IEEE Access
Multidisciplinary ⋮ Rapid Review ⋮ Open Access Journal

Received March 16, 2021, accepted March 25, 2021, date of publication March 30, 2021, date of current version April 9, 2021.

Digital Object Identifier 10.1109/ACCESS.2021.3069736

Performance Analysis of a Cache-Aided Wireless Heterogeneous Network With Secrecy Constraints

GEORGIOS SMPOKOS[1,4], **(Graduate Student Member, IEEE),**
ZHENG CHEN[2], **(Member, IEEE), PARTHAJIT MOHAPATRA**[3], **(Member, IEEE),**
AND NIKOLAOS PAPPAS[4], **(Member, IEEE)**

[1]Department of Group Network Engineering and Delivery, Vodafone, 11525 Athens, Greece
[2]Department of Electrical Engineering, Linköping University, S-581 83 Linköping, Sweden
[3]Department of Electrical Engineering, IIT Tirupati, Tirupati 517506, India
[4]Department of Science and Technology, Linköping University, SE-60174 Norrköping, Sweden

Corresponding author: Nikolaos Pappas (nikolaos.pappas@liu.se)

This work was supported in part by the Swedish Foundation for Strategic Research (SSF), in part by the Swedish Research Council (VR), in part by the Center for Industrial Information Technology (CENIIT), in part by the Excellence Center at Linköping-Lund in Information Technology (ELLIIT), and in part by the Indo-Sweden Project, Department of Science and Technology, India.

ABSTRACT In this paper, we analyze the impact of caching on the performance of a cache enabled system with heterogeneous traffic where one of the users need to be served with confidential data. In this setup, a wireless helper system always serves a dedicated user and it can also serve a user requesting cached content. A cellular network access point is also available to serve the latter user if it cannot retrieve the requested data from the helper's cache. The impact of caching and secrecy on throughput and delay performance for each user is then examined when the access point can deploy superposition coding to serve both users simultaneously. Two decoding schemes are considered in this work. The first decoding scheme treats interference from parallel transmissions as noise while the second one utilizes the parallel transmission to apply successive decoding for the intended data. Furthermore, network and cache related factors are identified and their impact on the overall performance of the system are analyzed. In order to find the optimal transmission power allocations, two distinct optimization problems are set in this context comparing the two decoding schemes. This will assist to identify the benefits of the considered decoding schemes for each user satisfying the secrecy requirements of the dedicated user and reducing its impact on the overall performance of the system.

INDEX TERMS Caching, delay analyis, secrecy, superposition coding.

I. INTRODUCTION

Video and image content has become the dominant type of wireless data traffic, and in most cases, this content can be requested several times which makes it reusable. Motivated by this, caching at the edge of the network has been identified as a promising approach to meet the high demand of reusable content [2], by the users. The key idea behind proactive caching is to store likely-to-be-requested content at the network edge nodes according to some caching policies (e.g. most popular content, random caching, coded caching) during off-peak hours [3]–[6]. When users request content that is already cached in their nearby nodes, the content delivery delay can be greatly reduced and the throughput of users requesting non-reusable content can vastly increase. Furthermore, users have different secrecy requirements, thus it is important to analyze the impact of caching on the system performance under secrecy constraints in heterogeneous traffic conditions.

In this work, we consider a network scenario, where two users have different secrecy requirements. A dedicated user receives external traffic with secrecy requirements while a non-dedicated user without secrecy requirements requests reusable (cacheable) content that can be stored at the edge node's (helper) cache. If the non-dedicated user requests for content that cannot be found at the helper's cache, then it can be served by the core network server through a base station. We consider the presence of a passive eavesdropper, which is not part of the network. The eavesdropper intends to decode the transmissions under secrecy requirements.

The associate editor coordinating the review of this manuscript and approving it for publication was Anton Kos.

When the helper needs to serve multiple users simultaneously, the decoding process and capability of the users can impact the system performance. Hence, it is important to investigate how the simultaneous transmission from the helper to different users can enhance or impede secure communication under different decoding schemes at the users. To measure the secrecy performance of the system, the notion of physical layer secrecy is considered which can exploit the randomness present in the wireless environment.

A. RELATED WORK

In [7], it was shown that is possible to send messages securely over a noisy channel without using any key between the legitimate nodes. The random nature of the physical wireless channel was studied for the case of the wiretap channel, where an eavesdropper tries to decode the messages intended to the appropriate receiver. A secure communication was maintained between the transmitter and the dedicated receiver without the need of cryptographic or other security coding techniques. The problem of secure communication over multiuser scenarios has been studied extensively under different settings [8]–[12]. The impact of fading on secure communication has been explored under various settings in [13]–[16]. It has been demonstrated that fading wireless channels can facilitate secure communication in contrast to the case of Gaussian wiretap channel [13], [15].

There is also a connection between the secure communication problem considered in this work and digital watermarking. Many of the communication techniques used for reliable and secure communication can be useful for watermarking. In the watermarking process, embedding information (known as a watermark) to the underlying signal results in limited distortion to the original signal. The work in [17] establishes an equivalence between watermarking game and a communication system with a jammer where the transmitter and jammer have access to different side-information. The code capacity is characterized for the watermarking game in the case of Gaussian covertext and squared-error distortion. Furthermore, the work in [18] provided an information-theoretic analysis of information hiding and characterized the achievable communication rate for the information hider. The work in [19] proposed a cognitive radio scheme that allows a secondary user to transmit over the same time-frequency slot of a primary user. It is shown that the secondary user can superimpose its information symbols on the primary user's signal without degrading the performance of the primary user and under certain conditions it can improve the performance of the system. Then, the previous work was extended by [20] by introducing the concept of convolutive superposition. In this case, multiple secondary user symbols are superimposed on the primary user received signal through a time-domain convolution which increases the achievable rate of the secondary user. The problem considered in this paper has similarities with watermarking where the superposition coding performed at the helper node can help to protect confidential data under certain conditions. Considering the proposed setup with a distortion constraint will be an interesting problem for future research, where the user needs to hide some information in the original message.

The analysis in [14] considered the secure broadcasting in the presence of fading channel. Some other findings regarding multiple user cases with secrecy requirements are demonstrated in [9], [10]. Data arrival at the service nodes in wireless networks is bursty however, a large number of research works in information theory assumes the presence of backlogged users [7]–[13]. In [21] the stability region of a broadcast channel for two users was examined taking into consideration the security of a single link and the bursty nature of packet arrival at the sender node. The secrecy constraints of a wireless broadcast channel were exploited in [22] and [23], where a confidential broadcast transmission is directed to multiple users that need to decode their dedicated packets to remain secret from the other users. Finally, network utility optimization in terms of reliability, stability, and secrecy was examined in [22] and [23].

Lately, there is an increasing focus on the delay analysis and the combined delay-throughput analysis for cache-enabled wireless networks providing useful insights in this growing research area as in [24], [25]. However, the delay analysis conducted in that research work considers the backhaul delay and the packet transmission delay for the saturated request cases. The study in [26] provided analysis on stable throughput and delay performance for single bottleneck cache enabled networks using stochastic request arrivals at different nodes. In [27], the authors investigated how bursty traffic and random caching availability of a small cell node affects the delay and throughput performance of a wireless caching system with two users. The works in [28]–[30] consider jointly physical layer security and caching.

B. CONTRIBUTIONS

This work explores the role of caching on the performance of a system under different traffic characteristics where a user need to be served with secret data. The wireless helper system that offers caching capabilities is a type of dynamic access node that could be seen as a small cell base station. This node can serve users with and without secrecy requirements and can hand over traffic to cellular network access nodes. To the best of our knowledge, physical layer secrecy in conjunction with caching where users need to be served with heterogeneous traffic characteristics has not been examined in the existing literature. The main contributions of this work are summarized below.

- The work derives probability of successful decoding for various decoding schemes for the cache-enabled wireless network where one of the users needs to be served with secure data and the channel between different nodes undergo Rayleigh fading. The work considers two approaches for decoding of a packet at the receivers: *treating interference as noise* and *successive decoding*. The probability of successful decoding for various schemes takes account of secrecy and reliability criteria.

- The work also derives the average service rate and delay for the considered system model with and without secrecy constraint for various decoding schemes. These performance metrics take account of the event that specific content can be found in the cache or not. The derived results also take account of the congestion level of the backhaul. The probability of successful decoding for various schemes helps to characterize these performance metrics under heterogeneous traffic characteristics.
- The derived results are utilized to optimize the performance in terms of throughput and delay of the system under different constraints on the parameters.

The findings of this work on different decoding schemes and network availability statistics, provide us insights into the network performance and service disruption while keeping the dedicated communications secure within a specific area. In real-life, this type of scenario can arise in cellular networks for different types of users. These users could have different subscription settings, where some have higher security requirements than other subscribers. Confidentiality could be very important for IoT network implementations where sensors and other devices collect data and monitor their respective environment. IoT devices that transmit and receive non-confidential data could play the role of the undedicated user requesting some updates and data from the helper system. The helper node stores the data into different queues transmitting into a single channel both the confidential and non-confidential data applying superposition coding. The ability of the transmitter to send non-confidential data to the undedicated user and hide the confidential data is eventually affected by its caching capabilities.

II. NETWORK MODEL

We consider a network consisting of one access point S with caching capability and two legitimate users with different traffic characteristics and secrecy requirements. The data traffic intended for the dedicated user, labeled as D, arrives at the access point S according to a Bernoulli process with arrival rate λ. Let Q represent the size of the queue at S that contains all the data packets waiting to be delivered to D. The access point S is equipped with cache memory, which can proactively store reusable content to be distributed. In addition to handling the traffic intended for D, the access point S can also serve as a caching helper to the non-dedicated user U, which occasionally requests for some content. The generated request by U will be first directed to the caching helper S. If the requested file is stored in the cache of S, then the file will be transmitted from S to U. Otherwise, the request will be re-directed to a nearby base station, and the file will be retrieved from a remote data center (DC) through the base station. In case both D and U are actively receiving their data from S, superposition coding (SC) is used to serve both the users [31]. The data communicated to the dedicated user D needs to be kept secret from a passive eavesdropper E, which is not part of the network as depicted in Fig. 1.

