15)$$

Proposition 1: Given a transmitter, the number N_r of BUEs in its TSZ increases with the increase of coverage range of the TSZ, while this increase has an upper bound. Moreover, the upper bound $N_r^{u.b.}$ of N_r can be expressed by

$$N_r^{u.b.} = \frac{\pi \lambda_r}{\kappa \sum_{l=1}^N \frac{1}{l^\beta}} \sum_{n=1}^N f_c(n) f_s(n) \frac{1}{n^\beta}. \quad (16)$$

Proof: Taking the derivative of N_r in (15) with respect to d_0 , we can obtain the derivation v as follows:

$$\begin{aligned} v &= \frac{\partial}{\partial d_0} \frac{\pi \lambda_r}{\kappa \sum_{l=1}^N \frac{1}{l^\beta}} \left[1 - \exp\left(-\kappa d_0^2\right)\right] \sum_{n=1}^N f_c(n) f_s(n) \frac{1}{n^\beta} \\ &= 2\kappa d_0 \exp\left(-\kappa d_0^2\right) \frac{\pi \lambda_r}{\kappa \sum_{l=1}^N \frac{1}{l^\beta}} \sum_{n=1}^N f_c(n) f_s(n) \frac{1}{n^\beta} > 0. \end{aligned} \quad (17)$$

Thus, N_r increases with d_0 , i.e., the coverage range of the TSZ. Since

$$\lim_{d_0 \rightarrow \infty} v = 0, \quad (18)$$

N_r has an upper bound $N_r^{u.b.}$, which can be calculated by

$$\begin{aligned} N_r^{u.b.} &= \lim_{d_0 \rightarrow \infty} \frac{\pi \lambda_r}{\kappa \sum_{l=1}^N \frac{1}{l^\beta}} \left[1 - \exp\left(-\kappa d_0^2\right)\right] \sum_{n=1}^N f_c(n) f_s(n) \frac{1}{n^\beta} \\ &= \frac{\pi \lambda_r}{\kappa \sum_{l=1}^N \frac{1}{l^\beta}} \sum_{n=1}^N f_c(n) f_s(n) \frac{1}{n^\beta}. \end{aligned} \quad (19)$$

For a targeted coverage with radius d_0 , the transmission power needs to be $E_i = \eta_0 d_0^{k\alpha}$ according to the power control strategy. That is, the delivery range is determined by the transmission power as follows:

$$d_0 = \left(\frac{E_i}{\eta_0}\right)^{\frac{1}{k\alpha}}. \quad (20)$$

Considering that d_0 increases with the transmission power, we can conclude from above analyses that N_r increases with the D2D transmission power before achieving the upper

bound. However, it is not energy-efficient to increase N_r at the cost of increasing energy consumption due to the low energy capacity of UE and the energy conservation requirement of green networks.

To increase N_r with less energy consumption, we firstly evaluate the efficiency of content delivery in cellular networks. Denote the transmission efficiency coefficient ρ as the average energy consumption of a successful acquisition for required files at receivers in a TSZ, i.e.,

$$\rho = \frac{E_i}{N_r}. \quad (21)$$

Through applying (15) into (21), we can obtain ρ as follows:

$$\rho = \frac{E_i}{\frac{\pi \lambda_r}{\kappa \sum_{l=1}^N \frac{1}{l^\beta}} [1 - \exp(-\kappa d_0^2)] \sum_{n=1}^N f_c(n) f_s(n) \frac{1}{n^\beta}}. \quad (22)$$

Since the energy consumption and interference among transmitters also increase with E_i , there is a trade-off between E_i and N_r . Concerning the optimal transmission power and its corresponding TSZ of a given transmitter, we have the following proposition.

Proposition 2: To minimize the transmission efficiency coefficient ρ under the transmission power constraint E_{\max} of a UE, i.e., minimizing the average energy consumption for a successful file delivery from transmitter i to its receivers, the optimal transmission power E_i^ should satisfy*

$$E_i^* = \begin{cases} \tilde{E}_i & \tilde{E}_i \leq E_{\max} \\ E_{\max} & \tilde{E}_i > E_{\max}, \end{cases} \quad (23)$$

where \tilde{E}_i needs to satisfy

$$\exp\left(-\frac{\kappa}{\eta_0^{m'}} \tilde{E}_i^{m'}\right) + \frac{\kappa m' \tilde{E}_i^{m'}}{\eta_0^{m'}} \exp\left(-\frac{\kappa}{\eta_0^{m'}} \tilde{E}_i^{m'}\right) - 1 = 0. \quad (24)$$

Proof: To minimize the average energy consumption of file delivery for a BUE, the optimal problem can be formulated as follows:

$$\begin{aligned} \min \quad & \rho = \frac{E_i}{\frac{\pi \lambda_r}{\kappa \sum_{l=1}^N \frac{1}{l^\beta}} [1 - \exp(-\kappa d_0^2)] \sum_{n=1}^N f_c(n) f_s(n) \frac{1}{n^\beta}} \\ \text{s.t.} \quad & 0 < E_i \leq E_{\max}. \end{aligned} \quad (25)$$

To solve (25), we include $d_0 = \left(\frac{E_i}{\eta_0}\right)^{\frac{1}{\kappa \alpha}}$ into the objective function (25), and then take the derivative of ρ with respect

to E_i , i.e.,

$$\begin{aligned} v' &= \frac{\partial}{\partial E_i} \frac{E_i}{\frac{\pi \lambda_r}{\kappa \sum_{l=1}^N \frac{1}{l^\beta}} [1 - \exp(-\kappa d_0^2)] \sum_{n=1}^N f_c(n) f_s(n) \frac{1}{n^\beta}} \\ &= \left[\frac{\pi \lambda_r}{\kappa \sum_{l=1}^N \frac{1}{l^\beta}} \sum_{n=1}^N f_c(n) f_s(n) \frac{1}{n^\beta} \right]^{-1} \frac{\partial}{\partial E_i} \left[\frac{E_i}{1 - \exp\left(-\frac{\kappa}{\eta_0^{m'}} E_i^{m'}\right)} \right] \\ &= \mathcal{A} \frac{\left[1 - \exp\left(-\frac{\kappa}{\eta_0^{m'}} E_i^{m'}\right) \right] - \frac{\kappa m' E_i^{m'}}{\eta_0^{m'}} \exp\left(-\frac{\kappa}{\eta_0^{m'}} E_i^{m'}\right)}{\left[1 - \exp\left(-\frac{\kappa}{\eta_0^{m'}} E_i^{m'}\right) \right]^2}, \end{aligned} \quad (26)$$

where $m' = \frac{2}{\kappa \alpha}$ and $\mathcal{A} = \left[\frac{\pi \lambda_r}{\kappa \sum_{l=1}^N \frac{1}{l^\beta}} \sum_{n=1}^N f_c(n) f_s(n) \frac{1}{n^\beta} \right]^{-1}$.

