View metadata, citation and similar papers at core.ac.uk

brought to you by CORE

provided by Kent Academic Repository

Kent Academic Repository

Full text document (pdf)

Citation for published version

Daghal, Asaad and Zhu, Huiling and Wang, Jiangzhou (2018) Content Delivery Analysis in Multiple Devices to Single Device Communications. IEEE Transactions on Vehicular Technology . ISSN 0018-9545.

DOI

https://doi.org/10.1109/TVT.2018.2865012

Link to record in KAR

https://kar.kent.ac.uk/69371/

Document Version

Author's Accepted Manuscript

Copyright & reuse

Content in the Kent Academic Repository is made available for research purposes. Unless otherwise stated all content is protected by copyright and in the absence of an open licence (eg Creative Commons), permissions for further reuse of content should be sought from the publisher, author or other copyright holder.

Versions of research

The version in the Kent Academic Repository may differ from the final published version.
Users are advised to check **http://kar.kent.ac.uk** for the status of the paper. **Users should always cite the published version of record.**

Enquiries

For any further enquiries regarding the licence status of this document, please contact:
researchsupport@kent.ac.uk

If you believe this document infringes copyright then please contact the KAR admin team with the take-down information provided at **http://kar.kent.ac.uk/contact.html**

KAR Kent Academic Repository

University of Kent

Content Delivery Analysis in Multiple Devices to Single Device Communications

Asaad. S. Daghal, Huiling Zhu, *Senior Member, IEEE* and Jiangzhou Wang, *Fellow, IEEE*

School of Engineering and Digital Arts, University of Kent, UK.
Email: {ad466, H.Zhu, J.Z.Wang}@kent.ac.uk

Abstract

Content caching at the terminal user device (UD) utilizing the device to device (D2D) communications is a promising technology to enhance the performance of mobile networks in terms of latency, throughput, energy consumption and so on. In this paper, a novel method of content delivery using multiple devices to single device (MDSD) communications through D2D links is presented. In this method, the Zipf distribution with exponent shape parameter is adopted to model the content caching popularity which affects the achievable signal to interference plus noise ratio (SINR). In order to investigate the advantage of the proposed MDSD method, firstly, a closed-form expression of the outage probability is theoretically derived for a single D2D communication to evaluate the success of content delivery to the UD. Secondly, the expression of the outage probability for MDSD communication is derived, where the outage probability is modeled as a function of content caching popularity, the density of UDs, and the size of cooperative area. The research work is further extended to address the frequency reuse among different UDs in one cell, where a frequency band factor is introduced, and the optimal radius of the cooperative area is defined and analysed. The analytical results are validated by the simulation results and show that the outage probability decreases drastically when the popularity of the content increases, or the radius of the cooperative area increases. Using the derived closed-form expression of the outage probability, the area spectral efficiency (ASE) of the system is presented. Furthermore, the results show that as the frequency band factor increases, the outage probability decreases. Finally, it is shown that the MDSD outperforms the single D2D-based method.

Index Terms

D2D communication, Outage probability, Stochastic geometry, Poisson point process, Caching, Zipf distribution, Lomax distribution.

I. INTRODUCTION

The proliferation of smartphones has boosted wireless data traffic substantially during the last decade [1]. With the increase in wireless data traffic, the fourth-generation (4G) cellular system has already reached its theoretical capacity [2]–[5]. Therefore, in order to provide high quality of services to users, dealing with growing amounts of data traffic of great importance. According to [1], the major contributor towards data traffic is video, which accounts for more than half of total mobile data traffic, and is often generated from duplicated requests for a few popular video streams. For example, 10% of the videos on Youtube account for nearly 80% of viewings [6]. This fact led to important solution to reduce data traffic, which utilizes the storage unit in the user device (UD) to store content temporarily and allowing other UDs to download from the UD storing the content. By enabling a local UD to communicate with others via device to device (D2D) links, content caching also becomes a useful method to offload the network data traffic, decrease the average access latency and make the base station (BS) more scalable [7].

D2D communication is considered a crucial technology for achieving high data rate, due to the ability of transmitting data directly amongst devices in proximity [8]. As D2D provides an opportunity for local communications and the storage capacity of devices increases, caching popular content in mobile devices along with using D2D communication for content delivery has been investigated as a promising method to unleash the ultimate benefits of content caching [9].

In [10], a femtocell BS with the storage unit has been considered as a helper, where the communication between a UD and the femtocell BS can be widely considered as D2D. [10] showed that each UD throughput can be increased if the content can be stored at the femto BS and reused by the UDs. In [11], a central caching placement method was considered. [11] demonstrated that the spectral efficiency can be improved up to two order of magnitude when the central caching method with D2D communication was adopted. In [12], frequency reuse has been proposed among multiple D2D pairs in order to reduce transmission power. In [12] the whole cell was divided into equally small square clusters, and frequency can be used among different D2D pairs in different clusters to transmit the same content while only one content can be received via one D2D link in one cluster to avoid intra-cluster interference. However, since the assumptions in [12] are too simplified and did not consider channel fading and inter cluster interference, this approach is not practical. Moreover, the possibility of finding and downloading the desired content from a neighbor UD using D2D communication is very likely to be low, and the outage probability could be high for several reasons. For examples, due to the privacy concern, only limited number of helpers can have the desired content, their channel fading may not be good, or the interference

caused by simultaneous transmission using the same frequency band. These reasons motivate us to explore the transmission diversity (TD) in multiple-devices-to-single-device (MDSD) communications in D2D communications in order to reduce the outage probability. The impact of TD has been widely investigated in cellular networks in order to combat the effects of fading by transmitting the same data over different antennas [13], [14]. In [15], the distribution of the signal to interference ratio (SIR) was derived by applying a Toeplitz matrix, in which a multi-antenna transmitters were equipped in a base station (BS). In [16], a closed-form expression of the outage probability was derived, where the distribution of the power of the received signal from two transmitters was obtained. In addition, [16] assumed that the interference is modeled as a Gaussian distribution with fixed power value (variance). Moreover, [16] only presented analytical results which showed that the outage probability depends on the popularity of content and the storage size of the UE. However, the distances from the transmitters to the receiver in the TD technique were assumed fixed in [13]–[15]. The analysis of random distance models with TD techniques are much more challenging than fixed distance.

This paper focuses on exploring MDSD communications for content caching and delivery. And, the outage probability is analyzed for UDs in MDSD communications by considering the randomness of the distances between transmitters and the receiver. At each UD, a receiving signal to interference plus noise ratio SINR threshold is adopted to guarantee successfully decoding the received signal. In order to satisfy this constraint, we investigate the impact of the TD in the MDSD communication-based method to cache and deliver contents in an environment with a high density of UEs, e.g. stadiums and shopping centers. The area spectral efficiency (ASE) of the proposed method is theoretically analyzed and verified by simulation results. The main contributions of this paper are listed below:

1) Based on stochastic geometry, a general formula of outage probability for content delivery only based on communication between one D2D pair, called single D2D communication, for a specific area is derived. Due to tractability, a closed-form expression of outage probability for special cases with path loss exponent $\alpha = 4$ and $\alpha = 2$ are derived, respectively. Under the assumption of Poisson Point Process (PPP) and $\alpha = 2$, the interference component of the success probability for the specific area is not going to be zero $(\mathbb{P}_{sec}^{\mathcal{I}_a} \neq 0)$ is proved.
2) Using stochastic geometry, we investigate outage probability performance of the proposed MDSD method. An expression of the outage probability is derived by considering the randomness of the distances between the transmitters and the receiver. The distribution of the power of received desired signal from multiple transmitters is derived when $\alpha = 2$. First, the probability density function (PDF) of the power of desired signal for a single link is defined as a special case of the Lomax

distribution. Then, a Laplace transform (LT) is used to find the distribution for the summation of the power of received desired signals, where a Bromwich integral and residue theorem are used to implement the inversion of the LT.

3) The analytical results of the proposed MDSD method are compared with that of a single D2D based method and validated by the simulation results. It is shown that the outage probability can be reduced by 90% for MDSD compared to the single D2D based method.
4) The ASE is analyzed, where the optimal value of the cooperative area is obtained, which gives us a good indication to the trade-off between outage probability and ASE.

Some of important symbols that used through this paper are listed in **Table** I.

Table I: Symbols used in the paper

Symbol	Definition
bps	bits per second
$\mathbb{E}(x)$	Expectation of random variable x
$\mathcal{L}_X(s)$	Laplace transform of a random variable X
$\mathbb{R}^2$	Two dimensional space
$\mathbb{P}(x)$	Probability of random variable x

The remainder of this paper is organized as follows. System model and assumptions are described in Section II. In Section III, the performance metrics and the main analytical results related to the system model are provided. More specific assumptions are made in the same section, yielding much simpler expressions. The discussion of the theoretical results which are validated by the simulation results are presented in Section IV, while the paper is concluded in Section V.

II. System model of the content delivery

A. Topology

Consider a cellular network with UDs randomly located in the cell as shown in Fig. 1. The UDs are independently distributed within the whole cell area $\mathcal{B}$ with radius R and modeled as a stationary homogeneous Poisson Point Process (PPP) Φ with intensity λ in the two-dimensional space $\mathbb{R}^2$. In other words, the process $\Phi = \{x_i\} \subset \mathbb{R}^2$, where x_i is the i^{th} UD location, which is identically and independently distributed (i.i.d) in the Euclidean plane. The number of UDs within the cell $\mathcal{B}$ is denoted

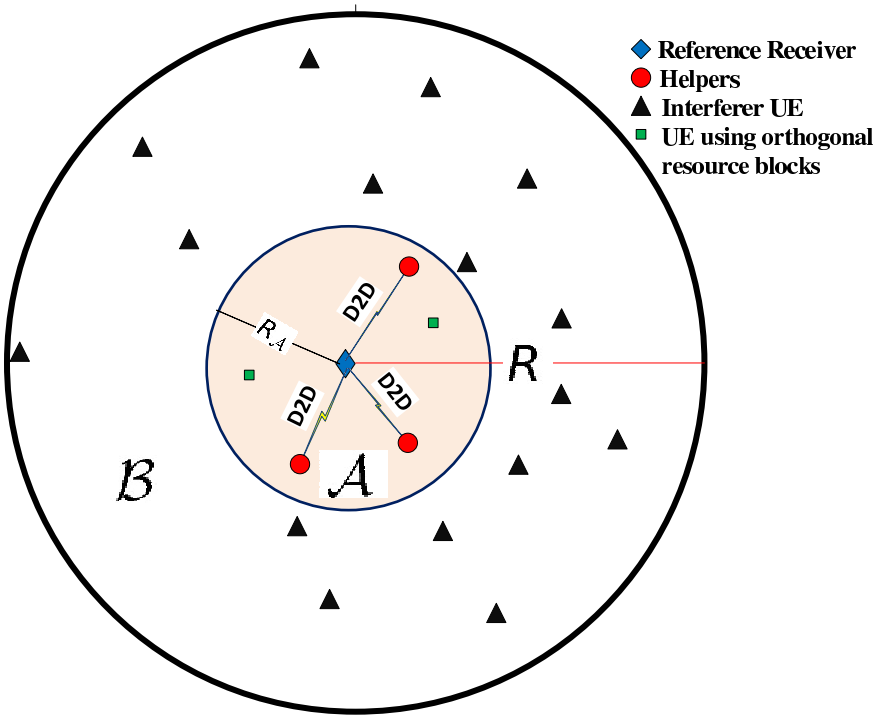

Figure 1: System model, the UDs are randomly and independently distributed according to PPP model

by $K_{\mathcal{B}}$, and its expectation $\mathbb{E}(K_{\mathcal{B}}) = \lambda|\mathcal{B}|$, where $|\mathcal{B}| = \pi R^2$. It is assumed that each UD has a cache unit used to store contents temporarily. Each content could be a text, image, or a part of or complete video. For simplicity, it is assumed that the cache unit size is the same for all UEs and at least there is one popular content to be stored in each UD. In this system, a reference receiver UD is assumed to be located at the origin (o) of the space $\mathbb{R}^2$ and receives the desired content from k_h transmitter devices, called helpers, simultaneously within the area $\mathcal{A}$. The locations of the transmitters are also modeled by a homogeneous PPP of intensity λ_h, which is denoted as Φ_h. It is assumed that all UDs are equipped with single Omni-directional antenna.

