MDS Coded Caching for Device-to-Device Content Sharing Against Eavesdropping

Xin Shi[1], **Dan Wu**[1*], **Meng Wang**[1], **Lianxin Yang**[1] and **Yan Wu**[1]
[1]College of Communications Engineering, Army Engineering University of PLA
Nanjing, Jiangsu 210007, China
[e-mail: shi_xin1995@126.com]
*Corresponding author: Dan Wu

Received December 21, 2018; revised March 5, 2019; accepted April 14, 2019; published September 30, 2019

Abstract

In this paper, we put forward a delay-aware secure maximum distance separable (MDS) coded caching scheme to resist the eavesdropping attacks for device-to-device (D2D) content sharing by combining MDS coding with distributed caching. In particular, we define the average system delay to show the potential coupling of delay-content awareness, and learn the secure constraints to ensure that randomly distributed eavesdroppers cannot obtain enough encoded packets to recover their desired contents. Accordingly, we model such a caching problem as an optimization problem to minimize the average system delay with secure constraints and simplify it to its convex relaxation. Then we develop a delay-aware secure MDS coded caching algorithm to obtain the optimal caching policy. Extensive numerical results are provided to demonstrate the excellent performance of our proposed algorithm. Compared with the random coded caching scheme, uniform coded caching scheme and popularity based coded caching scheme, our proposed scheme has 3.7%, 3.3% and 0.7% performance gains, respectively.

Keywords: D2D content sharing, MDS coded caching, secure constraints, eavesdropping, convex relaxation

This work is supported by the Jiangsu Provincial Natural Science Fund for Outstanding Young Scholars (No. BK20180028), and the National Natural Science Foundation of China (Grant No. 61671474), and Jiangsu Provincial Natural Science Fund for Excellent Young Scholars (No. BK20170089).

http://doi.org/10.3837/tiis.2019.09.009 **ISSN : 1976-7277**

1. Introduction

Device-to-device (D2D) content sharing is considered as a proper solution to deal with the irreconcilable contradiction between increasing enormous data traffic and scarce radio spectrum resources [1]. By making full use of the limited storage, computing and transmission capabilities of mobile user devices, D2D content sharing allows user devices to share contents directly through D2D links instead of cellular links [2]. In a typical D2D content sharing scenario, the content providers (CP) selectively store some contents obtained from the base station (BS), and the content requesters (CR) prefer to ask their adjacent CPs to transmit their desired contents via D2D communications [3-5]. Meanwhile, a proper distributed content caching plays a crucial role in ensuring advantages brought by D2D content sharing, e.g., high spectral efficiency, short transmission delay, and low energy consumption [6].

Unfortunately, D2D content sharing is more vulnerable to the eavesdropping issues than traditional cellular communications [7]. That is due to the following facts: i) There exist more open wireless channels by adding D2D links. ii) The lack of central authority results in additional functionalities undertaken by user devices such as auditing and logging [8]. iii) User devices have limited storage, computing and transmission capabilities for security-related works. iv) The distributed caching extends the scope of content sources, which increases the possibility of information leakage. As such, in this paper we will focus on the D2D content sharing against eavesdropping.

Generally, there are two main categories of techniques for preventing eavesdropping: i) One is the end-to-end encryption technique. However, the decentralized and large-scale nature of D2D content sharing makes key management difficult [9]. Besides, due to the limited computing power of user devices, they could not take more care in complex encryption and decryption works. More critically, the encrypted contents are uniquely for certain CRs and cannot be reused to serve other CRs. Thus the benefits of distributed content caching may well vanish. ii) The other is physical layer security technique, which exploits the wireless channel characteristics to prevent eavesdroppers based on the information theoretic secrecy analysis of Shannon [10]. Physical layer security technique has attracted wide attention recently. Authors in [11] investigate the resource allocation problem in D2D communications while the secrecy communications for both cellular users and D2D pairs are ensured by utilizing physical layer security technique. In [12], an access selection scheme is proposed to protect cellular users from multiple eavesdroppers by exploiting the interference generated by selected D2D user devices. Authors in [13] solve the resource management problem for D2D underlaying cellular networks from a physical layer security perspective in order to provide security for cellular users and improve spectral efficiency for D2D communications. However, in spite of the compatibility between distributed caching and physical layer security technique, most of the existing works need at least partial channel state information (CSI) of the eavesdropping nodes, which are difficult to obtain in practice.

Note that, the CPs have limited storage and transmission capabilities, and the eavesdroppers are able to wiretap the contents stored and transmitted by the CPs within their D2D communication range by impersonating to be legitimate users. Hence, it motivates us whether we could exploit the flexibility offered by local caches, so as to not only conform to the Quality of Service (QoS) demands of certain CRs, but also deliver the contents securely. Moreover, the Maximum Distance Separable (MDS) coding gains increasing attention in heterogeneous small cell networks. In particular, it mainly plays an active role in making full use of devices' cache capacities, by which the contents are divided into small encoded packets, then stored in different CPs [14-16]. In fact, MDS coding has another advantage from a secure backhaul for cache placement perspective. If the number of the encoded packets collected by eavesdroppers is not enough, the original content cannot be recovered [17]. With this regard, [14] gives us a clear inspiration. However, it considers an insecure backhaul for heterogeneous cellular networks. The eavesdroppers tapping the insecure backhaul can be prevented from obtaining a sufficient number of encoded packets for successful recovery of contents. In this way, the combination of distributed caching and MDS coding will open up a new avenue for protecting D2D content sharing from eavesdropping threats. This hence gives rise to new challenges in terms of how to fully tap the potential of such combination, with the goal of pursuing successful and rapid content delivery while satisfying the secure constraints.

Based on the above analyses, we provide a new solution to the eavesdropping attacks for D2D content sharing in this paper by combining MDS coding with distributed caching. Formally, we conclude our main contributions as follows.

i) Without loss of generality, we model a D2D content sharing scenario consisting of low-mobility user devices and randomly distributed eavesdroppers. In this context, we combine MDS coding with distributed caching, and further define the average system delay as a performance metric, which taps the potential relationship between the QoS performance, in terms of delay, and the content awareness.

ii) Then, we obtain the secure constraints which prevent any eavesdropper from sniffing enough encoded packets to recover the original contents. Guided by these analyses, we propose a delay-aware secure MDS coded caching scheme. That is, we formulate it as a constrained optimization problem to minimize the average system delay while guaranteeing D2D content sharing security under arbitrarily distributed eavesdroppers.

iii) In order to solve the proposed optimization problem, we simplify it to its convex programming relaxation, and propose a delay-aware secure MDS coded caching algorithm to obtain an optimal caching policy. Numerical results demonstrate the effectiveness of our proposed algorithm.

The remainder of this paper is organized as follows. We introduce the considering D2D content sharing scenario and system model in Section 2. In Section 3, a D2D content sharing paradigm with MDS coded caching is given to explain the distributed caching. Then, the average system delay and secure constraints based on MDS coding and distribute caching are proposed in Section 4. Moreover, we formulate the delay-aware secure MDS coded caching as a constrained optimization problem and simplify it to its convex relaxation, followed by our proposed algorithm. In Section 5, numerical results are presented to illustrate the excellent performance of our proposed algorithm. Finally, Section 6 concludes this work. The most used notations and symbols in this paper are presented in **Table 1**.

