Towards a Platform for Testing and Developing Privacy-Preserving Data Mining Applications for Smart Grids

Andrew Koster[1], Gabriel de Oliveira Ramos[1],
Ana L.C. Bazzan[1], and Fernando Koch[2]

[1] Institute of Informatics, Federal University of Rio Grande do Sul (UFRGS)
[2] IBM Research Brazil

Abstract. In this paper we analyse the trade-off between privacy-preservation methods and the quality of data mining applications, within the specific context of the smart grid. The use of smart meters to automate data collection is set to solve the problem of electricity theft, which is a serious concern in developing nations. Nevertheless, the unlimited use of data from smart meters allows for potentially private information to be discovered. There is a demand for methods to quantify the trade-off between privacy-preservation and quality of a classification model. We describe the research and development of an agent-based simulation platform to evaluate the balance between privacy-preservation mechanisms and methods for electricity theft detection. We have implemented a proof-of-concept model and validated it against real data collected from smart meters.

1 Introduction

A smart grid is an electricity and utilities grid instrumented to automatically collect and act upon information about the behaviour of suppliers and consumers. This technology aims to improve the efficiency, reliability, economics, and sustainability of the production and distribution of electricity. A common element in smart grids are the *smart meters*, used to collect information about household utilisation. Nonetheless, Cohen [9] alerts that "[smarter meters are] vulnerable to remote exploitation, viruses, worms, malicious upgrades, and all manner of other attacks". Hence, the utilisation of such devices represent a concern to individual privacy, despite the obvious benefits to electricity companies.

The interest of this research is to mitigate the aforementioned concerns without compromising the rationale for applying smarter meters in the fist place. To this end, we need to quantify the trade-off between (i) collecting detailed data through smart meters and (ii) concerns about privacy issues by individuals and groups. In particular we focus on the use of smart meters to detect electricity theft, a major problem in many developing nations, and one which information collection from smart meters can combat effectively.

In this paper, we introduce a framework for measuring the trade-off between privacy-preserving methods and data mining techniques for detecting electricity

M. Klusch, M. Thimm, and M. Paprzycki (Eds.): MATES 2013, LNAI 8076, pp. 292–305, 2013.
© Springer-Verlag Berlin Heidelberg 2013

theft. It encompasses an agent-based simulation model to (i) represent different trust and privacy profiles with regards to sharing information, considering that different people have different levels of trust and different preferences when it comes to what data is kept private, and; (ii) generate households' electricity load profiles, by applying publicly available data to create reference models, and concepts of social behaviour to create extended data sets that simulate community diversity. Then, we are able to apply probabilistic approach and data mining techniques to detect the expected behaviours in the different community setups, allowing us to measure the trade-off between privacy-preserving methods and approaches for collecting data to detect electricity thievery.

The outcome of this research provides new ways to understand the impact that the introduction of smarter grids will have in development communities. We are particularly interested in the situation of Brazil, where smarter grids are being implemented to help combat electricity theft. According to ANEEL, Brazil's national electricity agency, in some regions this problem compromises up to 25% of the total electricity production. The side-effects go beyond economical balance for energy companies, as ultimately this cost is shifted to consumers in terms of raising energy prices [16].

Moreover, this research is of interest to providers of smart grid technology. It allows for a better understanding of the social impact of applying the smart meters in different communities. In this context, we are building on the solutions and case scenarios in the *IBM Smart Grid* program [12]. This development envisages "a layer of intelligence throughout the grid to enhance system reliability and efficiency, improve management of supply and demand, optimize operations and streamline costs". We will contribute to this program with a layer of understanding about the impact and acceptability of the technology by different communities.

The paper is structured as follows. In Section 2 we discuss our motivation and present the related work in the areas of data mining methods for theft detection and privacy preservation methods. Section 3 describes our framework and Section 4 provides a proof-of-concept validation of this framework. The paper concludes with a discussion and an analysis of future work in Section 5.