Letting $v' = 0$, we can obtain the optimal \tilde{E}_i satisfying

$$\exp\left(-\frac{\kappa}{\eta_0^{m'}} \tilde{E}_i^{m'}\right) + \frac{\kappa m' \tilde{E}_i^{m'}}{\eta_0^{m'}} \exp\left(-\frac{\kappa}{\eta_0^{m'}} \tilde{E}_i^{m'}\right) - 1 = 0, \quad (27)$$

which can be resolved easily using numerical methods. ■

Based on Proposition 2, the optimal radius d_0^* of a TSZ for a transmitter can be calculated by using (20), i.e., $d_0^* = \left(\frac{E_i^*}{\eta_0}\right)^{\frac{1}{\kappa \alpha}}$. In addition, from this proposition, we can find that E_i^* is determined by wireless environment parameters such as transmitter density, path loss factor, decoding threshold, and so on, and is independent to the caching distribution and content delivery distributions (f_c, f_s). However, f_c and f_s affect the probability that BUEs in the TSZ receive their required files. Without considering energy consumption, the coverage radius d_0 and N_r can increase infinitely with transmission power. Although more UEs may benefit from once multicast, the benefit of D2D content delivery is reduced compared with BS content delivery, and interference becomes severer due to longer transmission distance.

B. COMPARISON OF ENERGY EFFICIENCY FOR D2D MULTICAST AND UNICAST

With the push strategy studied in this work, more than one BUE may exit for once multicast. To understand the advantages and constraints of energy efficiency for this push strategy, we compare the energy efficiency of D2D multicast and unicast (request-and-response strategy) here. For fair comparison, we evaluate the average transmission energy consumption per request for two strategies to show which one is more energy efficient via taking content request, user distribution and wireless transmission environment into consideration.

For the unicast content delivery, the transmission power can be expressed by

$$E_i = \eta_0 d_{ij}^\alpha. \quad (28)$$

For a constructed unicast link $i \rightarrow j$, the energy consumption of a successful delivery can be calculated by

$$\varepsilon_u^{succ} = \frac{E_i}{\Pr\{\gamma_{ij} \geq \gamma_{th}|i\}} = \frac{\eta_0 d_{ij}^\alpha}{\exp(-\kappa d_{ij}^2)}. \quad (29)$$

By averaging the location of transmitters, the average consumed energy of serving a request, ρ_u , can be calculated by

$$\begin{aligned} \rho_u &= \int_s \varepsilon_u^{succ} d\sigma \\ &= \int_0^{2\pi} \int_0^{d_0} \lambda_t \frac{\eta_0 \tau^\alpha}{\exp(-\kappa \tau^2)} \tau d\tau \\ &= \lambda_t B\left(1, \frac{\alpha}{2} + 1\right) d_0^2 {}_1F_1\left(\frac{\alpha}{2} + 1; \frac{\alpha}{2} + 2; \kappa d_0^2\right), \end{aligned} \quad (30)$$

where σ is an infinite small area, $B(x, y)$ is the Beta function and ${}_1F_1(x; y; z)$ is the Degenerate hypergeometric function.

For the multicast content delivery, the transmission power is $\eta_0 d_0^\alpha$ for once transmission since the targeted delivery distance is d_0 . Taking power control into consideration, unicast may be more energy efficient than multicast. For example, although the farthest receiver in the targeted coverage of multicast does not need the transmitted content, the transmission power is also fixed as ε_m . Even if the transmission power can be adaptively adjusted according to the distance of the farthest receiver requiring this content, energy is also consumed by overhead. However, multiple UEs may benefit from once multicast by receiving required or to be required contents. In this perspective, multicast is more energy efficient than unicast. As analyzed in the last subsection, the average energy consumed by a successful content requirement for a BUE is expressed by ρ , which is re-written as

$$\rho = E_i \left[\frac{\pi \lambda_r}{\kappa \sum_{l=1}^N \frac{1}{l^\beta}} \left[1 - \exp(-\kappa d_0^2) \right] \sum_{n=1}^N f_c(n) f_s(n) \frac{1}{n^\beta} \right]^{-1}. \quad (31)$$

Hence, we can resort to (30) and (31) to compare the energy efficiency of two strategies. From the comparison we observe that not only the UE densities (receiver density and transmitter density) but also caching and delivery strategies will affect the final efficiency of D2D multicast delivery.

C. CACHING AND DELIVERY STRATEGIES

From (15), we can see that, besides the transmission power E_i , the number of BUEs that receive required files (N_r) also depends on the caching and delivery schemes, i.e., the probability for the transmitter to cache and disseminate files. In addition, they also impact the advantages of the push

strategy. Here we discuss the caching and delivery distributions (f_c and f_s) of a transmitter to improve the D2D content delivery efficiency.

Let ζ denote a caching and delivery efficiency coefficient which characterizes the impact from caching and delivery distributions on the delivery performance, i.e.,

$$\zeta = \sum_{n=1}^N f_c(n) f_s(n) \frac{1}{n^\beta}. \quad (32)$$

The derivation of (15) shows that N_r grows with ζ for a given transmission power. Thus, to maximize N_r , i.e., maximizing the average number of receivers that obtain the required files, ζ needs to be maximized. The optimal problem can be formulated as follows:

$$\begin{aligned} \max \quad & \zeta = \sum_{n=1}^N f_c(n) f_s(n) \frac{1}{n^\beta} \\ \text{s.t.} \quad & \sum_{n=1}^N f_c(n) = 1 \\ & \sum_{n=1}^N f_s(n) = 1. \end{aligned} \quad (33)$$

With the Lagrange multipliers method, the objective function of this optimal problem can be written as

$$\begin{aligned} \mathcal{L}(\mathbf{g}, \delta) &= \sum_{n=1}^N f_c(n) f_s(n) \frac{1}{n^\beta} + \delta_1 \left(\sum_{n=1}^N f_c(n) - 1 \right) \\ &\quad + \delta_2 \left(\sum_{n=1}^N f_s(n) - 1 \right), \end{aligned} \quad (34)$$

where δ_1 and δ_2 are the Lagrange multipliers. To maximize the objective function in (33), we calculate the partial derivation of $\mathcal{L}(\mathbf{g}, \delta)$ with respect to $f_s(n)$ and $f_c(n)$ by

$$\frac{\partial}{\partial f_c(n)} \mathcal{L}(\mathbf{g}, \delta) = f_s(n) \frac{1}{n^\beta} + \delta_1, \quad n = 1, 2 \dots N, \quad (35a)$$

$$\frac{\partial}{\partial f_s(n)} \mathcal{L}(\mathbf{g}, \delta) = f_c(n) \frac{1}{n^\beta} + \delta_2, \quad n = 1, 2 \dots N. \quad (35b)$$