Table 1. The summary of notations

Notations	Meanings
$\mathcal{F}$, $\mathcal{CP}$, $\mathcal{CR}$	Set of contents, CPs and CRs
J , K , I	Number of contents, CPs, and CRs
$\mathcal{CP}^i$	Set of CPs that serve CR_i
s_j	Size of content F_j
C_k	Cache capacity of CP_k
d	Distance between adjacent CPs
r	D2D communication range

n , l	Number of fragments and encoded packets in a MDS coding
$\mathcal{F}_j$	Encoded packets set of content F_j
$m_j^{CP_k}$	Number of encoded packets stored in CP_k about content F_j
$q_j^{CP_k}$	Ratio of content F_j stored in CP_k
R	Achievable rate
B	Channel bandwidth
P	Transmission power
H	Channel gain
σ^2	Noise power spectrum density
p_j	Request probability of content F_j
α	Skewness of a Zipf distribution
D_{CR_i,F_j}	Delay for CR_i to obtain content F_j
D_{CR_i}	Average delay for CR_i to obtain its desired content
η	Average system delay
S	Maximum number of CPs that can be wiretapped by the eavesdroppers

2. System Model

We investigate D2D content sharing in a single-cell cellular network without loss of generality, where a number of users who carry smart devices are randomly distributed, and a central BS is reachable from any user within the cell. The BS can access a data center so that it stores a library of J contents, denoted by $\mathcal{F} = \{F_1, \cdots, F_j, \cdots, F_J\}$. The size of each content F_j is s_j. The D2D content sharing model may involve in some typical scenarios, e.g., school teaching buildings, administration buildings and stadiums. In this kind of D2D content sharing scenario, most user devices have similar content preferences, i.e., they are interested in the same content such as hot news, popular music and videos. More importantly, user devices are of high density and low mobility in such scenarios. In this regard, to make security-related content placement convenient, user devices located on a regular grid on the unit square can be selected as CPs [18]. As thus, all the CPs in the single cell are spaced out a distance d apart, shown as **Fig. 1(a)**, and they comprise a set $\mathcal{CP} = \{CP_1, \cdots, CP_k, \cdots CP_K\}$. These CPs obtain their desired contents from the BS. Then, due to the fact that each CP_k has a cache capacity C_k, it can selectively store some contents according to a certain caching policy. Moreover, other user devices act as CRs, denoted by $\mathcal{CR} = \{CR_1, \cdots, CR_i, \cdots, CR_I\}$. They are able to obtain their desired contents from the adjacent CPs by establishing D2D links. In practical terms, each CR_i can be served by multiple CPs, to which the distances from CR_i are not greater than D2D communication range r. Hence, let $\mathcal{CP}^i$ denote the set of CPs that enable to provide the desired contents for CR_i. For instance, CR_i in **Fig. 1(b)** is able to obtain its desired contents from CP_1 and CP_2 ,thus $\mathcal{CP}^i = \{CP_1, CP_2\}$.

Moreover, some illegal devices who impersonate to be legitimate users are randomly distributed in the cell, and that they may well access the contents stored by the CPs within the D2D communication rage r. In this work, the illegal devices are considered as eavesdroppers due to the fact that they want to obtain their desired contents in an illegal manner, and such behavior is considered as eavesdropping. For instance, the distances from the eavesdropper in **Fig. 1(b)** to CP_1, CP_2 and CP_3 are smaller than the D2D communication range r so that it could eavesdrop on the contents stored by these three CPs. Particularly, the eavesdroppers only work passively in the eavesdropping mode without sending any signals. Thus, it is difficult to learn their CSI. Note that, the potential eavesdropping attacks of cellular links and D2D links occur independently. Thus, due to the fact that security techniques of traditional cellular communications have been studied widely and deeply for networks without caching [19-21], this paper is mainly restricted to the eavesdropping attacks of D2D links. Here, we exploit the security framework in [22] to prevent the eavesdropping attacks of cellular links. Besides, the incentive mechanism in [23] is exploited to motivate CPs to provide contents to CRs, and thus we focus on caching strategy in this paper.

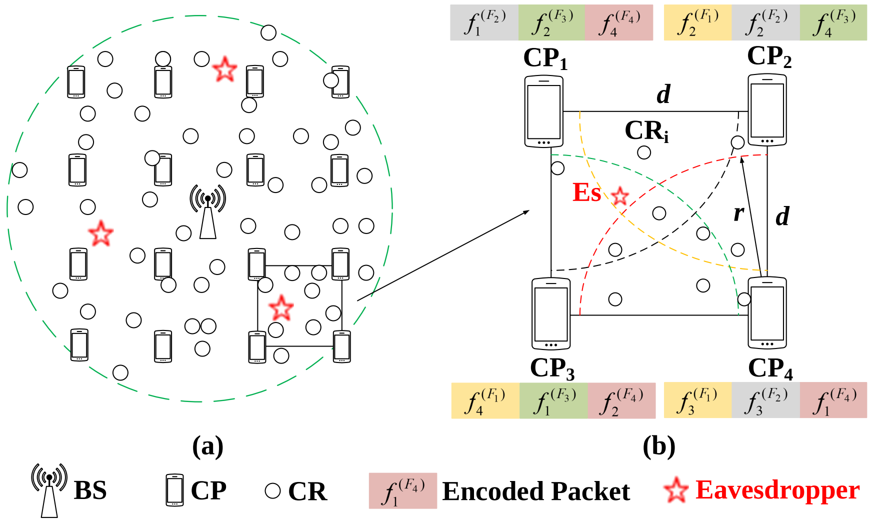

$f_1^{(F_2)}$ $f_2^{(F_3)}$ $f_4^{(F_4)}$
$f_2^{(F_1)}$ $f_2^{(F_2)}$ $f_4^{(F_3)}$
CP$_1$
CP$_2$
d
CR$_i$
Es
r
d
CP$_3$
CP$_4$
$f_4^{(F_1)}$ $f_1^{(F_3)}$ $f_2^{(F_4)}$
$f_3^{(F_1)}$ $f_3^{(F_2)}$ $f_1^{(F_4)}$
(a)
(b)
BS
CP
CR
$f_1^{(F_4)}$ Encoded Packet
Eavesdropper

Fig. 1. System model of D2D content sharing with eavesdroppers

3. D2D Content Sharing Paradigm with MDS Coded Caching

Using MDS coding, a D2D content sharing paradigm is given in this section. In a MDS (l,n) coding, each content is separated into n fragments firstly and then coded into l encoded packets [24]. In a word, by utilizing MDS coding, two-fold advantages are brought. i) Any set consisting of more than n encoded packets can be used to recover the original content. In other words, once just less than n encoded packets are received, the original content cannot be recovered. Accordingly, we can exploit this point to propose the secure constraints, that is, by tactfully placing the encoded packets, the number of encoded packets that may be wiretapped by eavesdroppers is less than n, so that they cannot recover the original content. ii) All the encoded packets are independent and non-overlapping, which allows researchers to calculate the system performance like backhaul rate and transmission delay only by the number of encoded packets stored in each CP_k rather than the details.

