Amer, R., Butt, M. M., ElSawy, H., Bennis, M., Kibiłda, J. and Marchetti, N. (2019) On Minimizing Energy Consumption for D2D Clustered Caching Networks. In: IEEE GLOBECOM 2018, Abu Dhabi, United Arab Emirates, 09-13 Dec 2018, ISBN 9781538647271.

http://eprints.gla.ac.uk/166272/

Deposited on: 7 August 2018

On Minimizing Energy Consumption for D2D Clustered Caching Networks

Ramy Amer*, M. Majid Butt*†, Hesham ElSawy‡, Mehdi Bennis§, Jacek Kibiłda*, Nicola Marchetti*

*CONNECT, Trinity College, University of Dublin, Ireland
†Nokia Bell Labs, France
‡King Abdullah University for Science and Technology (KAUST)
§Centre for Wireless Communications, University of Oulu, Finland
email:{ramyr, majid.butt, kibildj, nicola.marchetti}@tcd.ie, bennis@ee.oulu.fi, hesham.elsawy@kaust.edu.sa

Abstract—We formulate and solve the energy minimization problem for a clustered device-to-device (D2D) network with cache-enabled mobile devices. Devices are distributed according to a Poisson cluster process (PCP) and are assumed to have a surplus memory which is exploited to proactively cache files from a library. Devices can retrieve the requested files from their caches, from neighboring devices in their proximity (cluster), or from the base station as a last resort. We minimize the energy consumption of the proposed network under a random probabilistic caching scheme, where files are independently cached according to a specific probability distribution. A closed form expression for the D2D coverage probability is obtained. The energy consumption problem is then formulated as a function of the caching distribution, and the optimal probabilistic caching distribution is obtained. Results reveal that the proposed caching distribution reduces energy consumption up to 33% as compared to caching popular files scheme.

Index Terms—D2D caching, energy consumption, clustered process.

I. INTRODUCTION

Caching at mobile devices significantly improves system performance by facilitating D2D communications, which enhances the spectral efficiency and alleviates the heavy burden on backhaul links [1]. Modeling the cache-enabled heterogeneous networks, including small cell base station (BS) and mobile devices, follows two main directions in the literature. The first line of work focuses on the fundamental throughput scaling results by assuming a simple protocol channel model [1]–[4], where two devices can communicate if they are within certain distance. The second line of work, which is relevant to our work, considers a more realistic model for the underlying physical layer [5], [6]. We review some of the works relevant to the second line, focusing mainly on the energy efficiency (EE) of wireless caching networks.

EE in wireless caching networks is widely studied in the literature. For example, an optimal caching problem is formulated in [7] to minimize energy consumption in a wireless network. The authors consider a cooperative wireless caching network where relay nodes cooperate with the devices to cache the most popular files in order to minimize energy consumption. In [8], the authors quantify the effect of caching on the EE of wireless access networks. It is shown that EE gain from caching is a function of content popularity, backhaul capacity, and the interference level. However, studies in [7], [8] focus on caching the files at the BSs or access points.

Content placement at the device level in D2D clustered networks is a viable approach to improve the network performance [9]–[11]. The authors in [9] developed a stochastic geometry based model to characterize the performance of cluster-centric content placement in a D2D network. The authors characterize the optimal number of D2D transmitters that must be simultaneously activated in each cluster to maximize the area spectral efficiency. The performance of cluster-centric content placement is characterized in [10], where the content of interest in each cluster is cached closer to the cluster center, such that the collective performance of all the devices in each cluster is optimized. Inspired by the Matern hard-core point process that captures pairwise interactions between nodes, the authors in [11] devised a novel spatially correlated caching strategy called hard-core placement (HCP) such that the D2D devices caching the same content are never closer to each other than the exclusion radius. However, the problem of EE for D2D based caching is not yet addressed in the literature.

In this paper, we aim at minimizing the energy consumption of a clustered cache-enabled D2D network, where devices have unused memory to cache files of interest. Devices are assumed to cache files according to a random probabilistic caching scheme. We consider a PCP where the devices are spatially distributed as groups in clusters. The cluster centers are drawn from a Poisson point process (PPP), and the cluster members are normally distributed around the centers. We derive the D2D coverage probability and calculate the average achievable rate. The energy consumption problem is then formulated and shown to be convex, and the optimal caching distribution is obtained. Results unveil that the obtained optimal probabilistic caching scheme yields the lowest energy consumption among all conventional benchmark caching schemes.

