Secure Cache-Aided Multi-Relay Networks in the Presence of Multiple Eavesdroppers

Junjuan Xia, Lisheng Fan, Wei Xu, Xianfu Lei, Xiang Chen, George K. Karagiannidis, *Fellow, IEEE*, and Arumugam Nallanathan, *Fellow, IEEE*

***Abstract*—In this paper, we investigate the security of a cache-aided multi-relay communication network in the presence of multiple eavesdroppers, where each relay can pre-store a part of the requested files in order to assist secure data transmission from source to destination. If the relays have cached the requested file, then they can directly send it to the destination; otherwise, traditional dual-hop data transmission is used. For both cases, relay selection is performed to assist the secure data transmission. We analyze the network secrecy performance in both scenarios of *non-colluding* and *colluding* eavesdroppers, and obtain a closed-form expression for the average secrecy outage probability (SOP), as well as an asymptotic expression for the high main-to-eavesdropper ratio (MER). Through minimizing the network SOP, we further optimize the cache placement by proposing a stochastic sampling based cache learning (SacLe) strategy, which can be implemented in parallel and thus reduces the implementation latency substantially. Numerical and simulation results are finally presented to verify the proposed analysis, and show that the caching strategy has a significant impact on the network secrecy performance through affecting the caching diversity gain and signal cooperation gain at the relays. The proposed SacLe strategy is shown to be able to achieve the optimal performance obtained by the brute force (BF) algorithm.**

***Index Terms*—Secure communication, relay selection, cache, secrecy diversity order, cache placement.**

I. Introduction

Because of the broadcasting nature, wireless transmission may be wiretapped by non-intended receivers, which brings out the severe problem of information leakage. Thus, it is very important to safeguard the wireless networks, from physical to application layers [1], [2]. Compared with complicated encryption and decryption algorithms, physical-layer security (PLS) has less complexity and it is easy to implement. Thus, it can serve as a very good complement to the security of the application layer [3]. PLS can be backdated to the Shannon's work [4], and the classic wiretap model, proposed by Wyner in [5]. Based on this model, several researchers studied the security of wireless networks through several perspectives, including secrecy data rate as well as secrecy outage probability (SOP).

A. Literature

To improve the PLS of wireless networks, several relaying techniques have been proposed. There are two essential relaying protocols, amplify-and-forward (AF) and decode-and-forward (DF) [6], [7]. On the other side, the security of relaying networks was extensively studied in the literature. For example, the authors in [8]–[10] evaluated the security of relaying systems, from the viewpoints of the secrecy data rate and SOP. For a multi-relay network, channel fluctuation among relays can be exploited to enhance security. In [11], the authors proposed several relay selection criteria for secure communication in networks with multiple AF relays, and they evaluated the secrecy performance by deriving analytical expressions for the SOP as well as an asymptotic formula for the high main-to-eavesdropper ratio (MER). Furthermore, for multiple DF relaying networks, the authors in [12] employed relay selection to enhance the channel quality of the main links and derived a closed-form formula for the network SOP. Besides the relay selection, beamforming can be also employed to improve network security, through strengthening the main links and weakening the eavesdropping links. In [13]–[15], the authors applied beamforming to enhance the security of multiple-input single-output (MISO) cognitive radio networks or nonregenerative multiple-input multiple-output (MIMO) relay systems. When smart attackers appear in the relaying network, learning methods, such as Q-learning, can be used to increase the communication security [16]–[18].

Recently, wireless caching has emerged as a promising technique in wireless networks, especially for applications related to mobile edge computing and internet of things (IoT) [19]–[21]. Caching can improve the network performance and the quality of user experience, by pre-storing popular files at the nodes nearby users [22]–[24]. There are two traditional caching strategies: *most popular content (MPC)*

J. Xia and L. Fan are both with the School of Computer Science, Guangzhou University, China (e-mail: {xiajunjuan,lsfan}@gzhu.edu.cn).

W. Xu is with the National Mobile Communications Research Laboratory, Southeast University, Nanjing 210096, China (e-mail: wxu@seu.edu.cn).

X. Lei is with the School of Information Science and Technology, Institute of Mobile Communications, Southwest Jiaotong University, Chengdu 610031, China (e-mail: xflei@home.swjtu.edu.cn).

X. Chen is with the School of Electronics and Information Technology, Sun Yat-sen University, Guangzhou, China (e-mail: chenxiang@mail.sysu.edu.cn).

G. K. Karagiannidis is with Aristotle University of Thessaloniki, Thessaloniki 54636, Greece (e-mail: geokarag@auth.gr).

A. Nallanathan is with the School of Electronic Engineering and Computer Science, Queen Mary University of London, London, U.K (e-mail: a.nallanathan@qmul.ac.uk)

Lisheng Fan and Xianfu Lei are corresponding author of this paper.

This work was supported in part by the NSFC under Grant 61871139/61871109, by the Guangdong Natural Science Funds for Distinguished Young Scholar under Grant 2014A030306027, by the Innovation Team Project of Guangdong Province University under Grant 2016KCXTD017, by the Science and Technology Program of Guangzhou under Grant 201807010103, by the State's Key Project of Research and Development Plan (No. 2017YFE0121300-6), by the Sichuan Science and Techology Program under Grant 2017HH0035, the NSFC project under Grant 61501382, and the Fundamental Research Funds for the Central Universities under Grant 2682018CX27.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TCOMM.2019.2935047, IEEE Transactions on Communications

and *largest content diversity (LCD)*, which achieves optimal signal cooperation gain and optimal caching gain, respectively. Based on the MPC and LCD, a hybrid caching strategy was proposed to optimize the cache placement for collaborative relaying networks, aiming at the balance between signal cooperation and caching gains [25]. The authors in [26] studied the impact of caching on the relay selection in multi-relay networks, and optimized the cache placement by relaxing the integer-constraint in the optimization problem. Besides the impact on the main links, wireless caching also has a significant impact on the physical-layer security of the wireless networks. In this direction, the authors in [27], [28] proposed a novel wireless caching scheme to enhance the security of the backhaul-limited cellular networks and heterogeneous small cell networks, and pointed out that both the signal transmission and caching strategy should be optimized to guarantee the network security. The authors in [29] further studied the security of cache-aided relaying networks by cluttering the relays and performing the maximal ratio transmission among cluttered relays, and then employed a hybrid caching strategy based on LCD and MPC. In [21], Zhao et al. proposed a fundamental framework to guarantee the the security for UAV assisted hyper-dense networks via caching. However, although the aforementioned works have studied the impact of caching on wireless networks and security, there has been little research on the impact of caching on the security of relay selection for relaying networks.

B. Contribution

In this paper, we investigate the security of a multi-DF relaying network in the presence of multiple eavesdroppers, where each relay is equipped with a cache to assist the secure transmission. We start with the critical question: "*How cache affects the secure communication of a multi-relay network?*". To answer this question, we consider two cache status, where the requested file is pre-stored or not in the relay. If the relays have cached the file, then they can directly send it to the destination; otherwise, the destination has to fetch the file from the source through a dual-hop link. For the two cases, we then select the best relay to assist data transmission, based on the second-hop and dual-hop of the main relaying links, respectively. Assuming a wiretap scenario of either *non-colluding* or *colluding* eavesdroppers, we characterize the network secrecy performance by deriving a closed-form expression for the SOP as well as an asymptotic SOP in the high regime of MER.

Another critical question is: "*How to optimize the cache placement for secure communication of a cache-aided multi-relay network?*". To tackle this problem, we try to optimize the cache placement among relays by minimizing the network SOP. Since this optimization problem is integer-constraint, it is in general hard to obtain an analytical solution. Thus, we propose a stochastic sampling based cache learning (SacLe) strategy based on file popularity, which can be implemented in parallel. The proposed SacLe can achieve almost the optimal secrecy performance, and it surpasses the traditional LCD and MPC strategies. Simulations and numerical results are finally demonstrated to confirm the presented analysis.

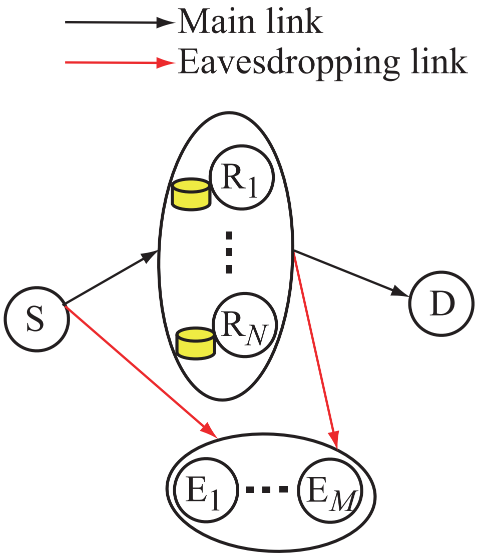


Fig. 1. Cache-aided multi-DF relaying network with multiple eavesdroppers.

C. Structure

The remaining parts of this paper are organized as follows. Section II presents the system model and describes the secure data transmission along with the relay selection criterion. Section III provides a closed-form expression of SOP and an asymptotic formula, under the wiretap scenario of either non-colluding or colluding eavesdroppers. Section IV studies the cache placement optimization and proposes the SacLe strategy. Numerical and simulation results are provided and discussed in Section V, and we conclude this work in Section VI.

D. Notations

Let $\mathcal{CN}(0, \sigma^2)$ be a random variable (RV) with zero mean and variance σ^2, subject to circularly symmetric complex Gaussian. Let $f_X(\cdot)$ and $F_X(\cdot)$ denote the probability density function (PDF) and cumulative density function (CDF) of the RV X, respectively. The function, $\Gamma(x)$, is the Gamma function [30], and the operation $\Pr(\cdot)$ returns probability.

