DEFENDING AGAINST PACKET-SIZE SIDE-CHANNEL ATTACKS IN IOT NETWORKS

Sijie Xiong, Anand D. Sarwate, Narayan B. Mandayam

Rutgers, The State University of New Jersey

ABSTRACT

Motivated by privacy issues in the Internet of Things (IoT), we generalize a previously proposed privacy-preserving packet obfuscation scheme to guarantee differential privacy. We propose a locally differentially private packet obfuscation mechanism as a defense against packet-size side-channel attacks in IoT networks. We formulate the problem as an optimization over a conditional probability distribution (channel) between the original and obfuscated packet sizes and show that the optimal set of obfuscated packet sizes is a strict subset of the set of original packet sizes. We study the optimal mechanisms for minimizing the (average or min-max) bandwidth overhead subject to a privacy constraint by solving the corresponding (linear or convex) program. We demonstrate our methods on synthetic and real data to illustrate privacy-bandwidth tradeoffs in different settings. Systems with many bandwidth-intensive devices can easily mask low-bandwidth devices. For data collected from actual smart home IoT devices, we show how the packet size distributions become increasingly indistinguishable as the level of privacy protection increases. The proposed mechanism highlights the possibility for bandwidth-constrained users to optimally tune their privacy preferences and trade off privacy with bandwidth.

Index Terms— Internet of Things, side-channel attacks, packet obfuscation, local differential privacy, linear/convex program.

1. INTRODUCTION

Systems in the Internet of Things (IoT) use ubiquitous devices to continuously sense and monitor users in a diverse set of applications, such as smart homes, smart health care, intelligent transportation systems, and environmental/industrial monitoring. Supporting these applications greatly benefits and improves the users' quality of life. The data collected from these devices will often be sent over wireless links: the volume of data imposes huge challenges for privacy protection [1] and bandwidth conservation [2].

Fine-grained monitoring data creates a security loophole that allows adversarial inference on private information. More recently, traffic analysis attacks [3, 4, 5, 6] (also known as *side-channel* attacks) have demonstrated that statistical network traffic patterns from IoT devices are highly correlated with the underlying sensing data. A passive wireless eavesdropper is able to extract sensitive information even if the data content is well protected under cryptographic techniques. For example, Apthorpe et al. [3] argue that a last-mile network observer or a WiFi eavesdropper can observe features of network traffic generated by smart home IoT devices. They use a clustering method to identify device types and operating states even if the devices use TLS/SSL for communication. Das et al. [4] show that an encrypted packet stream from wearable (e.g., a fitness tracker) Bluetooth Low Energy (BLE) signal allows a BLE sniffer to identify user activities (e.g., whether the person is at rest/working/walking/running) based on the packet-size and inter-arrival time distributions, without payload decryption. Buttyan and Holczer [5] apply simple signal processing techniques (Discrete Fourier Transform and Welch Averaged Periodogram) on traffic rates from wireless body area sensor network to reveal the types of medical sensors mounted on the patient, and hence the patient's health conditions. Finally, supervised machine learning can distinguish between network traffic generated by non-IoT and IoT devices, and further between specific types of IoT devices [6].

Early works on traffic privacy focused on building anonymous networks. An adversary can observe the correspondences between the size and timing of messages going into and out of an intermediate network node (e.g., a router), and make inferences about source-destination pairs in the network. Chaum [7] proposed using a cascade of anonymity *mixes* to hide the correspondences by outputting at each *mix* uniformly sized messages with padded bits. Similarly, a prototypical realization of the IoT system [8] consists of a gateway system that manages networked IoT devices. An adversary can observe the incoming and outgoing packets of the gateway and figure out which device is in operation and communicating with its application server. Even with deployment of the anonymity mixes, as the side-channel attacks imply, the adversary can still look at the packet sizes of an IoT device locally to extract contextual information and cause privacy breach. Wright et al. [9] and Apthorpe et al. [10] investigated packet obfuscation schemes under perfect privacy. However, with a limited bandwidth budget such as in low-power and low-memory IoT networks, perfect privacy is often not achievable and users would prefer tunable privacy. Mathur and Trappe [11] studied the fundamental tradeoff between mutual-information privacy and bandwidth but did not provide any empirical evaluation.

Our goal is to generalize a packet obfuscation mechanism against packet-size side-channel attacks to guarantee *differential privacy* [12], a stronger privacy guarantee than mutual-information privacy [13, 14]. Particularly, we adopt the *local differential privacy* model [15] where individual IoT devices must obfuscate their packets prior to transmission in the presence of a local eavesdropper. We study interesting privacy-bandwidth tradeoffs by formulating the minimization of the *on-average* or *worst-case* bandwidth overhead under given privacy levels as constrained linear or convex programs. We empirically evaluate the optimal solutions on synthetic and real data, and show how various user preferences of bandwidth-intensive and low-bandwidth devices affect the privacy-bandwidth tradeoffs differently. We believe that this work is a first step towards building a comprehensive on-device traffic obfuscation system with strong, formal and tunable privacy guarantees. The tradeoffs provide meaningful system indications for designing future IoT networks.

