Array 14 (2022) 100131

Contents lists available at ScienceDirect

Array

journal homepage: www.elsevier.com/locate/array

A generalized approach to estimation of memoryless covert channel information leakage capacity[☆]

Baki Berkay Yilmaz[a,*], Nader Sehatbakhsh[b], Moumita Dey[a], Chia-Lin Cheng[a], Milos Prvulovic[c], Alenka Zajić[a]

[a] Electrical and Computer Engineering, Georgia Institute of Technology, Atlanta, GA, 30332, USA
[b] Electrical and Computer Engineering, University of California, Los Angeles, CA, 90095, USA
[c] Computer Science, Georgia Institute of Technology, Atlanta, GA, 30332, USA

ARTICLE INFO

Keywords:
Asynchronous wireless communication
Digital/analog covert channels
Leakage capacity
Security
Side-channels

ABSTRACT

Estimating the amount of information that is leaked by covert channels is a necessity to comprehend and mitigate the severity of attacks exploiting these channels. Having such an estimation in *design-state* provides an opportunity for designers to adjust their systems to minimize information leakage. In this paper, we propose a methodology to estimate the worst-case information leakage (or capacity for the information leakage) through various memoryless covert channels – both analog and digital ones – exhibiting on–off keying structure. In that respect, we first model the communication channel as a *deletion–insertion channel* to account for the information losses due to software activities. Then, we derive the effective noise in covert channels as the combination of jitter noise caused by signaling time variation and Additive White Gaussian Noise (AWGN). Considering this effective noise, we propose a communication model which can be *generalized* for various covert channels and takes insertions, deletions, and asynchronous nature of covert channels into account. By leveraging the link between this communication model and information theory literature, we obtain the information leakage capacity that reveals the leakage limit for the worst-case scenarios. Finally, we provide experimental results to demonstrate that the proposed model is an effective and a generalized methodology to score the resilience of a given system to covert channel attacks.

1. Introduction

With the breakthrough improvements in computer systems and immense inter-connectivity of Internet-of-Things (IoT) devices, computing systems, especially mobile and embedded systems, have become the main storing space to keep sensitive data (e.g., financial data, personal information, etc.). Unfortunately, this creates opportunities for attackers to steal *secret* data by exploiting side-channel signals from these computing devices. In that respect, an extensive research is performed on cryptographic devices and their vulnerability against side channels [1,2]. They prove that side-channel attacks can circumvent the existing defense mechanisms and has to be considered while designing systems.

Side-channels are unintentionally generated due to digital and/or analog characteristics of computers while computers execute programs. If these channels are generated deliberately by manipulating an application's software, they are called *covert channels* [3]. The most important feature of covert channels is that they can easily circumvent existing protection mechanisms (e.g., isolation) deployed on a computer to *exfiltrate* sensitive data. For example, it is shown that a rogue employee can exfiltrate sensitive information about a project (stored in a secured laptop) by leveraging side-channels to establish a covert channel even if the computer's Input/Output (I/O) ports are fully monitored, and the computer is completely isolated from the network [4].

Depending on the exploited side-channel, covert channels can be classified into two broad categories of *digital/micro-architectural* and *analog/physical* covert channels. In both cases, the attacker uses a malicious process (called *source*) which, in one way or another (as will be discussed later), has access to secret data but *does not* have access to the outside world (i.e., the source cannot "extract" the sensitive information by itself through conventional channels without being scrutinized). To circumvent the protection mechanisms, the *source* exploits

[☆] This work has been supported, in part, by National Science Foundation, USA grant 1563991 and Defense Advanced Research Projects Agency, USA LADS contract FA8650-16-C-7620. The views and findings in this paper are those of the authors and do not necessarily reflect the views of NSF and DARPA.

[*] Corresponding author.
E-mail address: byilmaz7@gatech.edu (B.B. Yilmaz).

https://doi.org/10.1016/j.array.2022.100131
Received 18 July 2021; Received in revised form 21 November 2021; Accepted 23 January 2022
Available online 5 February 2022

an existing side-channel (either digital such as cache [5] or analog such as electromagnetic emanations [6]), to communicate the information to a *sink* (i.e., either to another process or to the outside world).

In the literature, there are many examples illustrating that different digital and/or analog characteristics of computers can be exploited to establish a reliable covert channel (e.g., caches [5,7–12], electromagnetic emanations [6,13–16], variation in power consumption [17–25], etc.). Given this extensive set of vulnerabilities and their non-negligible threat to the system security, it is important to measure the severity of leakage through covert channels (in terms of channel capacity) for a given design in both modern systems (e.g., laptops) and embedded/IoT devices.

Millen was the first to propose a method to establish a connection between information theory and flow models [26]. With such a model, Millen defined severity of covert channels in terms of conventional channel capacity assuming covert channels are synchronous communication channels. Similarly, a Gaussian channel model is proposed in [27] for template attacks based on power side channels. The model assumes each template takes equal amount of time and there is no deviation in template timing. In [28], another Gaussian channel model is proposed for the leakages to calculate the success rate of an attack targeting the implementations that are protected by masking. These channel models (which exhibit the footprints of an ideal communication system) are proposed to designate the best distinguisher for a side channel attack to cryptographic devices. Unlike conventional communication systems which are carefully designed such that transmitters and receivers are well-synchronized, covert channels are unintended, hidden, and undesired channels that do not exhibit such characteristics. Furthermore, they are corrupted not only by additive channel noise but also by deviation in signaling time, insertions, and deletions of transmitted bits.

In the communication literature, many papers discuss *(a)* bounds on the capacity of channels corrupted by synchronization errors [29–33] and *(b)* bounds on the capacity of channels corrupted by synchronization and substitution errors [34,35]. Likewise, information theoretic metrics are applied to side channel attacks to calculate the available information to an adversary under the assumption that leakages are Gaussian and all signal traces belong to the targeted cryptographic function [36]. In [37], the success rate estimation is easily computed by utilizing SNR and bounds for success rate is proposed. The effect of shuffling and higher order masking on the information leakage has been also investigated by utilizing information theoretic models. In [38], the security analysis of a cryptographic device has been given when higher order masking is used in Differential Power Analysis (DPA). Mutual information between the leakages and security keys of a cryptographic device with shuffled implementations are considered for worst-case security analysis in [39]. The noisy information leakage models and information theoretic approach also lead to security proofs for masked implementation of block ciphers and divide-and-conquer side channel attacks [40,41]. Although all these approaches aim to quantify and utilize the available information to an attacker, applying these ideas to measure covert channel leakage capacity has shortcomings because they do not consider the problems that covert channels face due to unprecedented computer activities, i.e. insertions, deletions, the variation in signaling time.

Furthermore, a *micro-level* investigation on side-channel capacities was conducted in prior work [42,43], assuming instructions (which are the lowest level order to computer processor) are the transmitted symbols between the transmitter (source) and the receiver (sink). These papers exploit emanated electromagnetic (EM) signal power while executing instructions and model the communication based on the differences in the signal power levels of different instructions. However, employing the *micro-level* capacity definitions to covert channels *overestimates* the leakage capacities because they do not consider insertions and deletions that are encountered in *macro-level* (program level) scenarios. In addition, these models cannot be generalized to measure the leakage capacity of other side-channels because the models are specifically designed for EM side-channels based on individual or pairwise signal power.

To avoid these problems, recent work in [44,45] proposed a program-level communication model including insertions and deviation in signaling time while transmitting signals. Furthermore, leakage capacity bounds were defined for only EM covert channels under the assumption that signaling time deviation can be modeled by changing the position of the pulses and by keeping the pulse width fixed. However, the scope of these papers is limited to EM based covert channels, ignores the losses due to deletions, fixes the pulse width while modeling the channel, and provides capacity bounds instead of actual capacity values. Moreover, these papers consider only the scenario where the signal is silent while transmitting a zero bit, and the signaling time distribution has zero mean and the same standard deviation irrespective of the bit. Interestingly, as we experimentally illustrate in this paper, for different covert channels these assumptions are too optimistic because the bit signals are generated by running different parts of a program or completely different benchmarks [46].

To address these issues, in this paper, we propose a *generalized* communication model for various covert channels, which considers insertions, deletions and their asynchronous nature, to calculate actual leakage capacity. The main contributions of the paper can be listed as follows:

- We propose a communication model for covert channels which considers insertions and deletions to comply with software activities.
- We experimentally demonstrate that the distribution of signaling time variation exhibits a **Gaussian** behavior, hence, can be characterized as a *Gaussian* distribution with mean μ and standard deviation σ.
- We mathematically show that jitter error (error due to variation in signaling time) can be combined with additive channel noise. We call it effective channel noise and experimentally demonstrate that the behavior of this effective channel noise can change for different symbols.
- Based on the communication model and combined effective channel noise, we propose a discrete memoryless channel model to calculate the worst-case leakage through a covert channel. With this channel model, we obtain actual leakage capacity.
- The proposed model can be generalized to various covert channels, therefore, the same structure can be utilized by system designers to assess security of their systems against different types of *already existing* covert channels.

The rest of the paper is organized as follows: In Section 2, we introduce the model for transmitted signal, receiver, and communication channel. Section 3 provides the derivation of effective channel noise and leakage capacity. Section 4 demonstrates how the proposed model can be used for various types of covert channels. Section 5 provides experimental setup and results, and Section 6 provides concluding remarks.

2. Overall communication model

In this section, we first describe the proposed model for the transmitter of a covert channel considering its asynchronous nature and the variation in signaling time. Then, we explain the underlying reasons to model the receiver as a pulse-shape filter. This is followed by deriving effective channel noise which is a combination of additive and jitter noise caused by the variation in signaling time. Finally, we propose discrete memoryless channel model by considering the effective channel noise, insertions, and deletions.

In the literature, side-channel analysis has been generally considered for cryptographic devices to secure the privacy of users. Many researchers have studied different implementations and their vulnerabilities to side channel attacks [47–49]. To obtain the worst-case information leakage that can be achieved by covert channels, we first need to model the transmitter and emanated signal. Here, we need to note that the covert channels considered in this paper do not seek for relevant information in a system. These channels are exploited as a bridge between a trusted insider or a Trojan and adversarial outsider. Also note that this paper does not propose a new covert channel. The goal of the paper is to analyze *already existing* covert channels and evaluate their severity.

