This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TVT.2018.2797047, IEEE Transactions on Vehicular Technology

Performance Analysis of Cache-Enabled Millimeter Wave Small Cell Networks

Yongxu Zhu, Gan Zheng, *Senior Member, IEEE*, Kai-Kit Wong, *Fellow, IEEE*, Shi Jin, *Member, IEEE*, and Sangarapillai Lambotharan, *Senior Member, IEEE*

***Abstract*—Millimeter wave (mmWave) small-cell networks can provide high regional throughput, but the backhaul requirement has become a performance bottleneck. This paper proposes a hybrid system that combines traditional backhaul-connected small base stations (SBSs) and cache-enabled SBSs to achieve the maximum area spectral efficiency (ASE) while saving backhaul consumption in mmWave small cell networks. We derive and compare the ASE results for both the traditional and hybrid networks, and also show that the optimal content placement to maximize ASE is to cache the most popular contents. Numerical results demonstrate the performance improvement of deploying cache-enabled SBSs. Furthermore, given a total caching capacity, it is revealed that there is a tradeoff between the cache-enabled SBSs density and individual cache size to maximize the ASE.**

***Index Terms*—Millimeter wave, cache-enabled network, backhaul, area spectral efficiency**

I. Introduction

Deploying small base stations (SBS) using the millimeter wave (mmWave) band is a promising solution to meet the growing demand for data and support emerging technologies in next generation mobile communications. However, the expensive high-capacity backhaul links have become a major challenge for realizing the benefits of small cell networks. Based on the observation that mobile data are dominated by multimedia contents and especially videos which are reusable, edge caching at SBSs is a viable solution that will minimize unnecessary data transport costs, latency and increase the effective system throughput [1, 2].

Proactive caching at mmWave SBSs for video streaming has been studied in [3] and it is shown to significantly reduce the handoff delays and connection latency. Area spectral efficiency (ASE) defined as average throughput per unit area, which is one of the most visualized performance metrics to analyze the impact of the path loss, shadowing, and multipath fading simultaneously for a radio mobile communication link than spectral efficiency, because it also considering the congestion and load within the cell area. [4, 5]. The ASE of cache-enabled heterogeneous networks (HetNets) in the sub-6 GHz band has been investigated in [6] where helper nodes with large storage but no backhaul are introduced. It is found that there is an optimal helper density to achieve the maximum ASE. A coded caching scheme was investigates in a large-scale small-cell network in [7], where each SBS stores a certain number of different coded packets of each file and the typical user to be associated with the set of nearest SBSs to obtain the required number of different packets for decoding.The deployment of unmanned aerial vehicles (UAVs) and cache placement at UAVs are studied in [8] based on predicted user requests in a cloud radio access network.

There have been very few works towards understanding the benefits and design tradeoffs in cache-enabled mmWave small cell networks which have distinctive channel characteristics. In this correspondence, we introduce a hybrid downlink mmWave small cell network, in which cache-enabled SBSs will replace some traditional backhual-connected SBSs, to achieve the benefits of both cache and backhaul. We adopt stochastic geometry to analyze and compare the ASEs of the traditional network with the proposed hybrid mmWave network including both the line of sight (LoS) and non-line of sight (NLoS) components. The main contributions of this paper are summarized as follows:

- We derive the analytical results of ASEs for both traditional SBSs and cache-enabled SBSs in mmWave networks;
- We prove that storing the most popular contents (MPC) at each cache-enabled SBS will achieve the maximum ASE performance;
- It is shown that given the total caching capacity, the cache-enabled SBSs density needs to be carefully optimized to combine the benefits of cache and backhaul.

II. System Description

A. Hybrid mmWave Small Cell Network Layout

The proposed hybrid downlink cache-enabled mmWave small cell cellular network is illustrated in Fig. 1. We define the path loss model including both LoS and NLoS links. Blockages effects are modeled as random process of rectangles in the outdoor scenario, where all mmwave SBSs are distributed as independent homogeneous Poisson point processes (PPPs) with density λ_o, denoted as $\Phi_{\mathcal{O}}$, which is composed of the

Manuscript received June 23, 2017; revised September 18, 2018; accepted January 14, 2018. This work was supported by the U.K. Engineering and Physical Sciences Research Council (EPSRC) under Grant EP/N007840/1 and EP/N008219/1. The work of S. Jin was supported by the National Natural Science Foundation of China under Grant 61531011. The review of this paper was coordinated by Dr. Z. Cai. (*Corresponding author: Yongxu Zhu.*)

Y. Zhu, G. Zheng and S. Lambotharan are with the Wolfson School of Mechanical, Electrical and Manufacturing Engineering, Loughborough University, Leicestershire, LE11 3TU, UK (Email: {y.zhu4, g.zheng, S.Lambotharan}@lboro.ac.uk).

K.-K. Wong are with the Department of Electronic and Electrical Engineering, University College London, London, WC1E 6BT, UK (Email: kai-kit.wong@ucl.ac.uk).

S. Jin is with the National Mobile Communications Research Laboratory, Southeast University, Nanjing 210096, P. R. China (E-mail:jinshi@seu.edu.cn).

