Enhancing Privacy in Participatory Sensing Applications with Multidimensional Data

Michael M. Groat*, Benjamin Edwards*, James Horey†, Wenbo He‡, and Stephanie Forrest*

*Department of Computer Science, University of New Mexico, Albuquerque, New Mexico 87131, USA
†Computational Sciences & Engineering, Oak Ridge National Laboratory, Oak Ridge, Tennessee 37831, USA
‡School of Computer Science, McGill University, Montreal, Quebec H3A 2T5, Canada

***Abstract*—Participatory sensing applications rely on individuals to share local and personal data with others to produce aggregated models and knowledge. In this setting, privacy is an important consideration, and lack of privacy could discourage widespread adoption of many exciting applications. We present a privacy-preserving participatory sensing scheme for multidimensional data which uses *negative surveys*. Multidimensional data, such as vectors of attributes that include location and environment fields, are challenging for privacy protection and are common in participatory sensing applications. When reporting data in a negative survey, an individual participant randomly selects a value from the set complement of the sensed data value, once for each dimension, and returns the negative values to a central collection server. Using algorithms described in this paper, the server can reconstruct the probability density functions of the original distributions of sensed values, without knowing the participants' actual data. Our algorithms avoid computationally expensive encryption and key management schemes, conserving energy. We study trade-offs between accuracy and privacy, and their relationships to the number of dimensions, categories, and participants. We introduce *dimensional adjustment*, a method that reduces the magnification of error associated with earlier work. Two simulation scenarios illustrate how the approach can protect the privacy of a participant's multidimensional data while allowing useful aggregate information to be collected.**

***Index Terms*—multidimensional data; negative surveys; privacy protection; participatory sensing applications**

I. Introduction

Participatory sensing applications [7] sense, collect, analyze, and share local information or knowledge collected from a large population of people, enabling a wide range of applications such as urban planning [8], public health [11], and vehicular transportation monitoring [24], [31]. In these applications, the privacy of those carrying sensing devices who are willing to share their information should be respected, especially when information travels across open wireless networks. On the other hand, it is desirable to generate high quality data for policymakers, researchers, and the public. Hence, trade-offs exist between protecting the privacy of the participants' data and the utility gained from examining this content.

We seek to preserve the privacy of multidimensional data where all dimensions are sensitive. For example, we present a radiation detection scenario that determines the distribution of radiation levels at various locations. Participants disguise both dimensions: their geographic location, and their local radiation level. Non-sensitive dimensions can remain un-perturbed if (1) the data cannot be linked to a particular individual and (2) there are no correlations between sensitive and non-sensitive dimensions.

Existing approaches for protecting privacy of multidimensional data [1], [19], [30] are designed for database applications, where large numbers of records from different users are available to a centralized server that summarizes statistics about the records [1], [30], [33]. However, in participatory sensing applications, individual nodes typically have access only to their own sensed values. Participants might not be willing to share information with other participants or trust a central collection server.

Our approach applies *negative surveys* [15], [16], [24] to categorical multidimensional data, where the categories might be symbolic values (e.g., gender) or a coarse-graining of numerical data into bins. A set of categories forms a proper partition over each dimension. Individual participants disguise data by reporting for each dimension a category from the set complement of the sensed category. A base station is then able to reconstruct the original distribution of sensed categories from this disguised data [24]. This approach avoids complicated encryption and key management schemes, thus conserving energy on the nodes.

Using privacy and utility metrics taken from Huang et al. [26], we quantify the trade-offs between the accuracy of this reconstruction and the amount of privacy protected. These metrics and some terminology are borrowed from the privacy-preserving data mining field. We use the terms, disguise, perturb, and negate interchangeably.

Our threat model treats the base station as an *honest but curious* [6], [21] entity. That is, we assume it faithfully follows the network protocols but could mischievously try to collect information to use against the nodes. Additional threats come from eavesdroppers listening to radio communications who try to intercept packets. We assume that all nodes are equipped with sensors for data capture.

One of the limitations of previous work with negative surveys was the requirement for a large number of participants to reconstruct the data accurately [24], [35]. A slight increase in the number of categories requires a significant increase in the number of participants to maintain a given level of utility. The problem is compounded when the data are multidimensional. We present a method called *dimensional adjust-*

ment that controls this error, reducing the number of required participants and improving utility 2.5 times more than the loss of privacy.

We illustrate our algorithms with two simulations. In the first simulation, a cell phone radiation detection scenario locates radiation threats such as unexploded dirty bombs, escaped radiation from a nuclear reactor accident, or lost or stolen medical waste, while not revealing individuals' locations. The second simulation reconstructs the underlying probability density function of continuous data, suggesting an alternative approach to *random data perturbation* [2].

The main contributions of this paper include: (1) We present the use of negative surveys on multivariate categorical data to protect privacy in participatory sensing applications. This includes a novel efficient reconstruction algorithm. To our knowledge, this is the first work addressing privacy of multidimensional data in participatory sensing applications. (2) We analyze the benefits of using negative surveys in participatory sensing applications compared to other perturbation approaches. (3) We introduce dimensional adjustment which reduces the needed number of participants to maintain a given level of utility, at the expense of a small amount of privacy. And, (4) we study usability in terms of reconstruction error and the strength of privacy through theoretical analysis and simulations.

The remainder of this paper is structured as follows. Section II gives background information on negative surveys and randomized response techniques. Our protocols are presented in Section III, and Section IV describes the privacy and utility metrics used in the analysis. Section V gives the benefits of using negative surveys. Dimensional adjustment is introduced and analyzed in Section VI. Section VII describes two simulations, reporting accuracy and privacy results for each. We discuss our simulations and speculate on how to improve their performance in Section VIII. Section IX discusses related work, and Section X gives future work and our conclusions.

II. Background

This section reviews background material on randomized response techniques and a specific instance of these, negative surveys in their single dimensional form.

Randomized response techniques (RRTs) disguise data by perturbing a categorical value to another value. For example, if race is Hispanic, it could be perturbed to Asian. A *perturbation matrix*, denoted M, gives the probabilities of perturbing category i to category j. It is an α by α square matrix where the columns sum to one and α is the number of categories.

Finding the optimal M that balances both privacy and utility has been the subject of earlier research [4], [26]. Warner described the RRT for binary data [34], however, it can be extended to categorical data [3] using the following perturbation matrix, which gives an initial suggestion for M:

$$M = \begin{pmatrix} p & \frac{1-p}{\alpha-1} & \cdots \\ \frac{1-p}{\alpha-1} & p & \cdots \\ \vdots & \vdots & \ddots \end{pmatrix}, \tag{1}$$

where p is the probability of a category remaining unchanged.