2.1. Transmitted signal and receiver model

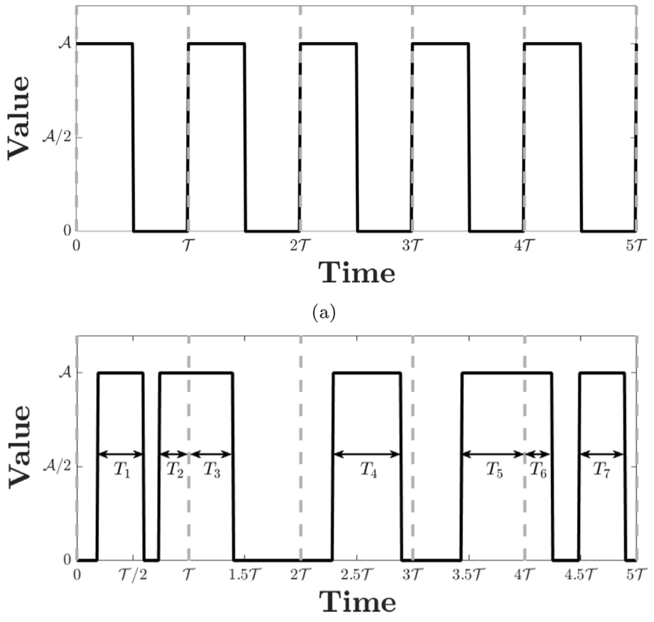
Transmitted Signal: Since transmitted signal models are well-studied in conventional communication theory, a natural approach to model the signals in covert channels is to establish a connection between covert and conventional communication systems. In that respect, we first investigate different modulation schemes (e.g., pulse amplitude-width modulation, pulse width modulation, and pulse amplitude modulation) and their suitability for covert channels. However, employing such modulation schemes with various width and amplitude choices are not practical for covert channels since the width and amplitude of the transmitted signals deviate due to other program activities [50]. To avoid these difficulties, the general practice in covert channel community is to employ modulation schemes that can only transmit zeros and ones [4,45,51]. Therefore, we follow the assumptions and notations below for the model of the transmitted signal:

A1: The Transmitted Signal Assumptions and Notations

- The receiver samples the signal at every $\mathcal{T}$ seconds under the assumption that the transmission time of the covert channel is $\mathcal{T}$.
- On–Off-Keying (OOK) is considered as the modulation scheme to transmit information signal (i.e., the source exploits a specific side-channel such as cache to transmit bits).
- The targeted duty cycle changes based on the implementation of the covert channel. In other words, the ratio between the width of the pulse and the transmission time $\mathcal{T}$ can vary for different channels.
- There is no overlapping among the bits transmitted by the covert channel transmitter. Therefore, a bit is transmitted only if the transmission of the previous bit is complete.
- The signal is on for T^0 or T^1 seconds if the transmitted bit is zero or one, respectively. Please note that T^0 can be zero, which represents the transmission period when nothing is transmitted.

In Fig. 1(a), an ideal OOK modulated signal is shown when the transmitted bit sequence is all ones (Although the behavior of any bit sequence is the same, all-one-bit-sequence is used for better explanation of the process). Unfortunately, obtaining such a signal is not possible with an unintentional channel. Because of the delays and synchronization problems in covert channels, the system experiences shifts in transmitted signals as shown in Fig. 1(b). This undesired behavior appears in almost all of the covert channels, and needs to be modeled to comprehend and estimate leakage capacity.

Receiver Model: For the receiver model, we need to combine the knowledge of conventional communication theory and some practices utilized in the security community. The common approach in conventional systems is to apply a match filter under the assumption that the system is well-synchronized [52]. Motivated by this approach, we make the following assumptions and observations:



(a)

(b)

Fig. 1. The received signal for (a) ideal conventional communication system, (b) covert channel communication system.

A2: The Receiver Model Assumptions

- The receiver employs a filter which mimics the match filter to capture the transmitted information.
- Since the modulation is a result of software–hardware activities, the transmitted keys encounter problems. These problems can be a result of changes in duty cycle, non-synchronization and delays while transmitting modulated signal.

Since synchronization problems limit the capability of the match filter for covert channels, the first step to design the receiver is to relax these synchronization requirements in order to handle shifts in signaling-time. To capture the transmitted bit sequence, the receiver employs a *matched* filter with 100% duty cycle (irrespective of the actual duty cycle of the transmission) to capture the changes in signaling time as

$$m_c(t) = \frac{1}{\sqrt{\mathcal{T}}}\texttt{rect}\left(\frac{t}{\mathcal{T}}\right), \tag{1}$$

where rect(t) is a function with amplitude one, and has only nonzero values between −0.5 and 0.5. Assuming $r(t)$ is the received signal, we can write the match filtered sequence after sampling with period $\mathcal{T}$ as

$$\begin{aligned} y(n\mathcal{T}) &= r(t) * m_c(t - n\mathcal{T})\big|_{t=n\mathcal{T}} \\ &= \frac{1}{\sqrt{\mathcal{T}}}\int_{(n-1)\mathcal{T}}^{n\mathcal{T}} r(n\mathcal{T} - \tau)\texttt{rect}\left(\frac{\tau}{\mathcal{T}} - (n - 0.5)\right)d\tau \\ &= y_n \end{aligned} \tag{2}$$

where $*$ is the convolution operation and n is the sample index.

Under this receiver implementation, the received signal can be considered as a signal which only experiences variation in duty cycles. Please observe that raising time of the received signals can be altered such that the total area under the signaling period stays the same. Therefore, an equivalent version of the original signal in Fig. 1(b) can be presented as in Fig. 2. The equivalence of these two signal models stems from the fact that the receiver does not consider the shift within the sampling period, but the existence of the signal components.

Let us consider a noiseless environment where the received signals only experience asynchronous nature of the covert channel. Fig. 3

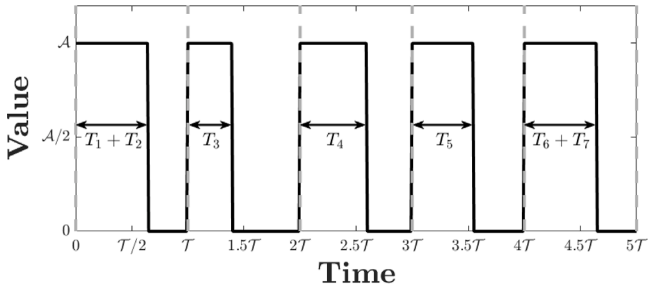


Fig. 2. The equivalent version of the received signal under the assumption that the receiver employs a modified matched filter in (1) .

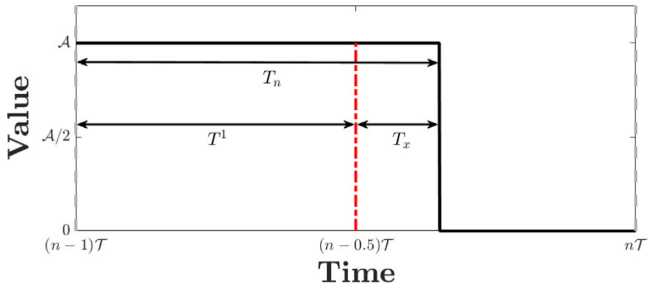


Fig. 3. One cycle corrupted received signal that was modified by signaling time variation, and modified such that the raising time is equivalent to $(n-1)\mathcal{T}$.

shows the received signal between $(n-1)\mathcal{T}$ and $n\mathcal{T}$ when, without loss of generality, the transmitted bit is one after modifying the raising time as given above (the same discussion can be made for bit-zero).

The received sample for this time slot can be written as

$$\begin{aligned} y_n &= \frac{1}{\sqrt{\mathcal{T}}}\int_{(n-1)\mathcal{T}}^{n\mathcal{T}} r(n\mathcal{T}-\tau)\texttt{rect}\left(\frac{\tau}{\mathcal{T}}-(n-0.5)\right)d\tau \\ &= \frac{\mathcal{A}}{\sqrt{\mathcal{T}}}T_n = \frac{\mathcal{A}}{\sqrt{\mathcal{T}}}\left(T^1+T_x\right) \end{aligned} \tag{3}$$

where T_x is a random variable for the width variation of the received signal. For the rest of the paper, we refer to T_x as the *effective variation*. Please note that T_x can also be negative and the outputs of the match filter can be smaller or larger than the expected output value.

2.2. Channel model

In this section, we introduce our discrete memoryless channel model for the covert communication. Having such a model is essential for establishing connection with information theory which allows us to calculate the channel capacity. For the channel model, we make the following assumptions:

A3: Bit-Deletion
The covert channel transmitter continuously sends information bits unless it encounters interrupts, stalls, etc. Due to other program activities that can run in parallel with the covert channel source, the received signal is masked and can be randomized because of the constructive and destructive interference. From the receiver perspective, the received bit has the highest entropy, hence, this scenario is considered as the deletion.

A4: Bit-Insertion
When stalls, interrupts, etc., force the source of covert channel to stop transmitting information, the insertion of random bits occurs. Since the receiver is not aware of such an interrupt, it keeps interpreting the sampled symbols as the actually transmitted symbols.

A5: Additive Gaussian White Noise (AWGN)
The transmitted signals are also corrupted by additive white Gaussian noise. We assume that this noise covers all unrelated signals that are produced by the environment and the system.