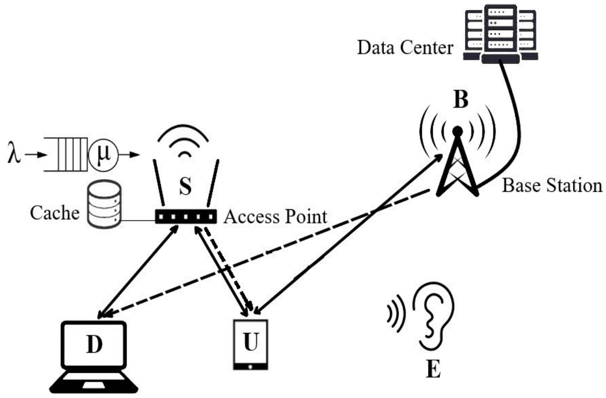

FIGURE 1. The system model. Solid lines represent intended transmissions and dotted lines represent interference.

TABLE 1. Probabilities notation.

Probability	Description
q_S	S transmits a packet to D
q_U	U will make a request
p_h/p_m	cache hit/miss from S
α	DC can serve U

In this work, an SNR/SINR based secrecy metric is used to measure the secrecy performance of the system under different decoding assumptions at the users and eavesdropper. In physical layer secrecy, some of the other commonly used secrecy metrics for fading scenarios are secrecy outage probability or ergodic secrecy rate [32]. However, these metrics do not take into account of decoding ability of the eavesdropper. On the other hand, SINR based metric can take into account of decoding ability of the eavesdropper and can also be used when packet length is short.

We assume that in each timeslot, the non-dedicated user U makes a content request with probability q_U. If the requested file is located inside the cache of S (cache hit), then S can deliver the file to U directly. In the meanwhile, if the queue at S is non-empty, S will transmit one packet to its dedicated user D with probability q_S, either through single transmission or parallel transmission to both U and D using SC. Due to limited storage capacity, the requested file by U can be found within the cache of S with probability p_h, which depends on the caching policy at the helper S and request pattern of U. In case of a cache miss event, with probability $p_m = 1 - p_h$, the content will be delivered to U from the remote data center through the base station B. Additionally, we assume that in each timeslot, the data center is available with probability α. α is a way to model congestion of the backhaul link that can be caused by several factors but it is outside the scope of this work to consider them in more detail.

Table 1 summarizes the meaning of these probabilities.

When treating interference as noise, whether a packet can be decoded correctly depends on if the received SINR or SNR exceeds a certain threshold. We consider Rayleigh fading and the power-law path loss model. The received SINR for the transmission link $i \to j$ is given by

$$\mathrm{SINR}_{ij/L} = \frac{P_{ij}|h_{ij}|^2 r_{ij}^{-\gamma}}{\sigma^2 + \sum_{k\in T\setminus\{i\}} P_{kj}|h_{kj}|^2 r_{kj}^{-\gamma}} \geq \theta_j, \qquad (1)$$

where $i \in \{S, B\}$ and $j \in \{D, U\}$, $L \subseteq \{SD, SU, BU\}$ is the set of active links. T represents the set of active transmitters, P_{ij} denotes the transmit power of the link $i \to j$. h_{ij} denotes the small-scale channel fading for the link $i \to j$, which follows $\mathcal{CN}(0, 1)$, r_{ij} is the distance of the link, σ^2 is the thermal noise power.

When SC is used at the access point S, powers P_{SD} and P_{SU} are allocated for the transmissions to the dedicated user and non-dedicated user respectively, such that

$$P_{SD} + P_{SU} = P_{\max}, \qquad (2)$$

where $P_{\max}$ is the maximum transmit power of S. Let P_B represent the transmit power of the base station B. For both cases, it is assumed that the eavesdropper decodes the message of the dedicated user by treating interference as noise. This kind of scenario can arise in practice when the eavesdropper has limited decoding ability or does not have access to the codebook used by the non-dedicated user.[1]

III. SUCCESS PROBABILITIES WITH SECRECY CONSTRAINTS

In this section, we obtain the probabilities of successful decoding for the two users with and without secrecy constraints. We consider two different decoding schemes, since in wireless networks, different users may have different decoding capabilities based on their hardware and software limitations (e.g. IoT applications). A conventional decoding scheme is to treat interference as noise (TIN), where parallel transmissions interfering with the intended signals will be treated as noise. A more advanced receiving scheme is successive decoding (SD), where the receiver first tries to decode the packet for the unintended user, then uses the decoded signal to cancel the interference caused to its transmission [33]. Based on the analysis in [21], we cannot perform successive decoding at both users simultaneously because this will result in infeasible power allocations for the packet transmissions. We consider the following scenarios based on different decoding ability at the users

1) Both the users D and U treat interference as noise.
2) The dedicated user D performs successive decoding and user U treats interference as noise.
3) The dedicated user D treats interference as noise and user U performs successive decoding.

TABLE 2. Cases notation.

Cases	Description
Case 1	Parallel transmissions $S \to D$ and $S \to U$
Case 2	Parallel transmissions $S \to D$ and $B \to U$
Case 3	Single transmission $S \to D$
Case 4	Single transmission $S \to U$
Case 5	Single transmission $B \to U$

A. BOTH THE USERS TREAT INTERFERENCE AS NOISE

Depending on the set of active links, we investigate the success probabilities in five different cases, as described in Table 2.

1) CASE 1

When there are two active links: $S \to D$ and $S \to U$, the event of successful decoding at D with secrecy constraint is defined as

$$\mathcal{D}^{\star}_{SD/SD,SU} = \left\{\mathrm{SINR}_{SD/SD,SU} \geq \theta_D, \mathrm{SINR}_{SE/SD,SU} < \theta_D\right\}, \qquad (3)$$

where θ_D is the SINR threshold for successful decoding. In this work, we consider packet with finite length and hence, SINR based secrecy metric is used [21], [34]–[36]. The commonly used metrics such as secrecy capacity, secrecy outage probability and ergodic secrecy rate are difficult to compute when length of the packet is short. The SINR based secrecy metric can take account of decoding ability of the users and can help to analyze the scenarios where the users or eavesdropper have varied decoding ability.

From (1), we have the success probability $\mathcal{P}(\mathcal{D}^{\star}_{SD/SD,SU})$ given by

$$\begin{aligned}
&\mathcal{P}(\mathcal{D}^{\star}_{SD/SD,SU})\\
&= \mathcal{P}\left\{\frac{P_{SD}|h_{SD}|^2\, r_{SD}^{-\gamma}}{1 + P_{SU}|h_{SD}|^2\, r_{SD}^{-\gamma}} \geq \theta_D, \frac{P_{SD}|h_{SE}|^2\, r_{SE}^{-\gamma}}{1 + P_{SU}|h_{SE}|^2\, r_{SE}^{-\gamma}} < \theta_D\right\}\\
&= \mathcal{P}\left\{(P_{SD} - \theta_D P_{SU})|h_{SD}|^2\, r_{SD}^{-\gamma} \geq \theta_D\right\}\\
&\quad \times \mathcal{P}\left\{(P_{SD} - \theta_D P_{SU})|h_{SE}|^2 r_{SE}^{-\gamma} < \theta_D\right\}\\
&= \exp\left(-\frac{\theta_D r_{SD}^{\gamma}}{P_{SD} - \theta_D P_{SU}}\right)\left[1 - \exp\left(-\frac{\theta_D r_{SE}^{\gamma}}{P_{SD} - \theta_D P_{SU}}\right)\right].
\end{aligned} \qquad (4)$$

From (4), we have that the event $\mathcal{D}^{\star}_{SD/SD,SU}$ occurs with non-zero probability if $\frac{P_{SD}}{P_{SU}} > \theta_D$.

The event of successful decoding at U is defined as

$$\mathcal{D}_{SU/SD,SU} = \left\{\mathrm{SINR}_{SU/SD,SU} \geq \theta_U\right\}, \qquad (5)$$

[1] The framework developed here can also be extended for the scenario where eavesdropper can also perform successive decoding using the SINR based metric.

and the success probability of this event is

$$\begin{aligned}\mathcal{P}(\mathcal{D}_{SU/SD,SU}) &= \mathcal{P}\left\{\frac{P_{SU}|h_{SU}|^2 r_{SU}^{-\gamma}}{1+P_{SD}|h_{SU}|^2 r_{SU}^{-\gamma}} \geq \theta_U\right\}\\ &= \exp\left(-\frac{\theta_U r_{SU}^{\gamma}}{P_{SU}-\theta_U P_{SD}}\right) \qquad (6)\end{aligned}$$

From (6) we get that the event $\mathcal{D}_{SU/SD,SU}$ occurs with non-zero probability if $\frac{P_{SU}}{P_{SD}} > \theta_U$.

2) CASE 2

In this case, the access point S transmits to the dedicated user D and the base station B transmits to the non-dedicated user U, while the eavesdropper E tries to decode the message from S to D.

