AN EVALUATION OF THE EFFECTIVENESS OF SMART METER DATA PERTURBATION MECHANISMS USING A UNIFIED STOCHASTIC FRAMEWORK

A Thesis by

Zoya Khan

Bachelor of Engineering, Pune University, 2011

Submitted to the Department of Electrical Engineering and Computer Science
and the faculty of the Graduate School of
Wichita State University
in partial fulfillment of
the requirements for the degree of
Master of Science

December 2015

AN EVALUATION OF THE EFFECTIVENESS OF SMART METER DATA PERTURBATION MECHANISMS USING A UNIFIED STOCHASTIC FRAMEWORK

The following faculty members have examined the final copy of this thesis for form and content, and recommend that it be accepted in partial fulfillment of the requirement for the degree of Master of Science with a major in Computer Networking.

_____

Murtuza Jadliwala, Committee Chair

_____

Kaushik Sinha, Committee Member

_____

Ehsan Salari, Committee Member

DEDICATION

To my loving parents, beautiful sister Ruba and immensely supportive husband Shifath

ACKNOWLEDGMENTS

I am truly grateful to my advisor, Dr. Murtuza Jadliwala, for his time, patience and guidance throughout the journey of my Master's thesis. Dr. Jadliwala motivated me to constantly improve my research and technical writing skills. I learnt a lot from him and I am sure this learning will help me in the future as well.

I would also like to thank my thesis committee members: Dr. Kaushik Sinha and Dr. Ehsan Salari, for their valuable feedback and always being willing to guide me at all stages of my research.

This thesis would never be completed without the help of my labmate Anindya Maiti. Anindya shared his experience as a Master's student with me and helped me through the hurdles I faced with my thesis presentation and documentation.

I would like to thank all my teachers and classmates who were a very essential part of my graduate school experience. A special thank you to all the wonderful friends I made in Wichita, who will remain an integral part of my life.

I would like to thank my family, especially my parents-in-law for always being very encouraging and supportive.

This work was partially supported by Power Systems Engineering Research Center (PSERC), under project S-54.

ABSTRACT

The growing energy needs are forcing utility companies towards the use of Advanced Metering Infrastructure (AMI). Smart meters are a part of the AMI, that capture fine grained electric consumption data from households and share it with the utility company, operations center, and other third party entities in the smart grid network, in order to make the grid more efficient and reliable. Yet, sharing such fine grained data gives rise to privacy concerns from the consumer's perspective. As a response to such privacy concerns, several smart meter data perturbation techniques have been proposed to fortify against extraction of sensitive personal information from the consumers' electric consumption data. However, due to the lack of a unified framework for assessing and comparing different perturbation techniques, consumers do not have access to a readily available tool to measure and evaluate the privacy provided by these perturbation techniques. We introduce a unified and practical data centric framework for evaluating and measuring the privacy offered by different smart meter data perturbation techniques. The framework trains multiple smart meter data models based on past smart meter data and other auxiliary information (e.g., time, temperature, location, etc.) that may be easily available to the adversary. The framework then evaluates the privacy offered by different perturbation techniques by carrying out reconstruction attacks (usually the adversary's first step before performing more advanced inference attacks) using trained models suitable for the characteristics of the perturbed data. Accuracy of reconstruction and the loss of privacy is measured in terms of well-known metrics, such as R-squared correlation and relative entropy. The framework is said to be unified because it considers all the elements required for evaluation such as the prior information available to the adversary, perturbation techniques, and the reconstruction strategy. To validate the effectiveness of our framework, for evaluating and measuring privacy offered by some of the popular perturbation techniques, we test the framework using real smart meter power consumption data collected from twenty-two households over a period of two-years.

TABLE OF CONTENTS

LIST OF FIGURES

Figure **Page**

LIST OF FIGURES (continued)

Figure **Page**

LIST OF ABBREVIATIONS

AMI Advanced Metering Infrastructure

SM Smart Meter

SGN Smart Grid Network

SDPM Smart Meter Data Perturbation Mechanisms

CHAPTER 1
INTRODUCTION

Smart Meters (SMs) are a part of the Advanced Metering Infrastructure (AMI), that capture fine grained electric consumption data from consumers and report it to the utility provider, operations center, and other third party entities in the Smart Grid Network (SGN). As many as 800 million SMs are expected to be installed globally by 2020 as a part of the SGN [10], and the information collected by these SMs will be beneficial for various operations such as real time monitoring, fault detection, time-of-use billing, load balancing, demand response, self healing [23][22] and peak shaving [35][30]. Despite the utility, collection of fine grained information gives rise to privacy concerns especially from the consumer's point-of-view. A common privacy threat is inference of private information about a consumer by eavesdropping on the data as it passes through the SGN. Another possible threat is the misuse of data by the utility company. The company could release, or even sell, power consumption information it has collected from its customers. In both cases, the information can be misused by advertising companies, nosy landlords, employers, or even stalkers. Various research efforts have attempted to study the private information that can be inferred from the fine-grained power consumption data [13][28]. Research shows that it is possible to estimate the number of residents in a household based on the frequency of power switches turned on and the number of appliances simultaneously in use [19]. One could infer information about the appliances used in a house by observing power consumption data, as shown in Figure 1.1 [25] or could even monitor the location of a resident inside the home based on the type of appliances being used [18]. By inferring information about appliance usage, the behavior of people living in the house can be inferred. For example, energy cycle of TV implies someone's presence at home, energy cycle of coffee pot tells the adversary when people wakeup, and energy cycles of water heater can be used to infer the number of occupants in a house. Sufficiently fine-

grained data even allows identifying the TV channel or movies being watched since television power consumption changes with the image being displayed on the television screen [33].

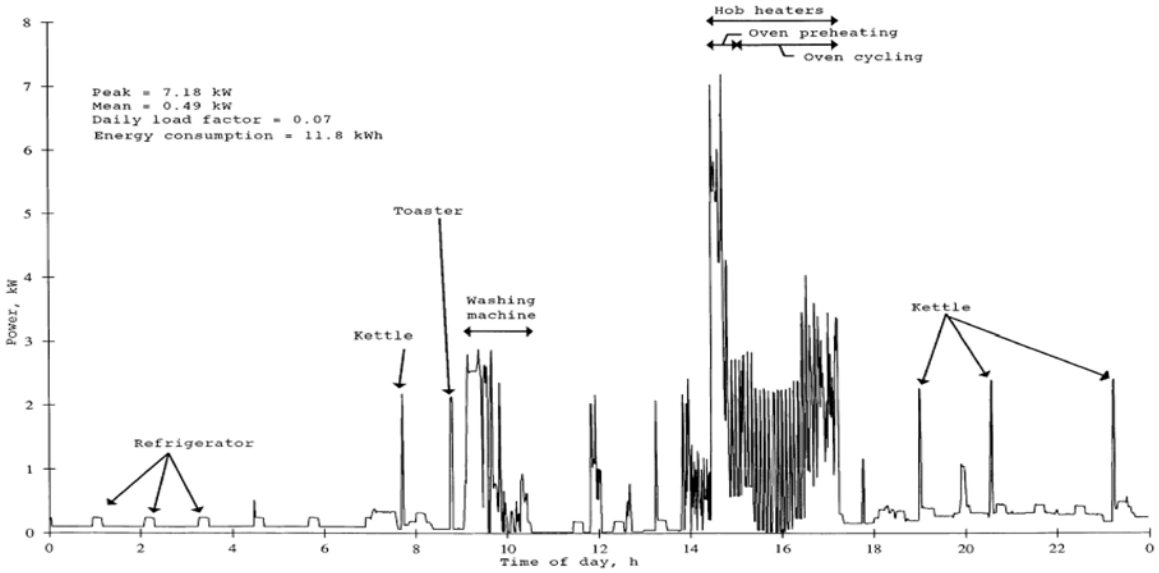

Figure 1.1: Mapping from power consumption to appliance usage [25]

To overcome these concerns, many privacy preserving mechanisms have been proposed, such as use of cryptographic techniques with homomorphic properties to aggregate data [9, 24, 19, 13, 17, 3], use of zero knowledge proofs and certified policies for private bill computation [26], distortion of SM data by addition of random noise [15] and use of large batteries [21, 4, 16] to "flatten" the output from SMs.