B. Content caching and delivery method

Content caching at terminal UDs is an efficient method to alleviate the data traffic from the network. The scenario of the content request and caching is explained as follows. It is assumed that each UD in the system requests its own desired content randomly and independently from a finite library with M different contents denoted as $\mathcal{F} = \{f_1, f_2, \cdots, f_M\}$, where f_m represents the m^{th} most popular content. By defining the popularity of a content as the probability that the content is requested, according

to statistics, the most common distribution has been used widely to model content popularity is Zipf distribution [17]. The Zipf law reveals that the probability of any content being requested is inversely proportional to the content's rank in the popularity table of contents. That is, the popularity of the content f_m with ranking index m is given by

$$p_m = \frac{m^{-\rho}}{\sum_{m=1}^{M} m^{-\rho}} \quad \rho > 0, \tag{1}$$

where ρ is the shape parameter indicating the rate of popularity. As shown in Fig. 2, when ρ becomes large, only a small number of contents are very popular and account for most of the requests. On the other hand, when $\rho = 0$, the popularity of each content is the same, which means the contents are uniformly distributed. In this system, it is also assumed that the UDs cache the contents downloaded

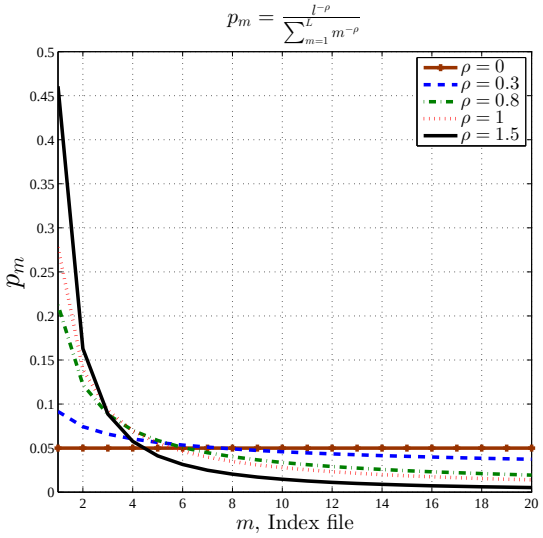

Figure 2: Zipf distribution of contents popularity

with a certain probability. Let $\mathrm{C} = \{\mathcal{C}_1, \mathcal{C}_2, \cdots, \mathcal{C}_L\}$ be the set of the contents that cached in all the UDs within $\mathcal{A}$, where $L \leq M$. The caching probability of a content $\mathcal{C}_l$ at a UD is also assumed to follow Zipf distribution as

$$\vartheta_l = \frac{l^{-\rho}}{\sum_{i=1}^{L} i^{-\rho}}, \quad \rho > 0. \tag{2}$$

The content caching and delivery can be divided into two stages. In the first stage as initialization stage, within the cell, contents are distributed from the core network to the UDs according to its popularity among all the contents requested. In the second stage, the UD who requests for a specific content $\mathcal{C}_l$ but does not receive it, will be serve by those UDs with this content via D2D communications. The main idea of the content caching and delivery method of the MDSD is described as follows: if the reference

UD requests for a desired content $\mathcal{C}_l$, which can be found and downloaded from its neighbors within the area $\mathcal{A}$, these neighbors holding the desired content will serve the reference UD via D2D links. Otherwise, the requested content will be retrieved from the core network via the BS. According to the caching probability given in (2), the density of transmitters is given by $\lambda_h = \vartheta_l \lambda$, and the number of UDs holding the content, denoted by K, and its expectation $\mathbb{E}(K)$=$\lambda|\mathcal{A}|$, where $|\mathcal{A}| = \pi \mathrm{R}_{\mathcal{A}}^2$ is the size of $\mathcal{A}$.

C. Signal to Interference plus Noise Ratio

As described above, in the proposed MDSD based content delivery system, the transmission method allows a set of multiple helpers cooperating jointly to transmit the same content to a given reference UD using the same time-frequency resource block. Therefore, the desired content could be received from multiple UDs. The strength of the received signal can then be improved to increase the successful probability of receiving the requested content by applying transmission diversity (TD), which allocates transmit power to each transmitter in the paper proportionally to the channel condition between the transmitter and receiver. However, in order to simplify the system design, it is assumed that all users in the system have the same transmit power (unit power signal). Within the area $\mathcal{A}$ which is centered by the reference UD, an other devices except the helpers use different resource blocks for their transmissions, while outside the area $\mathcal{A}$, the same resource block used for the reference UD's content transmission can be reused. Therefore, at the reference UD, the SINR is given by

$$\text{SINR} = \frac{\sum_{i=1}^{k_h} |h_i|^2 r_i^{-\alpha} P_t}{\sigma_n^2 + \mathcal{I}_a}, \tag{3}$$

where P_t is the transmit power of one UD, r_i is the distance between the reference UD and the i^{th} serving helper, k_h is the number of synchronized helpers holding the desired content $\mathcal{C}_l$ in their caches within $\mathcal{A}$, and σ_n^2 is the additive white Gaussian noise power. α is the path loss exponent depending on the carrier frequency and physical environment, which is approximated in the range of (1.6 , 6.5) [18]. A Rayleigh based small-scale fading, h_i, is assumed between the i^{th} helper and the reference UD, where $X = |h_i|^2$ is the power gain following an exponential distribution with unit mean. $\mathcal{I}_a$ represents the sum of power of the aggregated interfering signals at the reference UD due to the non-cooperative UDs from outside the area $\mathcal{A}$, and is given by

$$\mathcal{I}_a = \sum_{k \in \Phi / \mathcal{A}} |g_k|^2 r_k^{-\alpha} P_t \tag{4}$$

where r_k is the distance from the reference UD to the k^{th} interferer UD. $|g_k|^2$ is the power of the channel fading between the k^{th} interferer UD and reference UD, where g is the interferer's channel distribution follows exponential distribution with unit mean.

III. Performance Analysis

In this section, the outage probability of a reference UD is analyzed. For the developed content delivery method, outage happens if the received SINR falls below a given target threshold η within area $\mathcal{A}$. In general, the outage probability of the proposed MDSD method is given by

$$\mathbb{P}_{out} = \mathbb{E}\left[\left(\frac{\sum_{i=1}^{k_h}|h_i|^2 r_i^{-\alpha} P_t}{\sigma_n^2 + \mathcal{I}_a} < \eta | r_i\right)\right]. \tag{5}$$

By considering that the reference UD contacts with the nearest helper (not necessarily closest UD), the PDF of the distance r between the reference UD and the nearest helper, is given by [19]

$$f_r(r) = 2\pi\lambda_h r \exp(-\pi\lambda_h r^2). \tag{6}$$

where $\lambda_h = \lambda\vartheta_l$, and ϑ_l is the probability of the file $\mathcal{C}_l$ being cached, as considered in (2). In the following subsection, we analytically derive the outage probability in two cases, single D2D and MDSD communications, respectively.

A. Single D2D communication

For a single D2D communications, no matter whether $k_h \geq 1$, there is only one helper to transmit the desired content to the reference UD. The success probability is the probability that SINR exceeds the target threshold η, when the requested content $\mathcal{C}_l$ exists in the nearest neighbor helper. Then the outage probability can be given as the complementary of the success probability as follows.

Theorem 1. *Given the density of UDs, λ, path loss exponent α and the target threshold η, the outage probability of the single D2D communication is given by*

$$\mathbb{P}_{out,\alpha}^{D2D} = 1 - \int_0^R \pi\lambda_h \exp\left(-\pi\lambda_l\left(1 + \kappa\left(\eta,\alpha\right)\right)v - \frac{\eta}{\eta_o}v^{\alpha/2}\right)dv, \tag{7}$$

where

$$\kappa(\eta,\alpha) = \int_1^R \frac{\eta}{\eta + y^{\alpha/2}}dy. \tag{8}$$

Proof. The outage probability of the single D2D communication conditioned on distance r, which is denoted as $\mathbb{P}_{out,\alpha}^{D2D}$, is expressed as

$$\begin{aligned}\mathbb{P}_{out,\alpha}^{D2D} &= \mathbb{E}_r\left[\mathbb{P}\left(\text{SINR}^{D2D}<\eta|r\right)\right],\\ &= 1-\int_{r>0}\mathbb{P}\left[\text{SINR}^{D2D}>\eta|r\right]f_r(r)dr,\\ &= 1-\int_{r>0}\mathbb{P}\left[\frac{|h|^2r^{-\alpha}P_t}{\sigma_n^2+\mathcal{I}_a}>\eta|r\right]2\pi\lambda_h r\exp\left(-\pi\lambda_h r^2\right)dr, \end{aligned} \tag{9}$$

where $\mathbb{E}_r(\cdot)$ is the expectation with respect to r. Since the channel power gain is $|h|^2\sim\exp(1)$, the probability of exceeding $\frac{\eta r^\alpha}{P_t}\left(\sigma_n^2+\mathcal{I}_a\right)$ is given by

$$\begin{aligned}\mathbb{P}\left[|h|^2>\frac{\eta r^\alpha}{P_t}\left(\sigma_n^2+\mathcal{I}_a\right)|r\right] &= \mathbb{E}_{\mathcal{I}_a}\left[\mathbb{P}\left[|h|^2>\frac{\eta r^\alpha}{P_t}\left(\sigma_n^2+\mathcal{I}_a\right)|r,\mathcal{I}_a\right]\right]\\ &= \mathbb{E}_{\mathcal{I}_a}\left[\exp\left(-\frac{\eta r^\alpha}{P_t}\left(\sigma_n^2+\mathcal{I}_a\right)\right)|r\right]\\ &\overset{(a)}{=} \underbrace{\exp\left(-\eta r^\alpha/\eta_o\right)}_{\mathbb{P}_{sec}^{\sigma_n^2}}\underbrace{\mathcal{L}_{\mathcal{I}_n}(\eta r^\alpha)}_{\mathbb{P}_{sec}^{\mathcal{I}_n}}, \end{aligned} \tag{10}$$

where $\mathcal{I}_n=\frac{\mathcal{I}_a}{P_t}$ is the total interference, $\eta_o=\frac{P_t}{\sigma_n^2}$ is the signal to noise ratio (SNR), and (a) is the success probability which is a product of two factors, a noise term $\mathbb{P}_{sec}^{\sigma_n^2}$ which is independent of the interference, and an interference term $\mathbb{P}_{sec}^{\mathcal{I}_n}$ which is independent of the noise. $\mathbb{P}_{sec}^{\mathcal{I}_n}$ is derived in Appendix A, which is obtained by determining the LT of the interference $\mathcal{I}_n$ at $s=\eta r^\alpha$ and given by

$$\mathbb{P}_{sec}^{\mathcal{I}_n}=\mathcal{L}_{\mathcal{I}_n}\left(\eta r^\alpha\right)=\exp\left(-\pi\lambda_h r^2\kappa\left(\eta,\alpha\right)\right). \tag{11}$$

Since our system is modeled for the specific area within distance R, the expectation in (9) is considered for the distance $0<r<R$. Substituting (11) into (10), (10) into (9), and changing variable $v=r^2$, (7) is obtained. ■

For the generality of (7), in the following parts, a closed form expression is presented under different scenarios with the path loss exponent α =4 and α=2, respectively.