The event of successful decoding is defined by

$$\mathcal{D}^{\star}_{SD/SD,BU} = \left\{\text{SINR}_{SD/SD,BU} \geq \theta_D, \text{SINR}_{SE/SD,BU} < \theta_D\right\}. \qquad (7)$$

Similar to (4), the success probability in this case is given by

$$\begin{aligned}\mathcal{P}(\mathcal{D}^{\star}_{SD/SD,BU}) &= \mathcal{P}\left\{\frac{P_{SD}|h_{SD}|^2 r_{SD}^{-\gamma}}{1+P_B|h_{BD}|^2 r_{BD}^{-\gamma}} \geq \theta_D, \frac{P_{SD}|h_{SE}|^2 r_{SE}^{-\gamma}}{1+P_B|h_{BE}|^2 r_{BE}^{-\gamma}} < \theta_D\right\}\\ &= \exp\left(-\frac{\theta_D r_{SD}^{\gamma} r_{BD}^{\gamma}}{P_{SD} r_{BD}^{\gamma}-\theta_D P_B r_{SD}^{\gamma}}\right)\\ &\quad \times\left\{1-\exp\left(-\frac{\theta_D r_{SE}^{\gamma} r_{BE}^{\gamma}}{P_{SD} r_{BE}^{\gamma}-\theta_D P_B r_{SE}^{\gamma}}\right)\right\}. \qquad (8)\end{aligned}$$

From (8), we have that the event $\mathcal{D}^{\star}_{SD/SD,BU}$ occurs with non-zero probability for

$$\frac{P_{SD}}{P_B} > \max\left\{\frac{r_{SD}^{\gamma}}{r_{BD}^{\gamma}}\theta_D, \frac{r_{SE}^{\gamma}}{r_{BE}^{\gamma}}\theta_D\right\}. \qquad (9)$$

The event of successful decoding at U is defined by

$$\mathcal{D}_{BU/SD,BU} = \left\{\text{SINR}_{BU/SD,BU} \geq \theta_U\right\}, \qquad (10)$$

and the success probability is given by

$$\begin{aligned}\mathcal{P}(\mathcal{D}_{BU/SD,BU}) &= \mathcal{P}\left\{\frac{P_B|h_{BU}|^2 r_{BU}^{-\gamma}}{1+P_{SD}|h_{SU}|^2 r_{SU}^{-\gamma}} \geq \theta_U\right\},\\ &= \exp\left(-\frac{\theta_U}{P_B r_{BU}^{-\gamma}-\theta_U P_{SD} r_{SU}^{-\gamma}}\right). \qquad (11)\end{aligned}$$

From (11), we have that the event $\mathcal{D}_{BU/SD,BU}$ occurs with non-zero probability for

$$\frac{P_B}{P_{SD}} > \frac{r_{SU}^{-\gamma}}{r_{BU}^{-\gamma}}\theta_U. \qquad (12)$$

3) CASE 3

When there is a single transmission $S \rightarrow D$, the event of successful decoding at D is defined by

$$\mathcal{D}^{\star}_{SD/SD} = \left\{\text{SNR}_{SD/SD} \geq \theta_D, \text{SNR}_{SE/SD} < \theta_D\right\}. \qquad (13)$$

The success probability $\mathcal{P}(\mathcal{D}^{\star}_{SD/SD})$ is given by

$$\begin{aligned}&\mathcal{P}(\mathcal{D}^{\star}_{SD/SD})\\ &= \mathcal{P}\left\{P_{SD}|h_{SD}|^2 r_{SD}^{-\gamma} \geq \theta_D, P_{SD}|h_{SE}|^2 r_{SE}^{-\gamma} < \theta_D\right\}\\ &= \exp\left(-\frac{\theta_D r_{SD}^{\gamma}}{P_{SD}}\right)\left[1-\exp\left(-\frac{\theta_D r_{SE}^{\gamma}}{P_{SD}}\right)\right]. \qquad (14)\end{aligned}$$

4) CASE 4

There is a single transmission $S \rightarrow U$. The dedicated user D is not served. The event of successful decoding at U is defined by

$$\mathcal{D}_{SU/SU} = \left\{\text{SNR}_{SU/SU} \geq \theta_U\right\}, \qquad (15)$$

and the success probability is given by

$$\mathcal{P}(\mathcal{D}_{SU/SU}) = \mathcal{P}\left\{P_{SU}|h_{SU}|^2 r_{SU}^{-\gamma} \geq \theta_U\right\} = \exp\left(-\frac{\theta_U r_{SU}^{\gamma}}{P_{SU}}\right). \qquad (16)$$

5) CASE 5

There is a single transmission $B \rightarrow U$. The dedicated user D is not served. The event successful decoding at U is defined by

$$\mathcal{D}_{BU/BU} = \left\{\text{SNR}_{BU/BU} \geq \theta_U\right\}, \qquad (17)$$

and the success probability becomes:

$$\mathcal{P}(\mathcal{D}_{BU/BU}) = \mathcal{P}\left\{P_B|h_{BU}|^2 r_{BU}^{-\gamma} \geq \theta_U\right\} = \exp\left(-\frac{\theta_U r_{BU}^{\gamma}}{P_B}\right). \qquad (18)$$

B. USER D PERFORMS SUCCESSIVE DECODING

In this scenario, the dedicated user D needs to decode the intended message for U first, in order to remove it from the received signal and decode its own message. The eavesdropper E always treats interference as noise. Depending on the set of active links, we derive the success probabilities as follows.

1) CASE 1

When there are two active links: $S \rightarrow D$ and $S \rightarrow U$, the event of successful decoding at D with secrecy constraint is defined as

$$\mathcal{D}^{\star}_{SD/SD,SU} = \left\{ \frac{P_{SU}|h_{SD}|^2 r_{SD}^{-\gamma}}{1 + P_{SD}|h_{SD}|^2 r_{SD}^{-\gamma}} \geq \theta_U, P_{SD}|h_{SD}|^2 r_{SD}^{-\gamma} \geq \theta_D, \frac{P_{SD}|h_{SE}|^2 r_{SE}^{-\gamma}}{1 + P_{SU}|h_{SE}|^2 r_{SE}^{-\gamma}} < \theta_D \right\}, \quad (19)$$

where θ_U is the SINR threshold for successfully decoding the message intended for user U.

The success probability of the event can be expressed as

$$\mathcal{P}(\mathcal{D}^{\star}_{SD/SD,SU}) = \exp\left(-\max\left\{\frac{\theta_U r_{SD}^{\gamma}}{P_{SU} - \theta_U P_{SD}}, \frac{\theta_D r_{SD}^{\gamma}}{P_{SD}}\right\}\right) \times \left\{1 - \exp\left(-\frac{\theta_D r_{SE}^{\gamma}}{P_{SD} - \theta_D P_{SU}}\right)\right\}. \quad (20)$$

From (20), we have that the event $\mathcal{D}^{*}_{SD/SD,SU}$ occurs with non-zero probability if

$$\frac{P_{SU}}{P_{SD}} > \theta_U \quad \text{and} \quad \frac{P_{SD}}{P_{SU}} > \theta_D. \quad (21)$$

For the non-dedicated user U, the probability of successful decoding is the same as in (6).

2) CASE 2

When there are two transmission links $S \rightarrow D$ and $B \rightarrow U$, the event of successful decoding with secrecy constraint at D is defined as

$$\mathcal{D}^{\star}_{SD/SD,BU} = \left\{ \frac{P_B|h_{BD}|^2 r_{BD}^{-\gamma}}{1 + P_{SD}|h_{SD}|^2 r_{SD}^{-\gamma}} \geq \theta_U, P_{SD}|h_{SD}|^2 r_{SD}^{-\gamma} \geq \theta_D, \frac{P_{SD}|h_{SE}|^2 r_{SE}^{-\gamma}}{1 + P_B|h_{BE}|^2 r_{BE}^{-\gamma}} < \theta_D \right\}. \quad (22)$$

The success probability is given by

$$\mathcal{P}(\mathcal{D}^{\star}_{SD/SD,BU}) = \exp\left(- \max\left\{\frac{\theta_U}{P_B r_{BD}^{-\gamma} - \theta_U P_{SD} r_{SD}^{-\gamma}}, \frac{\theta_D r_{SD}^{\gamma}}{P_{SD}}\right\}\right) \times \left[1 - \exp\left(- \frac{\theta_D r_{SE}^{\gamma} r_{BE}^{\gamma}}{P_{SD} r_{BE}^{\gamma} - \theta_D P_B r_{SE}^{\gamma}}\right)\right]. \quad (23)$$

From (23), we have that the event $\mathcal{D}^{*}_{SD/SD,SU}$ occurs with non-zero probability if

$$\frac{P_B}{P_{SD}} > \frac{r_{SD}^{-\gamma}}{r_{BD}^{-\gamma}} \theta_U \quad \text{and} \quad \frac{P_{SD}}{P_B} > \frac{r_{SE}^{\gamma}}{r_{BE}^{\gamma}} \theta_D. \quad (24)$$

For the non-dedicated user U, the success probability is the same as in (11).

The success probabilities for the cases $3 - 5$ are the same as in Section III-A.

C. USER U PERFORMS SUCCESSIVE DECODING

In this case, the non-dedicated user U performs successive decoding. It first attempts to decode the information transmitted from S to D, then removes it from the received signal, and proceeds to decode its own message either from S or B.

1) CASE 1

Similar to Section III-A, the event of successful decoding at user U is defined as

$$\mathcal{D}_{SU/SD,SU} = \left\{ \frac{P_{SD}|h_{SU}|^2 r_{SU}^{-\gamma}}{1 + P_{SU}|h_{SU}|^2 r_{SU}^{-\gamma}} \geq \theta_D, P_{SU}|h_{SU}|^2 r_{SU}^{-\gamma} \geq \theta_U \right\}. \quad (25)$$

The success probability is given by

$$\mathcal{P}(\mathcal{D}_{SU/SD,SU}) = \exp\left(- \max\left\{\frac{\theta_D r_{SU}^{\gamma}}{P_{SD} - \theta_D P_{SU}}, \frac{\theta_U r_{SU}^{\gamma}}{P_{SU}}\right\}\right). \quad (26)$$

From (26), we have that the event of successful decoding at U occurs with non-zero probability if $\frac{P_{SD}}{P_{SU}} > \theta_D$. For the dedicated user D, the probability of successful decoding with secrecy constraint is the same as in (4).

2) CASE 2

The event of successful decoding at user U is defined as

$$\mathcal{D}_{BU/SD,BU} = \left\{ \frac{P_{SD}|h_{SU}|^2 r_{SU}^{-\gamma}}{1 + P_B|h_{BU}|^2 r_{BU}^{-\gamma}} \geq \theta_D, P_B|h_{BU}|^2 r_{BU}^{-\gamma} \geq \theta_U \right\}. \quad (27)$$

The success probability is given by

$$\mathcal{P}(\mathcal{D}_{BU/SD,BU}) = \exp\left(- \max\left\{\frac{\theta_D}{P_{SD} r_{SU}^{-\gamma} - \theta_D P_B r_{BU}^{-\gamma}}, \frac{\theta_U r_{BU}^{\gamma}}{P_B}\right\}\right). \quad (28)$$

From (28), the successful decoding event occurs with non-zero probability if

$$\frac{P_{SD}}{P_B} > \frac{r_{SU}^{-\gamma}}{r_{BU}^{-\gamma}} \theta_D. \quad (29)$$

For user D, the probability of successful decoding with secrecy constraint is the same as in (8).