Although there have been several research efforts on designing Smart Meter Data Perturbation Mechanisms (SDPM), there has been very limited research in quantifying the privacy provided by such SDPMs. Sankar et al. [29] proposed a generic framework for modeling privacy and utility of released SM data. However, their model abstracts away specific perturbation mechanisms and inference techniques, and considers a highly capable adversary with individual appliance-level data access. They use Hidden Markov Models to represent smart meter data, where the hidden state is the appliance information. Taking cues from efforts to quantify privacy in various other domains, such as location-based services [31],

databases [12], anonymity protocols [5] and RFID [34], we apprehend the need of a unified and practical framework for measuring and evaluating the effectiveness of different SDPMs.

In this paper, we introduce a *unified framework* for measuring and evaluating the effectiveness of different SDPMs. The proposed framework is highly customizable as it does not abstract away the specifications of the SDPMs being evaluated, and can evaluate the SDPM under different adversarial capabilities, reconstruction strategies and effectiveness metrics. As we consider an adversary who can only observe the aggregated (and perturbed) power consumption of the target household, and has no information about individual appliances inside the household, we represent the SM power consumption values in our framework as states of a *n-order Markov Chain*. The framework trains multiple Markov models based on combinations of past smart meter data and additional contextual information (such as time, temperature, location, etc.) assumed to be available to the adversary. The framework then evaluates the privacy offered by existing perturbation techniques by carrying out *reconstruction attacks* using the trained Markov models (suitable for the characteristics of the perturbed data) by using them as a state sequence generator. Accuracy of reconstruction or privacy loss is measured in terms of well known *metrics*, such as R-squared correlation and relative entropy. We implement our framework as a modular software tool using python, which can be easily extended to add new SDPMs, reconstruction strategies, adversarial strengths and privacy metrics.

CHAPTER 2
BACKGROUND AND RELATED WORK

A series of power surges over a twelve-second period triggered a cascade of shutdowns in the US and Ontario on August 14, 2003. The result was the biggest blackout in North American history. 61800 megawatts of power was lost and over 50 million people we affected. Studies showed that the outage was because of lack of real-time monitoring and diagnosis [1].

A group of experts at the US Department of Energy (DOE) proposed the term smart grid for the goal of extending intelligence to parts of the electric grid. Title XIII of the energy independence and security act of 2007 mentions ten features of a smart grid [14]:

1. Increased use of digital information and controls technology to improve reliability, security, and efficiency of the electric grid.

2. Dynamic optimization of grid operations and resources, with full cyber security.

3. Deployment and integration of distributed resources and generation, including renewable resources.

4. Development and incorporation of demand response, demand-side resources, and energy-efficiency resources.

5. Deployment of smart technologies (real time, automated interactive, that optimize the physical operation of appliances and consumer devices) for metering, communications concerning grid operations and status, and distribution automation.

6. Integration of smart appliances and consumer devices.

7. Deployment and integration of advanced electricity storage and peak-shaving technologies, including plug-in electric and hybrid electrical vehicles, and thermal-storage air conditioning.

8. Provision to consumers of timely information and control options.

9. Development of standards for communication and interoperability of appliances and equipment connected to the electric grid, including the infrastructure serving the grid.

10. Identification and lowering of unreasonable or unnecessary barriers to adoption of smart grid technologies, practices, and services.

While there are many advantages of the SGN as described in [8] and the references therein, previous research literature such as [6] uncovers security and privacy vulnerabilities associated with the SGN. Cohen [6] studied the possible attacks on the SGN and analyzed existing solution strategies. Kalogridis et. al [16] identify the need for privacy of load signatures from households as they can be used to carry out inference attacks. They propose the use of batteries and alternative power sources to reduce the usefulness of load signatures from an intruder's perspective. They also propose three different privacy metrics: relative entropy, clustering classification, and a correlation/regression metric. In the same area of research, Sultanem [32] shows that local measurement of variations in active and reactive power is sufficient in most cases to identify an appliance. In SGN, data-oriented privacy is of greater interest, as compared to context-oriented privacy, because it deals with private consumer data. The security threats to the smart grid can target the confidentiality and the integrity of the gathered fine-grained user data. They can also threaten the availability of the power grid. It should go without saying that appropriate security and privacy-preserving technique are needed for large-scale deployment and consumer-acceptance of the SGN.

Sankar et al. [29] propose a framework for modeling privacy and utility of the released smart meter data. The authors model the privacy-utility trade-off by abstracting away specific perturbation mechanisms and inference techniques. The goal of our work is to quantify the privacy afforded by existing data perturbation techniques when they are applied to real smart meter data. We propose a model that considers the strength of the adversary such as the past data it has and the knowledge about the target user or household. Dong

et al. [11] quantify the utility-privacy trade-off in the smart grid and show how frequency of sampling affects performance of the smart grid operations. The goal of our work is not to compare the privacy-utility trade-off of data perturbation techniques (which can be done using standard utility measurement metrics). Our goal is to develop a framework, and its corresponding modular implementation, to enable consumers/utility companies to assess the effectiveness of the data perturbation schemes used (or envisioned to be used) by household SMs.

CHAPTER 3
SYSTEM AND ADVERSARY MODEL

3.1 System Model

An illustration of a typical smart grid communication network is shown in Figure 3.1. Smart meters are associated with customers who can be residential or commercial users of power from a service provider. Smart meters measure aggregate power consumption of all appliances used by the consumer and share it with the service provider or utility center, operations and markets. The data is sent from the smart meter to utility company using a public network such as the Internet. In order for this communication to be secure against basic forms of eavesdropping, protocols that offer a high level of data confidentiality and integrity, such as, Internet Protocol Security (IPSec), Secure Socket Layer (SSL), Transport Layer Security (TLS) and Secure Shell (SSH) are used. These protocols use standard cryptographic techniques to provide data confidentiality, integrity and authenticity between the smart meter and the utility provider. Our system model assumes that such protocols are already in place when the data is shared by the smart meter to the other entities in the smart grid network and we quantify the loss of privacy of the user, once the data is with its legitimate receiver.