Since for all files, i.e., $n = 1, 2 \dots N$, (35a) needs to be satisfied, we can obtain the relationship of the cached probabilities among these different files, which can be expressed as

$$\frac{f_c(p)}{f_c(q)} = \left(\frac{q}{p}\right)^\beta \quad p \neq q. \quad (36)$$

Due to $\sum_{n=1}^N f_c(n) = 1$, the distribution of f_c also follows the Zipf distribution, i.e.,

$$f_c(n) = \frac{1}{n^\beta} \frac{1}{\sum_{l=1}^N \frac{1}{l^\beta}}, \quad n = 1, 2 \dots N. \quad (37)$$

Similarly, from (35b) we can obtain the relationship of transmission probabilities among these different files, which can be expressed as follows:

$$\frac{f_s(p)}{f_s(q)} = \left(\frac{q}{p}\right)^\beta, \quad p \neq q. \quad (38)$$

The distribution of f_s also follows Zipf distribution as the distribution of required files, i.e.,

$$f_s(n) = \frac{\frac{1}{n^\beta}}{\sum_{l=1}^N \frac{1}{l^\beta}}, \quad n = 1, 2, \dots, N. \quad (39)$$

Theorem 1: To maximize the delivery efficiency, i.e., maximize the number of BUEs in the TSZ, the optimal file delivery distribution and the file caching distribution are Zipf when the required file distribution follows Zipf.

Based on above Theorem, more UEs can benefit from once multicast. The caching and delivery at UEs can be controlled by BSs to make UEs friendly and mutually beneficial.

IV. DELIVERY MODE SELECTION

The intention of D2D delivery is to improve resource efficiency for the cellular content delivery by using proactive cache and proximal transmission. However, this efficiency gain depends on many factors, such as UE density, file caching and delivery distributions at UEs, channel quality and so on. In this section, we will compare D2D and BS content delivery in term of energy efficiency to understand the benefits and practical limits of the D2D delivery. Moreover, we discuss the criteria of mode selection for energy efficiency delivery, which is critical to integrate D2D into cellular networks [40]–[42]. Here we use the expected energy consumption for a successful file acquisition at a receiver as the evaluation metric. For a receiver j which requires a file, the probability $\xi_D(i)$ of receiving its required file, called meeting probability (MP), which can be written as

$$\xi_D(i) = \sum_{n=1}^N \exp(-\kappa d_{ij}^2) f_s(n) f_c(n) f_r(n), \quad (40)$$

Then the expectation of $\xi_D(i)$, ξ_D , for a receiver under the D2D delivery mode can be calculated by averaging the location of i as follows:

$$\begin{aligned} \xi_D &= \int_0^{2\pi} \int_0^{d_0} \lambda_t \xi_D(i) \tau d\tau \\ &= 2\pi \lambda_t \int_0^{d_0} \frac{\exp(-\kappa \tau^2)}{\sum_{l=1}^N \frac{1}{l^\beta}} \sum_{n=1}^N f_s(n) f_c(n) \frac{1}{n^\beta} \tau d\tau \\ &= \frac{\pi \lambda_t [1 - \exp(-\kappa d_0^2)]}{\kappa \sum_{l=1}^N \frac{1}{l^\beta}} \sum_{n=1}^N f_s(n) f_c(n) \frac{1}{n^\beta}. \end{aligned} \quad (41)$$

With the BS content delivery, the probability that UE j receives the required file from its serving BS only depends

on the channel status and delivery distribution under the assumption that BS always stores the required files of UEs. Then MP with the BS delivery can be expressed as

$$\xi_C(i) = \sum_{n=1}^N \Pr\{\gamma_j \geq \gamma_0\} f_s(n) f_r(n), \quad (42)$$

where

$$\Pr\{\gamma_j \geq \gamma_{th}\} = \exp(-\kappa' d_{ij}^2), \quad (43)$$

and d_R is the coverage radius of the BS with $\kappa' = \pi \lambda_b m \Gamma(m) \Gamma(1-m) \gamma_{th}^m$. The derivation of (43) is similar to that of (8) and the difference is that the BS density is λ_b . By averaging locations of all potential serving BSs of the UE j , the expectation of $\xi_C(i)$, ξ_C , can be written as

$$\xi_C = \frac{\sum_{n=1}^N f_s(n) \frac{1}{n^\beta}}{\sum_{l=1}^N \frac{1}{l^\beta}} \int_0^{d_R} \exp(-\kappa' \tau^2) f(\tau) d\tau, \quad (44)$$

where $f(\tau)$ is the distribution of the distance between a receiver and its serving BS. In cellular networks, a UE usually selects its attached BS based on RSSI. Hence, the distribution of the distance between a UE and its serving BS $f(\tau)$ can be expressed as

$$f(\tau) = 2\pi \lambda_b \tau \exp(-\lambda_b \pi \tau^2). \quad (45)$$

Including (45) into (44), ξ_C can be written as

$$\begin{aligned} \xi_C &= \frac{\sum_{n=1}^N f_s(n) \frac{1}{n^\beta}}{\sum_{l=1}^N \frac{1}{l^\beta}} \int_0^{d_R} \exp(-\kappa' \tau^2) 2\pi \lambda_b \tau \exp(-\lambda_b \pi \tau^2) d\tau \\ &= \frac{\sum_{n=1}^N f_s(n) \frac{1}{n^\beta}}{\sum_{l=1}^N \frac{1}{l^\beta}} \frac{\pi \lambda_b}{\kappa' + \pi \lambda_b} \left[1 - \exp\left(-(\kappa' - \pi \lambda_b) d_R^2\right) \right]. \end{aligned} \quad (46)$$

For a fair comparison between cellular and D2D delivery in term of energy efficiency, we compare the MPs under different modes with the same transmission energy and delivery coverage. Let the ratio between ξ_C and ξ_D be ε_e . With (41) and (46), we can obtain ε_e as follows:

$$\begin{aligned} \varepsilon_e &= \frac{\xi_C}{\xi_D} \\ &= \frac{\sum_{n=1}^N \frac{f_s(n)}{n^\beta} \left[\sum_{l=1}^N \frac{1}{l^\beta} \right]^{-1} \frac{\pi \lambda_b}{\kappa' + \pi \lambda_b} \left[1 - \exp\left(-(\kappa' - \pi \lambda_b) d_R^2\right) \right]}{\pi \lambda_t \left[1 - \exp(-\kappa d_0^2) \right] \left[\kappa \sum_{l=1}^N \frac{1}{l^\beta} \right]^{-1} \sum_{n=1}^N f_s(n) f_c(n) \frac{1}{n^\beta}} \\ &= \frac{\kappa \lambda_b \left\{ 1 - \exp\left[-(\kappa' - \pi \lambda_b) d_R^2\right] \right\} \sum_{n=1}^N f_s(n) \frac{1}{n^\beta}}{\lambda_t (\kappa' + \pi \lambda_b) \left[1 - \exp(-\kappa d_0^2) \right] \sum_{n=1}^N f_s(n) f_c(n) \frac{1}{n^\beta}}. \end{aligned} \quad (47)$$

TABLE 2. Mode selection algorithm for content delivery.