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TVT.2018.2797047, IEEE Transactions on Vehicular Technology

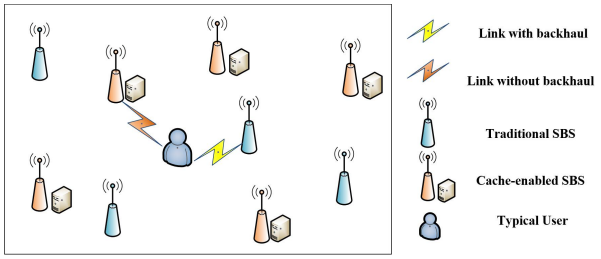

Fig. 1. System model.

LoS region $\Phi_{\mathcal{O}}^{\mathrm{L}}$ and the NLoS region $\Phi_{\mathcal{O}}^{\mathrm{N}}$, respectively [9]. We model the millimeter wave channel without small scale fading because of the sparse scattering mmWave environment [10]. Cache is deployed at a portion of SBSs denoted as $\Phi_{\mathcal{C}}$ with density $\eta\lambda_o$ to replace backhaul, where $0 \leq \eta \leq 1$ is the cache-enabled SBSs density factor. The remaining traditional SBSs are connected to backhaul with limited capacity, denoted as $\Phi'_{\mathcal{O}}$ ($\Phi_{\mathcal{C}} \cup \Phi'_{\mathcal{O}} = \Phi_{\mathcal{O}}$) with density $(1-\eta)\lambda_o$. Notice that $\Phi'_{\mathcal{O}} = \Phi_{\mathcal{O}}$ when $\eta = 0$ and $\Phi_{\mathcal{C}} = \Phi_{\mathcal{O}}$ when $\eta = 1$.

We use a sectored antenna pattern to approximate the mmWave beam, and the antenna gain is expressed as

$$G_{\ell k}^{tr} = \begin{cases} G_{\mathrm{M}}^{t} G_{\mathrm{M}}^{r}, & \Pr_{\mathrm{MM}}^{tr} = \frac{\theta\vartheta}{(2\pi)^2} \\ G_{\mathrm{M}}^{t} G_{\mathrm{m}}^{r}, & \Pr_{\mathrm{Mm}}^{tr} = \frac{(2\pi-\theta)\vartheta}{(2\pi)^2} \\ G_{\mathrm{m}}^{t} G_{\mathrm{M}}^{r}, & \Pr_{\mathrm{mM}}^{tr} = \frac{\theta(2\pi-\vartheta)}{(2\pi)^2} \\ G_{\mathrm{m}}^{t} G_{\mathrm{m}}^{r}, & \Pr_{\mathrm{mm}}^{tr} = \frac{(2\pi-\theta)(2\pi-\vartheta)}{(2\pi)^2} \end{cases}, \quad (1)$$

where $G_{lk}^{tr} = G_{\ell}^{t}(\theta) G_{k}^{r}(\vartheta)$ $(\ell, k \in \{\mathrm{M}, \mathrm{m}\})$ denotes the total directivity gain with beamwidth θ and ϑ in the links from the transmitter and the receiver, respectively. $G_{\mathrm{M}}^{r\&t}$ is the main lobe array gain with probability $\Pr_{\mathrm{M}}^{r\&t} = \frac{\upsilon}{2\pi}$ $(\upsilon \in \{\theta, \vartheta\})$ and $G_{\mathrm{m}}^{r\&t}$ is the sidelobe array gain with probability $\Pr_{\mathrm{m}}^{r\&t} = \frac{2\pi-\upsilon}{2\pi}$. $\Pr_{\ell k}^{tr}$ denotes the probability with the antenna gain G_{ℓ}^{t} and G_{k}^{r}.

B. Caching Placement Strategies

We assume that users request contents from the library $\mathcal{F} = \{f_1, f_2, ..., f_J\}$ randomly, where the content library cardinality is J. Each file j is requested according to the probability a_j. We assume that the request distribution for each content f_j is a_j where $a_1 > a_2 > \cdots > a_J$, and $\sum_{j=1}^{J} a_j = 1$. Assume that all contents have the same size normalized to 1. Each SBS has a cache capacity of K contents ($1 < K \ll J$), and can store up to $X \in [K, J]$ contents with the random caching probability which is denoted as b_j for the j-th content, $X = K$ is a special case when only the most popular contents are stored at each SBS. Therefore, we have $\sum_{j=1}^{X} b_j \leq K$. The remaining uncached contents are served by the traditional SBSs via backhaul.

$\Phi_{\mathcal{C},j}$ denotes the set of the cache-enabled SBSs with density $b_j\eta\lambda_o$ who can support the typical user when the j-th content is requested, and $\Phi_{\mathcal{C}} = \cup_{j\in J}\Phi_{\mathcal{C},j}$.

C. Transmission Scheme

Following the analytical framework of stochastic geometry, we consider the user association based on the closest distance to the active SBS [9]. To minimize the backhaul usage, we assume the typical user is associated with a cache-enabled SBS, as long as the requested file can be found in the cache; otherwise, the user is associated with the nearest traditional SBS. In the following, we will define the signal-to-interference-plus-noise ratios (SINRs) for different scenarios which are necessary for the ASE analysis.