3.2 Adversary Model

The adversary considered in our model is the legitimate recipient of the perturbed data sent by the smart meter. Such an adversary may be the utility company or other third party entities in the SGNs with whom the data is shared. The adversary is also able to obtain some contextual information about the household and its environment, e.g., temperature, geographic location, etc. We assume that the adversary has some past ground truth data from similar (or the same) users. Such an assumption is based on the possibility that sometimes data may be available from before a perturbation technique is applied or an

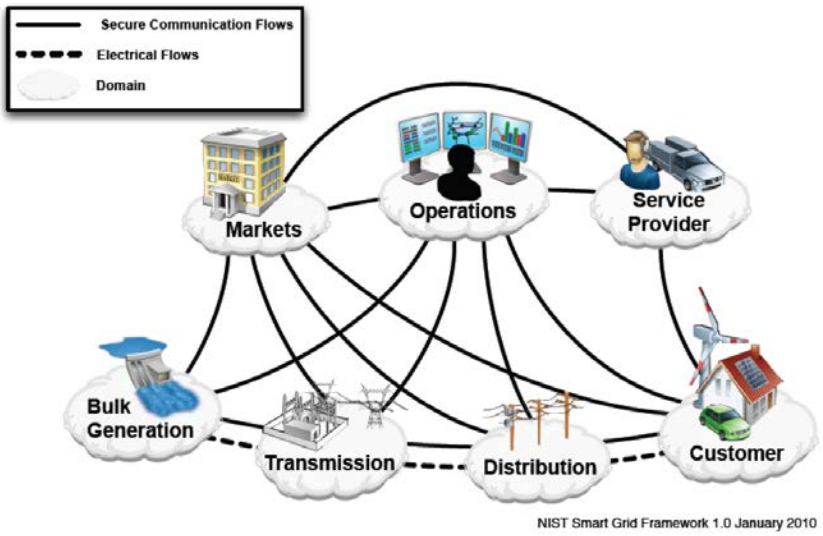

Figure 3.1: Illustration of the smart grid network [20]

adversary can have some homes with similar location, size and occupancy as that of its target user, from which it is able to collect data. The goal of the adversary is to infer fine-grained information about the target user or household from the power consumption data, archived from the SMs. One of the initial steps of the adversary prior to performing such inference attacks is to get back the original data from the perturbed data (i.e., reconstruction of the original data from the perturbed data obtained from the SMs). We refer to this process of getting back the original data from the perturbed data as a *reconstruction attack*.

CHAPTER 4
FRAMEWORK

Our proposed framework consists of a tuple of inseparable elements:

$$(\mathbb{D}, SDPM, \mathbb{O}, ADV, METRIC) \tag{4.1}$$

where $\mathbb{D}$ is the actual power consumption data of the target user or household, represented by a time ordered set of real numbers, SDPM stands for smart meter data perturbation mechanism, which is a mapping from a set of time ordered real numbers $\mathbb{D}$ to another time ordered set of real numbers $\mathbb{O}$, where $\mathbb{O}$ is the perturbed data of the target user shared by the smart meter. We refer to $\mathbb{O}$ as the observed power consumption data sequence as this is the data that the adversary can observe. The adversary ADV is an entity who has knowledge about the SDPM used, some past generated power consumption data from similar (or the same) users and some additional information about the customer or user with whom the smart meter data is associated. The goal of the adversary is to implement a reconstruction attack on the observed data $\mathbb{O}$, where a reconstruction attack, given $\mathbb{O}$, is a technique that is used to obtain a time ordered sequence of power consumption values $\mathbb{R}$, which has almost identical to $\mathbb{D}$, i.e., has identical information as contained in $\mathbb{D}$. The success of the adversary can be defined by quantifying the similarity between the reconstructed data $\mathbb{R}$ and the actual data $\mathbb{D}$. By similarity we mean that using the reconstructed data, same (or almost same) information about a user can be extracted, as would be possible using the actual data $\mathbb{D}$. The quantification of adversary's success and user's loss of privacy is captured by an evaluation metric METRIC.

In the following sections we describe all the entities of our framework and their inter-relationship.

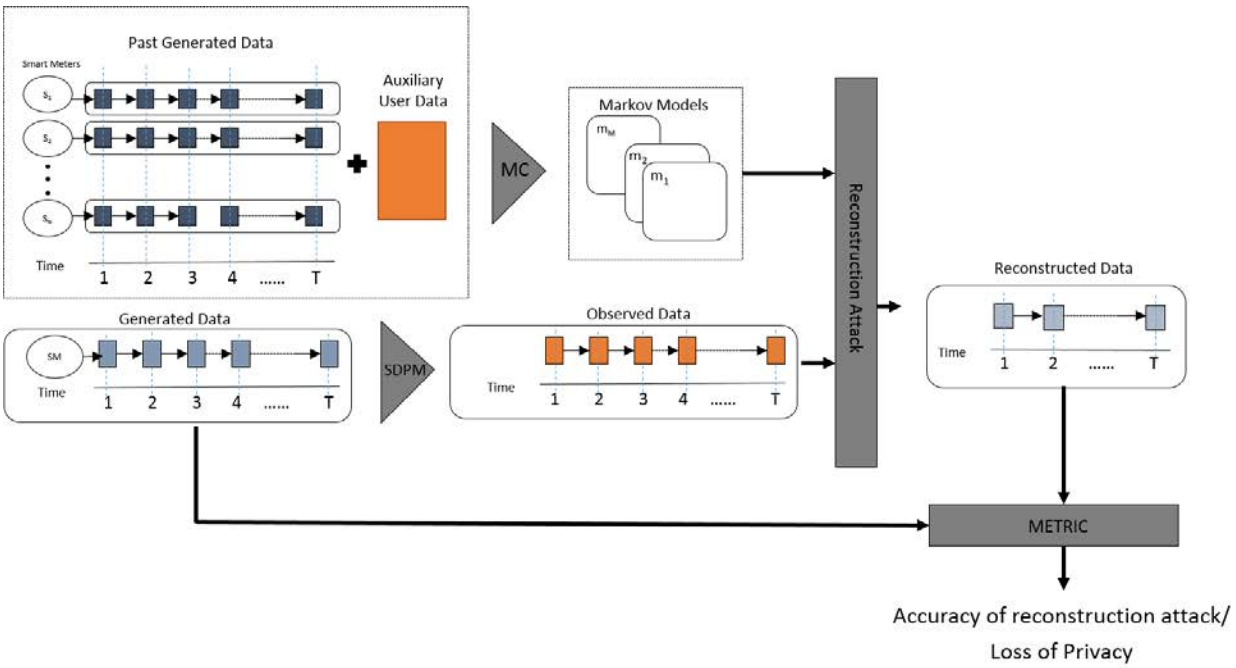

Figure 4.1: Elements of proposed smart grid privacy framework. The user's power consumption data is perturbed to preserve privacy. The adversary has some past power consumption data and some additional information about the user, which he uses to contruct different Markov models. We refer to the process of creating different models as model construction (MC).

4.1 Smart meter data as a n-order Markov chain

In our framework, we model the SM data and reconstruction attacks considering that the adversary can only see the observed data $\mathbb{O}$. Ideally, SM and its data could be represented as a Hidden Markov Model [29] where each state could be modeled as one of the combinations of appliances in use by the customer. Unfortunately, our adversary cannot construct such a model as he has no knowledge about the appliances the user has and cannot relate the observed data with the state transition of the appliances. Such an assumption makes our model more practical, as not all adversaries can have access to such fine-grained information about the user. As a consequence, we model smart meter data $\mathbb{D}$, as a discrete n-order Markov chain, where the present state of the system can be probabilistically described using n previous states. Each state in our model represents a power consumption value that lies

in a range that belongs to the set $\mathbb{L} = \{l_1, l_2, ., l_N\}$. Each element l_i of set $\mathbb{L}$ denotes a range of power values. Time is divided into discrete instants and the set of time instants under consideration is $\mathbb{T} = \{1, 2, .., T\}$. The granularity in terms of power levels and time depends on the precision desired by the user of the framework. Smart meters capture power consumption data at instants of time in $\mathbb{T}$. The sequence of actual power consumption values for a user is a T-size vector $\mathbb{D} = (d(1), d(2), .., d(T))$.

A Markov model is characterized as follows:

1. Let N be the finite and countable number of states in the system. We denote the set of states as $S = \{S_1, S_2, ..., S_N\}$ and the state at time t is represented by a random variable q_t. Each state is our model represents a region of power in $\mathbb{L}$.