2 Motivation and Related Work

The use of smart meters brings many benefits; one of which is the ability to detect and shut down electricity theft, but there are many other uses that benefit both electricity providers (such as the use of this data in forecasting electricity use) and end-users (time-of-use billing can lower overall expenses). However, the data collection that allows these uses is a double-edged sword. Load-monitoring allows the identification of specific electrical appliances, such as medical equipment, or baby monitors [14]. Over time, the data collected can be used to discover patterns of use, from which it can deduced whether a family is at home, or even when particular family members are at home. The European Union's Data Protection Supervisor recommends that, in addition to other measures, privacy-enhancing

technologies be made mandatory in the deployment of smart electricity meters [7]. While most users are not concerned with what their electricity provider can do with this information, the information is valuable. There are significant security concerns with the storage of such data, and there are no regulations in place to prevent the sale of the data to untrusted third parties.

There is a considerable amount of work on both privacy preservation for data mining techniques, as well as in the detection of electricity theft using smart meter technology. However, insofar as we know there is no work that attempts to quantify the trade-off between privacy and a specific knowledge extraction goal. We discuss some of the related work in both these separate domains below.

2.1 Detection of Electricity Theft

We emphasise that the data collection from traditional meters, which do not cause any privacy concerns, is not detailed enough to detect theft reliably through data mining techniques. For instance, Cabral and Gontijo [8] use so-called rough sets to derive rules that allow the classification of fraudulent customers, using data from traditional meters, and achieve an accuracy of 20%, which at the time was considered good. Smart meters, however, collect, and automatically send, far more fine-grained readings to the utilities companies, and using this data there are many novel approaches for detecting fraudulent activities. Kadurek et al [13] have proposed a methodology for automated detection of electricity theft, which does not require data collection and is performed in real-time at the substation, but they do not present data on how well it works, particularly in a market where electricity theft is more prevalent than the Netherlands (where they deploy their prototype).

Most state-of-the-art work in the area approaches the problem with data mining techniques. For instance, Nagi et al. use Support Vector Machines (SVMs) to detect customers with irregular consumption patterns, which are associated with fraudulent profiles [15], and Ramos et al. use a graph-based classifier, optimum-path forest (OPF) in order to detect NTLs [18]. It is not our intention to give a complete overview of the techniques used: rather, it should be clear that such methods require extensive sets of detailed information to distinguish between larcenous and honest households. Nagi et al. use a set of data from 500 households, whose electricity readings were recorded hourly for two years. Ramos et al. used data that was collected at 15 minute intervals. This kind of data can be used to uncover privacy-sensitive information and thus some form of privacy protection must necessarily be employed.

2.2 Privacy Protection in Smart Grid Technology

We are not the first to recognise the need for privacy-preserving methods for use in the smart grid. Erkin et al. survey a number of approaches designed specifically to prevent revealing sensitive information [11]. These methods all use secure signal processing techniques to encrypt an individual household's load data.

Because of specific properties of the cryptographic methods used, the aggregation can be performed on the encrypted data, meaning that when the data is decrypted, individual households' load data cannot be retrieved. While this is effective if the use of the smart meter is primarily to build predictive models for load balancing or for adapting the price in a time-of-use billing mechanism, it denies any possibility of using the data to discover fraudulent individuals.

A more promising alternative is to build on the privacy-preservation techniques in data mining. Aggarwal and Yu [1] identify various methods for this: randomisation, the k-anonymity model, distributed privacy-preserving data mining and downgrading the classifier effectiveness. We intend to test the functioning of a number of these methods. In particular randomisation seems promising: by adding random noise to the data it may be possible to sufficiently hide privacy-sensitive information while still allowing fraudulent individuals to be detected. One concern with this is obviously that real data must still be sent for the purpose of billing, but this is only needed once per billable period.

3 Methodology

The main aim of this paper is to lay out a clear methodology for evaluating the inevitable trade-off between detecting electricity theft and preserving households' privacy. While our long-term goal is to develop methods that allow theft-detection algorithms to maintain an adequate level of performance while preserving an adequate level of privacy, we need a method for making explicit what an "adequate level" is in both these cases and measure the trade-off explicitly. While we focus on theft detection as the principal application for smart meters in this paper, we wish to emphasise that the same techniques are usable for other applications of machine learning techniques on data from smart meters, such as load prediction.

The two main problems in measuring the trade-off are:

1. Testing different preferences when it comes to privacy.
2. Quantifying the trade-off between privacy-preservation and accuracy of theft detection.

In Section 3.1 we propose an agent-based simulation to solve the former problem. In Section 3.2 we present a statistical method for solving the latter.