1. For a D2D UE i scheduled to transmit in a content delivery period
2. BS calculates ε_e based on (47)
3. If $\varepsilon_e \leq 1$,
4. UE i transmits cached file n with probability $f_s(n)$
5. BS schedules UEs within T_i to listen to the multicast file
6. else
7. BS transmits cached file n with probability $f_s(n)$
8. BS schedules attached UEs to listen to the multicast file
9. end

TABLE 3. Simulation parameter settings.

Simulation parameter	Value
Transmitter UE density λ_t	$10^{-7}, 10^{-6}, 10^{-5}$
BS density λ_b	$10^{-7}, 10^{-6}, 10^{-5}$
Receiver UE density λ_r	10^{-2}
SINR threshold γ_{th}	10 dB
Targeted receiving power η_0	-80 dBm
Maximum transmission power E_{Tmax}	31 dBm
Path loss compensatory factor k	1
Path loss factor α	3
Zipf exponent β	2
Files number N	100
Simulation times T	1000

From (47), we can find that ε_e is affected by multiple factors such as the density of BS, D2D transmitters, the targeted coverage range of delivery, the distributions of delivery and caching. The transmission mode of a UE is determined by its serving BS based on ε_e . That is, BS selects the cellular mode when $\varepsilon_e > 1$. Otherwise, BS selects the D2D mode. The detailed mode selection algorithm is listed in Table 2. Here we can see that the mode selection is based on the spatial average performance of communication links which is independent of the location of a transmitter. Then all transmitters will perform the same mode selection decision to achieve a better average performance without much signalling overhead cost. Thus the distribution of transmitters still follows homogenous PPP after mode selection.

V. PERFORMANCE EVALUATION

To evaluate proposed schemes, we present some numerical and simulation results in this section. The default parameters are listed in Table 3. In the simulation, each UE has a single role, i.e., transmitter or receiver during a simulation round. The network performance can be well reflected due to the randomness of their locations. More detailed simulation and relating parameters setup for D2D communication can be referred to [35]. Due to different system model and problem formulation, we will not compare with existing works such as [43] in the simulation. Instead, we widely simulate various transmission environment parameters to check their impacts on D2D multicast delivery. In addition, we indicate that the proposed D2D multicast strategy achieve the best energy efficiency coefficient by simulation.

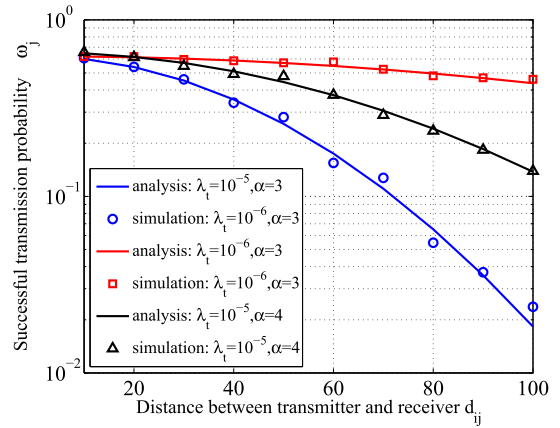


FIGURE 1. Successfully receiving probability ω_j against the D2D link distance d_{ij} .

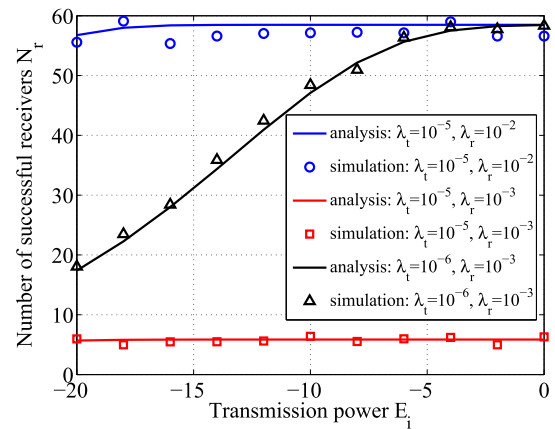


FIGURE 2. The number of UEs successfully receiving their required files N_r against transmission power E_t .

Firstly, we compare successfully receiving probability (ω_j) at a receiver against the distance between receiver and transmitter (d_{ij}) under different settings of transmitter densities (λ_t) and path loss factors (α) in Fig.1. We can find that ω_j decreases with d_{ij} under different λ_t and α . Thus, a larger TSZ (leading to larger link distances for more transmissions) may not increase N_r . Moreover, comparing lines with circles and squares, respectively, we can observe that the higher transmitter density leads to a smaller successfully receiving probability. Larger α leads to larger ω_j at the higher transmitter density, say $\lambda_t = 10^{-5}$, since interference from simultaneous transmitters decreases with α .

As shown in Fig.2 for N_r against d_0 , N_r increases with the radius of TSZ (d_0). However, this increase is trivial for the higher transmitter density (say $\lambda_t = 10^{-5}$) due to interference. In particular, for the lower receiver density, N_r increases more slowly. Thus, the transmission power needs to be adjusted based on nodes density to avoid energy waste.

Fig.3 compares N_r and $N_r^{u,b}$ against d_0 for $\lambda_t = 10^{-5}$. As shown in this figure, N_r increases with d_0 . However, N_r tends to be a constant when d_0 is larger than a threshold, i.e., N_r has

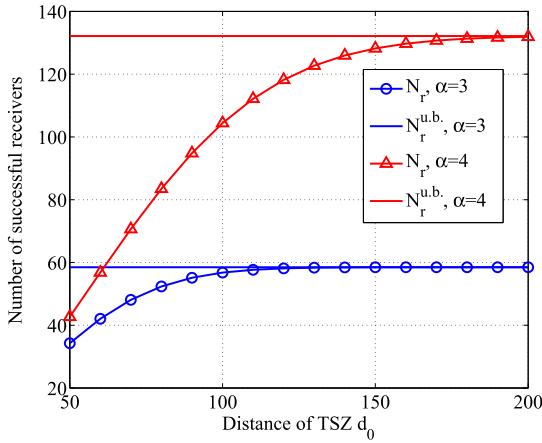


FIGURE 3. The number of UEs successfully receiving their required files N_r and corresponding upper bound $N_r^{u.b.}$ against radiuses of TSZ d_0 .