The success probabilities for the cases $3 - 5$ are the same as in Section III-A.

IV. THROUGHPUT AND DELAY ANALYSIS

In this section, we derive the throughput and delay performance of the considered network.

A. NETWORK THROUGHPUT

Recall that the queue of the access point S contains information packets intended for the dedicated user D as described in Section II. When the queue is stable, the throughput of user D is equal to its arrival rate λ. The queue stability condition is satisfied if the arrival rate is smaller than the service rate of S.

The average service rate of the link $S \rightarrow D$ is obtained as

$$\mu = q_S(1-q_U)\mathcal{P}(\mathcal{D}^{\star}_{SD/SD}) + q_S q_U p_h \mathcal{P}(\mathcal{D}^{\star}_{SD/SD,SU}) + q_S q_U p_m \alpha \mathcal{P}(\mathcal{D}^{\star}_{SD/SD,BU}) + q_S q_U p_m (1-\alpha)\mathcal{P}(\mathcal{D}^{\star}_{SD/SD}), \tag{30}$$

which considers all the cases described in Table 2.

The queue at S is stable if $\lambda < \mu$. When the queue stability condition is satisfied, the probability that the queue is non-empty is

$$\mathcal{P}(Q \neq 0) = \frac{\lambda}{\mu}, \tag{31}$$

where μ is given by (30).

The average throughput of the non-dedicated user U is

$$T_U = q_S \frac{\lambda}{\mu} q_U \left[p_h \mathcal{P}(\mathcal{D}_{SU/SD,SU}) + p_m \alpha \mathcal{P}(\mathcal{D}_{BU/SD,BU})\right] + \left(1 - q_S \frac{\lambda}{\mu}\right) q_U \left[p_h \mathcal{P}(\mathcal{D}_{SU/SU}) + p_m \alpha \mathcal{P}(\mathcal{D}_{BU/BU})\right]. \tag{32}$$

B. DELAY

When the queue at S is stable, the average delay experienced by the dedicated user D can be obtained as

$$D_D = \frac{1}{\mu - \lambda}(1-\lambda) + \frac{1}{\mu}, \tag{33}$$

where the first term is the queueing delay and the second term is the transmission delay.

For the non-dedicated user U, the delay is only characterized by the transmission delay, considering all the possible cases for finding the content, as follows

$$\begin{aligned} D_U = {} & p_h q_s \frac{\lambda}{\mu} \mathcal{P}(\mathcal{D}_{SU/SD,SU}) + p_h \left(1 - q_s \frac{\lambda}{\mu}\right) \mathcal{P}(\mathcal{D}_{SU/SU}) \\ & + p_h q_s \frac{\lambda}{\mu} [1 - \mathcal{P}(\mathcal{D}_{SU/SD,SU})](1 + D_S) \\ & + p_h \left(1 - q_s \frac{\lambda}{\mu}\right) [1 - \mathcal{P}(\mathcal{D}_{SU/SU})](1 + D_S) \\ & + p_m q_s \frac{\lambda}{\mu} \alpha \mathcal{P}(\mathcal{D}_{BU/SD,BU}) + p_m \left(1 - q_s \frac{\lambda}{\mu}\right) \alpha \mathcal{P}(\mathcal{D}_{BU/BU}) \\ & + p_m q_s \frac{\lambda}{\mu} [1 - \alpha \mathcal{P}(\mathcal{D}_{BU/SD,BU})](1 + D_B) \\ & + p_m \left(1 - q_s \frac{\lambda}{\mu}\right) [1 - \alpha \mathcal{P}(\mathcal{D}_{BU/BU})](1 + D_B), \end{aligned} \tag{34}$$

where D_S and D_B represent the transmission delay from the access point S and from the base station B, respectively. The transmission delay is inversely proportional to the average success probability. We have

$$D_S = \frac{1}{q_s \frac{\lambda}{\mu} \mathcal{P}(\mathcal{D}_{SU/SD,SU}) + [1 - q_s \frac{\lambda}{\mu}] \mathcal{P}(\mathcal{D}_{SU/SU})}, \tag{35}$$

$$D_B = \frac{1}{q_s \frac{\lambda}{\mu} \alpha \mathcal{P}(\mathcal{D}_{BU/SD,BU}) + [1 - q_s \frac{\lambda}{\mu}] \alpha \mathcal{P}(\mathcal{D}_{BU/BU})}. \tag{36}$$

V. THROUGHPUT AND DELAY PERFORMANCE OPTIMIZATION

In this section, we formulate two optimization problems which jointly consider the throughput and delay performance of both users D and U. The variables to be determined are the allocated transmit powers P_{SD} and P_{SU} at the access point S. The transmit power of the base station, P_B, is considered fixed.

A. THROUGHPUT OPTIMIZATION WITH DELAY CONSTRAINTS

We aim at finding the optimal power allocation that maximizes the average throughput perceived by the non-dedicated user U, while satisfying the delay requirement at the dedicated user D. The problem is defined as follows

$$\begin{aligned} \max_{\mathbf{P}=[P_{SD},P_{SU}]} \quad & T_U(\mathbf{P}) \\ \text{subject to} \quad & \lambda < \mu, \\ & P_{SD} + P_{SU} = P_{\max}, \\ & D_D(\mathbf{P}) \leq D_{Dmax}. \end{aligned} \tag{37}$$

Here, T_U and μ are given in (32) in (30), respectively. The first constraint ensures that the queue at S is stable. The second constraint comes from the limited total transmit power of S. The last constraint sets a maximum tolerable delay D_{Dmax} experienced by the dedicated user D.

B. DELAY OPTIMIZATION WITH THROUGHPUT CONSTRAINT

The objective of this optimization problem is to minimize the average delay D_D perceived by the dedicated user D, while achieving a minimum throughput for the non-dedicated user U.

$$\begin{aligned} \min_{\mathbf{P}=[P_{SD},P_{SU}]} \quad & D_D(\mathbf{P}) \\ \text{subject to} \quad & \lambda < \mu, \\ & P_{SD} + P_{SU} = P_{\max}, \\ & T_U(\mathbf{P}) \geq T_{Umin} \end{aligned} \tag{38}$$

where D_D is given in (33), and T_{Umin} represents the lower limit of the average throughput for user U.

The optimization problems in (37) and (38) are nonlinear due to the exponential factors in the equations for the success probabilities, thus, we resort to numerical optimization since we cannot derive a closed form expression for the considered problem.

VI. NUMERICAL RESULTS

In this section, we present the numerical results for the system performance analysis covered in the previous sections. We use path-loss exponents $\gamma_S = 2$ for all the links from the helper S ($S \rightarrow D$, $S \rightarrow U$, $S \rightarrow E$) and $\gamma_B = 4$ for all the links from the base station B ($B \rightarrow D$, $B \rightarrow U$, $B \rightarrow E$). For simplicity, we normalize the noise power to one, and set the transmission power level of the helper as

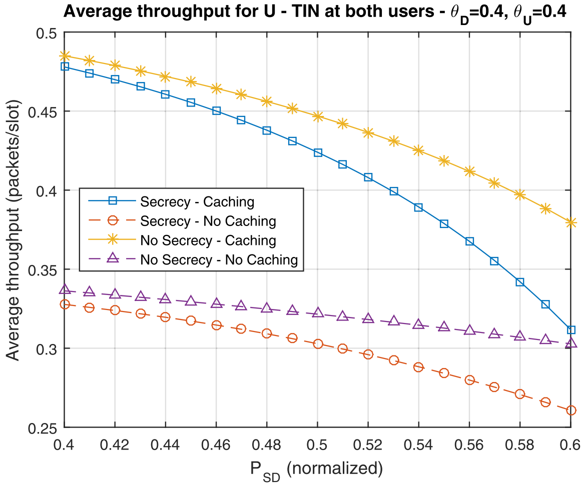

FIGURE 2. Average throughput (packets/slot) for user U T_U versus transmit power from helper S to the dedicated user D (normalized). In this setup both users treat interference as noise (TIN). The plot was generated with: $\lambda = 0.4$, $q_S = 0.9$, $q_U = 0.8$, $p_h = 0.6$, $\alpha = 0.7$, $\theta_D = 0.4$ and $\theta_U = 0.4$.

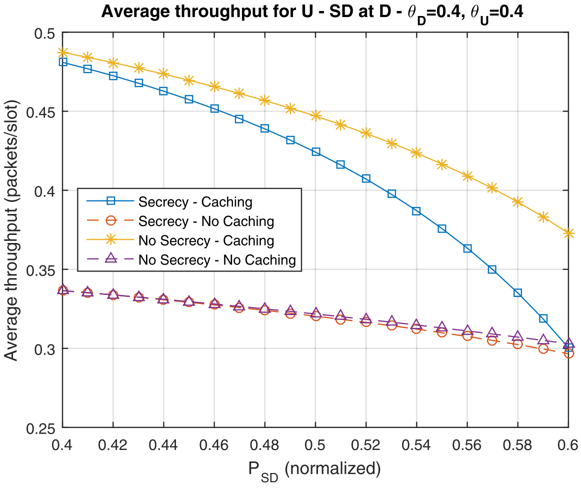

FIGURE 3. Average throughput (packets/slot) for user U T_U versus transmit power from helper S to the dedicated user D (normalized). In this setup user D performs successive decoding (SD). The plot was generated with: $\lambda = 0.4$, $q_S = 0.9$, $q_U = 0.8$, $p_h = 0.6$, $\alpha = 0.7$, $\theta_D = 0.4$ and $\theta_U = 0.4$.