In general, there exist two phases for D2D content sharing, i.e., placement phase and delivery phase. In the placement phase, the BS encodes the contents using MDS coding and according to a certain caching policy, each CP_k obtains some encoded packets from the BS to store in their caches. In the delivery phase, the CRs prefer to request their desired contents from the CPs within their D2D communication range. Once the CRs cannot receive enough encoded packets to recover the content, they will ask the BS for the rest encoded packets.

3.1 Placement Phase

Using a MDS (l,n) coding, $\forall F_j \in \mathcal{F}$ is cut into n_j fragments firstly and then coded into l_j independent and non-overlapping encoded packets by the BS, denoted by $\mathcal{F}_j = \{f_1^{(F_j)}, f_2^{(F_j)}, \cdots, f_{l_j}^{(F_j)}\}$. To ensure the encoded packets received by each CR_i are non-overlapping, the number of encoded packets for content F_j should be at least

$$l_j = n_j + \max_{CR_i \in \mathcal{CR}} \sum_{CP_k \in \mathcal{CP}^i} m_j^{CP_k}, \tag{1}$$

where n_j denotes the number of encoded packets stored in the BS, $m_j^{CP_k}$ denotes the number of encoded packets that CP_k receives from the BS so that $\max_{CR_i \in \mathcal{CR}} \sum_{CP_k \in \mathcal{CP}^i} m_j^{CP_k}$ denotes the largest number of encoded packets which are stored by the CPs in the set $\mathcal{CP}^i$ for $\forall CR_i \in \mathcal{CR}$. After that, CP_k stores the $m_j^{CP_k}$ encoded packets about F_j in its cache. By this means, we are committed to determine the optimal $m_j^{CP_k}$, so that each CR_i can obtain the required contents with satisfactory QoS performance, and not to be eavesdropped. An example is given in **Fig. 1(b)**, where $\mathcal{F} = \{F_1, F_2, F_3, F_4\}$. Each content is coded into small encoded packets using a MDS coding, and each CP_k selectively stores some encoded packets.

3.2 Delivery Phase

In the delivery phase, CR_i requests content $F_j \in \mathcal{F}$ with a probability p_j at a time. Firstly, CR_i receives all the $\sum_{CP_k \in \mathcal{CP}^i} m_j^{CP_k}$ non-overlapping encoded packets stored in the CPs within its D2D communication range r , i.e., all the $CP_k \in \mathcal{CP}^i$. The achievable rate of CR_i receiving the encoded packets from CP_k can be characterized as

$$R_{CR_i}^{CP_k} = B \log_2 (1 + \frac{P_{CP_k} H_{CP_k, CR_i}}{B\sigma^2}), \tag{2}$$

where B represents the channel bandwidth, P_{CP_k} is the transmission power of CP_k, H_{CP_k, CR_i} denotes the channel gain from CP_k to CR_i, and σ^2 is the noise power spectrum density. Then, CR_i follows the following rules: i) If $\sum_{CP_k \in \mathcal{CP}^i} m_j^{CP_k} \geq n_j$, CR_i can recover the content F_j successfully. ii) Otherwise, CR_i asks the BS for sending the remaining $n_j - \sum_{CP_k \in \mathcal{CP}^i} m_j^{CP_k}$ encoded packets so that the content F_j can be well recovered. Similarly, the achievable rate of CR_i receiving the encoded packets from the BS can be represented as

$$R_{CR_i}^{BS} = B \log_2 (1 + \frac{P_{BS} H_{BS, CR_i}}{B\sigma^2}), \tag{3}$$

in which P_{BS} is the transmission power of the BS, and H_{BS,CR_i} denotes the channel gain from the BS to CR_i. Here, we employ the underlying D2D communications with the dedicated mode. Then, we emploit the proposed scheme in [25] to schedule the link resources and coordinate the power to manage the interference, and exploit the constellation rotation technique in [26] to further eliminate the interference, so that the interference does not have impact on the achievable rate.

4. Delay-Aware Secure MDS Coded Caching

We focus on the MDS coded caching to satisfy QoS of certain CRs as well as preventing eavesdropping. Delay, as one of the key requirements for D2D content sharing, is increasingly important due to the explosion of delay-sensitive and real-time applications, e.g., online videos and virtual reality. Hence, we define average system delay η as the performance metric that indicates the average time taken for each CR_i to obtain its desired content. Moreover, the security issues of the eavesdroppers who impersonate to be legitimate users to wiretap the contents stored in the CPs within their D2D communication range trouble us, i.e., how to ensure the contents shared by D2D links not to be eavesdropped. A reasonable way is to ensure the security of content sharing by putting forward secure constraints based on the properties of MDS coding, one of the most used fountain codes, i.e., a sufficient number of encoded packets can be obtained by eavesdroppers by tactfully encoded packets placement [27]. In this regard, we formulate the delay-aware secure MDS coded caching problem as an optimization problem to minimize the average system delay with secure constraints.

Firstly, the delay for CR_i to get enough encoded packets of F_j can be calculated by

$$D_{CR_i,F_j} = [\sum_{CP_k \in \mathcal{CP}^i} \frac{m_j^{CP_k}}{R_{CR_i}^{CP_k}} + \frac{(n_j - \min(n_j, \sum_{CP_k \in \mathcal{CP}^i} m_j^{CP_k}))^+}{R_{CR_i}^{BS}}]\frac{s_j}{n_j}, \tag{4}$$

where $(\cdot)^+ = \max(\cdot, 0)$, the first part represents the delay for CR_i getting the encoded packets from the adjacent $CP_k \in \mathcal{CP}^i$, and the latter part represents the delay for CR_i obtaining the encoded packets from the BS if all the encoded packets obtained from $CP_k \in \mathcal{CP}^i$ are not enough to recover the original content. Then, we define the average delay for CR_i as

$$D_{CR_i} = \sum_{j=1}^{J} p_j D_{CR_i,F_j}, \tag{5}$$

where $p_j \in [0,1]$ is the request probability of F_j, which can be obtained from a record of historical requests over a period of time in practice. Formally, it obeys a Zipf distribution [28], denoted as

$$p_j = \frac{1/j^{\alpha}}{\sum_{j=1}^{J} 1/j^{\alpha}}, \tag{6}$$

where α denotes the skewness of the distribution. The essence of the average system delay is to tap the potential relationship between the delay and the content awareness. That is, the more popular content may well be requested by more CRs, and thus, the delay for obtaining it has the more powerful effect. Accordingly, the average system delay η can be expressed as

$$\eta = \frac{1}{I}\sum_{i=1}^{I} D_{CR_i} = \frac{1}{I}\sum_{i=1}^{I}\sum_{j=1}^{J} p_j[\sum_{CP_k \in \mathcal{CP}^i} \frac{m_j^{CP_k}}{R_{CR_i}^{CP_k}} + \frac{(n_j - \min(n_j, \sum_{CP_k \in \mathcal{CP}^i} m_j^{CP_k}))^+}{R_{CR_i}^{BS}}]\frac{s_j}{n_j}. \quad (7)$$

Delay for the BS downloading contents from data center, delay for CPs downloading encoded packets from the BS, and propagation delay for wireless signals in free space are not considered in this paper due to the fact that compared with content sharing process, they are negligible.

