Analysis of A Distance-Based Pairing Scheme for Collaborative Content Distribution via Device-to-Device Communications

Wei Song, Senior Member, IEEE

Abstract—With the increasing penetration of smart devices, device-to-device (D2D) communications offer a promising paradigm to accommodate the ever-growing mobile video traffic and unremitting demands for fast content distribution. The redundant storage and communication capacities of smart devices can be exploited for collaborative content caching and distribution. Such united resources at the network edge can serve end users with low delivery cost and high performance. In this paper, we analyze a heuristic distance-based scheme, which appropriately pairs a device requesting a content item with a cache device within a collaboration distance and providing the requested content. An analytical framework is developed to derive the optimal collaboration distance so as to maximize the likelihood that the paired cache devices successfully fulfill the content requests. The simulation results validate the analytical framework and the effectiveness of the heuristic distance-based scheme for D2D-assisted content distribution. We also investigate the effects of various system parameters on the performance in terms of expected satisfaction probability.

Index Terms—D2D communications, collaborative content distribution, D2D pairing, satisfaction probability, coverage probability.

I. INTRODUCTION

With the ubiquitous wireless infrastructure and pervasive smart devices, video traffic becomes dominant for global mobile networks. Many mobile videos feature short durations and may quickly spread with the aid of social media. Traditionally, content providers often deliver their content via a distributed content delivery network (CDN). The pervasive smart devices introduce abundant resources at the network edge. Their redundant storage and communication capacities can be exploited for collaborative content caching and distribution, which can complement traditional CDNs with lower delivery cost and higher performance.

Particularly, content distribution can be assisted with device-to-device (D2D) communications, which offer various benefits such as expanding coverage, improving energy efficiency, and offloading traffic [1,2]. Depending on how the spectrum is shared between D2D and cellular users, D2D communications can adopt overlay spectrum sharing and underlay spectrum sharing. In overlay spectrum sharing, the licensed spectrum is divided into two parts and respectively assigned to D2D and cellular users. Intra-cell interference between D2D and cellular users can thus be completely avoided through orthogonal channel assignment. Underlay spectrum sharing allows D2D and cellular users to access the same spectrum simultaneously, which may result in non-negligible interference but can achieve high spectral efficiency.

In the literature, there have been many studies on the key issues of D2D communications, such as interference management, power control, and resource allocation [3]–[7]. The co-channel interference is widely considered in these studies on D2D underlaying networks in a general service context. Meanwhile, D2D communications can facilitate a variety of specialized applications, including proximity-based social networking, public safety services, intelligent transport systems, and content distribution. In particular, this work focuses on a D2D-assisted content distribution scenario, where a helper device can send a file to a request device within a collaboration distance. In a general D2D context, the collaboration distance mainly depends on the interference among the D2D links and the target quality-of-service (QoS) requirement, which is usually defined in terms of the received signal-to-interference-plus-noise ratio (SINR). Nonetheless, in the cached-enabled D2D networks, the collaboration distance is also affected by content caching in nearby devices. Because the D2D interference effect is complicated for the special content distribution scenario, many studies in this area (e.g., [8,9]) neglect the intra-cell interference or consider a simplified interference model. For example, the work in [9] captures the interference effect by dividing a cell into equal-area square clusters and allowing only one D2D link per cluster.

In this paper, we consider a more realistic interference model and analyze a distance-based pairing scheme for D2D-assisted content distribution. The D2D pairing problem needs to appropriately pair a device requesting a content item with a cache device within a collaboration distance and providing the requested content. On one hand, a larger collaboration distance increases the possibility that an available cache device that contains the requested content exists within the collaboration region. On the other hand, there are potentially more D2D pairs within a larger collaboration region, which can cause higher interference to each other. Moreover, the D2D pairing scheme has a significant impact on the resulting geometric distributions of the successfully paired devices.

To determine the optimal collaboration distance, we de-
velop an analytical framework to evaluate the expectation of the probability that a content request is served satisfactorily (termed satisfaction probability in this paper). There have been some studies on optimizing the communication distance for D2D communications in a general context, and the design objectives are often defined in terms of standard performance metrics, such as sum rate, SINR, and coverage probability. For instance, the coverage probability is analyzed in [10] for a similar D2D underlying network, but it focuses on the power control algorithms and does not involve any content distribution constraint. In contrast, our analytical framework further incorporates the effects of content caching in addition to the interference among D2D links. Therefore, we focus on a more specialized performance metric, i.e., the expectation of the satisfaction probability, which also takes content availability into account. Furthermore, our analysis needs to properly address the stochastic attributes of the device distributions that result from the D2D pairing scheme. As the expected satisfaction probability depends on the coverage probability, we extend the analysis in [10] to the content distribution scenario and provide a closed-form approximate expression for the coverage probability. The approximate analysis can effectively expedite the search for the best collaboration distance that maximizes the expected satisfaction probability.

Specifically, our main contributions lie in the following aspects.

- Considering a distance-based algorithm that pairs request devices with cache devices, we investigate the effects of the collaboration distance and the D2D device pairing on the achievable performance for content distribution.
- Assuming that D2D pairing follows the distance-based algorithm, we further develop an analytical framework to determine the optimal collaboration distance, within which a best available cache device is identified for a request device. Here, our analysis properly characterizes the content availability in cache devices and the interference among D2D pairs, such that a maximum number of request devices are satisfied while the D2D interference is still tolerable.
- We conduct simulation experiments to validate the analysis accuracy, examine the impact of various system parameters, and show the performance gap of the distance-based scheme to the optimal solution.

The remainder of this paper is structured as follows. In Section II, we review related works on D2D-assisted content distribution. Section III describes the collaborative content distribution scenario and the research problem under study. In Section IV, we present an analytical framework to derive the optimal collaboration distance. In Section V, we present numerical results to validate the analysis accuracy and evaluate the performance of the proposed scheme with various settings. Finally, Section VI concludes this paper.

II. RELATED WORK

D2D communications can offer various benefits and facilitate many interesting applications, among which content distribution is one typical example. In the literature, there have been some studies on how to exploit D2D communications to assist content distribution. In [11], a coalition formation game is formulated to exchange content pieces among the on-board units (OBUs) of vehicles in VANETs. Cognitive radio is further considered in [12] for direct communications between OBUs. Consequently, the utility function becomes dependent on the coalitions’ graph structure, as the vehicle-to-vehicle (V2V) links between OBUs interfere with the vehicle-to-roadside (V2R) links between the roadside unit (RSU) and OBUs. Thus, a coalitional graph game is used to analyze the content distribution in VANETs. Similar to D2D communications, a promising technique, non-orthogonal multiple access (NOMA), is considered in [13] to support vehicle-to-everything (V2X) services. The key aspects such as spectrum management, power control, and scheduling are investigated to achieve low latency and high reliability.