2. SYSTEM MODEL

We model an IoT device's outgoing packet stream as a discrete-time sequence of variable-sized packets $X_1, X_2, \ldots, X_n$ where $X_i \in$

Work supported by NSF awards SaTC-1617849 and CCF-1453432.

$\mathcal{X} = \{a_1, a_2, \ldots, a_{|\mathcal{X}|}\}$; $\mathcal{X}$ is the set of all possible packet sizes (e.g., in bytes) and we assume that $a_1 < a_2 < \cdots < a_{|\mathcal{X}|}$ without loss of generality (w.l.o.g.). We think of the packets as independent and identically distributed (i.i.d.) with distribution $p_\nu(x)$ on $\mathcal{X}$. Specifically, we assume that p_ν comes from a family of possible probability mass functions (PMFs) $\mathcal{P} = \{p_\nu\}_{\nu=1}^{s}$, and we denote the prior probability that the IoT device generates packets from p_ν as $P_{p_\nu}(X) \triangleq P(X \sim p_\nu)$. This family $\mathcal{P}$ can represent IoT device types so that each ν corresponds to a different type, or $\mathcal{P}$ can represent the operating states of a single IoT device so that ν is the current state. We assume $\mathcal{P}$ and $P_{p_\nu}(X)$ are known by both the system designer and the adversary. When the device sends a packet with size $x \in \mathcal{X}$ over a wireless link, the eavesdropping adversary can observe x and infer something about p_ν. Therefore, the privacy-aware source must obfuscate x to conceal the true p_ν.

Packet Padding Obfuscator. Built on the idea of our previous work [16], we design a packet obfuscator acting as a discrete memoryless channel (DMC). This DMC defines a conditional probability distribution $q(\hat{x}|x)$ and randomly maps the original packet size x to an obfuscated size $\hat{x} \in \hat{\mathcal{X}} = \{d_1, d_2, \ldots, d_{|\hat{\mathcal{X}}|}\}$; $\hat{\mathcal{X}}$ is the set of all possible obfuscated packet sizes and assuming $d_1 < d_2 < \cdots < d_{|\hat{\mathcal{X}}|}$ (w.l.o.g.). To ensure no loss in data, we only allow the DMC to pad the packets by restricting the structure of the channel matrix such that $q(\hat{x}_j|x_i) = 0, \forall x_i > \hat{x}_j, i \in [|\mathcal{X}|], j \in [|\hat{\mathcal{X}}|]$. Particularly, if $\hat{\mathcal{X}} = \mathcal{X}$, the $|\mathcal{X}|^2$-dimensional channel matrix q will only have non-zero entries in the upper triangle. In a special case where q has all ones in the last column and zeros elsewhere, it essentially pads all the packets to the largest possible packet size $d_{|\hat{\mathcal{X}}|}$. This guarantees perfect privacy but meanwhile requires maximum extra bandwidth. Our goal is to tune/optimize the transitional probabilities to trade off the privacy level and bandwidth overhead.

2.1. Privacy Model

We require a quantifiable measure of privacy risk to meaningfully protect privacy and trade off privacy and bandwidth. *Differential privacy* [12] has emerged over the last decade as a compelling framework for measuring privacy risk in various applications. We use a more stringent privacy model – *local differential privacy* (LDP) [17, 18] to protect individual IoT devices. They can obfuscate the packet sizes prior to transmission by passing the packets through the aforementioned channel $q(\hat{x}|x)$ while satisfying LDP.

Definition 1. *A channel $q : \mathcal{X} \to \hat{\mathcal{X}}$ satisfies ϵ-local differential privacy* [17] *if* $\max\{q(\hat{x}|x)/q(\hat{x}|\tilde{x})\} \leq e^\epsilon, \forall (x, \tilde{x}, \hat{x}) \in \mathcal{X}^2 \times \hat{\mathcal{X}}$.