2. The state transition probability distribution denoted by $A = \{a_{ij}\}$ where

$$a_{ij} = P[q_{t+1} = S_j | q_t = S_i] q_{t-3}, q_{t-2}, q_{t-1}. \tag{4.2}$$

   with the state transition coefficients having the following

$$a_{ij} \geq 0 \tag{4.3a}$$

$$\sum_{j=1}^{N} a_{ij} = 1 \tag{4.3b}$$

   since they obey standard stochastic constraints.

3. The initial state distribution $\pi = \{\pi_i\}$ where

$$\pi_i = P[q_1 = S_i], \quad 1 \leq i \leq N \tag{4.4}$$

Though Markov chains are memoryless by definition, Markov chains can be created to have multiple time-step memory. In first-order Markov chains, the current state depends

only on the previous state, but it is possible to create n-order Markov chain where current state depends on n previous states. For example in a third-order Markov chain ($n = 3$) the current state q_t would depend on the three previous states $q_{t-3}, q_{t-2}, q_{t-1}$. The equation below represents the state transition probability for a third-order Markov chain:

$$a_{i(j,k,l)} = P[q_t = S_i | q_{t-3} = S_j, q_{t-2} = S_k, q_{t-1} = S_l] \quad 1 \leq i, j, k, l \leq N \tag{4.5}$$

In our model we use n-order Markov chains hence there will be N^n previous states. Therefore, A will be a $N^n \times N^n$ matrix, where reach row represents a tuple of n previous states and each column represents a tuple of $n-1$ previous states and the current state.

The complete specification of a Markov model requires parameter N, two probability measures A, π and n. For convenience, we use the following compact notation to represent model parameters:

$$\lambda = \{N, A, \pi, n\} \tag{4.6}$$

to indicate complete parameter set of the model.

Calculation of A : In order to calculate A, we need to estimate the value of $a_{(k,l,\ldots,m)j}, 1 \leq j, k, l, \ldots m \leq N$.

$$\hat{a}_{(k,l,\ldots,m)j} = \frac{\sum_{t=1}^{T} 1\{q_t = S_j, q_{t-n} = S_k, q_{t-n+1} = S_l, \ldots, q_{t-1} = S_m\}}{\sum_{t=1}^{T} 1\{q_{t-n} = S_k, q_{t-n+1} = S_l, \ldots, q_{t-1} = S_m\}} \tag{4.7}$$

where, $1\{\}$ represents an identity function.

Calculation of π: Let us consider that we are using V training sequences. The initial probability is the ratio of the number of times the initial n states were $S_k, S_l, ...$, where $1 \leq k, l, m \ldots \leq N$, to the total number of training sequences. As estimation of the initial probability is represented by $\hat{\pi}_{(k,l,\ldots,m)}$ and calculated as follows:

$$\hat{\pi}_{(k,l,\ldots,m)} = \frac{\sum_{v=1}^{V} 1\{q_1 = S_k, q_2 = S_l, \ldots, q_n = S_m\}}{V} \tag{4.8}$$

Given the definition of a Markov model for smart meter data as above, we define a set Markov models for the smart meter under consideration. One of the way to define these models is by considering different combinations of time durations and climate conditions determined by using average temperature. The set of all such Markov models is defined as $Models = \{m_1, m_2, \ldots, m_M\}$. For each model m_i, generated for a particular combination of time and climate information associated with the target user (or geographical location of the target household), we define the state transition probability matrix as A and initial probability distribution π, as discussed before.

4.2 Smart Meter Data Perturbation Techniques

A perturbation function f in a SDPM, is a function that maps the actual data $\mathbb{D}$ to the observed data $\mathbb{O}$.

$$f: \mathbb{D} \mapsto \mathbb{O} \tag{4.9}$$

The goal of perturbation is to preserve the privacy of users by not sharing very fine grained power consumption data. There are several perturbation techniques such as data hiding using cryptographic techniques, random perturbation, flattening of power consumption data using batteries and down sampling. Down sampling is one of the commonly used perturbation techniques, mainly because the perturbed data contains samples from the actual power consumption without any modifications and such information is useful to the utility company. For the purpose of evaluation we consider down sampling as a perturbation technique, but our framework is not restricted to down sampling. Any perturbation technique can be incorporated into our framework.

If K is the down sampling factor then from the actual data $\mathbb{D}$ every K^{th} sample is transmitted as a part of the observed data $\mathbb{O}$, all other samples of the observed data are null values. We consider the following downsampling techniques:

- Uniform Down Sampling: In this down sampling technique, the down sampling factor K is a fixed value that remains constant for the entire duration of time under consideration, $\mathbb{T}$.

- Random Down Sampling: This is a perturbation technique in which the down sampling factor K is not constant throughout the duration under consideration and is calculated before sending each perturbed data sample. K is random number generated from a uniform probability distribution having range $(1, \mathbb{T} - t)$.

- Probabilistic Down Sampling: In this technique, a Gaussian function with mean and standard deviation represented by μ and σ respectively is used. Based on a threshold applied to values sampled from the Gaussian distribution, it is determined whether to send a power consumption value in the observed data $\mathbb{O}$ at a particular instance of time. Instead of a Gaussian function any other probabilistic distribution can also be used.

4.3 Adversary

In order to evaluate an SDPM accurately, it is important to model the adversary ADV against whom the protection is placed. The strength of an adversary is defined by his knowledge about the target user and the attacks he performs in order to reconstruct the user's original smart meter data sequence from the observed data sequence. We assume that the adversary has the data shared by the smart meters i.e. $\mathbb{O}$. Additionally, we assume that the adversary knows the details of the SDPM, used to perturb the actual data. The adversary may have some past generated actual data from the smart meters. An example of such a situation is; a utility company starts implementing a perturbation technique recently but it has

the unperturbed values of power consumption from a user (or household), before it started applying the perturbation techniques. It is important to point that the adversary is just using its knowledge about the user and observed data because he does not know about the appliances present in a household and hence cannot model the data based on the power consumption information of the appliances. The adversary may also have additional contextual information about the target user or household, which it could use for reconstruction.

The adversary may be interested in reconstructing the original power values only for a certain duration of time in day, a particular day of the week or some length of days. There can be a separate Markov model for each day of the week or for each part of the day such as morning, afternoon, evening and night, since power consumption of a user may follow a repetitive pattern. Based on his intentions the adversary will use past generated power consumption values and appropriate Markov models to reconstruct the actual data from the SM.

4.4 Reconstruction Technique

We recall from section 4.3 that the goal of the adversary is to reconstruct the original data sequence for a period of interest. For such a reconstruction, the adversary can use the appropriate Markov model (constructed earlier by him) as a state sequence generator. In order to generate a state sequence of length T, the following steps need to be done;

1. If no perturbed data sample exists for the fist n samples, choose initial state sequence $\{q_1, q_2, ...q_n\}$ according to the initial state distribution π. Else, choose from the highest probability distribution among the initial state sequences, that best fit the available perturbed data.

2. Set $t = n + 1$

3. If perturbed data exists at time t, q_t will be the state associated with perturbed value i.e O(t). Else, transit to a new state $q_t = S_j$ according to the state transition probability matrix A.

4. Set $t = t + 1$; return to step 3 if $t < T$; otherwise terminate the procedure.

Let us denote the reconstructed power consumption for a user as a vector of size T, represented as $\mathbb{R} = (r(1), r(2), .., r(T))$. Once the state sequence is known, the reconstructed data $\mathbb{R}$ at any instance of time $t \in T$, denoted by r(t), is the power value associated with state q_t.

This technique is chosen for our evaluation because it works well with down sampling, our perturbation technique of choice. This is because in down sampling the perturbed data contains samples from the actual data itself and such data can be directly used in the reconstructed data sequence. For other perturbation techniques where the actual data may be unavailable in the perturbed data, the actual value corresponding to instances for which the perturbed data is available will also need to be predicted. Readers should note that our framework is flexible enough to easily incorporate other, more complex, reconstruction or prediction techniques as well.