II. System Model

In Fig. 1, the system model of a cache-aided multi-relay network is depicted in the presence of multiple eavesdroppers, where the N DF relays[1] $\{R_n | 1 \le n \le N\}$ assist the secure data transmission from source S to legitimate destination D. Each relay is equipped with a cache of a finite size C, and it can pre-store a part of requested files from the S. There are K files in total requested by the destination D, and the file popularity follows by Zipf distribution [25]. For the k-th file, its popularity is characterized by

$$\mu_k = \frac{k^{-\eta}}{\sum_{k_1=1}^{K} k_1^{-\eta}}, \tag{1}$$

[1] Compared with the AF and other relaying protocols, DF relaying protocol does not amplify the received noise at the relays, and it can provide a better reliability performance. Hence, we adopt the DF relaying protocol in this paper.

where $\eta > 0$ is the popularity factor. Note that a larger value of η leads to a more concentrated popularity. If the requested file falls in the cache of relays, then the relays can directly send the file to D; otherwise, the destination D has to fetch the file from the source S, through the two-hop relaying links. It is assumed that there are M eavesdroppers [2] $\{E_m | 1 \leq m \leq M\}$ in the network, which can be untrusted users who can overhear confidential messages or files from the source. An example of the confidential content is the video files with some popularity profile [27], [28], [35]. By analyzing the wiretapped video files, the eavesdroppers may obtain some critical private information, such as video access history and subscription details [3]. Hence, it is crucial to safeguard the transmission of confidential content in the physical-layer. Note that multiple users can exist in the wireless networks, and when some of users act as the eavesdroppers to overhear the confidential messages, multiple eavesdroppers appear. Moreover, it is a general case to consider multiple eavesdroppers in the networks, and the scenario of a single eavesdropper can be viewed as a special case of multiple eavesdroppers, by setting the number of eavesdroppers, M, to one. In this paper, the direct link from the source to the destination D does not exist due to the severe shadowing, while the direct links from the source to eavesdroppers exist, and hence, the M eavesdroppers can overhear the message from both the source and relays [40]. This wiretap model can be viewed as a worse-case eavesdropping scenario, which can serve as a reference for scenarios without direct eavesdropping links. Moreover, the N relays form a cluster, which has the same distance from the other nodes in the network, and this assumption also holds for the M eavesdroppers. In addition, each node has a single antenna, due to size limitation.

Let Ω_k denote a set of relays, which have cached the requested k-th file, and $g_k = |\Omega_k| \in \{0, \cdots, N\}$ is the cardinality of Ω_k. As the cache status affects the secure data transmission process, next, we discuss the two cases where the relays have cached the requested file or not.

A. $g_k \geq 1$

In this case, there is at least one relay which has pre-stored the requested k-th file, and the g_k relays can directly send the file to the destination D. Suppose that the n-th relay is selected from Ω_k to assist the data transmission. Then, the relay R_n transmits the normalized signal x_s with power P, while the signals received at D and E_m are,

$$y_{n,D} = \sqrt{P} h_{R_n,D} x_s + n_D, \tag{2}$$

$$y_{n,E_m} = \sqrt{P} h_{R_n,E_m} x_s + n_E, \tag{3}$$

where $h_{R_n,D} \sim \mathcal{CN}(0, \beta_1)$ and $h_{R_n,E_m} \sim \mathcal{CN}(0, \beta_2)$ are the channel parameters of the R_n–D and R_n–E_m links, respectively. The terms $n_D \sim \mathcal{CN}(0, \sigma^2)$ and $n_E \sim \mathcal{CN}(0, \sigma^2)$ are the additive white Gaussian noise (AWGN) at the D and eavesdroppers, respectively. From (2) and (3), the received SNRs at the D and E_m are

$$\text{SNR}_{n,D} = \tilde{P} v_{1n}, \tag{4}$$

$$\text{SNR}_{n,E_m} = \tilde{P} v_{2n,m}, \tag{5}$$

where $\tilde{P} = \frac{P}{\sigma^2}$ denotes the transmit SNR, and $v_{1n} = |h_{R_n,D}|^2$ and $v_{2n,m} = |h_{R_n,E_m}|^2$ are the instantaneous channel gains of the R_n–D and R_n–E_m links, respectively.

According to the wiretap model, the M eavesdroppers can work in a non-colluding way, if they cannot share the received information among them. In this case, each eavesdropper decodes the message individually. In contrast, the M eavesdroppers can work in a colluding way, if they can share the received information among them. In this case, the eavesdroppers can use maximal ratio combining (MRC) to jointly decode the message, which represents the worse-case eavesdropping scenario. This is the main reason that we consider both the non-colluding and colluding wiretap scenarios for the considered system. For M non-colluding and colluding eavesdroppers, the equivalent SNRs are respectively given by

$$\text{SNR}_{n,E} = \tilde{P} \max(v_{2n,1}, v_{2n,2}, \cdots, v_{2n,M}), \tag{6}$$

and

$$\text{SNR}_{n,E} = \tilde{P}(v_{2n,1} + v_{2n,2} + \cdots + v_{2n,M}). \tag{7}$$

A secrecy outage event occurs when the difference in the data rate between the main and eavesdropping links is below a given secrecy data rate R_s, i.e.,

$$\log_2(1 + \tilde{P} v_{1n}) - \log_2(1 + \text{SNR}_{n,E}) < R_s, \tag{8}$$

which is equivalent to

$$\frac{1 + \tilde{P} v_{1n}}{1 + \text{SNR}_{n,E}} < \gamma_{1s}, \tag{9}$$

where $\gamma_{1s} = 2^{R_s}$ is the secrecy SNR threshold with cache.

In practice, it is difficult to obtain the instantaneous channel gains of the eavesdropping links, especially when the eavesdroppers are passive. In this work, we perform the relay selection by using the main links only, and the best relay R_{n^*} is chosen according to

$$n^* = \arg \max_{n \in \Omega_k} v_{1n}. \tag{10}$$

To implement the above relay selection, the relays belonging to the set Ω_k firstly send some pilot signals to the destination

[2] The estimate of the number of eavesdroppers in the networks depends on the communication scenarios. In some scenarios where the eavesdroppers are a part of the network, the eavesdroppers can be legitimate users in other applications different from the current one. In this case, the channel information of the eavesdroppers can be obtained, from which the system can analyze and estimate the total number of eavesdroppers in the networks. The examples of known number of eavesdroppers can be found in the literature, such as the works [31]–[34]. Even in other scenarios where it is hard to accurately estimate the number of eavesdroppers, the work in this paper can still serve as an important reference for the secure communication of cache-aided relaying networks.

[3] Note that the video access history and subscription details are not cached in the relays. As described in the literature such as the works [36]–[39], the file popularity can be estimated in advance in the networks, and hence it can be viewed as prior information in the system. If the personal videos are not frequently requested, i.e., the file popularity is low, the files may still need to be cached when the transmission secrecy is very poor and becomes the bottleneck of the network average security. In this case, the usage of caching can help pre-store the files on the nodes nearby the desired users, and provide some additional spatial secrecy diversity, through which the secrecy performance is improved.

D, at the beginning of each transmission slot. Then, the destination D can estimate the required channel gains v_{1n} with the help of pilot signals. Based on the estimated channel gains v_{1n}, the destination D performs the selection according to the criterion in (10), and broadcasts the selection result through some dedicated feedback links to the relays. In this way, the relay selection process is completed.

B. $g_k = 0$

In this case, none of the relays has cached the requested k-th file, and the conventional dual-hop relaying is used for secure data transmission, assisted by N fixed-DF relays. Suppose that the n-th relay is chosen among N ones for data transmission. Let the transmit power at the source S be equal to P, and let $h_{S,R_n} \sim \mathcal{CN}(0,\alpha)$ and $h_{S,E_m} \sim \mathcal{CN}(0,\varepsilon)$ be the channel parameters of the S–R_n and S–E_m links, respectively. We use $u_n = |h_{S,R_n}|^2$ and $w_m = |h_{S,E_m}|^2$ to denote the instantaneous channel gains of the S–R_n and S–E_m links, respectively. When the M eavesdroppers are non-colluding, the secrecy outage event occurs if the data rate difference between the main and eavesdropping links is below R_s, [12], [41]

$$\begin{aligned}&\frac{1}{2}\log_2[1+\tilde{P}\min(u_n,v_{1n})]\\&-\frac{1}{2}\log_2[1+\tilde{P}\max_{1\le m\le M}(v_{2n,m}+w_m)] < R_s,\end{aligned} \tag{11}$$

where the factor $\frac{1}{2}$ comes from the two-phase transmission. From (11), it holds that

$$\frac{1+\tilde{P}\min(u_n,v_{1n})}{1+\tilde{P}\max_{1\le m\le M}(v_{2n,m}+w_m)} < \gamma_{2s}, \tag{12}$$

where $\gamma_{2s} = 2^{2R_s}$ is the secrecy SNR threshold without cache.

When the M eavesdroppers are colluding, the secrecy outage event occurs when the following equation holds, [12], [41]

$$\begin{aligned}&\frac{1}{2}\log_2[1+\tilde{P}\min(u_n,v_{1n})]\\&-\frac{1}{2}\log_2\left[1+\tilde{P}\sum_{m=1}^{M}\Big(v_{2n,m}+w_m\Big)\right] < R_s,\end{aligned} \tag{13}$$

which is equivalent to

$$\frac{1+\tilde{P}\min(u_n,v_{1n})}{1+\tilde{P}\sum_{m=1}^{M}(v_{2n,m}+w_m)} < \gamma_{2s}. \tag{14}$$

Using (11)-(14), we can perform relay selection according to the conventional max-min criterion,

$$n^* = \arg\max_{1\le n\le N}\min(u_n,v_{1n}), \tag{15}$$

which depends only on the dual-hop channels of the main links. To implement the above relay selection, the source S firstly sends some pilot signals to the destination D through the help of relays, at the beginning of each transmission slot. Then, the destination D can estimate the required channel gains u_n and v_{1n} with the help of pilot signals. Based on the estimated channel gains u_n and v_{1n}, the destination D performs the selection according to the criterion in (15), and broadcasts the selection result through some dedicated feedback links to the relays and the source. In this way, the relay selection process is completed.