3.1 Simulating Households' Electricity Use

Different people have different requirements when it comes to privacy. Some households may not mind providing any amount of information, regardless of what sensitive details it reveals. On the other side of the spectrum are very private individuals who are uncomfortable giving out any more information than they do currently. Both are probably minority groups, with the majority of people willing to reveal some information as long as it does not reveal sensitive

information, and enough guarantees are given that the data will not be further distributed (willingly or due to security leaks) without their permission.

This all is closely related to trust. For someone to feel comfortable giving out (potentially) sensitive information, there must be a trust relationship that this information will not be mistreated. Different people have different levels of trust and different preferences when it comes to what data is kept private. This can be simulated using a multi-agent system in which agents have different trust and privacy profiles with regards to sharing information. This also allows for distributed privacy methods to be tested, where network proximity (using either a physical network or a social one) is used to do some preprocessing in order to preserve users' privacy at the global level. However, then trust is needed at a local level as well; between households and not just between the household and the electricity provider.

Furthermore, the agent-based simulation can be used to generate the load profiles. This alleviates the problem of obtaining real data with sufficient detail. The load data available publicly, such as that made available by the Remodece project [10], is only from honest households. Such data could be used to test privacy-preserving methods, but not to distinguish between larcenous and honest households. On the other hand, data such as the load profiles availabe in the work by Nagi et al. [15] makes a clear distinction between the profiles of honest and fraudulent households, but there is not enough detailed information availabe to discover any privacy-sensitive information. As such we propose a simulation-based approach in which each agent represents a household with a specific privacy profile and generating an electricity load profile as in the work by Paatero and Lund [17] or Armstrong et al. [3]. Larcenous agents, however, deviate from the profile and instead are given profiles based on the profiles of fraudulent agents in Nagi et al.'s work. An initial step to verify this simulation-based approach is described in Section 4.

3.2 Measures

The assumption in the various theft-detection algorithms is that a detailed load profile for each household is available, which allows an accurate distinction between profiles of larcenous and lawful households. However, such detailed information allows for infringement on users' privacy.

In this section we provide a framework for measuring the trade-off between privacy-preserving methods and data mining techniques for detecting electricity theft. While privacy is generally regarded in absolute measures (such as in the encryption approach discussed in Section 2, this is not the case here: we want to know how well a privacy-preservation method hides sensitive information, with respect to the original data. For instance, if a piece of private information is already hidden, then it is not necessary to use a privacy-preservation method. We will quantify how well a user's privacy is preserved using some privacy-preservation technique with such a relative measure.

Similarly, the measure for how well the difference between larcenous and lawful individuals can be learned using machine learning methods, is also relative: we

want to know how much using a privacy-preservation method impacts a ML algorithm's functioning with respect to the original dataset. By quantifying this property as well, we can quantify the trade-off between the two different interests, and this gives us some idea of whether it is a worthwile approach or not.

Quantifying Privacy Preservation. There are a number of different ways of quantifying privacy preservation, designed for different uses [5]. Most take a probabilistic approach and see how much "harder" it is to guess the right answer after performing the privacy-preserving operation. One of the most prominent, presented by Agrawal and Aggarwal [2], perform this quantification by using the conditional entropy. Using the entropy $H(A)$ of a random variable A, they define the privacy inherent in that variable as $\Pi(A) = 2^{H(A)}$. The privacy *lost* between two random variables is given in terms of conditional entropy. If we consider the data as a random variable A with domain Ω_A, and the privacy-preserved data B, the conditional entropy is given as follows:

$$\begin{aligned} H(A|B) &= \int_{y \in \Omega_A} f_B(y) \cdot H(A|B = y) dy \\ &= -\int_{y \in \Omega_A} \int_{x \in \Omega_A} f_{A,B}(x, y) \cdot log_2(f_{A|B=y}(x)) dx dy \end{aligned} \tag{1}$$

Where we assume the privacy preservation operation does not change the domain of the random variable. f_B is the probability density function for variable B. Similarly, $f_{A,B}$ is the density function for the joint probability of A, B and $f_{A|B=y}$ the density function for the conditional probability.