Lemma 1.1. *For the Rayleigh fading and the path loss exponent α=4, the outage probability in this case, denoted as $\mathbb{P}_{out,4}^{D2D}$, is given by*

$$\begin{aligned}\mathbb{P}_{out,4}^{D2D} &= 1-\frac{\pi^{\frac{3}{2}}\lambda_h}{\sqrt{4\frac{\eta}{\eta_o}}}\exp\left(\frac{\left(\pi\lambda_h\left(1+\kappa\left(\eta,4\right)\right)\right)^2}{4\frac{\eta}{\eta_o}}\right)\\ &\times \left[\text{erfc}\left(\frac{\pi\lambda_h\left(1+\kappa\left(\eta,4\right)\right)}{\sqrt{4\frac{\eta}{\eta_o}}}\right)-\text{erfc}\left(\frac{\pi\lambda_h\left(1+\kappa\left(\eta,4\right)\right)+2R\frac{\eta}{\eta_o}}{\sqrt{4\frac{\eta}{\eta_o}}}\right)\right]. \end{aligned} \tag{12}$$

Proof. By applying $\alpha = 4$ into (7), $\mathbb{P}_{out,4}^{D2D}$ is expressed as

$$\mathbb{P}_{out,4}^{D2D} = 1 - \pi\lambda_h \underbrace{\int_0^R \exp\left(-\pi\lambda_h(1+\kappa(\eta,4))v - \frac{\eta}{\eta_o}v^2\right) dv}_{(***)}. \quad (13)$$

where

$$\kappa(\eta,4) = \int_1^R \frac{\eta}{\eta+y^2}dy = \sqrt{\eta}\left(\cot^{-1}(\sqrt{\eta}/R) - \cot^{-1}(\sqrt{\eta})\right). \quad (14)$$

and $\cot^{-1}(\cdot)$ is the inverse cotangent function. Letting $\varrho = \pi\lambda_l(1+\kappa(\eta,4))$ and $\beta = \frac{\eta}{\eta_o}$, $(***)$ in (13) can be evaluated according to (1.8 page 108 [20])

$$\begin{aligned}\int_0^u \exp\left(-\beta x^2 - \varrho x\right)dx &= \sqrt{\frac{\pi}{4\beta}}\exp\left(\frac{\varrho^2}{4\beta}\right)\left[\text{erf}\left(\sqrt{\beta}u + \frac{\varrho}{2\sqrt{\beta}}\right) - \text{erf}\left(\frac{\varrho}{2\sqrt{\beta}}\right)\right], \\ &= \sqrt{\frac{\pi}{4\beta}}\exp\left(\frac{\varrho^2}{4\beta}\right)\left[\text{erfc}\left(\frac{\varrho}{2\sqrt{\beta}}\right) - \text{erfc}\left(\sqrt{\beta}u + \frac{\varrho}{2\sqrt{\beta}}\right)\right], \quad (15)\end{aligned}$$

where erf$(\cdot)$ is the standard error function, defined as erf(x) $= \frac{2}{\sqrt{\pi}}\int_0^x \exp(-t^2)dt$ and erfc$(\cdot) = 1 -$ erf$(\cdot)$ is the complementary error function. By substituting (18) into (16), one obtains

$$\mathbb{P}_{out,4}^{D2D} = 1 - \frac{\pi^{\frac{3}{2}}\lambda_h}{\sqrt{4\beta}}\exp\left(\frac{\varrho^2}{4\beta}\right)\left[\text{erfc}\left(\frac{\varrho}{\sqrt{4\beta}}\right) - \text{erfc}\left(\sqrt{\beta}R + \frac{\varrho}{2\sqrt{\beta}}\right)\right]. \quad (16)$$

■

Lemma 1.2. *For the Rayleigh fading and path loss exponent* α*=2, the outage probability of this case, denoted as* $\mathbb{P}_{out,2}^{D2D}$*, is given by*

$$\mathbb{P}_{out,2}^{D2D} = 1 - \frac{\pi\lambda_h\left(1 - \exp\left(-R\left(\pi\lambda_h(1+\eta\log(\frac{\eta+R}{\eta+1})) + \frac{\eta}{\eta_o}\right)\right)\right)}{\left(\pi\lambda_h(1+\eta\log(\frac{\eta+R}{\eta+1})) + \frac{\eta}{\eta_o}\right)}. \quad (17)$$

Proof. By applying α=2 into (7), $\mathbb{P}_{out,2}^{D2D}$ is expressed as

$$\mathbb{P}_{out,2}^{D2D} = 1 - \pi\lambda_h\int_0^R \exp\left(-\pi\lambda_h(1+\kappa(\eta,2))v - \frac{\eta}{\eta_o}v\right)dv, \quad (18)$$

where $\kappa(\eta,2)$ is given by

$$\kappa(\eta,2) = \int_1^R \frac{\eta}{\eta+y}dy = \eta\log\left(\frac{\eta+R}{\eta+1}\right). \quad (19)$$

Then, by substituting (19) into (18), the solution of (18) is derived as

$$\begin{aligned}\mathbb{P}_{out,2}^{D2D} &= 1 - \pi\lambda_h\int_0^R \exp\left(-\pi\lambda_h\left(1+\eta\log\left(\frac{\eta+R}{\eta+1}\right)\right)v - \frac{\eta}{\eta_o}v\right)dv, \\ &= 1 - \frac{\pi\lambda_h\left(1 - \exp\left(-R\left(\pi\lambda_h(1+\eta\log(\frac{\eta+R}{\eta+1})) + \frac{\eta}{\eta_o}\right)\right)\right)}{\left(\pi\lambda_h(1+\eta\log(\frac{\eta+R}{\eta+1})) + \frac{\eta}{\eta_o}\right)}. \quad (20)\end{aligned}$$

For PPP infinite network in the two-dimensional space $\mathbb{R}^2$ with $\alpha = 2$, [21] proved that the expected value of the interference in (10) is infinite ($\mathbb{E}_{\mathcal{I}_a}(\cdot) = \infty$), and the success probability of the interference component is equal to zero ($\mathbb{P}_{sec}^{\mathcal{I}_n} = 0$), thereby $\mathbb{P}_{out} = 1$. In contrast, we proved that for the specific area within radius R, the success probability of the interference component is not going to be zero ($\mathbb{P}_{sec}^{\mathcal{I}_n} \neq 0$), thereby $\mathbb{P}_{out} < 1$. ■

B. Multiple devices to single device (MDSD) communication

In order to improve the outage probability of the single D2D communication, MDSD based method is proposed in this paper, especially to be used in the scenario where the density of users is high. The purpose of this proposed scenario is to alleviate the data traffic from the network due to a large number of connected devices in a dense area. The outage probability, denoted as $\mathbb{P}_{out}^{MD}$, is evaluated as follows.

Theorem 2. *Given the density of UDs λ, path loss exponent $\alpha = 2$, and the target* SINR *threshold η, the outage probability of the MDSD communication is given by*

$$\mathbb{P}_{out,2}^{MD} = 1 - \sum_{K=1}^{K_{\mathcal{B}}} \sum_{k_h=1}^{K} \binom{K_{\mathcal{B}}}{K} \left(\frac{|\mathcal{A}|}{|\mathcal{B}|}\right)^{K} \left(1 - \frac{|\mathcal{A}|}{|\mathcal{B}|}\right)^{K_{\mathcal{B}}-K} \binom{K}{k_h} \vartheta_l^{k_h} (1-\vartheta_l)^{K-k_h} \int_0^{\infty} \frac{\exp\left(-x(k_h + \frac{\eta}{\eta_o \mu} + \eta\nu)\right)}{\pi x} \Upsilon_{k_h}(x) dx, \quad (21)$$

where $\nu = \log\left(\frac{x\eta+\mu R^2}{x\eta+\mu R_{\mathcal{A}}^2}\right)$, $\binom{K_{\mathcal{B}}}{K} = \frac{K_{\mathcal{B}}!}{(K_{\mathcal{B}}-K)!K!}$ and symbol $(!)$ represents the operator of factorial. $\Upsilon_{k_h}(x)$ is defined as

$$\Upsilon_{k_h}(x) = \sum_{n=1}^{k_h} \binom{k_h}{n} \sin\left(\frac{n\pi}{2}\right) \left(E_2(-x)\right)^{k_h-n} (\pi x)^n, \quad (22)$$

where $E_v(\cdot)$ is the generalized exponential integral function, which is defined as

$$E_v(x) = \int_1^{\infty} \frac{\exp(-xt)}{t^v} dt, \quad v = 0, 1, 2, \cdots \quad (23)$$

Proof. The outage probability of MDSD communication is expressed as

$$\mathbb{P}_{out,2}^{MD} = \mathbb{E}\left[\mathbb{P}\left(\text{SINR}^{MD} < \eta | r_i\right)\right], \quad (24)$$

where the received SINR^{MD} for the content $\mathcal{C}_l$ is conditioned on $R_{\mathcal{A}}$. (27) is written as

$$\mathbb{P}_{out,2}^{MD} = 1 - \sum_{K=1}^{K_{\mathcal{B}}} \sum_{k_h=1}^{K} \mathcal{H}_l \underbrace{\mathbb{E}\left[\mathbb{P}\left(\frac{\sum_{i=1}^{k_h} |h_i|^2 r_i^{-\alpha} P_t}{\sigma_n^2 + \mathcal{I}_a} > \eta | r_i\right)\right]}_{(\star)}, \quad (25)$$

where $\mathcal{H}_l$ is the hit probability that there are K UDs out of $K_{\mathcal{B}}$ inside area $\mathcal{A}$, and k_h helpers out of K UDs having a specific content $\mathcal{C}_l$, which is derived in Appendix B. Note that, $(\star)$ in (25) is the probability

that SINR^{MD} exceeds a target threshold value η which is explained as follows. Since the received signal is collected from multiple transmitters from random distances at the same time, $(\star)$ is expressed as

