

# Covert, Low-Delay, Coded Message Passing in Mobile (IoT) Networks

Pei Peng<sup>ID</sup> and Emina Soljanin<sup>ID</sup>, *Fellow, IEEE*

**Abstract**—We introduce a gossip-like protocol for covert message passing between Alice and Bob as they move in an area watched over by a warden Willie. The area hosts a multitude of Internet of (Battlefield) Things (Io $\beta$ T) objects. Alice and Bob perform random walks on a random regular graph. The Io $\beta$ T objects reside on the vertices of this graph, and some can serve as relays between Alice and Bob. The protocol starts with Alice splitting her message into small chunks, which she can covertly deposit to the relays she encounters. The protocol ends with Bob collecting the chunks. Alice may encode her data before the dissemination. Willie can either perform random walks as Alice and Bob do or conduct uniform surveillance of the area. In either case, he can only observe one relay at a time. We evaluate the system performance by the covertness probability and the message passing delay. In our protocol, Alice splits her message to increase the covertness probability and adds (coded) redundancy to reduce the transmission delay. The performance metrics depend on the graph, communications delay, and code parameters. We show that, in most scenarios, it is impossible to find the design parameters that simultaneously maximize the covertness probability and minimize the message delay.

**Index Terms**—Covert communications, random walks, Internet of (battlefield) Things, delay reduction.

## I. INTRODUCTION

HINGING various aspects of communications is often essential. In wartime, communication between the suspected parties can alert the adversary even if the message is unknown. In everyday life, revealing the identity of communicating parties affects the increasingly important anonymity and privacy. Several recent papers addressed covert communications at the physical layer. There, two parties, Alice and Bob, communicate while observed by the warden Willie. An information-theoretic approach to achieving covertness, roughly speaking, relies on camouflaging messages as noise (see, e.g., [1]–[5] and references therein). An extension of this model (see, e.g., [6]) involves a jammer that can help

Manuscript received August 7, 2021; revised December 21, 2021; accepted January 12, 2022. Date of publication February 4, 2022; date of current version February 14, 2022. This work was supported by the National Science Foundation under Grant SaTC-1816404. An earlier version of this paper was presented in part at the 2019 57th Annual Allerton Conference on Communication, Control, and Computing [32] [DOI: 10.1109/ALLERTON.2019.8919706]. The associate editor coordinating the review of this manuscript and approving it for publication was Dr. Alexey Vinel. (Corresponding author: Pei Peng.)

Pei Peng is with the Department of Electrical and Computer Engineering, Rutgers University, Piscataway, NJ 08854 USA, and also with the School of Telecommunication and Information Engineering, Nanjing University of Posts and Telecommunications, Nanjing 210003, China (e-mail: pei.peng@rutgers.edu).

Emina Soljanin is with the Department of Electrical and Computer Engineering, Rutgers University, Piscataway, NJ 08854 USA (e-mail: emina.soljanin@rutgers.edu).

Digital Object Identifier 10.1109/TIFS.2022.3147642

Alice to transmit covertly [7]. Another extension involves a third participant Carol which acts as a cover for Bob, [8], [9] studies the covert communication in one-way relay networks. We here propose a covert message-passing protocol for IoT environments. This protocol is complementary and can be used in conjunction with the previously proposed methods for covert transmission.

The last decade has seen a wide variety of novel communications systems. Future 5G systems are supposed to host a hundred times more devices than current 4G environments, and one can potentially harness the resources expected to be brought in by smart (battlefield) devices in the future Internet of (Battlefield) Things (Io $\beta$ T) environments. By exploiting these devices' storage and communication ability, Alice can covertly pass messages to Bob in a gossip-like manner as outlined below.

Alice and Bob communicate over a wide geographic area (e.g., battlefield or an occupied city) patrolled by a warden Willie. The area hosts a multitude of IoT objects capable of storing, sending, and receiving data. Alice splits her message into small chunks, which she can inconspicuously pass, one at a time, to IoT objects that appear in her proximity as she randomly moves through the area. We refer to such IoT objects as relays or helpers. Bob, who also randomly moves through the area, can then retrieve the stored data chunks. Because the IoT objects are distributed over a wide area, Willie can only periodically check if any of these objects is transmitting or receiving data.

The movements of Alice and Bob are modeled as random walks on graphs. Information gathering and dissemination on graphs is an interesting problem that naturally arises and is actively researched in many different contexts. Examples include: border control using unmanned aerial vehicles (UAVs) [10], measuring traffic, reporting road conditions and helping with emergency response using UAVs [11], monitoring the ocean [12], measuring air pollution [13], multi-agent systems [14], and more recently for timely exchange of information updates [15]. We consider two detection patterns of Willie: random patrolling and uniform surveillance. In the patrolling model, Willie performs random walks on graphs and can detect the communication when he happens to be on at the same relay as Alice or Bob. In the surveillance model, Willie scans graph nodes (e.g., while sitting in a control room). He can detect communication when he observes Alice or Bob.

Having to distribute and collect many chunks, as well as the unpredictability of mobility and availability of helpers can cause significant delays in our mobile information transfer. To increase the persistence of information in the unstable

environment, the agent may make the data chunks redundant by erasure correcting codes, requiring more data chunks to be distributed and collected. One would expect that would further increase the delay. However, that is not necessarily the case, and we will see that coding and some other forms of redundancy can, in fact, be used to reduce the delay, as previously shown to be the case in data download and straggler mitigation (see, e.g., [16]–[21] and references therein).

It may be helpful to the reader to consider our work in the context of the literature on throughput and delay in mobile ad-hoc communications. For example, [22] proposed a model for studying the throughput scaling of static ad hoc networks; [23] introduced the mobility into the communications model and assumed the source only transfers the data packet to a nearby relay which subsequently passes the packet to the destination. [24] introduced a random walk model for node mobility and analyzed the throughput and delay tradeoff, and [25] showed how coding improves this tradeoff, and [26] studied the throughput scaling of covert communication in ad-hoc networks introduced in [23]. Our work is different in multiple ways. It considers an IoT specific system model, which includes different communication protocols and performance metrics. However, our approach to deriving communications delay and the techniques for reducing it should be of interest to the mobile ad-hoc communications community.

The contributions of this paper are summarized as follows:

- 1) We propose a gossip-like protocol for a covert dissemination/collection of message chunks in IoT environments. We analyze the dissemination and collection delay for two transmission time models. For both models, our theoretical and numerical results show that introducing redundant IoT relays can reduce the dissemination time, and erasure coding of messages can reduce the collection time.
- 2) We extend our analysis to covert communication scenarios. We introduce two warden models and derive/analyze the covertness probability for each. The theoretical and numerical results show that coding always reduces the covertness probability while splitting the message into smaller chunks may increase the covertness probability.
- 3) We numerically analyze delay vs. the covertness probability tradeoff. We conclude that the tradeoff is very different for different system parameters. In some scenarios, there exists an optimal code rate that maximizes the covertness probability and minimizes delay. However, in most scenarios, simultaneously maximizing the covertness probability and minimizing the communications delay is impossible.

The paper is organized as follows: In Sec. II, we present the message passing model and two delay models. In Sec. III, we derive and analyze the expressions of dissemination/collection time for different delay models and numerically analyze the message passing delays. In Sec. IV, we propose two warden detection models and point out the tradeoff that exists between the message passing delay and covertness probability. In Sec. V, we derive and analyze the expressions of covertness probability for different detection models.

In Sec. VI, we present some numerical results for the delay vs. covertness probability tradeoff. Conclusions are given in Sec. VII-5.

## II. SYSTEM MODEL AND PROBLEM FORMULATION

We first describe the communication participants, their mobility, and message passing protocols and delay models. We then define two message passing performance metrics we will be studying. This section focuses the message passing delay. The message passing covertness and the delay/covertness tradeoff are discussed in Sec. IV.

### A. Communication Participants and Their Mobility

There are three types of participants present within a geographical area (e.g., a city or a battlefield): a mobile source Alice,  $r$  static relays, a mobile receiver Bob. Alice's and Bob's mobility is modeled as a simple random walk over a *mobility graph*. (Sec. IV extends this model to include a warden Willie.)

We model the mobility graph as a random  $d$ -regular graph, that is, a graph selected uniformly at random from the set of all  $d$ -regular graphs on  $n$  vertices. The related literature (see, e.g., [24], [25]) uses rectangular grid graphs over a torus to study communications in (mobile) ad-hoc networks. We adopt the random regular graph model for the following reasons. 1) Since these graphs are locally trees with high probability [27], the random walks on such graphs are reasonable mobility models. 2) Although random regular graphs are more realistic than the grid on torus graph for mobility modeling, random walks on these graphs are fairly well understood. Thus, the existing mathematical results on this topic can be used and easily extended to our scenarios, allowing us to concentrate on the communications and covertness problems we aim to solve. Other mobility graphs will be used in future research, especially irregular graphs with a few high degree hubs.

The  $r$  relays are IoT objects residing on  $r$  vertices of the mobility graph. Alice has a message to pass to Bob. Instead of communicating with Bob directly, Alice uses the IoT objects (relays) to store pieces of her message, which Bob can subsequently retrieve. Because IoT objects have limited power, Alice and Bob can only communicate with a relay when they reach the vertex where the relay resides.

### B. Communications Protocol & the Mobility Model

Alice's *message* has length  $m$  bits (symbols, packets). Because of the relay's storage constraints and to avoid long transmission time at a relay, she partitions the message into  $k$  *data chunks* each of length  $\ell = m/k$ . By using an MDS code, she encodes the data chunks into  $n(\leq r)$  *coded chunks* (also of length  $\ell$ ). Thus, Bob needs to collect any  $k$  out of  $n$  coded chunks to recover the message. An example of message partitioning and encoding, together with the mobility graph, is shown in Fig. 1. Here, Alice splits the message into  $k = 2$  data chunks and encodes them into  $n = 3$  coded chunks  $\{M_1, M_2, M_3\}$ .

Message passing from Alice to Bob has two phases: the dissemination phase and the collection phase. In the dissemination phase, Alice transfers  $n$  coded chunks to the first

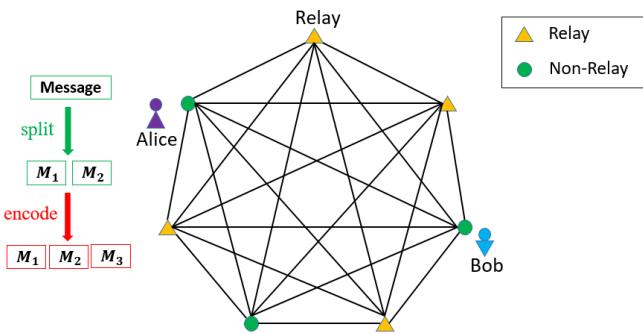


Fig. 1. An example of message passing over a complete graph with  $v = 7$  vertices. Each vertex contains an IoT object. Four among them (yellow triangles) are relays and three (green disks) are not. Alice and Bob can communicate with any object but only disseminate/collect message chunks to/from the relays.