Using the conditional entropy, Agrawal and Aggarwal define the "fraction of privacy loss", or the amount of privacy that is lost with regards to A by knowing B as $\mathcal{P}(A|B) = 1 - \Pi(A|B)/\Pi(A)$. We use this in a slightly modified form as our privacy preservation measure (PPM):

$$PPM(A|B) = H(A|B)/H(A) \tag{2}$$

Rather than using their measure of privacy, we use the entropy directly. The advantage of this measure over Agrawal and Aggarwal's measure is that our measure is 0 if A is entirely determined by B, whereas the original measure is $1 - 1/\Pi(A)$, which is only 0 if the entropy of A is 0. Our measure is undefined in this trivial case and we define it separately as $PPM(A|B) = 0$ if $H(A) = 0$. The original dataset has no privacy-sensitive properties to preserve. If A and B are independent, then both the measures are 1.

Example. We illustrate this with an example for discrete random variables. Assume a dataset collected from 500 homes, 450 with at least one child and 50 without children. We thus have $H(A) = \sum_{x \in \{c, \neg c\}} -p(x) \cdot log_2(p(x)) = 0.47$ bits.

Now assume we have a privacy preserving function that randomly labels 200 of the homes with at least one child as childless, resulting in set B. We have the following probabilities: $P(A = c|B = c) = 1$, $P(A = c|B = \neg c)) = 4/5$, $P(A =$

$\neg c|B = c) = 0$ and $P(A = \neg c|B = \neg c) = 1/5$. We thus have $H(A|B) = 0.36$ and $PPM(A|B) = 0.77$. This corresponds with what we would expect from a machine learning algorithm: if we train an algorithm on set B, it will learn to misclassify a large number of families with children as being childless and will thus be inaccurate on the original set A. Because the actual accuracy depends on specifics of the machine learning algorithm, it makes more sense to specify this in terms of relative entropy than in terms like precision or recall.

A second privacy-preserving function creates set C by simply removing 200 homes where at least one child is childless. We then have the following probabilities $P(A = c|C = c) = 1, P(A = c|C = \neg c) = 0$, $P(A = \neg c|C = c) = 0$ and $P(A = \neg c|C = \neg c) = 1$. This results in $H(A|C) = 0$ and also a $PPM(A|C) = 0$. It is clear why: if we learn which of the families in C have children, we can use that same classifier on set A and expect it to be accurate.

Quantifying Theft Detection. In contrast to the quantification of privacy-preservation, which we want to do independent of the specifics of the machine learning algorithm, for theft detection we are particularly interested in how well the machine learning algorithm performs. There are a number of measures that are traditionally used to quantify the functioning of machine learning algorithms, most notably *precision* and *recall* [4].

Precision can be seen as the probability that a positively classified household is a true positive. Recall, on the other hand is the probability that a true positive is correctly classified as positive. These two measures are combined into the F_β measure as follows:

$$F_\beta = (1 + \beta^2) \cdot \frac{precision \cdot recall}{(\beta^2 \cdot precision) + recall} \tag{3}$$

This can be interpreted as a weighted average (the harmonic mean) of precision and recall, with weighting factor β. If $\beta > 1$, then recall is given more importance than precision and $\beta < 1$ the reverse. Common values for β are $0.5, 1$ and 2. We suggest to use 2, because in detecting theft we are interested in casting a fairly wide net: every positive hit will need to be verified in any case to decide whether legal action should be taken. Of course, the net shouldn't be too wide or it is useless, so precision of the method should not be ignored.

Note that the F_β measure is always between 0 and 1. If either recall or precision are 0, then F_β is 0, and if both are 1, then F_β is also 1. Anything else results in an intermediate value. The measure is undefined if $F_\beta(A) = 0$.

However, we are not interested in the absolute performance of a machine learning method, but rather in the relative performance on a dataset that has been modified by a privacy-preserving method with respect to the original performance. As such we consider A the dataset before privacy-preserving measures and B after, with corresponding performance measures of a machine learning method $F_\beta(A)$ and $F_\beta(B)$ respectively. The theft detection measure is then:

$$TDM(A, B) = F_\beta(B)/F_\beta(A) \tag{4}$$

While this ratio is not limited to the $[0, 1]$ range, it is only greater than 1 if the privacy-preserving method actually improves the performance of the theft detection method. If we are afraid of this happening we can simply take the minimum value between the TDM as calculated in Eq. (4) and 1. We define $TDM(A, B) = 1$ if $F_\beta(A) = 0$, because if the recall and precision of the machine learning algorithm are both 0 on the original data set, no privacy-preservation method is going to make the method perform any worse.