The rest of this paper is organized as follows. Section II and Section III present the system model and energy consumption problem formulation, respectively. The coverage probability analysis is discussed in Section IV and the energy minimization problem is solved in Section V. Numerical results are then presented in Section VI, and Section VII concludes the paper.

II. System Model

A. System Setup

We model the location of the mobile devices with a PCP in which the parent points are drawn from a PPP Φ_p with density λ_p, and the daughter points are normally scattered with variance $\sigma^2 \in \mathbb{R}^2$ around each parent point [12]. The parent points and offspring are referred to as cluster centers and cluster members, respectively. The number of cluster members in each cluster is considered constant, and denoted as n. It is worth highlighting that the well-known Thomas cluster process (TCP), see [13], is similar to the considered PCP but the number of cluster members follows a Poisson distribution. Therefore, our setup can be interpreted as a variant of TCP. Note that the choice of such a variant of TCP is in fact for ease of analysis. The density function of the location of a cluster member relative to its cluster center, $y \in \mathbb{R}^2$, is

$$f_Y(y) = \frac{1}{2\pi\sigma^2} \exp\left(-\frac{\|y\|^2}{2\sigma^2}\right), \tag{1}$$

where $\|.\|$ is the Euclidean norm. The intensity function of a cluster is given by $\lambda_c(y) = \frac{n}{2\pi\sigma^2}\exp\left(-\frac{\|y\|^2}{2\sigma^2}\right)$. Therefore, the overall intensity of the process is given by $\lambda = n\lambda_p$. We assume that the BSs' distribution follows another PPP Φ_{bs} with density λ_{bs}, which is independent of Φ_p.

B. Content Popularity and Probabilistic Caching Placement

We assume that each device has a surplus memory of size M designated for caching files. The total number of files is $N_f > M$ and the set (library) of content indices is denoted as $\mathcal{F} = \{1, 2, \ldots, N_f\}$. These files represent the content catalog that all the devices in a cell may request, which are indexed in a descending order of popularity. The probability that the $i-th$ file is requested follows a Zipf's distribution [14] given by,

$$q_i = \frac{i^{-\beta}}{\sum_{k=1}^{N_f} k^{-\beta}}, \tag{2}$$

where β is a parameter that reflects how skewed the popularity distribution is. For example, if $\beta = 0$, the popularity of the files has a uniform distribution. Increasing β increases the disparity among the files popularity such that lower indexed files have higher popularity. By definition, $\sum_{i=1}^{N_f} q_i = 1$. It is assumed that different clusters may have different interests, and hence, different library content and/or popularity distribution. We use Zipf's distribution to model the popularity of files per cluster.

D2D communication is enabled within each cluster to deliver popular content. A probabilistic caching model is considered, where the content is randomly and independently placed in the cache memories of different devices in the same cluster, according to the same distribution. The probability that a generic device stores a particular file i is denoted as b_i, $0 \leq b_i \leq 1$ for all $i \in \mathcal{F}$. To avoid duplicate caching of the same content within the memory of the same device, we follow the caching approach proposed in [15] as illustrated in Fig. 1.
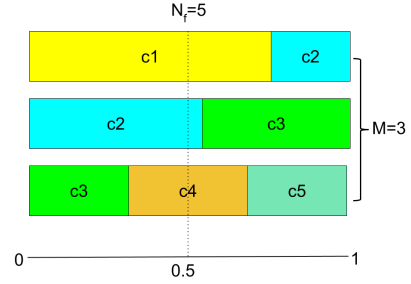

Fig. 1. The cache memory of size $M = 3$ is equally divided into 3 blocks of unit size. Starting from content $i = 1$ to $i = N_f$, each content sequentially fills these 3 memory blocks by an amount b_i. The amounts (probabilities) b_i eventually fill all 3 blocks since $\sum_{i=1}^{N_f} b_i = M$ [15]. Then a random number $\in [0, 1]$ is generated, and a content i is chosen from each block, whose b_i fills the part intersecting with the generated random number. In this way, in the given example, the contents $\{1, 2, 4\}$ are chosen to be cached.

If a device caches the desired file in its own memory, the device directly retrieves the content. However, if the device does not cache the desired file, it downloads the requested file from one neighboring device in the same cluster (that caches the desired file) via D2D communication. Otherwise, the device attaches to the nearest BS as a last resort to download the content which is not cached entirely within the device's cluster. We assume that the D2D communication is operating as out-of-band D2D. Also, it is assumed that all the D2D transmitters transmit with the same power P_d, and all BSs transmit with the same power P_b. We assume that devices' requests in each cluster are served in a round-robin manner where only one active D2D link can be scheduled within each cluster at a time.