In [14], Wang et al. considered D2D offloading and edge caching at base stations (BSs). Content requests that are not satisfied by nearby devices with feasible D2D links are forwarded to the BSs, which may fulfill the requests by the local caching. Taking into account the impact of D2D content delivery, this work further develops a Markov decision process and a Q-learning based strategy for the BS caching placement. In [15], Lan et al. studied data offloading via D2D communications with proactive caching. An NP-hard data caching problem is formulated to address different mobility and cache capacities of mobile nodes, as well as different popularity and sizes of content data. The problem is solved by a $(1 + \alpha)$-approximation algorithm, based on which two data offloading algorithms are designed. In [16], Jiang et al. formulated D2D local caching as a Knapsack problem to address the limited storage capacities of devices, and developed a distributed caching algorithm based on message popularity estimates. In addition, they explored sender-receiver pairing as a maximum weighted bipartite matching problem, and designed a matching algorithm as an auction.

Social-awareness is an important aspect that is widely considered for content distribution through D2D communications. In [17], a hypergraph framework is used to represent multidimensional information from three layers: the physical layer, the social ties layer, and the common interests layer. Then, the hypergraph is used to study the D2D caching strategies. Though the hypergraph can capture various physical communication conditions such as path loss, shadowing, and fading, the simulation results only consider path loss.

In [18], a randomized auction is proposed to offload message requests from the BS to nearby devices, so that multiple message requests are fulfilled via D2D broadcast with low costs. However, the work in [18] neglects D2D interference but only respects the classic half-duplex constraint of wireless communications. In [19], we considered a similar content distribution scenario with D2D multicast but further restricted the total cost that the BS can afford. The proposed reverse auction selects one message multicast by each cache device so that the BS maximizes the gain in cost saving. Moreover, in [20], we focused on D2D unicast to accommodate variable resource costs of individual devices, and attempted to efficiently direct requests to cache devices. A one-to-many matching is derived
between cache devices and requests to minimize the total resource cost or maximize the cost saving, while satisfying the individual cost budget of each cache device. The proposed randomized mechanism maintains an approximation guarantee for the worst case while subject to a polynomial computation time. For simplicity, in [19,20], the co-channel interference is approximated by a Gaussian process similar to noise.

In [21], we modified the request direction problem in [20] to a one-to-one device pairing problem, and further characterized the co-channel interference among D2D links in a more realistic manner based on the Rayleigh fading model. The device pairing problem aims to obtain a maximum number of device pairs that satisfy the SINR requirement. The problem is solved by a three-step approximation algorithm, in which the first step obtains an upper bound solution via Lagrangian relaxation, the second step derives a feasible solution and further augments it, and the last step refines the pairing solution to guarantee its stability. The proposed approach is proved to converge to a two-sided exchange stable matching.

A similar pairing problem is also studied in [22] but for spectrum sharing between D2D users and cellular users. The authors focused on how to match D2D links with cellular channels so as to maximize the overall performance in terms of the sum rate of all users. The proposed pairing approach includes two steps, in which the first step narrows the candidate D2D links for each cellular user with QoS assurance for the minimum SINR, and the second step optimizes the pairing from the pre-established candidate set and the power allocation for each respective pair. The proposed algorithm performs closely to the well-known Hungarian algorithm but with lower computation cost.

The work in [8] focuses on a D2D content distribution scenario, where a helper device can send a file to a request device within a collaboration distance. The authors studied a probabilistic caching policy in which each device caches a file according to a probability distribution. The optimal caching distribution is obtained to maximize the offloading ratio. It is shown that the maximal offloading ratio increases with the collaboration distance. Nonetheless, this conclusion is conditioned on the assumption that the interference among D2D links is treated as noise. If this assumption is relaxed, the collaboration distance actually results in two conflicting effects. On one hand, a larger collaboration distance increases the request hit ratio over helper devices. On the other hand, more D2D pairs are introduced within a larger collaboration distance, which obviously leads to higher interference among the D2D links and thus decreases the decoding success probability or the transmission rate.

In [9], Golrezaei et al. considered such conflicting effects and analyzed the optimal collaboration distance. Unfortunately, their work is based on a simplified physical-layer communication model, where a cell is divided into equal-area square clusters and only one D2D link is allowed per cluster. This simplification is reasonable since they mainly examined a deterministic caching policy and a random caching policy. However, a more realistic interference-aware D2D communication model is essential for D2D pairing since the interference directly affects the receiving QoS and therefore the pairing feasibility. The work in [23] considers a more realistic interference-aware D2D channel model and studies the pairing of request devices and cache devices for collaborative content distribution. The D2D pairing problem is solved by a heuristic channel-aware algorithm.

### III. System Model and Problem Formulation

#### A. Content Distribution Scenario

In this paper, we consider a content distribution scenario depicted in Fig. 1. A set of request devices, $D$, are requesting content items (referred to as “files” henceforth for simplicity) from a “library” of size $M$, denoted by $M$. The request devices are randomly distributed in a circular disk $C$ of radius $R$, following a homogeneous Poisson point process (PPP), $\Phi_r$, with an intensity function $\lambda_r$. The circular disk $C$ denotes the coverage region of a BS centered at the origin. Each device $k \in D$ requests one file from the library $M$ independently according to a Zipf popularity distribution, which has been shown to be a good model that captures the popularity of video clips [24]. Specifically, each request device demands file $i$ with a probability

$$q_i = \frac{\gamma_r}{\sum_{\ell=1}^m \frac{1}{\ell^{\gamma_r}}}, \quad 1 \leq i \leq m, \gamma_r > 0$$

where the exponent $\gamma_r$ characterizes the relative popularity of files. A larger value of $\gamma_r$ implies that more requests are concentrated on fewer files. Previous studies on real video traces find that a typical exponent $\gamma_r$ ranges from 0.9 to 0.97 [25]. Let $m^*_k$ denote the file requested by device $k$.

In addition, there are another set of cache devices, $S$, which are distributed within circular disk $C$ as another homogeneous PPP $\Phi_c$ with an intensity function $\lambda_c$. Considering the resource limitation of mobile devices, we assume that each cache device $j \in S$ stores one video file in the library $M$. Thus, the cache devices can flexibly and frequently replace the cached content. Specifically, following the random caching policy studied in [9], each cache device stores file $i$ with a probability $h_i$.
according to a Zipf distribution with exponent $\gamma_c$, given by

$$h_i = \frac{1}{\sum_{\ell=1}^{m} \ell^{-\gamma_c}}, \quad 1 \leq i \leq m, \gamma_c > 0.$$ 

Here, we denote the file stored at cache device $j$ by $m_j$. The cache devices can be pooled together and form the fog nodes in fog computing.