Note that each CP_k will send all the encoded packets stored in its cache to certain CR once it receives content request from the CR. Randomly distributed eavesdroppers in the network may impersonate to be legitimate users to eavesdrop the encoded packets transmitted by the CPs within their D2D communication range r. Then, considering the worst case that all the encoded packets stored by CPs are transmitted over D2D links if they receive content requests, and all the encoded packets transmitted through D2D links can be obtained by eavesdroppers, we put forward secure constraints by defining S as the maximum number of CPs that can be wiretapped by the eavesdroppers randomly distributed in the network without loss of generality, i.e., $S = \max_{CR_i \in \mathcal{CR}} |\mathcal{CP}^i|$, where $|\mathcal{CP}^i|$ denotes the number of elements in the set $\mathcal{CP}^i$. In particular, S is relevant with the distance between adjacent CPs d and D2D communication range r as far as our system model is concerned.

Then, based on the properties of MDS coding, the secure constraints are given as

$$m_j^{CP_k} S < n_j, \forall CP_k \in \mathcal{CP}, F_j \in \mathcal{F}, \quad (8)$$

so that the encoded packets wiretapped by any eavesdropper in the network within its D2D communication range r are not enough to recover the original content.

As such, the delay-aware secure MDS coded caching problem can be formulated as the following constrained optimization problem, i.e.,

$$\begin{aligned} &\min_{\mathbf{m}} \quad \eta \\ &s.t. \quad \sum_{j=1}^{J} \frac{m_j^{CP_k}}{n_j} s_j \le C_k, \forall CP_k \in \mathcal{CP} \\ &\qquad\ 0 \le m_j^{CP_k} S < n_j, \forall CP_k \in \mathcal{CP}, F_j \in \mathcal{F}, \end{aligned} \quad (9)$$

where $\mathbf{m} \in \mathbb{R}^{K \times J}$ and $[\mathbf{m}]_{k,j} = m_j^{CP_k}$. The first constraint means that each CP_k can store no more than its cache capacity C_k encoded packets totally. The second constraint indicates the secure constraints. Note that optimization problem (9) is a complex nonconvex optimization problem due to the non-convex objective function (7). In order to simplify it, we propose Proposition 1 shown as follows.

Proposition 1. *Given any secure MDS coded caching policy satisfying constraint (8), the encoded packets received from all the $CP_k \in \mathcal{CP}^i$ for each CR_i are not enough to recover the original content, i.e.,*

$$\sum_{CP_k \in \mathcal{CP}^i} m_j^{CP_k} < n_j, \forall CR_i \in \mathcal{CR}, F_j \in \mathcal{F}. \quad (10)$$

Proof. In order to prove (10), we denote $\{CP_k, CP_{k+1}, \cdots CP_{k+|\mathcal{CP}^i|-1}\}$ as the set of CPs that belong to $\mathcal{CP}^i, \forall CR_i \in \mathcal{CR}$. It can be rewritten as $\sum_{CP_k \in \mathcal{CP}^i} m_j^{CP_k} = m_j^{CP_k} + m_j^{CP_{k+1}} + \cdots + m_j^{CP_{k+|\mathcal{CP}^i|-1}}$.

According to (8), it can be seen that $m_j^{CP_k} < \frac{n_j}{S}, \forall CP_k \in \mathcal{CP}, F_j \in \mathcal{F}$, thus we have the following inequation, $\max\{m_j^{CP_k}, m_j^{CP_{k+1}}, \cdots, m_j^{CP_{k+|\mathcal{CP}^i|-1}}\} < \frac{n_j}{S}, \forall F_j \in \mathcal{F}$. Then, due to the equation $S = \max_{CR_i \in \mathcal{CR}} |\mathcal{CP}^i|$, we have $|\mathcal{CP}^i| \le S, \forall CR_i \in \mathcal{CR}$. Based on the above analyses, we come to the result $\sum_{CP_k \in \mathcal{CP}^i} m_j^{CP_k} \le \max\{m_j^{CP_k}, m_j^{CP_{k+1}}, \cdots, m_j^{CP_{k+|\mathcal{CP}^i|-1}}\} S < n_j, \forall CR_i \in \mathcal{CR}, F_j \in \mathcal{F}$. Hence, Proposition 1 is proved.

Moreover, we define $q_j^{CP_k} = m_j^{CP_k} / n_j$ as the ratio of content F_j stored in CP_k to simplify (7) and (8). Based on Proposition 1 and the definition of $q_j^{CP_k}$, we denote $\hat{\eta}$ as the average system delay while constraint (8) is satisfied, then (7) can be reformulated as

$$\hat{\eta} = \frac{1}{I}\sum_{i=1}^{I}\sum_{j=1}^{J} p_j [\frac{1}{R_{CR_i}^{BS}} + (\frac{1}{R_{CR_i}^{CP_k}} - \frac{1}{R_{CR_i}^{BS}}) \sum_{CP_k \in \mathcal{CP}^i} q_j^{CP_k}] s_j . \tag{11}$$

Besides, (8) can be rewritten as $q_j^{CP_k} < \frac{1}{S}$. It can be easily seen that (11) is a linear, and of course, convex fucntion. Then, the delay-aware secure MDS coded caching problem (9) can be reformulated as the following linear programming and convex optimization problem according to Chapter 4.3 in [29], i.e.,

$$\begin{aligned} \min_{\mathbf{q}} \quad & \hat{\eta} \\ s.t. \quad & \sum_{j=1}^{J} q_j^{CP_k} s_j \le C_k, \forall CP_k \in \mathcal{CP} \\ & 0 \le q_j^{CP_k} < \frac{1}{S}, \forall CP_k \in \mathcal{CP}, F_j \in \mathcal{F}, \end{aligned} \tag{12}$$

where $\mathbf{q} \in \mathbb{R}^{K \times J}$ and $[\mathbf{q}]_{k,j} = q_j^{CP_k}$. In order to solve problem (12), a delay-aware secure MDS coded caching algorithm is summarized in Algorithm 1 and described in detail as follows.

With given contents sizes $\mathbf{s}$ and cache capacities $\mathbf{C}$, we firstly define V_k to represent the current used caching capacity of CP_k. For all the $CP_k \in \mathcal{CP}$ and $F_j \in \mathcal{F}$, we calculate the partial derivative of $\hat{\eta}$ with respect to $q_j^{CP_k}$, i.e., the linear coefficient of $q_j^{CP_k}$, denoted as $z(q_j^{CP_k})$ and store it in a set $\mathcal{Z}$. Then, we perform as follows until $\mathcal{Z}$ is an empty set. Specifically, we select the smallest $z(q_j^{CP_k})$ from $\mathcal{Z}$ and denote it as $\hat{z}(q_j^{CP_k})$. If $\hat{z}(q_j^{CP_k}) < 0$ and the cache capacity of CP_k has not been fully utilized, i.e., $C_k - V_k > 0$, $q_j^{CP_k}$ is the minimum of $\frac{1}{S}$ and $\frac{C_k - V_k}{s_j}$. Otherwise, CP_k will not store encoded packets of content F_j, i.e., $q_j^{CP_k} = 0$. Finally, the algorithm outputs a caching policy $\mathbf{q}$.