LDP ensures that the distribution of the output $\hat{x}$ reveals limited information about the input x: for any other input $\tilde{x}$, the output under $\tilde{x}$ has a similar distribution to that under x. Smaller ϵ means greater indistinguishability and hence less privacy risk. In our context, the goal of designing an ϵ-LDP packet obfuscator q is to ensure that the adversary's likelihood of guessing that the device's packet size distribution was p_ν over $p_{\nu'}$ does not increase, multiplicatively, more than e^ϵ after seeing the obfuscated packet size $\hat{x}$. Formally,

$$\frac{P(X \sim p_\nu | \hat{X} = \hat{x})}{P(X \sim p_{\nu'} | \hat{X} = \hat{x})} \leq \frac{P_{p_\nu}(X)}{P_{p_{\nu'}}(X)} \cdot e^\epsilon, \begin{pmatrix} \forall \nu, \nu' \in [s] \\ \forall \hat{x} \in \hat{\mathcal{X}} \end{pmatrix}, \tag{1}$$

using Bayes' rule and plugging in $q(\hat{x}_j|x_i)$, we have

$$\frac{\sum_{i=1}^{|\mathcal{X}|} p_\nu(x_i) \cdot q(\hat{x}_j|x_i)}{\sum_{i=1}^{|\mathcal{X}|} p_{\nu'}(x_i) \cdot q(\hat{x}_j|x_i)} \leq e^\epsilon, \begin{pmatrix} \forall \nu, \nu' \in [s] \\ \forall j \in [|\hat{\mathcal{X}}|] \end{pmatrix}. \tag{2}$$

We observe that this privacy model is the same as the *distributional privacy* [19] in the case of discrete outcomes.

2.2. Bandwidth Overhead

We model the bandwidth overhead $\hat{W}$ after applying the channel q in two ways: i) *on-average*, the expected number of bytes per packet (over all priors $P_{p_\nu}(X)$ and types $p_\nu, \forall \nu \in [s]$) needed to send the obfuscated packets $\hat{X}$, and ii) *worst-case*, the maximum expected number of bytes per packet among all types $p_\nu, \forall \nu \in [s]$. Formally,

$$\begin{aligned} \text{i) } \hat{W}_{\text{avg}}(q) &= \mathbb{E}_{P_{p_\nu}(X),\ X \sim p_\nu,\ q(\hat{X}|X)}[\hat{X}] \\ &= \sum\nolimits_\nu \sum\nolimits_i \sum\nolimits_j P_{p_\nu}(X) p_\nu(x_i) q(\hat{x}_j|x_i) \hat{x}_j, \quad (3) \\ \text{ii) } \hat{W}_{\text{worst}}(q) &= \max_\nu \mathbb{E}_{X \sim p_\nu,\ q(\hat{X}|X)}[\hat{X}] \\ &= \max_\nu \sum\nolimits_i \sum\nolimits_j p_\nu(x_i) q(\hat{x}_j|x_i) \hat{x}_j \quad (4) \end{aligned}$$

3. OPTIMAL CHANNEL

Finding the optimal channel $q(\hat{x}|x)$ given a fixed $\hat{\mathcal{X}}$ can now be solved by the following optimization problem:

$$\begin{aligned} \min_{q(\hat{x}|x)} \quad & \hat{W}_{\text{avg}}(q) \text{ in (3) or } \hat{W}_{\text{worst}}(q) \text{ in (4)} && (5) \\ \text{s.t.} \quad & 0 \leq q(\hat{x}_j|x_i) \leq 1, \quad \forall i, j && (6) \\ & \sum\nolimits_{j=1}^{|\hat{\mathcal{X}}|} q(\hat{x}_j|x_i) = 1, \quad \forall i && (7) \\ & q(\hat{x}_j|x_i) = 0, \quad \forall x_i > \hat{x}_j && (8) \\ & q(\hat{x}_j|x_i) \text{ satisfies (2)} && (9) \end{aligned}$$

where the goal is to minimize the *on-average* or *worst-case* bandwidth overhead subject to the privacy constraint (2). Constraints (6) and (7) ensure that the mapping $q(\hat{x}|x)$ is a stochastic matrix, and constraint (8) enforces the packet padding strategy.

Proposition 1. *The optimization problem in* (5)-(9) *with objective $\hat{W}_{\text{avg}}(q)$ or $\hat{W}_{\text{worst}}(q)$ is a linear or convex program, respectively.*

Proof. The objective function $\hat{W}_{\text{avg}}(q)$ (3) is linear in $q(\hat{x}|x)$, and $\hat{W}_{\text{worst}}(q)$ (4) is convex in $q(\hat{x}|x)$ since point-wise maximum preserves convexity [20]. The privacy constraint (2) is equivalent to

$$\sum\nolimits_{i=1}^{|\mathcal{X}|} (p_\nu(x_i) - e^\epsilon p_{\nu'}(x_i)) \cdot q(\hat{x}_j|x_i) \leq 0, \tag{10}$$

$\forall \nu, \nu' \in [s],\ \forall j \in [|\hat{\mathcal{X}}|]$, which imposes $s^2 \times |\hat{\mathcal{X}}|$ linear inequality constraints on $q(\hat{x}|x)$. Along with other linear constraints (6), (7) and (8), the optimization (5)-(9) with objectives $\hat{W}_{\text{avg}}(q)$ and $\hat{W}_{\text{worst}}(q)$ constitute a linear and convex program, respectively. □

The optimal channel given a fixed $\hat{\mathcal{X}}$ can now be solved efficiently by linear or convex programming. However, the size of $\hat{\mathcal{X}}$ can be arbitrarily large. We now show in the following proposition that an optimal channel will actually have $\hat{\mathcal{X}} \subset \mathcal{X}$.