$$\begin{aligned}\mathbb{E}\left[\mathbb{P}(\text{SINR}^{MD}>\eta|r_i)\right]&=\int_{r_{k_h}}\cdots\int_{r_2}\int_{r_1}\mathbb{P}\left[\frac{\sum_{i=1}^{k_h}|h_i|^2r_i^{-\alpha}P_t}{\sigma_n^2+\mathcal{I}_a}>\eta|r_i\right]f(r_1,\cdots,r_{k_h})dr_1\cdots dr_{k_h},\\&\overset{(a)}{=}\int_{r_{k_h}}\cdots\int_{r_2}\int_{r_1}\mathbb{P}\left[\frac{\sum_{i=1}^{k_h}|h_i|^2r_i^{-\alpha}P_t}{\sigma_n^2+\mathcal{I}_a}>\eta|r_i\right]f(r_1)\cdots f(r_{k_h})dr_1\cdots dr_{k_h},\\&\overset{(b)}{=}\int_{r_{k_h}}\cdots\int_{r_2}\int_{r_1}\mathbb{P}\left[\frac{\sum_{i=1}^{k_h}|h_i|^2r_i^{-\alpha}P_t}{\sigma_n^2+\mathcal{I}_a}>\eta|r_i\right]\prod_{i=1}^{k_h}(2\pi\lambda_h)^i r_i\exp\left(-\pi\lambda_h\sum_{i=1}^{k_h}r_i\right)dr_i,\end{aligned}\tag{26}$$

where $f_{r_i}(r)$ is the PDF of the distance between the r_i^{th} nearest neighbor helper and the reference UD, defined in (6). Step (a) is derived based on properties of the independent joint probability of different distances from different helper, and step (b) is derived based on properties of exponential functions. Since the multiple integration in (26) is complicated and cannot be translated into a product of a one-dimensional integral, (26) can be solved in an alternative method. First, the distribution of the power of a single received desired signal, which is a quotient of the two random variables, channel gain to the path loss function is derived. Then, the distribution of the power of the summation received desired signals is derived. However, $(\star)$ in (25) is derived as follow

$$\begin{aligned}\mathbb{E}\left[\mathbb{P}\left(\text{SINR}^{MD}>\eta|r_i\right)\right]&=\mathbb{E}\left[\mathbb{P}\left(\frac{\sum_{i=1}^{k_h}|h_i|^2r_i^{-\alpha}P_t}{\sigma_n^2+\mathcal{I}_a}>\eta|r_i\right)\right],\\&=\mathbb{E}\left[\mathbb{P}\left(\sum_{i=1}^{k_h}\frac{|h_i|^2}{r_i^{\alpha}}>\eta\left(\frac{\sigma_n^2}{P_t}+\frac{\mathcal{I}_a}{P_t}\right)|r_i\right)\right].\end{aligned}\tag{27}$$

Let $U_i=r_i^{\alpha}$. Given the PDF of r_i in (6), the PDF of U_i is obtained as

$$f_{U_i}(u)=\frac{2\pi\lambda_h}{\alpha}u^{\frac{2-\alpha}{\alpha}}\exp\left(-\pi\lambda_h u^{\frac{2}{\alpha}}\right),\quad u>0\tag{28}$$

Let $T=\sum_{i=1}^{k_h}t_i$ is defined as the sum of the random variables t_i, where $t_i=\frac{|h_i|^2}{U_i}$ is the ratio of two independent random variables having the respective PDFs, $f_{X_i}(x)$ and $f_{U_i}(u)$. Then the cumulative distribution function (CDF) $F_{t_i}(t)$ of the quotient is defined as

$$F_{t_i}(t)=\mathbb{P}\left(\frac{X_i}{U_i}\le t\right)=\mathbb{P}\left(X_i\le tU_i\right)=\int_0^{\infty}\left[\int_0^{tu}f_{X_i}(x)dx\right]f_{U_i}(u)du\tag{29}$$

By differentiating (29), $f_{t_i}(t)$ is obtained. For a special case $\alpha=4$, the transformation of $U_i=r_i^4$ is given by

$$f_{U_i}(u)=\frac{\pi\lambda_h}{2}u^{-\frac{1}{2}}\exp\left(-\pi\lambda_h u^{\frac{1}{2}}\right).\quad u>0\tag{30}$$

By applying (30) and the PDF of the channel gain h_i, into (29), the CDF of t_i is obtained as

$$F_{t_i}(t) = 1 - \frac{\pi^{3/2}\lambda_h \exp\left(\frac{\pi^2\lambda_h^2}{\sqrt{t}}\right)\text{erfc}\left(\frac{\pi\lambda_h}{2\sqrt{t}}\right)}{2\sqrt{t}}, \quad t > 0. \tag{31}$$

By differentiating (31), the PDF of t_i is obtained as

$$f_{t_i}(t) = \left(\frac{\pi^{7/2}\lambda_h^3}{8t^{5/2}} + \frac{\pi^{3/2}\lambda_h}{4t^{3/2}}\right)\exp\left(\frac{\pi^2\lambda_h^2}{4t}\right)\text{erfc}\left(\frac{\pi\lambda_h}{2\sqrt{t}}\right) - \frac{\pi^2\lambda_h^2}{4t^2}, \quad t > 0. \tag{32}$$

It can be seen from (29), it is hard to derive a simple close form for the PDF of T when $\alpha > 2$. For the sake of simplicity, it is assumed that $\alpha = 2$ in the following analysis, which is also adopted as a typical value of path loss exponent in different environments [18]. The PDF of $U_i = r_i^2$ is given by

$$f_{U_i}(u) = \pi\lambda_h \exp\left(-\pi\lambda_h u\right), \quad u > 0. \tag{33}$$

Then the PDF of t_i is obtained as

$$f_{t_i}(t) \;=\; \frac{\pi\lambda_h}{(\pi\lambda_h + t)^2}, \quad t > 0. \tag{34}$$

which is a special case of the Lomax distribution (Pareto type II), defined as

$$f(x) = \frac{c\mu^c}{(\mu + x)^{c+1}}, \quad x > 0. \tag{35}$$

where c is a positive integer shape parameter, and $\mu > 0$ is the scale parameter. (34) is similar to (35) when the shape parameter $c = 1$, and the scale parameter $\mu = \pi\lambda_h$. To determine the PDF of T, which is the sum of independent random variables t_i, the expectation $\mathbb{E}\left[\exp(-sT)\right]$ denoting the moment generating function or LT of T can be written as the product of individual LT functions. Therefore, the LT of (34) denoted as $\mathcal{L}[f_{t_i}(t)]$, is given by

$$\mathcal{L}\left[f_{t_i}(t)\right] = \int_0^\infty \frac{\mu\exp\left(-st\right)}{\left(\mu + t\right)^2}dt = \exp\left(\mu s\right)E_2(\mu s), \tag{36}$$

where $\mathcal{L}\left[f_{t_i}(t)\right]$ is an analytical function in the right half plane $\Re(s) > 0$, and the LT of T is written as

$$\begin{aligned}\mathcal{L}\left[f_T(t)\right] &= \int_0^\infty \exp\left(-st\right)f_T(t)dt = \left(\mathcal{L}\left[f_{t_i}(t)\right]\right)^{k_h} \\ &= \left(\exp\left(\mu s\right)E_2(\mu s)\right)^{k_h}. \end{aligned} \tag{37}$$

In order to determine $f_T(t)$, the inversion of (37) is derived in Appendix C, and $f_T(t)$ is obtaind as

$$f_T(t) = \frac{1}{\pi\mu}\int_0^\infty \Upsilon_{k_h}(x)\exp\left(-x\left(k_h + \frac{t}{\mu}\right)\right)dx. \tag{38}$$

Therefore, the solution of $(\star)$ in (25) is given by

$$\begin{aligned}\mathbb{E}[\mathbb{P}\left(\text{SINR}^{MD} > \eta\right)] &= \frac{1}{\pi\mu}\int_{V}^{\infty}\int_{0}^{\infty}\Upsilon_{k_h}(x)\exp\left(-x\left(k_h+\frac{t}{\mu}\right)\right)dxdt,\\ &= \int_{0}^{\infty}\frac{\Upsilon_{k_h}(x)}{\pi x}\exp\left(-x\left(k_h+\frac{V}{\mu}\right)\right)dx,\\ &= \int_{0}^{\infty}\frac{\Upsilon_{k_h}(x)}{\pi x}\exp\left(-x\left(k_h+\frac{\eta}{\mu\eta_o}+\frac{\eta}{\mu}\mathcal{I}_n\right)\right)dx,\\ &= \int_{0}^{\infty}\frac{\Upsilon_{k_h}(x)}{\pi x}\exp\left(-x\left(k_h+\frac{\eta}{\mu\eta_o}\right)\right)\mathcal{L}_{\mathcal{I}_n}\left(\frac{x\eta}{\mu}\right)dx, \end{aligned} \tag{39}$$

where $V=\eta(\eta_o^{-1}+\mathcal{I}_n)$, and $\mathcal{L}_{\mathcal{I}_n}\left(\frac{x\eta}{\mu}\right)$ is the LT of the aggregation interference which is derived in Appendix D and given by

$$\mathcal{L}_{\mathcal{I}_n}\left(\frac{x\eta}{\mu}\right)=\exp\left(-x\eta\log\left(\frac{x\eta+\mu R^2}{x\eta+\mu R_{\mathcal{A}}^2}\right)\right). \tag{40}$$

By substituting (40) into (39), and (39) into (25), $\mathbb{P}_{out,2}^{MD}$ is obtained. ■

C. Frequency Reuse

Frequency reuse by allowing multiple content-transmissions UD-groups using the same frequency can improve the capacity and the spectral efficiency [22]. In our proposed MDSD method, in the aim of reducing the interference caused by the UDs that use the same frequency allocation, frequency band factor $\omega \geq 1$, is defined as the number of different frequency bands that can be used orthogonally in the system. In other words, ω indicates the number of different UD-groups serving delivering same content with the different frequency band. By dividing the whole cell area $\mathcal{B}$ into multiple small cooperative areas with size of $\mathcal{A}$, the maximum number of different frequency band will be introduce as $\omega=\frac{|\mathcal{B}|}{|\mathcal{A}|}$, which indicates the number of UD-groups serving delivering same content with the different frequency band. The increase of ω leads to decrease of the outage probability, since the interference from the interfering UDs that use the same frequency band decreases. On the other hand, increasing ω causes the decreases of ASE at the same time since the spectrum will be divided among UD-groups without reusing, which will be the trade-off between ASE and the outage probability will results in

- The larger the value of ω indicating to the lower of the ASE. However, the cooperative area having the same frequency allocations will be far from each other resulting in lower interference, thereby the lower outage probability.
- The lower the value of ω indicating to the higher of the ASE. However the cooperative area having the same frequency allocations will be closer from each other resulting in higher interference, thereby a higher outage probability.

The density of the interfering UDs using the same frequency band is defined as λ_h/ω [22].

