Towards Accuracy Aware Minimally Invasive Monitoring (MiM)

Sameera Ghayyur
University of California,Irvine
sghayyur@uci.edu

Dhrubajyoti Ghosh
University of California,Irvine
dhrubajg@uci.edu

Xi He
University of Waterloo
xi.he@uwaterloo.ca

Sharad Mehrotra
University of California, Irvine
sharad@ics.uci.edu

ABSTRACT

This paper proposes a privacy preserving framework for monitoring applications that embodies the principle of minimal data disclosure. Utility in monitoring context is measured as number of events (e.g., anomalous behavior, violation of building fire code, etc.) that went undetected. Depending upon the context, often such applications to be useful require bounded guarantees on the number of false negatives. Differential privacy, wherein one adds noise based on a predetermined ϵ, may lead to unacceptable level of accuracy (if ϵ is small) or significant loss of privacy (if ϵ is high), and, furthermore, offers no guarantee on accuracy of detection. We propose a minimally invasive framework framework that offers a guaranteed bound on event that may go undetected (false negatives), while minimizing the privacy loss.

1 INTRODUCTION

Much of the prior work on privacy has been motivated by the need for data sharing while ensuring privacy of sensitive data. Examples include privacy-preserving sharing of demographic data (e.g., US Census), medical data to support research (e.g., cancer registries), or collecting click-stream data for vulnerability analysis (e.g., from browsers). Over the past decade, differential privacy [1] has emerged as one of the most popular privacy notion. It provides a formal mathematical guarantee that individual records are hidden even with the release of aggregate statistics and it is possible to bind the information leakage by a total privacy budget across multiple data releases. This has led to a wide range of adoption of differential privacy in a number of products at the US Census Bureau [2], Google [7], and Uber [5].

With advances in sensing, computation, and communication capabilities, a new class of IoT applications are emerging that provide the ability to monitor, in real-time, physical spaces such as nursing home, office buildings, homes, etc. through diverse sensors [8]. As part of our ongoing effort, we have built one such sensor-enabled monitoring testbed, entitled TIPPERS [8] that exploits WiFi connectivity data to compute occupancy levels at different spatial granularities (building, floors, regions, and rooms) at UCI. TIPPERS has been used to build a diverse set of applications such as building analytics to understand the space utilization and real-time monitoring to understand dynamic occupancy density. A sample occupancy heat-map of a floor inside the building is shown in Figure 1. Such data is used to detect anomalies of occupancy inside the building (e.g., significant violation of the fire code, abnormally high/low occupancy levels). A monitoring application requires that every event of interest remains detectable with high probability - i.e, the false negative rate is bounded.

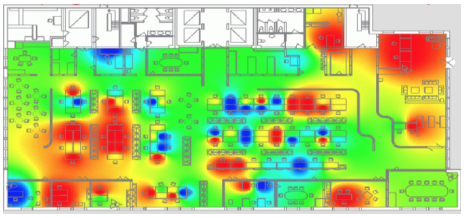

Figure 1: Occupancy Heatmap of a Building in UCI.

While continuously captured sensor data provides novel opportunities to build new functionalities, such data increases the risk of sensitive inferences about individuals [4] significantly. Our prior work in the context of TIPPERS has shown that occupancy data, with enough background knowledge, can lead to inferences about location of individuals, which, in turn, can leak sensitive information (e.g., faculty arriving late to classes, staff consistently leaving work early, smoking habits of individuals, etc.). Our goal in this work is to explore a privacy preserving framework for monitoring applications that ensures privacy but does not interfere with the purpose for which monitoring application is built in the first place.

We note that the traditional approach of using differential privacy, as in the case of data sharing applications, does not suffice in our context. Differential privacy, used directly, will introduce noise into the query answers and the noisy occupancy levels can lead to undetected events of interest. Such a strategy would not offer any bound on the level of false negatives. What is needed is a monitoring system which (i) meets the accuracy requirement of the applications and (ii) minimizes the privacy loss of individuals. We denote such a system by *accuracy aware minimally invasive monitoring* (MiM). We explore MiM in the context of *iceberg counting query* that, in the monitoring context, identify rooms whose occupancy are over a prespecified threshold.