***Definition** 1:* When the typical user is associated with a traditional mmWave SBS with only the backhaul link, the downlink SINR is given by

$$\mathrm{SINR}'_o = \frac{P_t G_{\mathrm{M}}^{t} G_{\mathrm{M}}^{r} L\left(|X_o|\right)}{\mathcal{I}_o + \overline{\mathcal{I}}_o + \sigma_o^2}, \quad (2)$$

where the P_t is transmit power which is the same for all SBSs, G_{M}^{t} and G_{M}^{r} is the associated SBS and the typical user's maximum array gain. We employ a short-range propagation model with corresponding distance $|X_o|$ from the associated o-th SBS to the typical user. The path loss function is denoted as $L\left(|X_o|\right) = \beta|X_o|^{-\alpha_o}$, where β is the frequency dependent constant parameter, and α_o is the path loss exponent given by $\alpha_o = \alpha_{\mathrm{L}}$ for the LoS area and $\alpha_o = \alpha_{\mathrm{N}}$ for the NLoS area, respectively. We use the approximation of an equivalent LoS ball [9] and define LoS reference area less than D, i.e., the link is considered LoS if $|X_o| \leq D$, and NLoS otherwise. Note that we neglect the small-scale fading since it has limited impact due to the lack of spatial scattering [11]. We have the interference from the traditional SBSs $\mathcal{I}_o = \sum_{i\in\Phi'_{\mathcal{O}}\setminus o} P_t G_{\ell}^{t} G_{k}^{r} L\left(|X_i|\right)$ and the interference from cache-enabled SBSs $\overline{\mathcal{I}}_o = \sum_{g\in\Phi_{\mathcal{C}}} P_t G_{\ell}^{t} G_{k}^{r} L\left(|X_g|\right)$ in (2) where $|X_i|$ and $|X_g|$ are the distances between the typical user and the traditional SBS i, and the cache-enabled SBS g, respectively. When $\eta = 0$, $\mathcal{I}_o = \sum_{i\in\Phi_{\mathcal{O}}\setminus o} P_t G_{\ell}^{t} G_{k}^{r} L\left(|X_i|\right)$, and $\overline{\mathcal{I}}_o = 0$. σ_o^2 denotes the noise power.

***Definition** 2:* When the typical user is associated with a cache-enabled SBS and requests the j-th content, the downlink SINR is defined as:

$$\mathrm{SINR}_{c,j} = \frac{P_t G_{\mathrm{M}}^{t} G_{\mathrm{M}}^{r} L\left(|X_{n,j}|\right)}{\mathcal{I}_{c,j} + \overline{\mathcal{I}}_{c,j} + \mathcal{I}'_o + \sigma_o^2}, \quad (3)$$

where $\mathcal{I}_{c,j} = \sum_{h\in\Phi_{\mathcal{C},j}\setminus\{n\}} P_t G_{\ell}^{t} G_{k}^{r} L\left(|X_h|\right)$ is the sum interference from the active cache-enabled SBSs who store the j-th content in $\Phi_{\mathcal{C},j}$, $\overline{\mathcal{I}}_{c,j} = \sum_{g\in\Phi_{\mathcal{C}}\setminus\{\Phi_{\mathrm{C},j}\}} P_t G_{\ell}^{t} G_{k}^{r} L\left(|X_g|\right)$ is the sum interference from the active cache-enabled SBSs who do not store the j-th content, and $\mathcal{I}'_o = \sum_{i\in\Phi'_{\mathcal{O}}} P_t G_{\ell}^{t} G_{k}^{r} L\left(|X_i|\right)$ is the sum interference from the active traditional SBSs without cache. $|X_h|$, $|X_g|$, and $|X_i|$ denote the distances between the typical user and the corresponding interfering SBSs in the above three categories, respectively.

We use ASE as the performance indicator to measure the average network capacity which is defined as the average throughput per unit spectrum and area [4, 6]. For the traditional mmWave small cell networks without caching, ASE can be obtained as

$$\mathcal{A}_o = \lambda_o \mathcal{R}_{\mathcal{O}}, \quad (4)$$

where $R_{\mathcal{O}} = \min\left\{\mathbb{E}\left[\log_2(1+\mathrm{SINR}_o)\right], \mathcal{R}_{\mathcal{B}}\right\}$, SINR_o is the special case of SINR'_o in (2) when $\eta = 0$, $\mathcal{R}_{\mathcal{B}}$ is the backhaul capacity limit and $\mathbb{E}[\cdot]$ denotes the expectation operation.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TVT.2018.2797047, IEEE Transactions on Vehicular Technology

For the proposed hybrid systems with mixed cache-enabled SBSs and traditional SBSs, ASE can be obtained as

$$\mathcal{A}_c = \underbrace{\sum_{j=1}^{X} a_j \left(b_j \eta \lambda_o\right) \mathcal{R}_{\mathcal{C},j}}_{\mathcal{A}_c^{\mathrm{C}}} + \underbrace{\sum_{j=X+1}^{J} a_j (1-\eta)\lambda_o \mathcal{R}'_{\mathcal{O}}}_{\mathcal{A}_c^{\mathrm{B}}}, \quad (5)$$

where $\mathcal{A}_c^{\mathrm{C}}$ denotes ASE of cache-enabled SBSs aggregated in $\Phi_{\mathcal{C}}$. $R_{\mathcal{C},j} = \mathbb{E}\left[\log_2(1+\mathrm{SINR}_{c,j})\right]$ is the achievable rate of the j-th requested content. $R'_{\mathcal{O}} = \min\left\{\mathbb{E}\left[\log_2(1+\mathrm{SINR}'_{\mathrm{o}})\right], R_{\mathcal{B}}\right\}$ is the achievable rate of the traditional network with limited backhaul capacity.