III. Secrecy Outage Probability

In this section, we analyze the network secrecy performance with either non-colluding or colluding eavesdroppers, by deriving a closed-form expression for the SOP as well as an asymptotic expression in the high MER region. Let $p_r(g_k)$ denote the secrecy outage probability of transmitting the k-th file, which has been cached by the g_k relays. The network average SOP P_{out} for the total number of K files is

$$P_{out} = \sum_{k=1}^{K}\mu_k p_r(g_k). \tag{16}$$

According to the Zipf distribution in (1), the file popularity decreases with the file index k. Thus, only the first $K_1 (1 \le K_1 \le K)$ files can be cached at the relays, while the residual $(K-K_1)$ files cannot. Then, it holds that

$$P_{out} = \sum_{k=1}^{K_1}\mu_k p_r(g_k) + \sum_{k=K_1+1}^{K}\mu_k p_r(0). \tag{17}$$

Next, we derive closed-form and asymptotic expressions for the $p_r(g_k)$ in the cases of non-colluding and colluding eavesdroppers.

A. Non-colluding eavesdroppers

In the following theorem, we first provide a closed-form expression for $p_r(g_k)$ in the two cases: $g_k \ge 1$ and $g_k = 0$,

Theorem 1: A closed-form expression for $p_r(g_k)$ in the case of non-colluding eavesdroppers is shown in (18), where $\lambda = \frac{\beta_1}{\beta_2}$ denotes the MER, and

$$\begin{cases} b_1 = \dfrac{\gamma_{1s}-1}{\tilde{P}}, \quad b_2 = \dfrac{\gamma_{2s}-1}{\tilde{P}}, \\ \tau_{m_1m_2} = \dbinom{M}{m_1}\dbinom{m_1}{m_2}\dfrac{(-1)^{m_2-1}\varepsilon^{m_2}\beta_2^{m_1-m_2}}{(\varepsilon-\beta_2)^{m_1}}. \end{cases} . \tag{19}$$

Proof: *See Appendix A.*

By using the results of Theorem 1 into (17), we obtain a closed-form expression for the SOP of the cache-aided multi-DF relaying network with non-colluding eavesdroppers, as shown in (20). Note that (20) is composed of elementary functions only, and hence it is easily to be evaluated.

Next, we extend the previous analysis to provide an asymptotic expression for P_{out} in the high MER, in order to obtain more insights on the system design. We first study the asymptotic expression of $p_r(g_k)$. Specifically, the CDF of v_{1n}, $F_{v_{1n}}(x)$, can be asymptotically given by $\frac{x}{\beta_1}$ by using the approximation of $e^{-x} \simeq 1-x$ [30]. Then, from the theory of order statistics [42], the CDF of v_{1n^*} has the form of $(\frac{x}{\beta_1})^{g_k}$ with $g_k \ge 1$. Accordingly, the asymptotic $p_r(g_k)$ should be of the order of g_k with respect to $\frac{1}{\lambda}$, when $g_k \ge 1$, as there are g_k cache-aided relays which can assist the second-hop secure transmission. Similar analysis can be applied to the case of

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TCOMM.2019.2935047, IEEE Transactions on Communications

$$p_r(g_k) = \begin{cases} \sum_{m=1}^{M}\sum_{n=0}^{g_k}\binom{M}{m}\binom{g_k}{n}(-1)^{m+n-1}e^{-\frac{nb_1}{\beta_1}}\Big(1+\frac{n\gamma_{1s}}{m\lambda}\Big)^{-1}, \text{ If } g_k \geq 1, \\ \sum_{n=0}^{N}\sum_{m_1=0}^{M}\sum_{m_2=0}^{m_1}\binom{N}{n}(-1)^n\Big(\frac{m_1-m_2}{\beta_2}+\frac{m_2}{\varepsilon}\Big)\tau_{m_1m_2}e^{-(\frac{1}{\alpha}+\frac{1}{\beta_1})nb_2}\Big[\frac{m_1-m_2}{\beta_2}+\frac{m_2}{\varepsilon} \\ +n\gamma_{2s}(\frac{1}{\alpha}+\frac{1}{\beta_1})\Big]^{-1}, \qquad \text{If } g_k = 0. \end{cases} \tag{18}$$

$$P_{out} = \sum_{k=1}^{K_1}\sum_{m=1}^{M}\sum_{n=0}^{g_k}\binom{M}{m}\binom{g_k}{n}(-1)^{m+n-1}\mu_k e^{-\frac{nb_1}{\beta_1}}\Big(1+\frac{n\gamma_{1s}}{m\lambda}\Big)^{-1} + \sum_{n=0}^{N}\sum_{m_1=0}^{M}\sum_{m_2=0}^{m_1}\binom{N}{n}(-1)^n$$
$$\times\Big(\frac{m_1-m_2}{\beta_2}+\frac{m_2}{\varepsilon}\Big)\tau_{m_1m_2}e^{-(\frac{1}{\alpha}+\frac{1}{\beta_1})nb_2}\Big[\frac{m_1-m_2}{\beta_2}+\frac{m_2}{\varepsilon}+n\gamma_{2s}(\frac{1}{\alpha}+\frac{1}{\beta_1})\Big]^{-1}\Big(\sum_{k=K_1+1}^{K}\mu_k\Big). \tag{20}$$

$g_k = 0$. The specific expression of the asymptotic $p_r(g_k)$ is given by the following theorem,

Theorem 2: An asymptotic expression for $p_r(g_k)$ in the case of non-colluding eavesdroppers is

$$p_r(g_k) \simeq \begin{cases} \big(\frac{\gamma_{1s}}{\lambda}\big)^{g_k}\zeta_{1k}, & \text{If } g_k \geq 1, \\ \frac{\gamma_{2s}^N}{\lambda^N}\zeta_2, & \text{If } g_k = 0. \end{cases}, \tag{21}$$

with

$$\begin{cases} \zeta_{1k} = g_k!\sum_{m=1}^{M}\binom{M}{m}(-1)^{m-1}m^{-g_k}, \\ \zeta_2 = \big(1+\frac{\beta_1}{\alpha}\big)^N N!\Big(\sum_{m_1=0}^{M}\sum_{m_2=0}^{m_1}\tau_{m_1m_2} \\ \quad\times\big[m_1+(\frac{\beta_2}{\varepsilon}-1)m_2\big]^{-N}\Big). \end{cases} \tag{22}$$

Proof: *See Appendix B.*

By using the results of Theorem 2 into (17), we obtain an asymptotic expression for P_{out} for the cache-aided multi-DF-relay network in the presence of non-colluding eavesdroppers, as follows:

$$P_{out} \simeq \sum_{k=1}^{K_1}\mu_k\zeta_{1k}\big(\frac{\gamma_{1s}}{\lambda}\big)^{g_k} + \frac{\gamma_{2s}^N}{\lambda^N}\zeta_2\Big(\sum_{k=K_1+1}^{K}\mu_k\Big). \tag{23}$$

From (23), we can conclude the following insights for the system design:

Remark 1: From (21): When the relays have cached the requested k-th file, the secrecy diversity order of transmitting this file is equal to g_k, as the relay selection in (10) can exploit the g_k relays for secure data transmission.

Remark 2: From (21): When none of the relays has cached the requested k-th file, the secrecy diversity order is equal to N, indicating that in dual-hop relaying process, all of the relays can be fully exploited for the secure data transmission.

Remark 3: As ζ_{1k} and ζ_2 increase with M, the secrecy performance degrades with the increasing number of eavesdroppers.

Remark 4: From Remarks 1-3, we conclude that the secrecy diversity order of transmitting the K files is equal to $\min(g_1, g_2, \cdots, g_{K_1})$. This indicates that the minimum g_k among K_1 ones leads to the worst SOP, which becomes the bottleneck of the whole network secrecy performance.

Remark 5: The MPC strategy approaches to the optimal caching strategy in the high MER region, as the secrecy diversity order becomes the main factor that regulates the network secrecy performance.

B. Colluding eavesdroppers

Theorem 3: A closed-form expression for $p_r(g_k)$ in the case of colluding eavesdroppers is shown in (24), with

$$J(n,x) = \big[1+(\frac{x}{\alpha}+\frac{x}{\beta_1})n\gamma_{2s}\big]^{-M}. \tag{25}$$

Proof: *See Appendix C.*

By applying the results of Theorem 3 into (17), we can obtain a closed-form expression on the network SOP for the case of colluding eavesdroppers, as shown in (26), which consists of elementary functions only, and hence we can easily evaluate the network SOP performance.

Next, we provide an asymptotic formula for P_{out} of the cache-aided multi-DF-relay network in the presence of colluding eavesdroppers. The asymptotic expression of $p_r(g_k)$ with colluding eavesdroppers is given in the following theorem,

Theorem 4: A closed-form expression of $p_r(g_k)$ with colluding eavesdroppers is

$$p_r(g_k) \simeq \begin{cases} \big(\frac{\gamma_{1s}}{\lambda}\big)^{g_k}\zeta_{3k}, & \text{If } g_k \geq 1, \\ \frac{\gamma_{2s}^N}{\lambda^N}\zeta_4, & \text{If } g_k = 0. \end{cases}, \tag{27}$$

$$p_r(g_k) = \begin{cases} \sum_{n=0}^{g_k} \binom{g_k}{n} (-1)^n e^{-\frac{nb_1}{\beta_1}} \Big(1 + \frac{n\gamma_{1s}}{\lambda}\Big)^{-M}, & \text{If } g_k \geq 1, \\ \sum_{n=0}^{N} \binom{N}{n} (-1)^n e^{-(\frac{1}{\alpha}+\frac{1}{\beta_1})nb_2} J(n,\varepsilon) J(n,\beta_2). & \text{If } g_k = 0. \end{cases} \qquad (24)$$

$$P_{out} = \sum_{k=1}^{K_1} \sum_{n=0}^{g_k} \binom{g_k}{n} (-1)^n \mu_k e^{-\frac{nb_1}{\beta_1}} \Big(1 + \frac{n\gamma_{1s}}{\lambda}\Big)^{-M} + \sum_{n=0}^{N} \binom{N}{n} (-1)^n e^{-(\frac{1}{\alpha}+\frac{1}{\beta_1})nb_2} \times J(n,\varepsilon) J(n,\beta_2) \Big(\sum_{k=K_1+1}^{K} \mu_k \Big). \qquad (26)$$

with

$$\begin{cases} \zeta_{3k} = \frac{(M-1+g_k)!}{\Gamma(M)}, \\ \zeta_4 = \Big(1 + \frac{\beta_1}{\alpha}\Big)^N \sum_{n=0}^{N} \frac{(M+n-1)!(M+N-n+1)!}{\Gamma(M)\Gamma(M)} \\ \quad \times \binom{N}{n} \big(\frac{\varepsilon}{\beta_2}\big)^n. \end{cases} \qquad (28)$$