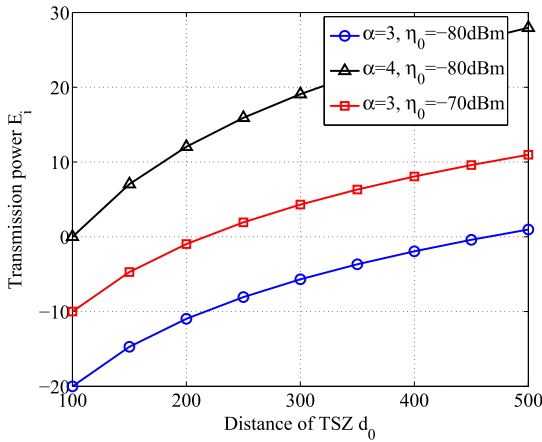


FIGURE 4. The transmission power E_i against radius of TSZ d_0 under different path loss factor α and targeted receiving signal power η_0 .

an upper bound as analyzed. In addition, N_r tends to reach the upper bound more quickly with smaller α (say $\alpha = 3$) due to severer interference from simultaneous transmitters, which match the theoretical analyses, i.e., it is not energy-efficient to increase N_r by enlarging the TSZ of a transmitter.

As shown in Fig.4, the energy consumption (E_i) increases with d_0 under different α and η_0 . In addition, higher η_0 and α lead to higher energy consumptions for the same d_0 , because they need higher transmitter powers to achieve the targeted receiving power and compensate path loss, respectively. Thus, d_0 needs to be reduced to save energy. Fig.3 also shows that N_r increases with d_0 with an upper bound, i.e., N_r does not increase any more when d_0 is larger than a threshold. Thus, an optimal d_0 exists especially for a worse transmission environment as analyzed.

The comparisons of transmission efficiency coefficient (ρ) with different settings of d_0 , α and γ_{th} are shown in Fig.5. It is shown that ρ firstly decreases with the increase of d_0 for small d_0 with ranging from 0 to 50, and then increases with d_0 after achieving its minimal value. Hence there is an optimal d_0

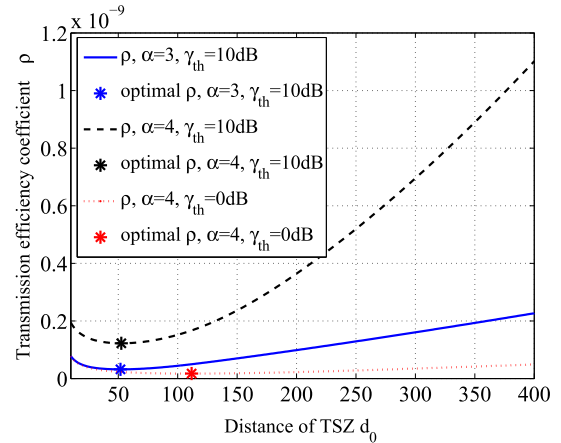


FIGURE 5. The transmission efficiency coefficient ρ with different radius of TSZ d_0 , path loss factor α and decoding threshold γ_{th} .

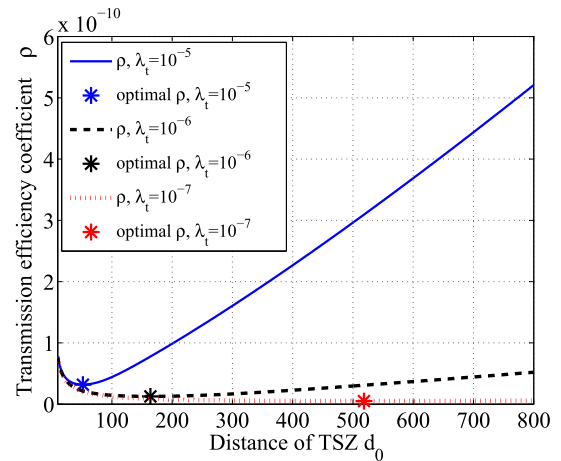


FIGURE 6. The transmission efficiency coefficient ρ with different TSZ radius d_0 and transmitter density λ_t .

for the TSZ given transmission environment parameters such as transmitter density, receiver density and channel quality. In addition, the optimal ρ can achieve the minimal value under different transmission environments, which verifies the effectiveness of the proposed energy-efficiency oriented D2D delivery scheme. Moreover, we can find that ρ increases with path loss and decoding threshold while the best energy efficiency can always be achieved by the proposed delivery power scheme.

Similarly as shown in Fig.6, ρ of the proposed scheme (marked by star points) can obtain the optimal energy efficiency coefficient at different transmitter densities with the smallest transmission power consumption for the same number of successful decoding receivers. In addition, a higher transmitter density yields a higher ρ , i.e., a lower energy efficiency of the D2D delivery due to severer interference.

Fig.7 compares ρ with different d_0 and path loss compensation factor k . We can see that ρ increases with the increase of k since larger path loss leads to a smaller number of successful receiving UEs and the lower energy efficiency.

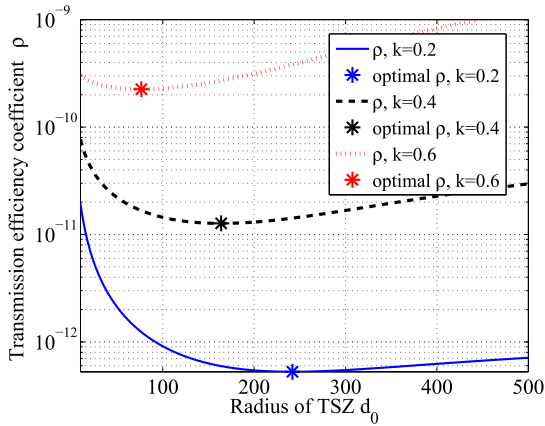


FIGURE 7. The transmission efficiency coefficient ρ with different TSZ radius d_0 and path loss compensation factor k .

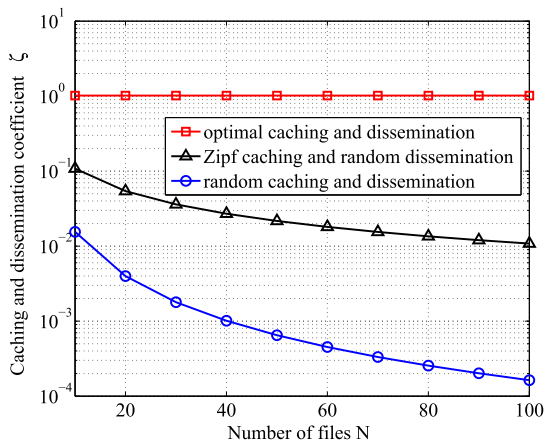


FIGURE 8. Caching and delivery coefficient ζ with different files number N .