III. Performance Analysis

In this section, in order to derive the ASEs of the hybrid cache-enabled network and the traditional network, we first derive the achievable rate of $\mathcal{R}_{\mathcal{C},j}$ and $\mathcal{R}'_{\mathcal{O}}$, respectively. Then the optimal content placement probability the maximizes that ASE is proved.

A. Achievable Rate Through Cache-Enabled SBSs

***Theorem** 1:* The average delivery rate when the typical user requests the j-th content and is associated with a cache-enabled SBS is given by:

$$\mathcal{R}_{\mathcal{C},j} = \frac{1}{\ln 2}\int_0^{\infty}\int_0^{D} \frac{1}{z}\Theta_{\mathrm{L,j}}(z,x,b_j\eta)e^{-z\sigma_o^2}dxdz + \frac{1}{\ln 2}\int_0^{\infty}\int_D^{\infty} \frac{1}{z}\Theta_{\mathrm{N,j}}(z,x,b_j\eta)e^{-z\sigma_o^2}dxdz, \quad (6)$$

where $\Theta_{\mathrm{L}}(z,x,b_j\eta)$ and $\Theta_{\mathrm{N}}(z,x,b_j\eta)$ are given by (7) and (9) at the top of the next page, respectively.

***Proof** 1:* The average rate of the j-th content served by a cache-enabled SBS can be expressed as

$$\mathcal{R}_{\mathcal{C},j} = \mathbb{E}\left[\log_2\left(1+\mathrm{SINR}_{c,j}\right)\right] \overset{(a)}{=} \frac{1}{\ln 2}\times \int_0^{\infty} \frac{1}{z} \underbrace{\mathbb{E}\left[(1-e^{-z\mathcal{S}_{c,j}})e^{-z\mathcal{I}_{c,j}}\right]}_{\Delta_j(z)} \underbrace{\mathbb{E}\left[e^{-z(\overline{\mathcal{I}}_{c,j}+\mathcal{I}'_c)}\right]}_{\overline{\Delta}_j(z)} e^{-z\sigma_o^2}dz, \quad (11)$$

where (a) comes from [12] which is valid for $\mathrm{SINR}_{c,j} \geq 0$. $\mathcal{S}_{c,j} = P_t G_{\mathrm{M}}^t G_{\mathrm{M}}^r L\left(|X_{n,j}|\right)$ depends on the nearest active cache-enabled SBS with distance $|X_{n,j}| = x$. First, we can directly obtain $\Delta_j(z)$ with signal and interference from the cache-enabled SBS who serves the j-th file

$$\Delta_j(z) = \int_0^{D}\left(1-e^{-zP_tG_{\mathrm{M}}^tG_{\mathrm{M}}^r\beta x^{-\alpha_{\mathrm{L}}}}\right)\underbrace{e^{-z\mathcal{I}_{c,j}^{\mathrm{L}}(z,x)}}_{\Delta_{\mathrm{L},j}(z,x)} f_{|X_{n,j}|}(x)dx + \int_D^{\infty}\left(1-e^{-zP_tG_{\mathrm{M}}^tG_{\mathrm{M}}^r\beta x^{-\alpha_{\mathrm{N}}}}\right)\underbrace{e^{-z\mathcal{I}_{c,j}^{\mathrm{N}}(z,x)}}_{\Delta_{\mathrm{N},j}(z,x)} f_{|X_{n,j}|}(x)dx, \quad (12)$$

where $f_{|X_{n,j}|}(x)$ is the PDF of the distance $|X_{n,j}|$ between the typical user and its nearest cache-enabled serving SBS that stores the j-th content, which is given by $f_{|X_{n,j}|}(x) = 2\pi b_j\eta\lambda_o x e^{-\pi b_j\eta\lambda_o x^2}$.

$\Delta_{\mathrm{L,j}}(z,x)$ is the interference inside the LoS area with the radius D and $\Delta_{\mathrm{N,j}}(z,x)$ is the interference of the NLoS area. Based on the law of total expectation, we can directly obtain $\Delta_{\mathrm{L,j}}(z,x)$ and $\Delta_{\mathrm{N,j}}(z,x)$ as

$$\Delta_{\mathrm{L},j}(z,x) = \mathbb{E}_{\Phi^{\mathrm{L}}_{\mathcal{C},j}/\{n\}}\left[e^{-z\mathcal{I}^{\mathrm{L}}_{c,j}(x)}\right] = \sum_{\ell,k\in\{\mathrm{M,m}\}} \Pr{}_{\ell k}^{tr} \times \exp(-z2\pi b_j\eta\lambda_o P_t G_{\ell k}^{tr}\beta \times \left(\int_{\min(x,D)}^{D} r^{1-\alpha_{\mathrm{L}}}dr + \int_{\max(x,D)}^{\infty} r^{1-\alpha_{\mathrm{N}}}dr\right)), \quad (13)$$

