Cache-Aided Cooperative Device-to-Device (D2D) Networks: A Stochastic Geometry View

Junchao Ma, *Student Member, IEEE*, Lingjia Liu, *Senior Member, IEEE*, Bodong Shang, *Student Member, IEEE*, and Pingzhi Fan, *Fellow, IEEE*

***Abstract*— Caching is a promising technique for 5G networks to reduce the backhual traffic and increase the overall network efficiency. In this paper, we study the caching placement policy with consideration of cooperative transmission for a two-hop relay-enabled device-to-device (D2D) network. In the caching placement phase, the probabilistic caching placement policy is considered, and in the content transmission phase, the hybrid automatic repeat request (HARQ) scheme with soft information combining [i.e., energy accumulation (EA) and mutual-information accumulation (MIA)] is utilized to improve the content retrieval experience. Cache-aided successful transmission probability (CSTP) is adopted as the main performance metric in this paper. By using tools from stochastic geometry, analytical expressions for the CSTP under different transmission schemes are derived. In moderate or high SIR regime, the optimal caching placement policy is identified based on the analytical expression and the CSTP performance is maximized accordingly. Evaluation results suggest that the MIA-based caching strategy performs better as opposed to existing strategies in most cases, but this outperformance vanishes when the transmission environment becomes severe or when users' requests become concentrated.**

***Index Terms*— Caching, cache-aided successful transmission probability, device-to-device network, mutual-information accumulation, stochastic geometry.**

I. Introduction

A. Motivation

IT IS predicted that the monthly world mobile data traffic will approach 49 exabytes (1 exabyte = 1 billion gigabyte) per month in 2021 which is 6 times higher than 7.9 exabytes per month in 2016 [1]. Providing realible service for such a huge volume of mobile data traffic becomes a challenging problem for the implementation and development of 5G networks. As a promising technique, caching enables intermediate nodes between the content server and the end consumers to prefetch relevant contents according to different caching algorithms. Two major benefits can be achieved through the use of caching. From the network's perspective, the data traffic between the content server and the end user can be reduced. From the end user's perspective, the delay to fetch such content can be reduced, and thus the corresponding quality of experience (QoE) of retrieving the underlying content increases. In fact, the essence of caching can be regarded as saving the precious transmission resource through paying the price for the storage resource. Since the storage resource is relatively cheap, caching in wireless devices provides both positive technical and economical impacts for future 5G networks.

One of the major challenges of caching is that the storage capacity of the intermediate node is usually limited compared to the overall predicted mobile data. Therefore, the design of the caching algorithm, i.e., the caching placement policy is of crucial importance to the caching gains, and attracts a lot of attention in recent years. On the other hand, instead of focusing on the caching placement policy exclusively, transmission strategies should also be considered jointly. This is due to the fact that the content still needs to be delivered to the end user through radio transmission to complete the content retrieval process. Accordingly, transmission schemes are also important to maximize the underlying performance gain achieved by caching. From the end user's perspective, a content cannot be retrieved locally via caching if a cache miss occurs or the transmission from the caching node fails. Therefore, caching placement policy and transmission strategies are coupled in nature for caching networks and should be considered jointly for the design of caching placement policy.

In order to improve the overall performance, relay-enabled cooperation technique can be used in the transmission process [2]. More specifically, if the transmission from the caching node to the end user fails but a relay node near the receiver decodes the content successfully, then the relay can re-encode the content and transmit it to the end user [3]. By carefully selecting a relay user, the performance of the relay-enabled cooperation scheme always outperforms the conventional direct transmission [4], [5]. Together with relay-enabled cooperation technique, the hybrid automatic repeat request (HARQ) scheme with soft information combining can be adopted at the end user. Equipped with this technique,

Manuscript received February 12, 2019; revised June 6, 2019; accepted July 17, 2019. Date of publication July 29, 2019; date of current version November 19, 2019. The work of J. Ma and P. Fan is supported by NSFC project No.61731017, NSFC-NRF project No.61661146003, and the 111 Project No.111-2-14. The work of L. Liu and B. Shang is supported in part by National Science Foundation under grants ECCS-1802710, ECCS-1811497, and CNS-1811720. The associate editor coordinating the review of this article and approving it for publication was S. Bhashyam. *(Corresponding author: Lingjia Liu.)*

J. Ma is with the Key Lab of Information Coding and Transmission, Southwest Jiaotong University, Chengdu 610031, China, and also with Virginia Tech, Blacksburg, VA 24061 USA.

L. Liu and B. Shang are with the Bradley Department of Electrical and Computer Engineering, Virginia Tech, Blacksburg, VA 24061 USA (e-mail: ljliu@ieee.org).

P. Fan is with the Key Lab of Information Coding and Transmission, Southwest Jiaotong University, Chengdu 610031, China.

Color versions of one or more of the figures in this article are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TCOMM.2019.2931556

the receiver of the end user will have the ability to accumulate information from multiple receptions. Depending on the applied coding schemes utilized in the multiple transmissions, energy accumulation (EA) or mutual-information accumulation (MIA) can be achieved at the end user. In EA, the multiple transmissions use the same coding scheme (such as repetition coding in [6]) to encode the same content, and the messages in different transmissions are identical. At the receiver side, the end user receives redundant copies of the content. The user can decode the content if the accumulated energy exceeds the energy threshold which highly depends on the transmission power and channel condition. On the other hand, if different coding schemes are used in multiple transmissions (such as rateless fountain coding in [6]) to encode the same content, the messages in different transmissions are independent, and it is a kind of MIA at the receiver side. The end user can decode the content if the accumulated mutual-information exceeds the entropy of the source [7].

The difference between MIA and EA can be easily understood from a simple example that two transmitters send one message to a destination, each via an erasure channel with the erasure probability p_e. If repetition coding is used (two transmitters send redundant message), the destination can receive $(1-p_e^2)$ bits per transmission on average. On the contrary, if the two transmitters use different fountain codes, the transmitted messages are independent and the destination can receive $2(1-p_e)$ (always $\geq 1-p_e^2$) bits in one transmission,

In this paper, we focus on analyzing the performance of caching in two-hop device-to-device (D2D) networks [8], [9]. In D2D networks, users can fetch content from nearby peers directly without going through the BS. This can improve the end user's QoE in content retrieval because the transmission can achieve higher data rate and suffer lower delay due to the proximity of communicating devices [10]. At the same time, the transmission burden of the BS is alleviated. Depending on the caching placement policy and the underlying transmission scheme, a user can retrieve its desired content from itself (self-caching), from nearby peers (D2D caching), or from the BS (cache miss or D2D transmission failure). The caching gain under the D2D network can be maximized by optimizing the caching placement policy with consideration of various transmission schemes.

B. Related Work

Applications of caching technique have been widely studied in different scenarios. As an early work to apply caching in wireless network, [11] introduces the femtocaching, i.e., caching is deployed at the femto base stations to reduce the traffic of the backbone network. Femtocaching and D2D caching are jointly considered to further increase the caching gain in [12]. In addition, caching placement policy for heterogeneous networks is studied in [13], [14]. Assuming the cache can be retrieved as long as it is found locally, these work mainly focus on designing the caching placement policy without considering the impact of transmission failure. D2D caching with the consideration of mobile users' remain battery capacity is investigated in [15]. And [16] studies the cooperation of multiple network operators when determining what to cache. Channel assignment policy in D2D networks to minimize the average content delivery delay is characterized in [17].

It is important to note that [18] jointly studies the caching policy and transmission under D2D networks with the consideration of interference and noise. The closed-form expression for the optimal caching policy is found in some special cases. Further, [19] extends the work by considering the gain of self-caching. These work mainly focus on the transmission in a single-hop network while [20] extends the single-hop transmission into multi-hop applying retransmission. The information theory aspect of multi-hop transmission in caching area is studied in [21].

On the other hand, EA and MIA are important cooperative transmission strategies for wireless networks and receive a lot of attention in the literature. To be specific, the performance of EA and MIA in the multiple relay scenario are compared in [6]. [7] uses MIA to assist the resource allocation in multi-hop transmission network to minimize the end-to-end relay under limited energy and bandwidth budget. Later, Hao *et. al.* in [22] applies MIA in cognitive radio network to improve the delay performance of the secondary users. Also, in [23] EA and MIA are applied to increase secondary user transmission opportunity and to increase the primary user throughput. In addition, the benefit of MIA in coordination joint transmission (CoMP) is studied in [24].

C. Main Contributions

Different from the literature that caching placement policy and cooperative transmission scheme, especially the EA and MIA, are analyzed separately, in this paper we study the interplay between the caching placement policy and the cooperative transmission in relay-enabled D2D networks. Depending on the applied caching placement policy and transmission scheme, a user can fetch its desired content from itself, nearby D2D users, or the base station according to the placed content availability and the channel condition of the content transmission [19]. Our purpose is to maximize the probability that the end users' requests can be satisfied locally. This is achieved through optimizing the caching placement policy with the consideration of the transmission scheme. Here for relay-based networks, different cooperative transmission schemes should have the same performance in the first time slot because they share the same transmission process. However, in the second time slot, due to different accumulation strategies are used, these schemes will have different performance. Cache-aided successful transmission probability (CSTP) is utilized as the performance metric to reflect the mutual impact of caching and cooperative transmission [18]–[20]. Mathematically, CSTP can be defined as the sum of the probability that a receiver can retrieve its desired content from cache (including self-caching and D2D caching) weighted by the corresponding request probability. Using tools from stochastic geometry, we obtain the expressions for the CSTP with respect to the caching placement policy. And the analytical expressions of optimal caching placement policy under different transmission strategies can also be derived in the moderate or high SIR regime (i.e., when the SIR threshold $\theta > 1$ dB).

The detailed contributions of this paper can be listed in the following. Firstly, although a lot of papers have been carried out to study the caching and transmission, to the best of our knowledge, this is one of the first works to jointly analyze the performance of the caching and relay-based cooperative transmissions in D2D networks on the consideration of different HARQ techniques (EA and MIA) and optimize the caching placement policy. The joint consideration of caching and cooperative transmission could significantly improve the caching performance. Secondly, using tools from stochastic geometry, analytical expressions for the CSTP are characterized. In addition, optimal caching placement policies are derived based on analytical expressions in various operation regimes. Also, the analytical results are verified using Monte Carlo simulations and numerical evaluations. The results demonstrate that the CSTP performance is influenced by both the underlying caching placement policy and the transmission scheme. In addition, MIA based caching strategy performs better than existing non-cooperative strategies in most cases. However, the performance gain of MIA will vanish when the required SIR threshold becomes high and when the users' requests become concentrated to the most popular contents.