Proof: *See Appendix D.*

By applying the results of Theorem 4 into (17), we obtain an asymptotic expression of P_{out} for the cache-aided multi-DF-relay network in the presence of colluding eavesdroppers, as follow,

$$P_{out} \simeq \sum_{k=1}^{K_1} \mu_k \zeta_{3k} \big(\frac{\gamma_{1s}}{\lambda}\big)^{g_k} + \frac{\gamma_{2s}^N}{\lambda^N} \zeta_4 \Big(\sum_{k=K_1+1}^{K} \mu_k \Big). \qquad (29)$$

From (29), we can obtain the following insights on the system:

Remark 6: From (27), we can find that the secrecy diversity order of transmitting the k-th file is g_k, when there are g_k relays which have cached the requested file. In addition, from (27), we see that the secrecy diversity order of transmitting the k-th file becomes N, when none of the relays has cached the file.

Remark 7: From Remark 5, we conclude that the average secrecy diversity order of transmitting the total K files is equivalent to $\min(g_1, g_2, \cdots, g_{K_1})$. This indicates that the minimum g_k among K_1 ones leads to the worst SOP and hence becomes the bottleneck of the whole network secrecy performance.

Remark 8: The MPC strategy will converge to the optimal cache placement which achieves the secrecy diversity order of N, in the high MER regime.

Remark 9: As ζ_{3k} and ζ_4 become larger with increasing M, we conclude that the network secrecy performance deteriorates with more eavesdroppers. Moreover, as $\zeta_{3k} \geq \zeta_{1k}$ and $\zeta_4 \geq \zeta_2$ hold, the secrecy performance with colluding eavesdroppers is worse than that with non-colluding eavesdroppers. This is because that the cooperation among eavesdroppers can help strengthen the eavesdropping links.

IV. Cache Placement Optimization

As the caching parameters $\{g_k | 1 \leq k \leq K\}$ have a significant impact on the network secrecy performance, next, we investigate the cache placement strategy for the considered system. For the non-colluding or colluding eavesdroppers, we optimize the cache placement by minimizing P_{out} in (16) as,

$$\min_{\{g_k | 1 \leq k \leq K\}} \sum_{k=1}^{K} \mu_k p_r(g_k), \qquad (30)$$

$$\text{s.t.} \quad g_k \in \{0, \cdots, N\}, \forall k \in \{1, K\}, \qquad (31)$$

$$\sum_{k=1}^{K} g_k \leq NC. \qquad (32)$$

From the above two equations, we can find that the cache placement depends on both the file popularity and the average channel gains of the network links. Specifically, the most popular file has the largest probability to be cached at all relays, since its secrecy outage event imposes the most severe impact on the network secrecy outage probability. Moreover, the most popular file should be cached at the relay with the lowest SOP, when the relays have different SOPs of the secure transmission. In further, the average channel gain of the R_n–D link plays a more important role in optimizing the cache placement than that of the S–R_n link, since the former affects the SOP $p_r(g_k)$ with all values of g_k, while the latter affects the SOP $p_r(0)$ only.

Since the optimization problem in (30) is an integer-constraint one, it is in general hard to obtain an analytical solution for g_k. We can use some efficient software packages, such as Lingo, to solve this integer-constraint optimization problem [26], [43]. However, this is a brute-force (BF) solution, and requires exponential computational complexity with respect to N and K, which imposes a large implementation latency, in practice.

In order to reduce the implementation latency, we propose a stochastic sampling based cache learning (SacLe) method to solve the cache placement optimization problem in (30), which can be implemented in parallel. In this method, L stochastic seeds are generated, where each of them is an implementation of the caching sequence $[g_1, g_2, \cdots, g_K]$. Next, we discuss how to randomly generate a seed of $[g_1, g_2, \cdots, g_K]$. By considering that the popularity of the $(k+1)$-th file is smaller than that of the k-th file, the $(k+1)$-th file should have a lower priority to be cached, i.e., $g_{k+1} \le g_k$. For a given g_k, the value of g_{k+1} can be computed as

$$g_{k+1} = g_k - \rho_k, \tag{33}$$

where ρ_k is an integer, and it is randomly distributed in the range of $[0, g_k - 1]$. Let $\Pr(\rho_k = n)$ denotes the probability that ρ_k is equal to n, where $n \in \{0, \cdots, g_k - 1\}$. The probability $\Pr(\rho_k = n)$ depends on the ratio of the file popularity between the $(k+1)$-th and k-th files, given by

$$\frac{\mu_{k+1}}{\mu_k} = \frac{k^\eta}{(k+1)^\eta} = \Big(1 - \frac{1}{k+1}\Big)^\eta. \tag{34}$$

The popularity ratio $\frac{\mu_{k+1}}{\mu_k}$ approaches to 1 for a small η or a large k. In this case, the $(k+1)$-th file has almost the same popularity with the k-th file, and it should be cached with the same priority with the k-th file. Hence, the popularity ratio can reflect the probability that g_{k+1} is equal to g_k, and we have

$$\Pr(\rho_k = 0) = \Big(1 - \frac{1}{k+1}\Big)^\eta. \tag{35}$$

We now compute $\Pr(\rho_k = n)$ for $n = 1, \cdots, g_k - 1$. As $\rho_k = n$ means that there exist n relays which have cached the k-th files, while have not cached the $(k+1)$-th file. In other words, n relays have changed the cache status. Considering that the cache status at the relays is independent of each other, we can set $\Pr(\rho_k = n)$ for $n = 1, \cdots, g_k - 1$ as

$$\Pr(\rho_k = n) = q^n, \tag{36}$$

in which the variable $q \in \{0, 1\}$ should be subject to

$$q + q^2 + \cdots + q^{g_k - 1} = 1 - \Big(1 - \frac{1}{k+1}\Big)^\eta, \tag{37}$$

where q can be efficiently solved by some numerical methods. After collecting the probabilities $\Pr(\rho_k = n)$ for $n \in \{0, \cdots, g_k - 1\}$, we can recursively generate a sequence of $[g_1, g_2, \cdots, g_K]$ at random, which constitutes a sampled seed. After generating the L seeds, we can choose the best seed, which has the minimum SOP among the L seeds. The proposed SacLe strategy improves with a larger number of L, and it can achieve the optimal performance as long as L is large enough. Note that as the SacLe strategy can be implemented in parallel, increasing the number of seeds L will not lead to a larger latency. The details about the convergence and the implementation of such stochastic learning-based methods can be found in the literature, such as [44].

Algorithm 1: Proposed SacLe strategy

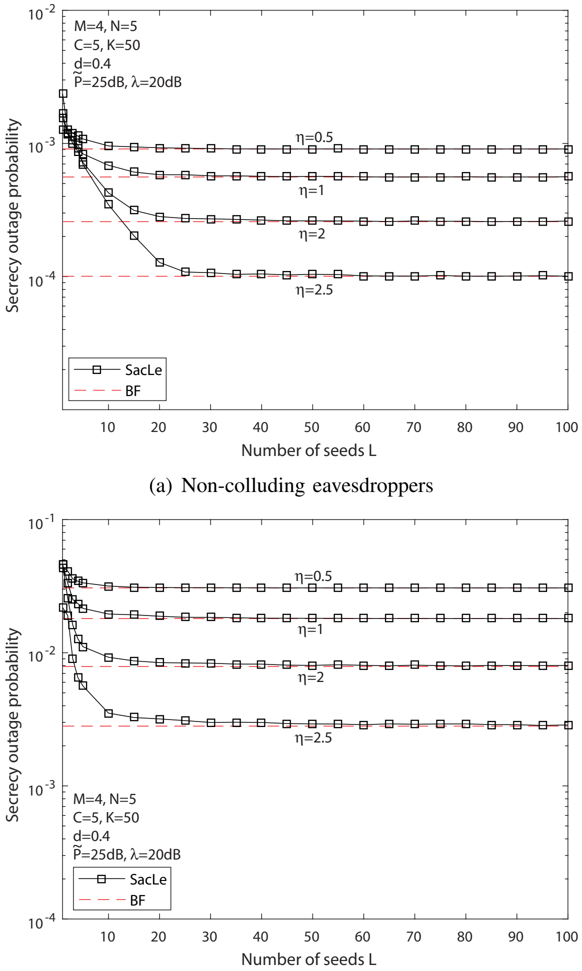
```
1:  Input the parameters $K, N, C, \eta$ and $L$.
2:  Calculate the probabilities of $\Pr(\rho_k = n)$ for $k \in \{0, \cdots, NC\}$ and $n \in \{0, \cdots, g_k - 1\}$.
3:  for Seed=1 : $L$ do
4:      $g_1$ is randomly chosen from $\{1, \cdots, N\}$.
5:      for $k = 1 : K$ do
6:          if $(g_k > 1)$ then
7:              Randomly generate an integer $\rho_k$ according to $\Pr(\rho_k = n)$.
8:              $g_{k+1} = g_k - \rho_k$.
9:          else
10:             $g_{k+1} = \min(g_k, NC - \sum_{k_1=1}^{k} g_{k_1})$.
11:         end if
12:     end for
13: end for
14: Compute $P_{out}$ of the $L$ seeds.
15: Find the best seed among $L$ ones which has the minimum $P_{out}$.
```

Algorithm 1 summarizes the process of the proposed SacLe strategy. After the input of the system parameters in line 1 and then calculating the probabilities $\Pr(\rho_k = n)$ in line 2, line 3-13 generate the L seeds at random according to the probabilities $\Pr(\rho_k = n)$. Specifically, we firstly randomly chooses an integer from $\{1, \cdots, N\}$ for g_1, as shown in line 4, and then recursively compute g_{k+1} based on g_k. If $g_k > 1$ holds, we randomly generate ρ_k, and compute g_{k+1}, as shown in line 6-8; otherwise, g_{k+1} is set to the minimum between g_k and the residual cache space $NC - \sum_{k_1=1}^{k} g_{k_1}$, as shown in line 9-10. After collecting these L seeds, we compute the associated P_{out} in line 14, and then find the best seed which has the minimum P_{out} among L ones in line 15. Note that the seed generation in line 3-13 and the associated SOP computation in line 14 can be implemented in parallel, which reduces the implementation latency substantially.