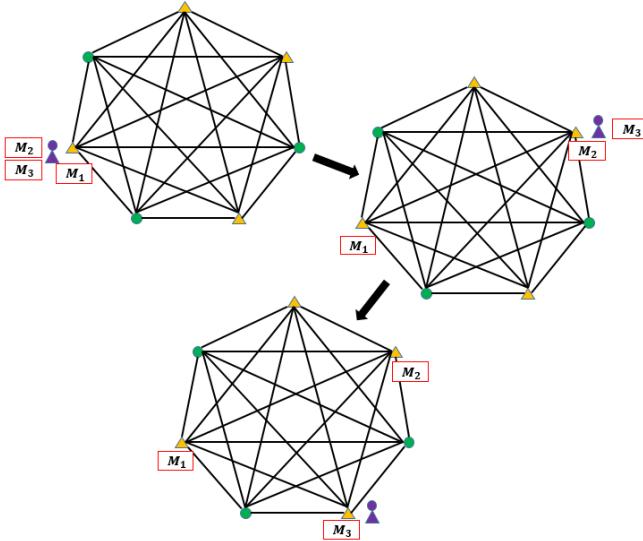


Fig. 2. Alice needs to disseminate three coded chunks to any three out of four relays. Every time when she meets a relay, she stores one coded chunk, until all three coded chunks are disseminated. Each relay can only store one coded chunk.

$n$  relays she encounters as she randomly walks through the graph. This model is similar to a classic predator-prey model, e.g., as analyzed in [27]. Fig. 2 shows an example. When she meets the first relay, she deposits the coded chunk  $M_1$  to the relay. She then deposits  $M_2$  to the second relay, and so on. The dissemination is complete when Alice deposits the last chunk. Fig. 2 shows the snapshots of Alice at the three relays. She may need to walk through several non-relay vertices to reach all the relays. Similarly, Bob recovers the message in the collection phase by collecting  $k$  coded chunks from the  $n$  relays as he randomly walks through the graph.

### C. Message Chunk Passing Time

The time to pass a chunk between Alice (or Bob) and a relay has two components. The first is the time Alice (or Bob) needs to spend randomly walking to meet a relay. The second is the transmission time of a data chunk between Alice/Bob and a relay. We refer to Alice's move from a vertex to one of its neighbors as *one step*. The time to take a step is, in general, a random variable  $\eta$ . We consider the following two chunk passing time models:

*Model 1 (Constant Transmission Time):* In this classic model for random walks (see, e.g., [27], [28] and references therein), the chunk passing time is measured in the number of steps Alice/Bob needs to make to encounter a relay. It is applicable here under the following circumstances: 1) the time of a step  $\eta$  is a constant, say 1 and 2) the chunks are sufficiently small so that their transmission time can be neglected.

*Model 2 (Random Transmission Time):* In this general model, the step times are independent and identically distributed (i.i.d.) random variables. The transmission time of a data chunk between Alice/Bob and a relay follows a shifted exponential distribution whose shift is proportional to the length of a chunk  $\ell$  and whose tail accounts for various disturbances (noise) in the system.

### D. Covert Message Passing

The covert message passing is defined as Alice successfully passing a message to Bob through some relays without Willie's detection. We consider two detection patterns of Willie: random patrolling and uniform surveillance. In the patrolling model, Willie performs random walks on graphs and can detect the communication when he happens to be on at the same relay as Alice or Bob. In the surveillance model, Willie scans graph nodes (e.g., while sitting in a control room). He can detect communication when he observes Alice or Bob. The details of the detection models are given in Sec. IV.

### E. Performance Metrics and Problem Formulation

Alice's goal is to covertly and quickly pass the message to Bob. Thus, the performance metrics of interest are *covertness probability* ( $P_c$ ), the *expected dissemination time* ( $\mathbb{E}[T_A]$ ) of  $n$  coded chunks by Alice, the *expected collection time* ( $\mathbb{E}[T_B]$ ) of  $k$  coded chunks by Bob, and the *expected message passing delay* ( $\mathbb{E}[T_{A+B}]$ ) (dissemination plus collection). We evaluate the above metrics in terms of the design parameters  $k$  and  $n$ , where  $k/n$  is the code rate. Our goal is to find the optimal  $k$  and  $n$  that maximize  $P_c$ , and minimize  $\mathbb{E}[T_A]$ ,  $\mathbb{E}[T_B]$ , and  $\mathbb{E}[T_{A+B}]$ . In the following section, we will derive the formulas for each metric and analyze the optimal  $k$  and  $n$ . The covertness probability will be addressed in Sec. IV.

Recall that a *message* of length  $m$  data bits, and is split into  $k$  *data chunks*. Applying an  $[n, k]$  MDS code,  $k$  data chunks are encoded into  $n$  *coded chunks*. The length of each data/coded chunk is  $\ell = m/k$  bits. Therefore, Alice needs to meet  $n$  relays to deposit her coded chunks. We denote by  $S_A$  the random number of steps Alice needs to make to meet  $n$  out of  $r$  relays. Bob needs to meet  $k$  relays with coded chunks. We denote by  $S_B$  the random number of steps Bob needs to make to meet  $k$  out of  $n$  relays storing coded chunks.

### F. Parameters and Notation

- $v$  - number vertices in the mobility graph
- $d$  - mobility graph vertex degree
- $r$  - number of relays in the system
- $m$  - message (data) length in bits
- $k$  - number of message (data) chunks
- $n$  - number of coded chunks

$k/n$ -	code rate
$\ell$ -	length of a data (or coded) chunk
$\eta$ -	random walking time between two vertices
$T_A$ -	random dissemination time of $n$ coded chunks
$T_B$ -	random collection time of $k$ coded chunks
$T_{A+B}$ -	random message passing time
$S_A$ -	random # of Alice's steps to meet $n$ relays
$S_B$ -	random # of Bob's steps to meet $k$ relays

Observe that  $v$ ,  $d$ ,  $r$ ,  $m$  and  $\mathbb{E}[\eta]$  are given system parameters.  $\mathbb{E}[T_A]$ ,  $\mathbb{E}[T_B]$ ,  $\mathbb{E}[T_{A+B}]$ ,  $\mathbb{E}[S_A]$  and  $\mathbb{E}[S_B]$  are the performance metrics.  $n$ ,  $k$  and  $\ell$  are the design parameters which we can select to optimize the performances metrics.

### III. MESSAGE PASSING DELAYS

#### A. Preliminary Reasoning

Message passing delay includes dissemination and the collection time. It is the time during which Alice and Bob communications with relays could be discovered by Willie. Observe that the message passing delay is not the time that the message spends in the system from the beginning of Alice's dissemination to the end of Bob's collection. That time depends on the starting times of the dissemination and collection, and is beyond the scope of this paper.

The message passing delay depends on how many steps Alice (Bob) needs to make to disseminate (collect) the coded chunks. To understand the message passing delay dependencies on the system and design parameters, we first consider a special case when the mobility graph is a complete graph. Here, Alice needs to disseminate  $n$  coded chunks to  $r$  relays residing on a  $v$ -vertex complete graph. Her first chunk can be deposited on any of the  $r$  relays. In a single step, Alice arrives at a relay with the probability  $r/v$ , and thus she needs to make on average  $v/r$  steps to find a relay to deposit her first chunk. For depositing her second chunk, Alice needs to arrive to one of the remaining  $r-1$  relays, which happens in a single step with probability  $(r-1)/v$ . Thus, she has to make another  $v/(r-1)$  steps on average to deposit her second chunk. Therefore, to deposit her first two chunks, Alice will make  $\frac{v}{r} + \frac{v}{r-1}$  steps on average. Following this reasoning, we see that for  $n$  chunks, Alice needs to make on average  $v(H_r - H_{r-n})$  steps, where  $H_r = \sum_{i=1}^r 1/i$  is the  $r$ -th harmonic number. We often use the approximation  $H_n = \log n + \gamma + \mathcal{O}(n^{-1})$ , where  $\gamma = 0.577$  is Euler's constant. Bob needs to collect  $k$  chunks from  $n$  relays storing coded chunks. Therefore, Bob needs to make on average  $v(H_n - H_{n-k})$  steps.

The above reasoning does not extend to general graphs. However, reference [27] provides useful results for large, random regular graphs (which are good mobility graphs models). We adapt the findings of [27] to our setting, and get the following results. Let  $S$  be the number of steps that the source (or collector) needs to make to meet a relay. When we randomly choose a graph from the set of all  $d$ -regular graphs with  $v$  vertices, with high probability, 1) the expected number of steps is

$$\mathbb{E}[S] \sim_v \frac{\theta_d v}{r} \quad (1)$$

and 2) the probability that the source meets a relay  $P_{meet}$  is

$$P_{meet} \sim_v \frac{r}{\theta_d v}. \quad (2)$$

where  $X \sim_Z Y$  means  $\lim_{Z \rightarrow \infty} X/Y = 1$ , and  $\theta_d = \frac{d-1}{d-2}$ .

We can now find the expected message passing delay. According to the definition of  $T_{A+B}$ , we know that its expectation is  $\mathbb{E}[T_{A+B}] = \mathbb{E}[T_A] + \mathbb{E}[T_B]$ . The dissemination time  $T_A$  is the sum of  $n$  message chunk passing times (discussed in Sec. II-C). For the constant transmission time model with the step time  $\eta = 1$ ,  $T_A$  is equal to the random number of steps  $S_A$  that Alice needs to make to meet  $n$  out of  $r$  relays to deposit her  $n$  coded chunks. Therefore, the expected dissemination time is  $\mathbb{E}[T_A] = \mathbb{E}[S_A]$ . Similarly, the collection time  $T_B$  is equivalent to the total number of steps  $S_B$  that Bob needs to make to meet  $k$  out of  $n$  relays to collect  $k$  coded chunks. Therefore, the expected dissemination time is  $\mathbb{E}[T_B] = \mathbb{E}[S_B]$ .

For the random transmission time model, the dissemination time  $T_A$  (collection time  $T_B$ ) is again equal to the sum of  $n$  ( $k$ ) chunk passing times. Recall that the chunk transmission time can not be neglected in this model, and it follows a shifted exponential distribution. Thus the calculations of  $\mathbb{E}[T_A]$  and  $\mathbb{E}[T_B]$  are more complicated and will be discussed later.

In the following, we focus on the delay analysis of a random regular graph. We will find the expression of message passing delay by deriving the dissemination time and collection time separately under two different delay models.

#### B. Constant Transmission Time

Under the constant transmission time model, the message passing delay  $T_{A+B} = S_A + S_B$ , where  $S_A$  and  $S_B$  are the number of steps that the source and collector need to spend to meet enough relays.

1) *Dissemination Time*: From the above discussion, it follows that in a  $d$ -regular graph with  $v$  vertices, the source needs to make on average  $\frac{\theta_d v}{r}$  steps to disseminate the first coded chunk. After the first dissemination, the number of available relays reduces to  $r-1$ , and then the source needs to spend on average  $\frac{\theta_d v}{r-1}$  steps to disseminate the second chunk. Finally, we can get the expected number of steps to disseminate  $n$  coded chunks

$$\mathbb{E}[S_A] \sim_v \theta_d v (H_r - H_{r-n}).$$

It is obvious that  $\mathbb{E}[S_A]$  decreases with increasing  $r$  and increases with  $n$ . Since  $n = \varphi r$  for some  $\varphi \in (0, 1)$ , it is not hard to see  $\mathcal{O}(\mathbb{E}[S_A]) = \mathcal{O}(v)$ .