Algorithm 1: Delay-Aware Secure MDS Coded Caching Algorithm

```
1: Input: $\mathcal{CP}$, $\mathcal{CR}$, $\mathbf{s}=[s_1,\cdots,s_j,\cdots,s_J]$, $\mathbf{C}=[C_1,\cdots,C_k,\cdots,C_K]$;
2: Initialization: $\mathcal{Z}=\varnothing$, $\mathbf{V}=[V_1,\cdots,V_k,\cdots V_K]=\mathbf{0}$;
3: while $CP_k \in \mathcal{CP}$ do
4:   while $F_j \in \mathcal{F}$ do
5:     Calculate $z(q_j^{CP_k})=\frac{\partial \hat{\eta}}{\partial q_j^{CP_k}}=\frac{1}{I}\sum_{i:CP_k\in\mathcal{CP}^i} p_j(\frac{1}{R_{CR_i}^{CP_k}}-\frac{1}{R_{CR_i}^{BS}})s_j$;
6:     $\mathcal{Z}=\mathcal{Z}\bigcup\{z(q_j^{CP_k})\}$;
7:   end while
8: end while
9: while $\mathcal{Z}\neq\varnothing$ do
10:   Search for the smallest $\hat{z}(q_j^{CP_k})$ in $\mathcal{Z}$;
11:   if $\hat{z}(q_j^{CP_k})<0$ and $C_k-V_k>0$
12:     $q_j^{CP_k}=\min(\frac{1}{S},\frac{C_k-V_k}{s_j})$;
13:   else
14:     $q_j^{CP_k}=0$;
15:   end if
16:   $V_k=V_k+q_j^{CP_k}s_j$, $\mathcal{Z}=\mathcal{Z}\setminus\{\hat{z}(q_j^{CP_k})\}$;
17: end while
18: Output: $\mathbf{q}$.
```

Remark 1. (Implementation) As shown in **Fig. 2** in the next page, we define time duration as a period of time, in which types of contents and popularity of contents remain unchanged. The length of time duration can be obtained through historical information. At the beginning of each time duration, the BS downloads and updates contents from data center, and codes them into multiple encoded packets by utilizing MDS coding. Then, the remaining part of time duration is divided into multiple time slots. In general, the corresponding frequency is a predetermined value. As a result, we focus on one time slot. In order to implement Algorithm 1, the signaling mechanism in [30] is exploited for each $CR_i \in \mathcal{CR}$ to broadcast pilots and report to the BS its location information at the beginning of each time slot. In this way, H_{BS,CR_i} and $H_{CP_k,CR_i}, \forall CP_k \in \mathcal{CP}^i$ can be estimated by the BS and each $CP_k \in \mathcal{CP}^i$. Then, considering the limited computing resources of CPs, the BS will carry out preliminary calculations to obtain the achievable rate between each $CR_i \in \mathcal{CR}$ and the BS, i.e., $R_{CR_i}^{BS}$, which is sent to the corresponding $CP_k, \forall CP_k \in \mathcal{CP}^i$, along with location of CR_i. As thus, each $CP_k \in \mathcal{CP}$ is able to perform Algorithm 1 (lines 2-18) based on the received information. After an optimal caching strategy is obtained by each CP, it downloads and updates the encoded packets from the BS. Note that contents sizes and popularity of contents are sent to each CP only in the first time slot of each time duration and kept during the whole time duration. Channel and location information is updated periodically by the given length of time slot. Once the time is up, the most outdated information is replaced by the newest. In this way, Algorithm 1 is a BS-assisted distributed algorithm and always keeps a history of recent values with low storage overhead.

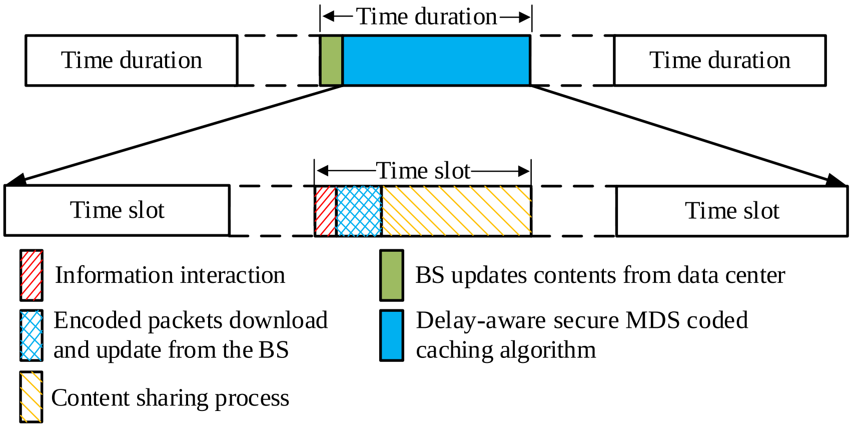

Time duration
Time duration
Time duration
Time slot
Time slot
Time slot
Information interaction
BS updates contents from data center
Encoded packets download and update from the BS
Delay-aware secure MDS coded caching algorithm
Content sharing process

Fig. 2. Schematic of the whole content sharing process

Remark 2. (Complexity) With K content providers and J contents, the complexity of lines 3-8 is $O(KJ)$. There is totally $K \times J$ elements in the set $\mathcal{Z}$, thus the complexity of lines 9-17 is $O(KJ)$. Hence, the complexity of Algorithm 1 is $O(KJ)$.

Table 2. Simulation parameters

Parameters	**Values**
Number of contents J	500
Cache capacity C_k	16GB
Skewness of Zipf distribution α	0.8
Distance between adjacent CPs d	20m
User density ρ	0.5users/m^2
D2D communication range r	20m
Transmission power of BS P_{BS}	36dBm
Transmission power of CPs P_{CP_k}	17dBm
Content size s_j	[0, 500]MB
Noise power spectrum density σ^2	-174dBm/Hz
Channel bandwidth B	10MHz

5. Numerical Results

In this section, we provide extensive numerical results to evaluate the performance of our proposed algorithm and illustrate the relationship between some vital parameters and the average system delay η. In order to evaluate the proposed delay-aware secure MDS coded caching scheme objectively, we compare it with other related literatures. All the schemes considered and simulated in this work are summarized and listed as below.

i) Our proposed scheme: Contents are coded into multiple encoded packets by utilizing MDS coding. Then encoded packets are stored by CPs according to our proposed delay-aware secure MDS coded caching algorithm.

ii) Uncoded caching scheme in [24]: Contents are stored by CPs without being coded. This scheme cannot prevent eavesdropping due to the lack of MDS coded caching.

iii) Random coded caching scheme in [31-33]: Encoded packets of contents generated by MDS coding are randomly stored by CPs.

iv) Popularity based coded caching scheme in [34]: Encoded packets are stored by CPs according to the popularity of contents.

v) Uniform coded caching scheme in [35]: This scheme performs encoded packets caching with uniform cache size allocation to each content.

Note that the uncoded caching scheme cannot prevent eavesdropping due to the lack of MDS coded caching. Besides, in order to resist the eavesdropping, we propose secure constraints based on the properties of MDS coding in the four kinds of coded caching schemes, i.e., our proposed scheme, random coded caching scheme, popularity based coded caching scheme, and uniform coded caching scheme. As a result, CRs cannot obtain enough encoded packets to recover their desired contents from their nearby CPs and they have to ask the BS for the remaining encoded packets in the four kinds of coded caching schemes.