and

$$\Delta_{\mathrm{N},j}(z,x) = \mathbb{E}_{\Phi^{\mathrm{N}}_{\mathcal{C},j}/\{n\}}\left[e^{-z\mathcal{I}^{\mathrm{N}}_{c,j}(x)}\right] = \sum_{\ell,k\in\{\mathrm{M,m}\}} \Pr{}_{\ell k}^{tr} \times \exp\left(-z2\pi b_j\eta\lambda_o P_t G_{\ell k}^{tr}\beta\int_x^{\infty} r^{1-\alpha_{\mathrm{N}}}dr\right). \quad (14)$$

Then $\Delta_j(z)$ in (12) can be derived with (13) and (14) available. Next, we derive $\overline{\Delta}_j(z)$. It includes two sources of interference: one comes from the cache-enabled SBSs who do not have the j-th content, and the other comes from the traditional SBSs. So $\overline{\Delta}_j(z)$ is given by

$$\overline{\Delta}_j(z) = \mathbb{E}_{\Phi_{\mathcal{C}}\setminus\{\Phi_{\mathrm{C},j}\}\cup\Phi'_{\mathcal{C}}}\left[e^{-z\left(\overline{\mathcal{I}}_{c,j}+\mathcal{I}'_c\right)}\right] = \sum_{\ell,k\in\{\mathrm{M,m}\}} \Pr{}_{\ell k}^{tr} \times \exp\left(-z2\pi(1-b_j\eta)\lambda_o \times P_t G_{\ell k}^{tr}\beta\left(\int_0^{D} r^{1-\alpha_{\mathrm{L}}} r dr + \int_D^{\infty} r^{1-\alpha_{\mathrm{N}}}dr\right)\right). \quad (15)$$

Substituting (12) and (15) into (11), we attain the desired results in (6), (7) and (9). This completes the proof.

B. Achievable Rate Through Traditional SBSs

For comparison, we derive the achievable rate of the traditional network, which is novel and necessary in order to obtain the average achievable rate of the typical user in the hybrid network with limited backhaul, which is summarized in the theorem below.

***Theorem** 2:* The average achievable rate of the typical user that is associated with a traditional SBS in hybrid caching networks with limited backhaul which can be quantified as

$$\mathcal{R}_{\mathcal{O}}{}' = \min\left\{\widetilde{\mathcal{R}}'_{\mathcal{O}}, \mathcal{R}_{\mathcal{B}}\right\}, \quad (16)$$

where $\widetilde{\mathcal{R}}'_{\mathcal{O}}$ is derived as

$$\widetilde{\mathcal{R}}'_{\mathcal{O}} = \frac{1}{\ln 2}\int_0^{\infty}\int_0^{D}\frac{1}{z}\widetilde{\Theta}_{\mathrm{L}}(z,x)e^{-z\sigma_o^2}dxdz + \frac{1}{\ln 2}\int_0^{\infty}\int_D^{\infty}\frac{1}{z}\widetilde{\Theta}_{\mathrm{N}}(z,x)e^{-z\sigma_o^2}dxdz, \quad (17)$$

and $\widetilde{\Theta}_{\mathrm{L}}(z,x)$ and $\widetilde{\Theta}_{\mathrm{N}}(z,x)$ are given by (18) and (19) on the next page, respectively.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TVT.2018.2797047, IEEE Transactions on Vehicular Technology

$$\Theta_{\mathrm{L},j}(z,x,b_j\eta)=\sum_{\ell,k\in\{\mathrm{M},\mathrm{m}\}}\Pr_{\ell k}^{tr}\times f_{|X_{n,j}|}(b_j\eta,x)\cdot\left(1-e^{-zP_tG_{\mathrm{M}}^tG_{\mathrm{M}}^r\beta x^{-\alpha_{\mathrm{L}}}}\right)\Psi_{\mathrm{L}}\left(z,x,b_j\eta\right), \tag{7}$$

$$\text{with}\quad \Psi_{\mathrm{L}}\left(z,x,\mu\right)=e^{-z2\pi\lambda_oP_tG_{\ell k}^{tr}\beta\left(\left(\frac{\min\left\{D^{2-\alpha_{\mathrm{L}}}-\mu x^{2-\alpha_{\mathrm{L}}},(1-\mu)D^{2-\alpha_{\mathrm{L}}}\right\}}{2-\alpha_{\mathrm{L}}}-\frac{\max\left\{D^{2-\alpha_{\mathrm{N}}},\mu x^{2-\alpha_{\mathrm{N}}}+(1-\mu)D^{2-\alpha_{\mathrm{N}}}\right\}}{2-\alpha_{\mathrm{N}}}\right)\right)} \tag{8}$$