III. PROBLEM FORMULATION

In this section, we formulate the energy consumption minimization problem for the clustered D2D caching network. We define the cost c_{d_i} as the time required to download a file i from a neighboring device in the same cluster. Considering the file size S_i of the i-th ranked content, the time c_{d_i} is computed as $c_{d_i} = S_i/\overline{R}_d$, where $\overline{R}_d$ denotes the average rate of intra-cluster D2D communication. Similarly, we have the time $c_{b_i} = S_i/\overline{R}_b$ when the i-th content is served by the BS with average rate $\overline{R}_b$. We assume that the library files have a mean size $\overline{S}$ Mbits. Hence, the total consumed energy for all files downloaded within a given cluster can be written similar to [7] as,

$$E = \sum_{j=1}^{n} \sum_{i=1}^{N_f} q_i \mathcal{P}_{i,j}^d P_d c_{d_i} + q_i \mathcal{P}_{i,j}^b P_b c_{b_i}, \tag{3}$$

where $\mathcal{P}_{i,j}^d$ and $\mathcal{P}_{i,j}^b$ represent the probability of obtaining the i-th file by the j-th device from the local cluster, i.e., via D2D communication, and the BS, respectively. Since we assume that the caching distribution is the same for all devices, the index j is henceforth dropped. For the self-caching probability, it is clear that $\mathcal{P}_i^s = b_i$ since a device

caches the i-th file with probability b_i. For the probability $\mathcal{P}^d_{i,j}$, we define two events:

Event A: The requesting device does not cache file i.

Event B: The requested file is cached by any other device within the same cluster. Then, the probability $\mathcal{P}^d_i$ represents the probability of intersection of these two events, written as

$$\mathcal{P}^d_i = \mathbb{P}\big(A \cap B\big), \tag{4}$$

Since the two events A and B are independent, then

$$\begin{aligned}\mathcal{P}^d_i &= \big(1 - b_i\big)\big(1 - (1 - b_i)^{n-1}\big), \\ &= \big(1 - b_i\big) - (1 - b_i)^n,\end{aligned} \tag{5}$$

where $(1 - b_i)$ is $\mathbb{P}(A)$, and $\big(1 - (1 - b_i)^{n-1}\big)$ is $\mathbb{P}(B)$. Therefore, we have,

$$\mathcal{P}^b_i = (1 - b_i)^n \tag{6}$$

In the following, the achievable rate analysis for the D2D and BS communication is presented. These rates are then used to derive a closed-form expression for the energy consumption in Section V.

IV. AVERAGE ACHIEVABLE RATE

A fixed rate transmission model is adopted in our study, where each TX (D2D or BS) transmits at a fixed rate $\log_2(1 + \theta)$ bits/s/Hz, where θ is a design parameter. Since, the rate is fixed, the transmission is subject to outage due to fading and interference fluctuations. Consequently, the de facto transmissions rate (i.e., throughput) is given by

$$\overline{R}_m = \mathrm{W}_m \log_2(1 + \theta) p_c, \tag{7}$$

where W_m is the allocated bandwidth for the communication link $m \in \{\mathrm{D2D}, \mathrm{BS}\}$, θ is a pre-determined threshold for successful reception, and p_c is the coverage probability. p_c is defined as the probability that the signal to interference plus noise ratio (SINR) of the link of interest at the receiver exceeds the required threshold for successful demodulation and decoding, i.e.,

$$p_c = \mathbb{E}(\mathbf{1}\{\mathrm{SINR} > \theta\}), \tag{8}$$

where $\mathbf{1}\{.\}$ is the indicator function.

A. Average Achievable Rate for D2D communication

Following the methodology in [16], we compute the D2D coverage probability p_{c_d}. Without loss of generality, we conduct the analysis for a cluster whose center is located at $x_0 \in \Phi_p$ (referred to as representative cluster), and the device that requests the content (henceforth called typical device) is located at the origin. We assume that the D2D-TX that caches the requested file is located at y_0 w.r.t. x_0, where $x_0, y_0 \in \mathbb{R}^2$. The distance from the D2D-TX to the typical device (D2D-RX of interest) is denoted as $r = \|x_0 + y_0\|$, which is a realization of a random variable R whose distribution is described later. This setup is illustrated in Fig. 2. The received power at the D2D-RX of interest is expressed as

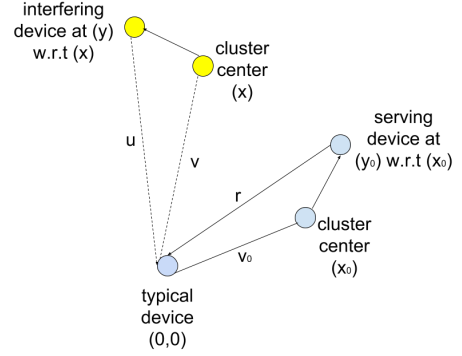$$P = P_d g_0 \|x_0 + y_0\|^{-\alpha} = P_d g_0 r^{-\alpha} \tag{9}$$

Fig. 2. Illustration of the representative cluster and one interfering cluster.

where $g_0 \sim \exp(1)$ is an exponential random variable which models Rayleigh fading and $\alpha > 2$ is the path loss exponent.