The simulation setups are listed as follows. Considering a circular cell with a radius of 100m, the BS is located in the center with a library of N contents whose sizes are generated randomly. Besides, user devices are also randomly distributed, with a user density ρ. Specifically, the related simulation parameter settings are shown in **Table 2** without special declaration. Also, we exploit the channel model as $H = c^{-\varepsilon}\left|h\right|^2$, where c is the transmission distance between the transmitter and the receiver, h is the unitary power, Rayleigh fading channel coefficient, and $\varepsilon = 4$ is the path loss exponent. Moreover, due to the randomness of both user devices locations and contents sizes, we have run hundreds of simulations to obtain accurate statistical average values.

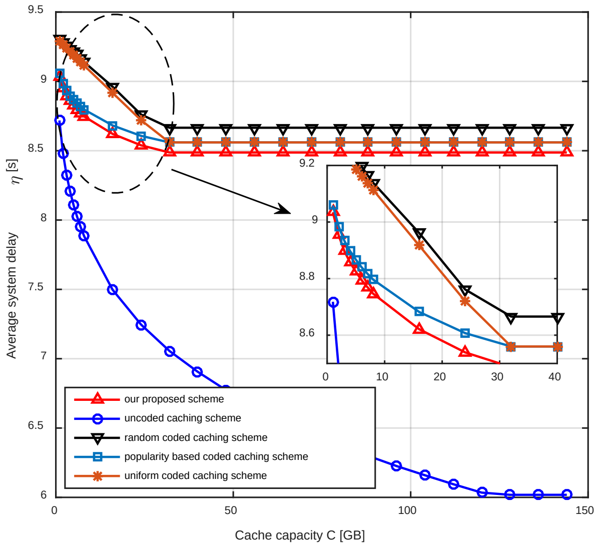


Fig. 3. The comparison of average system delay versus cache capacity

We plot **Fig. 3-6** to illustrate the comparison of average system delay versus cache capacity, number of contents and Zipf distribution skewness, respectively. As shown in the figures, our proposed algorithm always outperforms the other three coded caching scheme, i.e., random coded caching scheme, popularity coded caching scheme and uniform coded caching scheme, which verifies the correctness and effectiveness of our proposed algorithm. However, our proposed scheme performs worse than uncoded caching scheme because secure constraints restrict CPs from storing more encoded packets, thus CRs have to obtain some encoded packets to recover their desired contents from the BS rather than their nearby CPs. Meanwhile, random coded caching scheme always performs the worst because of the underutilized cache capacity.

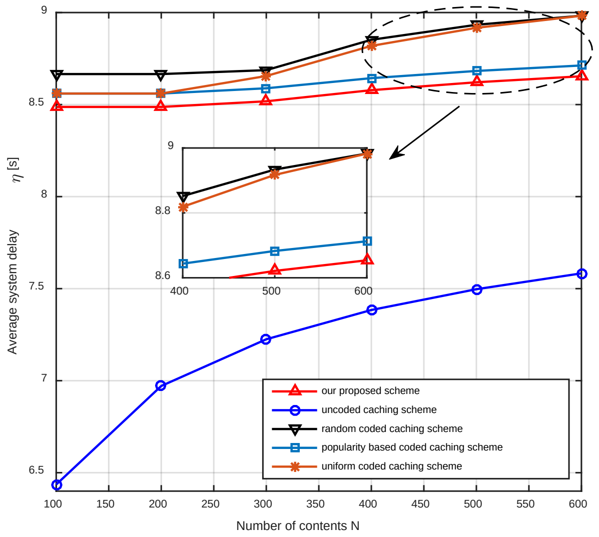


Fig. 4. The comparison of average system delay versus number of contents

Fig. 3 illustrates the relationship between caching capacity and average system delay. Specifically, when cache capacity is small, the average system delay decreases with cache capacity increasing in all the schemes. It is due to the fact that with the growing cache capacity, CPs have more space to store encoded packets so that CRs are able to receive more encoded packets from their nearby CPs, which leads to the decreasing average system delay. At this point, cache capacity is the main factor that restricts CPs from storing more encoded packets. As cache capacity grows bigger, secure constraints become the main restraining factor instead. In order to guarantee the security of content sharing, CPs are constrained from storing more encoded packets even their cache capacities are underutilized. In other word, $\forall CP_k \in \mathcal{CP}$ are constrained from storing more than certain number of encoded packets about content $F_j, \forall F_j \in \mathcal{F}$, in mathematics shown as $m_j^{CP_k} < n_j / S, \forall CP_k \in \mathcal{CP}, F_j \in \mathcal{F}$. Thus, average system delay in all the coded caching schemes finally come to the fixed values. In addition, average system delay in the uncoded caching scheme is eventually reduced to a fixed value when cache capacity is large enough, which indicates that all the contents have been stored by CPs. Besides, a noticeable phenomenon is that the average system delay in uniform coded caching scheme approximates that in random coded caching scheme at the beginning, which finally comes to the same fixed value as in popularity based coded caching scheme.

Fig. 4 shows the average system delay versus the number of contents. As shown in the figure, average system delay remains stable when the number of contents is small. This is due to the fact that when the number of contents is small, cache capacities of CPs are large enough to store all the encoded packets while secure constraints are satisfied. In other words, secure constraints, rather than cache capacity, are the main restraining factors to restrict CPs from storing more encoded packets. Thus, increasing number of contents has no influence on average system delay when it is small. Then, with the increasing number of contents, average system delay in all the schemes increases. At this point, the limited cache capacities of CPs become the main restraining factors. Due to the fact that CPs do not have large enough cache capacities to store all the encoded packets under the condition of satisfactory secure constraints, CRs have to obtain the remaining encoded packets from the BS to recover their desired contents. Thus, average system delay increases with the growing number of contents growing. Moreover, average system delay in uniform coded caching scheme is same as that in popularity based coded caching scheme at first, and finally approximates that in random coded caching scheme with the increasing number of contents, which is consistent with **Fig. 3**.

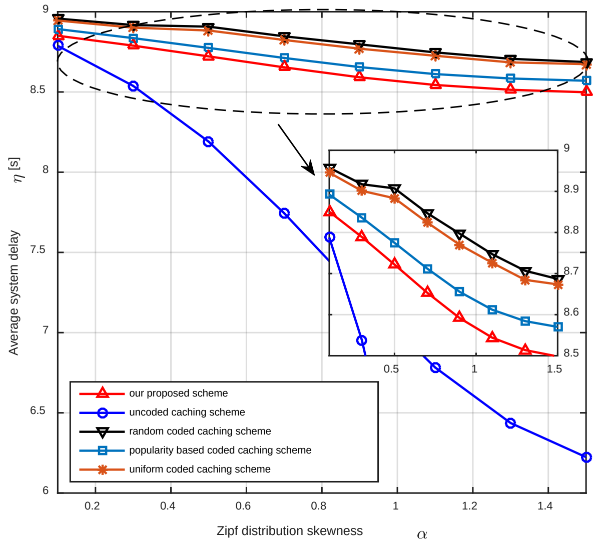


Fig. 5. Zipf distribution under different skewnesses α

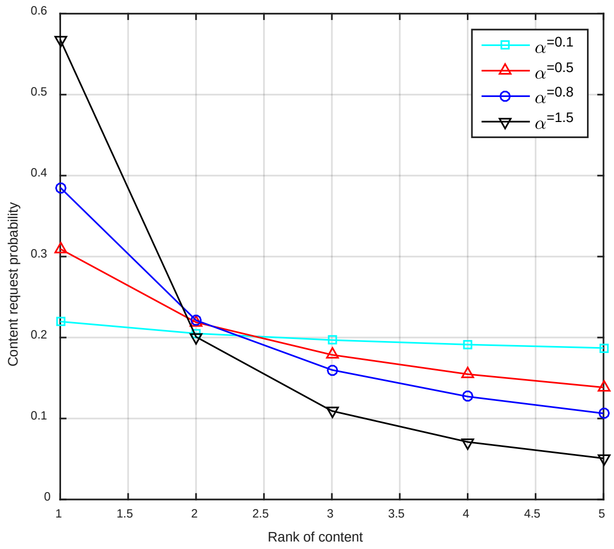


Fig. 6. The comparison of average system delay versus Zipf distribution skewness α