For a scenario without redundant relays, the number of relays  $r$  is equal to the number of coded chunks  $n$ . Thus, the expected number of steps to disseminate  $n$  coded chunks is  $\mathbb{E}[S_A] \sim_v \theta_d v H_n$ . Namely, we have  $\mathcal{O}(\mathbb{E}[S_A]) = \mathcal{O}(v \log n)$ .

2) *Collection Time*: The collector needs to collect any  $k$  coded chunks from  $n$  relays that store the chunks. Similarly, we can get the expected number of steps to collect any  $k$  chunks

$$\mathbb{E}[S_B] \sim_v \theta_d v (H_n - H_{n-k}).$$

It is obvious that  $\mathbb{E}[S_B]$  decreases with increasing  $n$  and increases with  $k$ . Since  $k = \xi n$  for some  $\xi \in (0, 1)$ , it is not hard to see  $\mathcal{O}(\mathbb{E}[S_B]) = \mathcal{O}(v)$ .

For a scenario without coding, the code rate  $k/n = 1$ , namely,  $k = n$ . Thus, the expected number of steps to collect  $n$  data chunks is  $\mathbb{E}[S_B] \sim_v \theta_d v H_n$ . Namely, we have  $\mathcal{O}(\mathbb{E}[S_B]) = \mathcal{O}(v \log n)$ .

3) *Message Passing Delay*: The message passing delay is the sum of dissemination time and collection time. Since both  $\mathbb{E}[S_A]$  and  $\mathbb{E}[S_B]$  are positive, the following asymptotic equivalence holds:

$$\mathbb{E}[T_{A+B}] \sim_v \theta_d v (H_r + H_n - H_{r-n} - H_{n-k}). \quad (3)$$

We know that  $\mathbb{E}[T_{A+B}]$  increases with  $k$ . Namely, when the number of coded chunks  $n$  is given,  $\mathbb{E}[T_{A+B}]$  reaches minimum at  $k = 1$ . However, since the IoT devices have limited storage, we sometimes have to split the original message into smaller chunks to store in the relays. Therefore, it is important to study when  $k$  is given, what is the value of  $n$  can minimize  $\mathbb{E}[T_{A+B}]$ . We find some results in Theorem 1.

**Theorem 1:** *By using the asymptotic equivalence (3), the expected message passing delay  $\mathbb{E}[T_{A+B}]$  reaches the minimum at  $n = \sqrt{rk + k} - 1$  given  $r$  and  $k$ . Notice that if the above value is not an integer, the optimal  $n$  is  $\lceil \sqrt{rk + k} - 1 \rceil$  or  $\lfloor \sqrt{rk + k} - 1 \rfloor$ .*

*Proof:* When  $k$  and  $r$  are given,  $\mathbb{E}[T_{A+B}]$  is a function of  $n$ . Since  $\mathbb{E}[T_{A+B}]$  is discrete, we can get its minimum by find an  $n^*$  where  $\mathbb{E}[T_{A+B}](n = n^*) \leq \mathbb{E}[T_{A+B}](n = n^* + 1)$  and  $\mathbb{E}[T_{A+B}](n = n^*) \leq \mathbb{E}[T_{A+B}](n = n^* - 1)$ . From (3), we have  $\mathbb{E}[T_{A+B}(n = i)] \sim_v \theta_d v (H_r + H_i - H_{r-i} - H_{i-k})$  and  $\mathbb{E}[T_{A+B}(n = i + 1)] \sim_v \theta_d v (H_r + H_{i+1} - H_{r-i-1} - H_{i+1-k})$ .

Define  $A_{n=i} = \mathbb{E}[T_{A+B}(n = i + 1)] - \mathbb{E}[T_{A+B}(n = i)]$ , then

$$\begin{aligned} A_{n=i} &\sim_v \theta_d v \left( \frac{1}{i+1} + \frac{1}{r-n} - \frac{1}{n+1-k} \right) \\ &= \frac{\theta_d v}{(i+1)(r-n)(n+1-k)} (n^2 + 2n + 1 - rk - k) \\ &= \frac{\theta_d v}{(i+1)(r-n)(n+1-k)} [(n+1)^2 - (rk+k)]. \end{aligned}$$

Since  $\frac{\theta_d v}{(i+1)(r-n)(n+1-k)} > 0$ ,  $A_{n=i} \geq 0$  for  $n \geq \sqrt{rk + k} - 1$  and  $A_{n=i} < 0$  for  $n < \sqrt{rk + k} - 1$ . Therefore,  $n = \sqrt{rk + k} - 1$  minimizes  $\mathbb{E}[T_{A+B}]$ .  $\square$

Observe that for a complete graph with  $v$  vertices, the expression for message passing delay becomes  $T_{A+B} = v(H_r + H_n - H_{r-n} - H_{n-k})$ , as previously derived.

**Numerical Analysis:** In Fig. 3, we evaluate the expected message passing delay  $\mathbb{E}[T_{A+B}]$  vs. the number of coded chunks  $n$  for both approximation and exact result. We consider a regular graph with 100 vertices and the degree is 5. The approximation is calculated from the expression of  $\mathbb{E}[T_{A+B}]$  given in (3), and the exact result is an average of 1000 sampled delay values. Some observations are made from the figure: when  $n$  is given,  $\mathbb{E}[T_{A+B}]$  increases with  $k$ . For example, when  $n = 7$ , the case “ $k = 2$ ” gives the minimum  $\mathbb{E}[T_{A+B}]$ . When  $n$  is large (e.g.,  $n = 10$ ), the gaps between these three cases are small. These observations are consistent with the theoretical analysis for  $k$ . We also observe that when  $k$

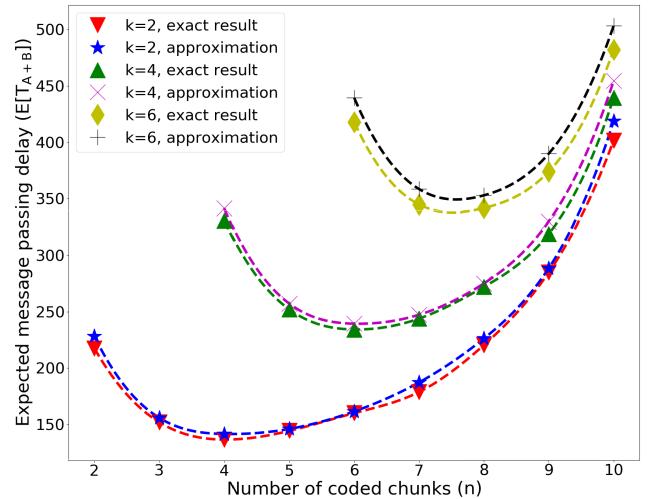


Fig. 3. Expected message passing delay  $\mathbb{E}[T_{A+B}]$  vs. the number of coded chunks  $n$  (cf. (3)). This is a regular graph with 100 vertices and the degree is 5. The number of relays is  $r = 10$ . Introducing proper data redundancy can reduce the message passing delay and the approximation is close to the exact result.

is given,  $\mathbb{E}[T_{A+B}]$  always reaches the minimum at  $n > k$ . For example, when  $k = 4$ , the optimal  $n$  is 6 and the code rate  $k/n = 2/3$ . We conclude that introducing proper data redundancy can reduce the message passing delay. This observation is consistent with the result in Theorem 1. Besides, we observe that for each  $k$ , the approximation is very close to the exact result. It confirms that (3) is good enough to approximate the expected message passing delay.

### C. Random Transmission Time

Under the random transmission time model, both the source and collector need to spend some walking time (totally  $\eta_1 + \dots + \eta_S$ ) to meet a relay, and then they need to spend the message chunk passing time (defined as  $t_t$ ) to delivery and collect the chunk. According to Subsection II-C,  $t_t$  follows a shifted exponential distribution  $S\text{-Exp}(\Delta, \lambda)$ , where  $\Delta = \ell$ .

1) *Dissemination Time*: The source needs to disseminate  $n$  coded chunks to  $r(\geq n)$  relays. When the source randomly walks on a  $d$ -regular graph with  $v$  vertices, the probability that it meets a relay is  $\frac{r}{\theta_d v}$  according to (2). After the source deposits the first chunk in one of the  $r$  relays, the second chunk can only be stored in one of the remaining  $r - 1$  relays. The probability of meeting an unoccupied relay decreases as the number of occupied relays grows. Therefore, in order to get the dissemination time  $T_A$ , we need to find the time  $T_i$  for the source to disseminate the  $i^{\text{th}}$  ( $i = \{1, 2, \dots, n\}$ ) coded chunk.

**Lemma 1:** *For the random transmission time model, the time  $T_i$  for the source to disseminate the  $i^{\text{th}}$  coded chunk to any one of  $r - i + 1$  relays is*

$$T_i = t_t + \sum_{j=1}^S \eta_j \quad \text{with probability } (1 - p_{r-i+1})^{S-1} p_{r-i+1}. \quad (4)$$

where  $S$  is the number of steps the source spends to meet a relay, and  $p_{r-i+1} \sim_v \frac{r-i+1}{\theta_d v}$  is the probability that the source

meets any one of  $r - i + 1$  relays. Then we have

$$\mathbb{E}[T_i] = \frac{1}{\lambda} + \frac{m}{k} + \frac{\mathbb{E}[\eta]}{p_{r-i+1}}. \quad (5)$$

*Proof:* Let  $p = p_{r-i+1}$ , we get the expectation of  $T_i$  as follows:

$$\begin{aligned} \mathbb{E}[T_i] &= \sum_{s=1}^{\infty} \mathbb{E} \left[ t_i + \sum_{j=1}^s \eta_j \right] (1-p)^{s-1} p \\ &= p \mathbb{E}[t_i] \sum_{s=1}^{\infty} (1-p)^{s-1} \\ &\quad + p \sum_{s=1}^{\infty} \mathbb{E} \left[ \sum_{j=1}^s \eta_j \right] (1-p)^{s-1}. \end{aligned}$$

Since  $\eta_1, \dots, \eta_s$  are i.i.d.,  $\mathbb{E} \left[ \sum_{j=1}^s \eta_j \right] = s \mathbb{E}[\eta]$ . Therefore,  $\mathbb{E}[T_i] = \mathbb{E}[t_i] + \frac{\mathbb{E}[\eta]}{p} = \frac{1}{\lambda} + \frac{m}{k} + \frac{\mathbb{E}[\eta]}{p_{r-i+1}}$ .  $\square$

Using Lemma 1, we can get the expected total dissemination time  $\mathbb{E}[T_A]$  in the following theorem.