Data exploration with accuracy guarantees for iceberg queries has also been recently proposed in APEx [3]. APEx, however, cannot be directly applied to MiM for several reasons. First, APEx considers a simpler, data-independent, but less intuitive accuracy requirement than MiM. Also, no existing differentially private algorithms are designed for iceberg counting query with a bounded false negative rate in mind. Furthermore, in APEx false negatives as well as false positives are both part of the application utility. In contrast, in MiM while false negatives characterize utility, false positives result in privacy loss. The reason is that the monitoring setting, depending upon the application in question, an anomalous event (e.g., violation of building policy) can warrant a deeper exploration of the space enabling the administrative staff, to investigate the space in question

more invasively (e.g., via a physical inspection). A false positive, thus, may result in not just the true occupancy data to be leaked, it may also leak the actual identities of the individuals in the location. Hence, limiting the privacy loss not only requires controlling the ϵ (as in differential privacy) with which occupancy data is released, but also minimizing the number of false positives.

In this work, we propose the first solution to achieve a realistic accuracy and privacy requirements in Minimally Invasive Monitoring applications and discuss interesting novel research directions. In summary the contribution of our paper is as follows:

- We introduce the concept of Minimally invasive Monitoring that has a wide applicability in building privacy preserving IoT applications.
- We introduce a new accuracy requirement for monitoring queries which is dependent on false negative rate.
- We propose a set of differentially private mechanisms that achieves bounded false negative rate with the same privacy budget and then present a model to choose one with the least false positives among these mechanisms.
- Experiments performed on real and synthetic datasets show the efficacy of our proposed approach.

The organization of this paper is as follows. In Section 2 we present our query model of monitoring queries and their accuracy requirement and privacy definition. We summarize our algorithms in Section 3. Section 4 shows our preliminary evaluation results using two synthetic datasets and one real dataset. Last, we discuss future directions in Section 5.

2 PROBLEM SETUP

We consider *iceberg counting query* which takes into an attribute of domain size L and a threshold c. Let $x_i(D)$ be the frequency of the $i-th$ value of this attribute in the dataset D. This query returns the attribute values with frequencies no less than equal to c. We refer to this query by

$$Q_{>c}^{\{1,\ldots,L\}}(D) = \{i \in [1, L] \mid x_i(D) > c\}.$$

Definition 2.1 (Accuracy Requirement (β-False Negative Rate)). The accuracy requirement of a mechanism $M : D \rightarrow O$ that answers an iceberg counting query $Q_{>c}^{1,\cdots,L}$ is defined as:

$$\forall i \in A, Pr(i \notin M(D) | i \in A) \leq \beta \tag{1}$$

where, A is the set of ids for the attribute values with frequency higher than c.

We use differential privacy as the measure of privacy in our setup. An algorithm is said to follow differential privacy, given an input dataset D, if the output of the algorithm does not change significantly, when a single tuple is added or removed from the input dataset. Formally it is defined as follows:

Definition 2.2 (Differential Privacy). A randomized mechanism M : $\mathcal{D} \rightarrow \mathcal{O}$ satisfies ϵ-differential privacy, if

$$Pr[M(D) \in O] \leq e^{\epsilon} Pr[M(D') \in O] \tag{2}$$

for any set of outputs O $\in \mathcal{O}$, and any pair of neighboring databases D, D' where D and D' differs by only one tuple.

Algorithm 1 Baseline Algorithm

procedure ThresholdShiftLM($Q_{>c}^{\{1,\ldots,L\}}, D, \beta, \epsilon$)
  $\alpha \leftarrow \ln(1/(2\beta))/\epsilon$
  return $O \leftarrow \{i \in [1, L] \mid x_i(D) + \eta_i > c - \alpha, \eta_i \sim Lap(1/\epsilon)\}$
end procedure

Definition 2.3 (False Positives). Let M be a randomized mechanism that answers an iceberg counting query $Q_{>c}^{\{1,\ldots,L\}}$ on dataset D, given an output O of M, the number of false positives is

$$n_{fp}(O) = |\{i \in O | i \notin Q_{>c}^{\{1,\ldots,L\}}(D)\}| \tag{3}$$

Definition 2.4 (Privacy Metric). We use (i) the total privacy budget (ϵ) used in answering an iceberg counting query $Q_{>c}^{\{1,\ldots,L\}}$ and (ii) the number of false positives n_{fp} in the query answer as the privacy metric used in the differentially private mechanism.