4.5 Metric

As reconstruction is generally the first step to advanced inference attacks, the success with which the original data can be recovered from the perturbed power consumption values determines the loss of privacy of the consumer, due to the perturbation technique. Even though the original data $\mathbb{D}$ and the reconstructed data $\mathbb{R}$ consist of discretized values, their correlation still measures the "closeness" or "similarity" of one of the datasets with the other, which is a good indication of the success of the adversary, and thus the loss of privacy to the consumer. Our goal is not to measure the exact difference between the original data and the reconstructed data, but to evaluate a perturbation technique in terms of loss of privacy to the consumer. We use the R-squared correlation coefficient to measure the linear

dependence between original data $\mathbb{D}$ and reconstructed data $\mathbb{R}$. It gives a value between 0 to 1 inclusive, where 1 is total correlation and 0 is no correlation. Higher correlation signifies greater probability that the two data sets comes from the same distribution, which could enable an adversary to draw similar conclusions from both data sets.

Alternatively, there could also be another measure of privacy when an adversary is able to identify a power consumption event within the home using the reconstructed data. This identification can be done by monitoring the change in power consumption values between successive power measurements. We represent the sequence of such change in power consumption values in original data $\mathbb{D}$ by P_D, where each element of the set of values in P_D is, $p_D(t) = d(t) - d(t-1)$. Similarly for the reconstructed data $\mathbb{R}$, the sequence of values representing the difference in power between successive values is denoted by P_R. Each element of this set $p_R(t)$, is calculated as $p_R(t) = r(t) - r(t-1)$. In order to measure loss of privacy in this case, we use relative entropy which is a well known information theoretic measure for quantifying the relation between two probability distribution functions (pdfs). Relative entropy of two identifical pdfs is zero and relative entropy is always positive. The higher value of relative entropy means lower loss of privacy. We assume that P_D and P_R can be modeled as stochastic processes with pdfs, ${P'}_D$ and ${P'}_R$ respectively. The relative entropy $D({P'}_D \| {P'}_R)$ is defined as follows: [7]:

$$D({P'}_D \| {P'}_R) = \sum_{t_{min}}^{t_{max}} {P'}_D(t) log \frac{{P'}_D(t)}{{P'}_R(t)} \tag{4.10}$$

We are using the two metrics discussed above, for our evaluation but the framework can be extended to support other metrics as well.

CHAPTER 5
EVALUATION

5.1 Software Tool

We developed a modular python-based, software tool for the proposed framework to evaluate the efficiency of smart meter data perturbation techniques. In this tool, users can input time-series power consumption data from a household for a time period of interest and select from predefined perturbation techniques or they can even add perturbation techniques of their choice to the tool. The efficiency of the various perturbation techniques is evaluated by employing statistical reconstruction techniques, and their effectiveness is quantified using an appropriate metric, that captures the accuracy of the reconstruction attack. Our tool has the flexibility to specify any additional contextual information that the adversary may have regarding the SM, household or conditions near the household. For simplicity, we currently restrict this additional information available to the adversary to the external temperature information of the geographical location of the smart meters. The goal of having such a software tool is to enable ease of comparison and evaluation of various smart meter data perturbation techniques using a unified and practical framework, such as the one proposed in this work.

5.2 Experimental Set-up

We conduct an empirical evaluation of perturbation techniques using real power consumption data from twenty-two households in Midlands, UK [27]. In this dataset, we have over 2 years (from 2008 and 2009) of power consumption data, recorded every minute. We also have the weather information in the Midland region of UK for the years of 2008 and 2009 [2]. The evaluation results presented below have been collected by executing our software tool on a traditional personal computer comprising of a 2.3 GHz processor and 8GB memory.

5.3 Results

The following initial evaluation is conducted to validate the performance of the proposed framework and the corresponding software tool by considering a small-sized data set and limited computation resources. However, in order to perform a large scale analysis, an adversary could take advantage of this framework by executing this tool over a much larger dataset and by using computational systems with higher resources.

5.3.1 Standard Deviation of Actual Data

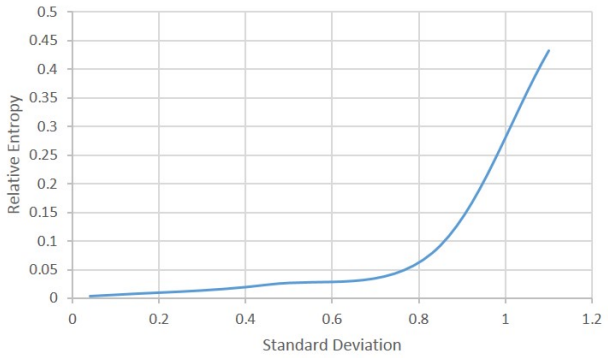

Figure 5.1: Variation in reconstruction accuracy in terms of relative entropy of actual data with respect to the reconstructed data, as the standard deviation of the actual data changes, when n = 3 and N = 8. The perturbation techniques applied is down sampling with a factor of 5 i.e. 12 samples/hour.

Table 5.1: Variation of reconstruction accuracy in terms of relative entropy of actual data with respect to the reconstructed data for different values of standard deviation of actual data. The same training data with standard deviation of 0.59 was used for all measurements.

	Order of Markov chains				
Std. Dev	1	2	3	4	5
1.1	0.157	0.433	0.433	0.0.433	0.433
0.83	0.104	0.104	0.525	0.078	0.078
0.45	0.023	0.023	0.023	0.023	0.023
0.04	0.003	0.003	0.003	0.003	0.003

The success of reconstruction and loss of user's privacy is measured in terms of R-squared correlation and relative entropy as discussed in section 4.5. For all the measurements in this

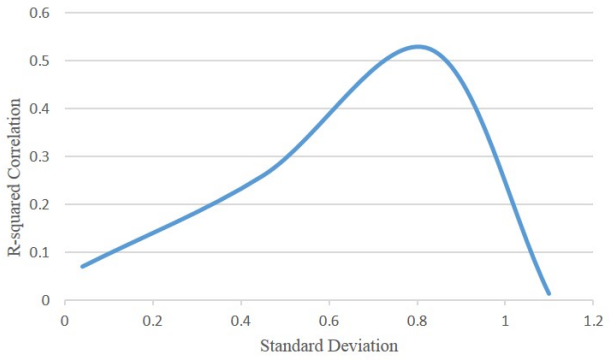

Figure 5.2: Reconstruction accuracy in terms of R-squared correlation of actual data with respect to the reconstructed data when n = 3 (third order Markov chains) and N=8 i.e. number of states in the model are 8. The perturbation techniques applied is down sampling with a factor of 5 i.e. 12 samples/hour.

Table 5.2: Variation of reconstruction accuracy in terms of R-squared correlation, with respect to standard deviation values of actual data. The same training data with standard deviation of 0.59 was used for all measurements.

	Order of Markov chains				
Std. Dev	1	2	3	4	5
1.1	0.373	0.014	0.014	0.070	0.070
0.83	0.547	0.467	0.525	0.525	0.525
0.45	0.261	0.261	0.261	0.261	0.261
0.04	0.071	0.071	0.071	0.071	0.071

subsection the training data was constant and had a standard deviation of 0.59. We used one month of training data to calculate the transition matrix and initial probability distribution and carry out reconstruction attacks on actual data $\mathbb{D}$. The adversary has knowledge about the average temperature on the day, for which the reconstruction is carried out and the temperature of the geographical location during the time the training data was collected. The plot for relative entropy of actual data with respect to the reconstructed data, seen in Figure 5.1 shows that, as the standard deviation of actual data increases, the value of relative entropy also increases, which signifies that as the variance in the actual data increases it is more difficult to carry out a reconstruction attack successfully. In the same figure, from the trend seen for R-squared correlation we observe that when the actual data $\mathbb{D}$ is similar to

the training data, in terms of standard deviation, the correlation between the actual and reconstructed data is higher. As the difference between standard deviation of the actual data and the training data increases, the reconstructed data is less correlated with respect to the actual data.