Measuring the Trade-off. We now have two measures that can be seen as a way of measuring how well the privacy-preserved data performs with regards to the original. If the PPM is 0, then the modified dataset preserves privacy equally well as the original. Similarly if the TDM is 1, the dataset allows for equally good theft detection as the original. It is necessary to be very careful in comparing these two values, because strictly speaking they are not comparable: a 0.1 increase in privacy protection does not mean the same as 0.1 increase in theft-detection. Nevertheless, the measure $PPM(A|B) + TDM(A, B)$ can give a rough estimate of how well we are accomplishing the trade-off between the two conflicting goals. $PPD(A|A) + TDM(A, A) = 1$, so if the measure drops below 1 this could indicate that we are losing performance: the amount of privacy we have gained, as quantified by the PPM is less than the accuracy in theft detection we have lost, as measured using the TDM. Similarly if the measure is greater than 1 it can indicate that the loss in accuracy is offset by a greater gain in privacy preservation.

In future work we aim to evaluate the usefulness of this measure simultaneously with privacy preservation methods.

3.3 Putting It All Together

The simulator, as described in Section 3.1 simulates the data that could be collected by an electricity provider, where different households have different privacy profiles. In addition, the simulator can generate the data with full access. This gives us our two sets A and B, generated in a way that respects individual wishes for privacy. The measures of Section 3.2 can then be used as an indicator of whether we can accurately detect larcenous households and whether the data has been protected against privacy infringements. This can guide us in our research into better methods for preserving users' privacy, which can be codeveloped with datamining techniques for detecting electricity theft in privacy-protected data, and possible other applications of data from smart meters.

4 Simulation and Experimentation

We have implemented a preliminary model to demonstrate the viability of the approach detailed in the previous section. This prototype model does not implement the full framework as presented in the previous section, because we do not incorporate privacy-sensitive information into households' load profiles.

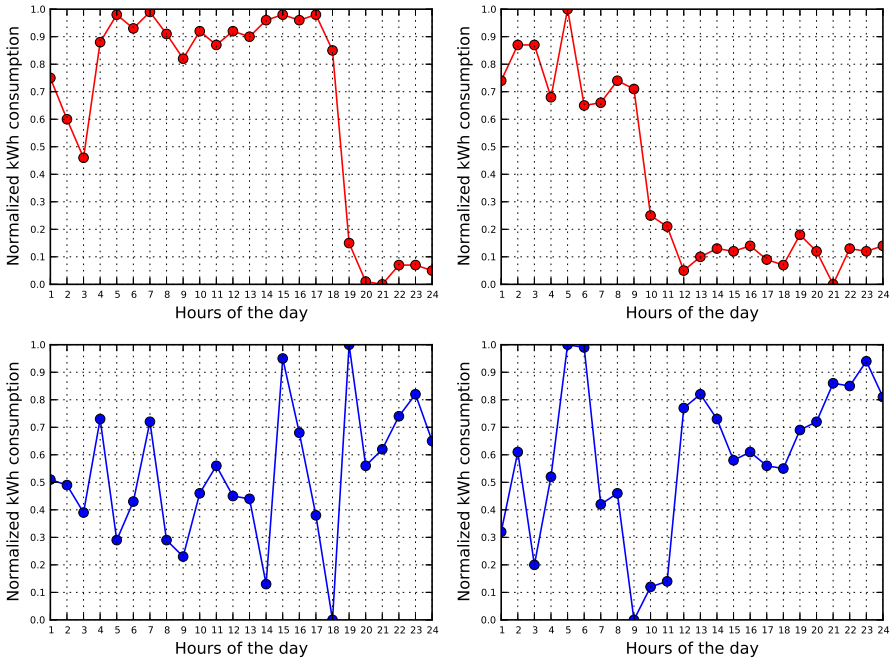

Fig. 1. The four (normalised) profiles from Nagi et al. [15], used as templates for modelling load profiles. The two top profiles (in red) are typical profiles of fraudulent households and the bottom ones (in blue) are of honest households.

Nevertheless, this prototype serves as a proof-of-concept for the model, and we demonstrate how the generation of honest and fraudulent households works. In Section 4.2 we validate the model empirically, but first we explain how it works.