Lemma 2.1. *For Rayleigh fading, a frequency band ω and the path loss exponent α=2, the outage probability, denoted as $\mathbb{P}^{MD}_{out,2,\omega}$, is given by*

$$\mathbb{P}^{MD}_{out,2,\omega}=1-\sum_{K=1}^{K_{\mathcal{B}}}\sum_{k_h=1}^{K}\binom{K_{\mathcal{B}}}{K}\left(\frac{|\mathcal{A}|}{|\mathcal{B}|}\right)^{K}\left(1-\frac{|\mathcal{A}|}{|\mathcal{B}|}\right)^{K_{\mathcal{B}}-K}\binom{K}{k_h}\vartheta_l{}^{k_h}(1-\vartheta_l)^{K-k_h}\int_0^{\infty}\frac{\exp\left(-x(k_h+\frac{\eta}{\eta_o\mu}+\frac{\eta}{\omega}\nu)\right)}{\pi x}\Upsilon_{k_h}(x)dx. \quad (41)$$

Proof. As the reference UD at the origin o would be served by its multiple closest helpers from the point process Φ within the distance $R_{\mathcal{A}}$, the interference part is obtained as (40)

$$\mathcal{L}_{\mathcal{I}_n}\left(\frac{x\eta}{\mu\omega}\right)=\exp\left(-\frac{x\eta}{\omega}\log\left(\frac{x\eta+\mu R^2}{x\eta+\mu R^2_{\mathcal{A}}}\right)\right) \quad (42)$$

The reset of the proof is the same procedure as deriving **Theorem 2** started from (27) to (40). ■

D. Area spectral efficiency

ASE is an important metric in wireless communication, which is defined as the average number of bits transmitted per second per Hz per square meter (m^2). From Shannon's capacity, we investigate ASE under two cases, i.e. guaranteed ASE with guaranteed data rate ($\mathcal{R}^{MD}_t$) and the achievable ASE based on the average achievable link data rate ($\mathcal{R}^{MD}_a$) as follows. The guaranteed ASE is defined as

$$\text{ASE}_t=\lambda\frac{|\mathcal{B}|}{|\mathcal{A}|}(1-\mathbb{P}^{MD}_{out})\mathcal{R}^{MD}_t \quad (43)$$

where $(1-\mathbb{P}^{MD}_{out})$ is the success transmission probability denoted as $\mathbb{P}^{MD}_{sec}$, and $\mathcal{R}^{MD}_t$ is a constant transmission data rate, calculated based on guaranteeing the SINR threshold η, and is given by

$$\mathcal{R}^{MD}_t=\log_2(1+\eta). \quad (44)$$

Then, the achievable ASE is obtained as

$$\text{ASE}_a=\lambda\frac{|\mathcal{B}|}{|\mathcal{A}|}(1-\mathbb{P}^{MD}_{out})\mathcal{R}^{MD}_a \quad (45)$$

where $\mathcal{R}^{MD}_a$ is the achievable data rate of the system, and is given by

$$\mathcal{R}^{MD}_a=\log_2(1+\text{SINR}^{MD}). \quad (46)$$

It is clear that the ASE is a function of SINR, $\mathbb{P}^{MD}_{out}$, and $R_{\mathcal{A}}$ since $|\mathcal{A}|=\pi R^2_{\mathcal{A}}$. The outage probability is an increasing function of η. On the other hand, the transmission data rate $\log_2(1+\eta)$ is an increasing function of η. Therefore, the ASE is a decreasing function of $\mathbb{P}^{MD}_{sec}$ and an increasing function of $\mathcal{R}^{MD}_t$. Furthermore, it is a decreasing function of $R_{\mathcal{A}}$. Anticipate the optimal value of ω for achieving the maximum ASE will be study in future work. Investigate and evaluated by numerical results in section IV.

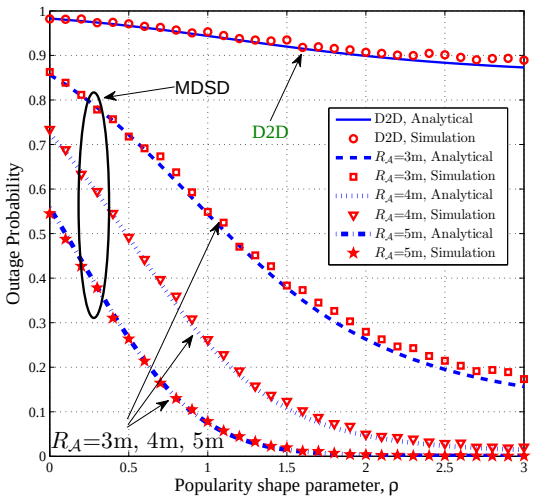

Figure 3: Outage probability versus Popularity shape parameter ρ

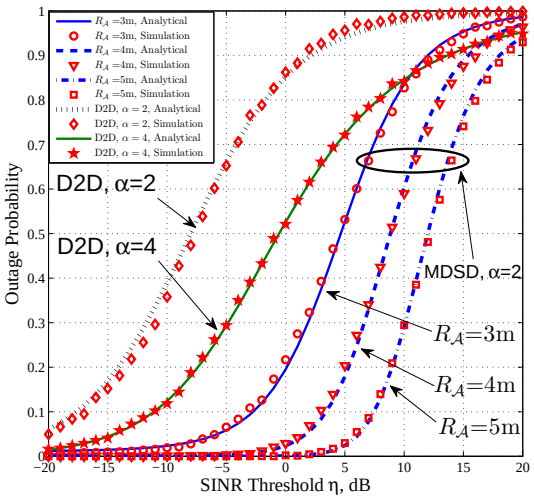

Figure 4: Outage probability versus SINR threshold value η

IV. Results and Discussions

In this section, numerical and simulation results are shown to evaluate outage probability performance and ASE in which the effects of the popularity shape parameter, ρ, and the SINR target threshold, η, are presented. The system performance is numerically evaluated and verified by simulation results. In the simulation, the UDs are randomly generated and distributed according to the Poisson distribution with intensity λ. An independent equivalence of the exponential random variable as the power of the Rayleigh fading channel is generated for each link. The contents are generated according to the Zipf distribution, and the desired content requested by the reference UD is a most popular one, i.e. $\mathcal{C}_1$ with caching probability ϑ_1. The radius of the whole cell area $\mathcal{B}$ is fixed to R=25 meter (m), and the transmission power P_t is normalized to one. The simulation results are carried out via a one million trials per point.

Fig. 3 depicts the effect of the popularity shape factor ρ on the outage probability performance, where $\eta = 0$ dB, $\lambda = 0.25$ UD/m^2, and $\eta_o = 20$ dB are assumed. It is seen from Fig. 3 that the outage probability decreases rapidly with the increase of ρ, as well as when $R_{\mathcal{A}}$ increases in MDSD. Increasing ρ means that a less number of contents are highly common requested, thereby it is probably there are multiple helpers holding the desired content in proximity. The performance gain of the outage probability is enhanced by approximately 30% when ρ= 0 and $R_{\mathcal{A}}$ increases from 3m to 5m, whereas the gain is enhanced by 42% when $\rho = 0.5$. In addition, the performance is enhanced by roughly 51%, 80%, and 90%, respectively for MDSD compared to single D2D when $\rho = 1.5$. In fact, when ρ= 0, all contents are uniformly distributed with equal probability and the outage probability will be higher than the Zipf

distribution with shape parameter $\rho > 0$. Furthermore, as $R_{\mathcal{A}}$ increases, outage probability decreases since the number of UDs inside $\mathcal{A}$ increases, thereby the chance of finding and downloading the desired content from the proximity UDs increases. However, the analytical results match well simulation results.

Fig. 4 shows the analytical and simulation results of the average outage probability as a function of SINR target threshold η in two cases: single D2D, and MDSD communications. ρ, λ, and η_o are assumed to be 2.5, 0.25 UD/m^2 and 20 dB, respectively. Different values of $R_{\mathcal{A}}$ (3m-5m) in MDSD, as well as the effect of different values of path loss exponent in a single D2D communication are considered. It can be seen that the performance is enhanced drastically for a single D2D communication with the increase of α. Although the outage probability is improved when α increases, it remains high in single D2D communication. For instance, in single D2D the performance is enhanced by 35% when η=0 dB with $\alpha = 4$ compared to $\alpha = 2$. The outage probability decreases significantly when the MDSD based method is used instead of single D2D based method, this is due to the benefit from the transmission diversity to combat the channel fading, as well as to download the desired content from neighbors. Moreover, outage probability decreases rapidly when $R_{\mathcal{A}}$ increases in MDSD method, since the number of helpers having the desired content increases, thereby it is probably increasing the chance of getting the desired content from neighbors. When $\eta = -5$ dB, the performance gain is enhanced around 60% for $R_{\mathcal{A}}$=3m and 65% for $R_{\mathcal{A}}$ =4m in MDSD compared to single D2D with $\alpha = 2$, whereas the outage probability gain is around 89% for MDSD with $R_{\mathcal{A}}$=5m and $\eta = 0$ dB compared to single D2D with $\alpha = 2$. In addition, the power gain is around 30 dB when $\mathbb{P}_{out} = 0.1$ and $R_{\mathcal{A}}$ =5m in MDSD compared to single D2D with $\alpha = 2$. It is seen that the simulation results match perfectly the analytical results, thereby validating the accuracy of the analysis.

Fig. 5 shows the performance of $\mathbb{P}_{out}$ when the different content popularity are considered in MDSD based method. It is assumed that $R_{\mathcal{A}}$ =5m, $\lambda = 0.25$ UD/m^2 and $\eta_o = 20$ dB. The results depict the performance for three content, least popular content $\mathcal{C}_{10}$ with probability ϑ_{10}, most popular content $\mathcal{C}_1$ with probability ϑ_1, and averaged popular content $\mathcal{C}_1 - \mathcal{C}_{10}$, i.e. any content from 1 to 10 when $\rho = 0$. It can be seen that when $\rho = 1$, the outage probability for ϑ_1 is less than ϑ_{10} and ϑ_{1-10} since it has a higher probability of occurrence (34.14%), while $\vartheta_{10} = 3.42\%$. Furthermore, $\vartheta_{1-10} = 10\%$ in uniform distribution of $\rho = 0$ since the number of different content cached in the different UDs is $L = 10$. As a comparison between highest popular content $\mathcal{C}_1$ and averaged popular one $\mathcal{C}_{1-10}$, it is seen that the average $\mathbb{P}_{out}$ is enhanced by roughly 15% and 44% for the target threshold $\eta = -20$ dB and $\eta = 0$ dB, respectively. However, the average gain of the content outage probability depends on the number of content that can be stored in the system, the density of UDs and the popularity shape parameter ρ.

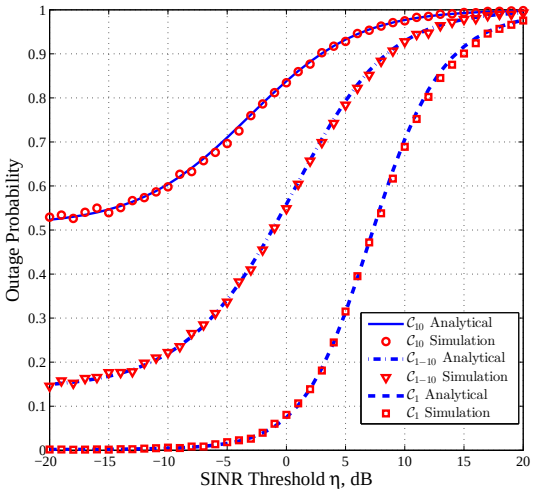

Figure 5: Outage probability for different content popularity ϑ_l

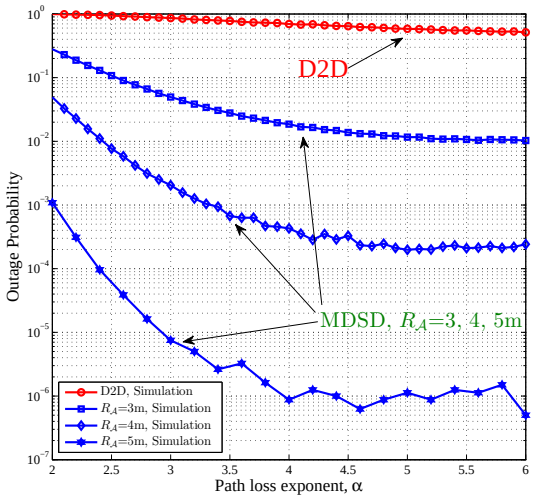

Figure 6: Outage probability versus path loss exponent α

In Fig. 6, the simulation results of the outage probability versus path loss exponent α are shown for different values of radius $R_{\mathcal{A}}$ in MDSD based method and the nearest helper in single D2D based method. $\eta_o = 20$ dB, $\eta = 0$ dB, $\lambda = 0.25$ UD/m^2, and $\rho = 2.5$ are assumed. It is seen from the figure that the outage probability decreases rapidly initially with the increase of the path loss exponent α, and then gradually get flat. This is because when α is large, the interference may be neglected so that channel noise dominates.