Proposition 2. *The optimal set of obfuscated packet sizes is a strict subset of the set of original packet sizes, that $\hat{\mathcal{X}} \subset \mathcal{X}$.*

Proof. We first show that the optimal channel $q^{|\mathcal{X}| \times |\hat{\mathcal{X}}|}$ will have $\hat{\mathcal{X}} \subseteq \mathcal{X}$. Suppose contrarily that $\hat{\mathcal{X}} = \mathcal{X} \cup \{d_k\}$, where $d_k \notin \mathcal{X}$ has non-null probability in $\hat{\mathcal{X}}$. With slight abuse of notation, we also

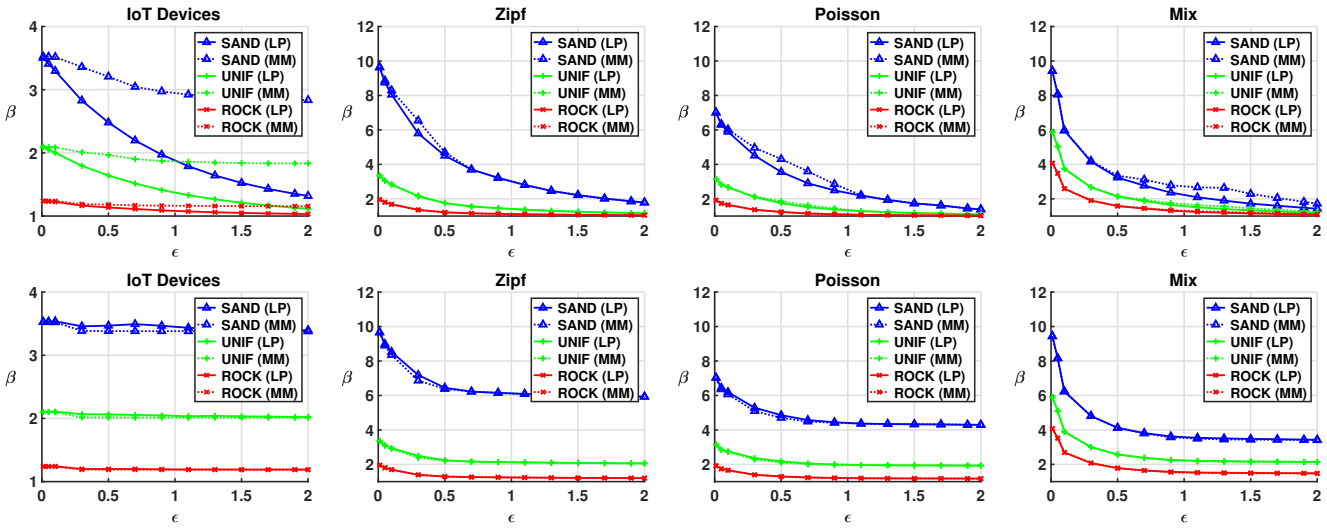


Fig. 1: On-Average and Worst-Case Privacy-Bandwidth Tradeoffs. Each of the 4 subfigures on the top row compares the *on-average* tradeoffs ϵ-$\hat{W}_{\text{avg}}(q^*_{\text{LP}})$ (solid line) and ϵ-$\hat{W}_{\text{avg}}(q^*_{\text{MM}})$ (dotted line), under SAND, UNIF, and ROCK priors for the corresponding family of PMFs. Each subfigure on the bottom row compares *worst-case* tradeoffs ϵ-$\hat{W}_{\text{worst}}(q^*_{\text{LP}})$ and ϵ-$\hat{W}_{\text{worst}}(q^*_{\text{MM}})$, accordingly.

assume $a_i = d_i < d_k < a_{i+1} = d_{i+1}, i \in \{1, \dots, |\mathcal{X}| - 1\}$, and let $k - 1 \triangleq \arg\max_j (d_j < d_k)$ and $q_{i,j} \triangleq q(d_j|a_i)$.