V. Simulations and Discussion

In this section, we provide simulation and numerical results to confirm the proposed analysis. All links in the network are subject to Rayleigh fading, with path loss factor of 3 [45]. The distance from S to D is normalized to unity, and the relays are assumed to be between. Let d stand for the distance between the relays and D, so that $\alpha = (1-d)^{-3}$ and $\beta_1 = d^{-3}$. From β_1 and a given MER λ, β_2 is given by $\frac{\beta_1}{\lambda}$. The average eavesdropping channel gain of the direct links is set to half of that of the relaying links, i.e., $\varepsilon = 0.5\beta_2$. The secrecy data rate is targeted to 0.5 bps/Hz, so that the associated secrecy SNR thresholds γ_{1s} and γ_{2s} are 1.41 and 2, respectively. There are totally 50 files to be transmitted from the source to the destination, and each relay can pre-store 5 files at most, so that $K = 50$ and $C = 5$.

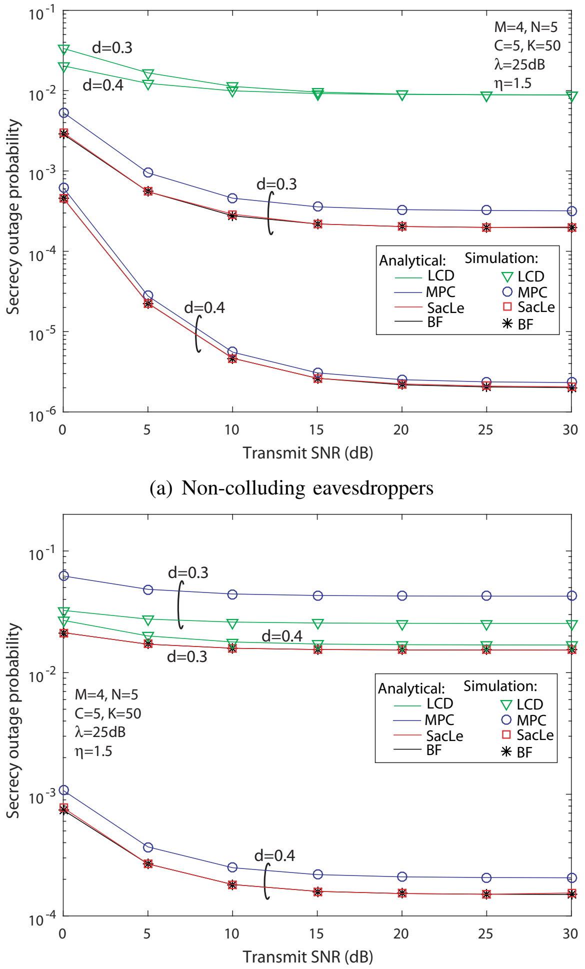
Fig. 2 depicts the effect of the number of seeds L on the simulated SOP of the proposed SacLe strategy, where $M = 4$, $N = 5$, $d = 0.4$, $\tilde{P} = 25$dB and $\lambda = 20$dB. The file popularity factor η varies from 0.5 to 2, where $\eta = 0.5$ represents a relatively flat file popularity, while $\eta = 2$ is a sharp file popularity. Specifically. Fig. 2 (a) and (b) are associated with the non-colluding and colluding eavesdroppers, respectively. As a benchmark, the secrecy performance of the BF caching strategy is also plotted, which is obtained through the Lingo software [26]. From this figure, we can conclude that for various values of the popularity factor η, with either non-colluding or colluding eavesdroppers, the secrecy performance of the SacLe becomes better with an increase in the number of



(a) Non-colluding eavesdroppers

(b) Colluding eavesdroppers

Fig. 2. Effect of the number of seeds on the secrecy outage probability for the proposed SacLe strategy.



(a) Non-colluding eavesdroppers

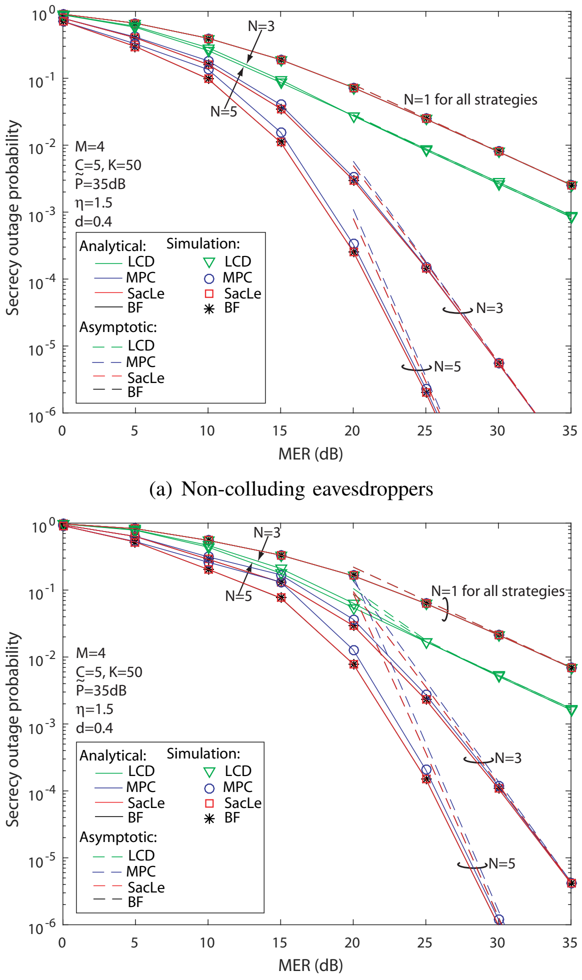
(b) Colluding eavesdroppers

Fig. 3. Secrecy outage probability of several caching strategies versus the transmit SNR.

seeds, and it converges to the optimal SOP of the BF strategy. Moreover, when L is sufficiently large, the SacLe can achieve the same SOP performance as the BF strategy. Specifically, to achieve the BF performance, when the eavesdroppers are non-colluding, η can be set to 20, 25, 35, 60 and 75, respectively; when the eavesdroppers are colluding, η can be set to 20, 35, 65 and 90, respectively. As the SacLe can be implemented in parallel, increasing the value of L will not impose a larger latency. Hence, next, we can set L to a sufficiently large value to obtain the optimal secrecy performance.

Fig. 3 demonstrates the closed-form and simulated SOP of the proposed SacLe, LCD, MPC and BF caching strategies versus the transmit SNR $\tilde{P}$, where $M = 4$, $N = 5$, $\lambda = 25$dB and $\eta = 1.5$. The value of d is set to 0.3 or 0.4, and the transmit SNR $\tilde{P}$ varies from 0dB to 30dB. Fig. 3 (a) and (b) correspond to the non-colluding and colluding eavesdroppers, respectively. We can observe from Fig. 3 that for all caching strategies with various values of $\tilde{P}$ and d, the closed-form SOP matches well with the simulated result, which validates the accuracy of the derived closed-form expressions for P_{out} in (20) and (26). Moreover, SOP improves with an increase in the value of $\tilde{P}$, as a larger transmit SNR can assist the data transmission. However, the improvement is saturated in the high region of the transmit SNR, as the MER becomes the bottleneck of the network security. Furthermore, the SacLe can achieve almost the same secrecy performance as the BF, and it outperforms the MPC and LCD strategies. Also, the MPC is even worse than the LCD when the eavesdroppers are non-colluding with $d = 0.3$. This is because that in this situation, the first relaying link is quite weak, and hence, the caching gain becomes more important than the signal cooperation gain in the network secrecy performance.

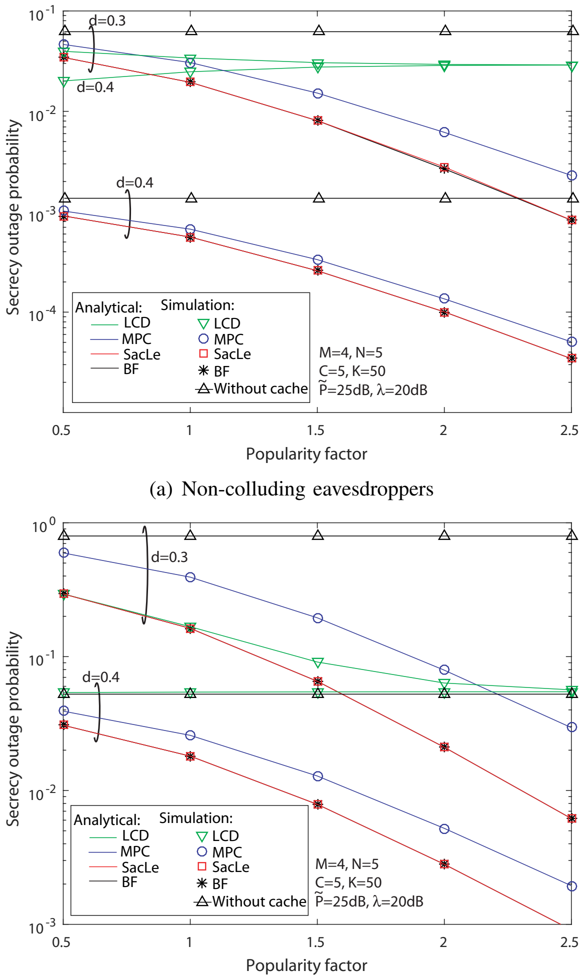
Fig. 4 illustrates the impact of MER λ on the closed-form, asymptotic and simulated SOP results of the several caching strategies, where $\tilde{P} = 35$dB, $\eta = 1.5$, $d = 0.4$, $M = 4$ and the relay number N is set to 1, 3 and 5. Specifically, Fig. 4 (a) and (b) are associated with the non-colluding and colluding eavesdroppers, respectively. We can observe from Fig. 4 that for either non-colluding or colluding eavesdroppers with various values of N, the closed-form SOP fits well with the simulation one in the entire MER region, and when the MER is high, the asymptotic SOP converges to the exact value, which validates the derived closed-form and asymptotic SOP expressions. Moreover, the SacLe outperforms the



(a) Non-colluding eavesdroppers

(b) Colluding eavesdroppers

Fig. 4. Secrecy outage probability of several cache placement strategies versus the MER, with different values of N.