Fig. 7 plots the simulation results of the average outage probability in semi-log y-axis versus the density of UDs, λ UD/m^2, in the x- axis for path loss exponent $\alpha = 2$ and α =4, respectively. It can be seen that the performance is better for $\alpha = 4$ than $\alpha = 2$ because the interference decreases drastically as α increases. The outage probability decreases almost linearly with the increase of the density of UDs λ, since the number of helpers increases with the density within the cooperative distance $R_{\mathcal{A}}$, leading to the decrease of the outage probability due to more opportunities of downloading the desired content from proximity neighbors.

Fig. 8 illustrates the outage probability versus the target threshold η for $R_{\mathcal{A}} = \{3\text{m}, 5\text{m}\}$, $\rho = 2.5$, $\eta_o = 20$ dB, $\lambda = 0.25$ UD/m^2 and for different frequency bands $\omega = \{1, 2, 4\}$. It is seen from Fig. 8 that as ω increases, the outage probability decreases since the interference due to the interferer UDs using the same frequency band decreases. In other words, for a given SINR threshold, the outage probability is lower for $\omega = 2$ as compared to $\omega = 1$. Similarly, the outage probability is lower for $\omega = 4$ as compared to $\omega = 2$. This is indicated by the curves in Fig. 8 shifting to the right. $\omega = 1$ means that the frequency

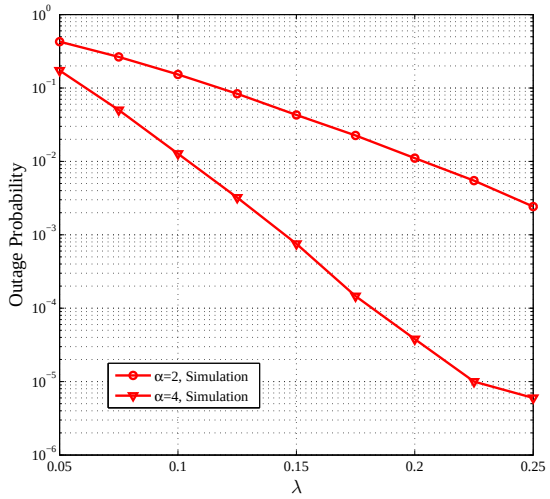

Figure 7: Outage probability versus density of UDs

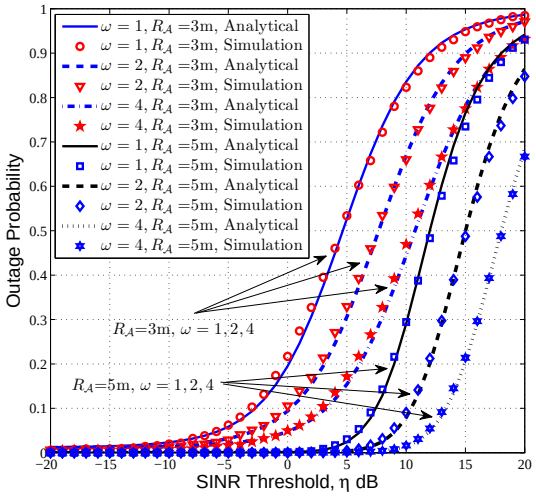

Figure 8: Outage probability versus η for different ω

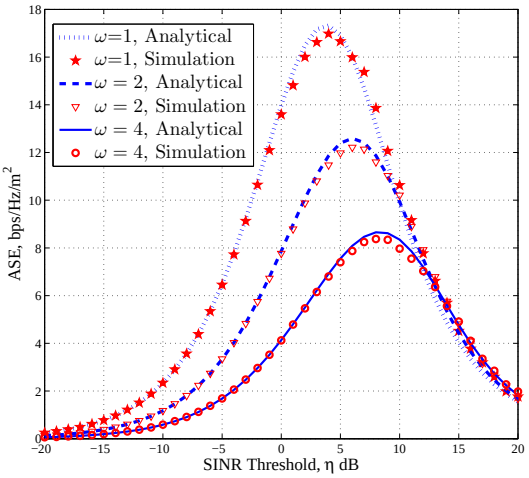

Figure 9: ASE versus η for different frequency band ω

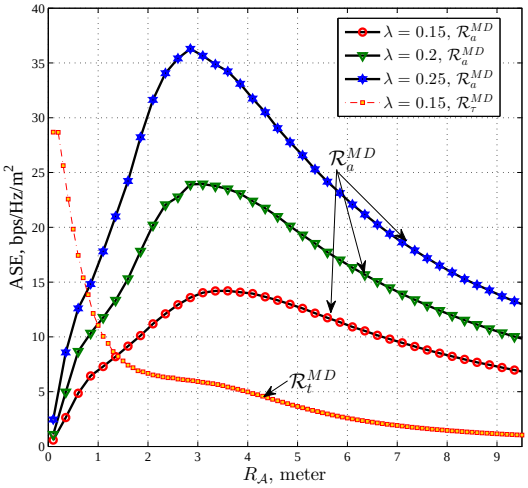

Figure 10: ASE versus $R_{\mathcal{A}}$ for different densities of UDs

used within $\mathcal{A}$ can be reused for all UDs outside $\mathcal{A}$, while $\omega = 2$ means that the frequency band is divided into two orthogonal sub-bands and each sub-band can be reused by $\frac{\omega}{2}$. In other words, when $\omega = 2$, the frequency band used in area $\mathcal{A}$ will be 1 or 2, and for the second radius will be 1,2,1,2, and so on. It can be seen that the gain performance is enhanced by 45% when $\eta = 10$ dB and frequency band $\omega = 4$ for $R_{\mathcal{A}}$ =5m compared to $R_{\mathcal{A}}$ =3m since the larger value of ω decreases the interference significantly.

Fig. 9 shows the impact of the frequency band ω on the ASE of the system when the fixed data rate $\mathcal{R}_t^{MD}$ is considered. $\rho = 2.5$, $\eta_o = 20$ dB, $\lambda = 0.25$ UD/m^2 and $R_{\mathcal{A}}$ =3m are assumed. It can be seen that as the number of frequency band ω increases, ASE decreases substantially. In fact, increasing the number of frequency bands will consume the system bandwidth via using different frequencies in different areas, which results in decreases of ASE. Furthermore, as ω increases, the interference due to the same resources decreases, which implies to improve the capability of the UD to detect the desired signal. That is the reason for the SINR threshold is shifted to the right when ω increases. It can be seen that the ASE increases with the increase of η until its optimal value of η and then decreases rapidly. This is due to the decrease of the success probability $(1 - \mathbb{P}_{out})$ at the same time of increasing $\mathcal{R}_t^{MD}$. The optimal value of SINR threshold takes from 4 dB to 8 dB.

In Fig. 10, simulation results of ASE versus a radius $R_{\mathcal{A}}$ of area $\mathcal{A}$ are shown. It is assumed that $\rho = 2.5$, $\eta_o = 20$ dB, $\eta = 0$ dB, $\lambda = \{0.15, 0.2, 0.25\}$ UD/m^2, and for equations (46), (48), respectively. It can be seen that ASE increases significantly to its optimal $R_{\mathcal{A}}$ for different densities and then decreases gradually after this value in case of achievable data rate $\mathcal{R}_a^{MD}$. In this case, the optimal value gives us an indication about the frequency reuse under a constraint outage probability since a trade-off between the ASE and the outage probability is required. In other words, under certain parameters, the selecting of the appropriate $R_{\mathcal{A}}$ is important to have high ASE. Moreover, it is seen that the optimal value $R_{\mathcal{A}}$ increases as the density of UDs decrease. This is because when the density of UD decreases, the possibility of finding the helpers who have the desired content decreases. Therefore, the optimal $R_{\mathcal{A}}$ increases in order to increase the number of helpers within the distance $R_{\mathcal{A}}$ thereby increases the success probability. It is also seen that the area spectral efficiency linearly decreases with the increase of $R_{\mathcal{A}}$ when the fixed data rate $\mathcal{R}_t^{MD}$ is considered. However, (43) decreases since the data rate is fixed to 1 bps/Hz and the factor $\frac{|\mathcal{B}|}{|\mathcal{A}|}$ decreases with the increase of $R_{\mathcal{A}}$.

Fig. 11 shows the ASE performance versus SINR target threshold η and for different values of $R_{\mathcal{A}}$. It is assumed that $\lambda = 0.25$ UD/m^2, $\eta_o = 20$ dB and $\rho = 2.5$. The most popular content $\mathcal{C}_1$ is considered with a probability of ϑ_1. It is seen from Fig. 11 that ASE is a non-linear function of η for fixed data rate $\log_2(1+\eta)$ and a linear decreasing for achievable data rate $\log_2(1+\text{SINR}^{MD})$. Generally speaking, ASE is a function of both outage probability and transmission data rate. As SINR threshold value η increases, the transmission data rate increases almost linearly. However, the outage probability also increases with SINR, i.e. the success probability decreases. Therefore, there is an optimal SINR threshold in which the ASE is the highest, and the peaks of the graph corresponding to different distances. It can be seen that ASE is approximately maximum around $R_{\mathcal{A}} = 3$m, which consider an optimum distance for frequency

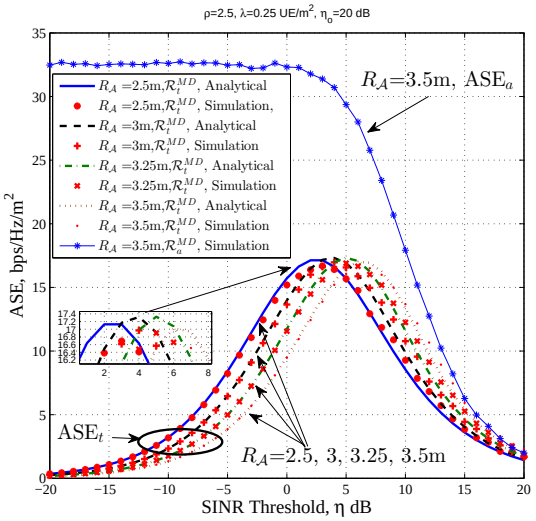

Figure 11: ASE versus SINR target threshold η

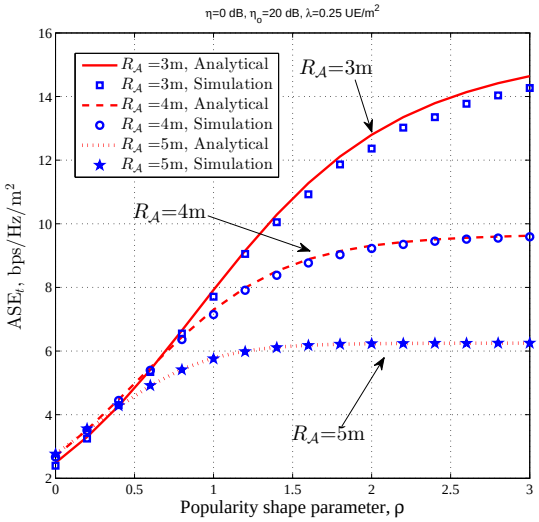

Figure 12: ASE versus popularity shape parameter ρ

reuse in this case, i.e. the optimum distance with the parameters that assumed. However, ASE depends on several parameters such as λ, ρ, $R_{\mathcal{A}}$ and SINR target threshold η.