We prove $\hat{\mathcal{X}} \subseteq \mathcal{X}$ by showing that removing d_k from $\hat{\mathcal{X}}$, in the act of merging the extra k-th column $q_{:,k}$ with the $(k-1)$-th column $q_{:,k-1}$, can further decrease the objectives in (5) without violating the privacy constraint (2). This merging in effect creates a new channel $\bar{q}^{|\mathcal{X}|^2}$ where

$$\begin{cases} \bar{q}_{i,k-1} = q_{i,k-1} + q_{i,k}, \\ \bar{q}_{i,k} = 0, \end{cases} \qquad \forall i \in [|\mathcal{X}|]. \tag{11}$$

The amount of bandwidth conservation going from q to $\bar{q}$ for any packet size PMF p_ν is

$$\Delta W(q, \bar{q}) \triangleq \mathbb{E}_{p_\nu, q}[\hat{X}] - \mathbb{E}_{p_\nu, \bar{q}}[\hat{X}] \tag{12}$$

$$= \sum\nolimits_i p_\nu(a_i) \left(\sum\nolimits_{j=1}^{|\hat{\mathcal{X}}|} q_{i,j} d_j - \sum\nolimits_{j=1}^{|\mathcal{X}|} \bar{q}_{i,j} d_j \right) \tag{13}$$

$$= \sum\nolimits_i p_\nu(a_i) \left[(q_{i,k-1} - \bar{q}_{i,k-1}) d_{k-1} + (q_{i,k} - \bar{q}_{i,k}) d_k \right] \tag{14}$$

$$= \sum\nolimits_i p_\nu(a_i) q_{i,k} (d_k - d_{k-1}) > 0. \tag{15}$$

We have (14) from (13) since the merging only affects the $(k-1)$-th and k-th columns. (15) results from (11) and that $d_{k-1} < d_k$. Because $\hat{W}_{\text{avg}}(q) - \hat{W}_{\text{avg}}(\bar{q}) = \mathbb{E}_{p_{p_\nu}(X)}[\Delta W(q, \bar{q})] > 0$ and $\hat{W}_{\text{worst}}(q) - \hat{W}_{\text{worst}}(\bar{q}) = \max_\nu [\Delta W(q, \bar{q})] > 0$, we conclude that removing d_k from $\hat{\mathcal{X}}$ can reduce the bandwidth in either case.

The privacy constraint (2), or equivalently (10), can be viewed as s^2 constraints on each column j of the optimal channel q. Since $\bar{q}$ differs from q only in the $(k-1)$-th column $\bar{q}_{:,k-1}$, showing that $\bar{q}_{:,k-1}$ satisfies (10) is sufficient to prove that $\bar{q}$ is still ϵ-LDP. This is obvious because $\sum_i c \cdot \bar{q}_{i,k-1} = \sum_i c \cdot (q_{i,k-1} + q_{i,k}) \leq 0$, where $c \triangleq p_\nu(a_i) - e^\epsilon p_{\nu'}(a_i)$.

Now we show $a_1 \notin \hat{\mathcal{X}}$ by contradiction. Suppose that $a_1 \in \hat{\mathcal{X}}$, with padding-only mechanism, we have $q_{1,1} = 1$ and $q_{i,1} = 0, i = \{2, \dots, |\mathcal{X}|\}$. However, $\sum_i c \cdot q_{i,1} = p_\nu(a_1) - e^\epsilon p_{\nu'}(a_1) \leq 0$ can not hold for all $\nu, \nu' \in [s]$. This violates the privacy constraint (10), and contradicts with q being optimal. Combined with $\hat{\mathcal{X}} \subseteq \mathcal{X}$, we have $\hat{\mathcal{X}} \subseteq \mathcal{X} \setminus \{a_1\}$, and therefore, $\hat{\mathcal{X}} \subset \mathcal{X}$. □