It is assumed that there is always one active D2D link per cluster. Similar to (9), the interference from the simultaneously active D2D-TXs outside the representative cluster $x_0 \in \Phi_p$ at the typical device is expressed as

$$I_{\Phi^!_p} = \sum_{x \in \Phi_p \setminus \{x_0\}} P_d g_{yx} \|x + y\|^{-\alpha} \tag{10}$$

$$= \sum_{x \in \Phi_p \setminus \{x_0\}} P_d g_u u^{-\alpha} \tag{11}$$

where y is the marginal distance from the potential interfering device in a given cluster to its cluster center at $x \in \Phi_p$. $u = \|x + y\|$ is a realization of a random variable U modeling the interfering distance (shown in Fig. 2), $g_{yx} \sim \exp(1)$ are independently and identically distributed exponential random variables modeling Rayleigh fading, and $g_u = g_{yx}$ for ease of notation.

Assuming that the thermal noise is neglected compared to inter-cluster interference, the signal to interference ratio (SIR) at the typical device is

$$\mathrm{SIR} = \frac{P}{I_{\Phi^!_p}} = \frac{P_d g_0 r^{-\alpha}}{I_{\Phi^!_p}} \tag{12}$$

For the D2D coverage probability, we proceed as follows,

$$\begin{aligned}p_{c_d} &= \mathbb{P}[\mathrm{SIR} > \theta] = \mathbb{E}_R \mathbb{P}\Big[\frac{g_0 P_d r^{-\alpha}}{I_{\Phi^!_p}} > \theta | R = r\Big] \\ &= \int_{r>0} \mathbb{P}\Big[\frac{g_0 P_d r^{-\alpha}}{I_{\Phi^!_p}} > \theta\Big] f_R(r) \mathrm{d}r \\ &= \int_{r>0} \mathbb{P}\Big[g_0 > \theta r^\alpha \frac{I_{\Phi^!_p}}{P_d}\Big] f_R(r) \mathrm{d}r,\end{aligned} \tag{13}$$

where $f_R(r)$ is the serving distance distribution. Since $g_0 \sim \exp(1)$, we have

$$\begin{aligned}p_{c_d} &= \int_{r>0} \mathbb{E}\Big[\exp\big(\frac{-\theta r^\alpha}{P_d} I_{\Phi^!_p}\big)\Big] f_R(r) \mathrm{d}r, \\ &\overset{(a)}{=} \int_{r>0} \mathscr{L}_{I_{\Phi^!_p}}\big(s = \frac{\theta r^\alpha}{P_d}\big) f_R(r) \mathrm{d}r,\end{aligned} \tag{14}$$

where (a) follows from the Laplace transform definition. Now, we proceed with the derivation of the Laplace transform of the inter-cluster interference to obtain the D2D coverage probability.

Lemma 1. *Laplace transform of the inter-cluster interference can be expressed as*

$$\mathscr{L}_{I_{\Phi_p^!}}(s) = \exp\Big(- \pi\lambda_p(sP_d)^{2/\alpha}\Gamma(1+2/\alpha)\Gamma(1-2/\alpha)\Big) \tag{15}$$

Proof. Please see Appendix A. $\square$

Lemma 2. *Substituting the obtained Laplace transform of the inter-cluster interference (15) into (14) yields the D2D coverage probability*

$$p_{c_d} = \frac{1}{4\sigma^2 \mathcal{Z}(\theta,\alpha,\sigma)}, \tag{16}$$

where $\mathcal{Z}(\theta,\alpha,\sigma) = (\pi\lambda_p\theta^{2/\alpha}\Gamma(1+2/\alpha)\Gamma(1-2/\alpha) + \frac{1}{4\sigma^2})$

Proof. Please see Appendix B. $\square$

Remark 1. The Laplace transform of interference of the PCP with one active D2D link per cluster in (15) is similar to that of the PPP [16]. This can be explained with the displacement theory of the PPP [17], as each interferer is a point of a PPP that is displaced randomly and independently of all other points.