Then, instead of fulfilling each content request by the BS, it is potentially beneficial to serve some request devices in $D$ via D2D communications with nearby cache devices in $S$. Considering that each device is equipped with a single omnidirectional antenna, each cache device can serve at most one request device. This D2D content distribution not only can offload traffic from the BS but also is more cost-effective in view of the close proximity. It is worth noting that, because each cache device is assumed to store only one video file temporarily to serve one request device, the short time scale can further reduce the chances that a pair of matched request and cache devices move out of each other’s transmission range during the content delivery. Even if such mobility changes the specific locations of these devices, we can use the new intensity functions of the resulting PPPs in the analysis.

### B. D2D Channel Model

Here, we consider a D2D underlaid cellular network, where the D2D links share the uplink spectrum of regular cellular users. There are several good reasons for favoring the use of uplink resources, such that the uplink resources are often less utilized, and the BS is more powerful in interference mitigation [26]. According to the reused cellular channels, we can split the D2D users into different groups. Each group of D2D users reuse an orthogonal cellular channel. The reused cellular channel should be appropriately selected by employing efficient channel allocation schemes [3]–[7], so that the integrated interference to the cellular user is minimized. With the spectrum orthogonality, we can analyze the performance integrated interference to the cellular user is minimized. With the D2D links sharing the uplink spectrum of regular cellular users. There are several good reasons for favoring the use of D2D underlaid cellular networks, where D2D users communicate only if they fall within a collaboration distance, $L$, where $0 < L < R$. Hence, a request device can only be paired with a feasible cache device, which contains the requested file and falls within the collaboration distance. The feasibility relationship between $S$ and $D$ can be modelled by a bipartite graph, depicted in Fig. 2. Then, the device pairing function $\varphi : S \mapsto D$, $\varphi(j')$ is to derive a one-to-one matching between set $S$ and set $D$, where $\varphi(j) = k$ indicates that cache device $j \in S$ is mapped to request device $k \in D$. Correspondingly, we use $D \subseteq D$ and $S \subseteq S$ to denote the subset of request devices and the subset of cache devices that are successfully paired, respectively.

Fig. 3 gives an example that illustrates the impact of D2D pairing. Here, the D2D pairing shown in Fig. 3(a) is worse than that in Fig. 3(b), because the shorter distances between the D2D pairs cause stronger interference. It is worth noting that, for clarity purpose, Fig. 3 only shows non-negligible interference signals in the interference range with the red dotted lines. As seen in Fig. 3(b), the D2D pairs are more spread out, which results in longer distances and weaker interference signals. Thus, the spectrum resources are shared more efficiently by the D2D pairs.

In fact, it is NP-hard [21] to derive an optimal device pairing, which ensures that the SINR at each D2D receiver is not less than a decoding threshold, $\beta$. An edge in Fig. 2 only indicates a feasibility condition. When an edge is included in a matching for the D2D pairing, it causes interference to all existing D2D pairs in the matching. That is, all edges in the matching are inter-dependent instead of independent. Therefore, we cannot obtain an optimal D2D pairing by finding a maximum or minimum weighted bipartite matching.

![Fig. 2: Bipartite graph modelling for D2D pairing.](image-url)
which is solvable in polynomial time. Hence, we consider a heuristic distance-based D2D pairing algorithm in Alg. 1. Here, each request device is assigned to a cache device, which is closest to the request device among those available (not paired to any device yet) and feasible cache devices that store the requested file and fall within the collaboration distance.

\[ P_s = P(k \in \tilde{D}) \cdot P(\xi_k \geq \beta | k \in \tilde{D}). \]  

\[ \max_{0 < L \leq R} E_{\Phi_r} [P_s] = E_{\Phi_r} \left[ P(k \in \tilde{D}) \cdot P(\xi_k \geq \beta | k \in \tilde{D}) \right]. \]  

\[ \varphi(j) \leftarrow k \text{ if } d_{j,k} < d_{\varphi^{-1}(k),k} \]  

Algorithm 1: A heuristic distance-based D2D pairing algorithm.

\begin{algorithm}
\caption{A heuristic distance-based D2D pairing algorithm.}
\begin{algorithmic}[1]
\STATE \textbf{Input:} \( S, D, L, P_d, P_c, \sigma^2, \{m_j^1 : j \in S\}, \) \( \{m_k^1 : k \in D\}, \{d_{j,k} : j \in S, k \in D\}, \) \( \{|h_{j,k}|^2 : j \in S, k \in D\}, \{d_{c,k} : k \in D\}, \) \( \{|h_{c,k}|^2 : k \in D\} \)
\STATE \textbf{Output:} Device pairing function \( \varphi \)
\FOR {\( j \in S \)}
\IF {\( m_j^1 = m_k^1 \) and \( d_{j,k} \leq L \)}
\IF {\( \varphi^{-1}(k) = \emptyset \) then // request device \( j \) is not assigned to any cache device}
\STATE \( \varphi(j) \leftarrow k \)
\ELSE 
\STATE \( \varphi(j) \leftarrow k \) // Choose the closest feasible cache device for request device \( k \)
\ENDIF
\ENDIF
\ENDFOR
\STATE \text{Return} \( \varphi \)
\end{algorithmic}
\end{algorithm}

This D2D-assisted content distribution scenario defined in Sections III-A, III-B and III-C is close to that considered in [9]. Differently, the system model in [9] divides the cell coverage into equal-area squares called clusters. To avoid intra-cluster interference among D2D links, only one communication pair is allowed to be active at one time, while inter-cluster interference is neglected assuming that an appropriate frequency reuse scheme is in place. In this work, however, we relax such constraints by allowing simultaneous active D2D pairs without a fixed clustering pattern while meeting certain minimum QoS in terms of SINR.

\section*{D. Optimal Collaboration Distance}

As seen in Alg. 1, a critical parameter for this distance-based pairing algorithm is the collaboration distance, \( L \). Intuitively, a larger collaboration distance introduces more interfering devices, while it also increases the possibility that a cache device that stores the requested file is located within the collaboration region for a request device. This trade-off shows that the collaboration distance should be determined properly to optimize the overall content distribution performance. On the other hand, the D2D pairing affects the spatial distributions of D2D transmitters and receivers, which also influences the perceived interference and the best collaboration distance.

To determine the optimal collaboration distance, we focus on the satisfaction probability, defined as

\begin{equation}
P_s = P(k \in \tilde{D}) \cdot P(\xi_k \geq \beta | k \in \tilde{D}).
\end{equation}  

\begin{equation}
\max_{0 < L \leq R} E_{\Phi_r} [P_s] = E_{\Phi_r} \left[ P(k \in \tilde{D}) \cdot P(\xi_k \geq \beta | k \in \tilde{D}) \right].
\end{equation}  

As seen, the expectation of the satisfaction probability captures the overall statistical performance in terms of fulfilling the content demands of all request devices. Also, the problem formulation in (3) takes into account the random spatial distributions of request devices and cache devices.