$P_S = P_{SD}+P_{SU} = 1000$ and P_B is restricted by the condition in (12). We assume that both D and U are within a predefined "restricted" area with a radius of 30 m, e.g., $r_{SD} = 10$ m and $r_{SU} = 20$ m. The distance between the eavesdropper E and the helper S is $r_{SE} = 30$ m. The distances from the base station B to the dedicated user D, the non-dedicated user U, and the eavesdropper E are set to be 1000 m ($r_{BD} = r_{BU} = r_{BE} = 1000$ m).

Initially, we will try to determine how caching at the edge-helper affects the overall performance of the wireless system. This will be captured through the throughput and delay performance for each user comparing the cases with different decoding schemes. It is important to clarify that caching enhances the performance and provides us with higher overall system performance. Afterwards, we will identify the decoding scheme TIN or SD applied at each user that improves the performance by solving the optimization problems that were previously set. Finally, we will illustrate how the system performance is affected by various network parameters concerning different power allocations.

Based on Figure 2 we observe the cases with and without secrecy for this scenario where both users treat interference as noise (TIN). For lower values of power allocated to the S-D transmission P_{SD}, caching increases the achieved throughput for user U and performs almost similarly with the system without any secrecy constraints for lower values of P_{SD}. As the power level P_{SD} increases (P_{SU} decreases) there is a steeper decrease in throughput performance for user U in the case of the system with secrecy constraints due to two main factors. First, the decrease in P_{SU} deteriorates the decoding capability of user U, and secondly as P_{SD} increases, the eavesdropper will have higher probability to decode the transmission which eventually will decrease the performance of the system with secrecy constraints.

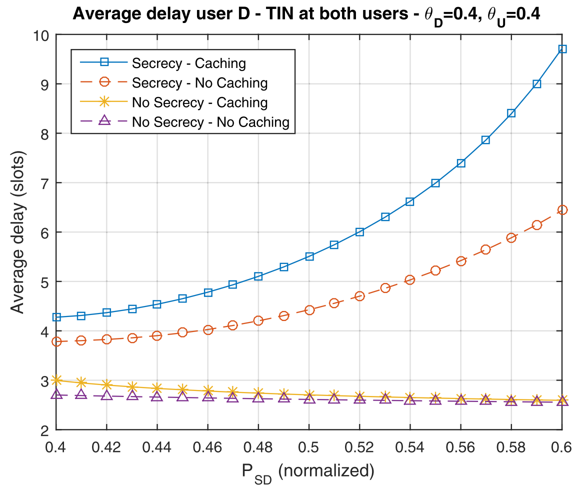

FIGURE 4. Average delay (slots) for user D D_D versus transmit power from helper S to the dedicated user D (normalized). In this setup both users treat interference as noise (TIN). The plot was generated with: $\lambda = 0.4$, $q_S = 0.9$, $q_U = 0.8$, $p_h = 0.6$, $\alpha = 0.7$, $\theta_D = 0.4$ and $\theta_U = 0.4$.

Similar results where SD was only applied at the dedicated user D are illustrated in Figure 3. Caching still increased the performance in terms of user's U average throughput, though we can point out that SD at user D eliminates the performance deterioration for user U with the with secrecy constraints compared to no secrecy especially for low P_{SD} values. This is important in the case of low hit rate probability p_h at the helper's cache i.e. not efficient caching scheme, which points out the SD as a more robust decoding scheme compared to the TIN.

Another useful system performance metric, namely the average delay for the dedicated user D was analyzed in Figure 4 again versus the normalized S to D transmission power level with both users applying TIN for decoding. As expected, the delay is increased for D when the other user

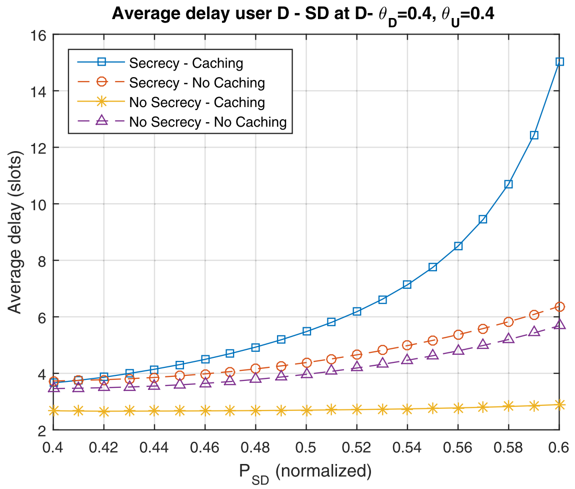

FIGURE 5. **Average delay (slots) for user *D* D_D versus transmit power from helper *S* to the dedicated user *D* (normalized). In this setup user *D* performs successive decoding (SD). The plot was generated with: $\lambda = 0.4$, $q_S = 0.9$, $q_U = 0.8$, $p_h = 0.6$, $\alpha = 0.7$, $\theta_D = 0.4$ and $\theta_U = 0.4$.**

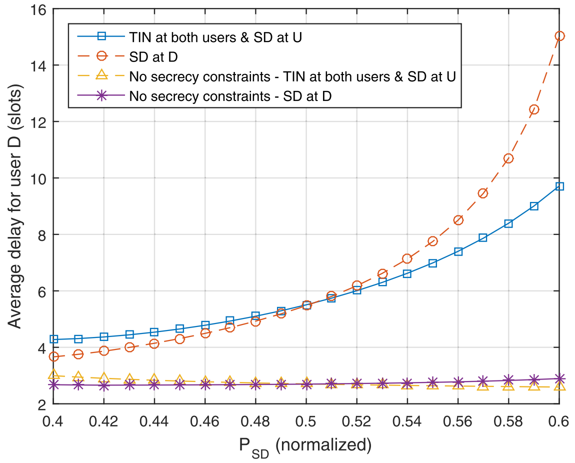

FIGURE 6. **Average delay (slots) for user *D* D_D versus transmit power from helper *S* to the dedicated user *D* (normalized). The plot was generated with: $\lambda = 0.4$, $q_S = 0.9$, $q_U = 0.8$, $p_h = 0.6$, $\alpha = 0.7$, $\theta_D = 0.4$ and $\theta_U = 0.4$.**

U is served by the helper. This is the effect of caching at the helper *S* on *D*'s average delay performance. For lower values of P_{SD}, and thus higher P_{SU}, the system performance with secrecy constraints is very close for the cases with and without caching. This is the result of fewer re-transmissions from helper *S* to the non-dedicated user *U* in the case of content cached at helper's cache.

In Figure 5 similar to the previous figure we observe the average delay for user *D* with the only difference that user *D* applies SD for decoding. From these results, we conclude that for the secrecy constraints and caching scenario the average delay performance for *D* is lower than the case where both users apply TIN. Caching at the helper *S* can even further reduce the effects of secrecy constraints for lower P_{SD} values providing similar performance for the cases with and without secrecy constraints.

A comparison of the two decoding schemes TIN and SD is illustrated in Figure 6 where is highlighted the fact that for low P_{SD} power levels, applying successive decoding at *D* surpasses the performance of the system where both users treat interference as noise. This happens because for low P_{SD} levels thus higher P_{SU} levels there are fewer re-transmissions for the *S*-*U* messages offering in parallel increased secrecy as the eavesdropper *E* receives lower power transmissions for decoding. In both cases, with and without secrecy constraints, applying SD at *D* leads to better delay performance for *D* for lower P_{SD} levels as there are fewer re-transmissions for both *S*-*D* and *S*-*U* messages making it an ideal setup irrespective the secrecy constraints.

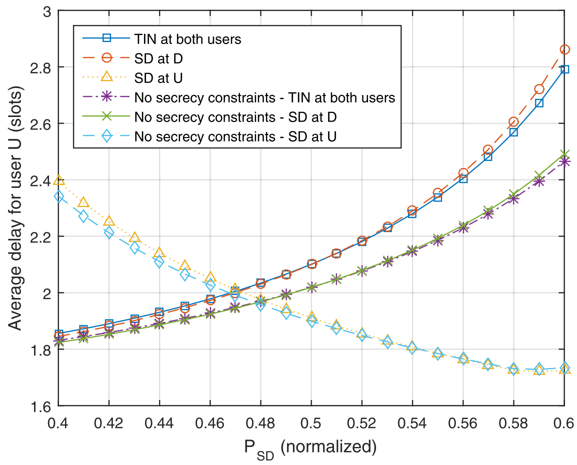

FIGURE 7. **Average delay (slots) for user *U* D_U versus transmit power from helper *S* to the dedicated user *D* (normalized). The plot was generated with: $\lambda = 0.4$, $q_S = 0.9$, $q_U = 0.8$, $p_h = 0.6$, $\alpha = 0.7$, $\theta_D = 0.4$ and $\theta_U = 0.4$.**

Focusing on user *U*, Figure 7 gives us an insight into the effects of the two decoding schemes and secrecy constraints on *U*'s average delay performance. Although there is a better overall performance for *U*'s average delay at higher P_{SD} levels applying SD at *U* this is not preferred as it will deteriorate the dedicated user *D*'s performance. The average delay performance for *U* applying SD at *D* for lower power levels of P_{SD} is close to the average delay performance for higher P_{SD} values (low P_{SU}) when applying SD at *U* thus it is the preferred setup for achieving better performance for both users.

Similarly, in Figure 8 we observe the average throughput for user *U* versus the normalized P_{SD} levels that follow the trend (inverse) of Figure 7, where for lower P_{SD} power levels the performance reaches a pick. As explained before, applying SD at *D* results in a better performance than TIN for both users achieving a performance closer to the no secrecy constraints scenario. Applying SD at *D* results in better performance than TIN because *D* can easily decode the *S*-*U* packets and then use this to decode the wicker *S*-*D* packet transmission thus increasing the security of the link.