4.1 Simulation Setup

In this prototype implementation, we model a neighbourhood, consisting of N households, with a percentage F of these households fraudulent. We assume that every smart meter reports its energy consumption on each hour period, and as stated above, do not yet generate privacy-sensitive information. The energy consumption, or load profile, that is generated, is based on the load profiles as presented in Nagi et al. [15], who present four typical load profiles from their set of historical consumption data in Malaysia (see Figure 1): their data set includes load profiles of fraudulent households and they disclose two typical profiles each for honest and fraudulent households. These load profiles serve as templates for generating the load profiles in our simulation.

We generate individual households' load profiles by starting with one of the templates and adding Gaussian noise, with a standard deviation σ to each of the datapoints. The choice between the four templates is decided by the percentage

F of fraudulent households and the percentage T of households using the profiles from the first column (and thus $1 - T$ using the second column).

All the experiments in the following section are run with the following parameters, unless stated differently. $N = 400$: this is the approximate size of the dataset Nagi et al. used, as well what we estimate is a typical size for a favela in Rio de Janeiro, based on recent census data. $F = 0.15$: LIGHT, the primary electricity provider in Rio de Janeiro, estimates that 15% of the electricity is lost in non-technical losses[1]. This number also corresponds with ANEEL's report, as well as Nagi et al.'s data. $T = 0.5$: we have no reason to favour one profile over another, so choose a uniform distribution. Finally, we determine the value for σ empirically in the next section, by comparing the simulated data for honest households to a dataset of load profiles from real households.

We do not incorporate privacy-sensitive information in the simulated households in this iteration of the paper, and thus do not use an agent-based model. Nevertheless, this model serves as a proof-of-concept for the method and we expect to extend the simulation with privacy-sensitive information, as described in Section 3.1.

4.2 Experimentation

In order to validate the simulation method, we perform two experiments. In the first, we show that the profiles for honest households, generated using our model, are similar to the load profiles of real households, collected in the Remodece project [10]. In the second, we show that machine learning algorithms are capable of detecting electricity theft in our simulated neighbourhood.

Experiment 1: Realism of Generated Profiles. In this experiment we show that the method for generating honest profiles generates profiles that are realistic and to determine what σ to use for the next experiment in order to keep the data as realistic as possible. For this, we compare the profiles generated for honest households to real data gathered from, insofar as anybody knows, honest households. The Remodece project collected data from smart meters in a number of countries in Europe, and we compare our generated load profiles against the load profiles in these datasets.

Because there is a high amount of variation between different households' use of electricity, it is not possible to compare load profiles with each other "directly" (even two randomly chosen profiles from the same data set may show no correlation at all with each other). We therefore compare the datasets statistically: for every hour, we test whether the real data and simulated data have similar distributions. Because neither our real data, nor the simulated data, are normally distributed (the simulated data follows a bimodal distribution: we add noise to two different template profiles), we need to use a non-parametric test. Two choices stand out, the Mann-Whitney U test and the Kolmogorov-Smirnov

[1] `http://www.relatoriolight.com.br/energia-cintica/distribuio/qualidade-na-distribuio?lang=en-US`

test. Both are applicable and test for slightly different things. The latter is generally less powerful for deciding whether two populations are similar, but has the added benefit of detecting differences in the shape of the distributions in addition to differences in the average rank. We thus simply use both and check the simulated data for different values of σ against the real data. We average the p-values for each test over the hours in the day, and the results can be found in Table 1.

Table 1. Average p-values of the Mann-Whitney U test and Kolmogorov-Smirnov test between simulated and real data. The minimum values are bolded.

σ	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1
MW U test	1.7e-2	5.0e-2	4.4e-2	2.6e-2	**7.3e-3**	4.6e-2	3.4e-2	3.2e-2	1.8e-2	6.4e-2
KS test	8.5e-3	4.2e-3	1.0e-2	1.9e-3	**9.2e-4**	**9.2e-4**	2.5e-3	2.9e-3	1.8e-3	2.6e-2

As expected given the characteristics of the test, the p-value for the Kolmogorov-Smirnov test is always smaller than for the Mann-Whitney U-test, and if we accept the hypothesis that the two distributions are similar at $p < 0.05$, this hypothesis is never rejected by it, while the Mann-Whitney U test does reject it for a number of values of σ. However, for both tests the p-value is smallest at $\sigma = 0.5$, which we will adopt in Experiment 2.