For further discussion, we remind the signal-to-noise-ratio (SNR) for the conventional communication systems which can be written as

$$\text{SNR} = \frac{P_s}{\sigma_n^2} \tag{4}$$

where P_s is the signal power and σ_n is the standard deviation of the noise after sampling. However, the conventional SNR definition does not reflect the variation in the width of the transmitted signal. Therefore, the first goal is to combine the channel noise and the noise due to signal timing variation called jitter noise. First we consider the noiseless scenario given in (3). If we define the ideally received sampled symbol, y_o, as

$$y_o = T^i\frac{\mathcal{A}}{\sqrt{\mathcal{T}}} \tag{5}$$

and the jitter noise term, n_x as

$$n_x = T_x\frac{\mathcal{A}}{\sqrt{\mathcal{T}}}, \tag{6}$$

the received sampled symbol can be written as

$$y_n = y_o + n_x \tag{7}$$

where $i \in \{0, 1\}$. Let T_x be normally distributed as

$$T_x \sim \mathcal{N}\left(\mu_x, \sigma_x^2|\text{Bit-}i \text{ is transmitted}\right).$$

This equation exposes two main intuitions about the characteristics of covert channels: (1) the mean and standard deviation of the signaling time variation could be different for different bits, and (2) the distributions can exhibit differences for each system since they can correspond to non-identical program activities. Hence, different jitter-noise schemes need to be considered. With this assumption, the distribution of n_x can be written as

$$\mathcal{N}\left(\frac{\mu_x\mathcal{A}}{\sqrt{\mathcal{T}}}, \frac{\left(\sigma_x\mathcal{A}\right)^2}{\mathcal{T}}\middle|\text{Bit-}i \text{ is transmitted}\right) = \mathcal{N}\left(\mu_{x,i}, \sigma_{x,i}^2\right).$$

Eq. (7) reveals that even with noiseless scenario assumption, the covert channel system still encounters noise due to jitter in the system. Including the additive channel noise (AWGN), the received symbol can be written as

$$y_n = y_o + n_x + n_o \tag{8}$$

where $n_o \sim \mathcal{N}\left(0, \sigma_n^2\right)$ is the channel noise sample. Here, jitter and channel noise can be combined as a single random variable because both have Gaussian distribution. Let us define n_c as the effective noise component which is given as

$$n_c = n_x + n_o.$$

Therefore, the distribution of n_c can be written as

$$n_c \sim \mathcal{N}\left(\mu_{x,i}, \sigma_n^2 + \sigma_{x,i}^2\right). \tag{9}$$

Please also note that bit-zero signals can be nonzero for a while if $T^0 > 0$. Therefore, the average transmitted signal can be written as

$$P_s = \frac{\mathcal{A}^2}{\mathcal{T}}\left(p_0\left(T^0\right)^2 + p_1\left(T^1\right)^2\right) \tag{10}$$

where p_j ($\{j \in \{0, 1\}\}$) represents the probability that bit-j is transmitted without deletion. Likewise, average effective noise power can be written as

$$P_n = \sigma_n^2 + p_0\left(\mu_{x,0}^2 + \sigma_{x,0}^2\right) + p_1\left(\mu_{x,1}^2 + \sigma_{x,1}^2\right). \tag{11}$$

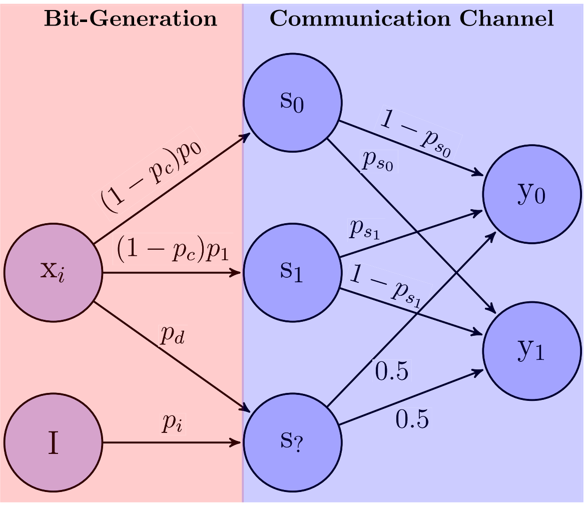


Fig. 4. Channel model for the communication system.

Therefore, the effective SNR (SNR_{eff}) in these covert channels can be defined as

$$\text{SNR}_{\text{eff}} = P_s/P_n. \tag{12}$$

These equations show that covert channels suffer not only from channel noise but also variations in signaling time. With the modeling assumptions from Section 2, we can observe that signaling time variation behaves like an extra source of channel noise. Therefore, for the simplicity of discussion, we assume additive channel noise has two components which are independent of each other: jitter and additive channel noise.

Having the model for the noise term in the system, the received samples given in (8) can be simply written as

$$y_n = y_o + n_c. \tag{13}$$

Therefore, overall channel between the transmitter and receiver can be modeled as a discrete memoryless channel since the transmitted bits show no dependency to other bits in the sequence. Another important point is that if the communication is insertion/deletion-free, the overall channel can be considered as a binary channel, but this leads to overestimation of information leakage through these systems. The channel model based on these assumptions is given in Fig. 4. To simplify explanations, we divide the model into two parts *Bit-Generation* and *Communication Channel*.

The *Bit-Generation* shows the probabilities of different types of signals that exist in the system. In this part, x_i represents the transmitter (the actual source of the covert channel). **I** represents the other activities that can cause insertions in the channel. The probabilities p_0 and $p_1(= 1 - p_0)$ are the probabilities to send bit zero and one, p_d is the deletion probability, p_i is the insertion probability, and p_c is equivalent to $p_i + p_d$. The second part, *Communication Channel*, presents the transition probabilities of different symbols. Here, s_0, s_1 and $s_?$ represent the transmitted bit zero, one, or an insertion/deletion. The received symbols corresponding to bits zero and one are denoted by y_0 and y_1, respectively.

The transmitter behaves like a ternary source, which generates zero, one or an insertion(or deletion), however, the receiver always interprets the received symbols as zero or one since it is unaware of insertion (or deletion) locations. Another observation is that the timing variation while transmitting bit-zero or bit-one could be different, which leads to differences in T_x distribution. This means the channel does not exhibit a binary-symmetric feature even if there are no insertions (or deletions). To incorporate this asymmetrical feature of the system, the substitution probabilities for bit-zero and bit-one are represented as p_{s_0} and p_{s_1}, respectively.

3. Leakage capacity

Having a model for the covert channels enables calculation of leakage capacity because leakage capacity corresponds to channel capacity of the model. In conventional communication systems, the channel capacity is calculated based on Shannon's theorem [53]. The channel capacity is defined as

$$C = \sup_{p(x)} I(X;Y) \tag{14}$$

where X and Y are the random variables for inputs and outputs, and $p(x)$ is the probability distribution for the inputs. In our scenario, X and Y represent the symbols sent from the transmitter (a *source*), and the received symbols (by an *sink*), respectively.

To calculate the leakage capacity, the first step is to obtain the transition probabilities, p_{si} where $i \in \{0,1\}$. We know that the threshold is calculated based on the posterior distribution of the inputs [52], and that the asymmetric nature of the system affects the threshold while calculating substitution probability. Combining these information, the threshold has to fulfill the following equations:

$$\begin{aligned} p_0 f\left(z_{thr} \middle| \hat{\mu}_0, \hat{\sigma}^2_{n,0}\right) &= p_1 f\left(z_{thr} \middle| \hat{\mu}_1, \hat{\sigma}^2_{n,1}\right) \\ p_0 f_0(z_{thr}) &= p_1 f_1(z_{thr}) \end{aligned} \tag{15}$$

where

$$\hat{\mu}_i = T^i \frac{\mathcal{A}}{\sqrt{T}} + \mu_{x,i} \quad \text{and} \quad \hat{\sigma}^2_{n|i} = \sigma^2_n + \sigma^2_{x,i},$$

which represent the effective symbol mean power and the effective noise variation when bit-i is transmitted, respectively, $f(x|\mu,\sigma^2)$ is the probability density function (pdf) for Gaussian distribution with mean μ and standard deviation σ, and z_{thr} is the threshold value to calculate substitution probabilities which preserve the equality in (15). Therefore, considering all these features of a covert channel, we define the leakage capacity as

$$\boxed{\begin{array}{lll} \underset{p(x)}{\text{maximize}} & I(X;Y) & \\ \text{subject to} & & \\ \quad p(x \text{ is insertion}) & = & p_i \\ \quad p_0 f_0(z_{thr}) & = & p_1 f_1(z_{thr}) \\ \quad p_{s_0} & = & P_0(Y > z_{thr}) \\ \quad p_{s_1} & = & P_1(Y \le z_{thr}) \end{array}} \tag{16}$$

where $P_i(\bullet)$ provides the probability of its event with respect to the pdf of the corresponding bit, $f_i(\bullet)$. The solution to this optimization problem provides the worst-case information leakage through covert channels.

We need to note here that the insertion (or the deletion) probability depends on the targeted computer system. Therefore, p_i and p_d have to be kept fixed while calculating the leakage capacity. A further analysis can be done to find the effect of insertion/deletion on the leakage capacity. We can observe that increase in these probabilities decreases the channel capacity. That means systems can be designed more chaotic (randomly pumping power to the systems, activating some random components, etc.) to increase insertion/deletion probability (which can be thought as a shielding strategy). Hence, investigating the effect of deletion/insertion on the channel capacity provides more knowledge and confidence about the required level of this chaotic regime.

4. Establishing connection between the proposed model and covert channels

In this section, we explain how the proposed framework can be used to model various covert channels. By establishing such a connection, we demonstrate that the model can be utilized to calculate leakage capacity of these channels, and define a metric measuring the resilience of any system to covert channel attacks.

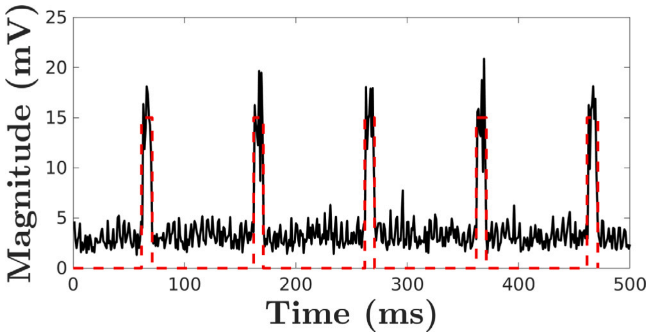


Fig. 5. Received signal generated by the covert channel in [19]. The black (solid) curve represents the measured signal and the red (dotted) curve is for the modeled signal.