The original data can then be estimated from the disguised data with the following equation:

$$\widehat{A} = M^{-1}\widehat{Y}, \tag{2}$$

where $\widehat{Y} = (Y_1, ..., Y_\alpha)^\tau$ and Y_i is the number of disguised values in the i^{th} category. Since this is an unbiased maximum likelihood estimate, $\widehat{A}$ approaches the original distribution as the population grows. Equation (2) is known as the matrix inversion approach. An iterative approach is given by Agrawal et al. [3] but is not developed for multiple dimensions.

We review a special case of the Warner scheme called negative surveys [15], [16], [24]. Negative surveys use a specialized perturbation matrix containing zeros on the diagonal entries and equal values everywhere else, with the columns summing to one, i.e., $p = 0$ in Equation (1). We call these matrices *negative survey perturbation matrices (NSPMs)*.

A negative survey consists of two protocols. The first, or *node protocol*, maps the sensed data into its negative representation. To do this, each node chooses a category it did not sense with uniform probability and returns that negative information to the base station. The second, or *base station protocol*, reconstructs the original data at the base station. Instead of Equation (2), the following simpler equation [15] can be used:

$$\forall i \mid A_i = N - (\alpha - 1) \cdot Y_i, \tag{3}$$

where A_i is the reconstructed number of values in category i, and Y_i is the reported perturbed number of values in category i, with $1 \leq i \leq \alpha$. N is the total number of sensed values. Equation (3) has time complexity $O(\alpha)$, compared to $O(\alpha^2)$ for Equation (2) (ignoring matrix inversion), while still remaining an unbiased maximum likelihood estimate.

III. Protocols

In our multidimensional protocols, the individual sensing device also always reports false data. Although, after the base station receives all the false reports, it reconstructs the data differently to approximate the true distribution of the original sensed values. Before we describe the multidimensional node and base station protocols, we introduce some notation.

For the entire population, X, Y, and A are D-dimensional matrices which represent the counts of the categories of the original, disguised, and reconstructed data sets respectively. For example, if D=3 then $X(a, b, c)$, $Y(a, b, c)$, and $A(a, b, c)$ are counts of all the values that occur in the a^{th}, b^{th}, and c^{th} category in the first, second, and third dimensions. Vectors such as $\vec{x}$=$<a, b, c>$ will indicate a specific index.

An individual participant senses vector $\vec{x}^+$=$<x_1^+, x_2^+, \cdots, x_D^+>$ from its environment. Sensed real values are quantized into categories, if necessary. Each $x_i^+ \in \vec{x}^+$ where $1 \leq i \leq D$, reflects that category x_i was sensed in dimension i. x_i is drawn from a set of categories C_i=$\{1, 2, \cdots, \alpha_i\}$, that form a proper partition over the data in dimension i, and α_i is the total number of categories in dimension i. The "+" in $\vec{x}^+$ denotes the positive or sensed categorical information, as

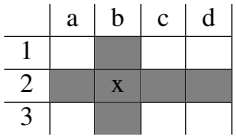

Fig. 1. A sensor that reads $< 2, b >$ selects among the white cells to report.

opposed to the negative or perturbed information represented as $\vec{x}^-$.

A. Node Protocol

There are three phases to the node protocol:

1) **Sensing:** A node senses a multidimensional value $\vec{x}^+$ from its environment and quantizes into categories if necessary.
2) **Negation:** For each $x_i^+ \in \vec{x}^+$, the node selects uniformly at random a category x_i^- to report to the base station from the set of possible categories C_i, such that $x_i^- \neq x_i^+$. It does this for each dimension, creating the perturbed vector $\vec{x}^-$. The probability of selecting a perturbed category is $\frac{1}{\alpha_i - 1}$, where α_i is the number of categories in dimension i. For example in Figure 1, a node has sensed $\vec{x}^+ = <2, b>$ from its environment, and must choose among the white cells, for instance $\vec{x}^- = <3, c>$, for a negative value to report back to the base station.
3) **Transmission:** After negation, the node sends $\vec{x}^-$ to the base station. We assume no data aggregation in the network.

Since the number of bits that is required to transmit either the positive or negative data is identical, there is only a slight increase in resource cost to compute and transmit the perturbed value. Hence, the node protocol saves resources compared to encryption methods with key management [24].

B. Base Station Protocol

The base station collects the reported data, Y, and estimates the original distributions of sensed values, A. In the single-dimensional case, Equation (3) is used to obtain this estimate [16], [24]. We give a method for a multidimensional approach and later present a time optimization.

1) Reconstruction: Each dimension must use a NSPM. If Equation (3) were extended to D dimensions, the reconstruction equation would be:

$$\forall \vec{x} \mid A(\vec{x}) = N + \sum_{k=1}^{D} (-1)^k \cdot \Gamma(\vec{x}, k), \tag{4}$$

where $\Gamma(\vec{x}, k)$ is given as:

$$\Gamma(\vec{x}, k) = \sum_{\substack{d \in \\ B(\{1,\ldots,D\},k)}} \left(\left[\prod_{j \in d} (\alpha_j - 1) \right] \cdot \sum_{\substack{\vec{y}\ s.t. \\ y_i \in \vec{x}, \\ \forall i \in d}} Y(\vec{y}) \right), \tag{5}$$

and $B(\{1, ..., D\}, k)$ is all the k length possible combinations of members of $\{1, ..., D\}$. For example, $B(\{1, 2, 3\}, 2)$ is $\{\{1, 2\}, \{1, 3\}, \{2, 3\}\}$. $Y(\vec{y})$ is the count of the reported disguised sensed values that have categories specified by d from $\vec{x}$. As an example, Equation (4) with D=3 is given as:

$$\begin{aligned} A(a,b,c) = {} & N - (\alpha_1 - 1) \sum_{\substack{\vec{y}\ s.t. \\ y_1 = a}} Y(\vec{y}) - (\alpha_2 - 1) \sum_{\substack{\vec{y}\ s.t. \\ y_2 = b}} Y(\vec{y}) \\ & - (\alpha_3 - 1) \sum_{\substack{\vec{y}\ s.t. \\ y_3 = c}} Y(\vec{y}) + (\alpha_1 - 1)(\alpha_2 - 1) \sum_{\substack{\vec{y}\ s.t. \\ y_1 = a, \\ y_2 = b}} Y(\vec{y}) \\ & + (\alpha_1 - 1)(\alpha_3 - 1) \sum_{\substack{\vec{y}\ s.t., \\ y_1 = a, \\ y_3 = c}} Y(\vec{y}) + (\alpha_2 - 1)(\alpha_3 - 1) \sum_{\substack{\vec{y}\ s.t. \\ y_2 = b, \\ y_3 = c}} Y(\vec{y}) \\ & - (\alpha_1 - 1)(\alpha_2 - 1)(\alpha_3 - 1) Y(a,b,c), \quad \forall a, b, c \end{aligned} \tag{6}$$