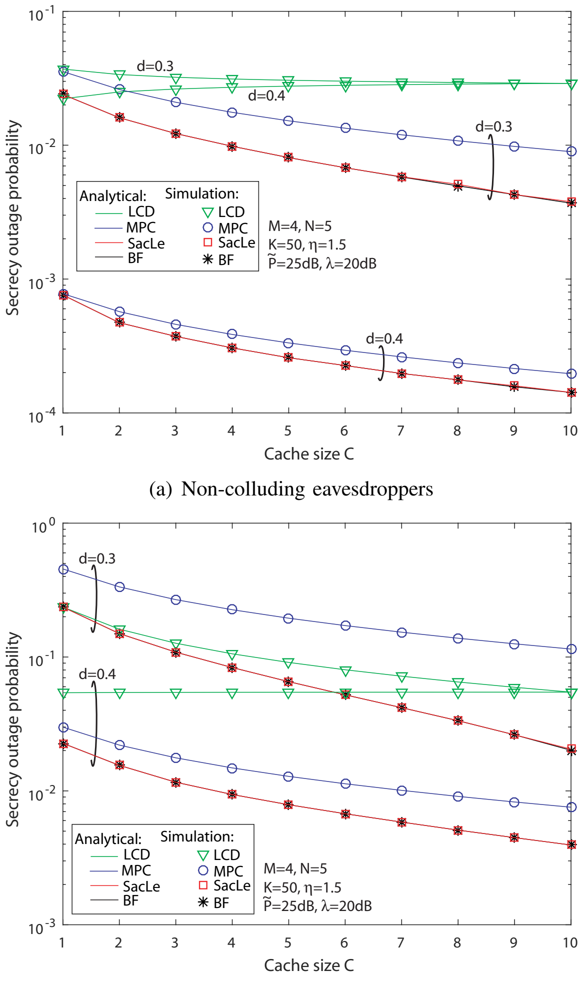
(a) Non-colluding eavesdroppers

(b) Colluding eavesdroppers

Fig. 5. Effect of the file popularity factor on the secrecy outage probability for several caching strategies and the secure transmission without cache.

traditional MPC and LCD, and it achieves the optimal secrecy performance of the BF. In the high MER region, the MPC can achieve the near-optimal secrecy performance. In further, the SacLe, BF and MPC improve rapidly with the number of relays in the high MER regime, and the associated curve slope is in parallel with N, which indicates that the system full secrecy diversity order is obtained by the three caching strategies. In contrast, although the secrecy performance of the LCD becomes better with an increase in the number of relays, the curve slope remains unchanged with N, as the LCD can only achieve the secrecy diversity order of unity, irrespective of the number of relays. Furthermore, by comparing the results in Fig. 4 (a) and (b), we can conclude that secrecy performance becomes worse when the eavesdroppers are colluding, as the cooperation between the eavesdroppers helps strengthen the eavesdropping links.

Fig. 5 shows the effect of the file popularity factor η on the secrecy outage probabilities of the several caching strategies, when $M = 4$, $N = 5$, $\tilde{P} = 25$dB and $\lambda = 20$dB. The value of d is set to 0.3 or 0.4, and the popularity factor η varies from 0.5 to 2.5, where 0.5 and 2.5 correspond to the relatively flat and quite concentrated popularity, respectively. Fig. 5 (a) and (b) are associated with the non-colluding and colluding eavesdroppers, respectively. To demonstrate the benefits of cache, we also provide the simulated secrecy outage probability of the secure transmission without cache in Fig. 5. As observed from Fig. 5, we can find that for either non-colluding or colluding eavesdroppers with various values of popularity factor, the performance of the secure transmission without cache is much worse than those of the SacLe, BF and MPC, which verifies the benefits of cache on the network security. When the popularity factor increases, the secrecy performances of the SacLe, BF and MPC improve profoundly, as a larger η leads to a higher popularity that the files are transmitted with an increasing secrecy diversity order. In contrast, the popularity factor has a less significant impact on the secrecy performance of LCD, as the secrecy diversity order of LCD is limited by unity. Moreover, the secrecy performances of all the caching strategies improve with a larger value of d, as the first-hop relaying links have been improved with d. In particular, when the value of d is equal to 0.3, MPC is even worse than LCD in some cases. This is because that the first-hop relaying links are not reliable when $d = 0.3$, and the signal cooperation gain from the MPC does not play an important



(a) Non-colluding eavesdroppers

(b) Colluding eavesdroppers

Fig. 6. Impact of cache size on the secrecy outage probability for several caching strategies.

role in the secure data transmission. In further, the secrecy performance in Fig. 5 (b) is much worse than that in Fig. 5 (a), since the cooperation among the colluding eavesdroppers can help strengthen the eavesdropping links. Finally, the closed-form SOP results are in good agreement with the simulation ones, for various values of η and d, which further validates the derived closed-form expressions for the network average secrecy outage probability.

Fig. 6 shows the impact of the cache size C on the secrecy outage probabilities of the several caching strategies, when $M = 4$, $N = 5$, $\tilde{P} = 25$dB, $\lambda = 20$dB and $\eta = 1.5$. The value of d is set to 0.3 or 0.4, and cache size C varies from 1 to 10. Fig. 6 (a) and (b) correspond to the non-colluding and colluding eavesdroppers, respectively. We can observe from Fig. 6 that for various values of C and d, the proposed SacLe can achieve the optimal performance of the BF strategy, and it outperforms the MPC and LCD strategies profoundly. Moreover, the secrecy outage probabilities of the SacLe, BF and MPC become much better with the increasing value of C, as larger cache size can help files pre-stored at more relays which can be exploited by the associated caching strategies. In further, the closed-form SOP results match well with the simulation ones, for various values of C and d, which validates the derived closed-form expressions for the network average secrecy outage probability.

VI. Conclusions

This paper investigated the communication security of a cache-aided multi-DF-relay network, where multiple eavesdroppers overheard the confidential message from both the source and relays. If the requested file was in the relay cache, then the relays directly send the file to the destination; otherwise, dual-hop secure data transmission was used. For the two cache status, relay selection was performed, based on the main channel parameters of the second-hop and dual-hop relaying links, respectively. We here analyzed the network security under the wiretap scenario of either non-colluding or colluding eavesdroppers, by providing closed-form and asymptotic expressions for the average SOP. To minimize the network SOP, we further optimize the cache placement by proposing the SacLe strategy, which can be implemented in parallel. Simulation and numerical results were demonstrated to confirm the proposed analysis. In particular, the caching strategy had a significant impact on the network secrecy performance through affecting the caching diversity gain and signal cooperation gain at the relays. Moreover, the proposed SacLe strategy outperformed the conventional MPC and LCD ones, and it was able to achieve the optimal performance obtained by the BF algorithm. In future works, we will further exploit the impact of the file popularity and average channels gains of the network links on the cache placement optimization, in order to develop some more efficient caching strategies for the considered system. Moreover, some intelligent techniques [46]–[48] such as the deep learning based algorithms will be utilized to optimize the considered system, in order to further enhance the network secrecy performance.

Appendix A
Proof of Theorem 1

According to (9) and (10), we first write $p_r(g_k)$ with $g_k \geq 1$ as

$$p_r(g_k) = \Pr\left(\frac{1+\tilde{P}v_{1n^*}}{1+\tilde{P}v_{2n^*}} < \gamma_{1s}\right), \tag{A.1}$$

where $v_{1n^*} = \max_{n\in\Omega_k} v_{1n}$ and $v_{2n^*} = \max_{1\leq m\leq M} v_{2n^*,m}$. By applying the PDF of v_{1n}, $f_{v_{1n}}(x) = \frac{1}{\beta_1}e^{-\frac{x}{\beta_1}}$, we obtain the CDF of v_{1n^*} as,

$$F_{v_{1n^*}}(x) = \left(1-e^{\frac{-x}{\beta_1}}\right)^{g_k} = \sum_{n=0}^{g_k}\binom{g_k}{n}(-1)^n e^{\frac{-nx}{\beta_1}}. \tag{A.2}$$

Then, $p_r(g_k)$ with $g_k \geq 1$ can be re-written as

$$p_r(g_k) = \sum_{n=0}^{g_k}\binom{g_k}{n}(-1)^n\int_0^\infty f_{v_{2n^*}}(x)e^{-\frac{n(\gamma_{1s}x+b_1)}{\beta_1}}dx, \tag{A.3}$$

where b_1 is given in (19) and the PDF of v_{2n^*} is given by [45]

$$f_{v_{2n^*}}(x) = \sum_{m=1}^{M}\binom{M}{m}(-1)^{m-1}\frac{m}{\beta_2}e^{-\frac{mx}{\beta_2}}. \tag{A.4}$$

By applying (A.4) into (A.3), and then tackling the resultant integral, we can get a closed-form expression for $p_r(g_k)$ with $g_k \geq 1$, when the eavesdroppers are non-colluding, as shown in (18).

