International Conference on Applied Energy 2019
Aug 12-15, 2019, Västerås, Sweden
Paper ID: 108

PRIVACY PROTECTION OF SMART METER USERS BASED ON RECHARGEABLE BATTERIES AND SOLAR PANELS

Yu Liang[1], Yang Liu[2*]

1 School of Electrical and Information Engineering, Tianjin University, Tianjin 300072, China (Corresponding Author)

2 School of Electrical and Information Engineering, Tianjin University, Tianjin 300072, China

ABSTRACT

In smart grid, smart meter (SM) samples the energy consumption of users at high frequency. By analyzing measurement data, users' behavior patterns can be identified, which threatens consumers' privacy. This paper studies the protection of SM privacy by solar power generation system and rechargeable battery (RB). Solar energy provides alternative power other than power grids for users. RB provides energy storage. Under the constraints of RB physical conditions and solar radiation, a "privacy-cost dual optimization model" is proposed to minimize the weighted sum of privacy leakage and power consumption costs. In order to solve the optimization problem, we propose a new optimization algorithm. The simulation results show the effectiveness of the algorithm. The effects of different target curves on privacy protection are studied.

Keywords: renewable energy, Smart Grid, privacy protection, rechargeable battery, optimal algorithm

1. INTRODUCTION

Smart Meter (SM) is a key part of smart grid, which records the users' load in seconds. Non-intrusive load monitor technology [1] can be used to dig out the use information of household appliances and infer consumer privacy. At present, there are three methods to protect privacy. The first method contains homomorphic encryption [2], anonymous authentication [3], etc. This method can ensure that the home power data package sent by SM to power companies, even if intercepted, must be correctly decrypted to obtain real data. However, the encryption method cannot guarantee that the user's data will not be dig out. The second is data tampering method [4]. Noise is added to disturb the original power consumption data before sending it to power company. Data tampering effectively limits privacy leaks, but it poses challenges to the correct charging of electricity bills. The third method is based on the Battery Load Hiding (BLH) method [5]. Charging and discharging of batteries can change the user's energy consumption, thus hiding the use of the original electrical appliances. In addition, the BLH method does not affect the normal charge. Recently, some scholars have considered using rechargeable batteries (RB) and other devices to protect privacy better. In [6] constant temperature control device and RB are used to hide load. Considering the existence of photovoltaic power generation system and RB in household microgrid. In this paper, privacy protection is achieved by using RB and solar energy to let electricity data recorded by SM track a set target. The ideal predetermined target value is found. From the perspective of information theory, mutual information is chosen as the standard of privacy. The optimization problem is decomposed into sub-optimal problems solved in each time period. An online optimization algorithm based on particle swarm optimization is proposed. The simulation results verify the effectiveness of the algorithm.

2. PRELIMINARIES

2.1 System model

This system is composed of RB, solar panel, household load and power router.

(1) Rechargeable battery model

The current residual energy of the RB is equal to the residual energy of the previous slot plus the energy consumed by the current slot as follow:

$$B(t+1) = B(t) + P(t) \quad (1)$$

RB's power should not be too large, so [7]:

$$-P^{min} \leq P(t) \leq P^{max} \quad (2)$$

Selection and peer-review under responsibility of the scientific committee of the 11th Int. Conf. on Applied Energy (ICAE2019).

P^{min}=0.5C, P^{max}=0.3C (3)

RB should not be over-discharged or over-charged, so:

$$0.2 \cdot B^{max} \le B(t) \le B^{max} \quad (4)$$

In general, the following limitations should be met [7]:

$$P(t) \ge -\min\{P^{min}, B(t)\} \quad (5)$$

$$P(t) \le \min\{P^{max}, B^{max} - B(t)\} \quad (6)$$

(2) Solar panel model

RB stores enough energy, so the intermittent nature of solar power has no effect. *s* represents the efficiency of solar panels (about 18%) [8], and $R_s = \{r_1, r_2, \dots, r_n\}$ represents the intensity of light. The power generated by solar panels in each time slot is:

$$P_s(\mathrm{t}) = s \times As \times R_S(t) \quad (7)$$

(3) Load model

L(t) represents the actual power, and it comes from Los Angeles (N33°, W118°), America in 1990. There is an upper limit on the electrical load:

$$0 \le L(t) \le L^{max} \quad (8)$$

(4) Household Power Router

The power router can receive user instructions and manage the energy of the microgrid by controlling the direction and size of the energy flow of each node, which provides technical basis for the power flow in the system.