It is also important to highlight the average service rate performance that represents the average throughput of D presented in Figure 10. In both cases, with and without any secrecy constraints, applying SD at *D* results in a higher service rate although with no secrecy constraints applying

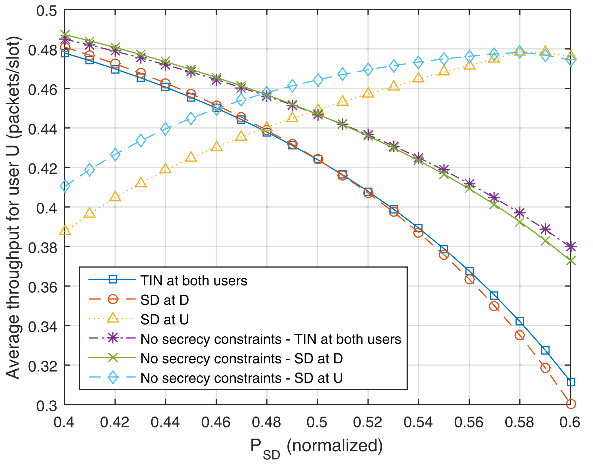

FIGURE 8. Average throughput (packets/slot) for user U T_U versus transmit power from helper S to the dedicated user D (normalized). The plot was generated with: $\lambda = 0.4$, $q_S = 0.9$, $q_U = 0.8$, $p_h = 0.6$, $\alpha = 0.7$, $\theta_D = 0.4$ and $\theta_U = 0.4$.

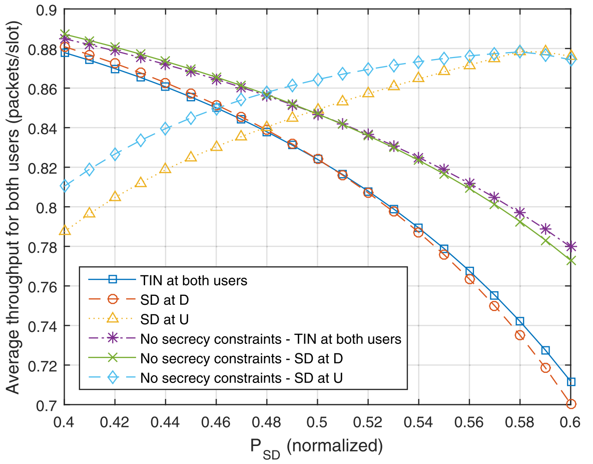

FIGURE 9. Cumulative average throughput (packets/slot) for both users D and U versus transmit power from helper S to the dedicated user D (normalized). The plot was generated with: $\lambda = 0.4$, $q_S = 0.9$, $q_U = 0.8$, $p_h = 0.6$, $\alpha = 0.7$, $\theta_D = 0.4$ and $\theta_U = 0.4$.

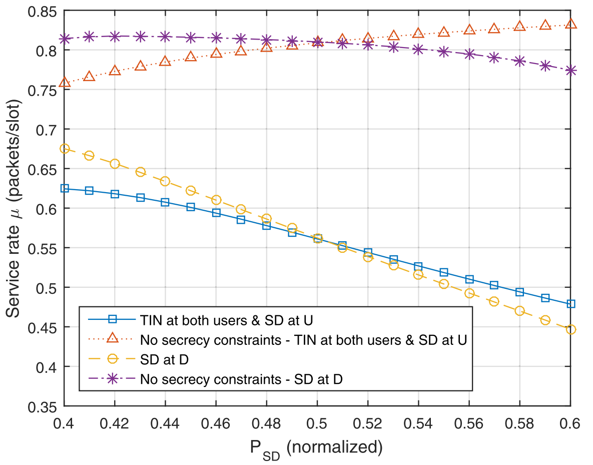

FIGURE 10. Service rate μ (packets/slot) at the helper S versus transmit power to the dedicated user D (normalized). The plot was generated with: $\lambda = 0.4$, $q_S = 0.9$, $q_U = 0.8$, $p_h = 0.6$, $\alpha = 0.7$, $\theta_D = 0.4$ and $\theta_U = 0.4$.

TIN or SD at U only outperforms SD at D for higher P_{SD} values. This happens as there is more error-free decoding at U for this power range thus less re-transmission resulting in less congestion at S.

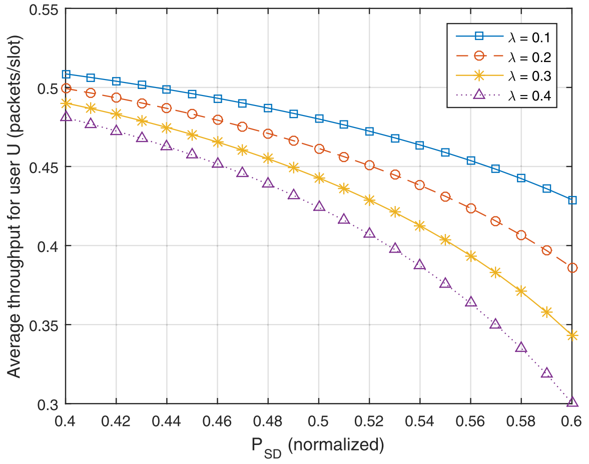

FIGURE 11. Average throughput (packets/slot) for user U T_U versus transmit power from helper S to the dedicated user D (normalized) applying SD at D for variable arrival rates λ at S. The plot was generated with: $q_S = 0.9$, $q_U = 0.8$, $p_h = 0.6$, $\alpha = 0.7$, $\theta_D = 0.4$ and $\theta_U = 0.4$.

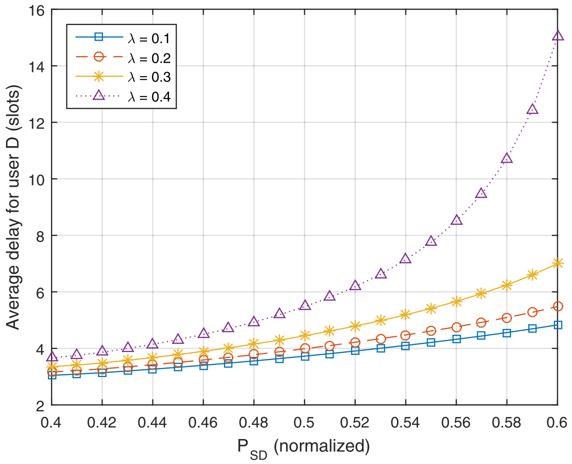

FIGURE 12. Average delay (slots) for user D D_D versus transmit power from helper S to the dedicated user D (normalized) applying SD at D for variable arrival rates λ at S. The plot was generated with: $q_S = 0.9$, $q_U = 0.8$, $p_h = 0.6$, $\alpha = 0.7$, $\theta_D = 0.4$ and $\theta_U = 0.4$.

Moving on with network characteristics and their significance in the system's performance, both Figure 11 and Figure 12 demonstrate U's throughput and D's delay performance respectively versus the P_{SD} power levels for different arrival rates λ at the helper S while applying SD at the dedicated user D. We observe that for lower transmission power levels P_{SD} we achieve the best performance and that for different arrival rates the performance is not varying significantly as it is for higher P_{SD} power levels. This means that by setting lower P_{SD} transmission power levels the performance can be unaffected when the arrival rate, thus packets sent to dedicated user D, is increased.

The next three figures (Figure 13-Figure 15) illustrate how the hit rate p_h at the helper S (probability U requests content from the helper S) affects the performance. Again,

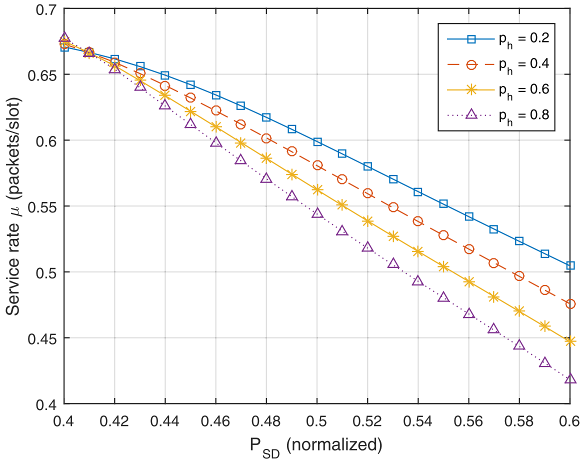

FIGURE 13. Service rate μ (packets/slot) versus transmit power from helper S to the dedicated user D (normalized) applying SD at D for variable hit rates p_h at S. The plot was generated with: $\lambda = 0.4$, $q_S = 0.9$, $q_U = 0.8$, $\alpha = 0.7$, $\theta_D = 0.4$ and $\theta_U = 0.4$.

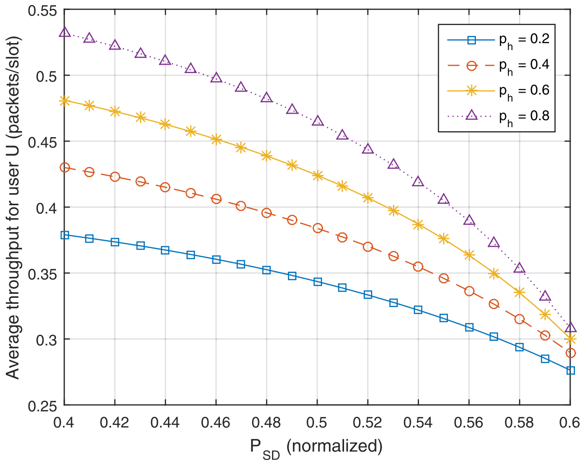

FIGURE 14. Average throughput (packets/slot) for user U T_U versus transmit power from helper S to the dedicated user D (normalized) applying SD at D for variable hit rates p_h at S. The plot was generated with: $\lambda = 0.4$, $q_S = 0.9$, $q_U = 0.8$, $\alpha = 0.7$, $\theta_D = 0.4$ and $\theta_U = 0.4$.