We now extend to deduce a closed-form expression for $p_r(g_k)$ with $g_k = 0$, i.e., $p_r(0)$. From (12) and (15), we can write $p_r(0)$ as

$$p_r(0) = \Pr\Big(\frac{1+\tilde{P}\min(u_{n^*}, v_{1n^*})}{1+\tilde{P}\max_{1\leq m\leq M}(v_{2n^*,m}+w_m)} < \gamma_{2s}\Big), \tag{A.5}$$

which is re-expressed as

$$p_r(0) = \Pr\Big(\frac{1+\tilde{P}Z_1}{1+\tilde{P}Z_2} < \gamma_{2s}\Big), \tag{A.6}$$

with $Z_1 = \min(u_{n^*}, v_{1n^*})$ and $Z_2 = \max_{1\leq m\leq M}(v_{2n^*,m} + w_m)$. The CDF of Z_1 is given by [45],

$$F_{Z_1}(x) = \sum_{n=0}^{N}\binom{N}{n}(-1)^n e^{-(\frac{1}{\alpha}+\frac{1}{\beta_1})nx}. \tag{A.7}$$

Then $p_r(0)$ in (A.6) is written as

$$p_r(0) = \sum_{n=0}^{N}\binom{N}{n}(-1)^n \int_0^\infty e^{-(\frac{1}{\alpha}+\frac{1}{\beta_1})n(\gamma_{2s}z+b_2)} f_{Z_2}(z)dz, \tag{A.8}$$

where b_2 is defined in (19). To solve this integral, we need to derive the PDF of Z_2. Let $Z_{2m} = w_m + v_{2n^*,m}$, and the CDF of Z_{2m} is written as

$$F_{Z_{2m}}(z) = \int_0^z \int_0^{z-v} f_{v_{2n^*,m}}(v) f_{w_m}(w)dwdv, \tag{A.9}$$

$$= 1 - \frac{\varepsilon}{\varepsilon-\beta_2}e^{-\frac{z}{\varepsilon}} + \frac{\beta_2}{\varepsilon-\beta_2}e^{-\frac{z}{\beta_2}}, \tag{A.10}$$

where the PDFs of $f_{v_{2n^*,m}}(v) = \frac{1}{\beta_2}e^{-\frac{v}{\beta_2}}$ and $f_{w_m}(w) = \frac{1}{\varepsilon}e^{-\frac{w}{\varepsilon}}$ are applied in the last equality. Based on $F_{Z_{2m}}(z)$, we can obtain the PDF of Z_2 as

$$f_{Z_2}(z) = \frac{d}{dz}\Big(F_{Z_{2m}}^M(z)\Big) \tag{A.11}$$

$$= \sum_{m_1=0}^{M}\sum_{m_2=0}^{m_1}\Big(\frac{m_1-m_2}{\beta_2}+\frac{m_2}{\varepsilon}\Big)\tau_{m_1m_2}e^{-(\frac{m_1-m_2}{\beta_2}+\frac{m_2}{\varepsilon})z}. \tag{A.12}$$

By applying (A.12) into (A.8), and then tackling the resultant integral, we can get a closed-form expression for $p_r(0)$, as shown in (18). Hence, Theorem 1 is proven.

Appendix B
Proof of Theorem 2

We first derive an asymptotic expression for $p_r(g_k)$, with $g_k \geq 1$. When the transmit SNR $\tilde{P}$ is large, $p_r(g_k)$ with $g_k \geq 1$ in (A.1) can be approximated as

$$p_r(g_k) \simeq \Pr(v_{1n^*} < \gamma_{1s}v_{2n^*}). \tag{B.1}$$

By using the exponential approximation, $e^{-x} \simeq 1-x$, where $|x|$ is small [30], the CDF of v_{1n^*} is asymptotically given by

$$F_{v_{1n^*}}(x) \simeq \big(\frac{x}{\beta_1}\big)^{g_k}. \tag{B.2}$$

By applying (B.2) into (B.1) and then tackling the resultant integral, we get an asymptotic formula for $p_r(g_k)$ with $g_k \geq 1$ as

$$p_r(g_k) \simeq \int_0^\infty \big(\frac{\gamma_{1s}v}{\beta_1}\big)^{g_k} f_{2n^*}(v)dv. \tag{B.3}$$

By solving the integral in (B.3), we can obtain a closed-form expression for $p_r(g_k)$ with $g_k \geq 1$.

We then derive an asymptotic expression of $p_r(0)$. From (A.6), the asymptotic $p_r(0)$ is given by

$$p_r(0) \simeq \Pr(Z_1 < \gamma_{2s}Z_2). \tag{B.4}$$

By employing the asymptotic result, $e^{-x} \simeq 1-x$, the CDF of Z_1 is asymptotically given by

$$F_{Z_1}(x) \simeq \big(\frac{1}{\alpha}+\frac{1}{\beta_1}\big)^N x^N. \tag{B.5}$$

Then, the asymptotic $p_r(0)$ is given by

$$p_r(0) \simeq \big(\frac{1}{\alpha}+\frac{1}{\beta_1}\big)^N \gamma_{2s}^N \int_0^\infty {Z_2}^N f_{Z_2}(Z_2)dZ_2. \tag{B.6}$$

By solving the integral in the above equation, we can get a closed-form expression in (21), and Theorem 2 is proven.

Appendix C
Proof of Theorem 3

We first derive a closed-form expression for $p_r(g_k)$, when $g_k \geq 1$. For the M colluding eavesdroppers, v_{2n^*} in (A.1) becomes equivalent to $v_{2n^*,1} + v_{2n^*,2} + \cdots + v_{2n^*,M}$, and its PDF is given by [45]

$$f_{v_{2n^*}}(x) = \frac{x^{M-1}}{\Gamma(M)\beta_2^M}e^{-\frac{x}{\beta_2}}. \tag{C.1}$$

By applying (C.1) into (A.3), and then solving the required integral, we can obtain (24).

We then compute $p_r(0)$ with colluding eavesdroppers. In this case, Z_2 in (A.6) becomes

$$Z_2 = \sum_{m=1}^{M}(w_m + v_{2n^*,m}) = \underbrace{\sum_{m=1}^{M} w_m}_{Z_{21}} + \underbrace{\sum_{m=1}^{M} v_{2n^*,m}}_{Z_{22}}. \tag{C.2}$$

The PDFs of Z_{21} and Z_{22} are

$$f_{Z_{21}}(z) = \frac{z^{M-1}}{\Gamma(M)\varepsilon^M}e^{-\frac{z}{\varepsilon}}, \tag{C.3}$$

$$f_{Z_{22}}(z) = \frac{z^{M-1}}{\Gamma(M)\beta_2^M}e^{-\frac{z}{\beta_2}}. \tag{C.4}$$

From (C.3) and (C.4), we can rewrite $p_r(0)$ in (A.8) as

$$p_r(0) = \sum_{n=0}^{N}\binom{N}{n}(-1)^n \int_0^\infty\int_0^\infty e^{-(\frac{1}{\alpha}+\frac{1}{\beta_1})n[\gamma_{2s}(z_{21}+z_{22})+b_2]} \times f_{Z_{21}}(z_{21})f_{Z_{22}}(z_{22})dz_{21}dz_{22}. \tag{C.5}$$

By solving the above integral, we can obtain a closed-form expression for $p_r(0)$ in (24), and Theorem 3 is proven.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TCOMM.2019.2935047, IEEE Transactions on Communications

Appendix D
Proof of Theorem 4

We first derive the asymptotic expression of $p_r(g_k)$, for $g_k \geq 1$. By using the PDF of v_{2n^*} in (C.1) into (B.3) and then tackling the resultant integral, we get (27).

We then derive the asymptotic expression of $p_r(0)$. As shown in (B.6), we need to evaluate

$$\int_0^\infty {Z_2}^N f_{Z_2}(Z_2) dZ_2, \tag{D.1}$$

where $Z_2 = Z_{21} + Z_{22}$ for the colluding eavesdroppers. The form Z_2^N is expanded as

$$Z_2^N = (Z_{21} + Z_{22})^N = \sum_{n=0}^{N} \binom{N}{n} Z_{21}^n Z_{22}^{N-n}. \tag{D.2}$$

From (D.2), we can derive an asymptotic formula for $p_r(0)$ as

$$p_r(0) \simeq \left(\frac{1}{\alpha} + \frac{1}{\beta_1}\right)^N \gamma_{2s}^N \sum_{n=0}^{N} \binom{N}{n} \int_0^\infty \int_0^\infty Z_{21}^n Z_{22}^{N-n} \\ \times f_{Z_{21}}(Z_{21}) f_{Z_{22}}(Z_{22}) dZ_{21} dZ_{22}. \tag{D.3}$$

By solving the integral in the above equation, we can obtain (27), and Theorem 4 is proven.

References

[1] C. Zhong, X. Jiang, F. Qu, and Z. Zhang, "Multi-antenna wireless legitimate surveillance systems: Design and performance analysis," *IEEE Trans. Wireless Commun.*, vol. 16, no. 7, pp. 4585–4599, 2017.

[2] J. Xu, W. Xu, J. Zhu, D. K. Ng, and A. L. Swindlehurst, "Secure massive MIMO communication with low-resolution DACs," *IEEE Trans. Commun.*, vol. PP, no. 99, pp. 1–10, 2019.

[3] 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 Trans. Vehic. Tech.*, vol. 65, no. 10, pp. 8767–8774, 2016.

[4] C. E. Shannon, "Communication theory of secrecy systems," *Bell Syst. Tech. J.*, vol. 28, no. 2, pp. 656–715, 1948.

[5] A. D. Wyner, "The wire-tap channel," *Bell Labs Technical Journal*, vol. 54, no. 8, pp. 1355–1367, Oct. 1975.

[6] C. Zhong, H. A. Suraweera, G. Zheng, I. Krikidis, and Z. Zhang, "Wireless information and power transfer with full duplex relaying," *IEEE Trans. Commun.*, vol. 62, no. 10, pp. 3447–3461, 2014.