4.1. Power based covert channels

The connection between covert channels based on Simple Power Analysis (SPA) and the proposed model can be established explicitly because power covert channel attacks utilize total power consumption of the system. For example, they exploit variation in power consumption while executing bits for signing operation in crypto-systems [18,54]. The main goal is to measure the total power, and to estimate whether the signed bit is zero or one, therefore, the system can be represented by OOK modulation.

To calculate the leakage capacity of power covert channel, we can assume that $\mathcal{T}$ is the average time required to sign a bit for a cryptosystem, or processing one bit of information. However, processing this information can take different amount of time due to other software activities, optimization, etc. Therefore, it can cause some shifts in time, and variation in processing time, which can be explained by the proposed model as long as the distribution of the effective variation is known. Also, due to stalls, interrupts, etc., some of the bits correspond to deletions or insertions. Both of these issues are covered by the proposed model, therefore, the power based covert channel can be analyzed theoretically.

An example of the received signal for power analysis is given in Fig. 5 when the microbenchmark in [16] is executed to transmit 0–1 sequence repetitively (Please note that these are not the signal traces while signing a key of cryptographic function). Because power channels are very noisy channels, we filter the signal with move-median filter that helps exposing the OOK structure of the received signal. Since the proposed framework is flexible to model any OOK signal with any duty cycle, it is possible to define and obtain the leakage capacity by collecting the statistics about timing variation, insertion, and deletions.

4.2. EM-based covert channels

EM covert channels are a consequence of computer activities and their effects on EM fields. By measuring the variation in the EM field, it is possible to steal information from a distance [4,45,55,56]. Requiring no direct access to the system and having larger available frequency band can be listed as the main advantages of these channels over other covert channels. In this section, we consider two channel types that exploit the variation in EM field caused by different units of a modern computer system.

4.2.1. EM-based covert channels due to processor activities

It is already shown that a covert channel can be generated by running a microbenchmark that causes systematic changes in the surrounding EM field, and a motivated attacker can monitor these changes to infer the transmitted bits [6,45]. An example of the generated signal is given in Fig. 6. The main observation is that the received signal displays OOK structure, but suffers from variation in signaling time.

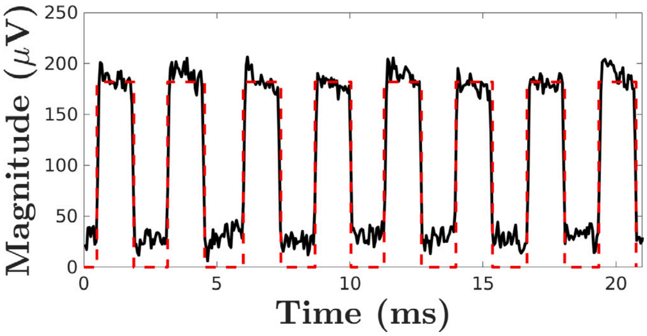


Fig. 6. Received signal generated by the covert channel in [45]. The black (solid) curve represents the measured signal and the red (dotted) curve is for the modeled signal.

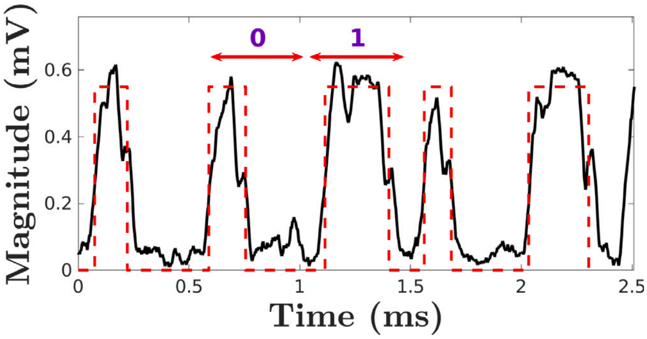


Fig. 7. Received signal generated by the covert channel in [4]. The black (solid) curve represents the measured signal and the red (dotted) curve is for the modeled signal.

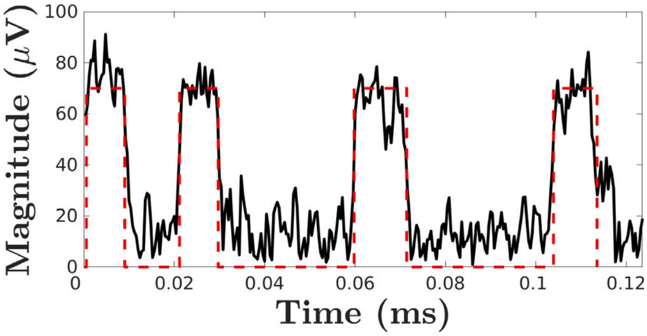
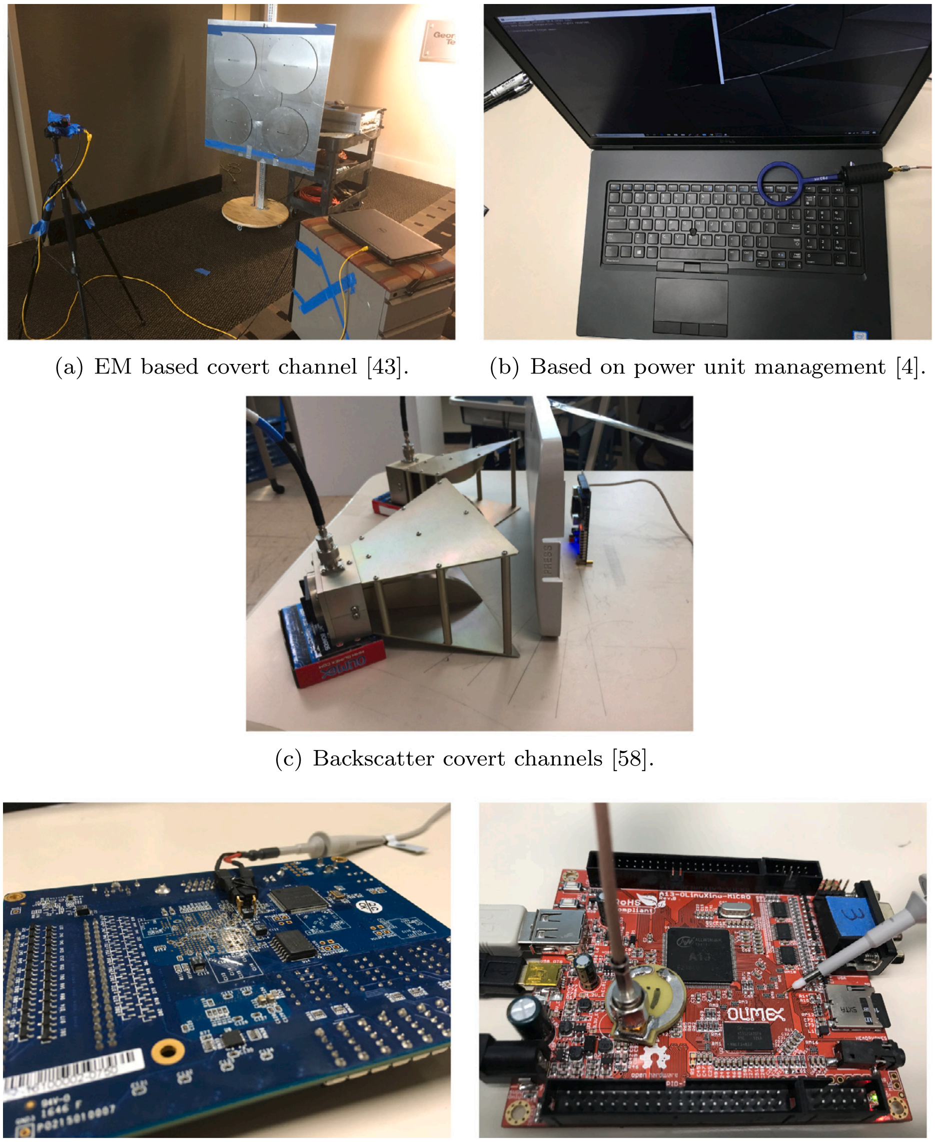


Fig. 8. Received signal generated by the covert channel in [57]. The black (solid) curve represents the measured signal and the red (dotted) curve is for the modeled signal.

Considering the same arguments with power based covert channels, we can see that the proposed model can explain the leakage in the worst-case scenario for a given design. To achieve our goal, the critical step is to obtain the variable values for deletion, insertion, $\mathcal{T}$, and the distribution of the effective variation. Once we obtain these parameters, the proposed methodology can be used to assess the resilience of a system against EM-based covert channels.

4.2.2. EM-covert channels based on power management units

These covert channels are generated by exploiting power management units (PMU) and voltage regulator module (VRM) of modern computers. PMU is responsible for power alteration to optimize the power consumption of a system. Since the priority of system designers is to minimize the consumption of power, they do not put enough effort on the security aspect of their design to covert/side-channel attacks.



(a) EM based covert channel [43].

(b) Based on power unit management [4].

(c) Backscatter covert channels [58].

(d) Power covert channels [19].

(e) Cache based covert channels [5, 59].

Fig. 9. Distributions for the signaling time for various covert channels.

By leveraging such a security flaw, a covert channel that transmits sensitive information from *air-gapped* computers is generated in [4].

To transmit information, a microbenchmark is designed causing changes in the power state of a system. An example of the demodulated signal for the covert channel is given in Fig. 7. The main observation here is that the received signals are active for a while even for the *off* case. However, this is also included in the proposed model since we do not restrict T^0 to be zero. Furthermore, this channel suffers from insertions, deletions and signaling time variation as previously considered covert channels. Although the received signal is not a perfect square signal, the proposed methodology can be exploited to calculate the leakage limit assuming distortions are due to additive channel noise.