The time complexity of Equation (4) is given as:

$$O\left(2^D \cdot \prod_{i=1}^{D} \alpha_i\right). \tag{7}$$

This is because there are $\binom{D}{0} + \binom{D}{1} + \cdots + \binom{D}{D}$, or 2^D, total Y terms in Equation (4). For example, Equation (6) has 8 total Y terms. N is another Y term that counts the entire number of participants. For each $Y(d)$ term, we need a count of the specific index d, whose time complexity involves the product of the number of categories in each dimension. This complexity is exponential with respect to the number of dimensions.

2) Optimization: We now present a more efficient alternative to the previous reconstruction process that uses dynamic programming and could be applied to any perturbation matrix, not just the NSPM. It is given in Algorithm 1, where α_δ is the number of categories for the δ^{th} dimension; ":" is an operation on a matrix designating every element in that dimension; τ is a function similar to transpose that takes a row, column, hyper-row, or hyper-column, and transforms it into a vector appropriate for matrix multiplication. M_δ is the α_δ by α_δ square perturbation matrix for the δ^{th} dimension.

The time complexity of Algorithm 1 is:

$$O\left(\sum_{i=1}^{D} \left[\prod_{j=1, j \neq i}^{D} \alpha_i^2 \alpha_j \right]\right) = O\left(\sum_{i=1}^{D} \alpha_i \cdot \prod_{i=1}^{D} \alpha_i\right), \tag{8}$$

ignoring the cost of matrix inversion for each M_δ. Intuitively, this complexity is based on a matrix multiplication with every possible vector in Y. However, if only NSPMs are used for each dimension, the cost of Algorithm 1 reduces to:

$$O\left(D \cdot \prod_{i=1}^{D} \alpha_i\right), \tag{9}$$

because line 10 in Algorithm 1 is replaced with the simpler Equation (3). This dynamic programming algorithm has the advantage that different parts of Y are updated through each iteration. Information is reused and does not need to be recalculated. The complexity denoted by Equation (9) is clearly an improvement over the complexity denoted by Equation (7).

Algorithm 1 Reconstruction Optimization for D Dimensions. Y is the D dimensional matrix of reported disguised values, $F = [\alpha_1, \ldots, \alpha_D]$ is a list of the number of categories for each dimension, and $M = [M_1, \ldots, M_D]$ contains the perturbation matrices for each dimension. The symbol : denotes the slice operator. $index$ is constructed to be a vector of length D, with a single instance of :. When used as an index into R, it returns a vector.

```
function reconstruct_matrix(Y, D, F, M)
    R = Y
    for δ ∈ [1 : D]
        update_dim(R, D, [], δ, F, M)
    end
    return R

function update_dim(R, D, index, δ, F, M)
    if length(index) = D
        R(index) ← M_δ^{-1} * R(index)^τ
    elseif len(index) + 1 = δ
        new_index ← index.append([:])
        update_dim(R, D, new_index, δ, F, M)
    else
        for i ∈ [1 : F(length(index) + 1)]
            new_index ← index.append([i])
            update_dim(R, D, new_index, δ, F, M)
        end
    end
```

IV. Privacy and Utility Metrics

We give metrics for the privacy and utility of multidimensional negative surveys which are an extended form of the one-dimensional case given by Huang and Du [26]. These formulations apply to any perturbation matrix, not necessarily a NSPM, and serve as a foundation for comparing different matrices. Both metrics range from 0 to 1, with the lower value being more desirable.

A. Privacy Metric

Privacy measures the probability of guessing the original data from the disguised values, and is based on the maximum a posteriori estimate. It is related to the Shannon Entropy of the underlying distribution. The higher the entropy, the better the privacy. If an underlying distribution contains more values in any particular category than another, it is easier to guess or predict that category. The privacy metric is:

$$Privacy = \sum_{\substack{\Upsilon \in Y(\vec{x}) \\ \forall \vec{x}}} P(\Upsilon | \widehat{X_\chi}) \cdot P(\widehat{X_\chi}), \tag{10}$$

where

$$\widehat{X_\chi} = \arg \max_{\substack{\chi \in X(\vec{x}) \\ \forall \vec{x}}} P(\chi | Y). \tag{11}$$

Equation (11) calculates for Equation (10) the optimal maximum a posteriori estimate for a given index of Y. This is the index that has the maximum probability in $P(X|Y)$ (the maximum index in each column of $P(X|Y)$). This metric assumes an adversary has no prior knowledge of the distribution of perturbed data.

B. Utility Metric

Utility, also known as accuracy or reconstruction error, measures the difference between the original and reconstructed data distributions. It is measured with the mean square error, calculated from the variance and co-variance as follows:

$$\begin{aligned} Utility &= E(P(A) - P(X))^2 \\ &= \frac{1}{\alpha_1 \cdot \ldots \cdot \alpha_D} \sum_{\vec{x}_1} E(P(A = \vec{x}_1) - P(X = \vec{x}_1))^2 \\ &= \frac{1}{\alpha_1 \cdot \ldots \cdot \alpha_D} \sum_{\vec{x}_1} \Bigg(\sum_{\vec{x}_2} \Big[\mu(\vec{x}_1, \vec{x}_2)^2 \cdot var(\vec{x}_2) \Big] \\ &+ \sum_{\substack{\vec{x}_3, \vec{x}_4 \ s.t. \\ c_i \in \vec{x}_3 \neq c_i \in \vec{x}_4}} \Big[2 \cdot \mu(\vec{x}_1, \vec{x}_3) \cdot \mu(\vec{x}_1, \vec{x}_4) \cdot cov(\vec{x}_3, \vec{x}_4) \Big] \Bigg), \end{aligned} \tag{12}$$

where

$$\mu(\vec{x}_i, \vec{x}_j) = \prod_{k=1}^{D} M_k^{-1}(c_k \in \vec{x}_i, c_k \in \vec{x}_j), \tag{13}$$

denotes the $c_k \in \vec{x}_i$ row and $c_k \in \vec{x}_j$ column in M_k^{-1}; and variance and covariance are:

$$var(\vec{x}_i) = \frac{1}{N} \cdot P(Y = \vec{x}_i) \cdot (1 - P(Y = \vec{x}_i)), \tag{14}$$

$$cov(\vec{x}_i, \vec{x}_j) = -\frac{1}{N} \cdot P(Y = \vec{x}_i) \cdot P(Y = \vec{x}_j). \tag{15}$$