$$\Theta_{\mathrm{N},j}(z,x,b_j\eta)=\sum_{\ell,k\in\{\mathrm{M},\mathrm{m}\}}\Pr_{\ell k}^{tr}\times f_{|X_{n,j}|}(b_j\eta,x)\cdot\left(1-e^{-zP_tG_{\mathrm{M}}^tG_{\mathrm{M}}^r\beta x^{-\alpha_{\mathrm{N}}}}\right)\Psi_{\mathrm{N}}\left(z,x,b_j\eta\right), \tag{9}$$

$$\text{with}\qquad \Psi_{\mathrm{N}}\left(z,x,\mu\right)=e^{-z2\pi\lambda_oP_tG_{\ell k}^{tr}\beta\left(\frac{(1-\mu)D^{2-\alpha_{\mathrm{L}}}}{2-\alpha_{\mathrm{L}}}-\frac{\mu x^{2-\alpha_{\mathrm{N}}}+(1-\mu)D^{2-\alpha_{\mathrm{N}}}}{2-\alpha_{\mathrm{N}}}\right)} \tag{10}$$

$$\widetilde{\Theta}_{\mathrm{L}}(z,x)=\sum_{\ell,k\in\{\mathrm{M},\mathrm{m}\}}\Pr_{\ell k}^{tr}\times f_{|X_o|}(x)\cdot\left(1-e^{-zP_tG_{\mathrm{M}}^tG_{\mathrm{M}}^r\beta x^{-\alpha_{\mathrm{L}}}}\right)\Psi_{\mathrm{L}}\left(z,x,1-\eta\right) \tag{18}$$

$$\widetilde{\Theta}_{\mathrm{N}}(z,x)=\sum_{\ell,k\in\{\mathrm{M},\mathrm{m}\}}\Pr_{\ell k}^{tr}\times f_{|X_o|}(x)\cdot\left(1-e^{-zP_tG_{\mathrm{M}}^tG_{\mathrm{M}}^r\beta x^{-\alpha_{\mathrm{N}}}}\right)\Psi_{\mathrm{N}}\left(z,x,1-\eta\right) \tag{19}$$

***Proof** 2:* Similar to Theorem 1, the achievable rate of the typical user that is associated with a traditional SBS is given by

$$\mathcal{R}'_{\mathcal{O}}=\frac{1}{\ln 2}\int_0^\infty\frac{1}{z}(1-e^{-z\mathrm{SINR}'_o})dz=\frac{1}{\ln 2}\times$$
$$\int_0^\infty\frac{1}{z}\underbrace{\mathbb{E}\left[e^{-z\overline{\mathcal{I}}_o}\right]}_{\overline{\Lambda}(z)}\left[\int_0^D\left(1-e^{-z\mathcal{S}_o}\right)\underbrace{\mathbb{E}\left[e^{-z\mathcal{I}_o^{\mathrm{L}}(x)}\right]}_{\Lambda_{\mathrm{L}}(z,x)}f_{|X_o|}(x)dx\right.$$
$$\left.+\int_D^\infty\left(1-e^{-z\mathcal{S}_o}\right)\underbrace{\mathbb{E}\left[e^{-z\mathcal{I}_o^{\mathrm{N}}(x)}\right]}_{\Lambda_{\mathrm{N}}(z,x)}f_{|X_o|}(x)dx\right]e^{-z\sigma_o^2}dz, \tag{20}$$

where $\Lambda_{\mathrm{L}}(z,x)$, $\Lambda_{\mathrm{N}}(z,x)$ and $\overline{\Lambda}_{\mathrm{L}}(z,x)$ are obtained from (13) and (14) by replacing the parameter $b_j\eta$ by $\rightarrow 1-\eta$, and changing $f_{|X_o|}(x)$ to be the PDF of the distance $|X_o|=x$ between the typical user and its serving traditional mmWave SBS, which is given by $f_{|X_o|}(1-\eta,x)=2\pi(1-\eta)\lambda_o x e^{-\pi(1-\eta)\lambda_o x^2}$.

Then we obtain the desired results in (16) and (17) and this completes the proof.

C. Optimal Content Placement

Here we aim to optimize the content placement probability $\{b_j\}$ to reduce the backhaul usage or maximize $\mathcal{A}_c^{\mathrm{C}}$. The optimization problem can be formulated as:

$$\mathtt{P}_1:\quad\max_{0\le b_j\le 1,X}\quad\mathcal{A}_c^{\mathrm{C}}=\sum_{j=1}^{X}a_j\left(b_j\eta\lambda_o\right)\mathcal{R}_{\mathcal{C},j}\ \text{ in (5)}$$
$$\text{s.t.}\quad\sum_{j=1}^{X}b_j\le K,\ \text{and}\ b_j=0,\forall j>X. \tag{21}$$

Its solution is given in Theorem 3 below.

***Theorem** 3:* To maximize $\mathcal{A}_c^{\mathrm{C}}$, caching the MPC is the optimal solution, i.e., the optimal caching placement probabilities are given by $b_j=1,\forall\ 0\le j\le K$ and $b_j=0$ otherwise. In this case, $X=K$.