**Theorem 2:** The expected dissemination time for the source to transmit all  $n$  coded chunks to any  $n$  out of  $r$  relays is

$$\mathbb{E}[T_A] \sim_v \frac{n}{\lambda} + \frac{nm}{k} + \theta_{dv} \mathbb{E}[\eta] (H_r - H_{r-n}). \quad (6)$$

*Proof:* Since  $\mathbb{E}[T_i] = \frac{1}{\lambda} + \frac{m}{k} + \frac{\mathbb{E}[\eta]}{p_{r-i+1}}$  for  $i \in \{1, 2, \dots, n\}$ , where  $p_{r-i+1} \sim_v \frac{r-i+1}{\theta_{dv}}$ , we have

$$\begin{aligned} \mathbb{E}[T_A] &= \mathbb{E} \left[ \sum_{i=1}^n T_i \right] = n \mathbb{E}[t_i] + \sum_{i=1}^n \frac{\mathbb{E}[\eta]}{p_{r-i+1}} \\ &\sim_v n \mathbb{E}[t_i] + \sum_{i=1}^n \frac{\theta_{dv} \mathbb{E}[\eta]}{r - i + 1} \\ &= \frac{n}{\lambda} + n\ell + \theta_{dv} \mathbb{E}[\eta] (H_r - H_{r-n}). \end{aligned}$$

Since  $\ell = \frac{m}{k}$ ,  $\mathbb{E}[T_A] \sim_v \frac{n}{\lambda} + \frac{nm}{k} + \theta_{dv} \mathbb{E}[\eta] (H_r - H_{r-n})$ .  $\square$

Notice that  $\eta$  is determined by the distance between two vertices and the speed of the source/collector. To study these two parameters is not the purpose of this paper. Therefore, we will directly assign a value to  $\mathbb{E}[\eta]$ . Therefore, from Theorem 2, we see that  $\mathbb{E}[T_A]$  increases with  $n$  and decreases with increasing  $k$ . Since  $r = \varphi n$  and  $k = \xi n$  for some  $\varphi, \xi \in (0, 1)$ , it is not hard to see  $\mathcal{O}(\mathbb{E}[T_A]) = \mathcal{O}(v + n)$ .

For a scenario without redundant relays, the number of relays  $r$  is equal to the number of coded chunks  $n$ . Thus, the expected dissemination time is  $\mathbb{E}[T_A] \sim_v \frac{n}{\lambda} + \frac{nm}{k} + \theta_{dv} H_n$ . Namely, we have  $\mathcal{O}(\mathbb{E}[T_A]) = \mathcal{O}(v \log n + n)$ .

**2) Collection Time:** The collector needs to collect at least  $k$  coded chunks to recover the message. Since there are only  $n$  ( $\leq r$ ) relays on the graph store the coded chunks, the probability that the collector meets a relay with a coded chunk is  $\frac{n}{\theta_{dv}}$ . After the collector retrieves the first chunk in one of the  $n$  relays, the second chunk can only be retrieved in one of the remaining  $n - 1$  relays. Therefore, similar to the dissemination time, we find the expected collection time  $\mathbb{E}[T_B]$  in the following corollary.

**Corollary 1:** The expected collection time for the collector to retrieve any  $k$  chunks from  $n$  relays is

$$\mathbb{E}[T_B] \sim_v \frac{k}{\lambda} + m + \theta_{dv} \mathbb{E}[\eta] (H_n - H_{n-k}).$$

The above result holds by applying the similar proofs of Lemma 1 and Theorem 2.

From Corollary 1, we see that  $\mathbb{E}[T_B]$  decreases with increasing  $n$  and increases with  $k$ . Since  $k = \xi n$  for some  $\varphi, \xi \in (0, 1)$ , it is not hard to see  $\mathcal{O}(\mathbb{E}[T_B]) = \mathcal{O}(v + k)$ .

For a scenario without coding, the code rate  $k/n = 1$ , namely,  $k = n$ . Thus, the expected collection time is  $\mathbb{E}[T_B] \sim_v \frac{k}{\lambda} + m + \theta_{dv} H_n$ . Namely, we have  $\mathcal{O}(\mathbb{E}[T_B]) = \mathcal{O}(v \log n + n)$ .

**3) Message Passing Delay:** The message passing delay is the sum of the dissemination time and collection time, and its expectation is in Corollary 2.

**Corollary 2:** For the random transmission time model, the expected message passing delay is:

$$\begin{aligned} \mathbb{E}[T_{A+B}] &\sim_v \frac{n+k}{\lambda} + \left( \frac{n}{k} + 1 \right) m \\ &\quad + \theta_{dv} \mathbb{E}[\eta] (H_r + H_n - H_{r-n} - H_{n-k}). \end{aligned} \quad (7)$$

From the conclusions of the dissemination time and the collection time, we see that there should be an optimal  $n$  and an optimal  $k$  which minimize the  $\mathbb{E}[T_{A+B}]$ . Assume  $k$  is given, we find the optimal  $n$  in the following theorem.

**Theorem 3:** For the random transmission time model, when  $k$  is given, the optimal  $n$  appear in the range  $[k, \sqrt{rk + k} - 1]$ .

*Proof:* Define  $A_{n=i} = \mathbb{E}[T_{A+B}(n = i+1)] - \mathbb{E}[T_{A+B}(n = i)]$ , we have  $A_{n=i} \sim_v \frac{1}{\lambda} + \frac{m}{k} + \theta_{dv} \mathbb{E}[\eta] \left( \frac{1}{i+1} + \frac{1}{r-n} - \frac{1}{n+1-k} \right)$ . Since  $\lambda, m, k > 0$ , the former terms  $\frac{1}{\lambda} + \frac{m}{k} > 0$ . According to the proof of Theorem 1, we know that the latter term is larger than 0 when  $n > \sqrt{rk + k} - 1$ , and smaller than 0 when  $n < \sqrt{rk + k} - 1$ . Therefore, the optimal  $n$ , as an integer, must appear in the range  $[k, \sqrt{rk + k} - 1]$ .  $\square$

From Theorem 3, we know that when  $\frac{1}{\lambda} + m \gg \theta_{dv} \mathbb{E}[\eta]$ , it is close to  $k$ . When  $\theta_{dv} \mathbb{E}[\eta] \gg \frac{1}{\lambda} + \frac{m}{k}$ , it is close to  $\sqrt{rk + k} - 1$ .

To analyze the optimal  $k$ , we assume  $n$  is given. From (7), we see that  $k$  appears in the terms  $\frac{mn}{k} + \frac{k}{\lambda} - \theta_{dv} \mathbb{E}[\eta] H_{n-k}$ , where  $\frac{mn}{k}$  decreases with increasing  $k$ , and  $\frac{k}{\lambda} - \theta_{dv} \mathbb{E}[\eta] H_{n-k}$  increases with  $k$ . Therefore, the optimal  $k$  changes with the values of system parameters, e.g.,  $\lambda, d, v, \eta$  and  $m$ .

Notice that for a complete graph with  $v$  vertices, we find the expression of message passing delay is  $T_{A+B} = \frac{n+k}{\lambda} + \left( \frac{n}{k} + 1 \right) m + v \mathbb{E}[\eta] (H_r + H_n - H_{r-n} - H_{n-k})$ . The results in Theorem 3 still hold.

**Numerical Analysis:** In Fig. 4, we evaluate the expected message passing delay  $\mathbb{E}[T_{A+B}]$  vs. the number of coded chunks  $n$  for both approximation and exact result. We consider a regular graph with 100 vertices and the degree is 5. The message chunk passing time  $t_i$  follows S-Exp( $\ell, \lambda$ ), where  $\ell = 100/k$  is the length of a chunk,  $\lambda = 1$  is the rate parameter. By sampling the shifted exponential distribution, we can get a message passing delay value. The exact result is an average of 1000 sampled delay values. The approximation is calculated from the expression of  $\mathbb{E}[T_{A+B}]$  given in (7). Some observations are made from the figure: when  $k$  is

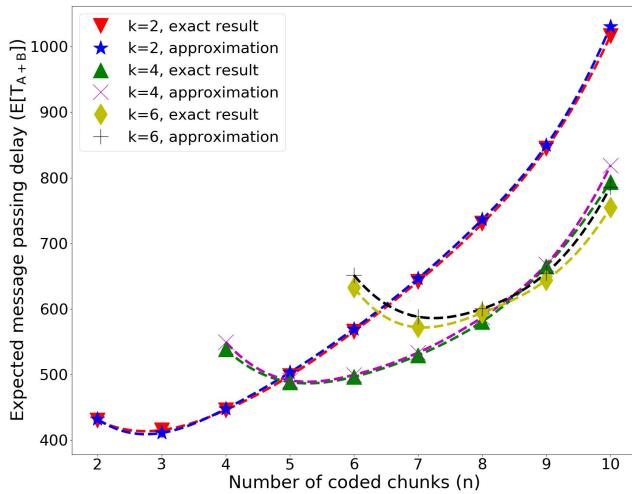


Fig. 4. Expected message passing delay  $\mathbb{E}[T_{A+B}]$  vs. the number of coded chunks  $n$  (cf. (7)). This is a regular graph with 100 vertices and the degree is 5. The number of relays is  $r = 10$  and the message chunk passing time follows  $\text{S-Exp}(100/k, 1)$ . Introducing proper data redundancy can reduce the message passing delay and the approximation is close to the exact result.

given,  $\mathbb{E}[T_{A+B}]$  still reaches the minimum at  $n > k$ . Which indicates that introducing proper data redundancy can reduce the message passing delay. However, comparing to Fig. 3, the optimal code rate  $k/n$  in Fig. 4 is more close to 1. For example, considering the case “ $k = 4$ ”, the optimal code rates in Fig. 3 and Fig. 4 are respectively  $2/3$  and  $4/5$ . This observation validates the result in Theorem 3 that when  $\frac{1}{\lambda} + \frac{m}{k} = 1 + 100/k$  is close to  $\theta_d v = 400/3$ , the optimal  $n$  appears in the range  $[1, \sqrt{11k} - 1]$ .

We also observe that when  $n$  is given,  $\mathbb{E}[T_{A+B}]$  no longer increases with  $k$ . For example, when  $n = 6$ , the case “ $k = 4$ ” has a smaller delay than the case “ $k = 2$ ”. Although our theoretical analysis also shows this result, we can not describe how the optimal  $k$  changes due to there are too many parameters can affect the result. Besides, we observe that the approximation is very close to the exact result. It confirms that (7) is good enough to approximate the expected message passing delay.

#### D. Comparisons and Conclusions

In Table I, we compare different scenarios with or without redundant relays (RR) and coding. We conclude that redundant relays can help to decrease the dissemination time and coding can decrease the collection time. For example, considering the constant transmission time model, with the help of redundant relays and coding, the average dissemination and collection time decrease from  $\mathcal{O}(v \log n)$  to  $\mathcal{O}(v)$ , respectively.

In Table II, we summarize the results of the expected message passing delay in both constant transmission time and random transmission time models. In the constant transmission time model, we deduce the optimal  $k$  and  $n$  from the expression of  $\mathbb{E}[T_{A+B}]$ . However, in the random transmission time model, the exact values of optimal  $k$  and  $n$  are hard to deduce from the expression of  $\mathbb{E}[T_{A+B}]$ .

#### E. Source and Collector’s Starting Time

In the previous subsections, we only consider the message passing delay which defined as the sum of the dissemination

time and collection time. In practice, we should also look into the source and collector’s dynamics on when to start each process. Thus, the overall message passing time, defined as the time from the beginning of the source’s dissemination to the end of the collector’s collection, is also an important performance metric for the message passing system.

To analyze the overall message passing time, in general, two scenarios should be considered: 1) the collector starts collecting chunks after the dissemination. Thus, the overall message passing time is larger than the message passing delay, since there will be some intermediate time between dissemination and collection. Nevertheless, the previous conclusions also hold for the overall message passing time analysis, because the intermediate time is unrelated to the dissemination and collection times. 2) The collector starts collecting chunks during the dissemination. Thus, the overall message passing time is smaller than the message passing delay and the previous conclusions do not always hold for the overall message passing time analysis. This is an interesting problem that needs to be studied in the future.