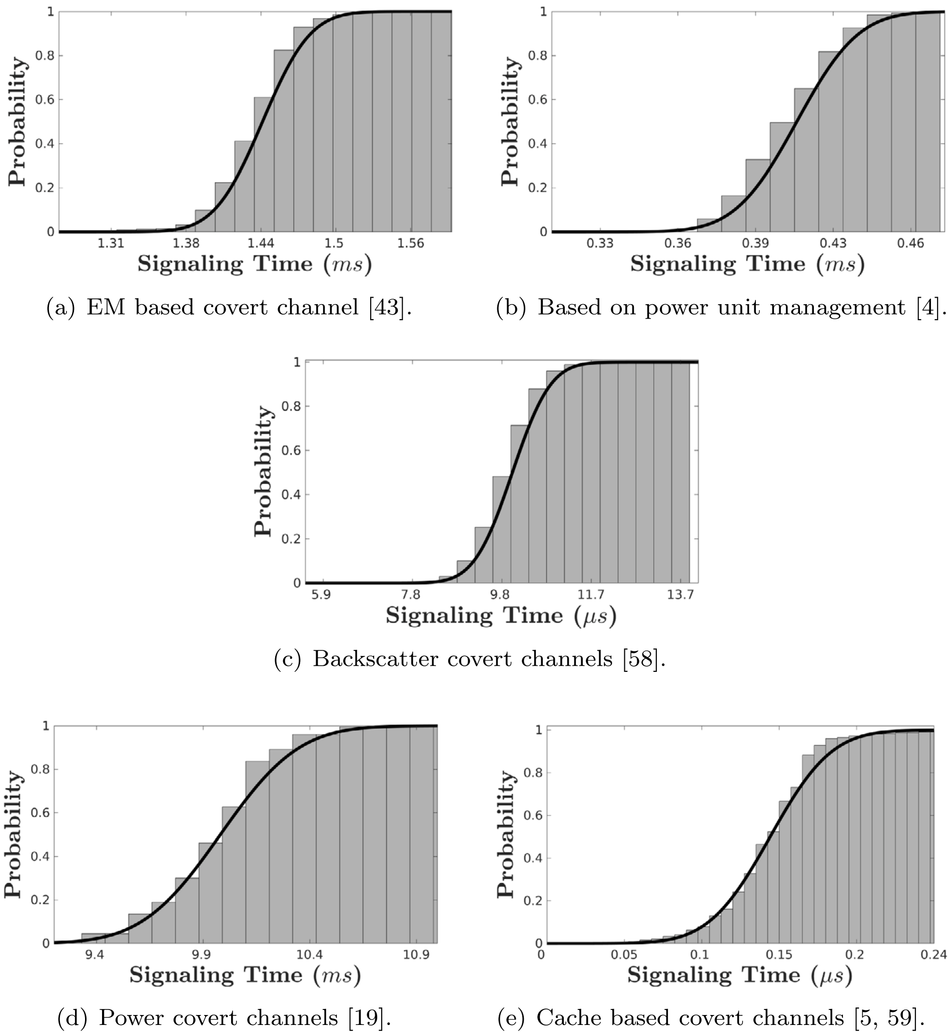
4.3. Backscattering covert channels

This covert channel is created by exploiting the recently introduced backscattering side-channel. It exploits circuits as a semi-passive RFID and relies on switching activities on the level of transistor gates (between low and high states). The switching activities change the impedance of circuits, hence, the circuit behaves as an RFID tag. For example, by utilizing this channel, circuits with Trojans are identified because the backscattered signals show different characteristics due to change in the impedance of the circuits [58]. Although there is no registered attack, or a paper investigating attack scenarios based on these channels, we still consider this channel to obtain its capacity because impedance change can result in exfiltration of some sensitive information.

An example of the received signal for the backscattering covert channel is given in Fig. 8. The same characteristic features with other covert channels, i.e., insertions, deletions, timing variation, are observed as well. Therefore, the proposed methodology can calculate the maximum leakage (or information transfer) given that the statistics about deletion, insertion, and timing distribution are known.

4.4. Cache-based covert channels

To improve the performance of a computer system by reducing the effective main memory latencies originating from data accesses, faster hardware caches are used for storing the frequently used data. Based on the speed of the caches, they are divided into levels i.e., L1, L2, etc.,



(a) EM based covert channel [43].

(b) Based on power unit management [4].

(c) Backscatter covert channels [58].

(d) Power covert channels [19].

(e) Cache based covert channels [5, 59].

Fig. 10. Distributions for the signaling time for various covert channels.

where the highest cache level i.e. L1 is the smallest, fastest and closest to the processor, and subsequent lower levels are placed closer to the main memory with varying higher latencies. Thus, more recent and/or frequent the data is, the higher it will be placed in the cache levels. The cache-based covert channel attacks exploit this difference in data access time to steal sensitive information of a victim. For example, based on the time differences in recalling cache entries, secret keys of different cryptosystems are broken [5,9]. We need to note here that the recall time can be used not only for *evil* purposes. For example, a methodology is proposed in [59] to profile the memory access that does not cause any overhead on the system. The method exploits emanated EM signals for performance analysis, and provides statistical information on the recall time of the system.

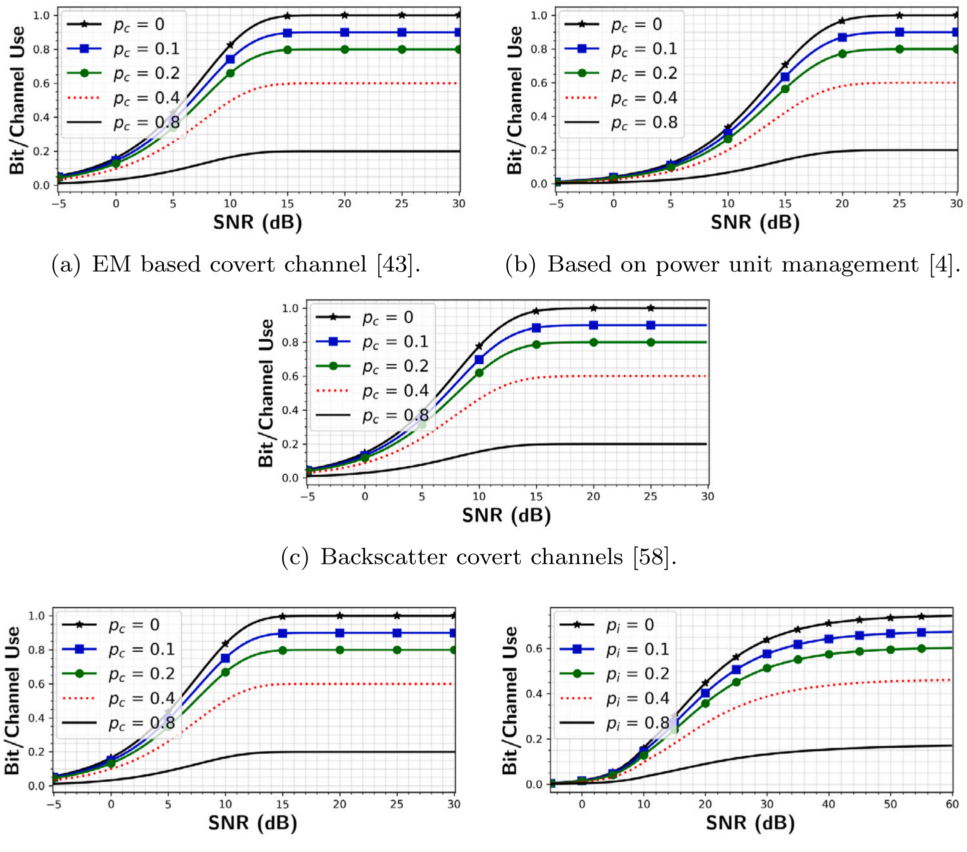
The question here is that how the proposed method can model the cache based covert channels since the only data collected during these attacks is the recall time. Let us start with the following observation: These attacks request recall time at every pre-defined time interval. In our case, this pre-defined interval is equivalent to $\mathcal{T}$. Moreover, the recall time can be considered as the output of the receiver after filtering with $m_c(t)$. Let T_R be the current recall time of an experiment or attack. If we assume the transmitted signal, $T_S(t)$, is a pulse function whose width and amplitude are equivalent to the recall time and $\sqrt{\mathcal{T}}$ with a random shift, we have

$$\begin{aligned} y(n\mathcal{T}) &= T_S(t) * m_c(t-n\mathcal{T})\big|_{t=n\mathcal{T}} \\ &= \int_{(n-1)\mathcal{T}}^{n\mathcal{T}} T_S(n\mathcal{T}-\tau)\frac{1}{\sqrt{\mathcal{T}}}\texttt{rect}\left(\frac{\tau}{\mathcal{T}}-(n-0.5)\right)d\tau \\ &= \int_{t'}^{t'+T_R} \sqrt{\mathcal{T}}\frac{1}{\sqrt{\mathcal{T}}}\texttt{rect}\left(\frac{\tau}{\mathcal{T}}-(n-0.5)\right)d\tau \\ &= \int_{t'}^{t'+T_R} \texttt{rect}\left(\frac{\tau}{\mathcal{T}}-(n-0.5)\right)d\tau \\ &= T_R \end{aligned} \tag{17}$$

where

$t' = \max((n-1)\mathcal{T}, (n-1)\mathcal{T}+t_d)$

and t_d is the time shift. Since the output is equivalent to T_R, our model can represent the cache based covert channels with the assumed transmitter and receiver. Please also observe that the recall time varies at each operation, and that is represented by the timing distribution in the model. Another observation is that based on these transmitter and receiver assumptions, the additive channel noise power is equivalent to zero, and all of the effective noise is represented by the jitter noise. Finally, cache-based covert channels generally suffer from deletions and insertions, which is also considered in the model. Therefore, our methodology can calculate the leakage capacity of these channels if the required statistics are available. Comparing to the literature, in this paper, we experimentally show that signaling time varies significantly and has a Gaussian structure. In Information Theory, the channel capacity is calculated under the assumption that the signaling time is



(a) EM based covert channel [43].