The rest of this paper is organized as follows. Section II presents the system model and the CSTP formulation. And an optimization problem is established to maximize the CSTP performance in this section. Section III then provides detailed analysis to maximize the CSTP performance. To be specific, the analytical expressions of the CSTP are derived at first. Then the CSTP is maximized by solving the optimization problem, and the corresponding closed-form expression of the optimal caching placement policy in the moderate and high SIR regime is characterized. The impacts of various system parameters on the performance of CSTP are investigated via a series of simulations in Section IV. Finally, the paper concludes in Section V.

II. CSTP Formulation

A. System Model

In this paper, we consider a cache-aided D2D network where wireless users are modeled by a homogeneous Poisson point process (HPPP) Φ_u with intensity of λ_u [25], [26]. The total content library these wireless users may require contains J contents, each of which is assumed to have a normalized size. Let f_j denote the index of the j-th content, and denote $p(j)$ as the popularity of content f_j. Along with the literature [17], [19], the popularity is modeled by Zipf distribution with skewness parameter γ, i.e., $p(j) = \frac{j^{-\gamma}}{\sum_{i=1}^{J} i^{-\gamma}}$.

The content retrieval in the cache-aided D2D network consists of a caching placement phase and a content transmission phase [17]. In the caching placement phase, each user independently and randomly caches relevant contents according to the content popularity. In this paper, the probabilistic caching policy with the probability vector $\mathbf{q} = \{q_1, \cdots, q_j, \cdots, q_J\} \in [0,1]^{1\times J}$ is assumed to be applied at each wireless user [18]–[20]. Thus, a content f_j is cached by a user with a specific probability $q_j \in [0,1]$. Each user is assumed to be equipped with a common caching capacity $M_S < J$ contents, and the caching placement policy should follow

$$\sum_{j=1}^{J} q_j \leq M_S. \tag{1}$$

In the content transmission phase, the slot-based protocol is assumed to be operated in the network. At a slot, each wireless user has the probability of ρ to be active. Active wireless users will initiate content requests to nearby wireless devices. Accordingly, the active wireless users are potential receivers of the underlying D2D network. On the other hand, wireless users have the probability of $1-\rho$ to be inactive. These inactive users will receive content requests from those active users and are the potential transmitters of the underlying D2D network. In this work we focus on a simplified case where wireless users cannot serve others while being served. Under this assumption, the potential receivers and potential transmitters respectively follow HPPP Φ_r with the intensity of $\lambda_r = \lambda_u \rho$ and Φ_t with the intensity of $\lambda_t = \lambda_u(1-\rho)$ based on the thinning rule of HPPP [27]. We restrict $\rho \in [0.2, 0.8]$ to avoid all users in the underlying D2D network being active ($\rho \to 1$) or inactive ($\rho \to 0$). To make our analysis more tractable, along with the literature [26], [28], the transmitters in the two time-slots are assumed to be independent and identically distributed, and the influence of spatial and temporal correlation of interference will be addressed in our future work.

In terms of radio access, the Rayleigh block fading is assumed for the underlying wireless channels, and shadow fading is not considered in this work. Furthermore, wireless users are assumed to use the same transmission power P across the network. Therefore, for a typical D2D receiver (destination user, DU) locating at the origin, the received signal-to-noise-plus-interference ratio (SINR) can be expressed as

$$\mathrm{SINR} = \frac{P|h_0|^2 r^{-\alpha}}{\delta_I^2 + I_k}. \tag{2}$$

In this equation, $h_0 \sim \mathcal{CN}(0,1)$ and r are respectively the Rayleigh fading coefficient, and the distance from X_0 to the origin, where X_0 is the location of the desired transmitter (source user, SU). $\alpha \geq 2$ is the underlying path loss exponent, and δ_I^2 is the noise power. $I_k = \sum_{X_k \in \Phi_t \setminus X_0} P|h_k|^2 D_k^{-\alpha}$ is the interference power suffered by the DU where D_k is the distance from the interfering user X_k to the origin. Since D2D networks usually operate in the interference limited regime, we can safely assume $I_k >> \delta_I^2 = 0$.

B. Mode of Operation

When the DU requires a content f_j, it first checks its self-caching status. If f_j is already cached by itself, then a self-caching hit event happens and the request can be satisfied immediately without any transmission or failure. Otherwise, if the self-caching misses, the DU attempts to fetch the content from nearby peers through D2D caching and transmission [18], [19]. Before establishing the D2D link, the DU should find an appropriate D2D transmitter who can provide the content f_j.

Like the named data networking (NDN), the D2D transmitter selection can be performed by DU broadcasting *Interest* packets to nearby users [29]. The users satisfying the following two conditions can respond the *Interest* and act as the D2D transmitters. Firstly, *the user should be inactive to ensure that it can act as a transmitter*; and the second condition is that *the required content f_j is cached in the user's memory.* If one or more potential transmitters satisfy the conditions, then the nearest transmitter is chosen as the SU to maximize the received SIR at the DU [17]. The SU selection process may fail if no D2D users cache the required content. In this work we treat this selection failure as D2D transmission failure by assuming the SU is infinity from the SU, i.e., $r \to \infty$. According to [27], the distance r between the SU and the DU follows the following distribution:

$$f_r(r) = 2\pi\lambda_t q_j r e^{-\pi\lambda_t q_j r^2}. \tag{3}$$

Due to the channel fading and path loss, the transmission may fail if the received SIR at the DU is lower than the decoding threshold θ. The delay constraint of the transmission is set to two time-slots in this paper, which means the D2D transmission is regarded as successful if the DU can decode the f_j successfully within two transmissions. Otherwise, the D2D transmission fails, and the DU's request cannot be satisfied by D2D caching.

To improve the overall network performance, the relay-enabled cooperative transmission is applied in the content transmission phase. And the decode-and-forward (DF) protocol is applied in the network. To be specific, if the transmission from SU to DU fails but an intermediate user near the DU decodes the content successfully in the first time-slot, such user re-encodes the content and transmits it to the DU in the second time-slot; and the intermediate user is called the relay user (RU). The selected RU needs to meet the following requirements. Firstly, *the potential RU should be active and requires f_j in the first time-slot*. Otherwise, the user cannot be selected as RU if it is inactive or requires a content other than f_j. For the former case that the user is inactive, the user should be selected as SU if f_j is cached because it satisfies the SU selection conditions, or acts as an interference transmitter if f_j is not cached. For the latter case that the user requires a content other than f_j, the user will treat the message related to content f_j as interference and discard it. The potential RUs satisfying this condition follow the HPPP with density of $\lambda_u \rho p(j)$. The second condition is that *the potential RU should locate in the relay area (RA)*. Here the RA is composed by a set of points that lay between the SU and the DU, i.e., the distance between the RU and the SU (r_1) and the distance between the RU and the DU (r_2) should not larger than that between the SU and the DU (r). This condition guarantees that the selected RU can provide a better channel condition in the second transmission, and it has a higher chance to successfully transmit f_j to the DU. As shown in Fig. 1, in this paper the RA is defined as a sector in which the vertex is the DU, the radius is r, and the intersection angle is $\frac{\pi}{2}$ which is bisected by the segment between DU and SU [27]. Define N_R as the number of potential RUs in the RA, and the probability that no RU can be found in the RA is [27]

$$\mathrm{P_r}(N_R = 0) = \exp(-\frac{\pi}{4}\lambda_u \rho p(j) r^2). \tag{4}$$

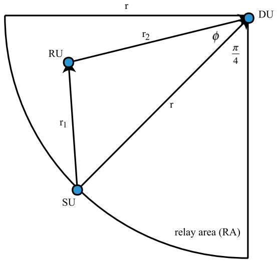


Fig. 1. Relay area in the RU selection process.

The third condition for the potential RUs is that *the selected RU decodes the content f_j successfully in the first time-slot.* This condition ensures that the RU can perform as a transmitter in the second time-slot. Otherwise, if the user fails to decode f_j, it will keep receiving the content to satisfy its requirement. If no RU can be selected in the RA, the SU retransmits the content f_j to the DU in the second time-slot. On the other hand, if one or more RUs are found, they can locate at any point in the RA, but it is challenging to formulate their exact locations. To make our analysis more tractable, the impact of RU selection on the overall network performance is analyzed by selecting the worst RU in the RA. The worst RU selection will be elaborated later. After a RU is selected, r_1, r_2, and the intersection angle between the bisection and r_2, ϕ, are all fixed and can be expressed as

$$\begin{aligned}
&0 < r_1, r_2 \le r,\\
&0 \le \phi \le \frac{\pi}{4},\\
&r_1^2 = r^2 + r_2^2 - 2rr_2\cos\phi.
\end{aligned} \tag{5}$$

For ease of illustration in the following derivations, (5) can be rewritten as $r_2 = \beta r$ where $\beta \in [0, 1]$ and $r_1^2 = (1 + \beta^2 - 2\beta\cos\phi)r^2$ if necessary [26] . At the DU side, the DU receives the transmitted messages and attempts to decode them based on the received SIR. If the decoding fails, it stores the messages and combines them with the subsequent received messages until the delay constraint is violated or the decoding succeeds. Based on the applied channel coding methods in the multiple transmissions, EA or MIA is achieved at the DU side to improve the decoding probability. If the decoding succeeds before the delay constraints, the DU sends an ACK feedback to the transmitter to stop transmission in the subsequent time-slots. The delay of decoding and the ACK feedback are assumed negligible in this paper.