#### F. Multiple Sources and Collectors With Network Coding for Further Reduction of Delay

With the help of redundant relays and coding, we reduced the dissemination and collection time for both the constant and random transmission time models. However, the effectiveness of redundant relays and coding is limited. For example, we can not reduce the average dissemination time beyond  $\mathcal{O}(v)$  under the constant transmission time model. Nevertheless, we can further reduce the delay by introducing multiple sources and collectors. We outline the main ideas below and leave the details of this problem for a follow-up study.

Based on the results of [28], we know that  $w$  random walks on a complete graph or a  $d$ -regular graph achieve  $\mathcal{O}(w)$  reduction of the single walk graph cover time. (This is true for some other graphs [29] as well.) Thus multiple sources will reach the relays faster than a single source but need to disseminate different message chunks to reduce the message passing delay. Thus, each source must know which chunks have been disseminated by the other sources. Such assumption is impractical and inadequate for covert communications scenarios. Without this assumption, having multiple sources and collectors may even worsen the delay. To see that, consider two sources that disseminate a message of 2 data chunks to 4 relays by randomly selecting a chunk each time they encounter a relay. If both sources disseminate only one data chunk, the dissemination time is significantly shorter than when one source disseminates 2 data chunks. However, the two sources will disseminate the same chunk with a probability 0.5. If both sources disseminate 2 data chunks, the dissemination time is larger than the one source scenario.

The sources can apply various strategies to ensure that they disseminate different chunks. For example, if there are two sources, one can disseminate the even chunks and the other odd chunks. However, we will have higher benefits if we do not fix the number of chunks that each source has to disseminate as long as the sum of the disseminated chunks is  $n$ .

TABLE I  
DISSEMINATION AND COLLECTION TIMES WITH OR WITHOUT REDUNDANT RELAYS AND CODING

	CONSTANT TT <sup>1</sup>		RANDOM TT	
	Dissemination	Collection	Dissemination	Collection
w/ RR <sup>2</sup> and coding	$\mathcal{O}(v)$	$\mathcal{O}(v)$	$\mathcal{O}(v + n)$	$\mathcal{O}(v + k)$
w/o RR and w/ coding	$\mathcal{O}(v \log n)$	$\mathcal{O}(v)$	$\mathcal{O}(v \log n + n)$	$\mathcal{O}(v + k)$
w/ RR and w/o coding	$\mathcal{O}(v)$	$\mathcal{O}(v \log n)$	$\mathcal{O}(v + n)$	$\mathcal{O}(v \log n + n)$
w/o RR and coding	$\mathcal{O}(v \log n)$	$\mathcal{O}(v \log n)$	$\mathcal{O}(v \log n + n)$	$\mathcal{O}(v \log n + n)$

<sup>1</sup> TT is transmission time (cf. II-C).

<sup>2</sup> RR is redundant relays.

TABLE II  
CONCLUSIONS OF IMPORTANT RESULTS IN BOTH DELAY MODELS

MODEL	$\mathbb{E}[\mathbf{T}_{\mathbf{A}+\mathbf{B}}]$	Optimal $k$ (given $n$ )	Optimal $n$ (given $k$ )
Constant TT <sup>1</sup>	$\theta_d v (H_r + H_n - H_{r-n} - H_{n-k})$	$k = 1$	$\lceil \sqrt{rk + k} - 1 \rceil$ or $\lfloor \sqrt{rk + k} - 1 \rfloor$
Random TT	$\frac{n+k}{\lambda} + \left(\frac{n}{k} + 1\right)m + \theta_d v \mathbb{E}[\eta] (H_r + H_n - H_{r-n} - H_{n-k})$	not clear	$k \leq n \leq \sqrt{rk + k} - 1$

<sup>1</sup> TT is transmission time (cf. II-C).

Moreover, the splitting strategy does not reduce the collection delay. We may use network coding to improve the multiple sources/collectors and reduce the collection delay. In this way, instead of sending the data chunks to the relays, the sources will send a random linear combination of the chunks.

For the most general scenario, the sources may not even be able to agree or know which chunks are disseminated before/during the dissemination. Here, using network coding is essential. Each time a source encounters a relay, it randomly generates a coded chunk as a linear combination of  $k$  data chunks. Thus, the sources only need to disseminate  $n$  coded chunks together. In the collection phase, the collectors need to collect totally  $k'$  out of  $n$  chunks, where  $k'$  is just slightly larger than or equal to  $k$ . Then, when the field size is sufficiently large, the  $k$  data chunks can be decoded from  $k'$  coded chunks with high probability. Using fountain codes, e.g., LT [30] or Raptor codes [31], instead of random network codes can simplify decoding.

#### IV. COVERT MESSAGE PASSING

We now consider message passing where in addition to Alice, Bob, and relays, there is another communication participant, Willie the warden. We are interested in the probability that the message passing from Alice and Bob is hidden from Willie who has certain mobility and detection capabilities.

The transmission is covert if Alice passes a message to Bob through relays without being detected by Willie. Although covert communication has been studied in many papers, covert message passing in mobile networks is recently proposed. Therefore, there are few warden detection models. In the following, we provide two possible models.

##### A. Two Warden Detection Models

1) *Random Patrolling Model*: In this model, see Fig. 5 (left), the warden Willie walks randomly on the same graph with Alice and Bob. We assume that all mobile

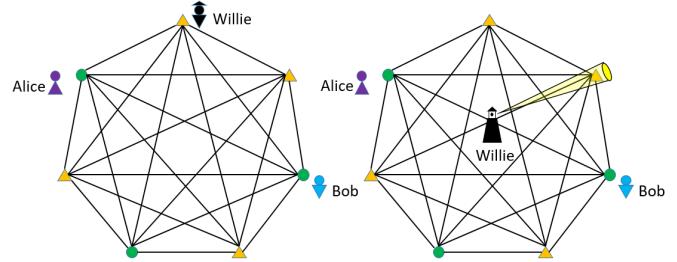


Fig. 5. Two warden models are proposed: random walk (left) and uniform surveillance (right). For the random walk model, Willie walks randomly on the graph with Alice and Bob. He can only detect the communication when he meets Alice or Bob at the relay. For the uniform surveillance model, Willie monitors each vertex uniformly. He can only detect the communication when he monitors the relay during Alice/Bob's transmission.

participants are moving synchronously. To establish that Alice and Bob are communicating, Willie has to 1) meet one of them at a relay and 2) detect the transmission to the relay is taking place before it was over. The more time Willie has to observe a chunk transmission, the higher his probability of detection will be. The detection probability is thus an increasing function of the chunk length  $\ell$ . However, decreasing the chunk length will result in having to pass more chunks, which in turn gives Willie more  $(m/\ell)$  opportunities to observe a transmission.

2) *Uniform Surveillance Model*: In this model, we base covertness on the assumption that the warden Willie can monitor part of the vertices for some given time. A straightforward and informal way to visualize this model is to imagine that the warden is stationed somewhere “in the middle” of the graph,<sup>1</sup> on top of a lighthouse, see Fig. 5 (right). This way, he can only check the part where the lighthouse sheds its light and can not see what is happening behind him. We can also imagine the

<sup>1</sup>This is a very informal statement. There is no need for an exact “middle”, we just need the warden to be at a place where he can observe different parts of the area at different times.

warden applies a Round-Robin protocol to check each vertex of the graph. Thus, he can monitor vertices uniformly.

To formalize this model, we assume that Alice and Bob transmit data to relays without implementing any covertness scheme. Therefore, if the warden happens to check a graph vertex while the data chunk transmission is taking place, he will detect it with probability 1. We further assume that when Alice or Bob starts transmitting on a relay, the warden's time of arrival at this relay follows a uniform distribution. Thus, if the warden monitors the relay during the transmission, he will detect it; otherwise, he will not. This model was first proposed in paper [32] which provided the analysis for a complete graph.

### B. Performance Metrics

1) *Covertness Probability*: Covertness probability is defined as the probability that Alice transmits a message to Bob without being detected by Willie. For example, assume that the message has 2 data chunks, Alice needs to transmit 2 times to 2 relays and Bob also needs 2 times to collect the chunks. If during each time, Willie will detect the transmission with a probability  $P_d$ , the covertness probability is  $P_c = (1 - P_d)^4$ . Notice that when Willie detects the transmission, it does not mean he will get the content of the message. The message may be camouflaged as noise to avoid detection. However, this is another covert communication problem and will not be studied in this paper.

2) *Covertness vs. Delay Tradeoff*: To optimize the covertness probability, we consider using splitting and data redundancy. From the definition of covertness probability, we know that, although splitting increases the number of Willie's detection chances Willie, it also decreases the detection probability  $P_d$  for each detection opportunity. Therefore, it is hard to tell how splitting affects the covertness probability. Meanwhile, introducing data redundancy decreases the covertness probability. Therefore, the tradeoff between the covertness probability and delay is apparent here.

On the one hand, if Alice delivers the data chunks without data redundancy to relays, then the probability of detection is small. This happens because the chance that Willie "sees" her is inversely proportional to the number of nodes. On the other hand, the delay is increased because Bob will have to visit many nodes until he meets all the relays that hold Alice's data chunks. At the other side of the spectrum, if Alice encodes the data chunks and delivers the coded chunks to more relays, then it will take Bob fewer steps to retrieve it. Thus, the message passing delay may reduce (see conclusions in Sec. III). However, the probability that they are caught increases significantly because of the more times that Alice has to deliver to the relays.

### C. Notation

- $P_d$  - detection probability for each transmission
- $P_c$  - covertness probability
- $\beta$  - the number of wardens
- $t_t$  - the message chunk passing time
- $t_a$  - the warden's arrival time

The parameters  $\beta$ ,  $t_t$  and  $t_a$  are the system parameters.  $P_d$  and  $P_c$  are the performance metrics.

### V. COVERTNESS PROBABILITY ANALYSIS

In Sec. IV-B.1, we introduced the covertness probability  $P_c$  and provided an example to calculate  $P_c$  given the transmission detection probability  $P_d$ . In practice,  $P_d$  is not always a constant. The warden Willie will detect the transmission with a higher probability when he "sees" the transmission for a longer time. In another word,  $P_d$  increases with the length of a chunk  $\ell$ . In this section, we provide the expression of  $P_d$  as a function of  $\ell$  for two different warden models, respectively. Meanwhile, we provide the general expression of  $P_c$ .

#### A. Transmission Detection Probability

1) *Random Patrolling Model*: In this model,  $\beta$  wardens, a source, and a collector walk randomly on a regular graph. As we introduced in Sec. IV-A, the warden firstly needs to meet the source or collector, and then detects the transmission. For a  $d$ -regular graph with  $v$  vertices, according to (1), the probability that the warden meets the source/collector at a relay is  $\frac{\beta}{\theta_d v}$ . After meeting the source or collector, we consider two scenarios for the warden: a) the warden will detect the transmission immediately. Thus, the detection probability is

$$P_d = \frac{\beta}{\theta_d v}. \quad (8)$$

The above formula shows that  $P_d$  only relates to the degree of graph  $d$  and the number of wardens  $\beta$ , thus we consider the detection probability as a constant.

b) The warden will fail to detect with some probability. In practice, the warden can not always detect the transmission successfully for some reasons, e.g., the source tries to hide the message, the noise of the detection channel, etc. It is intuitively to imagine that the longer time for the source or collector to transmit the coded chunk, the higher probability the warden detects successfully. Since the transmission time relates to the length of the coded chunk, here, we simply assume the detection probability is a linear function of  $\ell$ .

$$P_d(\ell) = \frac{\ell \beta}{m \theta_d v}. \quad (9)$$

Since  $\ell = \frac{m}{k}$ , we see that  $P_d$  is a function of  $k$ , and it decreases as  $k$  increases.