(b) Based on power unit management [4].

(c) Backscatter covert channels [58].

(d) Power covert channels [19].

(e) Cache based covert channels [5, 59].

Fig. 11. Bit/Channel use for various covert channels.

fixed and does not vary from one transmission to another. Also, we mathematically prove that the signal time variation can be modeled as an extra source for additive channel noise which causes decrease in leakage capacity. Another contribution is including the insertion–deletion channel model to consider the signal losses due to other program activities. We believe this is the first paper that combines all these features of memoryless covert channels to estimate their leakage capacity. After an exhausting research, we have not encountered any paper that combines all these parameters to propose a model for these unintended channels.

5. Experimental results and discussion

In this section, we first provide signaling time distributions for various covert channels to demonstrate that assuming Gaussian distribution with a specified mean and standard deviation is valid. Then, we provide the leakage capacity results for these channels.

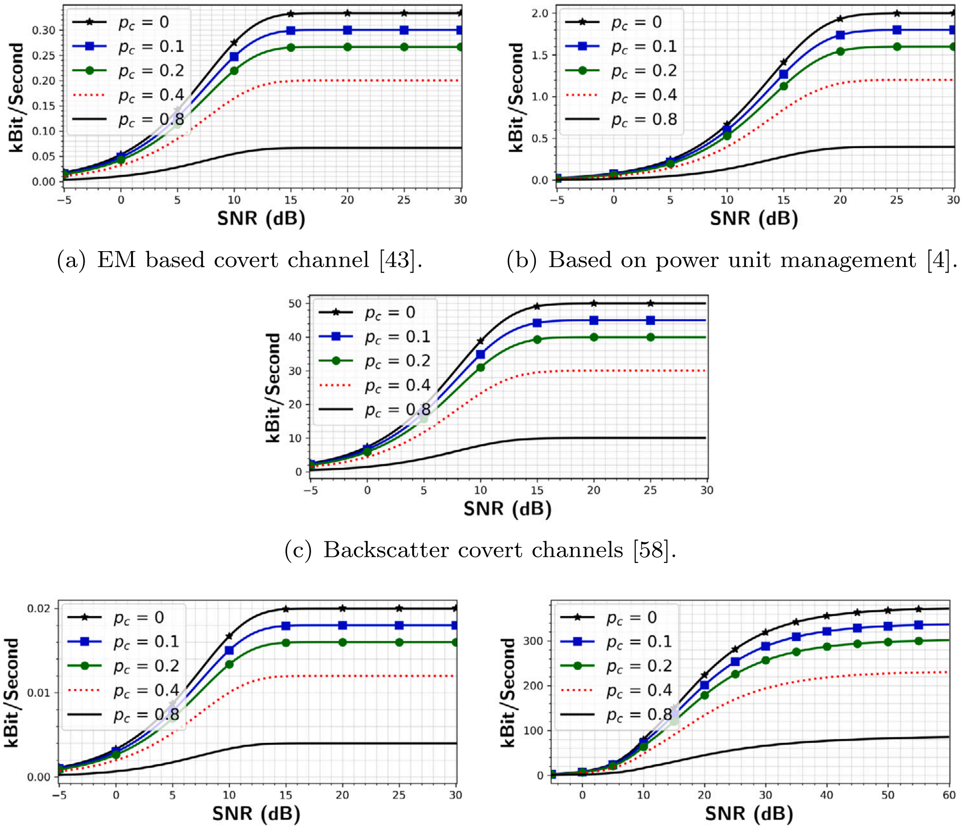
For the experiments, the devices we consider are an Altera NIOS-II processor with a commercial Terassic DE1 SoC board [60], an OlinuXino board [61] which has a modern Cortex A8 ARM core with two levels of caches, 4 MB main memory that is commonly used in factory lines, etc., and a Dell Precision 7730 laptop [62]. The antennas to collect emanated signals are a high-gain custom-made disk-array based antenna [63], near EM field probes [64], a power rail probe [65], a horn antenna [66] and a lab-made near field probe. We record the signals using a spectrum analyzer (Agilent MXA N9020 A) [67].

The first goal of these experiments is to collect data while transmitting bits to experimentally obtain the distribution of T_x for both bits. In that respect, we follow the experiments done in [6] for EM, [4] for power unit, [57] for backscattering, and [5,59] for cache-based covert channels. For the power covert channel, we collect the signal from a capacitor while running the code given in [46] and following the attack scenario given in [19]. The setup for all these measurements are given in Fig. 9. For more accurate distribution results of T_x, we collect signals closer to the device under inspection, and then perform an edge detection to obtain the width of the pulses. Empirical cumulative distribution functions (CDF) of bit-1 for various covert channels are given in Fig. 10. Furthermore, in this figure, we provide CDF of normal distributions that are fitted to these data. As seen from these figures, the signaling time distributions have a Gaussian characteristic, which means the assumption in Section 2 holds. Actually, the assumption is also supported by Law of Large numbers [68] because of ubiquitous software activities.

Table 1
Parameters utilized for the leakage capacities for covert channels.

Parameter	Power	Power unit	EM	Backscatter	Cache
$\mathcal{T}$	50	0.5	3	20	2
T^1	10	0.4	1.5	10	0.15
T^0	0	0.2	0	0	0
μ_0	0.12	0.03	0.14	1.02	≈0
μ_1	−0.02	0.02	−0.05	−0.01	−0.1
σ_0	0.05	0.01	0.03	0.66	≈0
σ_1	0.29	0.02	0.02	0.66	0.04
Unit	ms	ms	ms	μs	μs

The experimental variables required for the leakage capacity calculation for the experiments are given in Table 1. Here, σ_0 and σ_1 represent standard deviations of the signaling time (σ_x) when bit-0 and bit-1 are transmitted, respectively. Since the performance of these covert channels, in terms of bandwidth, noise characteristics, etc., are different from each other, we choose different transmission time for more reliable results. For example, in the literature, the reported transmission rates for various covert channel attacks vary from 5 bit/s to a



(a) EM based covert channel [43].

(b) Based on power unit management [4].

(c) Backscatter covert channels [58].

(d) Power covert channels [19].

(e) Cache based covert channels [5, 59].

Fig. 12. Bit per second (Bps) for various covert channels.

couple of kbits/s [4,51,69,70]. Please note that these parameters are not known beforehand and need to be measured carefully to calculate the possible leakage capacity. In that respect, an experimental setup needs to be designed which maximizes the SNR of the received signals and minimizes the interference caused by surrounding signals. These steps are required because we do not have any control on the power of the transmitted signal, and even have limited control on the signaling time of the signals for the covert channels like cache, etc. The leakage capacities, which are obtained by solving the optimization problem given in Section 3, are provided in Figs. 11 and 12. In the first figure, we provide the results in terms of Bit/Channel-Use [34] to show whether maximum gain from each bit transmission can be achieved at any *SNR*. The figures contain behavior of maximum leakage as the sum of deletion and insertion, p_c, changes. For a fair comparison of channel capacity for various deletion and insertion probabilities, *SNR* is defined as

$$SNR = \frac{\mathcal{A}^2\left(\left(\mathcal{T}^0\right)^2 + \left(\mathcal{T}^1\right)^2\right)}{2\mathcal{T}\sigma_n^2}. \tag{18}$$

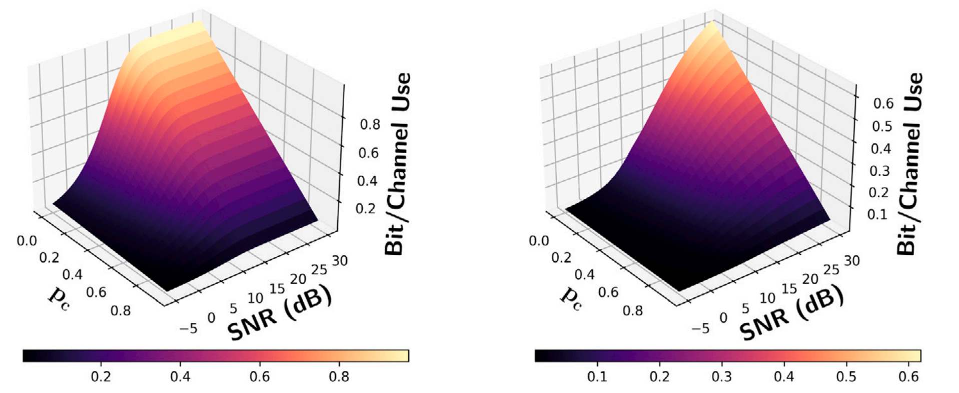
We observe that the maximum gain is only possible if there is no insertion and deletion, and the communication takes place in high *SNR*, which is an unrealistic scenario because of *unintentional* nature of covert channels. However, it does not mean that systems are secure enough to attacks exploiting these channels. If the attacker can establish a longer connection, even transmission with slow data rate could be a disaster. For example, if *SNR* is 5 dB, and $p_c = 0.8$, a communication with at least 0.1 Bit/Channel-Use could be possible for any given covert channel. Considering the attacker aims to steal some passwords, credit card information, etc., even this rate could be severe enough. In Fig. 12, the same results are given in terms of Bit-per-second (Bps). Our goal is to illustrate that hundreds of information bits can be transmitted per second through these channels even tough leakage capacity is small in terms of Bit/Channel-Use. For example, when Fig. 11(e) is compared with the rest of covert channels, we can conclude that it is the most inefficient channel for an attacker. However, Fig. 12(e) demonstrates that this channel can achieve higher data rates than others.

Another interesting observation here is that although all but cache-based covert channels achieve almost the maximum gain in terms of Bit/Channel-Use for high *SNR*, the cache-based covert channel converges to 0.6 Bit/Channel-Use. The reason is higher signaling time deviation when a cache-miss occurs. This introduces powerful jitter noise to the system, which could not be alleviated even the attacker measures the signal when *SNR* = ∞. This result shows that the signaling time variation causes a decisive additive noise which decreases the transmission rate further.