(5) Electricity price model

Let *C(t)* represents the different price in each time slot. Take Tianjin residential electricity standard as an example:

$$C(t) = \begin{cases} 0.49 & 6<\text{hour}<21 \\ 0.3 & 21<\text{hour}<6(\text{The next day}) \end{cases} \quad (9)$$

2.2 Mutual information Variable definition

The power router can be viewed as a communication channel, whose input sequence is user's energy consumption $X = \{X_t\}$ and output sequence is the smart meter measurement $Y = \{Y_t\}$. As a result, the problem of preserving privacy can be viewed as preserving the privacy between the input and output of the channel. In [9], we know that the mutual information is an important criterion for evaluating the correlation between two groups of data:

$$\begin{aligned} \mathrm{I(X;Y)} &= \mathrm{H(X)} - \mathrm{H(X|Y)} \\ &= H(X) + H(Y) - H(X,Y) \\ &= \sum_{x\in X} p(x) log \frac{1}{p(x)} + \sum_{y\in Y} p(y) log \frac{1}{p(y)} \\ &\quad - \sum_{x,y} p(x,y) log \frac{1}{p(x,y)} \\ &= \sum_{y\in Y}\sum_{x\in X} p(x,y) log\left(\frac{p(x,y)}{p(x)p(y)}\right) \end{aligned} \quad (10)$$

Smaller $I(X;Y)$ implies more privacy, so the goal of designing control policy is usually to minimize $I(X;Y)$.

John et al. classified the input space into three categories, i.e. strong correlation variable, weak correlation variable and irrelevant variable [10]. The self-information $I(Y;Y)$ of the variable *Y* represents all the information contained in *Y*. If the input variable X_i satisfies:

$$\mathrm{I}(X_i;Y) > \delta_1 I(Y;Y) \quad (11)$$

It indicates that X_i contains a certain amount of information about *Y*, that is, $X_i$is the relevant variable of *Y*. Among them, δ_1 ($\delta_1 \in [0,1]$) is the correlation threshold. If the input variable X_i does not satisfy the formula (11), it means that X_i does not contain *Y* information or contains only a small amount of information about *Y,* so it can be approximated that X_i has nothing to do with *Y*. After comprehensive consideration, this paper chooses δ_1 =0.3 in the simulation experiment [11]. $I(Y;Y)$ =1, so when $I(X_i;Y) \le 0.3$ (i.e. MI < 0.3), it can be considered that the user's electricity information leaks very little and privacy is well protected.

2.3 Variable definition

$C(t)$: Electricity price in each slot.
$X(t)$: Actual power in each slot.
$P(t)$: The power consumption of the battery.
$P_s(t)$: Power generated by solar panels.
$P_S{}'(t)$: solar panels' energy used by user.
$L'(t)$: Data tracked by smart meters.
β: Privacy Protection Coefficient.
$B(t)$: Residual capacity of storage battery.
P^{max}: Maximum Charging Power of Batteries.
P^{min}: Maximum Discharge Power of Battery.
B^{max}: Battery capacity
$R_S(t)$: The intensity of light in each time slot.
As: The effective area of solar panels.
L(t): The actual power consumed in the time slot t.