2) *Uniform Surveillance Model*: In this model, when a source or a collector arrives at a relay and starts transmitting the data chunk, the warden will arrive (monitor) at this relay following a uniform distribution  $U(0, W)$ . As we introduced in Sec. IV-A, whether the warden will detect the transmission successfully depends on the warden's arrival time  $t_a$  and the source/collector's message chunk transmission time  $t_t$ . We consider the warden's arrival time  $t_a$  follows a uniform distribution, i.e.,  $t_a \sim U(0, W)$ . And the message chunk passing time  $t_t$  follows a shifted exponential distribution i.e.,  $t_t \sim S\text{-Exp}(\Delta, \lambda)$  (the tail is given as  $\Pr\{t_t > x\} = e^{-\lambda(t_t - \Delta)}$  for  $t_t > \Delta$ ). Where  $\Delta = \ell$  indicates the time to transmit a chunk and the exponential tail is some inherent additive

system randomness at each relay, which does not depend on the chunk length  $\ell$ . Thus, the detection probability  $P_d$  is given by the following theorem:

*Theorem 4: For the uniform surveillance model, the probability that the warden arrives during the transmission (i.e., detects the transmission) is*

$$P_d(\ell) = \begin{cases} \frac{1}{\lambda W} + \frac{\ell}{W} - \frac{e^{-\lambda(W-\ell)}}{\lambda W} & \text{for } W \geq \ell \\ 1 & \text{for } W < \ell \end{cases} \quad (10)$$

*Proof:* Since  $t_t \sim \text{S-Exp}(\ell, \lambda)$ , it is obvious that  $t_t \geq \ell$ . We assume the warden's arrival time is  $t_a \sim U(0, W)$ , then we have  $t_a \leq W$ .

If  $W < \ell$ , the warden will definitely arrive before the transmission is done, the detection probability is  $P_d = 1$ .

If  $W > \ell$ , we calculate the detection probability as follows:

$$\begin{aligned} P_d(\ell) &= P(t_t \geq t_a) = \int_0^W \int_{t_a}^{\infty} f_t(t_t) f_a(t_a) dt_t dt_a \\ &= \int_{\ell}^W \int_{t_a}^{\infty} f_t(t_t) dt_t f_a(t_a) dt_a \\ &\quad + \int_0^{\ell} \int_{\ell}^{\infty} f_t(t_t) dt_t f_a(t_a) dt_a \\ &= \int_{\ell}^W e^{-\lambda(t_t-\ell)} f_a(t_a) dt_a + \frac{\ell}{W} \\ &= \int_{\ell}^W \frac{1}{W} e^{-\lambda(t_t-\ell)} dt_a + \frac{\ell}{W} \\ &= \frac{1}{\lambda W} + \frac{\ell}{W} - \frac{e^{-\lambda(W-\ell)}}{\lambda W}. \end{aligned}$$

□

### B. Covertness Probability

The communication between the source and collector stays covert only when all the coded chunk transmissions are undetected. Recall that the source needs to disseminate the  $n$  coded chunks and the collector needs to collect  $k$  coded chunks. Therefore, the total number of coded chunk transmissions is  $n+k$  and the covertness probability is

$$P_c(k) = (1 - P_d)^{n+k}. \quad (11)$$

When  $k$  is given, since  $P_d$  and  $n$  are independent and  $0 \leq P_d \leq 1$ , the covertness probability  $P_c$  decreases with increasing  $n$ . Since  $P_d$  is a function of  $\ell$  and  $\ell = \frac{m}{k}$ ,  $P_c$  is a function of  $k$ . Thus, when  $n$  is given, we need further analyze the optimal  $k$  for each detection model.

For the random patrolling model,  $P_d$  is a constant under the worst-case (see (8)). Otherwise,  $P_d$  increases linearly with  $\ell$  (see (9)). Thus, we can easily get the expression of  $P_c$  by substituting (8) or (9) to (11). From the expressions, we find the maximum of the covertness probability in Theorem 5.

*Theorem 5: For the random patrolling model, given the number of coded chunks  $n$ , the covertness probability  $P_c$  changes as follows:*

- 1) Under the worst-case,  $P_d = \frac{\beta}{\theta_{dv}}$  is a constant. Thus, for a given  $n$ ,  $P_c$  decreases with increasing  $k$ , which means  $P_c$  reaches the maximum at  $k = 1$ .

- 2) Otherwise,  $P_d(k) = \frac{\beta}{k\theta_{dv}}$  is a function of  $k$ . Thus,  $P_c$  increases with  $k$ , which means  $P_c$  reaches the maximum at  $k = n$ .

*Proof:* For property 1), since  $P_d$  is a constant and  $0 \leq 1 - P_d < 1$ , it is obvious that  $P_c$  decreases with  $k$ . Since  $k \geq 1$  is an integer, the maximum covertness probability is at  $k = 1$ , i.e.,  $P_c = (1 - P_d)^{n+1}$ .

For property 2), we know  $1 \leq k \leq n$  is an integer. Let  $a \in [1, n]$  be an integer. To prove  $P_c(k)$  increases with  $k$ , we only need to show  $P_c(k = a+1) > P_c(k = a)$  for all  $a$ .

According to (9) and  $\ell = \frac{m}{k}$ , we have  $P_d(k) = \frac{\beta}{k\theta_{dv}}$ . Define  $B = \frac{\beta}{\theta_{dv}}$ , then  $P_d(k) = \frac{B}{k}$ . Therefore, we only need to show  $(1 - B/(a+1))^{n+a+1} > (1 - B/a)^{n+a}$ .

Notice that  $0 < B \leq 1$  and  $1 \leq a \leq n$ , we can transform the inequality into  $\frac{a}{a-B} > \frac{a+1}{a+1-B} \sqrt[n+k]{\frac{a+1}{a+1-B}}$ . Thus, we get  $\frac{a(a+1-B)}{(a-B)(a+1)} > \sqrt[n+k]{\frac{a+1}{a+1-B}}$ .

For the left term, we have  $\frac{a(a+1-B)}{(a-B)(a+1)} = 1 + \frac{B}{(a-B)(a+1)} > 1 + \frac{B}{(a-B+1)(a+n)} = \frac{1}{n+a} \left( \frac{a+1}{a+1-B} + n + a - 1 \right)$ .

According to the arithmetic mean-geometric mean inequality,  $\frac{1}{n+a} \left( \frac{a+1}{a+1-B} + n + a - 1 \right) > \sqrt[n+k]{\frac{a+1}{a+1-B}}$ .

Finally, for any integer  $1 \leq a < n$ , we have  $P_c(k = a+1) > P_c(k = a)$ . □

For the uniform surveillance model, since  $P_d$  is a function of  $k$ ,  $P_c$  is also a function of  $k$ . From Theorem 4, the term  $\frac{\ell}{W} (= \frac{m}{kW})$  decreases as  $k$  increases, and the term  $-\frac{e^{-\lambda(W-\ell)}}{\lambda W} (= -\frac{e^{-\lambda(W-m/k)}}{\lambda W})$  increases with  $k$ . When  $\lambda$  is sufficiently large, the former term is much larger than the latter term. Thus,  $P_d$  also decreases as  $k$  increases. Since  $P_c$  decreases as  $k$  increases when  $P_d$  is given, there exists a tradeoff for  $k$  to maximize  $P_c$ . When  $\lambda$  is sufficiently small, the former term is smaller than the latter term. Thus,  $P_d$  increases with  $k$ , which leads to  $P_c$  reaches its maximum at  $k = 1$ .

## VI. DELAY VS. COVERTNESS PROBABILITY TRADEOFFS

In the previous sections, we know that introducing data redundancy reduces both the message passing delay and the covertness probability. Thus, it is important to study how much redundancy affects the delay and probability tradeoffs.

### A. Constant Transmission Time vs. Random Patrolling

Since both the random patrolling model and the constant transmission time model do not relate to the message chunk passing time, we combine these two models to analyze the tradeoff between the covertness probability  $P_c$  and the expected message passing delay  $\mathbb{E}[T_{A+B}]$ .

In Fig. 6, we evaluate (3) and (11) to see the tradeoff between the detection probability and the transmission delay for the worst case warden's detection probability  $P_d = \frac{\beta}{\theta_{dv}}$ . We consider a regular graph with  $v = 100$  vertices and the degree is  $d = 5$ . There are 15 relays distributed uniformly on vertices of the graph. 10 wardens, 1 source and 1 collector walk randomly on the graph. We evaluate  $P_c$  vs.  $\mathbb{E}[T_{A+B}]$  as the number of coded chunks  $n$  increases from 1 to 10. Each point in the figure is a different case with different values of  $(n, k)$ . Two different strategies are considered: minimum

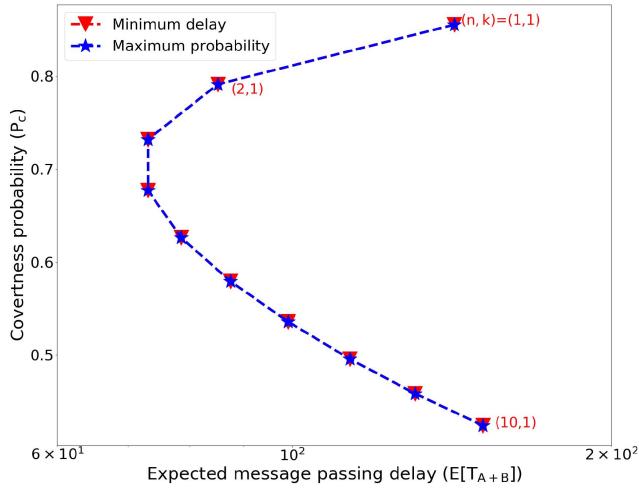


Fig. 6. Covertness probability  $P_c$  (cf. (3)) vs. the expected number of steps  $\mathbb{E}[T_{A+B}]$  (cf. (11)) as  $n$  increases from 1 to 10. We consider the worst-case, i.e., the detection probability is a constant (cf. (8)). This is a regular graph with 100 vertices and the degree is 5. The number of wardens is  $\beta = 10$  and the number of relays is  $r = 15$ . There exists an  $(n, k)$  which simultaneously maximizes  $P_c$  and minimizes  $\mathbb{E}[T_{A+B}]$ .