The results of this experiment seem to indicate that adding Gaussian noise to some template profiles allows us to generate realistic profiles, however we are hesitant to conclude this. Firstly, due to the large variation in profiles we have only been able to perform statistical tests per timestamp, rather than test whether the profiles truly are correlated. Secondly, it is possible that the template profiles we used happened to be "good" templates for the Remodece dataset, and this same may not be the case for other areas (such as the favelas of Rio de Janeiro). Nevertheless, the result is promising, and as we enhance the model of the households, so we may generate privacy-sensitive data, it is necessary to keep testing the generated profiles against real data in a similar manner.

Experiment 2: Detecting Fraudulent Households. In the first experiment we tested whether the simulated profiles of honest households are similar to real profiles of honest households. Due to the lack of real data regarding fraudulent households we cannot perform the same test for them. However, we do know that the ML techniques can learn to classify load profiles from real data as either honest or fraudulent. We can test our simulated data by verifying that the simulated profiles allow for a similar classification.

For this, we use the WEKA package. We tested a number of different ML settings: an SVM (similar to Nagi et al. [15]), a random forest (similar to the optimum-path forest classifier used by Ramos et al. [18], albeit less sophisticated), naive Bayes, and a multi-layer perceptron classifier [6].

We generate ten datasets with the same settings: 400 households, 15% of which are fraudulent, and using a 50/50 split of the profiles, adding Gaussian

Table 2. Average precision, recall and F_2-measure for four different ML methods over 10 simulated datasets

ML method	Precision	Recall	F_2-measure
SVM	0.98	0.87	0.89
Random Forest	0.96	0.55	0.60
Naive Bayes	0.97	0.90	0.91
Multilayer Perceptron	0.90	0.91	0.90

noise with $\sigma = 0.5$. We then use the four different learning algorithms with 10-fold cross validation. The average precision, recall and F_2-measure (calculated as described in Section 3.2) are in Table 2.

Other than the random forest method, all give very similar results, and in fact, quite significantly better than the results of either Nagi et al. or Ramos et al., who had a precision and recall around 0.8 on real data. This indicates that our dataset is actually "too easy" to accurately represent real profiles. By increasing σ we can make it harder to learn a correct classification. Simply increasing σ to 0.6 gives a precision and recall that is more in line with the results in the related work, at the cost of a decrease in realism of the honest profiles (per Table 1).

A remarkable side result of this exploratory experiment is that we did not expect naive Bayes, a significantly simpler learning method than using an SVM or a multilayer perceptron classifier, to perform so well. We have not seen naive Bayes applied on real data in any of the related work we studied, and it is worth investigating whether it gives similar results in practice.

5 Discussion and Future Work

Preservation of private information is a real concern for many modern applications of information technology. When deploying smart meters such privacy concerns should be addressed without compromising the benefits of using smart meters in the first place. The methodology we present in Section 3 makes explicit the trade-off between privacy-preservation and extraction of important information; in this case whether a household is fraudulent or not. Using a multi-agent system to model a neighbourhood allows for the flexible implementation of various different consumer profiles and their possible interactions.

The results of Section 4 validate our proof-of-concept implementation of the method, by comparing the simulated neighbourhood to real data. We argue that this demonstrates a novel metod for understanding the social impact of smart sensor technology. The methodology as presented allows for analysing the impact of smart grid technology on different communities.

The next step is to generate household profiles that contain privacy-sensitive information, and generate the household's electricity load profiles based on these, while verifying that this is still realistic, in the same sense as our currently

simulated data. These load profiles can then be used to verify that privacy-sensitive data can be discovered through data mining techniques, and with this simulation in place we can move on to both testing current privacy-preservation methods and designing new ones that optimise the trade-off between privacy and knowledge discovery.

While we are not there yet, this paper presents a step towards a privacy-conscious use of data from smart meters.

Acknowledgements. Gabriel Ramos and Ana Bazzan are partially supported by CNPq, Andrew Koster is supported by CAPES (PNPD). All three are supported in their work by FAPERGS. This research was done in collaboration with IBM Research Brazil.