Remark 2. The obtained D2D coverage probability in (16) for the PCP with one active D2D link per cluster shows two important insights when compared to an equivalent PPP. Firstly, it depends on the density of the parent PPP λ_p which directly affects the interfering distances. However, in the PPP, the effect of device density does not exist since it equally affects both the serving and interfering distances when a device associates to its nearest serving device. Secondly, it depends on the serving distance which is represented by the displacement standard deviation σ.

B. Average Achievable Rate for Cellular Communication

For the assumed PPP model for the BSs' distribution, and based on the nearest BS association principle, it is shown in [18] that the BS coverage probability can be expressed as

$$p_{c_b} = \frac{1}{{}_2F_1(1,-\delta;1-\delta;-\theta)}, \tag{17}$$

where ${}_2F_1(.)$ is the Gaussian hypergeometric function and $\delta = 2/\alpha$. By substituting the obtained coverage probabilities from (16) and (17) into (7), $\overline{R}_d$ and $\overline{R}_b$ are derived.

V. ENERGY CONSUMPTION MINIMIZATION

The energy minimization problem is formulated as

$$\min_{b_i} \quad E = n\sum_{i=1}^{N_f} q_i \mathcal{P}_i^d P_d \frac{S_i}{\overline{R}_d} + q_i \mathcal{P}_i^b P_b \frac{S_i}{\overline{R}_b} \tag{18}$$

$$\text{s.t.} \quad \sum_{i=1}^{N_f} b_i = M, \qquad b_i \in [0,1], \tag{19}$$

where (19) is the device cache size constraint to avoid duplicate caching of the same file (cf. Fig. 1). It is easy to show the convexity of the objective function E by confirming that the Hessian matrix of E w.r.t. the caching variables, $\mathbf{H}_{i,j} = \frac{\partial^2 E}{\partial b_i \partial b_j}$, $\forall i,j \in \mathcal{F}$, is positive semidefinite (and hence E is convex). Utilizing the convexity of E, the optimal caching distribution is computed as follows.

Lemma 3. *The optimal caching distribution b_i^* is given by*

$$b_i^* = \left[1 - \Big(\frac{v^* + k^2 q_i S_i \frac{P_d}{\overline{R}_d}}{k q_i S_i \big(\frac{P_d}{\overline{R}_d} - \frac{P_b}{\overline{R}_b} \big)} \Big)^{\frac{1}{n-1}} \right]^+ \tag{20}$$

where v^* satisfies the maximum cache constraint $\sum_{i=1}^{N_f} \Big[1 - \Big(\frac{v^* + k^2 q_i S_i \frac{P_d}{\overline{R}_d}}{k q_i S_i \big(\frac{P_d}{\overline{R}_d} - \frac{P_b}{\overline{R}_b} \big)} \Big)^{\frac{1}{n-1}} \Big]^+ = M$, and $[x]^+ = max(x,0)$.

Proof. Due to lack of space, we delegate this proof to the journal version of this paper [19]. $\square$

VI. NUMERICAL RESULTS

TABLE I
SIMULATION PARAMETERS

Description	Parameter	Value
BS-to-Device bandwidth	W_{bs}	20 MHz
D2D bandwidth	W_{d2d}	20 MHz
BS transmission power	P_b	43 dB$_m$
D2D transmission power	P_d	23 dB$_m$
Displacement standard deviation	σ	10 m
Popularity index	β	1
Path loss exponent	α	4
Library size	N_f	500 files
Cache size per device	M	10 files
Devices per cluster	n	10
Density of Φ_p	λ_p	50 clusters/km^2
Average content size	$\overline{S}$	100 Mbits
SIR threshold	θ	0 dB

At first, we validate the developed mathematical model via Monte Carlo simulations. Then we benchmark the proposed caching scheme against conventional caching schemes. Unless otherwise stated, the network parameters are selected as shown in Table I. In Fig. 3, we verify the accuracy of the analytical results for the D2D coverage probability. The close matching between the analytical and simulated results validates the developed mathematical model. We see that the coverage probability monotonically decreases with σ. As σ increases, the serving distance increases and the distance between the interferers and the typical device decreases, and equivalently, the complementary cumulative distribution function (CCDF) of SIR decreases. Similarly, it is observed that the D2D coverage probability decreases with θ owing to the decreasing CCDF of SIR.