3. PROBLEM DEFINITION

The smaller $I(X;Y)$ is, the less privacy leaks. In practical applications, $I(X;Y)$ is a function of probability vector P of random variable X and conditional probability matrix Q, denoted as $I(P;Q)$. Mutual information as a function of P and Q has the properties of convex function. When P is fixed, mutual information $I(P;Q)$ is the convex function of Q. So, when P is fixed, there is a conditional probability matrix Q to minimize mutual information $I(P;Q)$. In this paper, the value of set $\{X_t\}$ should be fixed in a statistical time, so $I(P;Q)$ has the minimum value when $\{Y_t\}$ is known, so the optimal solution exists. According to the formula of mutual information:

$$\mathrm{I(X;Y)} = \sum_{y\in Y}\sum_{x\in X} p(x,y) log\left(\frac{p(x,y)}{p(x)p(y)}\right) \quad (10)$$

The smaller the $p(x,y)$, the better privacy protected. $p(x,y)$ denotes the joint probability density of two sets data. It can be viewed as the joint probability density, which records the number of the (x_i, y_i) falling into $n \times n$ square lattices consisting of two sets of data. Assuming that *X* is the user's real power consumption data and is divided into *N* equal parts from the minimum to the maximum value, when the data *Y* is a fixed value or fluctuating near it, only the joint probability density of the fixed value is not zero. In this way, $I(X;Y)$ can be small enough to achieve protecting privacy.

In this paper, we set the target of electric data tracking as three curves: straight line, triangular wave and square wave which fluctuate near a fixed value.

To sum up, on the one hand, we need to make the data of SM follow the set curve by invoking the electric energy of RB and solar energy; on the other hand, we should save electricity in each time slot. As follow:

$$\mathrm{F} = \min\{(1-\beta)\cdot c(t)\cdot[X(t)+P(t)+P_s{}'(t)] + \beta\cdot [X(t)+P(t)+P_s{}'(t)-L'(t)]^2\} \quad (12)$$

The first item in the formula represents the electricity charges paid by users to the power grid.

$X(t)+P(t)+P_s{}'(t)$ represents the actual power. $X(t)$ represents the power of the household appliance, and $P(t)$ represents the power of RB. $P_s{}'(t)$ indicates the solar power used, which is not positive. Three formulas add up to the power family gets from the grid

The second item represents the leakage of users' privacy. The larger the value, the more privacy leaks.

Due to the limitations of battery, solar energy and other equipment, constraints are as follows:

$$P(t) \geq -\min\{P^{min}, B(t)\} \quad (13)$$

$$P(t) \leq \min\{P^{max}, B^{max} - B(t)\} \quad (14)$$

$$P_s{}'(t) \leq P_s(t) \quad (15)$$

4. STOCHASTIC OPTIMIZATION METHOD BASED ON PSO

Because of the existence of integer variables, the initial problem is NP-hard and cannot be solved in polynomial time. In this paper, a stochastic method based on PSO (Particle Swarm Optimization) is used to reduce the computational complexity and find a relatively ideal result with fewer iterations. Therefore, the convergence speed of our algorithm can be greatly improved, but the cost is to lose some optimality.

The process of the Random Optimal Search Method in this paper is as follows: Firstly, *m* groups of vectors are randomly generated. Each group of vectors has *S* dimension, where the group i of vectors can be expressed as $\overrightarrow{x_i} = (x_{i1}, x_{i2}, \dots, x_{iS}), i = 1,2,\dots,m,$ each set of vectors is a potential solution. By substituting $\overrightarrow{x_i}$ into the objective function, the corresponding function value can be calculated. In all these vectors, the objective function value corresponding to one vector is the smallest, that is, the current optimal vector. Each vector has a direction of change $\vec{V} = (V_{i1}, V_{i2}, \dots, V_{iS})$. Note that the optimal value of the first vector searched so far is $\overrightarrow{P_{is}}$, and the optimal value of the M vectors searched so far is $\overrightarrow{P_{gs}}$. Then iteration is carried out, in which the current optimal value of each vector is determined by the following formula.

$$p_i(t+1) = \begin{cases} p_i(t) \longrightarrow f(x_i(t+1)) \geq f(p_i(t)) \\ X_i(t+1) \longrightarrow f\big(x_i(t+1)\big) < f(p_i(t)) \end{cases} \quad (16)$$

The changes in each vector are as follows:

$$v_{is}(t+1) = v_{is}(t) + c_1 r_{1s}(t)[p_{is}(t) - x_{is}(t)] + c_2 r_{2s}(t)\big[p_{gs}(t) - x_{is}(t)\big] \quad (17)$$

$$x_{is}(t+1) = x_{is}(t) + v_{is}(t+1) \quad (18)$$

In the formula, $i = [1, \mathrm{m}], s = [1, \mathrm{S}]$; learning coefficients c_1 and c_2 are non-negative constants, and r_1 and r_2 are independent pseudo-random variables, which obey the uniform distribution on [0, 1]. $v_{is} \in [-v_{max}, v_{max}]$, v_{max} is a constant set by the user.