Let P_A where $A \in \{\mathrm{EA}, \mathrm{MIA}\}$ denotes the successful transmission probability within two transmissions when scheme A is applied in the content retrieval process. The analytical expressions of P_A regarding to different transmission schemes will be given in the subsequent section. Then the CSTP, denoted by $\mathcal{T}_A$ [20], can be mathematically expressed as

$$\mathcal{T}_A = \sum_{j=1}^{J} p(j)\left[q_j \times 1 + (1 - q_j)P_A\right]. \tag{6}$$

To maximize the CSTP performance which highly depends on the caching placement policy and transmission strategy, the optimization problem can be formulated by

$$\begin{aligned}\mathcal{P}_A : \underset{\mathbf{q}}{\text{maximize}} \quad & \mathcal{T}_A, \\ \text{subject to} \quad & (1), \\ & 0 \le q_j \le 1, \end{aligned} \tag{7}$$

where the constraint (1) indicates that the caching capacity of every user is limited to M_S contents, and (7) means the caching probability should not exceed one.

III. CSTP Performance Analysis

In this section, we analyze and maximize the CSTP performance through solving the optimization problem $\mathcal{P}_A$. To be specific, the expressions of the successful transmission probability P_A, $A \in \{\text{EA}, \text{MIA}\}$, are respectively given, and the maximal CSTP performance are achieved for different transmission schemes by obtaining the optimal caching placement policies.

A. Derivation of P_{EA}

The successful transmission event occurs if one of the following cases happens under EA:

CASE I: The first transmission from the SU to DU succeeds, i.e., the received SIR exceeds the decoding threshold θ. The probability of this case is

$$P_{\text{EA\_I}}(r) = \text{P}_\text{r}(\text{SIR}_{\text{SD},1} > \theta) = e^{-\pi\lambda_t K\theta^\delta r^2}, \tag{8}$$

where $\delta = \frac{2}{\alpha} \le 1$, and $K = \frac{\pi\delta}{\sin\pi\delta}$. The derivation is widely applied in the literature and can be found in [27] (eq. (3.29)). Here it should be noted that the $\text{SIR}_{\text{SD},1}$ means the received SIR at the DU from SU in the first transmission. Similar notations are also available in the rest of the paper. For example, $\text{SIR}_{\text{RD},2}$ indicates the received SIR from RU to DU in the second transmission.

CASE II: The first transmission from the SU to DU fails, and no RU can be selected in the RA. In the second time-slot, the SU keeps transmitting f_j using the same encoding method so that the DU can accumulate energy from multiple receptions. By conducting maximal ratio combing (MRC) technique at the DU, the received SIR after the two transmissions can be seen as the SIR summation of the two transmissions [6]. Thus, the second transmission succeeds if the total SIR exceeds the decoding threshold θ, and the successful transmission probability is

$$\begin{aligned} &P_{\text{EA\_II}}(r) \\ &\quad= \text{P}_\text{r}(\text{SIR}_{\text{SD},1} < \theta, N_R = 0, \text{SIR}_{\text{SD},1} + \text{SIR}_{\text{SD},2} > \theta) \\ &\quad= \text{P}_\text{r}\left(N_R = 0\right)\text{P}_\text{r}\left(\theta - \text{SIR}_{\text{SD},2} < \text{SIR}_{\text{SD},1} < \theta\right) \\ &\quad\overset{(a)}{=} e^{-\frac{\pi}{4}\lambda_u\rho p(j)r^2}\int_0^\infty f_{\text{SIR}_{\text{SD},2}}(y)\int_{\theta-y}^{\theta} f_{\text{SIR}_{\text{SD},1}}(x)dxdy \\ &\quad= e^{-\frac{\pi}{4}\lambda_u\rho p(j)r^2}\int_0^\theta \pi\lambda_t K\delta y^{\delta-1}r^2 e^{-\pi\lambda_t K\left[y^\delta+(\theta-y)^\delta\right]r^2}dy, \end{aligned} \tag{9}$$

where (a) applies the result of (4) and the distribution of the received SIR that

$$\begin{aligned} f_{\text{SIR}_{\text{SD}}}(\omega) &= \frac{d\text{P}_\text{r}\left(\text{SIR}_{\text{SD}} < \omega\right)}{d\omega} = \frac{d\left(1 - e^{-\pi K\lambda_t r^2\omega^\delta}\right)}{d\omega} \\ &= \delta\pi K\lambda_t r^2\omega^{\delta-1}e^{-\pi K\lambda_t r^2\omega^\delta}. \end{aligned}$$

CASE III: The first transmission from the SU to DU fails and there exists at least one RU in the RA. The selected RU decodes the f_j in the first transmission via its self-caching or via the transmission from SU to RU. In the second time-slot, the RU successfully transmits f_j to DU using the same coding scheme; and the DU decodes content f_j by combining the two receptions. The location of the selected RU inevitably impacts the performance of CASE III. Since it is difficult to get the exact expression of the RU location, in the following theorem we provide the lower bound performance by fixing the RU at the worst case as follows

Theorem 1: The worst RU in the RA is when $r_2 = r$ and $\phi = \frac{\pi}{4}$, and the successful transmission probability in this case is given by

$$\begin{aligned} &P_{\mathit{EA\_III}}(r) \\ &\quad= P_r(\mathit{SIR}_{\mathit{SD,1}} < \theta, N_R > 0, \mathit{SIR}_{\mathit{SR,1}} > \theta, \mathit{SIR}_{\mathit{SD,1}} + \mathit{SIR}_{\mathit{RD,2}} > \theta) \\ &\quad= \left[1 - e^{-\frac{\pi}{4}\lambda_u\rho p(j)r^2}\right]\left[q_j + (1-q_j)\,e^{-\pi K\lambda_t\theta^\delta\left(2-\sqrt{2}\right)r^2}\right] \\ &\qquad\times\int_0^\theta \pi K\lambda_t\delta y^{\delta-1}r^2 e^{-\pi\lambda_t Kr^2\left[y^\delta+(\theta-y)^\delta\right]}dy. \end{aligned} \tag{10}$$

Proof: The Proof is in Appendix A. ■

Then, P_{EA} can be derived by summing the probabilities of the three cases and calculating the expectation over r, i.e.,

$$\begin{aligned} P_{\text{EA}} &= \text{E}_r\left[P_{\text{EA\_I}}(r) + P_{\text{EA\_II}}(r) + P_{\text{EA\_III}}(r)\right] \\ &= q_j F_{\text{EA}1}\left(q_j, \delta, p\left(j\right), \rho, \theta\right) + q_j^2 F_{\text{EA}2}\left(q_j, \delta, p\left(j\right), \rho, \theta\right) \\ &\quad + q_j\left(1-q_j\right)F_{\text{EA}3}\left(q_j, \delta, p\left(j\right), \rho, \theta\right) + \frac{q_j}{q_j + K\theta^\delta}, \end{aligned} \tag{11}$$

where

$$\begin{aligned} &F_{\text{EA}1}\left(q_j, \delta, p\left(j\right), \rho, \theta\right) \\ &\quad= \int_0^\theta \frac{K\delta y^{\delta-1}}{\left\{q_j + \frac{1}{4}\frac{\rho}{1-\rho}p\left(j\right) + K\left[y^\delta + \left(\theta-y\right)^\delta\right]\right\}^2}dy, \\ &F_{\text{EA}2}\left(q_j, \delta, p\left(j\right), \rho, \theta\right) \\ &\quad= \int_0^\theta \frac{K\delta y^{\delta-1}}{\left\{q_j + K\left[y^\delta + \left(\theta-y\right)^\delta\right]\right\}^2}dy \\ &\qquad- \int_0^\theta \frac{K\delta y^{\delta-1}}{\left\{q_j + \frac{1}{4}\frac{\rho}{1-\rho}p\left(j\right) + K\left[y^\delta + \left(\theta-y\right)^\delta\right]\right\}^2}dy, \end{aligned}$$

and

$$\begin{aligned} &F_{\text{EA}3}\left(q_j, \delta, p\left(j\right), \rho, \theta\right) \\ &\quad= \int_0^\theta \frac{K\delta y^{\delta-1}}{\left\{q_j + K\left[y^\delta + \left(\theta-y\right)^\delta\right] + K\theta^\delta\left(2-\sqrt{2}\right)\right\}^2}dy \\ &\qquad- \int_0^\theta \frac{K\delta y^{\delta-1}dy}{\left\{q_j + K\left[y^\delta + \left(\theta-y\right)^\delta\right] + K\theta^\delta\left(2-\sqrt{2}\right) + \frac{1}{4}\frac{\rho}{1-\rho}p\left(j\right)\right\}^2}. \end{aligned}$$

The detailed derivation process is in Appendix B.

In the moderate or high SIR regime, the following approximations hold

$$\begin{cases} q_j + K\left[y^\delta + (\theta - y)^\delta\right] \approx K\left[y^\delta + (\theta - y)^\delta\right], \\ \dfrac{1}{4}\dfrac{\rho}{1-\rho}p\left(j\right) + K\left[y^\delta + (\theta - y)^\delta\right] \approx K\left[y^\delta + (\theta - y)^\delta\right], \end{cases} \tag{12}$$

where $K > 1, \delta \leq 1$, and $0 \leq q_j \leq 1$. Let $f(y) = K[y^\delta + (\theta - y)^\delta]$, and via first-order derivation, we can get that $f(y) \in (K\theta^\delta, 2K(\frac{\theta}{2})^\delta) \gg 1$. Then the q_j appeared in the F_{EA1}, F_{EA2}, and F_{EA3} can be safely removed with negligible influence. Similarly, the term $0 \leq \frac{1}{4}\frac{\rho}{1-\rho}p(j) \leq 1$ also can be removed if $\rho \in [0.2, 0.8]$ and $0 \leq p(j) \leq 1$. As a result, the P_{EA} in (11) can be simplified to

$$P_{\text{EA}} = q_j F_{\text{EA}}\left(K, \delta, \theta\right), \tag{13}$$

where

$$F_{\text{EA}}\left(K, \delta, \theta\right) = \int_0^\theta \frac{\delta y^{\delta-1}}{K\left[y^\delta + (\theta - y)^\delta\right]^2}dy + \frac{1}{K\theta^\delta}. \tag{14}$$

B. Derivation of P_{MIA}

Similar to the EA scheme, there are the following three cases ensuring that the user can fetch its desired content via caching when MIA is applied.