Graph 5.2 is plotted for $n = 3$ i.e. a third order Markov chain. The standard deviation of the actual data is varied and the accuracy of reconstruction is measured in terms of R-squared correlation. The trends in the Figure 5.2, show that, as the variance of the actual data differs from the training data, the correlation between the actual data and reconstructed data decreases. When the standard deviation of the actual data differs from the standard deviation training data by a considerable amount, the accuracy of reconstruction in terms of correlation is observed to be reduced. This is expected because if the training data does not have many power consumption values changes, i.e. it has low standard deviation and the actual data has many fluctuations, the model will not be able to accurately predict the next state for reconstruction. Similarly, if the training data has very high standard deviation and the actual data is comparatively flat, the model would have over learned and will inaccurately predict more fluctuations in power consumption values. Similar results for different values of order of Markov chains can be concluded from table 5.2.

By observing the results in Figure 5.1 and table 5.1, we see that when the variance of the actual data is very low, the accuracy of reconstruction in terms of relative entropy is high irrespective of the order of Markov chains. This is because relative entropy measurements capture the accuracy of reconstructing power consumption changes and when the standard deviation of actual data is low, there will be fewer changes in power consumption values. As the standard deviation of the actual data increases irrespective of the standard deviation of the training data, the accuracy of reconstruction in terms of relative entropy decreases. This is expected as higher standard deviation means significant changes in the power consumption values (or patterns) and the model may not have learnt of all such changes from the training data.

5.3.2 Order of Markov Chains and Number of States

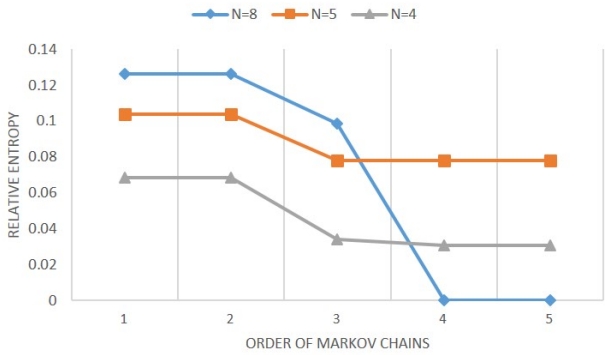

Figure 5.3: Reconstruction accuracy in terms of relative entropy when data is perturbed with a factor of 5 i.e. 12 samples/hour.

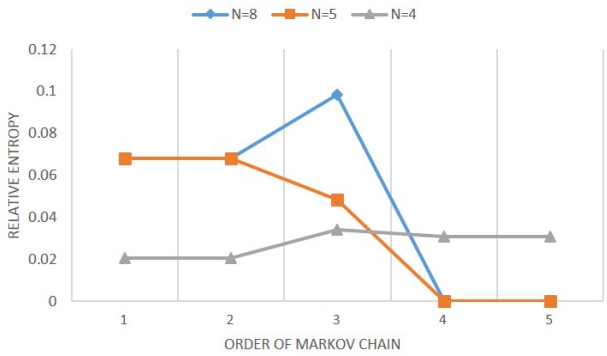

Figure 5.4: Reconstruction accuracy in terms of relative entropy when data is perturbed with a factor of 10 i.e. 6 samples/hour.

In this section, we discuss the significance of the order of Markov chains and the number of states in the model, on the reconstruction accuracy. Figures 5.3, 5.4 and 5.5 show the accuracy of reconstruction and loss of privacy in terms of relative entropy, for perturbation using down sampling factors of 5, 10 and 15, respectively. These plots represent a reconstruction duration of one hour using a training data of one month and relevant temperature information. It can be observed from the plots that as the order of Markov chains increases the relative entropy of the actual data with respect to the reconstructed data decreases, which signifies that the accuracy of reconstruction increases as the Markov chain

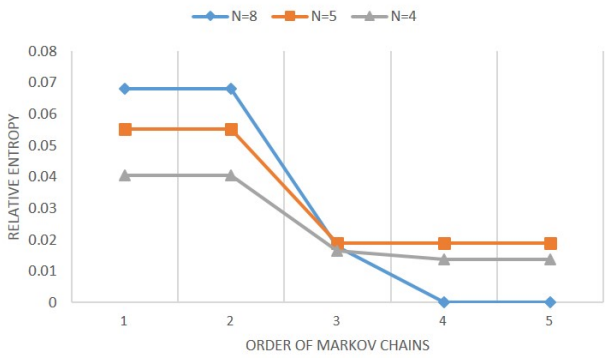

Figure 5.5: Reconstruction accuracy in terms of relative entropy when data is perturbed with a factor of 15 i.e. 4 samples/hour.

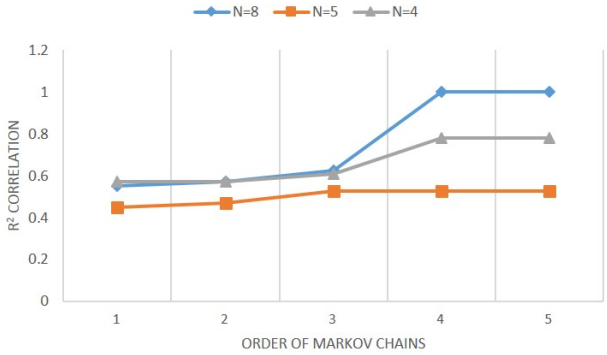

Figure 5.6: This figure shows accuracy of reconstruction in terms of R-squared correlation when data is perturbed with a down sampling factor of 5 i.e. 12 samples/hour.

order increases. Each legend in these graphs denotes results for a unique number of states. Based on these observations, we conclude that the reconstruction using higher order Markov chains is most accurate. This is intuitively expected since higher order Markov Chains are constructed using information of the n previous states, where n represents the order of the Markov chains and having more information increases the accuracy of prediction. Also, this would help capture and reconstruct more accurately, repetitive sets of power consumption values.

Figures 5.6, 5.7 and 5.8 represent accuracy of reconstruction in terms of R-squared correlation for data perturbed with sampling factors of 5, 10 and 15 respectively. It can be

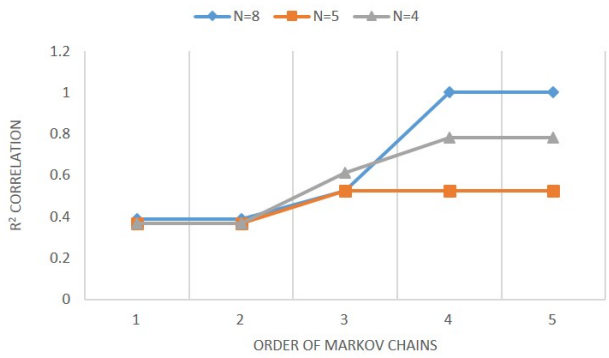

Figure 5.7: Reconstruction accuracy in terms of R-squared correlation when data is perturbed with a down sampling factor of 10 i.e. 6 samples/hour.