V. Benefits of a NSPM

NSPMs are well suited for participatory sensing applications, and we identify several reasons to base our scheme on them:

1) They are appropriate for resource-constrained devices because perturbation is simplified at the resource-constrained nodes. M does not need to be stored at a node, or used in the perturbation process. Negative values are simply chosen with equal probabilities, which can be beneficial if there are a large number of categories.

2) All samples are guaranteed to be perturbed for a NSPM, unlike RRTs. If the perturbation matrix has non-zero values on the diagonal, there is a small chance of a record maintaining all of its original values. This could be viewed as a privacy breach even if it only occurs in 1 record out of a million [17].

3) Sensors often do not know the prior distribution of their environment. However, in order to find the best perturbation matrix this distribution needs to be known *a priori* [26]. Our evidence empirically suggests that for a NSPM, utility is independent of the underlying distribution. This implies that utility can be known *a priori*, even if the underlying distribution is not.

4) Figure 2 suggests that an optimal value for p in Equation (1) is zero (a NSPM). The figure plots privacy and utility metrics against values of p for the perturbation matrices. Ten categories were used, and there are five different underlying data distributions. Each underlying distribution is drawn from an ideal distribution that contains a particular Shannon Entropy, labeled in the legend. The utility values for each

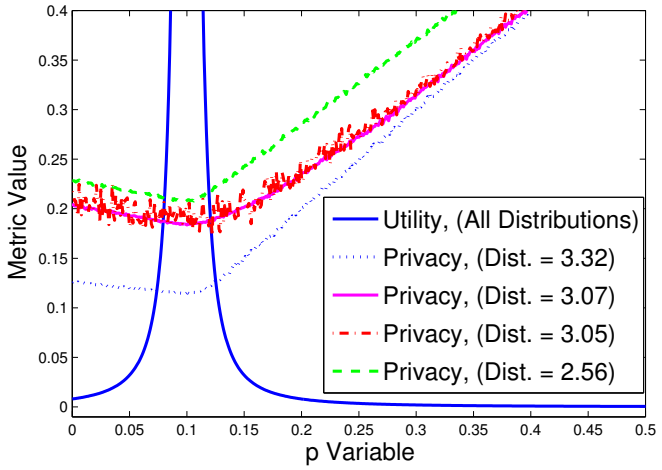

Fig. 2. Privacy and utility values using different values of p of Equation (1). Each "checkmark" curve represents the privacy value of a different underlying distribution, denoted in the legend by its Shannon entropy. Utility is nearly the same for all 4 different distributions and is given once.

distribution are similar enough that they appear as a single curve.

Trade-offs exist between privacy and utility: The best privacy gives the worst utility, and vice versa. For example in Figure 2, when p approaches 0.1, utility asymptotically increases. This is explained by the fact that there are 10 categories and every entry of the perturbation matrix is the same, $\frac{1}{10}$. At the other extreme, $p = 0.1$ provides the best privacy, as seen by examining the minimum of the "checkmark" curves.

Zero is the best value for p. Consider the points at $p = 0$ and $p = 0.2$ in Figure 2, where utility is equal but privacy is not. Privacy is better at $p = 0$, and has a better combined privacy and utility value than any other location on the graph, excepting the tails. If a certain level of privacy is desired, perhaps below 0.3, there is no better choice for p. As a caveat, we observe that the underlying distribution must have a sufficiently high Shannon entropy for NSPMs to be optimal.

VI. Dimensional Adjustment

A challenge of NSPMs observed in previous work for single dimensional data [24], [35] is that a slight increase in the number of categories significantly increases the required number of participants for a given utility value. This increase is compounded with each additional dimension and has limited negative surveys to small applications. In this paper, we propose the *dimensional adjustment (DA)* method to address this challenge.

DA increases utility by accepting a slight decrease in privacy for a given number of participants. It accomplishes this by distributing the same overall number of categories over an increasing number of dimensions. For example, if an original one-dimensional negative survey contains 64 categories, it can be remapped to: 2 dimensions of 8 categories each; 2 dimensions of 4 and 16 categories; or any number of dimensions where the product of the number of categories in each dimension equals 64. For space reasons, the details

TABLE I
Two negative surveys of 10,000 total categories and 1,000,000 participants. The second uses dimensional adjustment.

	1 dimension of 10,000 categories	6 dimensions of 5x5x5x5x4x4 categories
utility	0.00100	0.00014
privacy	0.01457	0.01960

for remapping dimensions have been left out, but are trivial to implement in the previous protocols.

Splitting data into multiple dimensions with a smaller number of categories for each dimension improves reconstruction accuracy (utility). Intuitively, accuracy is related to Figure 1 and the ratio of the white squares (negative information) to the total number of squares. As the number of dimensions grows, and the number of distinct categories remains the same, this ratio decreases, reducing the possible number of cells for perturbed data, which increases the accuracy of reconstruction.

There are trade-offs between a high number of dimensions with a low number of categories, versus a low number of dimensions with a high number of categories. A one-dimensional negative survey with 64 categories provides the best privacy but the worst utility, compared to 6 dimensions with 2 categories each, which provides the worst privacy but the best utility. The relationship between privacy and utility is usually nonlinear, providing an opportunity to sacrifice a small amount of one for a significant gain in the other. For example, in Table I with 1,000,000 samples and 10,000 categories, we see privacy degrades 34% while utility improves 86%.

Using Table I and modeling equations for privacy and utility, we further illustrate these trade-offs. Without loss of generality, the normal distribution is used as the original distribution, X, in Table I. The multidimensional negative survey that uses 6 dimensions and 1,000,000 participants is comparable to a single dimensional negative survey that uses 71,414,286 participants. This is calculated by setting the following utility modeling equation for a single dimension:

$$Utility_{model} = (\alpha - 2)/N, \tag{16}$$

to 1.40E-04 (from Table I), α to 10,000, and solving for N. The same multidimensional negative survey is also equivalent to a single dimension using 142 categories. This is calculated by setting Equation (16) to 1.40E-04, N to 1,000,000, and solving for α. Equation (16) has an R^2 value of 0.999.