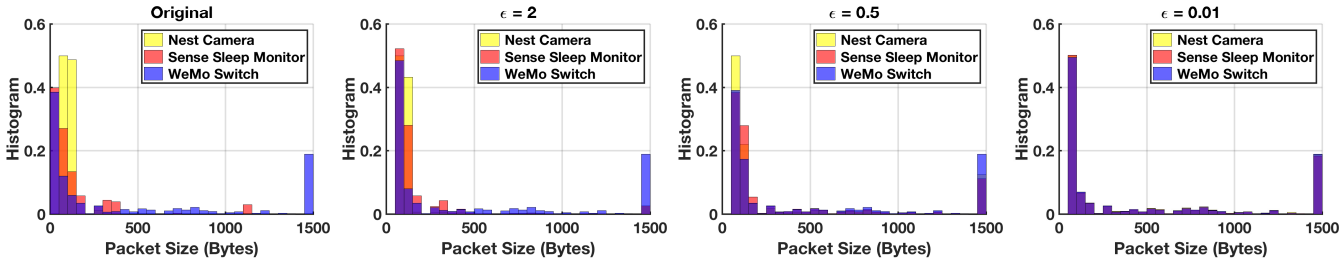
4. EXPERIMENTAL RESULTS

We experimented on synthetic data as well as packet size PMFs measured from 3 smart home IoT devices (Nest Camera, Sense Sleep Monitor and WeMo Switch) [21]. For IoT devices, we measured a total of 78 possible packet sizes ($40 - 1500$ bytes). For synthetic data, we simulated Zipf and Poisson distributed packet sizes with PMFs: $P_{\text{Zipf}}(k; \mu, N) = (1/k^\mu)/\sum_{n=1}^N (1/n)^\mu$ and $P_{\text{Poisson}}(k) = (\lambda^k/k!)e^{-\lambda}$, where the Zipf PMF characterizes the frequency of rank-k element out of a population of N elements. We assume that packet size a_k has rank k, with $k \in [N] = [|\mathcal{X}|]$. We choose the exponent $\mu \in \boldsymbol{\mu} = [5, 1, 0.01]$ for Zipf and $\lambda \in \boldsymbol{\lambda} = [0.5, 3.5, 5.5]$ for Poisson, and set the possible packet sizes to be $\mathcal{X} = \hat{\mathcal{X}} = [2, 4, 8, 16, 32, 64, 128, 256]$. We also include the mixture of Zipf and Poisson PMFs to represent the case in which there are increased number of sources.

Privacy-Bandwidth Tradeoff. Now we have $s = 4$ families of PMFs (IoT Devices, Zipf, Poisson and Mix). We want to study how different types of prior assumptions affect the privacy-bandwidth tradeoffs. We define 3 types of priors as follows:

- SAND: *sand* prior, with $P_{p_\nu}(X)$ highest on the least-bandwidth source and low on bandwidth-intensive sources (e.g., we assume $[0.8, 0.1, 0.1]$ on $\boldsymbol{\mu} = [5, 1, 0.01]$ for Zipfs);
- UNIF: *uniform* prior (e.g., we assume $[\frac{1}{3}, \frac{1}{3}, \frac{1}{3}]$ on $\boldsymbol{\mu}$);
- ROCK: *rock* prior, with $P_{p_\nu}(X)$ highest on the source consuming the most bandwidth and low on low-bandwidth sources (e.g., we assume $[0.1, 0.1, 0.8]$ on $\boldsymbol{\mu}$).

For different families of packet size PMFs and different priors, we solve (5)-(9) by using *linear programming* (LP) and *min-max* (MM) programming with respect to the $\hat{W}_{\text{avg}}(q)$ and $\hat{W}_{\text{worst}}(q)$ objectives. Denote q^*_{LP} and q^*_{MM} as the optimal solutions accordingly.



(a) Original (b) $\epsilon = 2$ (c) $\epsilon = 1$ (d) $\epsilon = 0.01$

Fig. 2: Packet Size Histograms Before/After Obfuscation. We used the $\hat{W}_{\text{avg}}(q)$-minimal channel q^*_{LM} with privacy levels $\epsilon = [2, 0.5, 0.01]$.

On the top row of Figure 1, we compare the tradeoffs, ϵ-$\hat{W}_{\text{avg}}(q^*_{\text{LP}})$ and ϵ-$\hat{W}_{\text{avg}}(q^*_{\text{MM}})$, between the privacy level ϵ and the *on-average* bandwidth. On the bottom row of Figure 1, we also compare the tradeoffs, ϵ-$\hat{W}_{\text{worst}}(q^*_{\text{LP}})$ and ϵ-$\hat{W}_{\text{worst}}(q^*_{\text{MM}})$, between the privacy level ϵ and the *worst-case* bandwidth. For ease of interpretation, we set the y axis in the figures to be $\beta = \hat{W}/W$, representing *multiples* of the average source bandwidth ($W = \mathbb{E}_{P_{p_\nu}(X),\ X\sim p_\nu}[X]$).

All curves suggest that more bandwidth provides better privacy protection. In the low privacy region ($\epsilon \to 2$), we need at least the average source bandwidth ($\beta \to 1$) using packet padding strategy. To guarantee perfect privacy ($\epsilon = 0$) for IoT devices under UNIF prior, we see that about twice the average source bandwidth is needed ($\beta \approx 2$). For any particular family of packet size PMFs, we see that SAND $>$ UNIF $>$ ROCK in terms of extra bandwidth requirement. Put in another way, if a user tends to use bandwidth-intensive devices more often, then that amount of bandwidth is sufficient to also conceal other low-bandwidth devices. Contrarily, a lot more extra bandwidth is required to hide a bandwidth-intensive device used only once in a while. This also explains the increasing gap between $\hat{W}_{\text{avg}}(q^*_{\text{LP}})$ and $\hat{W}_{\text{avg}}(q^*_{\text{MM}})$ as the prior changes from ROCK to UNIF and then to SAND.