\section*{IV. Analysis of the Distance-Based D2D Collaborative Content Distribution Scheme}

As discussed in Section III-D, the collaboration distance \( L \) as a key parameter of the D2D pairing algorithm in Alg. 1 should be determined appropriately to balance the trade-off between D2D interference and feasibility of cache devices. As
seen in problem (3), the optimal collaboration distance should maximize the expectation of the satisfaction probability. In this section, we present an analytical framework to derive this expectation for the content distribution system described in Section III. Then, we can find the optimal collaboration distance based on this analytical framework.

In the proposed analytical framework, we first obtain a general expression for the expectation of the satisfaction probability. After that, a conditional expectation is analyzed for a given D2D receiver. Last, we remove this condition and derive the unconditional expectation for all possible D2D receivers. As the analysis involves many notations, Table I lists the important symbols for easy reference.

A. Expectation of Satisfaction Probability

Given the original PPP of all potential receivers, \( \Phi_r \), we focus on a thinned PPP, \( \tilde{\Phi}_r \), which is derived from \( \Phi_r \) by retaining a request device according to a retention probability. The thinned PPP \( \tilde{\Phi}_r \) only includes the request devices that have at least one feasible cache device within the collaboration distance. The thinned PPP has an intensity function given by \( \lambda_\tilde{r} = (1 - P_f)\lambda_r \), where \( P_f \) is the probability that there does not exist such a feasible cache device in the collaboration region for a potential request device.

According to the definition of \( P_f \) for thinned PPP \( \tilde{\Phi}_r \), we can express \( P_f \) as

\[
P_f = \sum_{\ell=0}^{\infty} \frac{N_f(L)}{\ell!} e^{-\Lambda_t(L)} \sum_{i=1}^{m} q_i \cdot (1 - h_i)^\ell
\]

where \( n_t(L) \) is the number of cache devices within the collaboration distance \( L \), which follows a Poisson distribution of mean \( \Lambda_t(L) = \lambda_t \pi L^2 \). The second summation term in (4) gives the probability that none of these \( \ell \) cache devices stores the requested file. Expanding (4), we have

\[
P_f = \sum_{\ell=0}^{\infty} \frac{N_f(L)}{\ell!} e^{-\Lambda_t(L)} \sum_{i=1}^{m} q_i \cdot (1 - h_i)^\ell
\]

Then, the expectation of the satisfaction probability in (3) is written as

\[
E_{\Phi_r}[P]\big|\big.\{k \in \tilde{D}\} = (1 - P_f) \cdot E_{\tilde{\Phi}_r}\left[ P\{\xi_k \geq \beta | k \in \tilde{D}\} \right]
\]

TABLE I: Important notations.

<table>
<thead>
<tr>
<th>Symbol</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>( D/S )</td>
<td>Set of request devices / cache devices</td>
</tr>
<tr>
<td>( D/S )</td>
<td>Subset of paired request devices / cache devices</td>
</tr>
<tr>
<td>( M )</td>
<td>Set of content items (files)</td>
</tr>
<tr>
<td>( m )</td>
<td>Number of files in ( M )</td>
</tr>
<tr>
<td>( R )</td>
<td>Radius of cell coverage region</td>
</tr>
<tr>
<td>( L )</td>
<td>Collaboration distance</td>
</tr>
<tr>
<td>( P_d )</td>
<td>Transmit power of D2D transmitters</td>
</tr>
<tr>
<td>( P_c )</td>
<td>Transmit power of cellular uplink user</td>
</tr>
<tr>
<td>( P_f )</td>
<td>Thinning probability for D2D receivers</td>
</tr>
<tr>
<td>( d_{j,k} )</td>
<td>Distance between transmitter ( j ) and receiver ( k )</td>
</tr>
<tr>
<td>( h_{j,k} )</td>
<td>Channel gain between transmitter ( j ) and receiver ( k )</td>
</tr>
<tr>
<td>( d_{c,k} )</td>
<td>Distance between cellular uplink user and receiver ( k )</td>
</tr>
<tr>
<td>( h_{c,k} )</td>
<td>Channel gain between cellular uplink user and receiver ( k )</td>
</tr>
<tr>
<td>( \eta )</td>
<td>Path-loss exponent</td>
</tr>
<tr>
<td>( \beta )</td>
<td>Decoding SINR threshold</td>
</tr>
<tr>
<td>( \sigma^2 )</td>
<td>Additive noise power</td>
</tr>
<tr>
<td>( \gamma_c, \gamma_r )</td>
<td>Exponent of Zipf distribution for caching / requests</td>
</tr>
<tr>
<td>( \lambda_r )</td>
<td>Intensity function of ( \Phi_r ) for potential receivers</td>
</tr>
<tr>
<td>( \lambda_t )</td>
<td>Intensity function of ( \Phi_t ) for potential transmitters</td>
</tr>
<tr>
<td>( \lambda_c )</td>
<td>Intensity function of ( \Phi_c ) for feasible transmitters</td>
</tr>
<tr>
<td>( \lambda_c )</td>
<td>Intensity function of ( \Phi_c ) for feasible transmitters</td>
</tr>
<tr>
<td>( \xi_k )</td>
<td>SINR at D2D receiver ( k )</td>
</tr>
<tr>
<td>( \theta_j )</td>
<td>Binary indicator on whether transmitter ( j ) is selected</td>
</tr>
<tr>
<td>( \varphi )</td>
<td>Matching function for D2D pairing, ( \varphi : S \rightarrow D )</td>
</tr>
<tr>
<td>( \eta )</td>
<td>Analysis parameter ( \eta = P_d \cdot d_{c,k}^\beta )</td>
</tr>
<tr>
<td>( \zeta )</td>
<td>Analysis parameter ( \zeta = \eta P_d</td>
</tr>
<tr>
<td>( \rho )</td>
<td>Adaptation parameter for D2D distance distribution</td>
</tr>
</tbody>
</table>