The global optimal solution is found in the new m vectors after iteration. By analogy, the current optimal solution is found in every iteration. Then the current optimal solution obtained after a certain number of iterations can be regarded as the optimal solution of the objective function within the error allowable range.

Table 1: Optimal power allocation algorithm

1	For each period T
2	//Inputs: X(t),$P_s(t)$, B(t), P^{max}, P^{min}, *C(t)*
3	//Outputs: *Y(t)*, electricity fees, $I(X;Y)$
4	Initialize *M* group random solutions and their direction of change
5	For each iteration
6	Calculate the value of the objective function of each group and find the minimum value F_{best}.
7	If F_{best}<$F_{allbest}$
8	Then $F_{allbest} = F_{best}$
9	End if
10	Update the size and direction of each group according to formulation above.
11	End for

In this framework, the power router (represented by the smart meter) and multiple electrical equipment exchange information iteratively to calculate a final scheduling result [12], thus the final smart data reported to the utility company would be totally different from

that without scheduling. The residential user's privacy can be maintained.

5. SIMULATION RESULTS AND ANALYSIS

In this section, we use MATLAB to evaluate our proposed model. The electricity consumption values noted by *X(t)* are taken from a real house electricity consumption. The prevision of photovoltaic electricity productions is taken from [13]. The battery capacity is 3KWh. The number of time slots is 7200, which can cover entire day (24 hours) where each time slot represents 12 seconds. The electricity price model is taken from Tianjin.

We compare the effect of privacy protection (MI) and electricity fees when the reading of the smart meter follows the straight line, triangle wave and square wave respectively.

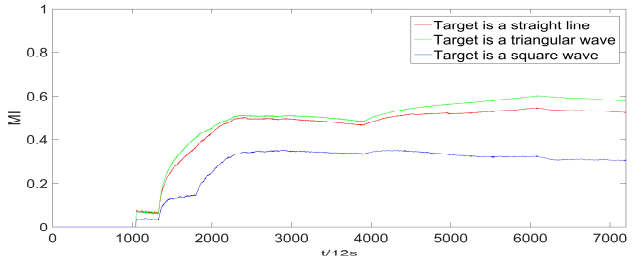

Fig 1 The effect of privacy protection of three target curves

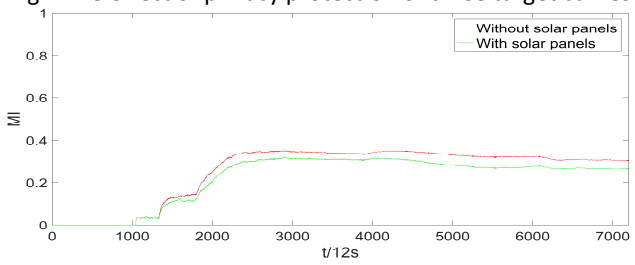

Fig 2 The privacy protection effect comparison

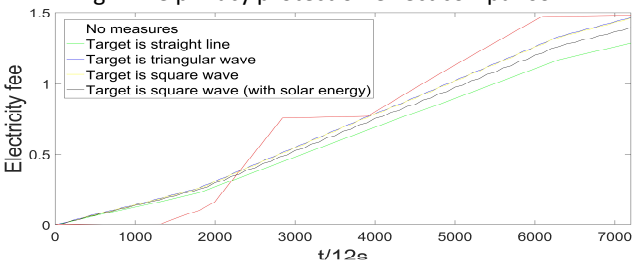

Fig 3 Electricity fee comparison

As mentioned in 2.2 above, when the mutual information (MI) of two groups of variables is no more than 0.3, the two groups of variables can be approximated as irrelevant variables. Therefore, in simulation, we regard MI < 0.3 as the criterion that two groups of variables are irrelevant.