In Fig. 12, ASE is plotted versus popularity shape parameter ρ for $\eta = 0$ dB, $\eta_o = 20$ dB, $\lambda = 0.25$ UD/m^2, for different $R_{\mathcal{A}}$ in MDSD, and for a fixed data rate $\mathcal{R}_t^{MD} = 1$ bps/Hz since $\eta = 0$ in (44). The dominant factor affecting the ASE is the outage probability that depends on the shape parameter ρ and the radius of $\mathcal{A}$. It can be seen that the ASE increases linearly with ρ since the probability ϑ_1 of the desired content $\mathcal{C}_1$ increases, i.e. the probability of finding the desired content increases. Increasing the distance $R_{\mathcal{A}}$ implies decreasing the outage probability leading to decreasing ASE in spite of increasing ρ. The ASE increases until its optimal value of ρ, and then stays flat. This is because increasing $R_{\mathcal{A}}$ increases the number of UDs inside the cooperative area $\mathcal{A}$, resulting in increase of the possibility of finding the desired content even though the probability of the desired content is small.

V. Conclusion

In this paper, we proposed a novel content delivery method based on D2D communication in which a spatial transmission diversity is applied. By using tools from stochastic geometry, a general theoretical expression of the outage probability for a single D2D communication was derived. A closed-form expression was further obtained for the outage probability in special cases where the path loss exponent $\alpha = 4$ and $\alpha = 2$, respectively. In addition, a closed-form expression of the outage probability as a single integral for MDSD based method was derived, where the distribution of the summed received the

desired signal from multiple helpers was obtained. The analysis was further extended to characterize a frequency reuse and the area spectral efficiency. Ultimately, extensive simulation and numerical results were presented to show the relationship between the outage probability, area spectral efficiency and various parameters, such as the SINR threshold η, the popularity shape parameter ρ, the frequency band factor ω, the radius of the cooperative area $R_{\mathcal{A}}$, etc. The conclusions are drawn as follows.

1) The impact of the popularity shape parameter ρ on the system performance has been shown, where the increase of ρ implies to decreases the outage probability significantly.
2) Reduces the interference via increase of ω implies to decreases the outage probability. On the other hand, the ASE will decreases since the available resources will be divided by ω.
3) The optimal SINR threshold value of the ASE was obtained when the different frequency band ω were applied. It can be concluded that the optimal SINR threshold is improved with the ω since the outage probability decreases.
4) The impact of selecting appropriate $R_{\mathcal{A}}$ on the ASE was shown for different density of UD, where the optimal value of $R_{\mathcal{A}}$ gives us an indication for a valuable frequency reuse.

APPENDIX A
PROOF OF THE LAPLACE TRANSFORM IN THEOREM 1

The LT of the aggregation interference $\mathcal{I}_n$ can be expressed as

$$\begin{aligned}
\mathcal{L}_{\mathcal{I}_n}(s) &= \mathbb{E}_{\mathcal{I}_n}\left[\exp\left(-s\mathcal{I}_a\right)\right] = \mathbb{E}_{\Phi,g_k}\left[\exp\left(-s\sum_{k\in\Phi/\mathcal{A}}|g_k|^2 r_k^{-\alpha}\right)\right], \\
&\overset{(a)}{=} \mathbb{E}_{\Phi,\{g_k\}}\left[\prod_{k\in\Phi/\mathcal{A}}\exp\left(-s|g_k|^2 r_k^{-\alpha}\right)\right], \\
&\overset{(b)}{=} \mathbb{E}_{\Phi}\left[\prod_{k\in\Phi/\mathcal{A}}\mathbb{E}_g\left[\exp\left(-s|g|^2 r_k^{-\alpha}\right)\right]\right], \\
&\overset{(c)}{=} \exp\left(-2\pi\lambda_h\int_r^{\infty}\left[1-\mathbb{E}_g\left[\exp\left(-s|g|^2 v^{-\alpha}\right)\right]\right]v dv\right). \qquad \text{(A.1)}
\end{aligned}$$

where (a) results from properties of exponential functions, (b) follows from the property of that all the g_k are i.i.d in the PPP, and (c) results from the definitions of the probability generating functional (PGFL) of the PPP Φ [21], where the PGFL of the function $f(x)$ is given by

$$\mathbb{E}\left[\prod_{x\in\Phi}f(x)\right] = \exp\left(-\lambda\int_{\mathbb{R}^2}\left(1-f(x)\right)dx\right). \qquad \text{(A.2)}$$

It is assumed that the interference channel is Rayleigh fading channel with channel gain PDF given by $f(g) = \exp(-g)$, therefore $\mathcal{L}_{I_n}(\eta r^{\alpha})$ is defined as

$$\begin{aligned}\mathcal{L}_{\mathcal{I}_n}(\eta r^{\alpha}) &= \exp\left(-2\pi\lambda_h \int_r^{\infty}\left(1-\frac{1}{1+sv^{-\alpha}}\right)vdv\right),\\ &= \exp\left(-2\pi\lambda_h \int_r^{\infty}\frac{\eta}{\eta+\left(\frac{v}{r}\right)^{\alpha}}vdv\right).\end{aligned} \tag{A.3}$$

By letting $y = \left(\frac{v}{r}\right)^2$, the limits of integral will be 1 and ∞. Since the interference is for limited area within R, therefore the integral is limited to R and (A.3) becomes

$$\mathcal{L}_{\mathcal{I}_n}(s=\eta r^{\alpha}) = \exp\left(-\pi\lambda_h r^2 \int_1^{R}\frac{\eta}{\eta+y^{\frac{\alpha}{2}}}dy\right) = \exp\left(-\pi\lambda_h r^2 \kappa(\eta,\alpha)\right), \tag{A.4}$$

where

$$\kappa(\eta,\alpha) = \int_1^{R}\frac{\eta}{\eta+y^{\frac{\alpha}{2}}}dy \tag{A.5}$$

APPENDIX B
PROOF OF THE HIT PROBABILITY IN THEOREM 2

For the hit probability $\mathcal{H}_l$ in (31), it is assumed that the reference receiver requests a specific content $\mathcal{C}_l$ with probability ϑ_l from k_h helpers at the same time, therefore $\mathcal{H}_l$ is given by

$$\mathcal{H}_l = \mathbb{P}_K \times \mathbb{P}_{k_h}, \tag{B.1}$$

where $\mathbb{P}_K$ is the probability that there are K UDs out of $K_{\mathcal{B}}$ inside the area $\mathcal{A}$, which is given by

$$\mathbb{P}_K = \binom{K_{\mathcal{B}}}{K}\left(\frac{|\mathcal{A}|}{|\mathcal{B}|}\right)^{K}\left(1-\frac{|\mathcal{A}|}{|\mathcal{B}|}\right)^{K_{\mathcal{B}}-K}, \tag{B.2}$$

where $\frac{|\mathcal{A}|}{|\mathcal{B}|} = \frac{\pi R_{\mathcal{A}}^2}{\pi R^2}$ is the ratio of the small area $\mathcal{A}$ to the whole cell area $\mathcal{B}$. $K_{\mathcal{B}}$ is the total number of UDs within the whole area $\mathcal{B}$, which is randomly generated and follow a Poisson distribution with probability as

$$\mathbb{P}(K_{\mathcal{B}} = n) = \frac{(\lambda\mathcal{B})^n \exp(-\lambda\mathcal{B})}{n!} \tag{B.3}$$

$\mathbb{P}_{k_h}$ is the probability that there are k_h helpers having the content $\mathcal{C}_l$ out of K UDs inside $\mathcal{A}$, which is given by

$$\mathbb{P}_{k_h} = \binom{K}{k_h}{\vartheta_l}^{k_h}(1-\vartheta_l)^{K-k_h}. \tag{B.4}$$

Substituting (B.4) and (B.2) into (B.1), one obtains

$$\mathcal{H}_l = \binom{K_{\mathcal{B}}}{K}\left(\frac{|\mathcal{A}|}{|\mathcal{B}|}\right)^{K}\left(1-\frac{|\mathcal{A}|}{|\mathcal{B}|}\right)^{K_{\mathcal{B}}-K}\binom{K}{k_h}{\vartheta_l}^{k_h}(1-\vartheta_l)^{K-k_h} \tag{B.5}$$

APPENDIX C
PROOF OF THE INVERSE LAPLACE TRANSFORM OF $f_T(s)$

To determine the inverse LT of (43), the Bromwich integral is used as follow

$$f_T(t) \;=\; \mathcal{L}^{-1}\left[f_T(s)\right]\frac{1}{2j\pi}\lim_{T\to\infty}\int_{\epsilon-jT}^{\epsilon+jT}\exp\left(st\right)\left(\exp\left(\mu s\right)E_2(\mu s)\right)^{k_h}ds, \tag{C.1}$$

where $j=\sqrt{-1}$ and $t>0$. The integral is evaluated along the path from $s=\epsilon-jT$ to $s=\epsilon+jT$ for an arbitrary real ϵ such that this path lies in the region of convergence. A Bromwich contour is shown in Fig. 13 to solve (C.1). In this case, the integrals along paths $B_oC_oD_o$ and $H_oI_oA_o$ go to zero as $R_o\to\infty$. However, the integral along the entire contour will be equal to the sum of the residues of all isolated singular points enclosed by the contour. In order to evaluate the integral along path A_oB_o (complex inversion integral), the residues and the integrals along the paths D_oE_o, G_oH_o and $E_oF_oG_o$ are given by

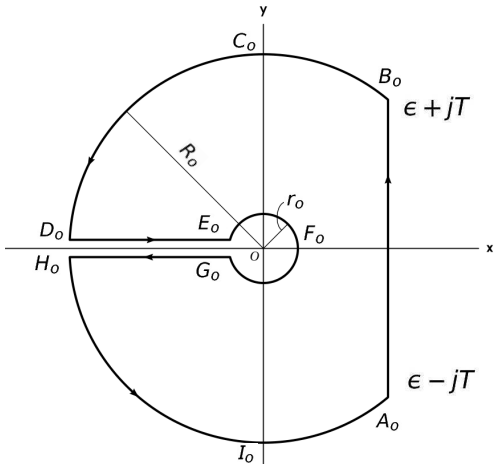

Figure 13: Bromwich contour

$$\int_{A_oB_o}+\int_{D_oE_o}+\int_{G_oH_o}+\int_{E_oF_oG_o}=\sum residues, \tag{C.2}$$