In Fig. 13, we provide the same results with Fig. 11 for backscattering and cache-based covert channels to observe the behavior in a 3-D surface plot. These figures reveal the general behavior of the worst-case leakage scenarios through different covert channels. Here, as the representative of other covert channels, we provide the results for backscattering covert channels. Our observation here is that the decrease in the leakage capacity of cache-based covert channel is sharper in both p_c and *SNR* directions than the backscattering channel due to jitter noise. However, both figures illustrate the possibility of severe information leakages through covert channels.

We compare our outcomes with the results given in [45]. In that paper, capacity bounds for the leakage capacity are provided. For a fair comparison, we assume

- The channels do not experience deletions but insertions,
- The signal deviation for both bits are same with zero mean,
- The standard deviation of signaling variation does not variate with respect to bit value.



(a) Backscatter covert channels [58].

(b) Cache based covert channels [5, 59].

Fig. 13. Bit/Channel use while p_c and *SNR* vary.

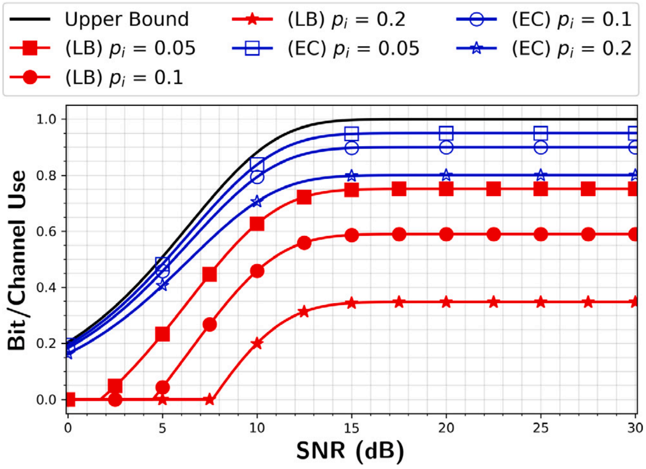


Fig. 14. The proposed leakage capacity and bounds given in [45] while *SNR* changes.

The results are given in Fig. 14. In this figure, we provide the upper and lower bounds in [45], and the proposed leakage capacity for the EM covert channel. In that paper, instead of providing an estimate for the leakage capacity, the model provides the leakage capacity bounds. The gap between these bounds increases as the insertion probability increases. We observe that our leakage capacity lies within the region defined by the bounds, and closer to the upper bound. We conclude that the decrease in the leakage capacity is much less than the decrease in the lower bound in [45] as the insertion probability increases. Please note that the proposed channel model can be used to calculate the mutual information for the leakage and success probability as given in [2,36,37,71] if deletion and insertion probabilities are zero and if signaling time shows no variation. Therefore, we believe that our model can be extended by including the assumptions related the distributions of leakages given in these papers.

Although all these results demonstrate the possible threat through these covert channels, the methodology in this paper can be used by designers to make their systems more resilient to covert channels. In the *design-stage*, designers can collect the statistics for p_c, and estimate signal power that can be generated by any covert channel attack. Then, *SNR* vs. leakage capacity analysis can be done by solving the optimization problem given in Section 3. If the leakage is zero or very close to zero at the targeted *SNR*, they can conclude their system is secure enough. Otherwise, they need to modify their design to protect privacy of their customer.

Please note that this paper considers the capacity of memoryless covert channels modulated with on–off keying. Since these channels are artifacts of legitimate program activities, estimating the parameters beforehand to calculate the channel capacity is overwhelming. Therefore, all these parameters are required to be obtained via experimentation. Also, for better assessment of system security to attacks based on these covert channels, people with expertise in covert channels need to perform the experiments. Lack of expertise can cause over/under-estimation of system security.

6. Conclusions

This paper considers the capacity of memoryless covert channels modulated with on–off keying. Since these channels are artifacts of legitimate program activities, estimating the parameters beforehand to calculate the channel capacity is overwhelming. Therefore, all these parameters are required to be obtained via experimentation. We proposed a methodology that can model multiple covert channels and estimate the available information to attackers for the worst-case scenarios through these covert channels. We showed that the method can be adopted to both analog and digital covert channels. To model the losses due to software activities, we first modeled the communication channel as a deletion–insertion channel. Then, we introduced the jitter noise that is an extra source for additive white noise. This jitter noise is a result of signaling time variation due to stalls, interrupts, optimization, etc., which conventional communication systems do not suffer. We showed that these noise sources can be combined and called effective additive noise. Secondly, based on the effective noise, we modeled the communication channel between the receiver (a *source*) and transmitter (a *sink*). Then, we defined the channel capacity as the maximum leakage for a given covert channel. Finally, we provide experimental results for various covert channels to show that the proposed model is an effective and a general method to attain the resilience of a given system.

CRediT authorship contribution statement

Baki Berkay Yilmaz: Conceptualization, Methodology/Study design, Validation, Formal analysis, Investigation, Data curation, Writing – original draft, Writing – review & editing, Visualization. **Nader Sehatbakhsh:** Methodology/Study design, Software, Validation, Investigation, Data curation, Writing – review & editing. **Moumita Dey:** Methodology/Study design, Validation, Data curation, Writing – review & editing. **Chia-Lin Cheng:** Methodology/Study design, Data curation, Writing – review & editing. **Milos Prvulovic:** Writing – review & editing, Supervision, Project administration, Funding acquisition. **Alenka Zajić:** Writing – review & editing, Supervision, Project administration, Funding acquisition.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

[1] Mangard S. Hardware countermeasures against DPA – a statistical analysis of their effectiveness. In: Cryptographers' track at the rsa conference. Springer; 2004, p. 222–35.

[2] Mangard S, Oswald E, Popp T. Power analysis attacks: revealing the secrets of smart cards, Vol. 31. Springer Science & Business Media; 2008.

[3] Lampson BW. A note on the confinement problem. Commun ACM 1973;16(10):613–5.

[4] Sehatbakhsh N, Yilmaz BB, Zajić A, Prvulovic M. A new side-channel vulnerability on modern computers by exploiting electromagnetic emanations from the power management unit. In: 2020 ieee international symposium on high performance computer architecture (hpca). IEEE; 2020, p. 123–38.

[5] Yarom Y, Falkner K. FlUSh+ RELOAD: a high resolution, low noise, L3 cache side-channel attack. In: 23rd usenix security symposium (usenix security 14). 2014, p. 719–32.

[6] Zajić A, Prvulovic M. Experimental demonstration of electromagnetic information leakage from modern processor-memory systems. IEEE Trans Electromagn Compat 2014;56(4):885–93.

[7] Wang Z, Lee RB. New cache designs for thwarting software cache-based side channel attacks. In: ACM SIGARCH Comput Archit News. 35, (2):ACM; 2007, p. 494–505.

[8] Tsunoo Y. Crypt-analysis of block ciphers implemented on computers with cache. 2002.

[9] Liu F, Yarom Y, Ge Q, Heiser G, Lee RB. Last-level cache side-channel attacks are practical. In: 2015 ieee symposium on security and privacy. IEEE; 2015, p. 605–22.

[10] Ristenpart T, Tromer E, Shacham H, Savage S. Hey, you, get off of my cloud: exploring information leakage in third-party compute clouds. In: Proceedings of the 16th acm conference on computer and communications security. ACM; 2009, p. 199–212.

[11] Yao F, Doroslovacki M, Venkataramani G. Are coherence protocol states vulnerable to information leakage? In: 2018 ieee international symposium on high performance computer architecture (hpca). IEEE; 2018, p. 168–79.

[12] Gullasch D, Bangerter E, Krenn S. Cache games–bringing access-based cache attacks on AES to practice. In: Security and privacy (sp), 2011 ieee symposium on. IEEE; 2011, p. 490–505.

[13] Agrawal D, Archambeault B, Rao JR, Rohatgi P. The EM side-channel(s). In: Cryptographic hardware and embedded systems - ches 2002, 4th international workshop, redwood shores, ca, usa, august 13-15, 2002, revised papers. 2002, p. 29–45.

[14] Kuhn MG. Compromising emanations: eavesdropping risks of computer displays (Ph.D. thesis), Citeseer; 2002.

[15] Backes M, Dürmuth M, Gerling S, Pinkal M, Sporleder C. Acoustic side-channel attacks on printers. In: 19th usenix security symposium, washington, dc, usa, august 11-13, 2010, proceedings. 2010, p. 307–22.

[16] Callan R, Zajić A, Prvulovic M. A practical methodology for measuring the side-channel signal available to the attacker for instruction-level events. In *Proceedings of the 47th international symposium on microarchitecture (micro)*, 2014.

[17] Kocher P. Timing attacks on implementations of diffie-hellman, rsa, dss, and other systems. In: Proceedings of crypto'96. Lecture notes in computer science, Springer; 1996, p. 104–13.

[18] Kocher P, Jaffe J, Jun B. Differential power analysis: leaking secrets. In: Proceedings of crypto'99. Lecture notes in computer science, Springer; 1999, p. 388–97.

[19] Boneh D, Brumley D. Remote timing attacks are practical. In *Proceedings of the usenix security symposium*, 2003.

[20] Bayrak AG, Regazzoni F, Brisk P, Standaert F-X, Ienne P. A first step towards automatic application of power analysis countermeasures. In *Proceedings of the 48th design automation conference (dac)* 2011.

[21] Chari S, Jutla CS, Rao JR, Rohatgi P. Towards sound countermeasures to counteract power-analysis attacks. In *Proceedings of crypto'99, springer, lecture notes in computer science*, 1999, pp. 398–412.

[22] Coppens B, Verbauwhede I, Bosschere KD, Sutter BD. Practical mitigations for timing-based side-channel attacks on modern x86 processors. In *Proceedings of the 30th ieee symposium on security and privacy* 2009, pp. 45–60.

[23] Goubin L, Patarin J. DES and Differential power analysis (the "duplication" method). In *Proceedings of cryptographic hardware and embedded systems - ches 1999*, 1999, pp. 158–172.

[24] Messerges TS, Dabbish EA, Sloan RH. Power analysis attacks of modular exponentiation in smart cards. In *Proceedings of cryptographic hardware and embedded systems - ches 1999*, 1999, pp. 144–157.