We perform a similar analysis for the privacy equivalence with Table I and the following privacy modeling equation where the values of the input distribution are normal:

$$Privacy_{model} = \frac{2.5}{(log_2(\alpha))^2 + 1.5}, \tag{17}$$

which has an R^2 value of 0.976. The multidimensional negative survey in Table I is equivalent in privacy to using a single-dimensional negative survey of 2,397 categories, yet previously it had the same utility as 142 categories.

VII. Applications and Simulations

A. Cell phone Radiation Threat Detection

Participatory sensing could help detect and locate radiation threats, for example a nuclear device or incident such as a dirty bomb, lost radioactive material, the spreading threat from a nuclear reactor accident, or medical waste dumps. In this scenario we assume that cell phones are equipped with radiation monitors and GPS devices. Locations are quantized into different groups, with a different label for each group. Individuals care about the privacy of their locations. We show that with reasonable parameter assumptions (number of locations, radiation levels, and participants), multidimensional negative surveys can maintain confidentiality and determine which locations contain radiation threats.

Cell phones are ideal for radiation detection, and the United States Department of Homeland Security has considered their use [18]. If radiation sensors were installed at fixed locations, they might be tampered with or avoided, which would be more difficult with cell phones because they are owned by individuals and exist in large numbers. As an incentive to promote participation, aggregate information could be disseminated freely to participants. Since readings from an individual cell phone might not be as accurate as the combined readings from a larger population, access to aggregate information would be advantageous. However, for an event such as the Fukushima Daiichi nuclear accident, participants might prefer to send the unperturbed data and receive more accurate readings. Either way, in such a situation, immediate feedback would be beneficial, especially to determine if radiation has spread further than publicly acknowledged.

1) Simulation Setup: Before we explain the simulation setup, we give a small example of a geographic area divided into a $3x3$ grid, shown in Figure 3. The total population of cell phones (participants) is 450,000 and is equally divided among the 9 locations. In the actual simulation, we do not assume a uniform population distribution and instead follow a more realistic model given by Bertaud et al. [5]. We simulate three radiation levels: low, medium, and high. Depending on the level of radiation, each location's distribution of reported levels will be shifted lower or higher. For example, in Figure 3, location 6 contains a *threat distribution*, illustrated by the black histogram. This distribution, exponentially shifted towards the higher range, contains 28,571 participants in the high radiation level, 14,286 in the medium radiation level, and 7,143 in the low level. Benign locations, characterized by the *non-threat distributions*, are shown in black at the other locations. These distributions are exponentially distributed in the reverse order.

San Francisco, which has roughly 46.7 square land miles, is our example city. We chose the number of distinct locations to be 48, which works well with DA due to its high number of composites. Each location roughly covers one square land mile, which is small enough for a response team with more powerful equipment (such as helicopters equipped with radiation detectors) to pinpoint the exact location of a threat.

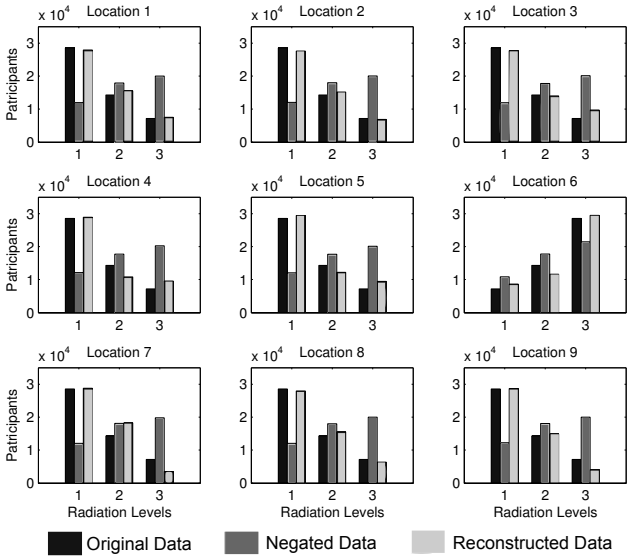

Fig. 3. Histograms for a multidimensional negative survey of 9 locations and 3 radiation levels. Location 6 is suspicious since its radiation levels form a threat distribution. The other locations have non-threat distributions.

San Francisco has a population of about 815,000. We vary the number of participants from 100,000 to 400,000 in increments of 100,000. The spatial distribution of people follows a standard urban model taken from Bertaud et al. [5]. The population is most concentrated at the central business district, and is reduced from the center.

We use 3 categories for the radiation levels. Although our experiments show that more categories would increase the granularity of the data, they would not improve accuracy. In fact, we were able to better determine threats with a lower number of radiation levels. There is a limit, however. If there were only 2 radiation levels, privacy would be lost, and adversaries could determine a user's location, if a threat existed.

Each participant's cell phone, when queried, samples the environment for the radiation level and notes its location. It then perturbs this information according to Section III-A and sends the perturbed values to the base station. After the base station collects the perturbed data (one sample from each cell phone), it reconstructs the original distribution.

The base station determines if a threat exists and if so, at which location. It computes a linear regression at each location from its reconstructed histogram of radiation levels, assuming that histogram values are one unit apart. Ideally, we expect that a location reporting elevated radiation levels will have a positive slope from the linear regression, and a location with a non-threat will have a negative slope. Although, we arbitrarily defined slope thresholds to distinguish threats from not threats, choosing values that minimized the overall numbers of false positives and false negatives. The thresholds could be adjusted to favor one error type over another. For example, one strategy might send response teams to investigate false positives, rather than allowing a false negative to slip through. We chose the thresholds *a posteriori*, but in a real deployment these values

TABLE II
RESULTS OF THE CELL PHONE RADIATION DETECTION SIMULATION. EACH TEST RAN 1,000 TIMES WITH 500 RUNS CONTAINING A RADIATION THREAT (TRUE POSITIVES) AND 500 RUNS CONTAINING NO THREAT (TRUE NEGATIVES).