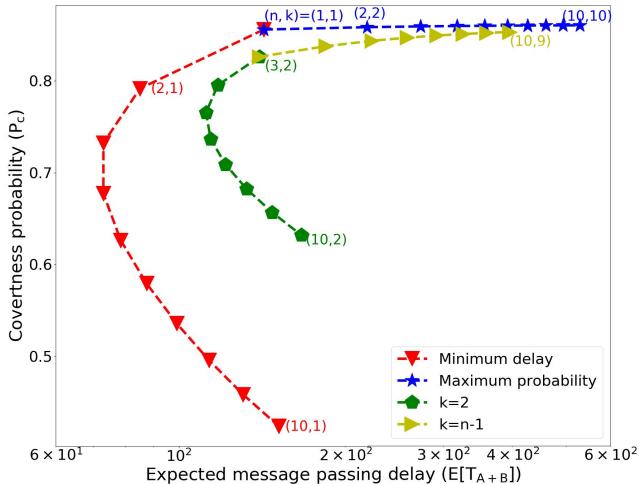


Fig. 7. Covertness probability  $P_c$  (cf. (3)) vs. the expected message passing delay  $\mathbb{E}[T_{A+B}]$  (cf. (11)) as  $n$  increases from 1 to 10. We consider the detection probability is a function of  $\ell$  (cf. 9). This is a regular graph with 100 vertices and the degree is 5. The number of wardens is  $\beta = 10$  and the number of relays is  $r = 15$ . It is impossible to reach the maximum  $P_c$  and the minimum  $\mathbb{E}[T_{A+B}]$  simultaneously.

delay and maximum probability. For the minimum delay, we firstly find the optimal  $k$  which minimizes  $\mathbb{E}[T_{A+B}]$  given each  $n$ . Then we calculate  $P_c$  by using the same  $k$ . For the maximum probability, on the contrary, we firstly find the optimal  $k$  for  $P_c$  given each  $n$ , and then calculate  $\mathbb{E}[T_{A+B}]$ . From Fig. 6, we observe that the results for both strategies coincide with each other. It means that the optimal  $k$  simultaneously maximizes  $P_c$  and minimizes  $\mathbb{E}[T_{A+B}]$ . We also observe that  $P_c$  decreases with increasing  $n$  and  $\mathbb{E}[T_{A+B}]$  reaches the minimum at  $n = 4$ . These observations are consistent with Theorems 1 and 5. Finally, we conclude that  $k = 1$ , i.e., no splitting, is the overall optimal strategy. Replication ( $n > 1$  and  $k = 1$ ) can reduce the transmission delay, but it also decreases the covertness probability.

In Fig. 7, we consider the warden's detection probability increases linearly with the length of data chunk  $\ell$ , i.e.,

$P_d = \frac{\ell\beta}{\ell\beta + \nu}$  (where  $\ell = \frac{m}{k}$ ). The same regular graph is adopted from Fig. 6. We evaluate  $P_c$  vs.  $\mathbb{E}[T_{A+B}]$  as  $n$  increases from 1 to 10. Four different strategies are considered: minimum delay, maximum probability, " $k = 2$ " and " $k = n - 1$ ". For the minimum delay, we have the same observations as them in Fig. 6. This is because the expression of  $\mathbb{E}[T_{A+B}]$  is unchanged, and it still reaches minimum at  $(n, k) = (4, 1)$ . For the maximum probability,  $P_c$  reaches minimum at  $k = n$ , which means splitting is optimal. It is obvious that the curves of the minimum delay and the maximum probability are far from each other. Therefore, we can not simultaneously decrease the message passing delay and increase the covertness probability. In the figure, we also see the results for " $k = 2$ " and " $k = n - 1$ " strategies. When we have a minimum requirement for  $P_c$  or  $\mathbb{E}[T_{A+B}]$ , then  $k = 2$  or  $k = n - 1$  strategies may provide a better performance. For example, if we require  $P_c > 0.8$ , then the case  $(n, k) = (4, 2)$  gives a relatively smaller  $\mathbb{E}[T_{A+B}]$ .

### B. Random Transmission Time vs. Uniform Surveillance

Since both the uniform surveillance detection model and the random transmission time model relate to the message chunk passing time, we combine these two models to analyze the tradeoff between the covertness probability  $P_c$  and the expected message passing delay  $\mathbb{E}[T_{A+B}]$ .

In Fig. 8, we evaluate (7), (10) and (11) to see the tradeoff between the detection probability and the expected message passing delay. The same regular graph is adopted from Fig. 6. The upper subfigure has a message chunk passing time  $t_t \sim S\text{-Exp}(10/k, 1)$  and the lower subfigure has a message chunk passing time  $t_t \sim S\text{-Exp}(10/k, 0.2)$ . The warden's arrival time follows  $U(0, 30)$ . We evaluate  $P_c$  vs.  $\mathbb{E}[T_{A+B}]$  as  $n$  increases from 1 to 10. Each point in the figure is a different case with a different value of  $(n, k)$ . Similarly, the minimum delay and maximum probability strategies are considered.

The upper subfigure shows that given the number of coded chunks  $n$ , minimum delay and maximum probability strategies have very different optimal  $k$  values. For minimum delay, the optimal  $k$  is small (i.e., 1 or 2); For maximum probability, the optimal  $k$  is equal to  $n$ . Therefore, we conclude that it is impossible to simultaneously increase the covertness probability and reduce the message passing delay. In practice, the appropriate  $(n, k)$  is also decided by the requirements for the performance metrics, e.g., the system requires  $P_c > 0.7$  or  $\mathbb{E}[T_{A+B}] < 170$ . The lower subfigure shows that given the number of coded chunks  $n$ , when  $n$  is small (e.g.,  $n < 5$ ), minimum delay and maximum probability strategies have very different optimal  $k$  values; When  $n$  is large (e.g.,  $n > 6$ ), they have very closed optimal  $k$  values. When  $n > 6$ , the optimal  $k$  for minimum delay is equal to 2 and the optimal  $k$  for maximum probability is equal to 5. Therefore, we conclude that as the variance of  $t_t$  increases ( $1/\lambda^2$  becomes smaller), the maximum probability and minimum delay strategies will finally be the same. The conclusion indicates that when  $\lambda$  is sufficiently small, we can simultaneously increase the covertness probability and reduce the message passing delay.

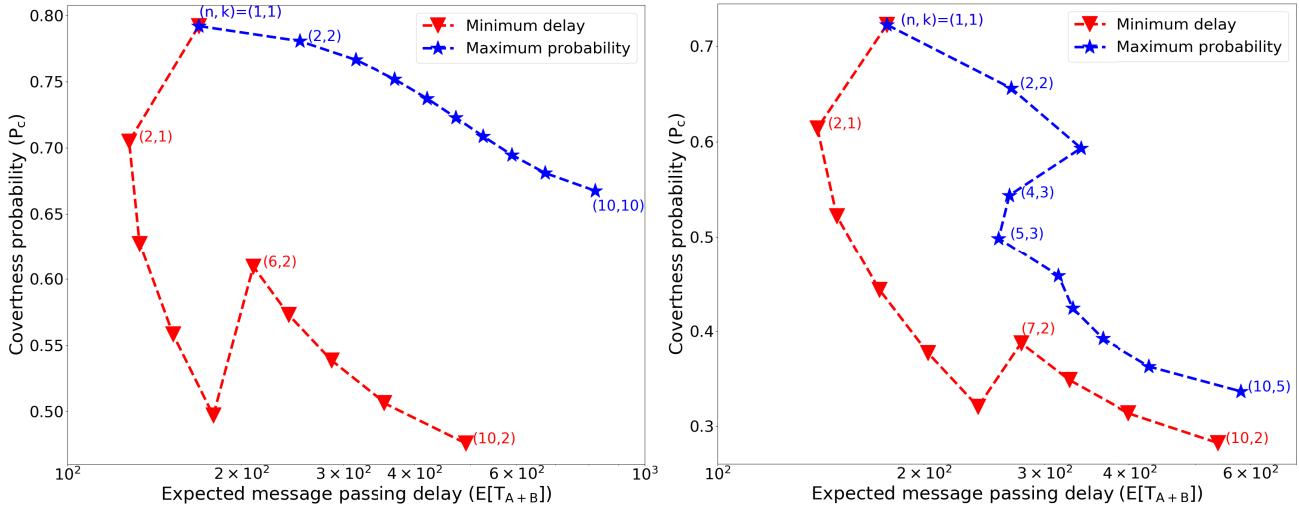


Fig. 8. Covertness probability  $P_c$  (cf. (7)) vs. the expected message passing delay  $E[T_{A+B}]$  (cf. (10) and (11)) as  $n$  increases from 1 to 10. The upper subfigure has a message chunk passing time  $t_l \sim S\text{-Exp}(10/k, 1)$  and the lower subfigure has the  $t_l \sim S\text{-Exp}(10/k, 0.2)$ . This is a regular graph with 100 vertices and the degree is 5. The number of relays is  $r = 10$ . The warden's arrival time follows  $U(0, 100)$ . It is impossible to reach maximum  $P_c$  and minimum  $E[T_{A+B}]$  simultaneously. Some sacrifice of covertness may result in a significant improvement in delay.

From both subfigures, we see that the minimum delay strategy leads to similar results, but the maximum probability strategy leads to very different results. This is because  $P_c$  is sensitive to the value of  $\lambda$ , but  $E[T_{A+B}]$  is not. We analyzed how  $P_c$  changes with  $\lambda$  in Sec. V-B. According to the expression of  $E[T_{A+B}]$  given in (7), although decreasing  $\lambda$  from 1 to 0.2 can increase the former term  $\frac{n+k}{\lambda}$ , the latter term  $\theta_{dv}(H_r + H_n - H_{r-n} - H_{n-k})$  is far larger than the former. Thus, the results will not change much with  $\lambda$ . Notice that both  $P_c$  and  $E[T_{A+B}]$  reach the optimal values when  $n$  and  $k$  are small. However, that may not be the case when the message length  $m$  is sufficiently large. Recall that the detection probability  $P_d = 1$  (in Theorem 4) when  $\ell > W$ , then the covertness probability  $P_c = 0$ . Since  $\ell = \frac{m}{k}$ , when  $m \gg W$  (e.g.,  $m = 3W$ ), a small  $(n, k)$  value (e.g., (4, 2)) leads to  $P_c = 0$ . Therefore, it is necessary to study how the  $P_c$  and  $E[T_{A+B}]$  trade-off changes when  $n$  and  $k$  are large in Fig. 8.

## VII. CONCLUSION AND FUTURE WORK

We introduced and studied a gossip-like protocol for covert passing messages between Alice and Bob as they move in an area hosting a multitude of IoT objects. Alice and Bob perform random walks on random regular graphs. The IoT objects reside on the vertices of this graph, and some can serve as relays between Alice and Bob. In our protocol, Alice splits her message into small chunks, which she can covertly deposit to the relays she encounters. Afterward, Bob collects the chunks. Alice may encode her data before the dissemination. The area where the message passing takes place is watched over by a warden Willie. Willie can either perform random walks as Alice and Bob do or conduct uniform surveillance of the area. In either case, he can only observe one relay at a time. We evaluated the system performance by the covertness probability and the message passing delay. These performance metrics depend on the graph, communications delay, and code parameters. We showed that, in most scenarios, it is impossible to choose the design parameters that maximize the covertness probability and minimize the message delay simultaneously.

This work sets the stage for many problems of interest to be studied in the future. We briefly describe five directions of immediate interest.

1) *Reducing Delay With Multiple Sources and Collectors:* According to Section III-F, although many random walks are faster than one [28], simply introducing multiple sources and collectors may not reduce the dissemination and collection time significantly. Fountain codes need further study as a possible way to improve the performance.