B. Decoding Success Probability

According to the SINR expression in (1), we can write the decoding success probability of a specific receiver \( k \) as

\[
P\{\{\xi_k \geq \beta | k \in \tilde{D}\} = P\left[ P_d d_{j,k}^{-\alpha} |h_{j,k}|^2 \geq \beta P_d d_{c,k}^{-\alpha} |h_{c,k}|^2 \right.
\]

\[
\left. + \sum_{j' \in \tilde{\Phi}_r, j' \neq j} P_d d_{j',k}^{-\alpha} |h_{j',k}|^2 \right]
\]

\[
\exp\left( -\eta \sigma^2 \right) \cdot \exp\left( -\eta P_d d_{c,k}^{-\alpha} |h_{c,k}|^2 \right)
\]

\[
\exp\left( -\eta \sum_{j' \in \tilde{\Phi}_r, j' \neq j} P_d d_{j',k}^{-\alpha} |h_{j',k}|^2 \right)
\]

where \( \tilde{\Phi}_r \) is a thinned PPP for the transmitters that are paired to the receivers in \( \Phi_r \). The properties of \( \tilde{\Phi}_r \) will be discussed further in Section IV-C. Simplifying the notations for the distance and channel gain in (7) by \( d_k \rightarrow d_{j,k}, d_c \rightarrow d_{c,k}, h_j \rightarrow h_{j,k}, h_c \rightarrow h_{c,k} \), we rewrite it as

\[
P\{\{\xi_k \geq \beta | k \in \tilde{D}\} = P\left[ |h_k|^2 \geq P_d^{-1} d_{c,k}^\alpha (\sigma^2 + P_d d_{c,k}^{-\alpha} |h_{c,k}|^2)
\]

\[
+ \sum_{j' \in \tilde{\Phi}_r, j' \neq j} P_d d_{j',k}^{-\alpha} |h_{j',k}|^2 \right]
\]

\[
\exp\left( -\eta \sigma^2 \right) \cdot \exp\left( -\eta P_d d_{c,k}^{-\alpha} |h_{c,k}|^2 \right)
\]

\[
\exp\left( -\eta \sum_{j' \in \tilde{\Phi}_r, j' \neq j} P_d d_{j',k}^{-\alpha} |h_{j',k}|^2 \right)
\]

where \( \eta \equiv P_d^{-1} d_{c,k}^\alpha \). The last equality in (8) is due to \( |h_k|^2 \sim \text{Exp}(1) \) with Rayleigh fading.
C. Conditional Coverage Probability

Next, we analyze the expectation of the decoding success probability, i.e., the coverage probability, by extending the stochastic geometry approach in [10]. Considering all receivers in \( \Phi_t \) and taking the expectation of the decoding success probability in (8), the unconditional coverage probability is obtained as

\[
E_{k \in \Phi_t} \left[ P(\xi_k \geq \beta) | k \in \tilde{D} \right] = E_{d_k} \left[ \exp \left( -\eta \sigma^2 \right) \cdot E_{|h_k|^2, d_k} \left[ \exp \left( -\eta P_c d_k^{-\alpha} |h_k|^2 \right) \right] \right] E_{|h_j|^2, d_j, d_k} \left[ \exp \left( -\eta \sum_{j' \in \Phi_t, j' \neq j} P_d d_j^{-\alpha} |h_j'|^2 \right) \right].
\]

Instead of directly deriving the unconditional coverage probability in (9), we first focus on the conditional coverage probability for a given receiver \( k \) that is \( d_k \) away from its transmitter, i.e.,

\[
E_{k \in \Phi_t} \left[ P(\xi_k \geq \beta) | k \in \tilde{D}, d_k \right] = \exp \left( -\eta \sigma^2 \right) \cdot E_{|h_k|^2, d_k} \left[ \exp \left( -\eta P_c d_k^{-\alpha} |h_k|^2 \right) \right]
\]

\[
E_{|h_j|^2, d_j, d_k} \left[ \exp \left( -\eta \sum_{j' \in \Phi_t, j' \neq j} P_d d_j^{-\alpha} |h_j'|^2 \right) \right].
\]

(10)

1) Second Expectation Term in (10): The second expectation term in (10) contains one important random variable \( r \), where \( d_c \) is the distance between a cellular user that is uniformly located within \( C \) and a D2D receiver in \( \Phi_t \). As derived in [10], the probability density function (pdf) of \( d_c \) is given by

\[
f_{d_c}(r) = \frac{2r}{\pi} \cos^{-1} \left( \frac{r}{2R} \right) - \frac{r}{\pi R} \sqrt{1 - \frac{r^2}{4R^2}}, \quad 0 \leq r \leq 2R
\]

where \( E[d_c] = \frac{128R}{\pi} \).

According to the analysis in [10], the second expectation in (10) is approximated by

\[
E_{|h_j|^2, d_j} \left[ \exp \left( -\eta P_c d_j^{-\alpha} |h_j|^2 \right) \right] = E_{d_j} \left[ \frac{1}{1 + \eta P_c d_j^{-\alpha}} \right] \approx \frac{1}{1 + \eta (qP/d_c)}.
\]

(12)

The first equality in (12) is because the Laplace transform of an exponentially distributed r.v. \( X \) is given by

\[
\mathcal{L}(s) = E[e^{-sx}] = \frac{1}{1 + s E[X]}.
\]

The approximate equality in (12) is based on the pdf of \( d_c \) given in (11).

2) Third Expectation Term in (10): The third expectation term in (10) is the most complex, and we extend the analysis in [10] to our scenario. First, we slightly transform the third expectation term in (10) as follows:

\[
E_{|h_j|^2, d_j} \left[ \exp \left( -\eta \sum_{j' \in \Phi_t, j' \neq j} P_d d_j^{-\alpha} |h_j'|^2 \right) \right] = E_{|h_j|^2, d_j} \left[ \prod_{j' \in \Phi_t, j' \neq j} \exp \left( -\eta P_d d_j^{-\alpha} |h_j'|^2 \right) \right].
\]

(14)

As (14) involves the thinned PPP \( \Phi_t \) for the cache devices that are paired to feasible request devices in \( \Phi_t \), we need to analyze \( \Phi_t \) in more detail. As a feasible transmitter must cache the same file as the requested file, the feasible transmitters within \( C \) actually forms a thinned PPP \( \Phi_t \) of an intensity function

\[
\tilde{\lambda}_t = \lambda_t \left[ 1 - \sum_{i=1}^m q_i (1 - h_i) \right]
\]

where the summation term gives the probability that the cached file is not the same as the requested one. Considering the D2D pairing algorithm in Alg. 1, when the collaboration distance \( L \) is larger and the request devices are more densely deployed within \( C \), there will be more competition toward the cache devices. As a result, a D2D receiver may not be paired with its nearest feasible transmitter within the collaboration region. Hence, applying the model fitting technique in [27], we incorporate an adaptation parameter \( \rho \) to \( \tilde{\lambda}_t \) in (15), i.e., \( \lambda_t \leftarrow \rho \lambda_t \), \( 0 < \rho \leq 1 \), in order to better fit the thinned PPP. In particular, when \( L \ll R, \rho \to 1 \).