***Proof** 3:* We can rewrite $\mathcal{R}_{\mathcal{C},j}$ as $\mathcal{R}_{\mathcal{C}}(b_j)$, and from (6) it can be seen that $\mathcal{R}_{\mathcal{C}}(b_j)$ is an increasing function of the probability b_j. Also notice that the optimal solution should satisfy $b_j\ge b_m,\forall\ j<m$, because $\{a_j\}$ is arranged in a decreasing order. Define a utility function $\mathcal{U}_j(b_j)\triangleq a_j\eta\lambda_o\mathcal{R}_{\mathcal{C}}(b_j)$, and it is obvious that $\mathcal{U}_j(b_j)\ge\mathcal{U}_m(b_m),\forall\ j<m$, and now we can see that $\mathcal{U}_j(b_j)$ is an increasing function of b_j but it is a decreasing function of j.

Next introduce the auxiliary linear programming problem below

$$\mathtt{P}_2:\quad\max_{0\le b_j\le 1}\quad\sum_{j=1}^{X}b_j\widetilde{\mathcal{U}}_j$$
$$\text{s.t.}\quad\sum_{j=1}^{X}b_j\le K,\ \text{and}\ b_j=0,\forall j>X, \tag{22}$$

where we have introduced an auxiliary constant $\widetilde{\mathcal{U}}_j$ which is non-increasing in j but not a function of b_j. The optimal solution of the linear problem $\mathtt{P}_2$ is given by $b_j=1$ for $0\le j\le K$ and $b_j=0$ otherwise.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TVT.2018.2797047, IEEE Transactions on Vehicular Technology

By comparing $\mathtt{P}_1$ and $\mathtt{P}_2$, we can see the only difference in $\mathtt{P}_1$ is that $\mathcal{U}_j(b_j)$ is a function of b_j, but it is still a non-increasing function in j regardless the solution b_j ($b_j \geq b_m, \forall\, j < m$ should be satisfied). Therefore the optimal solution of $\mathtt{P}_1$ remains the same as that of $\mathtt{P}_2$. This completes the proof.

Based on Theorem 3, MPC is the optimal cache placement policy, which simplifies $\mathcal{A}_c$ to

$$\mathcal{A}_c = p_c \eta \lambda_o \mathcal{R}_{\mathcal{C},j} + p_o(1-\eta)\lambda_o \mathcal{R}'_{\mathcal{O}}, \tag{23}$$

where $p_c = \sum_{j=1}^{K} a_j$ is the hit probability for the K MPCs, and p_o is the hit-miss probability for the remaining contents.

IV. Numerical Results and Discussions

In this section, we provide numerical evaluation of the ASE and compare the performance of the traditional and the proposed hybrid cache-enabled mmWave networks. The system parameters are shown in Table I. We assume that both the transmitter and the receiver have the same array gain and beamwidth.

TABLE I
System parameter values [13].

Parameters	Values
Main-lobe array gain $((G^t_{\mathrm{M}}, G^t_{\mathrm{m}}, \theta) = (G^r_{\mathrm{M}}, G^r_{\mathrm{m}}, \vartheta))$	$(3\ \mathrm{dB}, -3\ \mathrm{dB}, 45^o)$
LoS region (D)	15 m
Transmit power of each MW-SBS P_t	30 dBm
Path loss exponent f_c=60 GHz	α_{L}=2.25,α_{N}=3.76
Available bandwidth in mmwave (W)	1 GHz
SBS cache capacity (K)	100
Content library size (J)	10^4
Backhaul constraint ($\mathcal{R}_{\mathcal{B}}$)	0.4 bit/s/Hz

To generate the numerical results, the *Zipf* request distribution is used and consequently the probability that the content f_j is requested is equal to

$$a_j = j^{-\delta} / \sum_{k=1}^{J} k^{-\delta}, \tag{24}$$

where δ is the Zipf exponent that represents the popularity skewness.

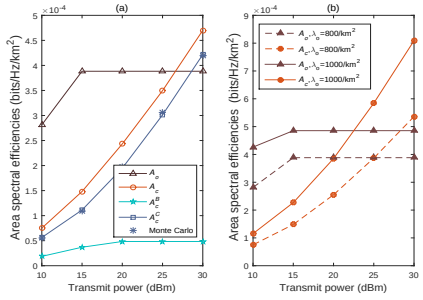

Fig. 2. Area spectral efficiencies, when $\eta = 0.5$. (a) $\lambda_o = 800/\mathrm{km}^2$, $\delta = 1.2$ and (b) $\delta = 1.5$.

In Fig. 2, we compare the traditional ASE $\mathcal{A}_o$ in (4) with the ASE $\mathcal{A}_c$ in (23) which is achieved by the hybrid cache-enabled network. The ASEs achieved by caching and backhaul only are also shown as $\mathcal{A}_c^C$ and $\mathcal{A}_c^B$, respectively. From Fig. 2 (a), we can see that as the transmit power increases, $\mathcal{A}_o$ is limited by the backhaul capacity, but $\mathcal{A}_c$ can keep increasing and achieve better performance than $\mathcal{A}_o$. From Fig. 2 (b), it is observed that the proposed cache-enabled SBSs achieve better performance in high-density networks.