Problem Statement. Given $Q_{>c}^{\{1,\ldots,L\}}$ and a dataset D, we would like to first (i) identify a set of ϵ-differentially private mechanisms that achieve β-false negative rate, denoted by $\mathcal{M}_{\epsilon,\beta}$, and then (ii) to find $M \in \mathcal{M}_{\epsilon,\beta}$ that has the smallest number of expected false positives, i.e., $minarg_{M \in \mathcal{M}_{\epsilon,\beta}} \mathbb{E}_{O \sim M(D)} n_{fp}(O)$.

3 OUR APPROACH

In this section, we will first present a set of ϵ-differentially private mechanisms that achieve β-false negative rate (Section 3.1) and then a cost model for estimating the expected number of false positives in the mechanism output to choose the optimal mechanism among the given set of mechanisms (Section 3.2).

3.1 Algorithm Design

Given $Q_{>c}^{\{1,\ldots,L\}}$, the objective is to bind β false negative rate of an ϵ-differentially private mechanism. Here is our first algorithm that achieves this goal.

Baseline Algorithm. As shown in Algorithm 1, we first relax the threshold in the iceberg counting query $Q_{>c}^{\{1,\ldots,L\}}$ to $Q_{>c-\alpha}^{\{1,\ldots,L\}}$, where $\alpha = \ln(1/\beta)/\epsilon$. Then we add η_i noise drawn from Laplace distribution to each count $x_i(D)$ with mean 0 and variance $2/\epsilon^2$ and return the set of ids for noisy counts higher than or equal to $c - \alpha$. This approach binds the probability of false negatives by β.

Theorem 3.1. *Algorithm 1 satisfies ϵ-differential privacy and achieves β-false negative rate.*

Proof. It is easy to see the algorithm satisfies ϵ-differential privacy. Then, we show that false negative rate is bounded,

$$\begin{aligned} &Pr[\{i \notin O | i \in Q_{>c}^{\{1,\ldots,L\}}(D)\}] \\ &= Pr[x_i(D) + \eta_i \leq c - \alpha | x_i(D) > c] \\ &\leq Pr[\eta_i > \alpha] \leq \frac{e^{-\ln(\frac{1}{2\beta})}}{2} \leq \beta \end{aligned}$$

□

We generalize the baseline algorithm by shifting the thresholds multiple times, and denote it by *progressive algorithm.*

Algorithm 2 Progressive Algorithm

1: **procedure** ProgressiveLM($Q_{>c}^{\{1,\ldots,L\}}$, D, β,$\{\epsilon_1 < \ldots < \epsilon_m = \epsilon\}$, $\{\beta_1, \cdots, \beta_m\}$, $\sum_{j=1}^{m} \beta_j = \beta$)
2: $[\eta_1, \ldots, \eta_L] \leftarrow Lap(1/\epsilon_1)^L$
3: $\alpha_1 \leftarrow \ln(1/(2\beta_1))/\epsilon_1$
4: $O \leftarrow \{i \in [1, L] \mid q_i(D) + \eta_i > c - \alpha_1\}$
5: **for** $j = 2, \ldots, m$ **do**
6: $\alpha_j \leftarrow \ln(1/(2\beta_j))/\epsilon_j$
7: $[\eta_1, \ldots, \eta_L] \leftarrow PrivacyRelax(\epsilon_{j-1}, \epsilon_j, [\eta_1, \ldots, \eta_L])$
8: $O \leftarrow \{i \in O \mid q_i(D) + \eta_i > c - \alpha_j\}$
9: **end for**
10: **return** O
11: **end procedure**