Further, we continue to derive the expression for (14) by using the probability generating functional (pgf) of a point process, defined as

\[
G(v) = E \left[ \prod_{x \in \Phi} \text{exp} \left( -\eta P_d d_j^{-\alpha} |h_j|^2 \right) \right]
\]

where \( \Phi \) is a point process in \( \mathbb{R}^2 \) and \( v \) is a measurable function: \( v : \mathbb{R}^2 \mapsto [0, 1] \). In particular, the pgf of a PPP is given by

\[
G(v) = E \left[ \prod_{x \in \Phi} \text{exp} \left( -\eta P_d d_j^{-\alpha} |h_j|^2 \right) \right] = \text{exp} \left( - \int_{\Omega} \left[ 1 - \text{exp} \left( -\eta P_d d_j^{-\alpha} |h_j|^2 \right) \right] \Lambda \cdot dx \right)
\]

where \( \Lambda \) is the intensity measure of the PPP \( \Phi \).

Applying the conclusion in (17) to (14), we have

\[
E_{|h_j|^2, d_j} \left[ \prod_{j' \in \Phi_t, j' \neq j} \exp \left( -\eta P_d d_j^{-\alpha} |h_j'|^2 \right) \right] = E_{d_j} \left[ \prod_{j' \in \Phi_t, j' \neq j} \exp \left( -\eta P_d d_j^{-\alpha} |h_j'|^2 \right) \right] \approx \text{exp} \left( - \int_{\Omega} \left[ 1 - \text{exp} \left( -\eta P_d d_j^{-\alpha} |h_j|^2 \right) \right] \tilde{\lambda}_t \cdot dx \right)
\]

(18)

In (18), \( \tilde{\lambda}_t \) is the intensity measure of thinned PPP \( \Phi_t \). The integral term in (18) can be further expanded as follows:

\[
\int_{\Omega} \left[ 1 - \text{exp} \left( -\eta P_d d_j^{-\alpha} |h_j|^2 \right) \right] \tilde{\lambda}_t \cdot dx
\]

\[
= \int_0^{2\pi} d\theta \int_0^R \left[ 1 - \text{exp} \left( -\eta P_d r^{-\alpha} |h_j|^2 \right) \right] \tilde{\lambda}_t r \cdot dr
\]

\[
= 2\pi \tilde{\lambda}_t \int_{|h_j|^2} \left[ 1 - \text{exp} \left( -\eta P_d r^{-\alpha} |h_j|^2 \right) \right] r \cdot dr
\]

The last equality in (19) is due to the additive property of expectation.

Now, we focus on the integral term in (19). Defining \( \zeta \triangleq \eta P_d |h_j|^2 = \beta d_k^\rho |h_j|^2 \) and \( y \triangleq r^{-\alpha} \), we have

\[
\int_0^R \left[ 1 - \text{exp} \left( -\eta P_d r^{-\alpha} |h_j|^2 \right) \right] r \cdot dr
\]

\[
= \int_{R^{-\alpha}}^{-1/2} \frac{1}{(1 - e^{-\zeta y})} \cdot dy
\]

\[
= \frac{1}{2} \left[ (1 - e^{-\zeta y}) y^{-\frac{1}{2}} \right]_{R^{-\alpha}}^{\infty} + \int_{R^{-\alpha}}^{\infty} \frac{1}{2} y^{-\frac{1}{2}} e^{-\zeta y} \cdot dy.
\]
Letting $z \triangleq \zeta y$, we further expand (20) to

$$
\int_0^R \left[ 1 - \exp \left( - \eta P_d r^{-\alpha} |h_j|^2 \right) \right] \cdot r \cdot dr \\
= \frac{1}{2} R^2 \left( 1 - e^{-R^{-\alpha} z} \right) + \frac{1}{2} \zeta \int_0^\infty z^{(1 - \frac{2}{\alpha}) - 1} e^{-z} \cdot dz \\
= \frac{1}{2} R^2 \left( 1 - e^{-R^{-\alpha} z} \right) + \frac{1}{2} \zeta \hat{\Gamma} \left( 1 - \frac{2}{\alpha}, \zeta R^{-\alpha} \right)
$$

(21)

where $\hat{\Gamma} (\cdot)$ is the upper incomplete gamma function.

Placing (21) back into (19), we have

$$
\int_{\mathbb{R}^2} \left( 1 - E_{|h_j|^2} \left[ \exp \left( - \eta P_d d_j^{-\alpha} |h_j|^2 \right) \right] \right) \lambda_k \cdot dx \\
= E_{|h_j|^2} \left[ \pi \lambda_k R^2 \left( 1 - e^{-R^{-\alpha} z} \right) + \pi \lambda_k \zeta \hat{\Gamma} \left( 1 - \frac{2}{\alpha}, \zeta R^{-\alpha} \right) \right] \\
= E_{|h_j|^2} \left[ \pi \lambda_k R^2 \left( 1 - e^{-\beta d_k^{-\alpha} |h_j|^2} \right) \right] \\
+ \pi \lambda_k \beta d_k^2 \left( 1 - \frac{2}{\alpha} \right) \left( 1 + \frac{2}{\alpha} \right) \\
= \pi \lambda_k R^2 \left( 1 - \frac{1}{1 + \beta d_k^{-\alpha} R^{-\alpha}} \right) \\
+ \pi \lambda_k \beta d_k^2 \left( 1 - \frac{2}{\alpha} \right) \left( 1 + \frac{2}{\alpha} \right)
$$

(22)

In (22), the third equality comes from the Laplace transform of the exponential distribution of $|h_j|^2$; the fourth approximate equality is valid when $d_k \leq L \ll R$ and thus the upper incomplete gamma function approaches the gamma function $\Gamma (\cdot)$; the fifth equality is due to the fact that $E[|h_j|^2] = \Gamma(1 + (2/\alpha))$; and the last equality comes from the definition of the normalized sinc function, given by

$$
\text{sinc}(x) = \frac{\sin(\pi x)}{\pi x} = \frac{1}{\Gamma(1 - x) \Gamma(1 + x)}
$$

Substituting the integral term in (18) by (25), we have a new expression for the third expectation term in (10). Correspondingly, an upper bound for the conditional coverage probability is obtained as

$$
\begin{align*}
E_{k \in \Phi_r} \left[ P(\xi_k \geq \beta) | k \in \hat{D}, d_k \right] \\
= \exp \left( - P_d^{-1} d_k^2 \beta \sigma^2 \right) \cdot \frac{1}{1 + \frac{d_k^2 (2 \beta P_d^{-1})^{2/\alpha}}{(128 R^4 / (45 \pi)^2)} \\
& \cdot \exp \left( - \frac{1}{2} \hat{\lambda}_k (2 R)^2 \frac{\beta d_k^2}{1 + \beta d_k^2 (2 R)^{-\alpha}} - \frac{1}{2} \pi \hat{\lambda}_k \hbar^2 d_k^2 \frac{2}{\sin(\alpha)} \right)
\end{align*}
$$

(24)

D. Unconditional Coverage Probability

In this section, we further take the expectation of the conditional coverage probability in (24) over distance $d_k$ to obtain the unconditional coverage probability in (9). In fact, the following analysis provides a lower bound for the unconditional coverage probability. Similarly, referring to the upper bound for the conditional coverage probability in (26), we can obtain an upper bound for the unconditional coverage probability.