In **Fig. 5**, we investigate the performance of average system delay versus Zipf distribution skewness. Specifically, we plot **Fig. 6** to show the content request probability under different skewnesses. It can be easily seen from the figure that with the distribution skewness increasing, content requests are more concentrated, i.e., more content requests focus on less contents. As shown in **Fig. 5**, average system delay decreases with the increasing Zipf distribution skewness. This is due to the fact that CPs are able to exploit their limited cache capacities to store the contents that are requested more frequently. Thus, average system delay decreases with growing Zipf distribution skewness.

6. Conclusions

In this paper, a delay-aware secure MDS coded caching scheme is proposed to guard against the eavesdroppers randomly distributed in a D2D content sharing scenario. Specifically, by combining the MDS coding with encoded packets distributed caching, the average system delay is defined to illustrate the coupling relationship of the delay-content awareness. Then, we put forward the secure constraints to ensure the security of content sharing even in the worst case. Correspondingly, the delay-aware secure MDS coded caching problem is formulated as a constrained optimization problem to minimize the average system delay with the secure constraints. To obtain an optimal caching policy, we simplify the problem to its convex relaxation and develop a caching algorithm. Extensive numerical results are provided to verify the effectiveness and excellent performance of our proposed algorithm. Energy efficiency is another important performance metric in D2D content sharing scenario due to the fact that smart devices are always energy constrained. Actually, there is a tradeoff between delay and energy efficiency, which will be our future focus.

References

[1] D. Wu, L. Zhou, and Y. Cai, "Social-Aware Rate Based Content Sharing Mode Selection for D2D Content Sharing Scenarios," *IEEE Transactions on Multimedia*, vol. 19, no. 11, pp. 2571-2582, Nov. 2017. Article (CrossRef Link)

[2] D. Wu, L. Zhou, Y. Cai, and Y. Qian, "Optimal Content Sharing Mode Selection for Social-Aware D2D Communications," *IEEE Wireless Communications Letters*, vol. 7, no. 6, pp. 910-913, Dec. 2018. Article (CrossRef Link)

[3] J. Qu, L. Zhou, G. Zhang, D. Wu, J. Zheng, and Y. Cai, "Secure Caching in D2D Content Sharing," in *Proc. of 2018 IEEE International Conference on Communications Workshops*, pp. 1-6, Kansas City, MO, 2018. Article (CrossRef Link)

[4] R. Wang, X. Peng, J. Zhang, and K. B. Letaief, "Mobility-aware caching for content-centric wireless networks: modeling and methodology," *IEEE Communications Magazine*, vol. 54, no. 8, pp. 77-83, Aug. 2016. Article (CrossRef Link)

[5] M. Sheng, C. Xu, J. Liu, J. Song, X. Ma, and J. Li, "Enhancement for content delivery with proximity communications in caching enabled wireless networks: architecture and challenges," *IEEE Communications Magazine*, vol. 54, no. 8, pp. 70-76, Aug. 2016. Article (CrossRef Link)

[6] D. Wu, L. Zhou, Y. Cai, and Y. Qian, "Collaborative Caching and Matching for D2D Content Sharing," *IEEE Wireless Communications*, vol. 25, no. 3, pp. 43-49, June 2018. Article (CrossRef Link)

[7] M. Ahmed, Y. Li, M. Waqas, M. Sheraz, D. Jin, and Z. Han, "A Survey on Socially Aware Device-to-Device Communications," *IEEE Communications Surveys & Tutorials*, vol. 20, no. 3, pp. 2169-2197, 3rd Quart., 2018. Article (CrossRef Link)

[8] M. Haus, M. Waqas, A. Y. Ding, Y. Li, S. Tarkoma, and J. Ott, "Security and Privacy in Device-to-Device (D2D) Communication: A Review," *IEEE Communications Surveys & Tutorials*, vol. 19, no. 2, pp. 1054-1079, 2nd Quart., 2017. Article (CrossRef Link)

[9] A. Zhang and X. Lin, "Security-Aware and Privacy-Preserving D2D Communications in 5G," *IEEE Network*, vol. 31, no. 4, pp. 70-77, 2017. Article (CrossRef Link)

[10] C. E. Shannon, "Communication theory of secrecy systems," *The Bell System Technical Journal*, vol. 28, no. 4, pp. 656-715, 1949. Article (CrossRef Link)

[11] M. Ahmed, H. Shi, X. Chen, Y. Li, M. Waqas, and D. Jin, "Socially Aware Secrecy-Ensured Resource Allocation in D2D Underlay Communication: An Overlapping Coalitional Game Scheme," *IEEE Transactions on Wireless Communications*, vol. 17, no. 6, pp. 4118-4133, June 2018. Article (CrossRef Link)

[12] L. Wang, J. Liu, M. Chen, G. Gui, and H. Sari, "Optimization-Based Access Assignment Scheme for Physical-Layer Security in D2D Communications Underlaying a Cellular Network," *IEEE Transactions on Vehicular Technology*, vol. 67, no. 7, pp. 5766-5777, July 2018. Article (CrossRef Link)

[13] J. Wang, Y. Huang, S. Jin, R. Schober, X. You, and C. Zhao, "Resource Management for Device-to-Device Communication: A Physical Layer Security Perspective," *IEEE Journal on Selected Areas in Communications*, vol. 36, no. 4, pp. 946-960, Apr. 2018. Article (CrossRef Link)

[14] F. Gabry, V. Bioglio, and I. Land, "On edge caching with secrecy constraints," in *Proc. of 2016 IEEE International Conference on Communications, Kuala Lumpur*, pp. 1-6, 2016. Article (CrossRef Link)

[15] V. Bioglio, F. Gabry, and I. Land, "Optimizing MDS Codes for Caching at the Edge," in *Proc. of 2015 IEEE Global Communications Conference, San Diego, CA*, pp. 1-6, 2015. Article (CrossRef Link)

[16] L. Wang, H. Wu, Y. Ding, W. Chen, and H. V. Poor, "Hypergraph-Based Wireless Distributed Storage Optimization for Cellular D2D Underlays," *IEEE Journal on Selected Areas in Communications*, vol. 34, no. 10, pp. 2650-2666, Oct. 2016. Article (CrossRef Link)