[25] Genkin D, Pipman I, Tromer E. Get your hands off my laptop: Physical side-channel key-extraction attacks on PCs. In: Batina L, Robshaw M, editors. Cryptographic hardware and embedded systems - ches 2014. Lecture notes in computer science, vol. 8731, Springer Berlin Heidelberg; 2014, p. 242–60.

[26] Millen JK. Covert channel capacity. In: Security and privacy, 1987 ieee symposium on. 1987, p. 60. http://dx.doi.org/10.1109/SP.1987.10013.

[27] Chari S, Rao JR, Rohatgi P. Template attacks. In: International workshop on cryptographic hardware and embedded systems. Springer; 2002, p. 13–28.

[28] Lomné V, Prouff E, Rivain M, Roche T, Thillard A. How to estimate the success rate of higher-order side-channel attacks. In: International workshop on cryptographic hardware and embedded systems. Springer; 2014, p. 35–54.

[29] Wang Z, Lee R. Capacity estimation of non-synchronous covert channels. In: Distributed computing systems workshops, 2005. 25th ieee international conference on. 2005, p. 170–6. http://dx.doi.org/10.1109/ICDCSW.2005.47.

[30] Davey MC, MacKay DJ. Reliable communication over channels with insertions, deletions, and substitutions. IEEE Trans Inform Theory 2001;47(2):687–98.

[31] Venkataramanan R, Tatikonda S, Ramchandran K. Achievable rates for channels with deletions and insertions. IEEE Trans Inform Theory 2013;59(11):6990–7013.

[32] Kirsch A, Drinea E. Directly lower bounding the information capacity for channels with iid deletions and duplications. IEEE Trans Inform Theory 2010;56(1):86–102.

[33] Hu J, Duman TM, Erden MF, Kavcic A. Achievable information rates for channels with insertions, deletions, and intersymbol interference with IID inputs. IEEE Trans Commun 2010;58(4).

[34] Rahmati M, Duman TM. Bounds on the capacity of random insertion and deletion-additive noise channels. IEEE Trans Inform Theory 2013;59(9):5534–46.

[35] Mercier H, Tarokh V, Labeau F. Bounds on the capacity of discrete memoryless channels corrupted by synchronization and substitution errors. IEEE Trans Inform Theory 2012;58(7):4306–30.

[36] Standaert F-X, Malkin TG, Yung M. A unified framework for the analysis of side-channel key recovery attacks. In: Annual international conference on the theory and applications of cryptographic techniques. Springer; 2009, p. 443–61.

[37] de Chérisey E, Guilley S, Rioul O, Piantanida P. Best information is most successful. IACR Trans Cryptogr Hardw Embedd Syst 2019;49–79.

[38] Standaert F-X, Veyrat-Charvillon N, Oswald E, Gierlichs B, Medwed M, Kasper M, Mangard S. The world is not enough: Another look on second-order DPA. In: International conference on the theory and application of cryptology and information security. Springer; 2010, p. 112–29.

[39] Veyrat-Charvillon N, Medwed M, Kerckhof S, Standaert F-X. Shuffling against side-channel attacks: A comprehensive study with cautionary note. In: International conference on the theory and application of cryptology and information security. Springer; 2012, p. 740–57.

[40] Duc A, Faust S, Standaert F-X. Making masking security proofs concrete. In: Annual international conference on the theory and applications of cryptographic techniques. Springer; 2015, p. 401–29.

[41] Prouff E, Rivain M. Masking against side-channel attacks: A formal security proof. In: Annual international conference on the theory and applications of cryptographic techniques. Springer; 2013, p. 142–59.

[42] Yilmaz BB, Callan R, Zajić A, Prvulovic M. Capacity of the EM covert/side-channel created by the execution of instructions in a processor. IEEE Trans Inf Forensics Secur 2018;13(3):605–20.

[43] Yilmaz BB, Prvulovic M, Zajić A. Electromagnetic side channel information leakage created by execution of series of instructions in a computer processor. IEEE Trans Inf Forensics Secur 2020;15:776–89. http://dx.doi.org/10.1109/TIFS.2019.2929018.

[44] Yilmaz BB, Zajić A, Prvulovic M. Modelling Jitter in wireless channel created by processor-memory activity. In: IEEE international conference on acoustics, speech and signal processing, icassp 2018. 2018, p. 2037–41. http://dx.doi.org/10.1109/ICASSP.2018.8461902.

[45] Yilmaz BB, Sehatbakhsh N, Zajić A, Prvulovic M. Communication model and capacity limits of covert channels created by software activities. IEEE Trans Inf Forensics Secur 2018;13(3):605–20.

[46] Callan R, Popovic N, Zajić A, Prvulovic M. A new approach for measuring electromagnetic side-channel energy available to the attacker in modern processor-memory systems. In: 2015 9th european conference on antennas and propagation (eucap). IEEE; 2015, p. 1–5.

[47] Prest T, Goudarzi D, Martinelli A, Passelègue A. Unifying leakage models on a Rényi day. In: Annual international cryptology conference. Springer; 2019, p. 683–712.

[48] Bronchain O, Hendrickx JM, Massart C, Olshevsky A, Standaert F-X. Leakage certification revisited: Bounding model errors in side-channel security evaluations. In: Annual international cryptology conference. Springer; 2019, p. 713–37.

[49] Heuser A, Rioul O, Guilley S. Good is not good enough. In: International workshop on cryptographic hardware and embedded systems. Springer; 2014, p. 55–74.

[50] Callan R, Behrang F, Zajić A, Prvulovic M, Orso A. Zero-overhead profiling via EM emanations. In: Proceedings of the 25th international symposium on software testing and analysis. ACM; 2016, p. 401–12.

[51] Guri M, Monitz M, Elovici Y. USbee: Air-gap covert-channel via electromagnetic emission from USb. 2016, CoRR abs/1608.08397 arXiv:1608.08397 URL http://arxiv.org/abs/1608.08397.

[52] Proakis J. Digital communications. Mcgraw-hill series in electrical and computer engineering. computer engineering, McGraw-Hill; 2001, URL https://books.google.com/books?id=sbr8QwAACAAJ.

[53] Shannon CE. A mathematical theory of communication. Bell Syst Tech J 1948;27(3):379–423.

[54] Mangard S. A simple power-analysis (SPA) attack on implementations of the AES key expansion. In: International conference on information security and cryptology. Springer; 2002, p. 343–58.

[55] Alam M, Khan HA, Dey M, Sinha N, Callan R, Zajić A, Prvulovic M. One&done: A single-decryption EM-based attack on OpenSSL's constant-time blinded $\{RSA\}$. In: 27th usenix security symposium (usenix security 18). 2018, p. 585–602.

[56] Genkin D, Pachmanov L, Pipman I, Tromer E. Stealing keys from PCs using a radio: Cheap electromagnetic attacks on windowed exponentiation. In: Cryptographic hardware and embedded systems - ches 2015 - 17th international workshop, saint-malo, france, september 13-16, 2015, proceedings. 2015, p. 207–28.

[57] Cheng C-L, Nguyen LN, Prvulovic M, Zajić A. Exploiting switching of transistors in digital electronics for RFID tag design. IEEE J Radio Freq Identif 2019;3(2):67–76.

[58] Nguyen LN, Cheng C-L, Prvulovic M, Zajić A. Creating a backscattering side channel to enable detection of dormant hardware Trojans. IEEE Trans Very Large Scale Integr (VLSI) Syst 2019.

[59] Dey M, Nazari A, Zajić A, Prvulovic M. Emprof: Memory profiling via EM-emanation in IoT and hand-held devices. In: 2018 51st annual ieee/acm international symposium on microarchitecture (micro). IEEE; 2018, p. 881–93.

[60] DE1 FPGA on NIOS Processor, https://www.terasic.com.tw/cgi-bin/page/archive.pl?Language=English&CategoryNo=53&No=83&PartNo=2.

[61] OlinuXino, https://www.olimex.com/Products/OLinuXino/A13/A13-OLinuXino/open-source-hardware.

[62] Precision D. https://www.dell.com/en-us/work/shop/workstations-isv-certified/sc/workstations/precision-laptops.

[63] Juyal P, Adibelli S, Sehatbakhsh N, Zajić A. A directive antenna based on conducting disks for detecting unintentional EM emissions at large distances. IEEE Trans Antennas and Propagation 2018;66(12):6751–61.

[64] AARONIA PBS, https://www.tequipment.net/Aaronia/PBS1-5/Standard/Passive-Oscilloscope-Probes/?rrec=true.

[65] Power Rail Probe, https://www.keysight.com/en/pd-2471132-pn-N7020A/power-rail-probe?&cc=US&lc=eng.

[66] AH-118, Double Ridge Horn Antenna, https://www.com-power.com/ah118_horn_antenna.html.

[67] Keysight Signal Analyzer, https://www.keysight.com/en/pdx-x202266-pn-N9020A/mxa-signal-analyzer-10-hz-to-265-ghz?pm=spc&nid=-32508.1150426&cc=US&lc=eng.

[68] Mendenhall W, Beaver RJ, Beaver BM. Introduction to probability and statistics. Cengage Learning; 2012.

[69] Backes M, Dürmuth M, Gerling S, Pinkal M, Sporleder C. Acoustic side-channel attacks on printers. In: 19th usenix security symposium, washington, dc, usa, august 11-13, 2010, proceedings. 2010, p. 307–22.

[70] Guri M, Kachlon A, Hasson O, Kedma G, Mirsky Y, Elovici Y. GSMem: DAta exfiltration from air-gapped computers over GSM frequencies. In: 24th usenix security symposium (usenix security 15). Washington, D.C.: USENIX Association; 2015, p. 849–64, URL https://www.usenix.org/conference/usenixsecurity15/technical-sessions/presentation/guri.

[71] Köpf B, Basin D. An information-theoretic model for adaptive side-channel attacks. In *Proceedings of the 14th acm conference on computer and communications security*, 2007, pp. 286–296.