CASE I: The first transmission from the SU to DU succeeds, i.e., the received SIR exceeds the threshold θ. The successful transmission probability of this case is the same as the CASE I of EA scheme that

$$P_{\text{MIA\_I}}(r) = P_{\text{EA\_I}}(r) = \text{P}_\text{r}(\text{SIR}_{\text{SD,1}} > \theta) = e^{-\pi\lambda_t K\theta^\delta r^2}. \tag{15}$$

CASE II: The first transmission from the SU to DU fails, and no RU can be selected in the RA. Thus the SU continues to transmit the file f_j to the DU in the second time-slot using rateless fountain coding. It would succeed finally if the accumulated mutual-information during the two transmissions exceeds the decoding threshold $\log_2(1+\theta)$ [22]. The successful transmission probability of this case is

$$\begin{aligned} &P_{\text{MIA\_II}}(r) \\ &= \text{P}_\text{r}\left[N_R = 0, \text{SIR}_{\text{SD,1}} < \theta,\right. \\ &\qquad \left.\log_2\left(1 + \text{SIR}_{\text{SD,1}}\right) + \log_2\left(1 + \text{SIR}_{\text{SD,2}}\right) \geq \log_2\left(1 + \theta\right)\right] \\ &= e^{-\frac{\pi}{4}\lambda_u\rho p(j)r^2} \int_0^\theta \pi\lambda_t K\delta y^{\delta-1} r^2 e^{-\pi\lambda_t K\left[y^\delta + \left(\frac{\theta-y}{1+y}\right)^\delta\right]r^2} dy. \end{aligned} \tag{16}$$

CASE III: The first transmission from SU to DU fails, and at least one RU can be selected in the RA. The selected RU decodes the f_j via self-caching or by transmission from SU. In the second time-slot, the RU re-encodes the f_j using rateless coding and successfully transmits it to the DU. It is worth noting that as stated in CASE III of the EA scheme, the location of RU ($r_2 = r$, $\phi = \frac{\pi}{4}$) provides the lower bound of successful transmission probability. This conclusion also holds for the MIA scheme. Thus, the successful transmission probability of this case under the worst RU is

$$\begin{aligned} &P_{\text{MIA\_III}}(r) \\ &= \text{P}_\text{r}\left[\text{SIR}_{\text{SD,1}} < \theta, N_R > 0, \text{SIR}_{\text{SR,1}} > \theta,\right. \\ &\qquad \left.\log_2\left(1 + \text{SIR}_{\text{SD,1}}\right) + \log_2\left(1 + \text{SIR}_{\text{RD,2}}\right) > \log_2\left(1 + \theta\right)\right] \\ &= \left[1 - e^{-\frac{\pi}{4}\lambda_u\rho p(j)r^2}\right]\left[q_j + \left(1 - q_j\right)e^{-\pi K\lambda_t\theta^\delta\left(2-\sqrt{2}\right)r^2}\right] \\ &\quad \int_0^\theta \pi K\lambda_t\delta y^{\delta-1} r^2 \ e^{-\pi\lambda_t K r^2\left[y^\delta + \left(\frac{\theta-y}{1+y}\right)^\delta\right]} dy. \end{aligned}$$

Overall, when MIA is applied, the successful transmission probability at the DU can be calculated by summing the three cases and then performing expectation over r, i.e.,

$$\begin{aligned} P_{\text{MIA}} &= \text{E}_r\left[P_{\text{MIA\_I}}(r) + P_{\text{MIA\_II}}(r) + P_{\text{MIA\_III}}(r)\right] \\ &= q_j F_{\text{MIA1}}\left(q_j, \delta, p\left(j\right), \rho, \theta\right) + q_j^2 F_{\text{MIA2}}\left(q_j, \delta, p\left(j\right), \rho, \theta\right) \\ &\quad + q_j\left(1 - q_j\right)F_{\text{MIA3}}\left(q_j, \delta, p\left(j\right), \rho, \theta\right), \end{aligned} \tag{17}$$

where

$$\begin{aligned} &F_{\text{MIA1}}\left(q_j, \delta, p\left(j\right), \rho, \theta\right) \\ &\quad = \frac{1}{\left[q_j + K\theta^\delta\right]} \\ &\qquad + \int_0^\theta \frac{K\delta y^{\delta-1}dy}{\left\{q_j + \frac{1}{4}\frac{\rho}{1-\rho}p\left(j\right) + K\left[y^\delta + \left(\frac{\theta-y}{1+y}\right)^\delta\right]\right\}^2}, \\ &F_{\text{MIA2}}\left(q_j, \delta, p\left(j\right), \rho, \theta\right) \\ &\quad = \int_0^\theta \frac{K\delta y^{\delta-1}dy}{\left\{q_j + K\left[y^\delta + \left(\frac{\theta-y}{1+y}\right)^\delta\right]\right\}^2} \\ &\qquad - \int_0^\theta \frac{K\delta y^{\delta-1}}{\left\{q_j + \frac{1}{4}\frac{\rho}{1-\rho}p\left(j\right) + K\left[y^\delta + \left(\frac{\theta-y}{1+y}\right)^\delta\right]\right\}^2}dy, \end{aligned}$$

and

$$\begin{aligned} &F_{\text{MIA3}}\left(q_j, \delta, p\left(j\right), \rho, \theta\right) \\ &= \int_0^\theta \frac{K\delta y^{\delta-1}}{\left\{q_j + K\left[y^\delta + \left(\frac{\theta-y}{1+y}\right)^\delta\right] + K\theta^\delta\left(2 - \sqrt{2}\right)\right\}^2}dy \\ &\quad - \int_0^\theta \frac{K\delta y^{\delta-1}dy}{\left\{q_j + K\left[y^\delta + \left(\frac{\theta-y}{1+y}\right)^\delta\right] + K\theta^\delta\left(2 - \sqrt{2}\right) + \frac{1}{4}\frac{\rho}{1-\rho}p\left(j\right)\right\}^2}. \end{aligned}$$

Similar with EA case, in the moderate or high SIR regime, the P_{MIA} can be simplified to

$$P_{\text{MIA}} = q_j F_{\text{MIA}}\left(K, \delta, \theta\right), \tag{18}$$

where

$$F_{\text{MIA}}\left(K, \delta, \theta\right) = \int_0^\theta \frac{\delta y^{\delta-1}}{K\left[y^\delta + \left(\frac{\theta-y}{1+y}\right)^\delta\right]^2}dy + \frac{1}{K\theta^\delta}. \tag{19}$$

C. CSTP Maximization

After deriving the P_A, the analytical expression of $\mathcal{T}_A$ can be given. Now we are in the position of solving the optimization problem $\mathcal{P}_A$. Due to the variable q_j in the denominator of the complex expression of P_A, the optimization problem $\mathcal{P}_A$ is hard to solve. However, we can get some instructive analysis in the moderate or high SIR regime. Take the EA case as an example. Since $\mathcal{T}_{\text{EA}}$ in this case is concave with respect to q_j, the Karush-Kuhn-Tucker (KKT) condition can be applied to solve the optimal caching placement policy [19]. Accordingly,

$$\mathcal{L}_{\text{EA}}(\mu_1, q_j) = -\mathcal{T}_{\text{EA}} + \mu_1 \sum_{j=1}^{J} q_j - M_S, \qquad (20)$$

and

$$\frac{\partial \mathcal{L}_{\text{EA}}(\mu_1, q_j)}{\partial q_j} = -p\,(j)\left[1 + (1 - 2q_j)\, F_{\text{EA}}\right] + \mu_1, \qquad (21)$$

where F_{EA} is shown in eq. (14). By $\frac{\partial \mathcal{L}_{\text{EA}}(\mu_1, q_j)}{\partial q_j} = 0$ we can yield the optimal caching policy, denoted by q^*_{EA}, as

$$q^*_{\text{EA}} = \frac{1}{2}\left[1 - \frac{1}{F_{\text{EA}}}\left(\frac{\mu_1}{p\,(j)} - 1\right)\right]. \qquad (22)$$

The value of μ_1 can be found by bi-section method with the constraint $\sum_{j=1}^{J} q_j \le M_S$ [19]. Similarly, the optimal caching policy in the MIA case, denoted by q^*_{MIA}, can be expressed as

$$q^*_{\text{MIA}} = \frac{1}{2}\left[1 - \frac{1}{F_{\text{MIA}}}\left(\frac{\mu_2}{p\,(j)} - 1\right)\right], \qquad (23)$$

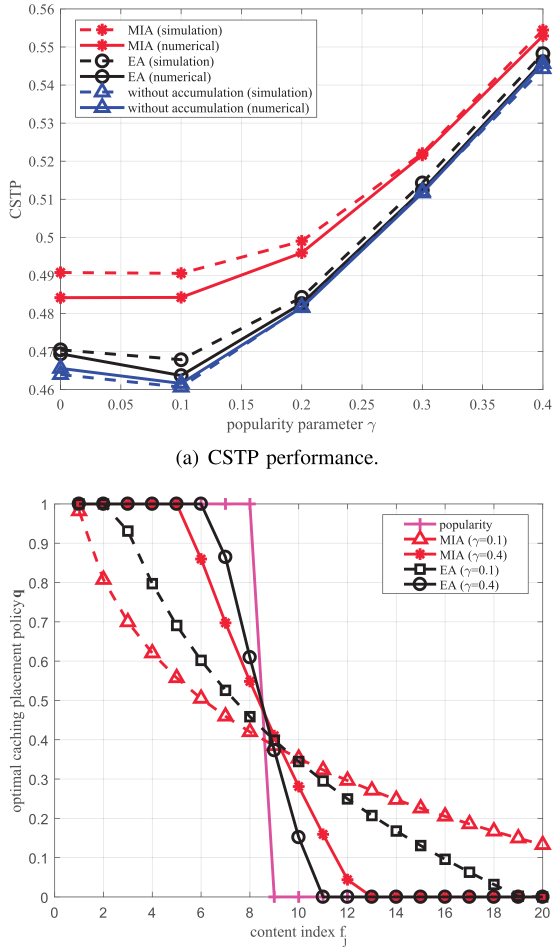
where F_{MIA} is shown in eq. (19).