1) General Expression: The above analysis on the conditional coverage probability depends on distance $d_k$ for a given D2D pair. We can remove the condition by considering the probability distribution of $d_k$. Consider the collaboration given receiver $k$ that is $d_k$ away from its transmitter as

$$
E_{k \in \Phi_r} \left[ P(\xi_k \geq \beta) | k \in \hat{D}, d_k \right]
$$

(25)
region as a circle of a radius $L$ with receiver $k$ centered at the origin. The probability that the distance between receiver $k$ and its nearest transmitter within the collaboration region falls within $[r, r + \Delta r]$, where $(r + \Delta r) \leq L$, is given by

$$\Pr[n_k(r) = 0] \cdot \Pr[n_k(r + \Delta r) = 1] = e^{-\lambda_t \pi r^2} \cdot \Pr[\pi(L^2 - r^2 - \pi^2 r^2)] = \frac{e^{-\lambda_t \pi r^2}}{1 - e^{\lambda_t \pi L^2}}.$$  

(27)

Taking the pdf by letting $\Delta r \to 0$, we have

$$\lim_{\Delta r \to 0} \frac{\Pr[n_k(r) = 0] \cdot \Pr[n_k(r + \Delta r) = 1]}{\Delta r} = e^{-\lambda_t \pi r^2} \frac{\lim_{\Delta r \to 0} \Pr[\pi(L^2 - r^2 - \pi^2 r^2)]}{\Delta r} = 2\pi r \lambda_t e^{-\lambda_t \pi r^2}, \quad 0 \leq r \leq L.$$  

(28)

The subset of paired transmitting devices is modelled by a thinned PPP $\Phi_t$ of an intensity function $\lambda_t$ given in (15). Therefore, when $L \ll R$, the pdf of $d_k$ is given by

$$f_{d_k}(r) = 2\pi r \lambda_t e^{-\lambda_t \pi r^2}, \quad 0 \leq r \leq L.$$  

(29)

Combining (24) and (29) into (9), we obtain the unconditional coverage probability as

$$E_{k \in \Phi_t} \left[ \Pr[\xi_k \geq \beta | k \in \tilde{D}] \right] = \int_0^L E_{k \in \Phi_t} \left[ \Pr[\xi_k \geq \beta] \right] f_{d_k}(r) \cdot dr.$$  

(30)

2) Closed-Form Approximate Expression: Generally, the coverage probability in (30) can be evaluated numerically. For specific scenarios, we can also obtain a closed-form approximate expression for the unconditional coverage probability. In particular, when $L \ll R$ and $\sigma^2 = 0$, the conditional coverage probability in (24) is reduced to

$$E_{k \in \Phi_t} \left[ \Pr[\xi_k \geq \beta | k \in \tilde{D}, d_k] \right] = \frac{1}{1 + \frac{\sigma^2 (2\rho_d P_t^2)^{2/\alpha}}{(128R/(45\pi))^2}} \cdot \exp \left( -\frac{\pi \lambda_t \beta^2 d_k^2}{\sin(\frac{2}{\alpha})} \right).$$  

(31)

which is consistent with the coverage probability derived in [10] for a fixed D2D distance $d_k$. Then, applying the pdf of $d_k$ in (29) into (31), we obtain a closed-form expression for the unconditional coverage probability as follows:

$$E_{k \in \Phi_t} \left[ \Pr[\xi_k \geq \beta | k \in \tilde{D}] \right] = \int_0^L E_{k \in \Phi_t} \left[ \Pr[\xi_k \geq \beta] \right] f_{d_k}(r) \cdot dr.$$  

(32)

Rearranging (32) by substitution $r^2 \to x$, we have

$$E_{k \in \Phi_t} \left[ \Pr[\xi_k \geq \beta | k \in \tilde{D}] \right] = \int_0^{L^2} \frac{1}{1 + \psi x^2} \cdot e^{-\psi x^2} \cdot \nu x \cdot dr = \int_0^L \frac{1}{x + \psi x^2} \cdot e^{-\psi x^2} \cdot \nu x \cdot dr.$$  

$$= \frac{\nu}{2\psi} \cdot e^{\psi} \cdot \frac{\psi^2}{e^{\psi}} \cdot \int e^{-\psi x^2} \cdot dx$$

(33)

where $\text{Ei}(\cdot)$ is the one-argument exponential integral function defined as

$$\text{Ei}(x) = -\int_{-x}^{\infty} e^{-t} \cdot dt.$$  

(34)

The last equality in (33) is due to the following equation

$$\int_0^{x_{\max}} e^{-\psi x^2} e^{-\mu x} \cdot dx = e^{-\psi x_{\max}} \cdot \text{Ei}(\cdotx_{\max}) - \text{Ei}(-\mu x_{\max}) - \text{Ei}(\cdot(-\mu x_{\max})).$$  

(35)

E. Final Remarks

According to the above derivation, we can place the unconditional coverage probability in (30) or (33) back to (6) to evaluate the expected satisfaction probability. Then, an optimal collaboration distance $L$ can be searched within the range $[0, R]$ to maximize the expected satisfaction probability. Logically, when $L$ increases from a short distance, a cache device can potentially serve more request devices within the collaboration range, thus resulting in a larger satisfaction probability. On the other hand, when $L$ reaches a sufficiently large value, the interference among the D2D pairs becomes so large that more D2D pairs cannot meet decoding threshold $\beta$ successfully. Consequently, the satisfaction probability begins to decrease with $L$. As seen, the satisfaction probability is a unimodal function of $L$. Hence, we can use a search method such as the bisection search and the golden-section search to find the optimal $L$ that maximizes the satisfaction probability. Now, problem (3) is solved based on the D2D pairing algorithm in Alg. 1 and the analysis in this section.

V. NUMERICAL RESULTS AND DISCUSSIONS

In this section, we present numerical results to validate the analytical framework, examine the effects of various system parameters, and evaluate the performance of the distance-based pairing scheme for D2D-assisted content distribution.