Without any measures, user's privacy is completely exposed (MI is 1). After protection, it can be seen from Figure 1 that when the target curve is square wave, the effect is the best, which can protect privacy (MI = 0.30) and save electricity (1.46) compared with straight line and triangular wave. As is shown in Figure 2, using solar panels could significantly protect privacy and save power (MI = 0.26, electricity fee is 1.39), so it is necessary to use solar panels in privacy protection system.

As is shown in Figure 3, without any measures, the user's electricity cost is the highest (1.4809). After protection, when the target is straight, the electricity cost is the lowest, but the privacy protection effect is not good, so the comprehensive effect is not ideal. When the target curve is square wave and solar energy is used meanwhile, not only the privacy is effectively protected, but also the electricity cost is significantly reduced, So the comprehensive effect of is ideal.

In summary, the method proposed in this paper, has protected privacy effectively, and save electricity fee. An ideal target curve has been found. The effectiveness of this method has been proved by simulation.

6. CONCLUSION

This paper discusses the problem of protecting users' privacy and save electricity fees in the context of the wide application of SM and solar panels, models the problem as a multi-constrained optimization problem, and proposes an optimization algorithm based on PSO. By making SM track the target curve to achieve privacy protection, an ideal target curve is found. The effect of the algorithm and the performance of the ideal curve are verified by simulation. This strategy not only protects privacy, but also saves electricity costs, so it is practical.

REFERENCE

[1] M. Marceau, and R. Zmeureanu. Nonintrusive load disaggregation computer program to estimate the energy consumption of major end uses in residential buildings. Energy Convers Manage 2000;41(13),1389-1403.

[2] R. Lu, X. Liang, X. Li, X. Lin, and X. Shen. EPPA: An Efficient and Privacy-Preserving Aggregation Scheme for Secure Smart Grid Communications. IEEE Transactions on Parallel and Distributed Systems 2012;23(9):1621-1631.

[3] C. Efthymiou, G. Kalogridis. Smart Grid Privacy via Anonymization of Smart Metering Data. IEEE International Conference on Smart Grid Communications, 2010: 238-243.

[4] S. Wang. A randomized response model for privacy preserving smart metering. IEEE Trans. Smart Grid 2012; 3(3): 1317–1324.

[5] J. Koo, X. Lin, and S. Bagchi. PRIVATUS: Wallet-friendly privacy protection for smart meters. Proc. 17th Eur. Symp. Res. Comput. Secur. 2012; 343–360.
[6] L. Endong, P. Cheng. Achieving Privacy Protection Using Distributed Load Scheduling: A Randomized Approach. IEEE Transactions on Smart Grid 2017; 8(5): 2460-2473.
[7] Xiaoze. P. Research on the Charging and Discharging System of High Power Lithium Ion Batteries. Beijing Jiaotong University, 2008.
[8] Anonymous. Conversion efficiency of P-type polycrystalline battery of Jingke Energy has set a world record again. Screen Printing Industry. 2017; (10): 66-67.
[9] L. ZhuX. Fundamentals of Applied Information Theory. Beijing: Tsinghua University Press; 2001, chapter 2.
[10] John G H, Kohavi R, Pflfleger K. Irrelevant features and the subset selection problem. In: Proceedings of the 11th International Conference on Machine Learning. San Francisco, USA: Morgan Kaufmann, 1994. 121−129
[11] M. Han, X. Liu. Multi-variable sequence prediction based on mutual information step-by-step input variable selection. Journal of Automation, 2012, 38 (6):999-1006.
[12] S. Zong, X. He, J. Wu. Research status and development of power routers based on power electronic conversion. China Journal of Electrical Engineering 2015; 35 (18): 4559-4570.
[13] National solar radiation data base 1961-1990. http://rredc.nrel.gov/solar/old_data/nsrdb/1961-1990