Since Zipf PMFs have "heavy tails" on large packet sizes, they consume higher bandwidth than Poisson PMFs. However, by augmenting the Zipf family with Poisson PMFs, the resulting Mix family consumes less extra bandwidth than Zipf alone under the SAND prior. Essentially, it becomes "easier" to hide a bandwidth-intensive device among increased number of low-bandwidth devices. However, this does not hold under ROCK and UNIF priors because we are merely adding in more devices with high bandwidth demands.

Applying Packet Obfuscation Channel. We pass the 3 IoT devices' packet streams through the optimal obfuscation channel q^*_{LP} with increasing privacy levels ($\epsilon = [2, 0.5, 0.01]$) and show the corresponding obfuscated packet size histograms in Figure 2b, 2c and 2d. Comparing to the original (Figure 2a), we see that as ϵ decreases to 0.01, the histograms of obfuscated packet sizes become increasingly identical. In light of this, one can alternatively design a 2-step packet obfuscation scheme for perfect privacy: i) choose a target packet size PMF, and ii) find the mapping from the source PMFs to the target PMF. Wright et al. [9] deals with ii) by optimally morphing one class of traffic to look like a target class. However, they didn't discuss the issue of how to choose the optimal target PMF. Our solution avoids solving i) explicitly.

Note in Figure 2 that the output space after obfuscation doesn't contain the smallest packet size anymore (the first bin of histogram disappeared). This validates $a_1 \notin \hat{\mathcal{X}}$ in Proposition 2. Additionally, the frequencies of moderate-to-large packet sizes (> 500 Bytes) don't change by much before and after obfuscation, whereas the number of small packet sizes (< 200 Bytes) decreases the most as ϵ decreases. The obfuscator seems to "only" pad the small packets to larger sizes. We also observed this from the structure of the optimal channel matrices that they have ones in the diagonal entries corresponding to moderate-to-large packet sizes, and the other non-zero entries only appear in the top few rows corresponding to small packet sizes. Exploring the optimal structural properties of these channel matrices can potentially reduce the complexity of solving (5)-(9).

Sequential Composition Attack. The privacy model in (2) assumes that the adversary only observes a single packet size to make inference about its source distribution (p_ν). This assumption is adequate for *event-driven* IoT devices, where the device only sends out a packet when it senses some event. It may become weak, however, if the device stays in the same state for longer time, during which multiple packets are sent out (which can be viewed as i.i.d. samples from the same p_ν). A dedicated adversary can then observe N i.i.d. samples to better infer about the device type/state. Using our current privacy model, the overall privacy leakage becomes $\epsilon \cdot N$ by the sequential composition [22] of differential privacy, which grows linearly with the sample size. To address this, one can extend the privacy model (2) by changing the source symbols in $\mathcal{X}$ to high-dimensional tuples in $\mathcal{X}^m$. However, the computational complexity of the problem now grows exponentially with m. We believe that by exploiting the aforementioned structural properties of the optimal channel, the exponential complexity can be greatly reduced. We defer a detailed investigation to the full version of this work.

5. CONCLUSIONS AND FUTURE WORK

In this work, we designed a packet-size obfuscation mechanism under LDP and empirically showed its effectiveness for hiding the packet size PMFs from different smart home IoT devices. The mechanism is also generally applicable in other IoT scenarios. This is a first step towards understanding and defending against more sophisticated traffic analysis attacks with strong privacy guarantee. Under various assumptions on the priors and source PMFs, we showed interesting fundamental tradeoffs between the privacy level and bandwidth requirement. These tradeoffs can advise users with different bandwidth limitations to optimally choose their privacy parameter.

One future extension is to also consider timing side-channel attacks. Privatizing the network traffic in this regard should also require extra delay. It would be interesting to see the impact of both bandwidth and delay, when used complimentarily, on the rate of privacy leakage.

Acknowledgements. Special thanks to Noah Apthorpe, Dillon Reisman and Nick Feamster for sharing the network traffic data of smart home IoT devices.

6. REFERENCES

[1] J. H. Ziegeldorf, O. G. Morchon, and K. Wehrle, "Privacy in the Internet of Things: Threats and Challenges," *Security and Communication Networks*, vol. 7, no. 12, pp. 2728–2742, 2014.

[2] A. Al-Fuqaha, M. Guizani, M. Mohammadi, M. Aledhari, and M. Ayyash, "Internet of Things: A Survey on Enabling Technologies, Protocols, and Applications," *IEEE Communications Surveys & Tutorials*, vol. 17, no. 4, pp. 2347–2376, 2015.

[3] N. Apthorpe, D. Reisman, and N. Feamster, "Closing the Blinds: Four Strategies for Protecting Smart Home Privacy from Network Observers," *arXiv preprint arXiv:1705.06809*, 2017.