Fig. 4 shows the normalized energy consumption per device versus β under different caching schemes, namely, proposed probabilistic caching (PC), random uniform caching (RC), caching popular files (CPF), and Zipf's caching (Zipf). We can see that the minimized consumed energy under PC scheme
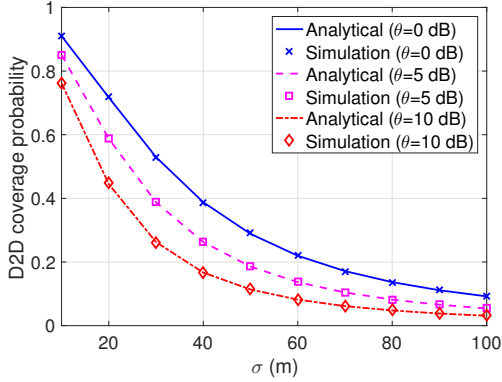
Fig. 3. D2D coverage probability versus displacement standard deviation σ.

Fig. 5. Normalized energy consumption versus the number of devices per cluster.

attains the best performance as compared to other schemes. Also, it is clear that, except for the RC, the consumed energy decreases with β. This can be justified by the fact that as β increases, fewer files are frequently requested which are more likely to be cached under PC, CPF, and the Zipf's caching schemes. In the RC scheme, files are uniformly chosen for caching independently of their popularity.

We plot the average device energy consumption per device versus number of devices per cluster in Fig. 5. First, we see that the normalized energy consumption decreases with the number of devices. As the number of devices per cluster increases, it is more probable to obtain requested files via low power D2D communication. When the number of devices per cluster is relatively large, the normalized energy consumption tends to flatten as most of the content becomes cached at the cluster devices.
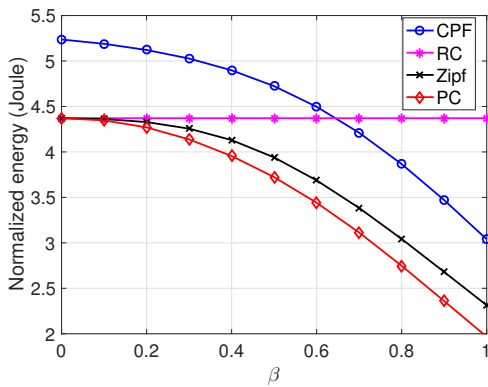

Fig. 4. Normalized energy consumption versus popularity exponent β.

VII. CONCLUSION

In this work, we formulate and solve an optimization problem to minimize the energy consumption of a clustered D2D network with random PC incorporated at the devices. Using tools from stochastic geometry, we get closed form expressions of the coverage probability and transmission rates, encompassing the underlying physical layer parameters, e.g.
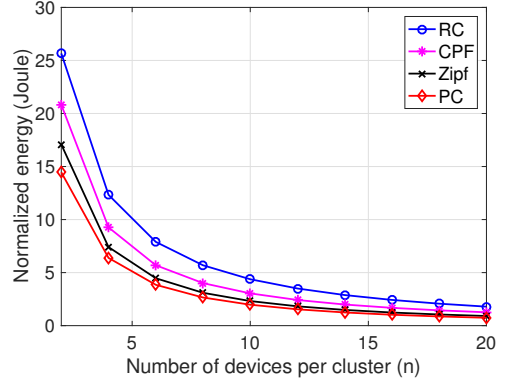
mutual interference and SIR distribution. We show the convexity of the formulated energy minimization problem and obtain a closed form solution for the file placement probabilities. Results reveal that the proposed PC scheme significantly reduce the consumed energy in the network when compared to conventional methods, e.g., CPF, RC and Zipf's caching.

APPENDIX A
PROOF OF LEMMA 1

Laplace transform of the inter-cluster aggregate interference $I_{\Phi_p^!}$ is expressed as

$$
\begin{aligned}
\mathscr{L}_{I_{\Phi_p^!}}(s) &= \mathbb{E}_{\Phi_p, g_u}\left[e^{-sP_d \sum_{\Phi_p \setminus \{x_0\}} g_u u^{-\alpha}} \right] \\
&\overset{(a)}{=} \mathbb{E}_{\Phi_p}\left[\mathbb{E}_{g_u}\left[\prod_{x \in \Phi_p \setminus \{x_0\}} e^{-sP_d g_u u^{-\alpha}} \right] \right] \\
&\overset{(b)}{=} \exp\left(-2\pi\lambda_p \mathbb{E}_{g_u} \int_{v=0}^{\infty} \mathbb{E}_{u|v}\left[1 - e^{-sP_d g_u u^{-\alpha}} \right] v \, dv \right),
\end{aligned}
\tag{21}
$$