Samples	False Neg.	False Pos.	Acc. of True Pos. % (Ratio)	Avg. Privacy	Avg. Utility
1 locational dimension with 48 categories					
100,000	246	246	5.5 (14/254)	0.0282	4.54E-04
200,000	244	245	7.8 (20/256)	0.0252	2.27E-04
300,000	244	244	18.0 (26/256)	0.0241	1.51E-04
400,000	241	242	18.9 (49/259)	0.0234	1.13E-04
2 locational dimensions with 8x6 categories					
100,000	246	248	11.8 (30/254)	0.0350	1.90E-04
200,000	250	251	21.2 (53/250)	0.0319	9.48E-05
300,000	222	222	31.3 (87/278)	0.0307	6.32E-05
400,000	199	199	39.2 (119/301)	0.0300	4.74E-05
3 locational dimensions with 4x4x3 categories					
100,000	203	205	56.6 (168/297)	0.0586	3.09E-05
200,000	139	139	81.7 (295/361)	0.0554	1.54E-05
300,000	87	87	92.5 (382/413)	0.0542	1.03E-05
400,000	58	58	94.3 (417/442)	0.0535	7.71E-06
4 locational dimensions with 2x2x4x3 categories					
100,000	17	17	99.4 (480/483)	0.1444	4.41E-06
200,000	0	0	100 (500/500)	0.1411	2.20E-06
300,000	0	0	100 (500/500)	0.1398	1.47E-06
400,000	0	0	100 (500/500)	0.1392	1.10E-06

(and better detection methods) could be chosen *a priori*, with additional domain knowledge.

We ran the simulation 1000 times for the various numbers of participants, assigning the threat distribution to a random location in 500 of the runs. In the other 500 runs we assigned a non-threat distribution to all locations.

2) Results and Analysis: Table II summarizes the results, showing the number of false positives and false negatives. Accuracy is the percentage of true positives that correctly determined the threat location. The average privacy and utility metrics are also shown. Since we are calculating an unbiased maximum likelihood estimate, more participants reduce the number of errors and increase reconstruction accuracy.

Because accuracy was low for a single dimension, shown in the first four rows of Table II, we used DA. We changed the location dimension of 48 categories to 2 dimensions of 6 and 8 categories; 3 dimensions of 4, 4, and 3 categories; and 4 dimensions of 2, 2, 4, and 3 categories. The results are shown in Table II. With 4 dimensions, we obtained 100% accuracy with 200,000 or more participants.

B. Reconstructing Continuous Values

In addition to categorical data such as locations and radiation levels, multidimensional negative surveys could be applied to continuous data such as temperature or humidity. We reconstruct the probability density functions of different underlying distributions and compare the parameters of these distributions to the original parameters. This could have implications in privacy-preserving data mining as an alternative to random data perturbation [2].

1) Simulation Setup: Any fixed point number can be represented as a collection of categories by labeling each digit's position (1's, 10's, 100's,...) with a value ranging from zero to nine. Thus, a fixed point number with n digits is treated as an n dimensional negative survey, with each dimension having ten potential categories; it is then straightforward to apply the protocols presented previously.

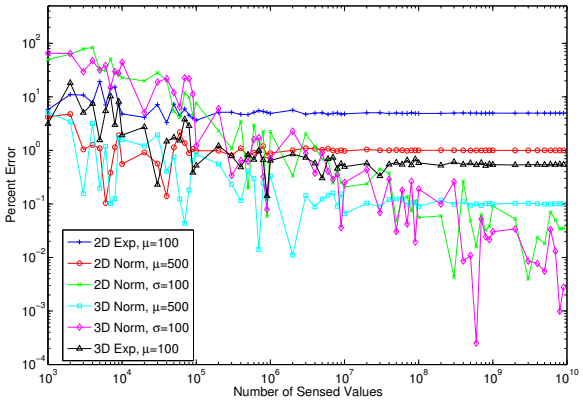

Fig. 4. Parameter reconstruction error from the continuous negative survey simulation measured as a percentage difference from the original parameter.

We generated values from two probability distributions, rounding each value to 2 and 3 significant digits. We used the following normal and exponential distributions:

$$\mathcal{N}(\mu,\sigma) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{\sigma}}, \qquad \mathcal{E}(\mu) = \frac{1}{\mu} e^{-\frac{x}{\mu}}.$$

Tests used $\mathcal{N}(500, 100)$ and $\mathcal{E}(100)$, and we truncated the tails of the distribution at 0 and 1000. To perturb a sensed value, a random digit not equal to the actual digit is reported for each position. We varied the number of samples from the two distributions from 1,000 to 9 billion in exponential increments.

The base station reconstructs the frequency of each number in the significant digit range (the probability density function) of the underlying data according to the protocols in Section III. The parameters from the reconstructed data are determined using a maximum likelihood estimate and then are compared to the original parameters used to construct the data.

2) Results and Analysis: We calculated the difference between the estimated and original parameter and divided by the original parameter. Each data point is an average of 20 runs. Figure 4 show the results. It suggests that a theoretical maximum accuracy depends on the parameter type, the distribution, and the number of dimensions. All parameters are within 5% of the original parameter values after 200,000 sensed values.

VIII. DISCUSSION

In the cell phone simulation, because the data are perturbed, it is almost impossible for the collection server to determine a participant's true location[1]. Most, if not all, encryption methods must eventually trust the final recipient of the data. In contrast, our method does not require such trust. Further, it does not incur the extra computational cost of encryption and

[1]While the cell phone tower could reveal the node's location, the base station cannot determine the location from its own information.

the additional communication overhead to transmit encrypted data, nor the extra cost of key distribution and management.

Sometimes nodes will be captured and masquerade as legitimate nodes, continually sending responses when queried. These nodes might become dishonest or rogue and try to corrupt the aggregate information by not following or altering the node protocol. Rogue nodes could either report the original sensed value or favor some categories over others. This can be addressed by adjusting in the reconstruction process the perturbation matrices, M_δ, for each dimension δ, with the correct probabilities that categories were favored.

Some participants' locations might either be constant (if they are not moving around) or follow regular patterns. If a single cell phone were to report its negative location regularly, an adversary might be able to infer its positive location through long term monitoring of the transmitted values. This is especially important if an ID is transmitted with the data. This threat can be addressed if participants are asked to respond to a base station query only if their location has changed since the last query, or to limit the amount of information sent to the base station.

The communications graph in the cell phone simulation has each node reporting directly to the base station. Routing in traditional wireless sensor networks usually follows a tree path. In situations like these, it could be possible to adopt an aggregation strategy similar to Castelluccia et al. [9], or the negative histograms could be aggregated using a min/max scheme from Groat et al. [22] where each value in the negative histogram is treated as a maximum.