Since the integrals along $B_oC_oD_o$ and $H_oI_oA_o$ approach to zero, thus

$$\int_{A_oB_o}+\int_{D_oE_o}+\int_{G_oH_o}+\int_{E_oF_oG_o}=0, \tag{C.3}$$

consequently

$$\int_{A_oB_o}=-\int_{D_oE_o}-\int_{G_oH_o}-\int_{E_oF_oG_o} \tag{C.4}$$

which is the solution to (C.1). Now, the integrals along the paths D_oE_o, G_oH_o and $E_oF_oG_o$ are shown as follow

A. Integral along D_oE_o

Letting $s = v\exp(j\pi)$, v goes from R_o to r_o as s goes from $-R_o$ to $-r_o$, thus

$$\begin{aligned}\int_{D_oE_o} &= \int_{-R_o}^{-r_o} \exp\left(s(t+k_h\mu)\right)\left(E_2(\mu s)\right)^{k_h} ds, \\ &= \int_{R_o}^{r_o} \exp\left(-v(t+k_h\mu)\right)\left(E_2(\mu v\exp(j\pi))\right)^{k_h}\exp(j\pi)dv. \end{aligned} \tag{C.5}$$

B. Integral along G_oH_o

Letting $s = v\exp(-j\pi)$, v goes from r_o to R_o as s goes from $-r_o$ to $-R_o$, thus

$$\begin{aligned}\int_{G_oH_o} &= \int_{-r_o}^{-R_o} \exp\left(s(t+k_h\mu)\right)\left(E_2(\mu s)\right)^{k_h} ds, \\ &= \int_{r_o}^{R_o} \exp\left(-v(t+k_h\mu)\right)\left(E_2(\mu v\exp(-j\pi))\right)^{k_h}\exp(-j\pi)dv. \end{aligned} \tag{C.6}$$

C. Integral along $E_oF_oG_o$

Letting $s = r_o\exp(j\theta)$

$$\begin{aligned}&\int_{E_oF_oG_o} \exp\left(s(t+k_h\mu)\right)\left(E_2(\mu s)\right)^{k_h} ds, \\ &\quad= \int_{\pi}^{-\pi} \exp\left(r_o\exp(j\theta)(t+k_h\mu)\right)\left(E_2(\mu r_o\exp(j\theta))\right)^{k_h} r_o\exp(j\theta)d\theta. \end{aligned} \tag{C.7}$$

When r_o goes to 0, (C.7) becomes zero. Substituting (C.7), (C.6) and (C.5) into (C.2), (C.1) can be expressed as

$$\begin{aligned} f_T(t) &= \lim_{\substack{R_o\to\infty \\ r_o\to 0}} \frac{1}{2j\pi}\int_{R_o}^{r_o} \exp\left(-v(t+k_h\mu)\right)\left(E_2(\mu v\exp(j\pi))\right)^{k_h} dv \\ &+ \lim_{\substack{R_o\to\infty \\ r_o\to 0}} \frac{1}{2j\pi}\int_{r_o}^{R_o} \exp\left(-v(t+k_h\mu)\right)\left(E_2(\mu v\exp(-j\pi))\right)^{k_h} dv, \\ &= \lim_{\substack{R_o\to\infty \\ r_o\to 0}} \frac{1}{2j\pi}\int_{r_o}^{R_o} \exp\left(-v(t+k_h\mu)\right)\left(\left(E_2(\mu v\exp(-j\pi))\right)^{k_h} - \left(E_2(\mu v\exp(j\pi))\right)^{k_h}\right) dv, \\ &= \frac{1}{2j\pi}\int_{0}^{\infty} \exp\left(-v(t+k_h\mu)\right)\left(\left(E_2(\mu v\exp(-j\pi))\right)^{k_h} - \left(E_2(\mu v\exp(j\pi))\right)^{k_h}\right) dv. \end{aligned} \tag{C.8}$$

From the definition (5.1.7) in [23], the generalized $E_a(\cdot)$ is written as

$$E_m(-v \pm j0) = E_m(-v) \mp j\pi\frac{v^{m-1}}{\Gamma(m)}, \tag{C.9}$$

where $\Gamma(\cdot)$ is a gamma function. By substituting (C.9) into (C.8), and setting the variable $x = \mu v$ yields

$$f_T(t) = \frac{-1}{2j\pi\mu}\int_0^{\infty} \exp\left(-x\left(k_h + \frac{t}{\mu}\right)\right)\left[\left(E_2(-x) - j\pi x\right)^{k_h} - \left(E_2(-x) + j\pi x\right)^{k_h}\right] dx. \tag{C.10}$$

From Binomial theory and for real a and b,

$$(a-jb)^k-(a+jb)^k \;=\; -2j\sum_{n=1}^{k}\binom{k}{n}\sin\left(\frac{n\pi}{2}\right)a^{k-n}b^n, \tag{C.11}$$

where

$$\sin\left(\frac{n\pi}{2}\right)=\begin{cases}\pm 1 & n=odd\\ 0 & n=even\end{cases}$$

and $\binom{k}{n}=\frac{k!}{n!(k-n)!}$ stands for binomial coefficient. As a result, we define $\Upsilon_{k_h}(x)$ as

$$\Upsilon_{k_h}(x) \;=\; \sum_{n=1}^{k_h}\binom{k_h}{n}\sin\left(\frac{n\pi}{2}\right)\left(E_2(-x)\right)^{k_h-n}(\pi x)^n. \tag{C.12}$$

By substituting (C.12) into (C.10), the PDF of T is obtained.

APPENDIX D
PROOF OF THE LAPLACE TRANSFORM OF THEOREM 2

The LT of the aggregation interference $\mathcal{I}_n$ evaluated at $s=\frac{x\eta}{\mu}$ is expressed as

$$\mathcal{L}_{\mathcal{I}_n}\left(\frac{x\eta}{\mu}\right)=\mathbb{E}_{\mathcal{I}_n}\left(\exp\left(\frac{-x\eta}{\mu}\mathcal{I}_n\right)\right) \tag{D.1}$$

By using the same derivation procedure as (A.1)-(A.3), and setting the limits of integral with $R_{\mathcal{A}}$ and R, (D.1) is derived as

$$\begin{aligned}\mathcal{L}_{\mathcal{I}_n}(s=\frac{x\eta}{\mu}) &= \exp\left(-2\pi\lambda_h\int_{R_{\mathcal{A}}}^{R}\frac{s}{s+v^2}vdv\right),\\ &= \exp\left(-\pi\lambda_h s\log\left(\frac{s+R^2}{s+R_{\mathcal{A}}^2}\right)\right),\\ &= \exp\left(-x\eta\log\left(\frac{x\eta+\mu R^2}{x\eta+\mu R_{\mathcal{A}}^2}\right)\right)\end{aligned} \tag{D.2}$$

REFERENCES

[1] C. V. N. I. Cisco, "Global mobile data traffic forecast update, 2015–2020," *white paper*, 2017.

[2] Y. Cao, T. Jiang, and C. Wang, "Cooperative device-to-device communications in cellular networks," *IEEE Wireless Communications*, vol. 22, no. 3, pp. 124–129, June 2015.

[3] H. Zhu, "Radio resource allocation for ofdma systems in high speed environments," *IEEE Journal on Selected Areas in Communications*, vol. 30, no. 4, pp. 748–759, 2012.

[4] ——, "Performance comparison between distributed antenna and microcellular systems," *IEEE Journal on Selected Areas in Communications*, vol. 29, no. 6, pp. 1151–1163, 2011.

[5] J. Wang, H. Zhu, and N. J. Gomes, "Distributed antenna systems for mobile communications in high speed trains," *IEEE Journal on Selected Areas in Communications*, vol. 30, no. 4, pp. 675–683, 2012.

[6] Y. Zhou, L. Chen, C. Yang, and D. M. Chiu, "Video popularity dynamics and its implication for replication," *Multimedia, IEEE Transactions on*, vol. 17, no. 8, pp. 1273–1285, 2015.

[7] K. Shanmugam, N. Golrezaei, A. G. Dimakis, A. F. Molisch, and G. Caire, "Femtocaching: Wireless content delivery through distributed caching helpers," *IEEE Transactions on, Info. Theory*, vol. 59, no. 12, pp. 8402–8413, 2013.

[8] S. Sun, Q. Gao, W. Chen, R. Zhao, and Y. Peng, "Recent progress of long-term evolution device-to-device in third-generation partnership project standardisation," *Communications, IET*, vol. 9, no. 3, pp. 412–420, 2015.

[9] M. Ji, G. Caire, and A. F. Molisch, "Wireless device-to-device caching networks: Basic principles and system performance," *Selected Areas in Com. IEEE Journal on*, vol. 34, no. 1, pp. 176–189, 2016.

[10] N. Golrezaei, A. Dimakis, and A. Molisch, "Wireless device-to-device communications with distributed caching," in *2012 IEEE International Symposium on, Info. Theory Proceedings (ISIT)*, July 2012, pp. 2781–2785.

[11] 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 Transactions on, Wireless Com.*, vol. 13, no. 7, pp. 3665–3676, 2014.

[12] N. Golrezaei, A. Dimakis, and A. Molisch, "Scaling behavior for device-to-device communications with distributed caching," *IEEE Transactions on, Info. Theory*, vol. 60, no. 7, pp. 4286–4298, July 2014.

[13] E. Erdogan and T. Gucluoglu, "Dual-hop amplify-and-forward multi-relay maximum ratio transmission," *Journal of Communications and Networks*, vol. 18, no. 1, pp. 19–26, Feb 2016.

[14] H. Min, S. Lee, K. Kwak, and D. Hong, "Effect of multiple antennas at the source on outage probability for amplify-and-forward relaying systems," *IEEE Transactions on Wireless Communications*, vol. 8, no. 2, pp. 633–637, Feb 2009.

[15] C. Li, J. Zhang, and K. B. Letaief, "Throughput and energy efficiency analysis of small cell networks with multi-antenna base stations," *IEEE Transactions on Wireless Communications*, vol. 13, no. 5, pp. 2505–2517, 2014.

[16] A. S. Daghal and Q. Z. Ahmed, "Video content delivery using multiple devices to single device communications," in *Vehicular Technology Conference (VTC Spring), 2016 IEEE 83rd.* IEEE, 2016, pp. 1–5.

[17] M. Cha, H. Kwak, P. Rodriguez, Y.-Y. Ahn, and S. Moon, "I tube, you tube, everybody tubes: analyzing the world's largest user generated content video system," in *Proceedings of the 7th ACM SIGCOMM conference on Internet measurement*. ACM, 2007, pp. 1–14.

[18] A. Goldsmith, *Wireless communications*. Cambridge University Press, 2005.

[19] D. Moltchanov, "Distance distributions in random networks," *Ad Hoc Networks*, vol. 10, no. 6, pp. 1146–1166, 2012.

[20] A. Jeffrey and D. Zwillinger, *Table of integrals, series, and products*. Academic Press, 2007.

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

[22] J. G. Andrews, F. Baccelli, and R. K. Ganti, "A tractable approach to coverage and rate in cellular networks," *IEEE Transactions on Communications*, vol. 59, no. 11, pp. 3122–3134, 2011.

[23] M. Abramowitz and I. A. Stegun, *Handbook of mathematical functions: with formulas, graphs, and mathematical tables,*. Courier Corporation, 1964, no. 55.