where (a) follows from the independence between devices' locations and the channel gains, (b) follows from the probability generating functional of the PPP [16]. It is proven in [9] that the interfering distance u, conditioned on the distance $v = \|x\|$ between the interfering cluster center $x \in \Phi_p$ and the typical device is given by $f_U(u|v) = \text{Rice}(u|v, \sigma)$, which represents the Rice's probability density function of parameter σ. By averaging over u, the Laplace transform of inter-cluster

interference can be expressed as

$$\mathscr{L}_{I_{\Phi_p^!}}(s) = \exp\Big(- 2\pi\lambda_p \mathbb{E}_{g_u} \Big[\int_{v=0}^{\infty} \int_{u=0}^{\infty} \big(1 - e^{-sP_d g_u u^{-\alpha}}\big)$$
$$\cdot f_U(u|v)\,\mathrm{d}u\,v\,\mathrm{d}v \Big] \Big)$$

$$\stackrel{(a)}{=} \exp\Big(- 2\pi\lambda_p \cdot$$

$$\mathbb{E}_{g_u} \underbrace{\int_{v=0}^{\infty} v\,\mathrm{d}v - \int_{v=0}^{\infty}\int_{u=0}^{\infty} e^{-sP_d g_u u^{-\alpha}} f_U(u|v)\,\mathrm{d}u\,v\,\mathrm{d}v}_{\mathcal{R}(s,\alpha)} \Big) \tag{22}$$

where (a) follows from $\int_{u=0}^{\infty} f_U(u|v)\,\mathrm{d}u = 1$. Now, we proceed by calculating the integrands of $\mathcal{R}(s,\alpha)$ as follows.

$$\mathcal{R}(s,\alpha) \stackrel{(a)}{=} \int_{v=0}^{\infty} v\,\mathrm{d}v - \int_{u=0}^{\infty} e^{-sP_d g_u u^{-\alpha}} \cdot$$
$$\int_{v=0}^{\infty} f_U(u|v)v\,\mathrm{d}v\,\mathrm{d}u$$
$$\stackrel{(b)}{=} \int_{v=0}^{\infty} v\,\mathrm{d}v - \int_{u=0}^{\infty} e^{-sP_d g_u u^{-\alpha}} u\,\mathrm{d}u$$
$$\stackrel{(c)}{=} \int_{u=0}^{\infty} (1 - e^{-sP_d g_u u^{-\alpha}}) u\,\mathrm{d}u$$
$$\stackrel{(d)}{=} \frac{(sP_d g_u)^{2/\alpha}}{\alpha} \int_{u=0}^{\infty} (1-e^{-t}) t^{-1-\frac{2}{\alpha}}\,\mathrm{d}u$$
$$\stackrel{(e)}{=} \frac{(sP_d)^{2/\alpha}}{2} g_u^{2/\alpha} \Gamma(1+2/\alpha), \tag{23}$$

where (a) follows from changing the order of integration, (b) follows from $\int_{v=0}^{\infty} f_U(u|v)v\,\mathrm{d}v = u$, (c) follows from changing the dummy variable u to v, (d) follows from changing the variables $t = sg_u u^{-\alpha}$, and (e) follows from solving the integration of (d) by parts. Substituting the obtained value for $\mathcal{R}(s,\alpha)$ into (22), and taking the expectation over the exponential random variable g_u, with the fact that $\mathbb{E}_{g_u} g_u^{2/\alpha} = \Gamma(1-2/\alpha)$, the lemma is proven.

APPENDIX B
PROOF OF LEMMA 2

Since the serving and typical devices' locations w.r.t. the cluster center x_0 are sampled from normal distribution with variance σ^2, the serving distance r is Rayleigh distributed with probability density function $f_R(r) = \mathrm{Rayleigh}(r, \sqrt{2}\sigma)$ with a scale parameter $\sqrt{2}\sigma$. Substituting the obtained Laplace transform of (15) and the probability density function $f_R(r)$ into (14) yields,

$$p_{c_d} = \int_{r=0}^{\infty} \mathrm{e}^{-\pi\lambda_p(sP_d)^{2/\alpha}\Gamma(1+2/\alpha)\Gamma(1-2/\alpha)} \frac{r}{2\sigma^2} \mathrm{e}^{\frac{-r^2}{4\sigma^2}}\,\mathrm{dr},$$
$$\stackrel{(a)}{=} \int_{r=0}^{\infty} \frac{r}{2\sigma^2} \mathrm{e}^{-\pi\lambda_p \theta^{2/\alpha} r^2 \Gamma(1+2/\alpha)\Gamma(1-2/\alpha)} \mathrm{e}^{\frac{-r^2}{4\sigma^2}}\,\mathrm{dr},$$
$$\stackrel{(b)}{=} \int_{r=0}^{\infty} \frac{r}{2\sigma^2} \mathrm{e}^{-r^2 \mathcal{Z}(\theta,\sigma,\alpha)}\,\mathrm{dr},$$
$$= \frac{1}{4\sigma^2 \mathcal{Z}(\theta,\alpha,\sigma)} \tag{24}$$