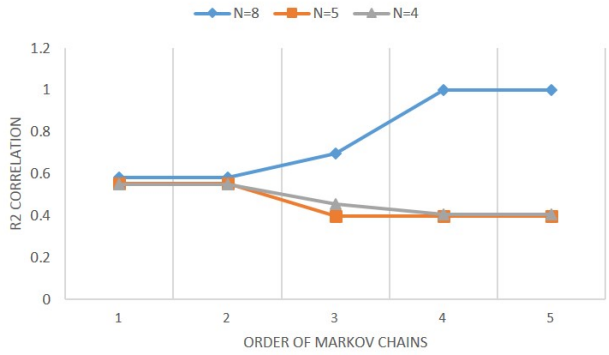

Figure 5.8: Reconstruction accuracy in terms of R-squared correlation when data is perturbed with a down sampling factor of 15 i.e. 4 samples/hour.

observed that, as the order of Markov chains increases, the correlation between the actual data and the reconstructed data increases for sampling factors of 5 and 10. For downsampling factor of 15, which means the perturbed data contains very few samples from the actual data, we see that the accuracy of reconstruction peaks at a certain order of Markov chains and then drops as we increase the order of Markov chains further, except for when we use a larger number of states in our Markov models i.e. $N = 8$. Based on all the observations in this section we can conclude that, as the order of Markov chains and the number of states in the model increases, in most cases, the accuracy of reconstruction increases. We have also identified special cases where this observation is not applicable.Considering all the results

presented above, we conclude that down sampling by itself does not provide privacy to the users and using very minimal computing power and training data, the actual data can be reconstructed by the adversary.

5.3.3 Random Perturbation

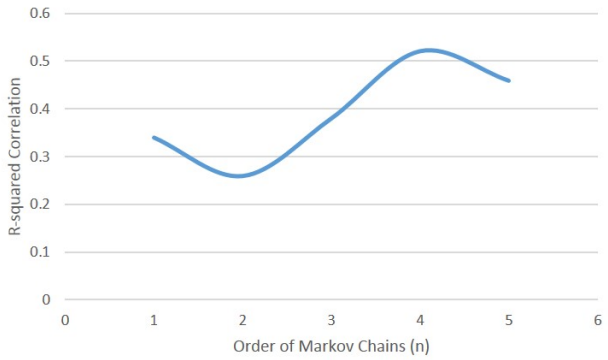

Figure 5.9: Reconstruction accuracy measured in terms of R-squared correlation when data was perturbed using a random sampling rate.

Figure 5.10: Reconstruction accuracy measured in terms of relative entropy when data was perturbed using a random sampling rate.

In the above sections, all the results were calculated considering down sampling as the perturbation technique. In this section we evaluate our model for random perturbation techniques.

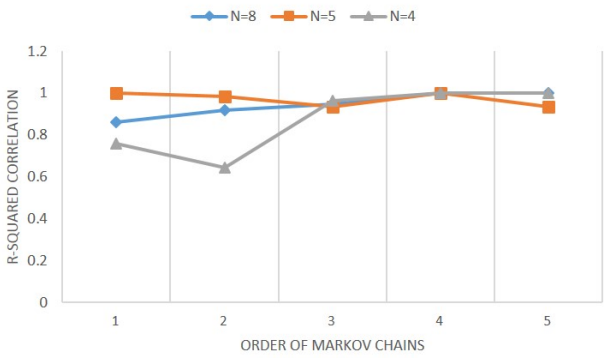

Figure 5.11: Reconstruction accuracy measured in terms of R-squared correlation when data was perturbed using a random samples from a Gaussian distribution with mean of 1, standard deviation of 0.5 and threshold of 0.75 to decide if data should be sent in a particular instant or not.

Figures 5.9 and 5.10 represent reconstruction accuracy, in terms of R-squared correlation and relative entropy, respectively, when data is down sampled at a random rate. We observe that, the results for accuracy of reconstruction do not follow any specific trend. Such a behavior is expected because the quantity of actual data samples present in the perturbed data varies randomly in such a technique and accuracy of reconstruction is affected by the occurrence of actual data samples in the perturbed data. The accuracy of reconstruction using our model is 0.09 in terms of relative entropy which signifies that power events (change in power consumption value between two consecutive instants of time) is captured with a good accuracy but the average correlation of the reconstructed data with respect to the actual data is low (0.309).

Similarly, when data is perturbed using random samples from a Gaussian distribution (for the purpose of evaluation we used a Gaussian distribution with mean 1 and variance 0.5, with a threshold of 0.75) to decide whether to send a value of power consumption data or not in a particular instant, we notice that the results do not follow any specific pattern. The accuracy of reconstruction varies randomly as we increase the number of states and order of Markov chains, but the model still reconstructs the actual data with very high accuracy.

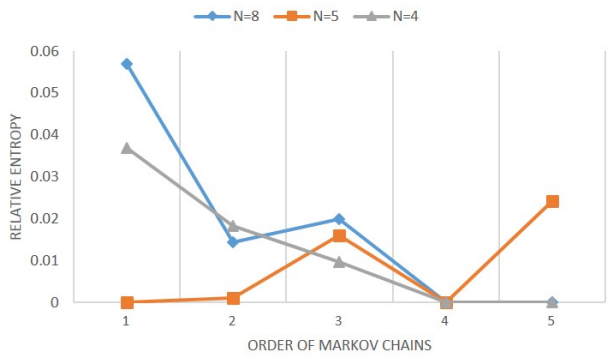

Figure 5.12: Reconstruction accuracy measured in terms of relative entropy when data was perturbed using a random samples from a Gaussian distribution with mean of 1, standard deviation of 0.5 and threshold of 0.75 to decide if data should be sent in a particular instant or not.

Especially for higher order Markov chains the accuracy captured in terms of relative entropy and R-squared correlation is high.

5.3.4 Performance

Figure 5.13: Time of reconstruction with respect to order of Markov chains.

In this section, we analyze the time taken for reconstruction with respect to the order of Markov chains. We measure the time required to generate the transition matrix A, initial probability distribution π and carry out the reconstruction attack. Such measurements were recorded by varying model parameters such as the number of states (N), variance of actual data and the quantity of training data. The results shown in Figure 5.13 represent an average

of 10 such measurements. We observe from the results of time measurements with respect to the order of Markov chains, as shown in Figure 5.13, that as the order of Markov chains increases the total reconstruction time also increases. This is because the number of rows in a transition matrix depends exponentially on the order of Markov chains, as discussed earlier in section 4.1.

CHAPTER 6
DISCUSSION AND FUTURE WORK

We verified the functionality of our proposed framework using real smart meter data. For this purpose we picked few well known perturbation techniques, metrics and reconstruction strategy. It is seen that the model performs well when the standard deviation of training data is similar to that of the actual data. The framework is not recommended for use with random perturbation techniques based on the results observed. We don't anticipate these techniques becoming very popular because the utility of the data after applying these techniques cannot be controlled and is pretty low. Because of reasonable accuracy achieved using lower order Markov chains and the complexity associated with implementing and executing higher order Markov chains, we do not evaluate the proposed framework for Markov chains higher than the 5^{th} order. For the purpose of evaluating the model we considered few sample scenarios but in the future more evaluations can be carried out for additional perturbation strategies, metrics and availability of other auxiliary information.

CHAPTER 7
CONCLUSION

The deployment of the future power grid or SGN is slow due to privacy concerns from the consumers' perspective. Several privacy preserving data perturbation techniques have been proposed to protect the user's private information. In order to give the users the flexibility to analyze the privacy offered to them by the perturbation technique used on their power consumption data, we propose a unified framework for evaluating the efficiency of smart meter data perturbation techniques. Our framework considers not only the perturbation technique, but also the strength and knowledge of the adversary. The results validate that the framework can be used for evaluation of different data perturbation techniques as applied to real smart meter data. The order of Markov chain required for a good reconstruction can also be determined using the proposed framework. This information will be helpful in analyzing the amount of computing power, training data and perturbed data the adversary would need for an accurate reconstruction. Such analysis can allow users' measure the privacy offered to them by the perturbation technique used to send their data to the utility company.