In the second simulation, if participant populations are not sufficiently large, a single participant could report multiple sensed values over time, and the base station could accumulate these multiple reported values to obtain a more accurate estimate of the parameters. Additionally, DA can improve accuracy by changing the samples to a lower base or radix.

IX. Related Work

Privacy-preserving algorithms have been developed for data mining [17], [27], [28], data aggregation [10], [20], [24], and other applications [29], [32]. There are four main classes of solutions: perturbation, k-anonymity, secure multi-party computation (SMC), and homomorphic encryption. The first class hides data values by perturbing individual data or query results [17], [27], [28]. These methods usually assume that the distribution of data/noise is known to obtain accurate results. However, as shown by Kargupta et al. [28] and Huang et al. [27], certain types of data perturbation might not preserve privacy well. The second class, k-anonymization [1], [30], [33], makes a data value or participant indistinguishable from $k-1$ other items. It was originally designed for privacy-preserving data mining, but in participatory sensing applications individual participants sense and share their own data. Hence, there is limited potential to mix individual participants' data with others' data. The third class, SMC techniques [12], [23], [25], rely on a joint computation among a set of involved peers. This is problematic in participatory sensing applications which may incur a high communication or computation overhead when the participant population is large. The fourth class aggregates data based on homomorphic encryption [10], [20], which allows a user to perform data aggregation on individual data without knowing the data. However, in order to interpret the final aggregation result, a server needs to know which users reported data, which is not always desirable.

Dwork et al. [14] introduced the term *pan-private* in the context of streaming algorithms which can protect the state of information inside a node. This is useful for node capture attacks that examine internal data. However, it assumes a secure stream as a precondition of the algorithm, while the work reported in our paper protects the stream of information in transit. Pan-private algorithms, however, work better for complex aggregates such as the t-incidence items, the t-cropped mean, and the fraction of k-heavy hitters [14].

Differential privacy [13] aims to provide the maximal accuracy of responses for users querying a statistical database, while minimizing the ability of these users to identify records in the database. Differential privacy assumes that a trusted server handles and responds to the queries, while negative surveys, on the other hand, do not assume that the server is trustworthy.

Gaussian negative surveys (GNSs) [35] also reduce the number of participants needed for accurate negative survey reconstruction. Xie et al. propose a special perturbation matrix where each column is represented as a Gaussian distribution with the mean centered over the original category, which is represented as zero. With location data, this perturbs an individual's location a Gaussian random distance away from the original location. This special perturbation matrix eliminates the need for reconstruction at the base station. However, GNSs with location data do not protect privacy as well as negative surveys. The privacy guarantee of an individual participant depends on the variance of the Gaussian distributions in the perturbation matrix. This variance must be small enough to maintain an acceptable level of utility and number of participants, however, smaller values do not perturb a location a sufficient amount of distance. This may make it easier for an adversary to determine the general location of an individual participant. It is not until the variance is increased to cover more than the entire column of the perturbation matrix that GNSs approach the same privacy guarantee as traditional negative surveys.

X. Conclusion and Future Work

Information such as physical locations, driving speeds, or medical information, can have devastating effects if intercepted by adversarial parties. Multidimensional negative surveys perturb data for participatory sensing applications, providing high levels of privacy. The privacy-preservation problem addressed here is challenging, because (1) users may not trust the information collection server, and (2) embedded or sensor devices may have limited resources. Thus, we do not rely on standard encryption schemes or key distribution and management. An advantage of our work is that privacy and accuracy can be

managed by simply tuning parameters of the protocols. Our method scales well because the communication and computation overhead is low for the sensor nodes, especially when compared to expensive encryption schemes.

Future work will examine the limits of dimensional adjustment on real-world data sets with large numbers of categories. Although histograms are useful, we are interested in reconstructing other aggregates.

ACKNOWLEDGMENT

The authors thank Rammohan, Williams, and Esponda for their suggestions, insights, and ideas. MG acknowledges support from Motorola, Inc, Eli Lilly and Company, and NSF grant HRD-0622930; WH acknowledges support from DOE NNSA grant DE-FG52-06NA27494; SF acknowledges the partial support of NSF (grants CCF-0621900, CCR-0331580, SHF-0905236), AFOSR MURI grant FA9550-07-1-0532, and DARPA grant P-1070-113237.

REFERENCES

[1] C. C. Aggarwal, "On k-Anonymity and the curse of dimensionality," in *Proc. 31st Int. Conf. Very Large Data Bases*, Trondheim, Norway, Aug. – Sep. 2005, pp. 901–909.

[2] R. Agrawal and R. Srikant, "Privacy-preserving data mining," in *Proc. 2000 ACM SIGMOD Int. Conf Management of Data*, Dallas, TX, Jun. 2000, pp. 439–450.

[3] R. Agrawal, R. Srikant, and D. Thomas, "Privacy preserving olap," in *Proc. 2005 ACM SIGMOD Int. Conf. Management of Data*, Baltimore, MD, Jun. 2005, pp. 251–262.

[4] S. Agrawal and J. R. Haritsa, "A framework for high-accuracy privacy-preserving mining," in *Proc. 21st Int. Conf. Data Engineering*, Tokyo, Japan, Apr. 2005, pp. 193–204.

[5] A. Bertaud and S. Malpezzi, "The spatial distribution of population in 48 world cities: Implications for economies in transition," 2003, unpublished manuscript.

[6] V. Bozovic, D. Socek, R. Steinwandt, and V. I. Villanyi, "Multi-authority attribute based encryption with honest-but-curious central authority," in *IACR Eprint Archive*, 2009. [Online]. Available: http://eprint.iacr.org/2009/083

[7] J. Burke, D. Estrin, M. Hansen, A. Parker, N. Ramanathan, S. Reddy, and M. B. Srivastava, "Participatory sensing," in *World Sensor Web Workshop, ACM Sensys*, Boulder, CO, Oct. 2006.

[8] A. T. Campbell, S. B. Eisenman, N. D. Lane, E. Miluzzo, and R. A. Peterson, "People-centric urban sensing," in *Proc. 2nd Annu. Int. Workshop on Wireless Internet*, Boston, MA, Aug. 2006, p. 18.

[9] C. Castelluccia and C. Soriente, "ABBA: A balls and bins approach to secure aggregation in WSNs," in *6th Int. Symp. Modeling and Optimization in Mobile, Ad Hoc, and Wireless Networks*, Berlin, Germany, Apr. 2008, pp. 185–191.