[7] C. Xing, F. Gao, and Y. Zhou, "A framework for transceiver designs for multi-hop communications with covariance shaping constraints," *IEEE Trans. Sig. Proc.*, vol. 63, no. 15, pp. 3930–3945, Aug. 2015.

[8] L. Sun, P. Ren, Q. Du, Y. Wang, and Z. Gao, "Security-aware relaying scheme for cooperative networks with untrusted relay nodes," *IEEE Commun. Lett.*, vol. 19, no. 3, pp. 463–466, 2015.

[9] J. Yao and Y. Liu, "Secrecy rate maximization with outage constraint in multihop relaying networks," *IEEE Commun. Lett.*, vol. 22, no. 2, pp. 304–307, 2018.

[10] L. Sun and H. Xu, "Unequal secrecy protection for untrusted two-way relaying systems: Constellation overlapping and noise aggregation," *IEEE Trans. Vehic. Tech.*, vol. 67, no. 10, pp. 9681–9695, 2018.

[11] L. Fan, X. Lei, N. Yang, T. Q. Duong, and G. K. Karagiannidis, "Secure multiple amplify-and-forward relaying with co-channel interference," *IEEE J. Sel. Topics Sig. Proc.*, vol. 10, no. 8, pp. 1494–1505, Dec. 2016.

[12] L. Fan, N. J. Yang, T. Q. Duong, M. Elkashlan, and G. K. Karagiannidis, "Exploiting direct links for physical layer security in multiuser multirelay networks," *IEEE Trans. Wireless Commun.*, vol. 15, no. 6, pp. 3856–3867, 2016.

[13] Q. Li and J. Qin, "Joint source and relay secure beamforming for nonregenerative MIMO relay systems with wireless information and power transfer," *IEEE Trans. Vehic. Tech.*, vol. 66, no. 7, pp. 5853–5865, 2017.

[14] X. Lin, "MARL-based distributed cache placement for wireless networks," *IEEE Access*, vol. 7, pp. 62 606–62 615, 2019.

[15] M. Zhang and Y. Liu, "Secure beamforming for untrusted MISO cognitive radio networks," *IEEE Trans. Wireless Commun.*, vol. 17, no. 7, pp. 4861–4872, 2018.

[16] L. Xiao, G. Han, D. Jiang, H. Zhu, Y. Zhang, and H. V. Poor, "Two-dimensional anti-jamming mobile communication based on reinforcement learning," *IEEE Trans. Vehic. Tech.*, vol. 67, no. 10, pp. 9499–9512, 2018.

[17] L. Xiao, Y. Li, C. Dai, H. Dai, and H. V. Poor, "Reinforcement learning-based NOMA power allocation in the presence of smart jamming," *IEEE Trans. Vehic. Tech.*, vol. 67, no. 4, pp. 3377–3389, 2018.

[18] L. Xiao, J. Liu, Q. Li, N. B. Mandayam, and H. V. Poor, "User-centric view of jamming games in cognitive radio networks," *IEEE Trans. Information Forensics and Security*, vol. 10, no. 12, pp. 2578–2590, 2015.

[19] J. Xia, "When distributed switch-and-stay combining meets buffer in IoT relaying networks," *Physical Commun.*, vol. PP, pp. 1–9, 2019.

[20] Z. Ding, P. Fan, G. K. Karagiannidis, R. Schober, and H. V. Poor, "NOMA assisted wireless caching: Strategies and performance analysis," *IEEE Trans. Commun.*, vol. 66, no. 10, pp. 4854–4876, 2018.

[21] N. Zhao, F. Cheng, F. R. Yu, J. Tang, Y. Chen, G. Gui, and H. Sari, "Caching UAV assisted secure transmission in hyper-dense networks based on interference alignment," *IEEE Trans. Commun.*, vol. 66, no. 5, pp. 2281–2294, 2018.

[22] J. Li, M. Liu, J. Lu, F. Shu, Y. Zhang, S. Bayat, and D. N. K. Jayakody, "On social-aware content caching for D2D-enabled cellular networks with matching theory," *IEEE Internet Things J.*, accepted to appear.

[23] Y. Liu, Q. Chen, X. Tang, and L. X. Cai, "On the buffer energy aware adaptive relaying in multiple relay network," *IEEE Trans. Wireless Commun.*, vol. 16, no. 9, pp. 6248–6263, Sept. 2017.

[24] N. Zhao, X. Liu, F. R. Yu, M. Li, and V. C. M. Leung, "Communications, caching, and computing oriented small cell networks with interference alignment," *IEEE Commun. Mag.*, vol. 54, no. 9, pp. 29–35, 2016.

[25] G. Zheng, H. A. Suraweera, and I. Krikidis, "Optimization of hybrid cache placement for collaborative relaying," *IEEE Commun. Lett.*, vol. 21, no. 2, pp. 442–445, Feb. 2017.

[26] L. Fan, N. Zhao, X. Lei, Q. Chen, N. Yang, and G. K. Karagiannidis, "Outage probability and optimal cache placement for multiple amplify-and-forward relay networks," *IEEE Trans. Veh. Technol.*, vol. 67, no. 12, pp. 12 373–12 378, 2018.

[27] L. Xiang, D. W. K. Ng, R. Schober, and V. W. S. Wong, "Cache-enabled physical layer security for video streaming in backhaul-limited cellular networks," *IEEE Trans. Wireless Commun.*, vol. 17, no. 2, pp. 736–751, 2018.

[28] ——, "Secure video streaming in heterogeneous small cell networks with untrusted cache helpers," *IEEE Trans. Wireless Commun.*, vol. 17, no. 4, pp. 2645–2661, 2018.

[29] J. Xia, "Intelligent secure communication for internet of things with statistical channel state information of attacker," *IEEE Access*, vol. PP, no. 99, pp. 1–7, 2019.

[30] I. S. Gradshteyn and I. M. Ryzhik, *Table of Integrals, Series, and Products*, 7th ed. San Diego, CA: Academic, 2007.

[31] K. Huang, H. Wang, Y. Wu, and R. Schober, "Pilot spoofing attack by multiple eavesdroppers," *IEEE Trans. Wireless Commun.*, vol. 17, no. 10, pp. 6433–6447, 2018.

[32] Y. Dong, A. E. Shafie, M. J. Hossain, J. Cheng, N. Al-Dhahir, and V. C. M. Leung, "Secure beamforming in full-duplex MISO-SWIPT systems with multiple eavesdroppers," *IEEE Trans. Wireless Commun.*, vol. 17, no. 10, pp. 6559–6574, 2018.

[33] J. Lee, "Confidential multicasting assisted by multi-hop multi-antenna DF relays in the presence of multiple eavesdroppers," *IEEE Trans. Commun.*, vol. 64, no. 10, pp. 4295–4304, 2016.

[34] Z. Chu, H. Xing, M. Johnston, and S. Y. L. Goff, "Secrecy rate optimizations for a MISO secrecy channel with multiple multiantenna eavesdroppers," *IEEE Trans. Wireless Commun.*, vol. 15, no. 1, pp. 283–297, 2016.

[35] X. Yuan, X. Wang, J. Wang, Y. Chu, C. Wang, J. Wang, M. Montpetit, and S. Liu, "Enabling secure and efficient video delivery through encrypted in-network caching," *IEEE J. Sel. Area Commun.*, vol. 34, no. 8, pp. 2077–2090, 2016.

[36] P. Blasco and D. Gündüz, "Learning-based optimization of cache content in a small cell base station," in *IEEE International Conference on Communications, ICC, Sydney, Australia*, 2014, pp. 1897–1903.

[37] A. Sengupta, S. Amuru, R. Tandon, R. M. Buehrer, and T. C. Clancy, "Learning distributed caching strategies in small cell networks," in *11th International Symposium on Wireless Communications Systems, ISWCS, Barcelona, Spain*, 2014, pp. 917–921.

[38] E. Bastug, M. Bennis, E. Zeydan, M. A. Kader, I. A. Karatepe, A. S. Er, and M. Debbah, “Big data meets telcos: A proactive caching perspective,” *Journal of Commun. and Net.*, vol. 17, no. 6, pp. 549–557, 2015.

[39] P. Yang, N. Zhang, S. Zhang, L. Yu, J. Zhang, and X. Shen, “Content popularity prediction towards location-aware mobile edge caching,” *IEEE Trans. Multimedia*, vol. 21, no. 4, pp. 915–929, 2019.

[40] X. Lai, L. Fan, X. Lei, J. Li, N. Yang, and G. K. Karagiannidis, “Distributed secure switch-and-stay combining over correlated fading channels,” *IEEE Trans. Information Forensics and Security*, vol. 14, no. 8, pp. 2088–2101, August 2019.

[41] C. Wang, H. Wang, and X. Xia, “Hybrid opportunistic relaying and jamming with power allocation for secure cooperative networks,” *IEEE Trans. Wireless Commun.*, vol. 14, no. 2, pp. 589–605, 2015.

[42] M. Gngor, Y. Bulut, and S. Calk, “Distribution of order statistics,” *Applied Mathematical Sciences*, vol. 3, no. 16, pp. 795–802, 2009.

[43] http://www.lingo.com.

[44] X. Huang, W. Xu, G. Xie, S. Jin, and X. You, “Learning oriented cross-entropy approach to user association in load-balanced HetNet,” *IEEE Wireless Commun. Lett.*, vol. 7, no. 6, pp. 1014–1017, 2018.

[45] M. K. Simon and M. S. Alouini, *Digital Communication over Fading Channels*, 2nd ed. John Wiley, 2005.

[46] C. Li, W. Zhou, and et.al, “Enhanced secure transmission against intelligent attacks,” *IEEE Access*, vol. 7, pp. 53 596–53 602, 2019.

[47] Z. Zhao and et.al, “A novel framework of three-hierarchical offloading optimization for mec in industrial IoT networks,” *IEEE Transactions on Industrial Informatics*, vol. PP, no. 99, pp. 1–12, 2019.

[48] G. Liu, “Deep learning based channel prediction for edge computing networks towards intelligent connected vehicles,” *IEEE Access*, vol. PP, pp. 1–10, 2019.