IV. Performance Evaluations

In this section, we evaluate the numerical results to illustrate the influence of network parameters on the CSTP performance. The numerical results are calculated based on (6) where P_{EA} (resp. P_{MIA}) is based on (11) (resp. (17)). The applied optimal caching placement policy is shown in (22) (resp. (23)). As a baseline scheme, the existing without accumulation scheme [20] is also given, in which the RU simply decodes and forwards the data to the DU in the second time slot. Furthermore, Monte Carlo simulations are performed to verify the accuracy of numerical results. In the simulations, a D2D network following HPPP with intensity $\lambda_u = 10^{-2}$ users per m^2 is considered. The path-loss exponent α is assumed to value 4 in the simulations, and the transmission is treated as successful if the received SIR exceeds the threshold $\theta = 2$ dB. Each user has an active probability $\rho = 0.5$ to initiate a content f_j from the content library containing $J = 20$ contents. Also, the request follows Zipf distribution with skewness parameter $\gamma = 0.1$.

A. Impact of Popularity Parameter γ

Fig. 2(a) plots the CSTP performance as a function of popularity parameter γ for all the candidate schemes. A larger γ means users' requests are more concentrated to the most popular contents. Such concentration feature indicates that an arbitrary content request is more likely to be satisfied through caching. From this figure, it is observed that the CSTP increases with γ under all the candidate schemes. Also, the MIA scheme outperforms the EA scheme, while EA shows the similar performance compared with the without accumulation scheme. This implies that the HARQ scheme, especially MIA, can improve the CSTP performance as opposed to existing schemes. In addition, one can find that the simulation results are a little higher than that of numerical results for the MIA case and EA case. The reason of the observation is that our derivations are based on the worst RU selection process, but this is not applied in the simulations.



(a) CSTP performance.

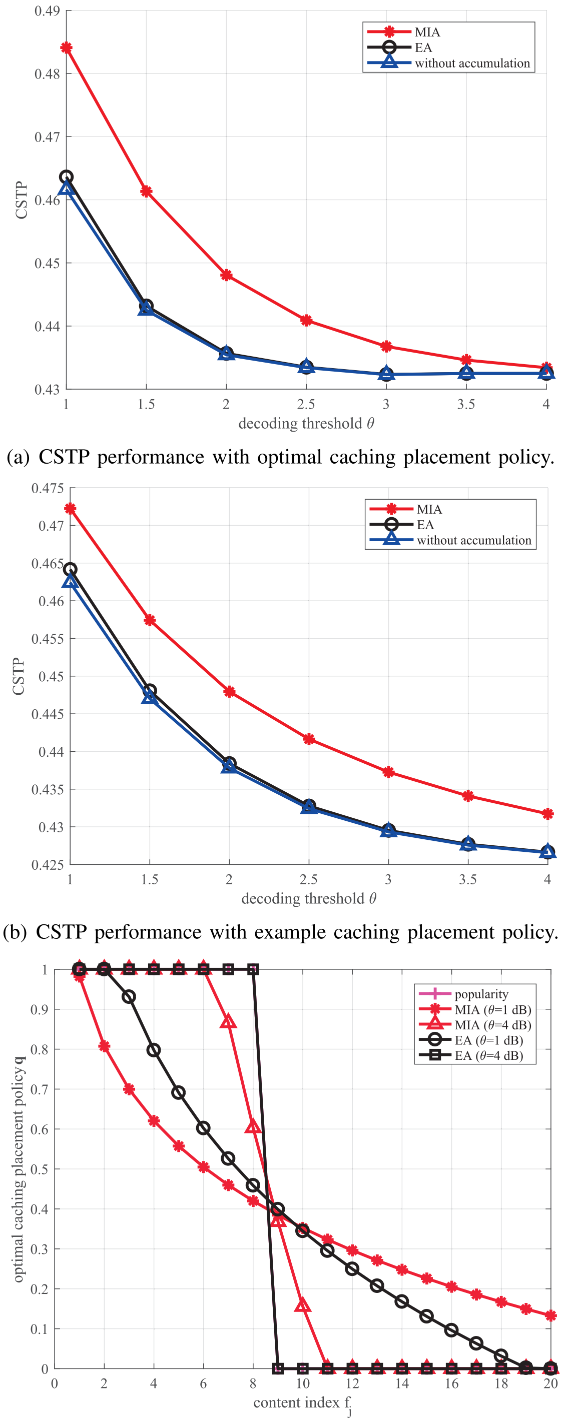
(b) Optimal caching placement policy.

Fig. 2. Performance comparison of CSTP under different popularity parameter γ.

Fig. 2(b) shows the optimal caching placement policy $\mathbf{q}$ as a function of content index f_j for EA and MIA schemes. The caching placement policy is a decreasing function in terms of f_j. It is obvious that contents with higher popularity should be cached with higher probabilities, while those with less popularities can be discarded if the storage capacity is limited. To have a better comparison, in this figure the popularity based caching scheme is shown as a benchmark. In popularity based caching scheme, the most M_S popular contents are cached with probability 1 and the others are cached with probability 0. If a content is not found by self-caching, it cannot be fetched from nearby peers via D2D caching either. Thus, this is the most "selfish" probabilistic caching placement policy without any caching diversity. For every scheme, the optimal caching placement policy when $\gamma = 0.4$ is more "selfish" than that when $\gamma = 0.1$. When users become "selfish", they allocate more caching space for the most popular contents. Accordingly, the portion of users' requests being satisfied by D2D caching are reduced. This may happen when the D2D transmission environment becomes severe, or when users' requests are concentrated to the most popular contents. This observation can also be seen in Fig. 2(a) that the performance gap of different schemes decreases with γ. In addition, as seen in Fig. 2(b), the optimal caching placement policy of EA is more "selfish" than that of the MIA scheme. This is because the EA achieves lower communication benefits compared with MIA, and the D2D caching contributes less to the overall CSTP performance accordingly. In this case, a more "selfish" caching placement policy is adopted to increase the self-caching benefit and to maximize the CSTP. Following the logic, we can conclude that the optimal caching placement policies of the without accumulation scheme and the single hop transmission are more "selfish" than EA scheme; and the illustration is omitted in this figure for brevity.

B. Impact of SIR Threshold θ

The impact of SIR threshold θ on the CSTP is analyzed in Fig. 3(a). From the figure, it can be observed that the performance of CSTP decreases with θ for all schemes. It is obvious that the decoding condition becomes rigorous with higher θ, and the successful transmission probability decreases accordingly. When θ is large enough (e.g., $\theta = 4$ dB), more rigorous channel condition is needed to decode the desired content, and almost no transmission gain remains. In this case, the D2D caching gain is negligible, and the self-caching dominates the CSTP performance. Thus, when θ is large enough, the performance of all the candidate schemes converge to the probability that the users' requests can be satisfied by self-caching. Such observation also can be seen in Fig. 3(c) that the optimal caching placement when $\theta = 4$ dB approaches to the popularity based scheme. However, as shown in Fig. 3(b), if only the transmission is considered and the influence of caching placement is excluded (this is done by applying a fixed caching placement policy to all the transmission schemes), the conclusion that MIA always performs better than EA can be drawn. The observation implies that such conclusions in the transmission area do not hold in the caching area anymore. Therefore, caching placement policy and cooperative transmission are coupled in nature when maximizing the caching performance.



(a) CSTP performance with optimal caching placement policy.

(b) CSTP performance with example caching placement policy.

(c) Optimal caching placement policy.

Fig. 3. Performance comparison of CSTP under different decoding threshold θ.

In addition, from the Fig. 3(a) one may observe that the outperformance of the MIA based scheme is not so significant compared with the baseline schemes. The explanation is as follows. As stated in the analysis part, the CSTP performance is determined by two factors: the caching placement policy and the transmission scheme. The following two conditions need to be met to contribute to the increase of the CSTP: 1. The required content should be cached by some D2D users (this depends on the caching probability vector); 2. The required and cached content should be successfully delivered to the DU. For the first condition, due to the fact that there is limited caching space at D2D users the probability that a particular content can be fetched nearby is low. That is why the CSTP performance of different transmission schemes cannot be differentiated significantly under the constraint of limited caching capacity. However, advanced transmission

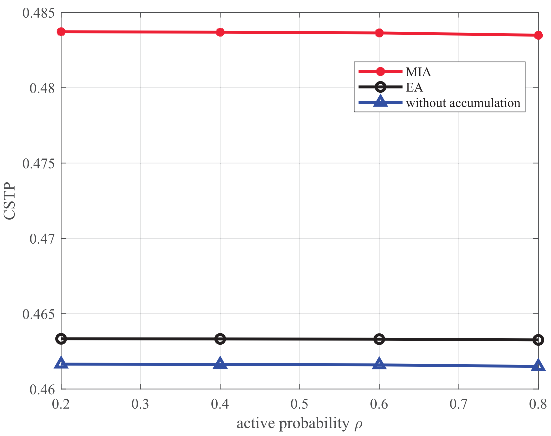


Fig. 4. CSTP performance comparison under different active probability ρ.

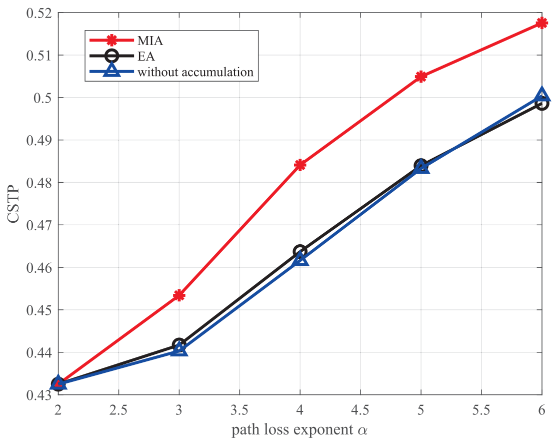


Fig. 5. Performance comparison of CSTP under different path loss exponent α.

schemes certainly improve the CSTP performance, and such improvement could be significant in practical content retrieval process under the caching capacity constraint.

C. Impact of User Active Probability ρ

The CSTP performance with user active probability ρ changing from 0.2 to 0.8 is analyzed in Fig. 4. For all the schemes shown in this figure, the change of ρ has negligible impact on the CSTP performance. Although the influence of ρ is omitted in deriving the optimal caching placement policy as shown in (22) (resp. (23)) for the EA (resp. MIA) case, its influence is considered when calculating P_{EA} (resp. P_{MIA}) as shown in (11) (resp. (17)). The stable performance indicates that the value of ρ does not impact the CSTP performance and that our approximation to remove its influence in deriving the expression of the optimal caching placement policy in accurate. The observation can be intuitively explained as follows. The parameter ρ influences the density of D2D transmitters ($\lambda_t = \lambda_u(1-\rho)$), including the desired transmitter (SU) and interference transmitters. When ρ is small, the transmitter density is dense and the received powers of both the desired signal and interference signal are high. On the contrary, in the large ρ case, the transmitters are very sparse, and the powers of the desired signal and interference signal are low. In both cases, the SIR, i.e., the ratio of the desired signal power and the interference signal power are comparative, thus the CSTP performance does not fluctuate too much with the change of ρ when other conditions are fixed. Similar conclusion can be achieved by studying the user density λ_u, and the illustration is omitted for brevity.