Progressive Algorithm. The key idea of this algorithm is to progressively relax the thresholds from a larger shift to a smaller shift, i.e., $\alpha_1 > \alpha_2 > \cdots > \alpha_m$, where each α_j depends on the privacy cost and false negative rate assigned to each iteration. We eliminate certain number of ids from the output O in every iteration, as summarized in Algorithm 2. This algorithm takes a sequence of $\{\epsilon_j\}$ with increasing values and a sequence of $\{\beta_j\}$ which has a sum of β. First, we relax c to $c - \alpha_1$, where $\alpha_1 = \ln(1/(2\beta_1))/\epsilon_1$ and finds the ids with noisy counts greater than this threshold. Then this process repeats. In i-th iteration (Line 7), instead of sampling independent noise, we draw correlated noise according to the algorithm proposed by the authors in [6] to correlate the noise of j-th iteration with the noise of $(j-1)$-th iteration. This *PrivacyRelax* function takes the inputs of ϵ_{j-1}, ϵ_j and the set of noise drawn in the previous iteration and outputs the new set of noise for iteration j. Noise drawn in this way ensures that the overall privacy loss of this algorithm is ϵ. This approach has the same property as the baseline algorithm.

Theorem 3.2. *Algorithm 2 satisfies ϵ-differential privacy and achieves β-false negative rate.*

Proof. (sketch) The privacy guarantee is based on the results of prior work [6] and the β-false negative rate is by union bound, as the sum of β_j is bounded by β. □

There are many possible settings for Algorithm 2, including the number of iterations m and the sequences for ϵ_j and β_j.

3.2 Algorithm Selection

Different algorithms result in different number of false negatives, even on the same dataset. We would like to choose the algorithms described in the previous section with the smallest expected number of false positives. We observe that the number of false positives in the output $n_{fp}(O)$ of a given algorithm depends on the data distribution. For example, when the number of iteration m is 1 in Algorithm 2, which is the baseline algorithm, the resulted number of false positives have different empirical averages among datasets of different distributions . Similarly, our progressive algorithm at $m > 1$ also results in different number of false positives when data distribution varies. For this paper, we reduce the scope of the problem by limiting the set of ϵ-differentially private mechanisms that achieve β-false negative rate to the Baseline Algorithm and the Progressive Algorithm with $m = 2$.

We also assume some knowledge about the data distribution at a low granularity is publicly available (we will discuss this in Section 5). For $m \leq 2$, we consider there are approximately h_1 number of counts that are less than $\overline{x_1}$ and h_2 number of counts that are less than $\overline{x_2}$ and more than $\overline{x_1}$, where $\overline{x_1} < \overline{x_2} < c$.

For the Baseline Algorithm M_B, we assume our query relaxation point $c - \alpha$ is between $\overline{x_2}$ and c to eliminate maximum number of false positives. In the worst case, the expected number of false positives is:

$$\begin{aligned}&\mathbb{E}_{O\sim M_B(D)} n_{fp}(O) = (h_1 + h_2) Pr(\eta > c - \alpha - \overline{x_2} | 0 < \overline{x_2} < c - \alpha) \\ &= (h_1 + h_2)\frac{1}{2}(2\beta)^{\frac{c-\alpha-\overline{x_2}}{\alpha}}\end{aligned} \tag{4}$$

For the Progressive algorithm M_P with $m = 2$, we assume our first query relaxation point $c - \alpha_1$ is between $\overline{x_1}$ and $\overline{x_2}$. In the worst case, the expected number of false positives after first step is:

$$\begin{aligned}&\mathbb{E}_{O\sim M_P{}^{step-1}(D)} n_{fp}(O) \\ &= h_1 Pr(\eta > c - \alpha_1 - \overline{x_1} | 0 < \overline{x_1} < c - \alpha_1) + h_2 \\ &= h_1 \frac{1}{2}(\beta)^{\frac{c-\alpha_1-\overline{x_1}}{\alpha_1}} + h_2\end{aligned} \tag{5}$$