[10] C. Castelluccia, E. Mykletum, and G. Tsudik, "Efficient aggregation of encrypted data in wireless sensor networks," in *2nd Annu. Int. Conf. Mobile and Ubiquitous Systems: Networking and Services*, San Diego, CA, Jul. 2005, pp. 109–117.

[11] J. Corburn, "Confronting the challenges in reconnecting urban planning and public health," *Amer. J. Public Health*, vol. 94, no. 4, pp. 541 – 549, Apr. 2004.

[12] R. Cramer, I. Damgård, and S. Dziembowski, "On the complexity of verifiable secret sharing and multiparty computation," in *Proc. 32nd Annu. ACM Symp. Theory of Computing*, Portland, OR, May 2000, pp. 325–334.

[13] C. Dwork, "Differential privacy," in *Proc. 33rd Int. Colloq. Automata, Languages and Programming, Part II*, Venice, Italy, Jul. 2006, pp. 1–12.

[14] C. Dwork, M. Naor, T. Pitassi, G. Rothblum, and S. Yekhanin, "Pan-private streaming algorithms," in *Proc. 1st Symp. Innovations in Computer Science*, Beijing, China, Jan. 2010, pp. 66–80.

[15] F. Esponda, "Negative surveys," *ArXiv Mathematics e-Prints*, Aug. 2006.

[16] F. Esponda and V. M. Guerrero, "Surveys with negative questions for sensitive items," *Statistics & Probability Letters*, vol. 79, no. 15, pp. 2456–2461, Dec. 2009.

[17] A. Evfimievski, R. Srikant, R. Agrawal, and J. Gehrke, "Privacy preserving mining of association rules (invited journal version)," *J. Information Systems*, vol. 29, no. 4, pp. 343–364, Jun. 2004.

[18] FOXNEWS.com, "Homeland security looking into cell phones as anti-terror device," 2007, *http://www.foxnews.com/story/0,2933,270033,00.html*.

[19] F. Furfaro, G. M. Mazzeo, and D. Saccà, "A probabilistic framework for building privacy-preserving synopses of multi-dimensional data," in *Proc. 20th Int. Conf. Scientific and Statistical Database Management*, Hong Kong, China, Jul. 2008, pp. 114–130.

[20] J. Girao, D. Westhoff, and M. Schneider, "CDA: Concealed data aggregation for reverse multicast traffic in wireless sensor networks," in *Proc. 40th IEEE Int. Conf. Communications*, Seoul, Korea, May 2005.

[21] O. Goldreich, *Foundations of Cryptography: Volume 2, Basic Applications*. New York, NY: Cambridge University Press, 2004.

[22] M. M. Groat, W. He, and S. Forrest, "KIPDA: k-Indistinguishable privacy-preserving data aggregation in wireless sensor networks," in *Proc. 30th IEEE Int. Conf. Computer Communications*, Shanghai, China, Apr. 2011, pp. 2024–2032.

[23] J. Halpern and V. Teague, "Rational secret sharing and multiparty computation: Extended abstract," in *Proc. 36th Annu. ACM Symp. Theory of Computing*, Chicago, IL, Jun. 2004, pp. 623–632.

[24] J. Horey, M. M. Groat, S. Forrest, and F. Esponda, "Anonymous data collection in sensor networks," in *4th Annu. Int. Conf. Mobile and Ubiquitous Systems: Computing, Networking and Services*, Philadelphia, PA, Aug. 2007, pp. 1–8.

[25] Q. Huang, H. J. Wang, and N. Borisov, "Privacy-preserving friends troubleshooting network," in *Proc. 12th Annu. Symp. Network and Distributed Systems Security*, San Diego, CA, Feb. 2005, pp. 245–257.

[26] Z. Huang and W. Du, "OptRR: Optimizing randomized response schemes for privacy-preserving data mining," in *Proc. IEEE 24th Int. Conf. Data Engineering*, Cancun, Mexico, Apr. 2008, pp. 705–714.

[27] Z. Huang, W. Du, and B. Chen, "Deriving private information from randomized data," in *Proc. 2005 ACM SIGMOD Int. Conf. Management of Data*, Baltimore, MD, Jun. 2005, pp. 37–48.

[28] H. Kargupta, S. Datta, Q. Wang, and K. Sivakumar, "On the privacy preserving properties of random data perturbation techniques," in *Proc. 3rd IEEE Int. Conf. Data Mining*, Melbourne, FL, Nov. 2003, pp. 99–106.

[29] L. Lu, J. Han, L. Hu, Y. Liu, and L. M. Ni, "Dynamic key-updating: Privacy-preserving authentication for RFID systems," in *Proc. 5th Annu. IEEE Int. Conf. Pervasive Computing and Communications*, White Plains, NY, Mar. 2007, pp. 13–22.

[30] A. Meyerson and R. Williams, "On the complexity of optimal K-anonymity," in *Proc. 23rd ACM SIGMOD-SIGACT-SIGART Symp. Principles of Database Systems*, Paris, France, Jun. 2004, pp. 223–228.

[31] C. Sharp, S. Schaffert, A. Woo, N. Sastry, C. Karlof, S. Sastry, and D. Culler, "Design and implementation of a sensor network system for vehicle tracking and autonomous interception," in *Proc. 2nd European Workshop Wireless Sensor Networks*, Istanbul, Turkey, Jan. – Feb. 2005, pp. 93–107.

[32] N. Subramanian, K. Yang, W. Zhang, and D. Qiao, "ElliPS: A privacy preserving scheme for sensor data storage and query," in *Proc. 28th IEEE Int. Conf. Computer Communication*, Rio de Janeiro, Brazil, Apr. 2009, pp. 936–944.

[33] L. Sweeney, "Achieving k-anonymity privacy protection using generalization and suppression," *Int. J. Uncertainty, Fuzziness and Knowledge-Based Systems*, vol. 10, no. 5, pp. 571–588, Oct. 2002.

[34] S. L. Warner, "Randomized response: A survey technique for eliminating evasive answer bias," *J. Amer. Statistical Association*, vol. 60, no. 309, pp. 63–69, Mar. 1965.

[35] H. Xie, L. Kulik, and E. Tanin, "Privacy-aware collection of aggregate spatial data," *Data & Knowledge Engineering*, vol. 70, no. 6, pp. 576–595, Jun. 2011.