BIBLIOGRAPHY

BIBLIOGRAPHY

[1] IESO, Blackout 2003. http://www.ieso.ca/imoweb/EmergencyPrep/black-out2003.

[2] WeatherOnline. http://www.weatheronline.co.uk, 2014.

[3] Ács, Gergely, and Castelluccia, Claude. I have a dream!(differentially private smart metering). In *Information Hiding* (2011), Springer, pp. 118–132.

[4] Acs, Gergely, Castelluccia, Claude, and Lecat, William. Protecting against physical resource monitoring. In *Proceedings of the 10th annual ACM workshop on Privacy in the electronic society* (2011), ACM, pp. 23–32.

[5] Chatzikokolakis, Konstantinos, Palamidessi, Catuscia, and Panangaden, Prakash. Anonymity protocols as noisy channels. In *Proceedings of the 2Nd International Conference on Trustworthy Global Computing* (Berlin, Heidelberg, 2007), TGC'06, Springer-Verlag, pp. 281–300.

[6] Cohen, F. The smarter grid. *Security Privacy, IEEE 8*, 1 (Jan 2010), 60–63.

[7] Cover, Thomas M., and Thomas, Joy A. *Elements of Information Theory*. Wiley-Interscience, New York, NY, USA, 1991.

[8] Deconinck, G., and Decroix, B. Smart metering tariff schemes combined with distributed energy resources. In *Critical Infrastructures, 2009. CRIS 2009. Fourth International Conference on* (March 2009), pp. 1–8.

[9] Defend, Benessa, and Kursawe, Klaus. Implementation of privacy-friendly aggregation for the smart grid. In *Proceedings of the First ACM Workshop on Smart Energy Grid Security* (New York, NY, USA, 2013), SEGS '13, ACM, pp. 65–74.

[10] Digital, Telefonica. The smart meter revolution - towards a smarter future, 2014.

[11] Dong, Roy, Cárdenas, Alvaro A., Ratliff, Lillian J., Ohlsson, Henrik, and Sastry, S. Shankar. Quantifying the utility-privacy tradeoff in the smart grid. *CoRR abs/1406.2568* (2014).

[12] Dwork, Cynthia. Differential privacy. In *Automata, languages and programming*. Springer, 2006, pp. 1–12.

[13] Erkin, Z., Troncoso-Pastoriza, J.R., Lagendijk, R.L., and Perez-Gonzalez, F. Privacy-preserving data aggregation in smart metering systems: an overview. *Signal Processing Magazine, IEEE 30*, 2 (March 2013), 75–86.

[14] Gabriel, Mark A. *Visions for a Sustainable Energy Future*. Fairmount Press, 2008.

[15] He, Xingze, Zhang, Xinwen, and Kuo, C.-C.J. A distortion-based approach to privacy-preserving metering in smart grids. *Access, IEEE 1* (2013), 67–78.

[16] Kalogridis, G., Efthymiou, C., Denic, S.Z., Lewis, T.A., and Cepeda, R. Privacy for smart meters: Towards undetectable appliance load signatures. In *Smart Grid Communications (SmartGridComm), 2010 First IEEE International Conference on* (Oct 2010), pp. 232–237.

[17] Li, Fenjun, Luo, Bo, and Liu, Peng. Secure information aggregation for smart grids using homomorphic encryption. In *Smart Grid Communications (SmartGridComm), 2010 First IEEE International Conference on* (2010), IEEE, pp. 327–332.

[18] Lisovich, M.A., Mulligan, D.K., and Wicker, S.B. Inferring personal information from demand-response systems. *Security Privacy, IEEE 8*, 1 (Jan 2010), 11–20.

[19] Lisovich, Mikhail, and Wicker, Stephen. Privacy concerns in upcoming residential and commercial demand-response systems. In *2008 Clemson University Power Systems Conference* (March 2008), Clemson University.

[20] Locke, G., and Gallagher, P. D. Nist framework and roadmap for smart grid interoperability standards, release 1.0. Tech. rep., Nat. Inst. Standards Technol.,MD, USA,, 2010.

[21] McLaughlin, Stephen, McDaniel, Patrick, and Aiello, William. Protecting consumer privacy from electric load monitoring. In *Proceedings of the 18th ACM conference on Computer and communications security* (2011), ACM, pp. 87–98.

[22] Moslehi, K., and Kumar, R. A reliability perspective of the smart grid. *Smart Grid, IEEE Transactions on 1*, 1 (June 2010), 57–64.

[23] Niyato, D., Wang, Ping, and Hossain, E. Reliability analysis and redundancy design of smart grid wireless communications system for demand side management. *Wireless Communications, IEEE 19*, 3 (June 2012), 38–46.

[24] Paillier, Pascal. Public-key cryptosystems based on composite degree residuosity classes. In *Advances in cryptologyEUROCRYPT99* (1999), Springer, pp. 223–238.

[25] Quinn, E. Smart metering and privacy: Existing laws and competing policies. Tech. Rep. SSRN Rep. 1462285, Univ. Colorado Law SchoolCEES, Boulder, CO, USA, 2009.

[26] Rial, Alfredo, and Danezis, George. Privacy-preserving smart metering. In *Proceedings of the 10th Annual ACM Workshop on Privacy in the Electronic Society* (New York, NY, USA, 2011), WPES '11, ACM, pp. 49–60.

[27] Richardson, I., and Thomson, M. One-minute resolution domestic electricity use data, 2008-2009. colchester, essex: Uk data archive, october 2010. sn: 6583.

[28] Rouf, Ishtiaq, Mustafa, Hossen, Xu, Miao, Xu, Wenyuan, Miller, Rob, and Gruteser, Marco. Neighborhood watch: Security and privacy analysis of automatic meter reading systems. In *Proceedings of the 2012 ACM Conference on Computer and Communications Security* (2012), CCS '12, pp. 462–473.

[29] Sankar, L., Rajagopalan, S.R., Mohajer, S., and Poor, H.V. Smart meter privacy: A theoretical framework. *Smart Grid, IEEE Transactions on 4*, 2 (June 2013), 837–846.

[30] Schuler, R.E. Electricity markets, reliability and the environment: Smartening-up the grid. In *System Sciences (HICSS), 2010 43rd Hawaii International Conference on* (Jan 2010), pp. 1–7.

[31] Shokri, R., Theodorakopoulos, G., Le Boudec, J.-Y., and Hubaux, J.-P. Quantifying location privacy. In *Security and Privacy (SP), 2011 IEEE Symposium on* (May 2011), pp. 247–262.

[32] Sultanem, F. Using appliance signatures for monitoring residential loads at meter panel level. *Power Delivery, IEEE Transactions on 6*, 4 (Oct 1991), 1380–1385.

[33] Ulrich Greveler, Benjamin Justus, and Lhr, Dennis. Multimedia content identification through smart meter power usage profiles. In *Proceedings of the International Conference on Information and Knowledge Engineering* (January 2012), IKE '12.

[34] Vaudenay, Serge. On privacy models for rfid. In *Proceedings of the Advances in Crypotology 13th International Conference on Theory and Application of Cryptology and Information Security* (Berlin, Heidelberg, 2007), ASIACRYPT'07, Springer-Verlag, pp. 68–87.

[35] Yu, F.R., Zhang, Peng, Xiao, Weidong, and Choudhury, P. Communication systems for grid integration of renewable energy resources. *Network, IEEE 25*, 5 (September 2011), 22–29.