The above discussions show that, under different transmission environments (path loss, transmitter density, and decoding threshold), the proposed delivery scheme can achieve the best energy efficiency.

For fair comparison among different caching and delivery strategies, the number of BUEs (N_r) should be compared for different schemes. To focus on caching and delivery strategies, communication parameters such as transmission power and channel quality are set the same for different schemes. From (15) we can see that N_r only depends on ζ when communication parameters are given. Therefore, we can resort to comparing ζ to determine which scheme is more energy efficiency. Fig.8 compares ζ with different files number (N) for three different content caching and delivery schemes: random caching and delivery, Zipf caching and random delivery, and the proposed optimal caching and delivery. We can find that the proposed scheme achieves the highest ζ (approaching to 1) as analyzed under any tested N , which will bring a largest number of BUEs from the delivery of a transmitter. Zipf caching is better than random caching with the random D2D delivery policy, and random caching and delivery performs worst especially with larger N : its

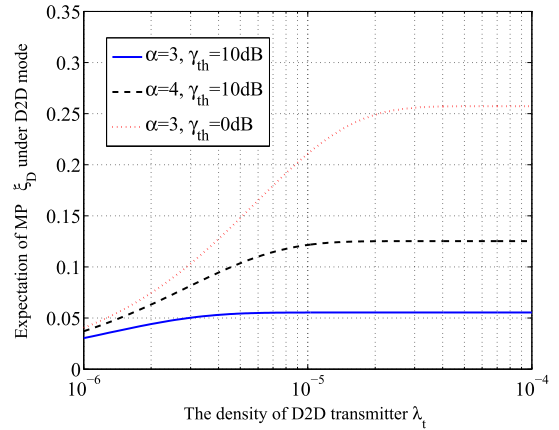


FIGURE 9. The expectation of MP ξ_D under D2D mode against D2D transmitter density λ_t .

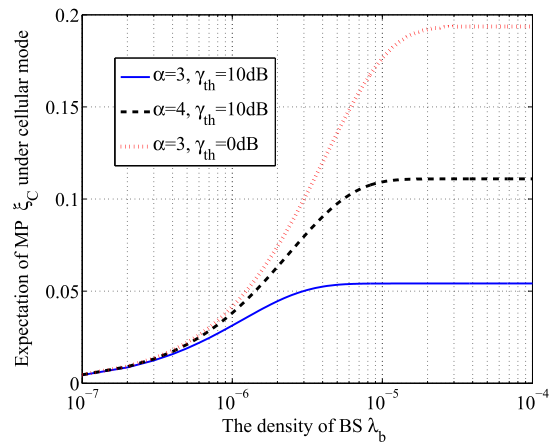


FIGURE 10. The expectation of MP ξ_C under cellular mode against BS density λ_b .

performance degrades significantly with N . On the other hand, the proposed optimal scheme almost performs the same when N changes, which shows the robustness of the proposed scheme.

Fig.9 shows the variation of expectation of MP (ξ_D) under the D2D mode against λ_t . ξ_D increases with the increase of λ_t since more transmitters have larger probability to cover a receiver and send the required file. In addition, the increase of ξ_D with λ_t tends to zero after λ_t goes beyond a threshold due to the severer interference introduced by more transmitters. ξ_D decreases with the α and γ_{th} due to worse file receiving condition.

Fig.10 shows the expectation of MP (ξ_C) under the cellular mode against λ_b . The targeted coverage distance of the BS, d_R , is set equal to d_0 for better comparison. Similar to Fig.9, ξ_C also increases with λ_b with an upper bound due to interference from simultaneous transmitters. In addition, ξ_C decreases with the increase of α and γ_{th} due to worse file receiving condition. Through comparing Figs. 9 and 10, we can find that under the same transmission environment, the BS delivery performs better than the D2D delivery, because BS stores all files instead of caching files opportunistically as the D2D UE does. Therefore, the D2D

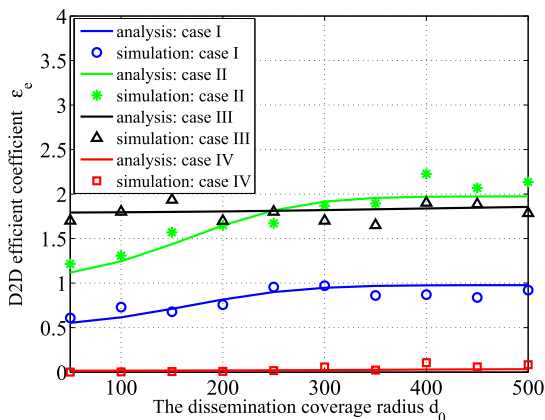


FIGURE 11. D2D energy efficiency coefficient ϵ_e under different cases. Case I: optimal delivery, $\beta = 2$ and $\lambda_t/\lambda_b = 2$, Case II: optimal delivery, $\beta = 0.5$ and $\lambda_t/\lambda_b = 2$, Case III: random delivery, $\beta = 0.5$ and $\lambda_t/\lambda_b = 2$, and Case IV: optimal delivery, $\beta = 0.5$ and $\lambda_t/\lambda_b = 5$.

file delivery can be a complement choice under certain conditions.

In Fig. 11, ϵ_e against d_0 are compared for different settings of λ_b , β and delivery strategies. The targeted delivery radiuses are set equal for the D2D transmitter and BS for fair comparison in term of energy efficiency. For the random delivery strategy, in which a D2D transmitter randomly caches files, we can observe that the D2D content delivery performs worse than the BS content delivery ($\epsilon_e > 1$). However, for the optimal delivery strategy in which D2D transmitter caches and transmits files optimally, the performance of D2D content delivery is improved. When the D2D transmitter density is much larger than BS density (say $\lambda_t = 5\lambda_b$), the D2D delivery may outperform the BS delivery ($\epsilon_e < 1$), since D2D transmitters bring content sharing closer to receivers due to the dense UE deployment. Thus, it is better to select the D2D delivery when available UEs for content multicast delivery are of higher density. For smaller β , the D2D delivery is worse than BS delivery due to limited caching capacity with not enough required files. Thus, D2D content delivery may perform better under some conditions. We can resort to ϵ_e to determine when to use the D2D delivery.

VI. CONCLUSION

In this paper, we focus on D2D content delivery strategies in order to make the content delivery in next-generation cellular networks more energy efficient, which have been proved by both theoretical and experimental results. The major outcomes are summarized as follows: i) an optimal transmission power scheme, based on which the UE cooperation range for proximal content delivery can be determined, ii) an energy efficient content delivery policy, which allows a node to organize the delivery and caching of files to minimize the energy consumption for a successful file delivery, and iii) a quantitative metric to evaluate the energy efficiency of using D2D content delivery, which provides an important insight into determining whether D2D would be a feasible option to be used for content delivery.