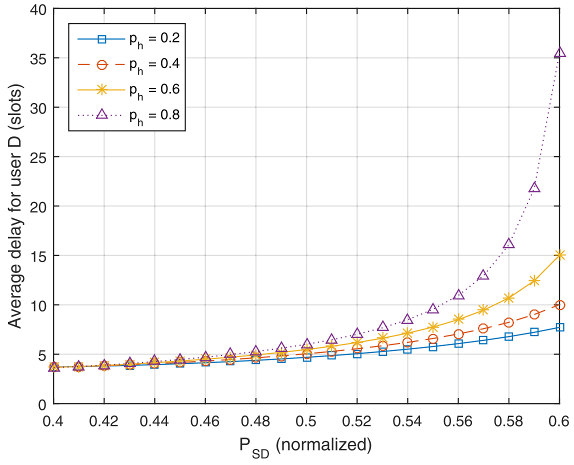

FIGURE 15. Average delay (slots) for user D D_D versus transmit power from helper S to the dedicated user D (normalized) applying SD at D for variable hit rates p_h at S. The plot was generated with: $\lambda = 0.4$, $q_S = 0.9$, $q_U = 0.8$, $\alpha = 0.7$, $\theta_D = 0.4$ and $\theta_U = 0.4$.

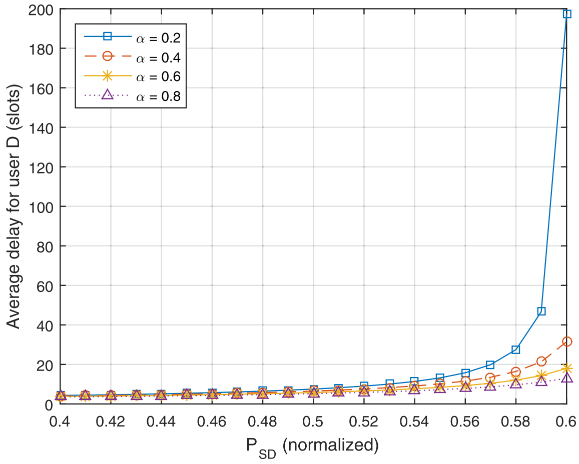

FIGURE 16. Average delay (slots) for user D D_D versus transmit power from helper S to the dedicated user D (normalized) applying SD at D for variable network availability α. The plot was generated with: $\lambda = 0.4$, $q_S = 0.9$, $q_U = 0.8$, $p_h = 0.6$, $\theta_D = 0.4$ and $\theta_U = 0.4$.

TABLE 3. Optimal average throughput (packets/slot) for user U and delay (slots) for user D values for variable arrival rates λ at S for TIN at both users, SD at D and U (SD-D, SD-U) and delay (D_D) and throughput (T_U) constraints. $p_h = 0.6$ $q_S = 0.9$, $q_U = 0.8$, $\alpha = 0.7$, $\theta_D = 0.4$, $\theta_U = 0.4$, $D_{Dmax} = 6$ slots, $T_{Umin} = 0.44$ packets/slot.

λ	0.1	0.2	0.3	0.4
$T^*_{U(TIN)}$	0.5	0.49	0.48	0.47
$T^*_{U(SD-D)}$	0.5	0.49	0.48	0.47
$T^*_{U(SD-U)}$	0.48	0.47	0.47	0.46
$D^*_{D(TIN)}$	3.5	3.71	4.05	4.76
$D^*_{D(SD-D)}$	3.21	3.36	3.6	4.03
$D^*_{D(SD-U)}$	3.5	3.71	4.12	5.24

we find out that for lower P_{SD} power levels while applying SD at D, we achieve the best performance in terms of service rate μ (Figure 13). Higher and lower hit rate values lead to similar performance for lower P_{SD} levels too. As expected in Figure 14 higher hit rate values result in higher average throughput for use U while for low P_{SD} values the performance difference is higher (higher P_{SU} values thus better decoding at U). Finally, Figure 15 demonstrates user D's average delay performance for various hit rate values p_h. In these results as indicated before we get optimal values (lower delay values) for low P_{SD} levels as the hit rate variations do not affect D's delay performance.

Solving the optimization problems introduced in (37) and (38), three tables were generated namely tables 3, 4, and 5. In each optimization problem producing these tables a different parameter is a variable namely the arrival rate λ, the hit rate p_h, and α. The only case that SD at U outperforms the case with SD at D in respect to the average throughput for U is that for low hit rate values p_h and that is because in this case user U is mainly receiving from the base station. In this case, the SD scheme increases it's decoding capabilities even if the B to U transmission is not very efficient. In every other

TABLE 4. Optimal average throughput (packets/slot) for user *U* and delay (slots) for user *D* values for variable hit rate p_h for TIN at both users, SD at *D* and *U* (SD-*D*, SD-*U*) and delay (D_D) and throughput (T_U) constraints. $\lambda = 0.4$ $q_S = 0.9$, $q_U = 0.8$, $\alpha = 0.7$, $\theta_D = 0.4$, $\theta_U = 0.4$, $D_{Dmax} = 6$ slots, $T_{Umin} = 0.44$ packets/slot.

p_h	**0.2**	**0.4**	**0.6**	**0.8**
$T^*_{U(TIN)}$	0.37	0.42	0.47	0.52
$T^*_{U(SD-D)}$	0.37	0.42	0.47	0.53
$T^*_{U(SD-U)}$	0.41	0.44	0.46	0.47
$D^*_{D(TIN)}$	4.49	4.63	4.76	4.93
$D^*_{D(SD-D)}$	4.24	4.13	4.03	3.93
$D^*_{D(SD-U)}$	5.68	5.82	5.24	5.13

TABLE 5. Optimal average throughput (packets/slot) for user *U* and delay (slots) for user *D* values for variable network availability α for TIN at both users, SD at *D* and *U* (SD-*D*, SD-*U*) and delay (D_D) and throughput (T_U) constraints. $\lambda = 0.4$ $q_S = 0.9$, $q_U = 0.8$, $p_h = 0.6$, $\theta_D = 0.4$, $\theta_U = 0.4$, $D_{Dmax} = 6$ slots, $T_{Umin} = 0.44$ packets/slot.

α	**0.2**	**0.4**	**0.6**	**0.8**
$T^*_{U(TIN)}$	0.38	0.41	0.45	0.49
$T^*_{U(SD-D)}$	0.38	0.42	0.45	0.49
$T^*_{U(SD-U)}$	0.28	0.36	0.43	0.48
$D^*_{D(TIN)}$	6.67	5.73	5.03	4.52
$D^*_{D(SD-D)}$	5.18	4.63	4.21	3.86
$D^*_{D(SD-U)}$	9.07	8.12	6.3	4.65

case, SD at D outperforms every other scenario and results in higher average throughput for U and lower delay for D.

VII. CONCLUSION

Throughout this work, the effect of caching on the transmission security of a helper system is studied while applying superposition coding for serving two users simultaneously with different secrecy requirements. Moreover, two distinct decoding schemes are compared namely treating interference as noise and successive decoding by introducing multiple network characteristics and caching capabilities on the helper's cache. The transmissions to a dedicated user must be kept secret inside a specific area defined by the locations of both users that are served. The initial findings of this analysis prove that caching can mitigate the effects of secrecy on the performance of the transmissions for both users that are served in this scheme. Successive decoding at the dedicated user while applying the lowest power allocation satisfying the stability condition, offers better overall performance compared to the treat as interference decoding scheme. Summarizing our findings, the optimal performance in terms of throughput and delay for both users, while keeping a dedicated link secure, is achieved when applying successive decoding at the dedicated user, and allocating the minimum power within the stability condition for the packets intended for that dedicated user.

REFERENCES

[1] G. Smpokos, N. Pappas, Z. Chen, and P. Mohapatra, "Wireless caching helper system with heterogeneous traffic and secrecy constraints," in *Proc. IEEE 20th Int. Workshop Signal Process. Adv. Wireless Commun. (SPAWC)*, Jul. 2019, pp. 1–5.

[2] E. Bastug, M. Bennis, and M. Debbah, "Living on the edge: The role of proactive caching in 5G wireless networks," *IEEE Commun. Mag.*, vol. 52, no. 8, pp. 82–89, Aug. 2014.

[3] N. Golrezaei, A. G. Dimakis, and A. F. Molisch, "Scaling behavior for device-to-device communications with distributed caching," *IEEE Trans. Inf. Theory*, vol. 60, no. 7, pp. 4286–4298, Jul. 2014.

[4] K. Shanmugam, N. Golrezaei, A. G. Dimakis, A. F. Molisch, and G. Caire, "FemtoCaching: Wireless content delivery through distributed caching helpers," *IEEE Trans. Inf. Theory*, vol. 59, no. 12, pp. 8402–8413, Dec. 2013.

[5] Z. Chen, J. Lee, T. Q. S. Quek, and M. Kountouris, "Cooperative caching and transmission design in cluster-centric small cell networks," *IEEE Trans. Wireless Commun.*, vol. 16, no. 5, pp. 3401–3415, May 2017.

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

[7] A. Wyner, "The wire-tap channel," *Bell Syst. Tech. J.*, vol. 54, no. 8, pp. 1334–1387, Oct. 1975.

[8] I. Csiszár and J. Korner, "Broadcast channels with confidential messages," *IEEE Trans. Inf. Theory*, vol. IT-24, no. 3, pp. 339–348, May 1978.

[9] R. Liu, I. Maric, P. Spasojević, and R. D. Yates, "Discrete memoryless interference and broadcast channels with confidential messages: Secrecy rate regions," *IEEE Trans. Inf. Theory*, vol. 54, no. 6, pp. 2493–2507, Jun. 2008.

[10] E. Ekrem and S. Ulukus, "Secrecy in cooperative relay broadcast channels," *IEEE Trans. Inf. Theory*, vol. 57, no. 1, pp. 137–155, Jan. 2011.

[11] E. Ekrem and S. Ulukus, "Effects of cooperation on the secrecy of multiple access channels with generalized feedback," in *Proc. 42nd Annu. Conf. Inf. Sci. Syst.*, Mar. 2008, pp. 791–796.

[12] O. O. Koyluoglu and H. El Gamal, "Cooperative encoding for secrecy in interference channels," *IEEE Trans. Inf. Theory*, vol. 57, no. 9, pp. 5682–5694, Sep. 2011.