Our second query relaxation point $c - \alpha_2$ is between $\overline{x_2}$ and c. The expected number of false positives for M_P after both steps becomes

$$\begin{aligned}&\mathbb{E}_{O\sim M_P(D)} n_{fp}(O) \\ &= \mathbb{E}_{O\sim M_P{}^{step-1}(D)} n_{fp}(O) Pr(\eta > c - \alpha_2 - \overline{x_2} | 0 < \overline{x_2} < c - \alpha_2) \\ &= \left(h_1 \frac{1}{2}(\beta)^{\frac{c-\alpha_1-\overline{x_1}}{\alpha_1}} + h_2\right)\frac{1}{2}(\beta)^{\frac{c-\alpha_2-\overline{x_2}}{\alpha_2}}\end{aligned} \tag{6}$$

If we choose our ϵ_1 as follows:

$$\frac{ln(\frac{1}{\beta})}{c - \overline{x_2}} < \epsilon_1 < \frac{ln(\frac{1}{\beta})\left(1 + ln(\frac{(h_1+h_2)(2\beta)^{(c-\overline{x_2})/\alpha} - 2h_2(\beta)^{(c-\overline{x_2})/\alpha_2}}{h_1\beta^{(c-\overline{x_1})/\alpha_2}})/ln(\beta)\right)}{c - \overline{x_1}} \tag{7}$$

where $\alpha = ln(1/(2\beta))/\epsilon$ and $\alpha_2 = ln(1/\beta)/\epsilon$, then the expected number of false positives for M_p of two iterations (Equation 6) is smaller than that for M_B (Equation 4). We show in Section 4, this condition effectively gives fewer empirical false positives.

Database	L	$\overline{x_1}$	h_1	$\overline{x_2}$	h_2	c	N_{TP}
D_1	268	10	240	25	28	100	2
D_2	364	10	300	40	60	100	4
D_3	503	10	400	40	103	100	3

Table 1: Datasets and Queries

4 PRELIMINARY RESULTS

In this section, we evaluate our proposed algorithms (Baseline Algorithm and Progressive Algorithm) using one real and two synthetically generated data sets. The real dataset is created from the data collected in the TIPPERS system. These datasets include location of users inside a building mimicking the set up of monitoring environments. We refer to the real data set as D_1 and the synthetically generated data sets as D_2, and D_3. The details of the datasets are

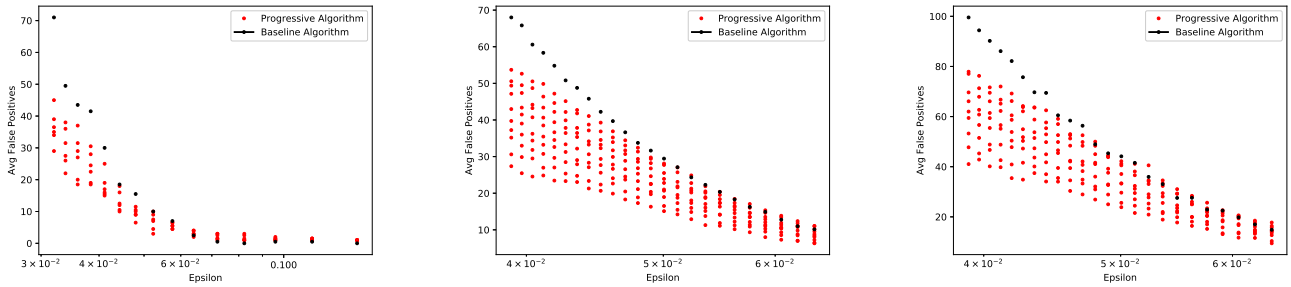

Figure 2: Average number of false positives for the Baseline Algorithm and the Progressive Algorithm at optimal settings (based on Equation 7). From left to right are results for D_1, D_2 and D_3 shown.

shown in Table 1. We choose an ICQ query for each dataset where the threshold values c and the resulted number of true positives (N_{TP}) is also shown in Table 1.

In all the experiments, we choose β-false negative rate to be $\beta = 0.05$. The number of times we ran each algorithm was 100 and in the results we plot average results of those 100 iterations.