D. Impact of Path Loss Exponent α

Fig. 5 plots the impact of path loss exponent α on the CSTP performance. When $\alpha = 2$, all schemes have the same performance. With the increase of α, the CSTP performance increases for all schemes. The observation can be explained as follows. For a lower α, the transmitted signal suffers more path loss than the interference signals, and the received SIR decreases gradually. When $\alpha = 2$, almost all transmission gain vanishes, thus the CSTP of different schemes converge to the situation that only considers self-caching gain. This is very similar with the case when $\theta > 4$ dB in Fig. 3(a). By accumulating mutual information in different transmissions, the MIA scheme can perform better to counter the deterioration of the transmission environment than other schemes. This is why the MIA scheme always performs better than its rival schemes when $\alpha > 2$. With α increasing as well as the transmission environment getting better, the outperformance of MIA is more significant compared with other schemes. This observation coincides with that in Fig. 3(a), indicating that MIA scheme could get higher gain and has better performance when higher transmission gain could be achieved.

V. Conclusion

Caching is a promising technique to mitigate the burden of ever-growing data transmission in 5G networks, and the applied transmission schemes highly impact the performance gain of caching. In this work, the caching placement policy was studied with consideration of cooperative transmission in a two-hop device-to-device (D2D) network. The hybrid automatic repeat request (HARQ) scheme with soft information combining (i.e., energy accumulation (EA) and mutual information accumulation (MIA)) was applied at the user side. Cache-aided successful transmission probability (CSTP) was adopted as the performance metric in this paper. Based on stochastic geometry tools, the analytical expressions for the CSTP were derived; and the corresponding optimal caching placement policy was given to maximize the CSTP in moderate or high SIR regime. Evaluation results indicate that the MIA based cooperative caching strategy performs better as opposed to existing strategies in most cases, but such outperformance vanishes when the transmission environment becomes severe or when the users' requests become concentrated.

Appendix A
Proof of Theorem 1

$$\begin{aligned} P_{\text{EA\_III}}(r) = \text{P}_\text{r}\left[N_R > 0\right] \text{P}_\text{r}\left[\text{SIR}_{\text{SR},1} > \theta\right] \\ \text{P}_\text{r}\left[\theta - \text{SIR}_{\text{RD},2} \leq \text{SIR}_{\text{SD},1} \leq \theta\right], \quad (24) \end{aligned}$$

where

$$\text{P}_\text{r}\left[N_R > 0\right] = 1 - \text{P}_\text{r}\left[N_R = 0\right] = 1 - e^{-\frac{\pi}{4}\lambda_r p(j) r^2}, \quad (25)$$

$$\mathrm{P_r}\left[\mathrm{SIR_{SR,1}} > \theta\right] = q_j + (1-q_j)\, e^{-\pi K \lambda_t r_1^2 \theta^\delta}, \tag{26}$$

$$\begin{aligned}
&\mathrm{P_r}\left[\theta - \mathrm{SIR_{RD,2}} \le \mathrm{SIR_{SD,1}} \le \theta\right] \\
&= \int_0^\infty \pi\lambda_t K \delta y^{\delta-1} r_2^2 e^{-\pi\lambda_t K y^\delta r_2^2} \\
&\quad \int_{\theta-y}^{\theta} \pi\lambda_t K \delta x^{\delta-1} r^2 e^{-\pi\lambda_t K x^\delta r^2} dx dy \\
&\overset{(a)}{=} \int_0^\theta \pi K \lambda_t \beta^2 r^2 \delta y^{\delta-1} \\
&\quad \times e^{-\pi K \lambda_t r^2 \left[\beta^2 y^\delta + (\theta-y)^\delta\right]} dy \\
&\quad + e^{-\pi K \lambda_t \beta^2 r^2 \theta^\delta} - e^{-\pi K \lambda_t r^2 \theta^\delta}.
\end{aligned} \tag{27}$$

In the derivation, operation (a) follows that $r_2 = \beta r$ ($\beta \in [0,1]$) and $r_1^2 = (1+\beta^2 - 2\beta\cos\phi)r^2$. Substituting (25), (26) and (27) into (24), we can yield the mathematical expression of $P_{\mathrm{EA\_III}}(r)$, and this is a function of β and ϕ while $0 \le \beta \le 1$ and $0 \le \phi \le \frac{\pi}{4}$. And we can yield the mathematical expression of $P_{\mathrm{EA\_III}}(r,\beta,\phi)$ as

$$\begin{aligned}
&P_{\mathrm{EA\_III}}(r,\beta,\phi) \\
&= \left[1 - e^{-\frac{\pi}{4}\lambda_r p(j) r^2}\right] \\
&\quad \times \left[q_j + (1-q_j)\, e^{-\pi K \lambda_t \left(1+\beta^2-2\beta\cos\phi\right) r^2 \theta^\delta}\right] \\
&\quad \times \left[\int_0^\theta \pi K \lambda_t \beta^2 r^2 \delta y^{\delta-1} e^{-\pi K \lambda_t r^2 \left[\beta^2 y^\delta + (\theta-y)^\delta\right]} dy \right. \\
&\quad \left. + e^{-\pi K \lambda_t \beta^2 r^2 \theta^\delta} - e^{-\pi K \lambda_t r^2 \theta^\delta}\right]
\end{aligned} \tag{28}$$

and

$$\begin{aligned}
\frac{\partial P_{\mathrm{EA\_III}}(r,\beta,\phi)}{\partial\phi} &\overset{(b)}{\simeq} -(1-q_j)\, \pi\lambda_t K \theta^\delta 2\beta \sin\phi \\
&\quad \times e^{-\pi K \lambda_t \left(1+\beta^2-2\beta\cos\phi\right) r^2 \theta^\delta} < 0.
\end{aligned} \tag{29}$$

Here (b) omits some terms unrelated to ϕ for ease of illustration and the values of these terms are larger than zero. Thus $P_{\mathrm{EA\_III}}(r,\phi)$ is a decreasing function with respect to ϕ. Similarly,

$$\frac{\partial P_{\mathrm{EA\_III}}(r,\beta,\phi)}{\partial\beta} = y_1(\beta) + y_2(\beta), \tag{30}$$

where

$$\begin{aligned}
y_1(\beta) &= \mathrm{P_r}(N_R > 0)\, \mathrm{P_r}(\theta - \mathrm{SIR_{RD,2}} \le \mathrm{SIR_{SD,1}} \le \theta)\, (1-q_j) \\
&\quad \times e^{-\pi\lambda_t K \theta^\delta r_1^2} \left[-2\pi\lambda_t K \theta^\delta r^2 (\beta - \cos\phi)\right], \\
y_2(\beta) &= \mathrm{P_r}(N_R > 0)\, \mathrm{P_r}(\mathrm{SIR_{SR,1}} \ge \theta) \\
&\quad \times \left[-2\beta\pi K \lambda_t \theta^\delta r^2 e^{-\pi K \lambda_t \theta^\delta \beta^2 r^2} + 2\pi\beta K \lambda_t \delta \right. \\
&\quad \quad \times \int_0^\theta y^{\delta-1} r^2 \left(1 - \pi\lambda_t K \beta^2 r^2 y^\delta\right) \\
&\quad \quad \left. \times e^{-\pi\lambda_t K r^2 \left[\beta^2 y^\delta + (\theta-y)^\delta\right]} dy\right].
\end{aligned}$$

The sign of $y_1(\beta)$ depends on $\beta - \cos\phi$, i.e., $y_1(\beta) < 0$ if $0 \le \beta < \cos\phi$ or $y_1(\beta) \ge 0$ otherwise. And the sign of $y_2(\beta)$ depends on the term $(1 - \pi\lambda_t K \beta^2 r^2 y^\delta)$. Since the sign of $y_2(\beta)$ relates to r and y, both of which are variables, we need to eliminate their influence by the following integrals. Here it should be noted that, we only care about the sign of $y_2(\beta)$ rather than its accurate expression.

$$\begin{aligned}
y_2(\beta) &= \int_0^\theta \mathrm{P_r}(N_R > 0)\, \mathrm{P_r}(\mathrm{SIR_{SR,1}} \ge \theta) 2\pi\beta K \lambda_t \delta y^{\delta-1} r^2 \\
&\quad \times \left(1 - \pi\lambda_t K \beta^2 r^2 y^\delta\right) e^{-\pi\lambda_t K r^2 \left[\beta^2 y^\delta + (\theta-y)^\delta\right]} dy + Y \\
&\overset{(c)}{=} y_2^A(\beta_0, r) \int_0^\theta \left(1 - \pi\lambda_t K \beta^2 r^2 y^\delta\right) dy + Y \\
&= y_2^A(\beta_0, r) \left(\theta - \frac{\pi\lambda_t K \beta^2 r^2 \theta^{1+\delta}}{1+\delta}\right) + Y,
\end{aligned}$$

where $Y = \mathrm{P_r}(N_R > 0)\mathrm{P_r}(\mathrm{SIR_{SR,1}} \ge 0)(-2\pi\beta K \lambda_t \theta^\delta r^2 e^{-\pi K \lambda_t \theta^\delta \beta^2 r^2}) < 0$ holds for arbitrary $\beta \in [0,1]$. And in (c), the mean value theorem of integral is applied that

$$\begin{aligned}
y_2^A(\beta_0, r) &= \mathrm{P_r}(N_R > 0)\, \mathrm{P_r}(\mathrm{SIR_{SR,1}} \ge \theta) \frac{\pi^2}{4} \lambda_t 2\beta r^2 y^{-\frac{1}{2}} \\
&\quad \times e^{-\frac{\pi^2}{2}\lambda_t r^2 \left[\beta^2 y^{\frac{1}{2}} + (\theta-y)^{\frac{1}{2}}\right]} \ge 0
\end{aligned}$$

also holds for arbitrary $\beta_0 \in [0,1]$. Further, to eliminate the variable r,

$$\begin{aligned}
&\mathrm{E}_r\left[y_2(\beta)\right] \\
&= \int_0^\infty f_r(r) \left[y_2^A(\beta_0, r)\left(\theta - \frac{\pi\lambda_t K \beta^2 r^2 \theta^{1+\delta}}{1+\delta}\right) + Y\right] dr \\
&\overset{(d)}{=} y_2^A(\beta_0, r_0) \int_0^\infty f_r(r) \left(\theta - \frac{\pi\lambda_t K \beta^2 r^2 \theta^{1+\delta}}{1+\delta}\right) dr + F(Y) \\
&= y_2^A(\beta_0, r_0)\, \theta \left(1 - \frac{K\beta^2\theta^\delta}{(\delta+1)\, q_j^2}\right) + F(Y).
\end{aligned}$$