where (a) comes from the substitution $s = \frac{\theta r^\alpha}{P_d}$, and (b) from $\mathcal{Z}(\theta,\alpha,\sigma) = (\pi\lambda_p \theta^{2/\alpha}\Gamma(1+2/\alpha)\Gamma(1-2/\alpha) + \frac{1}{4\sigma^2})$.

REFERENCES

[1] K. Shanmugam *et al.*, "Femtocaching: Wireless content delivery through distributed caching helpers," *IEEE Transactions on Information Theory*, vol. 59, no. 12, pp. 8402–8413, 2013.

[2] N. Golrezaei *et al.*, "Base-station assisted device-to-device communications for high-throughput wireless video networks," *IEEE Transactions on Wireless Communications*, vol. 13, no. 7, pp. 3665–3676, 2014.

[3] R. Amer, M. M. Butt, M. Bennis, and N. Marchetti, "Delay analysis for wireless D2D caching with inter-cluster cooperation," in *IEEE Global Communications Conference (GLOBECOM)*, Singapore, 2017, pp. 1–7.

[4] ——, "Inter-cluster cooperation for wireless D2D caching networks," *IEEE Transactions on Wireless Communications*, pp. 1–1, 2018.

[5] S. Andreev *et al.*, "Analyzing assisted offloading of cellular user sessions onto D2D links in unlicensed bands," *IEEE Journal on Selected Areas in Communications*, vol. 33, no. 1, pp. 67–80, 2015.

[6] 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.

[7] C. Yang *et al.*, "Energy efficiency in wireless cooperative caching networks," in *IEEE International Conference on Communications (ICC)*, Sydney, Australia, 2014.

[8] D. Liu and C. Yang, "Energy efficiency of downlink networks with caching at base stations," *IEEE Journal on Selected Areas in Communications*, vol. 34, no. 4, pp. 907–922, 2016.

[9] M. Afshang, H. S. Dhillon, and P. H. J. Chong, "Modeling and performance analysis of clustered device-to-device networks," *IEEE Transactions on Wireless Communications*, vol. 15, no. 7, pp. 4957–4972, 2016.

[10] ——, "Fundamentals of cluster-centric content placement in cache-enabled device-to-device networks," *IEEE Transactions on Communications*, vol. 64, no. 6, pp. 2511–2526, 2016.

[11] D. Malak, M. Al-Shalash, and J. G. Andrews, "Spatially correlated content caching for device-to-device communications," *IEEE Transactions on Wireless Communications*, vol. 17, no. 1, pp. 56–70, Jan 2018.

[12] M. Haenggi, *Stochastic geometry for wireless networks*. Cambridge University Press, 2012.

[13] R. K. Ganti and M. Haenggi, "Interference and outage in clustered wireless ad hoc networks," *IEEE Transactions on Information Theory*, vol. 55, no. 9, pp. 4067–4086, 2009.

[14] L. Breslau, P. Cao, L. Fan, G. Phillips, and S. Shenker, "Web caching and zipf-like distributions: Evidence and implications," in *Eighteenth Annual Joint Conference of the IEEE Computer and Communications Societies (INFOCOM)*, vol. 1, 1999, pp. 126–134.

[15] B. Blaszczyszyn and A. Giovanidis, "Optimal geographic caching in cellular networks," in *IEEE International Conference on Communications (ICC)*, 2015, pp. 3358–3363.

[16] J. G. Andrews *et al.*, "A tractable approach to coverage and rate in cellular networks," *IEEE Transactions on Communications*, vol. 59, no. 11, pp. 3122–3134, 2011.

[17] D. J. Daley and D. Vere-Jones, *An introduction to the theory of point processes: volume II: general theory and structure*. Springer Science & Business Media, 2007.

[18] R. K. Ganti and M. Haenggi, "Asymptotics and approximation of the SIR distribution in general cellular networks," *IEEE Transactions on Wireless Communications*, vol. 15, no. 3, pp. 2130–2143, 2016.

[19] R. Amer, M. M. Butt, H. ElSawy, M. Bennis, and N. Marchetti, "Optimizing joint probabilistic caching and communication for clustered D2D networks," *Submitted to IEEE Transactions on Wireless Communications*, 2018.