REFERENCES

- [1] Cisco, "Cisco visual networking index: Global mobile data traffic forecast update, 2015–2020 white paper," Cisco, San Jose, CA, USA, White Paper, 2016. [Online]. Available: <http://www.cisco.com/c/en/us/solutions/collateral/service-provider/visual-networking-index-vni/mobile-white-paper-c11-520862.html>
- [2] E. Baştuğ, M. Bennis, and M. Debbah, "Cache-enabled small cell networks: Modeling and tradeoffs," *EURASIP J. Wireless Commun. Netw.*, vol. 2015, no. 41, Feb. 2015.
- [3] J. Hoydis, M. Kobayashi, and M. Debbah, "Green small-cell networks," *IEEE Veh. Technol. Mag.*, vol. 6, no. 1, pp. 37–43, Mar. 2011.
- [4] E. Baştuğ, J.-L. Guenego, and M. Debbah, "Proactive small cell networks," in *Proc. 20th Int. Conf. Telecommun. (ICT)*, May 2013, pp. 1–5.
- [5] K. Shanmugam, N. Golrezaei, A. G. Dimakis, A. F. Molisch, and G. Caire, "FemtoCaching: Wireless content delivery through distributed caching helpers," *IEEE Trans. Inf. Theory*, vol. 59, no. 12, pp. 8402–8413, Dec. 2013.
- [6] N. Golrezaei, A. G. Dimakis, and A. F. Molisch, "Scaling behavior for device-to-device communications with distributed caching," *IEEE Trans. Inf. Theory*, vol. 60, no. 7, pp. 4286–4298, Jul. 2014.
- [7] G. Fodor et al., "Design aspects of network assisted device-to-device communications," *IEEE Commun. Mag.*, vol. 50, no. 3, pp. 170–177, Mar. 2012.
- [8] A. Asadi, Q. Wang, and V. Mancuso, "A survey on device-to-device communication in cellular networks," *IEEE Commun. Surveys Tuts.*, vol. 16, no. 4, pp. 1801–1819, 4th Quart., 2014.
- [9] A. Altieri, P. Piantanida, L. R. Vega, and C. G. Galarza, "A stochastic geometry approach to distributed caching in large wireless networks," in *Proc. 11th Int. Symp. Wireless Commun. Syst. (ISWCS)*, Aug. 2014, pp. 863–867.
- [10] N. Golrezaei, P. Mansourifard, A. F. Molisch, and A. G. Dimakis, "Base-station assisted device-to-device communications for high-throughput wireless video networks," *IEEE Trans. Wireless Commun.*, vol. 13, no. 7, pp. 3665–3676, Jul. 2014.
- [11] J. Paakkonen, C. Hollanti, and O. Tirkkonen, "Device-to-device data storage for mobile cellular systems," in *Proc. IEEE Globecom Workshops*, Dec. 2013, pp. 671–676.
- [12] M. Haenggi, J. G. Andrews, F. Baccelli, O. Dousse, and M. Franceschetti, "Stochastic geometry and random graphs for the analysis and design of wireless networks," *IEEE J. Sel. Areas Commun.*, vol. 27, no. 7, pp. 1029–1046, Sep. 2012.
- [13] N. Giatsoglou, K. Ntontin, E. Kartsakli, A. Antonopoulos, and C. Verikoukis, "D2D-aware device caching in mmWave-cellular networks," *IEEE J. Sel. Areas Commun.*, vol. 35, no. 9, pp. 2025–2037, Sep. 2017.
- [14] D. Malak and M. Al-Shalash, "Optimal caching for device-to-device content distribution in 5G networks," in *Proc. Globecom Workshops*, Dec. 2014, pp. 863–868.
- [15] H. J. Kang, K. Y. Park, K. Cho, and C. G. Kang, "Mobile caching policies for device-to-device (D2D) content delivery networking," in *Proc. IEEE Conf. Comput. Commun. Workshops*, Apr./May 2014, pp. 299–304.
- [16] Z. Zheng, T. Wang, L. Song, Z. Han, and J. Wu, "Social-aware multi-file dissemination in device-to-device overlay networks," in *Proc. IEEE Conf. Comput. Commun. Workshops*, Apr./May 2014, pp. 219–220.
- [17] A. Pyattaev, O. Galinina, S. Andreev, M. Katz, and Y. Koucheryavy, "Understanding practical limitations of network coding for assisted proximate communication," *IEEE J. Sel. Areas Commun.*, vol. 33, no. 2, pp. 156–170, Feb. 2015.
- [18] E. Baştuğ, M. Bennis, and M. Debbah, "Living on the edge: The role of proactive caching in 5G wireless networks," *IEEE Commun. Mag.*, vol. 52, no. 8, pp. 82–89, Aug. 2014.
- [19] Y. Cao and A. Maaref, "Device cooperation-assisted scalable video multicast with heterogeneous QoE guarantees," in *Proc. 11th Int. Symp. Wireless Commun. Syst. (ISWCS)*, Aug. 2014, pp. 733–738.
- [20] Y. Zhang, F. Li, X. Ma, K. Wang, and X. Liu, "Cooperative energy-efficient content dissemination using coalition formation game over device-to-device communications," *Can. J. Elect. Comput. Eng.*, vol. 39, no. 1, pp. 2–10, 2016.
- [21] Y. Xu and F. Liu, "Qos provisionings for device-to-device content delivery in cellular networks," *IEEE Trans. Multimedia*, vol. 19, no. 11, pp. 2597–2608, Nov. 2017.
- [22] X. Lin, R. Ratasuk, A. Ghosh, and J. G. Andrews, "Modeling, analysis, and optimization of multicast device-to-device transmissions," *IEEE Trans. Wireless Commun.*, vol. 13, no. 8, pp. 4346–4359, Apr. 2014.