In (d) the mean value of theorem of integral is applied and $y_2^A(\beta_0, r_0) \ge 0$ holds for arbitrary $r_0 \in (0,\infty)$. $F(Y) = \int_0^\infty f_r(r) Y dr > 0$. As a result, the sign of $y_2(\beta)$ depends on $1 - \frac{K\beta^2\theta^\delta}{(\delta+1)q_j^2}$. If $1 - \frac{K\beta^2\theta^\delta}{(\delta+1)q_j^2} < 0$, i.e., $\beta > \sqrt{\frac{(1+\delta)q_j}{K\theta^\delta}}$, then $y_2(\beta) < 0$. To ensure $\sqrt{\frac{(1+\delta)q_j}{K\theta^\delta}} < 1$, here we assume $\theta^\delta > \frac{\delta+1}{K}$ holds in the following derivation. And also, $y_1(\beta) < 0$ when $\beta > \cos\phi$. This $P_{\mathrm{EA\_III}}(r)$ reaches its minimal value when $\beta = 1$ and $\phi = \frac{\pi}{4}$. Then $P_{\mathrm{EA\_III}}(r)$ in this case is derived as shown in (10).

Appendix B
Derivation of P_{EA}

$$P_{\mathrm{EA}} = \mathrm{E}_r\left[P_{\mathrm{EA\_I}}(r)\right] + \mathrm{E}_r\left[P_{\mathrm{EA\_II}}(r)\right] + \mathrm{E}_r\left[P_{\mathrm{EA\_III}}(r)\right]. \tag{31}$$

$$\begin{aligned}
&\mathrm{E}_r\left[P_{\mathrm{EA\_I}}(r)\right] \\
&= \int_0^\infty 2\pi\lambda_t q_j r e^{-\pi\lambda_t q_j r^2} e^{-\pi\lambda_t K \theta^\delta r^2} dr \\
&= \frac{q_j}{q_j + K\theta^\delta},
\end{aligned} \tag{32}$$

$$\begin{aligned}
&\mathrm{E}_r\left[P_{\mathrm{EA\_II}}(r)\right] \\
&= \int_0^\infty 2\pi\lambda_t q_j r e^{-\pi\lambda_t q_j r^2} e^{-\frac{\pi}{4}\lambda_u \rho p(j) r^2} \\
&\quad \times \int_0^\theta \pi\lambda_t K \delta y^{\delta-1} r^2 e^{-\pi K \lambda_t \left[y^\delta + (\theta-y)^\delta\right] r^2} dy dr \\
&= \int_0^\theta \frac{\delta K q_j y^{\delta-1}}{\left\{q_j + \frac{1}{4}\frac{\rho}{1-\rho} p(j) + K\left[y^\delta + (\theta-y)^\delta\right]\right\}^2} dy \\
&= q_j F_{\mathrm{EA1}}\left(q_j, \delta, p(j), \rho, \theta\right),
\end{aligned} \tag{33}$$

where

$$
\begin{aligned}
&F_{\mathrm{EA1}}\left(q_j,\delta,p\left(j\right),\rho,\theta\right)\\
&=\int_0^\theta \frac{\delta K y^{\delta-1}}{\left\{q_j+\frac{1}{4}\frac{\rho}{1-\rho}p\left(j\right)+K\left[y^\delta+\left(\theta-y\right)^\delta\right]\right\}^2}dy,\\
&\mathrm{E}_r\left[P_{\mathrm{EA\_III}}(r)\right]\\
&=\int_0^\infty\int_0^\theta 2\pi^2\lambda_t^2 q_j r e^{-\pi\lambda_t q_j r^2}K\lambda_t r^2\delta y^{\delta-1}e^{-\pi K\lambda_t\left[y^\delta+(\theta-y)^\delta\right]r^2}\\
&\quad\times\left[1-e^{-\frac{\pi}{4}\lambda_u\rho p(j)r^2}\right]\left[q_j+\left(1-q_j\right)e^{-\pi K\lambda_t\theta^\delta\left(2-\sqrt{2}\right)r^2}\right]drdy\\
&=\int_0^\theta\int_0^\infty 2\pi^2\lambda_t^2 q_j^2 K r^3\delta y^{\delta-1}e^{-\pi\lambda_t\left\{q_j+K\left[y^\delta+(\theta-y)^\delta\right]\right\}r^2}drdy\\
&\quad+\int_0^\theta\int_0^\infty 2\pi^2\lambda_t^2 q_j\left(1-q_j\right)Kr^3\delta y^{\delta-1}\\
&\quad\times e^{-\pi\lambda_t\left\{q_j+K\left[y^\delta+(\theta-y)^\delta\right]+K\theta^\delta\left(2-\sqrt{2}\right)\right\}r^2}drdy\\
&\quad-\int_0^\theta\int_0^\infty 2\pi^2\lambda_t^2 q_j^2 K r^3\delta y^{\delta-1}\\
&\quad\times e^{-\pi\lambda_t\left\{q_j+K\left[y^\delta+(\theta-y)^\delta\right]+\frac{1}{4}\frac{\rho}{1-\rho}p(j)\right\}r^2}drdy\\
&\quad-\int_0^\theta\int_0^\infty e^{-\pi\lambda_t\left\{q_j+K\left[y^\delta+(\theta-y)^\delta\right]+\frac{1}{4}\frac{\rho}{1-\rho}p(j)+K\theta^\delta\left(2-\sqrt{2}\right)\right\}r^2}\\
&\quad\times 2\pi^2\lambda_t^2 q_j\left(1-q_j\right)Kr^3\delta y^{\delta-1}drdy\\
&=\int_0^\theta\frac{q_j^2 K\delta y^{\delta-1}}{\left\{q_j+K\left[y^\delta+\left(\theta-y\right)^\delta\right]\right\}^2}dy\\
&\quad-\int_0^\theta\frac{q_j^2 K\delta y^{\delta-1}}{\left\{q_j+K\left[y^\delta+\left(\theta-y\right)^\delta\right]+\frac{1}{4}\frac{\rho}{1-\rho}p\left(j\right)\right\}^2}dy\\
&\quad+\int_0^\theta\frac{q_j\left(1-q_j\right)K\delta y^{\delta-1}}{\left\{q_j+K\left[y^\delta+\left(\theta-y\right)^\delta\right]+K\theta^\delta\left(2-\sqrt{2}\right)\right\}^2}dy\\
&\quad-\int_0^\theta\frac{q_j\left(1-q_j\right)K\delta y^{\delta-1}dy}{\left\{q_j+K\left[y^\delta+\left(\theta-y\right)^\delta\right]+K\theta^\delta\left(2-\sqrt{2}\right)+\frac{1}{4}\frac{\rho}{1-\rho}p\left(j\right)\right\}^2}\\
&=q_j^2F_{\mathrm{EA1}}\left(q_j,\delta,p\left(j\right),\rho,\theta\right)+\left(q_j-q_j^2\right)F_{\mathrm{EA2}}\left(q_j,\delta,p\left(j\right),\rho,\theta\right),
\end{aligned}
\tag{34}
$$

where

$$
\begin{aligned}
&F_{\mathrm{EA1}}\left(q_j,\delta,p\left(j\right),\rho,\theta\right))\\
&=\int_0^\theta\frac{K\delta y^{\delta-1}}{\left\{q_j+K\left[y^\delta+\left(\theta-y\right)^\delta\right]\right\}^2}dy\\
&\quad-\int_0^\theta\frac{K\delta y^{\delta-1}}{\left\{q_j+K\left[y^\delta+\left(\theta-y\right)^\delta\right]+\frac{1}{4}\frac{\rho}{1-\rho}p\left(j\right)\right\}^2}dy,\\
&F_{\mathrm{EA2}}\left(q_j,\delta,p\left(j\right),\rho,\theta\right))\\
&=\int_0^\theta\frac{q_j\left(1-q_j\right)K\delta y^{\delta-1}}{\left\{q_j+K\left[y^\delta+\left(\theta-y\right)^\delta\right]+K\theta^\delta\left(2-\sqrt{2}\right)\right\}^2}dy\\
&\quad-\int_0^\theta\frac{q_j\left(1-q_j\right)K\delta y^{\delta-1}dy}{\left\{q_j+K\left[y^\delta+\left(\theta-y\right)^\delta\right]+K\theta^\delta\left(2-\sqrt{2}\right)+\frac{1}{4}\frac{\rho}{1-\rho}p\left(j\right)\right\}^2},
\end{aligned}
$$

Substituting (32), (33) and (34) into (31) we can obtain the expression of P_{EA} shown in (11).

References

[1] *Cisco Visual Networking Index: Global Mobile Data Traffic Forecast Update 2017–2022 White Paper*. Accessed: 2019. [Online]. Available: https://www.cisco.com/c/en/us/solutions/collateral/service-provider/visual-networking-index-vni/white-paper-c11-738429.html

[2] H. Chen, L. Liu, N. Mastronarde, L. Ma, and Y. Yi, "Cooperative retransmission for massive MTC under spatiotemporally correlated interference," in *Proc. IEEE GLOBECOM*, Dec. 2016, pp. 1–6.

[3] I. O. Nunes, P. O. S. vaz de Melo, and A. A. F. Loureiro, "Leveraging D2D multihop communication through social group meeting awareness," *IEEE Wireless Commun.*, vol. 23, no. 4, pp. 12–19, Aug. 2016.