[4] A. K. Das, P. H. Pathak, C.-N. Chuah, and P. Mohapatra, "Uncovering Privacy Leakage in BLE Network Traffic of Wearable Fitness Trackers," in *Proceedings of the 17th International Workshop on Mobile Computing Systems and Applications*. ACM, 2016, pp. 99–104.

[5] L. Buttyan and T. Holczer, "Traffic Analysis Attacks and Countermeasures in Wireless Body Area Sensor Networks," in *World of Wireless, Mobile and Multimedia Networks (WoWMoM), 2012 IEEE International Symposium on a*. IEEE, 2012, pp. 1–6.

[6] Y. Meidan, M. Bohadana, A. Shabtai, J. D. Guarnizo, M. Ochoa, N. O. Tippenhauer, and Y. Elovici, "ProfilIoT: A Machine Learning Approach for IoT Device Identification Based on Network Traffic Analysis," in *Proceedings of the Symposium on Applied Computing*. ACM, 2017, pp. 506–509.

[7] D. L. Chaum, "Untraceable Electronic Mail, Return Addresses, and Digital Pseudonyms," *Communications of the ACM*, vol. 24, no. 2, pp. 84–90, 1981.

[8] L. Atzori, A. Iera, and G. Morabito, "The Internet of Things: A Survey," *Computer networks*, vol. 54, no. 15, pp. 2787–2805, 2010.

[9] C. V. Wright, S. E. Coull, and F. Monrose, "Traffic Morphing: An Efficient Defense Against Statistical Traffic Analysis," in *NDSS*, 2009, vol. 9.

[10] N. Apthorpe, D. Reisman, S. Sundaresan, A. Narayanan, and N. Feamster, "Spying on the Smart Home: Privacy Attacks and Defenses on Encrypted IoT Traffic," *arXiv preprint arXiv:1708.05044*, 2017.

[11] S. Mathur and W. Trappe, "BIT-TRAPS: Building Information-Theoretic Traffic Privacy into Packet Streams," *IEEE Transactions on Information Forensics and Security*, vol. 6, no. 3, pp. 752–762, 2011.

[12] C. Dwork and A. Roth, "The Algorithmic Foundations of Differential Privacy," *Foundations and Trends® in Theoretical Computer Science*, vol. 9, no. 3–4, pp. 211–407, 2014.

[13] P. Cuff and L. Yu, "Differential Privacy as a Mutual Information Constraint," in *Proceedings of the 2016 ACM SIGSAC Conference on Computer and Communications Security*. ACM, 2016, pp. 43–54.

[14] A. D. Sarwate and L. Sankar, "A Rate-Disortion Perspective on Local Differential Privacy," in *Communication, Control, and Computing (Allerton), 2014 52nd Annual Allerton Conference on*. IEEE, 2014, pp. 903–908.

[15] J. C. Duchi, M. I. Jordan, and M. J. Wainwright, "Local Privacy and Statistical Minimax Rates," in *Foundations of Computer Science (FOCS), 2013 IEEE 54th Annual Symposium on*. IEEE, 2013, pp. 429–438.

[16] S. Xiong, A. D. Sarwate, and N. B. Mandayam, "Randomized Requantization with Local Differential Privacy," in *Acoustics, Speech and Signal Processing (ICASSP), 2016 IEEE International Conference on*. IEEE, 2016, pp. 2189–2193.

[17] S. L. Warner, "Randomized Response: A Survey Technique for Eliminating Evasive Answer Bias," *Journal of the American Statistical Association*, vol. 60, no. 309, pp. 63–69, 1965.

[18] K. Kalantari, L. Sankar, and A. D. Sarwate, "Optimal Differential Privacy Mechanisms Under Hamming Distortion for Structured Source Classes," in *Information Theory (ISIT), 2016 IEEE International Symposium on*. IEEE, 2016, pp. 2069–2073.

[19] S. Zhou, K. Ligett, and L. Wasserman, "Differential Privacy with Compression," in *Information Theory, 2009. ISIT 2009. IEEE International Symposium on*. IEEE, 2009, pp. 2718–2722.

[20] S. Boyd and L. Vandenberghe, *Convex Optimization*, Cambridge university press, 2004.

[21] N. Apthorpe, D. Reisman, and N. Feamster, "A Smart Home is No Castle: Privacy Vulnerabilities of Encrypted IoT Traffic," *arXiv preprint arXiv:1705.06805*, 2017.

[22] F. D. McSherry, "Privacy Integrated Queries: An Extensible Platform for Privacy-Preserving Data Analysis," in *Proceedings of the 2009 ACM SIGMOD International Conference on Management of data*. ACM, 2009, pp. 19–30.