[13] J. Barros and M. R. D. Rodrigues, "Secrecy capacity of wireless channels," in *Proc. IEEE Int. Symp. Inf. Theory*, Jul. 2006, pp. 356–360.

[14] A. Khisti, A. Tchamkerten, and G. W. Wornell, "Secrecy in cooperative relay broadcast channels," *IEEE Trans. Inf. Theory*, vol. 54, no. 6, pp. 2453–2469, Jun. 2008.

[15] P. K. Gopala, L. Lai, and H. E. Gamal, "On the secrecy capacity of fading channels," *IEEE Trans. Inf. Theory*, vol. 54, no. 10, pp. 4687–4698, Oct. 2008.

[16] K. Jiang, T. Jing, Z. Li, Y. Huo, and F. Zhang, "Analysis of secrecy performance in fading multiple access wiretap channel with SIC receiver," in *Proc. IEEE Conf. Comput. Commun.*, May 2017, pp. 1–9.

[17] A. S. Cohen and A. Lapidoth, "The Gaussian watermarking game," *IEEE Trans. Inf. Theory*, vol. 48, no. 6, pp. 1639–1667, Jun. 2002.

[18] P. Moulin and J. A. O'Sullivan, "Information-theoretic analysis of information hiding," *IEEE Trans. Inf. Theory*, vol. 49, no. 3, pp. 563–593, Mar. 2003.

[19] F. Verde, A. Scaglione, D. Darsena, and G. Gelli, "An Amplify-and-Forward scheme for spectrum sharing in cognitive radio channels," *IEEE Trans. Wireless Commun.*, vol. 14, no. 10, pp. 5629–5642, Oct. 2015.

[20] D. Darsena, G. Gelli, and F. Verde, "Convolutive superposition for multicarrier cognitive radio systems," *IEEE J. Sel. Areas Commun.*, vol. 34, no. 11, pp. 2951–2967, Nov. 2016.

[21] P. Mohapatra, N. Pappas, J. Lee, T. Q. S. Quek, and V. Angelakis, "Secure communications for the two-user broadcast channel with random traffic," *IEEE Trans. Inf. Forensics Security*, vol. 13, no. 9, pp. 2294–2309, Sep. 2018.

[22] Y. Liang, H. V. Poor, and L. Ying, "Secure communications over wireless broadcast networks: Stability and utility maximization," *IEEE Trans. Inf. Forensics Security*, vol. 6, no. 3, pp. 682–692, Sep. 2011.

[23] Y. Liang, H. V. Poor, and L. Ying, "Wireless broadcast networks: Reliability, security, and stability," in *Proc. Inf. Theory Appl. Workshop*, Feb. 2008, pp. 249–255.

[24] N. Karamchandani, S. Diggavi, G. Caire, and S. Shamai, "Rate and delay for coded caching with carrier aggregation," in *Proc. IEEE Int. Symp. Inf. Theory (ISIT)*, Jul. 2016, pp. 2724–2728.

[25] Y. Wang, X. Tao, X. Zhang, and G. Mao, "Joint caching placement and user association for minimizing user download delay," *IEEE Access*, vol. 4, pp. 8625–8633, 2016.

[26] F. Rezaei and B. H. Khalaj, "Stability, rate, and delay analysis of single bottleneck caching networks," *IEEE Trans. Commun.*, vol. 64, no. 1, pp. 300–313, Jan. 2016.

[27] N. Pappas, Z. Chen, and I. Dimitriou, "Throughput and delay analysis of wireless caching helper systems with random availability," *IEEE Access*, vol. 6, pp. 9667–9678, 2018.

[28] Q. Yang, H. Wang, and T. Zheng, "Delivery-secrecy tradeoff for cache-enabled stochastic networks: Content placement optimization," *IEEE Trans. Veh. Technol.*, vol. 67, no. 11, pp. 11309–11313, Sep. 2018.

[29] W. Zhao, Z. Chen, K. Li, N. Liu, B. Xia, and L. Luo, "Caching-aided physical layer security in wireless cache-enabled heterogeneous networks," *IEEE Access*, vol. 6, pp. 68920–68931, 2018.

[30] F. Shi, W. Tan, J. Xia, D. Xie, L. Fan, and X. Liu, "Hybrid cache placement for physical-layer security in cooperative networks," *IEEE Access*, vol. 6, pp. 8098–8108, 2018.

[31] N. Pappas, M. Kountouris, A. Ephremides, and V. Angelakis, "Stable throughput region of the two-user broadcast channel," *IEEE Trans. Commun.*, vol. 66, no. 10, pp. 4611–4621, Oct. 2018.

[32] B. He, X. Zhou, and A. L. Swindlehurst, "On secrecy metrics for physical layer security over quasi-static fading channels," *IEEE Trans. Wireless Commun.*, vol. 15, no. 10, pp. 6913–6924, Oct. 2016.

[33] A. E. Gamal and Y. H. Kim, *Networking Information Theory*. Cambridge, U.K.: Cambridge Univ. Press, 2011.

[34] A. K. Sadek, K. J. R. Liu, and A. Ephremides, "Cognitive multiple access via cooperation: Protocol design and performance analysis," *IEEE Trans. Inf. Theory*, vol. 53, no. 10, pp. 3677–3696, Oct. 2007.

[35] S. Sarma, S. Shukla, and J. Kuri, "Joint scheduling jamming for data secrecy in wireless networks," in *Proc. 11th Int. Symp. Workshops Model. Optim. Mobile, Ad Hoc Wireless Netw. (WiOpt)*, 2013, pp. 248–255.

[36] J. Lee, A. Conti, A. Rabbachin, and M. Z. Win, "Distributed network secrecy," *IEEE J. Sel. Areas Commun.*, vol. 31, no. 9, pp. 1889–1900, Sep. 2013.

GEORGIOS SMPOKOS (Graduate Student Member, IEEE) received the M.Eng. degree in electrical and computer engineering from the Aristotle University of Thessaloniki, Greece, in 2009, and the M.Sc. degree in wireless communications from the University of Southampton, U.K., in 2011. He is currently pursuing the Ph.D. degree with the Communications and Transportation Systems Group (KTS), Department of Science and Technology, Linköping University, Norrköping, Sweden. From 2011 to 2015, he was with Rohde & Schwarz, U.K. From 2015 to 2016, he was with Dialog Semiconductor, U.K. From 2016 to 2019, he was the Marie Curie Early-Stage Researcher with the Horizon 2020 MSCA ITN-EID Project WiVi-2020, under an industrial placement with Cyta Hellas, Athens, Greece. He is also working as an IP Development and Testing Engineer with the Group Network Engineering and Delivery Department, Vodafone, Athens. His main research interests include communication networks with emphasis on the access and core network optimization, network virtualization, and network slicing.

ZHENG CHEN (Member, IEEE) received the B.S. degree from the Huazhong University of Science and Technology, China, in 2011, and the M.S. and Ph.D. degrees from CentraleSupélec, University of Paris-Saclay, France, in 2013 and 2016, respectively. From June 2015 to November 2015, she was a Visiting Scholar with the Singapore University of Technology and Design, Singapore. Since January 2017, she has been a Postdoctoral Researcher with Linköping University, Sweden. Her research interests include stochastic modeling and optimization, wireless edge caching, machine-type communication, and distributed intelligent systems. She was a recipient of the 2020 IEEE Communications Society Young Author Best Paper Award. She was selected as an Exemplary Reviewer for IEEE COMMUNICATIONS LETTERS in 2016, IEEE TRANSACTIONS ON WIRELESS COMMUNICATIONS in 2017, and IEEE TRANSACTIONS ON COMMUNICATIONS in 2019.

PARTHAJIT MOHAPATRA (Member, IEEE) received the B.E. degree in electronics and communication engineering from the Biju Patnaik University of Technology (BPUT), India, in 2003, the M.Tech. degree in electronic systems and communications from the National Institute of Technology, Rourkela, India, in 2006, and the Ph.D. degree in electrical communication engineering from Indian Institute of Science, Bengaluru, India, in 2015. From March 2015 to July 2016, he was working as a Postdoctoral Research Fellow with the iTrust Center for Research in Cyber Security, Singapore University of Technology and Design (SUTD), Singapore. From August 2016 to July 2018, he was working as an Assistant Professor with the G. S. Sanyal School of Telecommunications, IIT Kharagpur, India. Since July 2018, he has been working as an Assistant Professor with the Department of Electrical Engineering, IIT Tirupati. His primary research interests include wireless communication and signal processing for communication, physical layer secrecy, finite block length information theory and its applications to wireless communication, and union of physical and network layer.

NIKOLAOS PAPPAS (Member, IEEE) received the B.Sc. and M.Sc. degrees in computer science, the second B.Sc. degree in mathematics, and the Ph.D. degree in computer science from the University of Crete, Greece, in 2005, 2007, 2012, and 2012, respectively. From 2005 to 2012, he was a Graduate Research Assistant with the Telecommunications and Networks Laboratory, Institute of Computer Science, Foundation for Research and Technology-Hellas, and a Visiting Scholar with the Institute of Systems Research, University of Maryland at College Park, College Park, MD, USA. From 2012 to 2014, he was a Postdoctoral Researcher with the Department of Telecommunications, Supélec, France. Since 2014, he has been with Linköping University, Norrköping, Sweden, as the Marie Curie Fellow (IAPP), where he is currently an Associate Professor of mobile telecommunications with the Department of Science and Technology. His main research interests include the field of wireless communication networks with emphasis on the stability analysis, energy harvesting networks, network-level cooperation, age-of-information, network coding, and stochastic geometry. From 2013 to 2018, he was an Editor of IEEE COMMUNICATIONS LETTERS. He is currently an Editor of IEEE TRANSACTIONS ON COMMUNICATIONS, IEEE/KICS JOURNAL OF COMMUNICATIONS AND NETWORKS, and IEEE OPEN JOURNAL OF COMMUNICATIONS SOCIETY. He is also a Guest Editor of IEEE INTERNET OF THINGS JOURNAL for the Special Issue "Age of Information and Data Semantics for Sensing, Communication and Control Co-Design in IoT."

• • •