[4] A. M. Akhtar, A. Behnad, and X. Wang, "Cooperative ARQ-based energy-efficient routing in multihop wireless networks," *IEEE Trans. Veh. Technol.*, vol. 64, no. 11, pp. 5187–5197, Nov. 2015.

[5] Y. Zhou and W. Zhuang, "Throughput analysis of cooperative communication in wireless ad hoc networks with frequency reuse," *IEEE Trans. Wireless Commun.*, vol. 14, no. 1, pp. 205–218, Jan. 2015.

[6] A. F. Molisch, N. B. Mehta, J. S. Yedidia, and J. Zhang, "Performance of fountain codes in collaborative relay networks," *IEEE Trans. Wireless Commun.*, vol. 6, no. 11, pp. 4108–4119, Nov. 2007.

[7] S. C. Draper, L. Liu, A. F. Molisch, and J. S. Yedidia, "Cooperative transmission for wireless networks using mutual-information accumulation," *IEEE Trans. Inf. Theory*, vol. 57, no. 8, pp. 5151–5162, Aug. 2011.

[8] H. Chen, L. Liu, T. Novlan, J. D. Matyja, B. L. Ng, and J. Zhang, "Spatial spectrum sensing-based device-to-device cellular networks," *IEEE Trans. Commun.*, vol. 15, no. 11, pp. 7299–7313, Nov. 2016.

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

[10] X. Lin, J. Andrews, A. Ghosh, and R. Ratasuk, "An overview of 3GPP device-to-device proximity services," *IEEE Commun. Mag.*, vol. 52, no. 4, pp. 40–48, Apr. 2014.

[11] N. Golrezaei, K. Shanmugam, A. G. Dimakis, A. F. Molisch, and G. Caire, "FemtoCaching: Wireless video content delivery through distributed caching helpers," in *Proc. INFOCOM*, Mar. 2012, pp. 1107–1115.

[12] N. Golrezaei, A. F. Molisch, A. G. Dimakis, and G. Caire, "Femtocaching and device-to-device collaboration: A new architecture for wireless video distribution," *IEEE Commun. Mag.*, vol. 51, no. 4, pp. 142–149, Apr. 2013.

[13] D. Liu and C. Yang, "Caching policy toward maximal success probability and area spectral efficiency of cache-enabled HetNets," *IEEE Trans. Commun.*, vol. 65, no. 6, pp. 2699–2714, Jun. 2017.

[14] J. Ma, J. Wang, and P. Fan, "A cooperation-based caching scheme for heterogeneous networks," *IEEE Access*, vol. 5, pp. 15013–15020, 2016.

[15] B. Chen, C. Yang, and A. F. Molisch, "Cache-enabled device-to-device communications: Offloading gain and energy cost," *IEEE Trans. Wireless Commun.*, vol. 16, no. 7, pp. 4519–4536, Jul. 2016.

[16] K. Poularakis, G. Iosifidis, A. Argyriou, I. Koutsopoulos, and L. Tassiulas, "Caching and operator cooperation policies for layered video content delivery," in *Proc. INFOCOM*, 2016, Apr. pp. 1–9.

[17] L. Zhang, M. Xiao, G. Wu, and S. Li, "Efficient scheduling and power allocation for D2D-assisted wireless caching networks," *IEEE J. Sel. Areas Commun.*, vol. 64, no. 6, pp. 2438–2452, Jun. 2016.

[18] D. Malak, M. Al-Shalash, and J. G. Andrews, "Optimizing content caching to maximize the density of successful receptions in device-to-device networking," *IEEE Trans. Commun.*, vol. 64, no. 10, pp. 4365–4380, Oct. 2016.

[19] Z. Chen, N. Pappas, and M. Kountouris, "Probabilistic caching in wireless D2D networks: Cache hit optimal versus throughput optimal," *IEEE Commun. Lett.*, vol. 21, no. 3, pp. 584–587, Mar. 2017.

[20] S. Krishnan, M. Afshang, and H. S. Dhillon, "Effect of retransmissions on optimal caching in cache-enabled small cell networks," *IEEE Trans. Veh. Technol.*, vol. 66, no. 12, pp. 11383–11387, Dec. 2017.

[21] S.-W. Jeon, S.-N. Hong, M. Ji, and G. Caire, "Caching in wireless multihop device-to-device networks," in *Proc. ICC*, Jun. 2015, pp. 6732–6737.

[22] H. Chen, L. Liu, J. D. Matyjas, and M. J. Medley, "Cooperative routing for underlay cognitive radio networks using mutual-information accumulation," *IEEE Trans. Wireless Commun.*, vol. 14, no. 12, pp. 7110–7122, Dec. 2015.

[23] R. Atat, J. Ma, H. Chen, U. Lee, J. Ashdown, and L. Liu, "Cognitive relay networks with energy and mutual-information accumulation," in *Proc. IEEE Conf. Comput. Commun.*, Apr. 2018, pp. 640–644.

[24] A. Rajanna and M. Haenggi, "Downlink coordinated joint transmission for mutual information accumulation," *IEEE Wireless Commun. Lett.*, vol. 6, no. 2, pp. 198–201, Apr. 2017.

[25] A. Zanella, A. Bazzi, and B. M. Masini, "Relay selection analysis for an opportunistic two-hop multi-user system in a poisson field of nodes," *IEEE Trans. Wireless Commun.*, vol. 16, no. 2, pp. 1281–1293, Feb. 2017.

[26] Y. J. Chun, G. B. Colombo, S. L. Cotton, W. G. Scanlon, R. M. Whitaker, and S. M. Allen, "Device-to-device communications: A performance analysis in the context of social comparison-based relaying," *IEEE Trans. Wireless Commun.*, vol. 16, no. 12, pp. 7733–7745, Dec. 2017.

[27] M. Haenggi and R. K. Ganti, "Interference in large wireless networks," *Found. Trends Netw.*, vol. 3, no. 2, pp. 127–248, Nov. 2009.

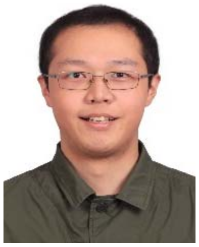
[28] D. Hu, J. Wu, and P. Fan, "Maximizing end-to-end throughput of interference-limited multihop networks," *IEEE Trans. Veh. Technol.*, vol. 67, no. 6, pp. 5465–5469, Jun. 2018.

[29] C. Xu, M. Wang, X. Chen, L. Zhong, and A. L. Grieco, "Optimal information centric caching in 5G device-to-device communications," *IEEE Trans. Mobile Comput.*, vol. 17, no. 9, pp. 2114–2126, Sep. 2018.

Junchao Ma (S'16) received the M.E. degree from Southwest Jiaotong University, Chengdu, China, in 2011, where he is currently pursuing the Ph.D. degree with the Key Laboratory of Information Coding and Transmission, School of Information Science and Technology. He was a Visiting Student at the Bradley Department of Electrical and Computer Engineering, Virginia Polytechnic Institute and State University, Blacksburg, VA, USA, from 2017 to 2019.

His current research interests include scalable video caching, random linear network coding (RLNC), and age of information (AoI).

Lingjia Liu (SM'15) received the B.S. degree in electronic engineering from Shanghai Jiao Tong University and the Ph.D. degree in electrical and computer engineering from Texas A&M University. He is currently an Associate Professor with the Bradley Department of Electrical and Computer Engineering, Virginia Tech. He is also the Associate Director of Affiliate Relations at Wireless@Virginia Tech. Prior to that, he was an Associate Professor at the EECS Department, University of Kansas (KU). From 2008 and 2011, he was with the Standards and Mobility Innovation Laboratory, Samsung Research America, where he received the Global Samsung Best Paper Award, in 2008 and 2010, respectively. He was leading Samsung's efforts on multiuser MIMO, CoMP, and HetNets in 3GPP LTE/LTE-advanced standards. His general research interests mainly lie in emerging technologies for 5G cellular networks, including machine learning for wireless networks, massive MIMO, massive machine-type communications, and mm-wave communications. He received the Air Force Summer Faculty Fellowship from 2013 to 2017, the Miller Scholarship at KU in 2014, the Miller Professional Development Award for Distinguished Research at KU in 2015, the 2016 IEEE GLOBECOM Best Paper Award, the 2018 IEEE ISQED Best Paper Award, the 2018 IEEE TCGCC Best Conference Paper Award, and the 2018 IEEE TAOS Best Paper Award.



Bodong Shang (S'16) received the B.S. degree in information science and technology from Northwest University, Xi'an, China, in 2015, and the M.S. degree in communication and information systems from Xidian University, Xi'an, China, in 2018. He is currently pursuing the Ph.D. degree with the Bradley Department of Electrical and Computer Engineering, Virginia Polytechnic Institute and State University, Blacksburg, VA, USA.

His recent research interests include several aspects of wireless communication, such as MIMO systems, device-to-device communication, unmanned aerial vehicle, and vehicle-to-everything.

Pingzhi Fan (M'93–SM'99–F'15) received the M.Sc. degree in computer science from Southwest Jiaotong University, China, in 1987, and the Ph.D. degree in electronic engineering from Hull University, U.K., in 1994.

He is currently a Distinguished Professor and the Director of the Institute of Mobile Communications, Southwest Jiaotong University, China. Since 1997, he has been a Visiting Professor with Leeds University U.K., and a Guest Professor of Shanghai Jiaotong University, since 1999. He was a recipient of the U.K. ORS Award in 1992, the NSFC Outstanding Young Scientist Award in 1998, and the IEEE VTS Jack Neubauer Memorial Award in 2018. He has over 290 research papers published in various international journals, and eight books (incl. edited), and is the Inventor of 23 granted patents. His research interests include vehicular communications, wireless networks for big data, and signal design and coding. He is a fellow of the IET, CIE, and CIC. He served as a Board Member of IEEE Region 10, the IET (IEE) Council, and the IET Asia–Pacific Region. He served as the General Chair or the TPC Chair for a number of international conferences, including the VTC2016Spring, IWSDA2017, and ITW2018. He is the Founding Chair of IEEE VTS BJ Chapter and IEEE ComSoc CD Chapter, and the IEEE Chengdu Section. He is an IEEE VTS Distinguished Lecturer from 2015 to 2019.