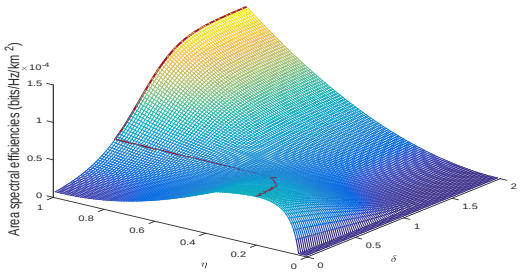

Fig. 3. Area spectral efficiencies, $\lambda_o = 200/\mathrm{km}^2$.

Next we investigate the tradeoff between the cache-enabled SBSs density η and the individual cache capacity limit K, given a total cache capacity of K_{t}, and we have $K = \min(J, \frac{K_t}{\eta \lambda_o \cdot |\mathcal{A}|})$. Fig. 3 shows the comparison of cache-enabled ASE in a square of side length $R_a = 1$ km given the total cache capacity $K_{\mathrm{t}} = 8 \times 10^4$. The Zipf skewness δ varies from $(0, 2)$. The red line shows the optimal value of the cache-enabled SBS density η when the Zipf skewness δ varies. The result shows that with high Zipf skewness, the ASE always increases with the cache-enabled SBS density. On the other hand, in the lower Zipf skewness case, there exists an optimal cache-enabled SBSs density η to maximize the ASE, and $\eta = 1$ is not necessarily the optimal strategy.

V. Conclusion

In this paper, we have derived the ASE performance of the proposed hybrid cache-enabled mmWave small cell networks and contrasted it with the traditional mmWave small cell networks without caching. We have shown that caching the MPCs is the optimal content placement strategy. Numerical results demonstrated that cache-enabled SBS networks can achieve higher ASE and save more backhaul consumption especially when the transmit power is high. For a given total cache capacity within an area, there is an optimal cache-enabled SBSs density that maximizes ASE. Our work is limited to the fixed network topologies without considering user mobility, we will consider in our future work.

References

[1] S. Woo, E. Jeong, S. Park, J. Lee, S. Ihm, and K. Park, "Comparison of caching strategies in modern cellular backhaul networks," in *Proceeding of the 11th annual international conference on Mobile systems, applications, and services*, Taipei, Taiwan, Jun. 2013, pp. 319–332.

[2] X. Wang, M. Chen, T. Taleb, A. Ksentini, and V. Leung, "Cache in the air: exploiting content caching and delivery techniques for 5g systems," *IEEE Commun. Mag.*, vol. 52, no. 2, pp. 131–139, Feb. 2014.

[3] J. Qiao, Y. He, and X. S. Shen, "Proactive caching for mobile video streaming in millimeter wave 5G networks," *IEEE Trans. Wireless Commun.*, vol. 15, no. 10, pp. 7187 – 7198, Aug. 2016.

[4] M.-S. Alouini and A. J. Goldsmith, "Area spectral efficiency of cellular mobile radio systems," *IEEE Trans. Veh. Technol.*, vol. 48, no. 4, pp. 1047–1066, Jul. 1999.
[5] J. G. Andrews, "Seven ways that hetnets are a cellular paradigm shift," *IEEE Commun. Mag.*, vol. 51, no. 3, pp. 136–144, Mar. 2013.
[6] D. Liu and C. Yang, "Caching policy toward maximal success probability and area spectral efficiency of cache-enabled Hetnets," *IEEE Trans. Commun.*, vol. PP, no. 99, pp. 1–1, Mar. 2017.
[7] X. Xu and M. Tao, "Modeling, analysis, and optimization of coded caching in small-cell networks," *IEEE Trans. Commun.*, vol. 65, no. 8, pp. 3415 – 3428, Aug. 2017.
[8] M. Chen, M. Mozaffari, W. Saad, C. Yin, M. Debbah, and C. S. Hong, "Caching in the sky: Proactive deployment of cache-enabled unmanned aerial vehicles for optimized quality-of-experience," *IEEE J. Sel. Areas Commun.*, vol. 35, no. 5, pp. 1046–1061, May 2017.
[9] T. Bai and R. W. Heath, "Coverage and rate analysis for millimeter-wave cellular networks," *IEEE Trans. Wireless Commun.*, vol. 14, no. 2, pp. 1100–1114, Feb. 2015.
[10] O. El Ayach, S. Rajagopal, S. Abu-Surra, Z. Pi, and R. W. Heath, "Spatially sparse precoding in millimeter wave mimo systems," *IEEE Trans. Wireless Commun.*, vol. 13, no. 3, pp. 1499–1513, 2014.
[11] T. S. Rappaport, S. Sun, R. Mayzus, H. Zhao, Y. Azar, K. Wang, G. N. Wong, J. K. Schulz, M. Samimi, and F. Gutierrez, "Millimeter wave mobile communications for 5g cellular: It will work!" *IEEE access*, vol. 1, pp. 335–349, May 2013.
[12] K. A. Hamdi, "Capacity of MRC on correlated Rician fading channels," *IEEE Trans. Commun.*, vol. 56, no. 5, pp. 708–711, May 2008.
[13] T. S. Rappaport, E. Ben-Dor, J. N. Murdock, and Y. Qiao, "38 GHz and 60 GHz angle-dependent propagation for cellular & peer-to-peer wireless communications," in *IEEE Int. Conf. Commun. (ICC)*, Ottawa, Canada, Nov. 2012, pp. 4568–4573.