References

1. Aggarwal, C.C., Yu, P.S.: A general survey of privacy-preserving data mining models and algorithms. In: Privacy-Preserving Data Mining. Advances in Database Systems, vol. 34, pp. 11–52. Springer, Heidelberg (2008)
2. Agrawal, D., Aggarwal, C.C.: On the design and quantification of privacy preserving data mining algorithms. In: Proceedings of the Twentieth ACM SIGMOD-SIGACT-SIGART Symposium on Principles of Database Systems, pp. 247–255. ACM (2001)
3. Armstrong, M.M., Swinton, M.C., Ribberink, H., Beausoleil-Morrison, I., Jocelyn, M.: Synthetically derived profiles for representing occupant-driven electric loads in canadian housing. Journal of Building Performance Simulation 2(1), 15–30 (2009)
4. Baeza-Yates, R., Ribeiro-Neto, B.: Modern Information Retrieval: The Concepts and Technology behind Search, 2nd edn. ACM Press (2011)
5. Bertino, E., Lin, D., Jiang, W.: A survey of quantification of privacy preserving data mining algorithms. In: Privacy-Preserving Data Mining. Advances in Database Systems, vol. 34, pp. 183–205. Springer, Heidelberg (2008)
6. Bishop, C.M.: Pattern Recognition and Machine Learning. Springer (2006)
7. Buttarelli, G.: Opinion of the European data protection supervisor on the commission recommendation on preparations for the roll-out of smart metering systems. EU Recommendation (2012)
8. Cabral, J., Gontijo, E.: Fraud detection in electrical energy consumers using rough sets. In: 2004 IEEE International Conference on Systems, Man and Cybernetics, pp. 3625–3629 (2004)
9. Cohen, F.: The smarter grid. IEEE Security Privacy 8(1), 60–63 (2010)
10. de Almeida, A., Fonseca, P., Bandeirinha, R., Fernandes, T., Araújo, R., Urbano, N., Dupret, M., Zimmermann, J.P., Schlomann, B., Gruber, E., Kofod, C., Feildberg, N., Grinden, B., Simeonov, K., Vorizek, T., Markogianis, G., Giakoymi, A., Lazar, I., Ticuta, C., Lima, P., Angioletti, R., Larsonneur, P., Dukhan, S., de Groote, W., de Smet, J., Vorsatz, D., Kiss, B., Ann-Claire, L., Pagliano, L., Roscetti, A., Valery, D.: REMODECE: Residential monitoring to decrease energy use and carbon emissions in europe. Technical report, IEEA Programme (2008)
11. Erkin, Z., Troncoso-Pastoriza, J.R., Lagendijk, R.L., Pérez-González, F.: Privacy-preserving data aggregation in smart meter systems. IEEE Signal Processing Magazine 30(2), 75–86 (2013)

12. IBM Corp.: IBM Smart Grid, http://www.ibm.com/smarterplanet/energy (last checked May 2013)
13. Kadurek, P., Blom, J., Cobben, J.F.G., Kling, W.: Theft detection and smart metering practices and expectations in the netherlands. In: 2010 IEEE PES Innovative Smart Grid Technologies Conference Europe (ISGT Europe), pp. 1–6 (2010)
14. Laughman, C., Lee, K., Cox, R., Shaw, S., Leeb, S., Norford, L., Armstrong, P.: Power signature analysis. IEEE Power and Energy Magazine 1(2), 56–63 (2003)
15. Nagi, J., Yap, K., Tiong, S.K., Ahmed, S., Mohamad, M.: Nontechnical loss detection for metered customers in power utility using support vector machines. IEEE Transactions on Power Delivery 25(2), 1162–1171 (2010)
16. Nagi, J.: An inteligent system for detection of non-technical losses in Tenaga Nasional Berhad (TNB) Malaysia low voltage distribution network. Master's thesis, University Tenaga Nasional (June 2009)
17. Paatero, J.V., Lund, P.D.: A model for generating household electricity load profiles. International Journal of Energy Research 30(5), 273–290 (2006)
18. Ramos, C.C.O., Souza, A.N., Papa, J.P., Falcão, A.X.: Learning to identify non-technical losses with optimum-path forest. In: Proceedings of the 17th International Conference on Systems, Signals and Image Processing (IWSSIP 2010), pp. 154–157 (2010)