2) *Computing the Overall Message Passing Time:* As we discussed in Section III-E, the overall time the message spends in the system depends on when the collection and dissemination start. For example, Alice and Bob can start their walks each day at some specified time, or collecting can start before the dissemination is over. Extending this work to include such dynamics is thus of interest.

3) *Extending Analysis to Other Detection Models and Mobility Patterns:* Some other detection models are reasonable but have not been studied yet. For example, regarding the warden, it is reasonable to assume that Willie needs to spend a specific time before detecting the transmission or that he can, over time, learn which nodes do not have relays. Regarding the mobility patterns, random walks on irregular graphs or some other area traversing models are of interest.

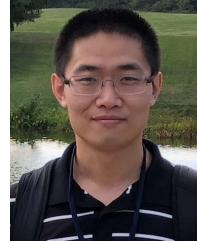
4) *Computing Trade-off Between False-Alarm and Missed-Detection:* Some IoT devices may help Alice and Bob achieve covert communication. For example, some devices can deceive the warden into wrongly accusing Alice and Bob of the message passing. In such scenarios, we should consider the tradeoff between false-alarm and missed-detection. Another scenario where false-alarm and missed-detection are of interest is when the transmission is organized through incremental data redundancy in a classical way (see, e.g., [33], [34]). Rather than assuming a perfect detection of the message chunk, we here assume that each relay sends a very noisy version of the entire message to reduce the detection probability. The

collector then recovers the message by, e.g., Chase combining of the received noisy versions.

5) *Studying Systems With Unreliable Relays:* In the IoT scenarios, we expect some of the relays to be adversarial. In general, these objects have power constraints, and we may not rely on all of them to provide the required storage service. Including such impairments is of interest for further study.

## REFERENCES

- [1] B. A. Bash, D. Goeckel, and D. Towsley, “Limits of reliable communication with low probability of detection on AWGN channels,” *IEEE J. Sel. Areas Commun.*, vol. 31, no. 9, pp. 1921–1930, Sep. 2013.
- [2] B. A. Bash, D. Goeckel, D. Towsley, and S. Guha, “Hiding information in noise: Fundamental limits of covert wireless communication,” *IEEE Commun. Mag.*, vol. 53, no. 12, pp. 26–31, Dec. 2015.
- [3] M. R. Bloch, “Covert communication over noisy channels: A resolvability perspective,” *IEEE Trans. Inf. Theory*, vol. 62, no. 5, pp. 2334–2354, May 2016.
- [4] S. Kadhe, S. Jaggi, M. Bakshi, and A. Sprintson, “Reliable, deniable, and hidable communication over multipath networks,” in *Proc. IEEE Int. Symp. Inf. Theory*, Jun. 2014, pp. 611–615.
- [5] L. Wang, G. W. Wornell, and L. Zheng, “Fundamental limits of communication with low probability of detection,” *IEEE Trans. Inf. Theory*, vol. 62, no. 6, pp. 3493–3503, Jun. 2016.
- [6] T. Sobers, B. A. Bash, S. Guha, D. Towsley, and D. Goeckel, “Covert communication in the presence of an uninformed jammer,” *IEEE Trans. Wireless Commun.*, vol. 16, no. 9, pp. 6193–6206, Sep. 2017.
- [7] K. Shahzad, X. Zhou, and S. Yan, “Covert communication in fading channels under channel uncertainty,” in *Proc. IEEE 85th Veh. Technol. Conf. (VTC Spring)*, Jun. 2017, pp. 1–5.
- [8] J. Hu, S. Yan, X. Zhou, F. Shu, J. Li, and J. Wang, “Covert communication achieved by a greedy relay in wireless networks,” *IEEE Trans. Wireless Commun.*, vol. 17, no. 7, pp. 4766–4779, Jul. 2018.
- [9] J. Hu, S. Yan, F. Shu, and J. Wang, “Covert transmission with a self-sustained relay,” *IEEE Trans. Wireless Commun.*, vol. 18, no. 8, pp. 4089–4102, Aug. 2019.
- [10] A. R. Girard, A. S. Howell, and J. K. Hedrick, “Border patrol and surveillance missions using multiple unmanned air vehicles,” in *Proc. 43rd IEEE Conf. Decis. Control (CDC)*, vol. 1, Dec. 2004, pp. 620–625.
- [11] A. Puri, “A survey of unmanned aerial vehicles (UAV) for traffic surveillance,” Dept. Comput. Sci. Eng., Univ. South Florida, Tampa, FL, USA, Tech. Rep., 2005, pp. 1–29.
- [12] D. A. Paley, F. Zhang, and N. E. Leonard, “Cooperative control for ocean sampling: The glider coordinated control system,” *IEEE Trans. Control Syst. Technol.*, vol. 16, no. 4, pp. 735–744, Jul. 2008.
- [13] T. F. Villa, F. Salimi, K. Morton, L. Morawska, and F. Gonzalez, “Development and validation of a UAV based system for air pollution measurements,” *Sensors*, vol. 16, no. 12, p. 2202, 2016.
- [14] E. Soljanin, “Reducing delay with coding in (mobile) multi-agent information transfer,” in *Proc. 48th Annu. Allerton Conf. Commun., Control, Comput. (Allerton)*, Sep. 2010, pp. 1428–1433.
- [15] V. Tripathi, R. Talak, and E. Modiano, “Age optimal information gathering and dissemination on graphs,” *IEEE Trans. Mobile Comput.*, early access, Apr. 30, 2021, doi: [10.1109/TMC.2021.3076755](https://doi.org/10.1109/TMC.2021.3076755).
- [16] G. Joshi, Y. Liu, and E. Soljanin, “Coding for fast content download,” in *Proc. 50th Annu. Allerton Conf. Commun., Control, Comput. (Allerton)*, Oct. 2012, pp. 326–333.
- [17] M. F. Aktas and E. Soljanin, “Straggler mitigation at scale,” *IEEE/ACM Trans. Netw.*, vol. 27, no. 6, pp. 2266–2279, Dec. 2019.
- [18] P. Peng, E. Soljanin, and P. Whiting, “Diversity vs. parallelism in distributed computing with redundancy,” in *Proc. IEEE Int. Symp. Inf. Theory (ISIT)*, Jun. 2020, pp. 257–262.
- [19] Q. Yu, M. A. Maddah-Ali, and A. S. Avestimehr, “Straggler mitigation in distributed matrix multiplication: Fundamental limits and optimal coding,” *IEEE Trans. Inf. Theory*, vol. 66, no. 3, pp. 1920–1933, Mar. 2020.
- [20] A. Badita, P. Parag, and V. Aggarwal, “Optimal server selection for straggler mitigation,” *IEEE/ACM Trans. Netw.*, vol. 28, no. 2, pp. 709–721, Apr. 2020.
- [21] A. Behrouzi-Far and E. Soljanin, “Efficient replication for fast and predictable performance in distributed computing,” *IEEE/ACM Trans. Netw.*, vol. 29, no. 4, pp. 1467–1476, Aug. 2021.
- [22] P. Gupta and P. R. Kumar, “The capacity of wireless networks,” *IEEE Trans. Inf. Theory*, vol. 46, no. 2, pp. 388–404, Mar. 2006.
- [23] M. Grossglauser and D. N. C. Tse, “Mobility increases the capacity of ad hoc wireless networks,” *IEEE/ACM Trans. Netw.*, vol. 10, no. 4, pp. 477–486, Aug. 2002.
- [24] A. E. Gamal, J. Mammen, B. Prabhakar, and D. Shah, “Optimal throughput-delay scaling in wireless networks—Part I: The fluid model,” *IEEE Trans. Inf. Theory*, vol. 52, no. 6, pp. 2568–2592, Jun. 2006.
- [25] Z. Kong, E. M. Yeh, and E. Soljanin, “Coding improves the throughput-delay tradeoff in mobile wireless networks,” *IEEE Trans. Inf. Theory*, vol. 58, no. 11, pp. 6894–6906, Nov. 2012.
- [26] H.-S. Im and S.-H. Lee, “Mobility-assisted covert communication over wireless ad hoc networks,” *IEEE Trans. Inf. Forensics Security*, vol. 16, pp. 1768–1781, 2021.
- [27] C. Cooper, A. Frieze, and T. Radzik, “Multiple random walks in random regular graphs,” *SIAM J. Discrete Math.*, vol. 23, no. 4, pp. 1738–1761, Jan. 2010.
- [28] N. Alon, C. Avin, M. Koucky, G. Kozma, Z. Lotker, and M. R. Tuttle, “Many random walks are faster than one,” in *Proc. 20th Annu. Symp. Parallelism Algorithms Archit. (SPAA)*, 2008, pp. 119–128.
- [29] R. Elsässer and T. Sauerwald, “Tight bounds for the cover time of multiple random walks,” *Theor. Comput. Sci.*, vol. 412, no. 24, pp. 2623–2641, May 2011.
- [30] M. Luby, “LT codes,” in *Proc. 43rd Annu. IEEE Symp. Found. Comput. Sci.*, 2002, p. 271.
- [31] A. Shokrollahi, “Raptor codes,” *IEEE Trans. Inf. Theory*, vol. 52, no. 6, pp. 2551–2567, Jun. 2006.
- [32] P. Peng, N. Melissaris, E. Soljanin, B. Lee, A. Maliev, and H. Fan, “Straggling for covert message passing on complete graphs,” in *Proc. 57th Annu. Allerton Conf. Commun., Control, Comput. (Allerton)*, Sep. 2019, pp. 453–459.
- [33] R. Liu, P. Spasojevic, and E. Sojanin, “Punctured turbo code ensembles,” in *Proc. IEEE Inf. Theory Workshop*, Paris, France, Apr. 2003, pp. 249–252.
- [34] N. Varnica, E. Soljanin, and P. Whiting, “LDPC code ensembles for incremental redundancy hybrid ARQ,” in *Proc. Int. Symp. Inf. Theory (ISIT)*, Adelaide, SA, Australia, Sep. 2005, pp. 995–999.



**Pei Peng** received the B.S. degree in information engineering from the South China University of Technology, Guangzhou, China, in 2011, the M.S. degree in electronics and communication engineering from the Shanghai Institute of Microsystem and Information Technology, Chinese Academy of Science, Shanghai, China, in 2014, and the Ph.D. degree in electrical and computer engineering from Rutgers, The State University of New Jersey, New Brunswick, NJ, USA, in 2022. He is currently an Assistant Professor at the Nanjing University of Posts and Telecommunications. His research interests include coding and allocation in distributed computing systems, covert communications, and machine learning.



**Emina Soljanin** (Fellow, IEEE) is currently a Professor at Rutgers University. Before moving to Rutgers in 2016, she was a (Distinguished) Member of Technical Staff for 21 years at the Mathematical Sciences Research of Bell Labs. Her interests and expertise are wide. Over the past quarter of the century, she has participated in numerous research and business projects, as diverse as power system optimization, magnetic recording, color space quantization, hybrid ARQ, network coding, data and network security, distributed systems performance analysis, and quantum information theory. She served as an Associate Editor for *Coding Techniques*, the IEEE TRANSACTIONS ON INFORMATION THEORY. She also served on the Information Theory Society Board of Governors and in various roles on other journal editorial boards and conference program committees. She is an Outstanding Alumnus of the Texas A&M School of Engineering, the 2011 Padovani Lecturer, a Distinguished Lecturer from 2016 to 2017, and the 2019 President of the IEEE Information Theory Society.