- [23] M. Afshang, H. S. Dhillon, and P. H. J. Chong, "Fundamentals of cluster-centric content placement in cache-enabled device-to-device networks," *IEEE Trans. Commun.*, vol. 64, no. 6, pp. 2511–2526, Jun. 2016.
- [24] Y. Xu, X. Li, and J. Zhang, "Device-to-device content delivery in cellular networks: Multicast or unicast," *IEEE Trans. Veh. Technol.*, to be published.
- [25] S. Shalmashi, E. Björnson, M. Kountouris, K. W. Sung, and M. Debbah, "Energy efficiency and sum rate when massive MIMO meets device-to-device communication," in *Proc. IEEE Int. Conf. Commun. Workshop (ICCW)*, Jun. 2015, pp. 627–632.
- [26] A. Antonopoulos and C. Verikoukis, "Multi-player game theoretic mac strategies for energy efficient data dissemination," *IEEE Trans. Wireless Commun.*, vol. 13, no. 2, pp. 592–603, Feb. 2014.
- [27] A. Antonopoulos, E. Kartsakli, and C. Verikoukis, "Game theoretic D2D content dissemination in 4G cellular networks," *IEEE Commun. Mag.*, vol. 52, no. 6, pp. 125–132, Jun. 2014.
- [28] E. Datsika, A. Antonopoulos, N. Zorba, and C. Verikoukis, "Green cooperative device-to-device communication: A social-aware perspective," *IEEE Access*, vol. 4, pp. 3697–3707, 2017.
- [29] C. Xu, L. Song, Z. Han, Q. Zhao, X. Wang, and B. Jiao, "Interference-aware resource allocation for device-to-device communications as an underlay using sequential second price auction," in *Proc. IEEE Int. Conf. Commun. (ICC)*, Jun. 2012, pp. 445–449.
- [30] D. Tsoalkas, E. Liotou, N. Passas, and L. Merakos, "A graph-coloring secondary resource allocation for D2D communications in LTE networks," in *Proc. IEEE 17th Int. Workshop Comput. Aided Modeling Design Commun. Links Netw. (CAMAD)*, Sep. 2012, pp. 56–60.
- [31] Y. Xu, Y. Liu, and D. Li, "Resource management for interference mitigation in device-to-device communication," *IET Commun.*, vol. 9, no. 9, pp. 1199–1207, 2015.
- [32] Z. Zhou, M. Dong, K. Ota, J. Wu, and T. Sato, "Energy efficiency and spectral efficiency tradeoff in device-to-device (D2D) communications," *IEEE Wireless Commun. Lett.*, vol. 3, no. 5, pp. 485–488, Oct. 2014.
- [33] M. Afshang, H. S. Dhillon, and P. H. J. Chong, "k-closest coverage probability and area spectral efficiency in clustered D2D networks," in *Proc. IEEE Int. Conf. Commun. (ICC)*, May 2016, pp. 1–6.
- [34] A. A. A. Haija and M. Vu, "Spectral efficiency and outage performance for hybrid d2d-infrastructure uplink cooperation," *IEEE Trans. Wireless Commun.*, vol. 14, no. 3, pp. 1183–1198, Mar. 2015.
- [35] 3GPP WG RAN1, "Tech. Rep. 36.843," 3GPP, Valbonne, France, Tech. Rep. 36.843, 2014. [Online]. Available: <http://www.3gpp.org/ftp/Specs/html-info/36843.htm>
- [36] N. Golrezaei, A. G. Dimakis, and A. F. Molisch, "Wireless device-to-device communications with distributed caching," in *Proc. IEEE Int. Symp. Inf. Theory Proc. (ISIT)*, Jul. 2012, pp. 2781–2785.
- [37] E. Baştuğ, M. Bennis, and M. Debbah, "Social and spatial proactive caching for mobile data offloading," in *Proc. IEEE Int. Conf. Commun. Workshops*, Jun. 2014, pp. 581–586.
- [38] K. Poularakis and L. Tassiulas, "Optimal selfishness-aware device-assisted content delivery in cellular networks," in *Proc. IEEE Wireless Commun. Netw. Conf. (WCNC)*, Apr. 2014, pp. 2288–2293.
- [39] Y. Xu, "On the performance of device-to-device communications with delay constraint," *IEEE Trans. Veh. Technol.*, vol. 65, no. 11, pp. 9330–9344, Nov. 2016.
- [40] H. ElSawy, E. Hossain, and M. S. Alouini, "Analytical modeling of mode selection and power control for underlay D2D communication in cellular networks," *IEEE Trans. Commun.*, vol. 62, no. 11, pp. 4147–4161, Nov. 2014.
- [41] A. H. Sakr and E. Hossain, "Cognitive and energy harvesting-based D2D communication in cellular networks: Stochastic geometry modeling and analysis," *IEEE Trans. Commun.*, vol. 63, no. 5, pp. 1867–1880, May 2015.
- [42] X. Lin, J. G. Andrews, and A. Ghosh, "Spectrum sharing for device-to-device communication in cellular networks," *IEEE Trans. Wireless Commun.*, vol. 13, no. 12, pp. 6727–6740, Dec. 2014.
- [43] Y. Li, Z. Wang, D. Jin, and S. Chen, "Optimal mobile content downloading in device-to-device communication underlying cellular networks," *IEEE Trans. Wireless Commun.*, vol. 13, no. 7, pp. 3596–3608, Jul. 2014.



YANLI XU (M'09) received the Ph.D. degree from the National Mobile Communications Research Laboratory, Southeast University, Nanjing, China, in 2012. From 2012 to 2015, she was a Research Scientist with Bell Labs of Alcatel-Lucent, China, with focus on the standardization of 5G cellular networks. Since 2015, she has been with the Shanghai Maritime University. Her research interests include wireless networks, ad-hoc networks, and green communications, with a current focus on D2D communications. She has served as a guest editor or a TPC member for reputational journals and conferences of communication field.



SHENGMING JIANG (SM'06) received the master's degree in computer science from the Pierreand-Marie-Curie University in 1992 and the Ph.D. degree from the University of Versailles Saint-Quentin-En-Yvelines, France, in 1995. He was a Member of Technical Staff with the Center for Wireless Communications, National University of Singapore, from 1997 to 2000, and a Senior Member of Technical Staff and the Acting Leader of the Network Strategic Group from 2001 to 2003. He has authored over 140 papers on related areas. His research interest includes networking issues in both wired and wireless networks and mobile computing.



JINSONG WU (SM'11) received the Ph.D. degree from the Department of Electrical and Computer Engineering, Queen's University, Kingston, ON, Canada. He is currently with the Department of Electrical Engineering, Universidad de Chile, Santiago, Chile. He was the Founder and the Founding Chair of the IEEE Technical Committee on Green Communications and Computing. He is an elected Vice-Chair for Technical Activities of the IEEE Environmental Engineering Initiative, a pan-IEEE effort under IEEE Technical Activities Board. He is also the Co-Founder and the founding Vice-Chair of the IEEE Technical Committee on Big Data. He received the 2017 IEEE Systems Journal Best Paper Award and three best paper awards for the IEEE conference papers. He received the IEEE Green Communications and Computing Technical Committee 2017 Excellent Services Award for Excellent Technical Leadership and Services in the Green Communications and Computing Community. He has been on editorial boards of the IEEE Series on Green Communication and Computing Networks in the *IEEE Communications Magazine*, the IEEE Transactions on Green Communications and Networking, the IEEE Journal of Selected Areas on Communications Series on Green Communications and Networking, the IEEE Systems Journal, the IEEE Communications Surveys and Tutorials, and the IEEE Access. He was a leading Editor and a co-author of the comprehensive book, entitled *Green Communications: Theoretical Fundamentals, Algorithms, and Applications*, (CRC Press, in 2012).

...