[17] F. Gabry, V. Bioglio, and I. Land, "On edge caching in the presence of malicious users," in *Proc. of 2016 IEEE International Conference on Communications Workshops, Kuala Lumpur*, pp. 278-283, 2016. Article (CrossRef Link)

[18] D. Wu, L. Zhou, Y. Cai, H. Chao, and Y. Qian, "Physical–Social-Aware D2D Content Sharing Networks: A Provider–Demander Matching Game," *IEEE Transactions on Vehicular Technology*, vol. 67, no. 8, pp. 7538-7549, Aug. 2018. Article (CrossRef Link)

[19] M. Baldi, M. Bianchi, N. Maturo, and F. Chiaraluce, "A Physical Layer Secured Key Distribution Technique for IEEE 802.11g Wireless Networks," *IEEE Wireless Communications Letters*, vol. 2, no. 2, pp. 183-186, Apr. 2013. Article (CrossRef Link)

[20] T. T. Tran and H. Y. Kong, "CSI-Secured Orthogonal Jamming Method for Wireless Physical Layer Security," *IEEE Communications Letters*, vol. 18, no. 5, pp. 841-844, May 2014. Article (CrossRef Link)

[21] Y. Liu, H. Chen, and L. Wang, "Physical Layer Security for Next Generation Wireless Networks: Theories, Technologies, and Challenges," *IEEE Communications Surveys & Tutorials*, vol. 19, no. 1, pp. 347-376, 1st Quart., 2017. Article (CrossRef Link)

[22] L. Zhou, D. Wu, B. Zheng, and M. Guizani, "Joint physical-application layer security for wireless multimedia delivery," *IEEE Communications Magazine*, vol. 52, no. 3, pp. 66-72, Mar. 2014. Article (CrossRef Link)

[23] Y. Wu, D. Wu, L. Yang, and S. Xu, "Incentive-based cluster formation for D2D multicast content sharing," in *Proc. of 2018 Asia-Pacific Conference on Communications, Ningbo, China*, pp. 125-130, 2018. Article (CrossRef Link)

[24] J. Liao, K. Wong, Y. Zhang, Z. Zheng, and K. Yang, "MDS Coded Cooperative Caching for Heterogeneous Small Cell Networks," in *Proc. of 2017 IEEE Global Communications Conference, Singapore*, pp. 1-7, 2017. Article (CrossRef Link)

[25] W. Zhang, W. He, D. Wu, and Y. Cai, "Joint Mode Selection, Link Allocation and Power Control in Underlaying D2D Communication," *KSII Transactions on Internet and Information Systems*, vol. 10, no. 11, pp. 5209-5228, Nov. 2016.

[26] L. Sun, Q. Du, P. Ren, and Y. Wang, "Two Birds With One Stone: Towards Secure and Interference-Free D2D Transmissions via Constellation Rotation," *IEEE Transactions on Vehicular Technology*, vol. 65, no. 10, pp. 8767-8774, Oct. 2016. Article (CrossRef Link)

[27] L. Sun and H. Xu, "Fountain-Coding-Based Secure Communications Exploiting Outage Prediction and Limited Feedback," *IEEE Transactions on Vehicular Technology*, vol. 68, no. 1, pp. 740-753, Jan. 2019. Article (CrossRef Link)

[28] M. Ji, G. Caire, and A. F. Molisch, "Wireless Device-to-Device Caching Networks: Basic Principles and System Performance," *IEEE Journal on Selected Areas in Communications*, vol. 34, no. 1, pp. 176-189, Jan. 2016. Article (CrossRef Link)

[29] S. Boyd and L. Vandenberghe, *Convex Optimization*, Cambridge University Press, 2004. Article (CrossRef Link)

[30] B. S. C. Choi and M. Gerla, "Wireless Interrupt: Inter-Device Signaling in Next Generation Wireless Networks," in *Proc. of 2010 IEEE Conference on Computer Communications Workshops, San Diego, CA*, pp. 1-5, 2010. Article (CrossRef Link)

[31] M. Ji, A. M. Tulino, J. Llorca, and G. Caire, "On the average performance of caching and coded multicasting with random demands," in *Proc. of 2014 International Symposium on Wireless Communications Systems, Barcelona*, pp. 922-926, 2014. Article (CrossRef Link)

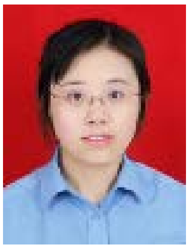
[32] S. A. Saberali, L. Lampe, and I. F. Blake, "Decentralized Coded Caching Without File Splitting," *IEEE Transactions on Wireless Communications*, vol. 18, no. 2, pp. 1289-1303, Feb. 2019. Article (CrossRef Link)

[33] J. Pedersen, A. G. i. Amat, I. Andriyanova, and F. Brännström, "Optimizing MDS Coded Caching in Wireless Networks With Device-to-Device Communication," *IEEE Transactions on Wireless Communications*, vol. 18, no. 1, pp. 286-295, Jan. 2019. Article (CrossRef Link)

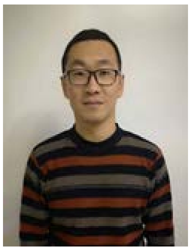
[34] T. D. Tran, T. D. Hoang, and L. B. Le, "Caching for Heterogeneous Small-Cell Networks With Bandwidth Allocation and Caching-Aware BS Association," *IEEE Wireless Communications Letters*, vol. 8, no. 1, pp. 49-52, Feb. 2019. Article (CrossRef Link)

[35] J. Liao, K. Wong, Y. Zhang, Z. Zheng, and K. Yang, "Coding, Multicast, and Cooperation for Cache- Enabled Heterogeneous Small Cell Networks," *IEEE Transactions on Wireless Communications*, vol. 16, no. 10, pp. 6838-6853, Oct. 2017. Article (CrossRef Link)

Xin Shi received the B.S. degree in electronic engineering from Peking University, China, in 2017. He is currently pursuing the M.S. degree with College of Communications Engineering, Army Engineering University of PLA, Nanjing, China. His current research interests include D2D communications, resource management, content security, and game theory.



Dan Wu received the B.S., M.S., and Ph.D. degrees from the Institute of Communications Engineering, PLA University of Science and Technology, Nanjing, China, in 2006, 2009, and 2012, respectively. She is currently an Associate Professor with the Army Engineering University of PLA, Nanjing, China. Her research interests are mainly in resource allocation and management, game theory, cooperative communications, and wireless sensor networks.



Meng Wang received the B.S. and M.S. degrees in instructional technology from the Nanjing Normal University, Nanjing, China, in 2006 and 2009, respectively. His current interests include wireless network security, D2D communications, and game theory.

Lianxin Yang received her B.S. and M.S. degrees from PLA University of Science and Technology, Nanjing, China, in 2015 and 2017, respectively. She is currently a Ph.D candidate in Army Engineering University of PLA, Nanjing, China. Her current research interests include social-aware D2D communications, resource management, game theory and user clustering.

Yan Wu received the B.S. degree from Nanjing University of Posts and Telecommunications, Nanjing, China, in 2017. He is currently pursuing the M.S. degree with the Army Engineering University of PLA, Nanjing, China. His current research interests include D2D communications, content sharing, game theory, and user clustering.