Accuracy Requirement. We ran our algorithms for multiple values of epsilon for all the three data sets and showed their accuracy guarantees. Figure 3 plots that the average number of false negatives (solid lines) for the Baseline Algorithm and the bound $\beta \times N_{TP}$ where N_{TP} is the number of true positives (dashed line). The average values are within the bound for all three datasets. We found similar results for the Progressive Algorithm.

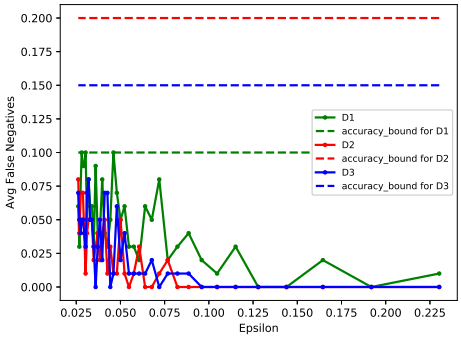

Figure 3: Average number of false negatives vs Epsilon for the Baseline Algorithm

Expected False Positives. Figure 2 shows the average number of false positives for the Baseline Algorithm (blue dots) and the Progressive Algorithm (red dots) at various total privacy budget ϵ. The values for ϵ_1 in the Progressive Algorithm are sampled from the range based on Equation 7. For all the three data sets, it can be observed that the number of false positives for the Progressive Algorithm is smaller than the number of false positives for baseline algorithm for most settings. There are few cases where the chosen values for $\{\epsilon_1, \epsilon_2 = \epsilon\}$ result in more false positives than the Baseline Algorithm, due to the approximation error in the model for the expected false positives. Improving this model is a future direction for our work.

5 DISCUSSION

In this section we discuss some of the challenges that we face as we extend our initial ideas towards a more comprehensive solution.

- In our existing work we have assumed that an accurate/close to accurate public knowledge is available about the data distribution. This is a valid assumption in certain settings where occupancy levels and fluctuations are predictable (e.g., class rooms holding classes with known enrollments). For situations when distributions are not known, we need to learn the distribution privately possibly using a small amount of privacy budget. Extending our approach to an optimal two-phase algorithm (e.g., how to split the privacy budget across phases) remains a challenge.
- We need to determine how to improve the cost model for the expected number of false positives of different strategies.
- We need to develop more formally a privacy metric that can enable comparison between two strategies with the same privacy budget and the same number of false positives.
- In our initial development of the idea, we have fixed the number of steps in the progressive approach. The number of steps can have impact on the quality of the strategy. We, thus, need to explore optimal mechanisms to decide on the number of steps to use in the progressive strategy as a function of the distribution.
- Our development has focused on simple iceberg queries. We will need to generalize the work to a more general class of event detectors.

REFERENCES

[1] C. Dwork and A. Roth. The algorithmic foundations of differential privacy. *Found. Trends Theor. Comput. Sci.*, 9(3–4):211–407, Aug. 2014.

[2] U. Erlingsson et al. Rappor: Randomized aggregatable privacy-preserving ordinal response. In *ACM SIGSAC*, 2014.

[3] C. Ge et al. Apex: Accuracy-aware differentially private data exploration. SIGMOD, 2019.

[4] S. Ghayyur et al. Iot-detective: Analyzing iot data under differential privacy. SIGMOD '18, pages 1725–1728, New York, NY, USA, 2018. ACM.

[5] N. Johnson et al. Towards practical differential privacy for sql queries. *Proc. VLDB Endow.*, 11(5):526–539, Jan. 2018.

[6] F. Koufogiannis et al. Gradual release of sensitive data under differential privacy. *CoRR*, abs/1504.00429, 2015.

[7] A. Machanavajjhala et al. Privacy: Theory meets practice on the map. In *2008 IEEE 24th International Conference on Data Engineering*, pages 277–286, April 2008.

[8] S. Mehrotra et al. Tippers: A privacy cognizant iot environment. In *2016 IEEE PerCom Workshops*, March 2016.