Here, we consider that the request devices are placed according to a PPP of an intensity function $\lambda_r = 0.00001$ within a circular disk $C$ of a radius $R = 1000$ m [28]. The cache devices are distributed within $C$ as another PPP of a varying intensity function $\lambda_t \in [0.00001, 0.000033]$. The cache devices reuse the uplink of a cellular user that is uniformly distributed within $C$. The content requests from the request devices follow a Zipf distribution of an exponent $\gamma_r = 0.9$ [25] toward a library of size within $[10, 20]$. The cache devices store the files in the library according to another Zipf distribution of an exponent $\gamma_c \in [0.2, 2.4]$. 


The transmit powers of the underlaid cellular user and D2D transmitters are \( P_c = 100 \) mW and \( P_d = 10 \) mW, respectively [28]. The decoding threshold and the path-loss exponent are taken to be \( \beta = 3 \) dB [29] and \( \alpha = 4 \) [4], respectively. The D2D links and the interference signals from the cellular user are assumed subject to Rayleigh fading. Considering the randomness in channel fading, requested and cached files at D2D devices, and their physical locations, we evaluate the expected satisfaction probability by averaging 1000 independent realizations.

### A. Validation of Analytical Framework

First, we conduct experiments to evaluate the accuracy of the analytical framework given in Section IV. Here, we take \( \lambda_t = 0.00005 \), \( \gamma_c = 1.3 \), and \( m = 10 \). Fig. 4 shows the variation of the expected satisfaction probability with the collaboration distance \( L \). As seen in Fig. 4, when \( L \ll R \), both the lower bound and the upper bound obtained from our analytical framework provide quite accurate evaluation for the expected satisfaction probability. In contrast, when \( L \) approaches \( R \), the gap with the simulation results increases with a larger value of \( L \), which is due to the cell edge effect. Hence, it is more desirable to average the lower bound and the upper bound to obtain a more accurate estimate for the expected satisfaction probability.

In addition, as observed in Fig. 4, the expected satisfaction probability first increases with the collaboration distance \( L \) since more feasible cache devices are available within the collaboration region of a request device. After reaching a maximum, the expected satisfaction probability further decreases with \( L \), because it is possible that the paired D2D transmitters become more clustered rather than spread out, thereby causing more interference to each other.

### B. Effect of Cache Device Density

Next, we show how the expected satisfaction probability varies with the intensity of the cache devices. Here, the intensity function of cache devices \( \lambda_t \) is varied from 0.00001 to 0.00033. We first derive an optimal collaboration distance using the analytical framework given in Section IV, then apply the D2D pairing algorithm described in Alg. 1, and obtain the expected satisfaction probability by averaging the results of 1000 realizations. As seen in Fig. 5, the expected satisfaction probability increases with the intensity of cache devices. This is because with a higher intensity of D2D transmitters it is likely that more request devices are able to be paired with closer cache devices in good channel conditions. In addition, it can be seen that the simulation results for the expected satisfaction probability match well the analytical results with the optimal collaboration distance.

### C. Effect of Additive Noise

Fig. 6 compares the expected satisfaction probability of the heuristic distance-based D2D pairing algorithm in Alg. 1 with that of the optimal D2D pairing solution. This optimal pairing solution is based on the integer linear program (ILP) formulation in [21] for the D2D pairing problem, which aims to maximize the total number of D2D pairs that meet the successful decoding condition. As seen, when the variance of the additive noise power is low, the gap between the heuristic pairing algorithm and the optimal solution from the ILP formulation can be as large as 11%. However, one main advantage of the heuristic algorithm is its low complexity. In contrast, a major concern with the ILP formulation is that it can be much more time-consuming and even computationally infeasible to obtain the optimal solution due to the NP-hardness, especially when there are a large number of devices. Moreover, the heuristic pairing algorithm leads to tractable spatial distributions for the devices that enable feasible stochastic analysis of the satisfaction probability. Conversely, it is difficult to model the distances between the devices that result from the optimal pairing solution. Furthermore, it is observed that the performance gap between the heuristic algorithm and the optimal solution becomes smaller with the increase of the...
Variance of additive noise power $\sigma^2$ (dBm)

Expected satisfaction probability

0.15 0.2 0.25 0.3 0.35 0.4

-100 -95 -90 -85 -80 -75 -70

Heuristic pairing: analysis
Heuristic pairing: simulation
Optimal pairing: analysis

Fig. 6: Expected satisfaction probability of D2D pairing schemes with different variance of additive noise power $\sigma^2$.

D. Effect of Caching Parameters

Last, we examine the impact of the caching parameters on the achievable performance in the D2D content distribution. Here, we vary the exponent $\gamma_c$ for the Zipf distribution of cached files and the size of the file library $m$. As seen in Fig. 7(a), the expected satisfaction probability first increases with $\gamma_c$ and then decreases slowly with $\gamma_c$. An optimal value of around 1.3 for $\gamma_c$ maximizes the expected satisfaction probability. A value lower or higher than the optimal value implies that the files stored at the cache devices are too scattered or too concentrated as opposed to the requests. Therefore, such a caching policy is not the best since the setting for $\gamma_c$ does not maximize the hit ratio at the feasible cache devices.

On the other hand, it is observed in Fig. 7(a) that the expected satisfaction probability decreases with the library size. This is because with a larger library size there is obviously a lower possibility that a file request can be fulfilled by a cache device in the neighbourhood. Correspondingly, as shown in Fig. 7(b), with a larger library size, the optimal collaboration distance derived by the analysis is larger to increase the likelihood for locating feasible cache devices within the collaboration region.

In addition, Fig. 7(b) shows that the optimal collaboration distance decreases with the caching exponent $\gamma_c$. This observation can be interpreted as follows. A larger value of $\gamma_c$ means that the cache devices are concentrated on fewer stored files. As a consequence, a request device is able to find a feasible cache device within a smaller collaboration region such that more request devices can successfully fit in while still achieving satisfactory QoS.

VI. CONCLUDING REMARKS

In this paper, we analyze a heuristic distance-based pairing scheme for collaborative content distribution via D2D communications. The pairing scheme matches a request device with a cache device within a collaboration distance. The collaboration distance is a critical parameter, as it affects the interference between the D2D pairs and the availability of cache device within the collaboration region. To balance the trade-off and determine an optimal collaboration distance, we develop an analytical framework to evaluate the expected satisfaction probability for content requests. Our analysis takes into account a realistic interference model and derives the unconditional coverage probability. In addition, a closed-form approximate expression is provided, if the additive noise is negligible in comparison with the co-channel interference from the regular cellular user and D2D users. The simulation results validate the analytical framework and demonstrate the effects of various system parameters, such as the collaboration distance, cache device density, additive noise power, and caching exponent. It is also shown that the performance of the distance-based algorithm is fairly close to that of the optimal solution in some content distribution scenarios.