Vulnerability Assessment and Defense Technology for Smart Home Cybersecurity Considering Pricing Cyberattacks

Yang Liu[1], Shiyan Hu[1], Tsung-Yi Ho[2]
[1] Department of Electrical and Computer Engineering,
Michigan Technological University, Houghton, MI, USA
[2] Department of Computer Science,
National Chiao Tung University, Hsinchu, Taiwan
Email: {yliu18, shiyan}@mtu.edu, tyho@csie.ncku.edu.tw

***Abstract*— Smart home, which controls the end use of the power grid, has become a critical component in the smart grid infrastructure. In a smart home system, the advanced metering infrastructure (AMI) is used to connect smart meters with the power system and the communication system of a smart grid. The electricity pricing information is transmitted from the utility to the local community, and then broadcast through wired or wireless networks to each smart meter within AMI. In this work, the vulnerability of the above process is assessed. Two closely related pricing cyberattacks which manipulate the guideline electricity prices received at smart meters are considered and they aim at reducing the expense of the cyberattacker and increasing the peak energy usage in the local community. A countermeasure technique which uses support vector regression and impact difference for detecting anomaly pricing is then proposed. These pricing cyberattacks explore the interdependance between the transmitted electricity pricing in the communication system and the energy load in the power system, which are the first such cyberattacks in the smart home context. Our simulation results demonstrate that the pricing cyberattack can reduce the attacker's bill by 34.3% at the cost of the increase of others' bill by 7.9% on average. In addition, the pricing cyberattack can unbalance the energy load of the local power system as it increases the peak to average ratio by 35.7%. Furthermore, our simulation results show that the proposed countermeasure technique can effectively detect the electricity pricing manipulation.**

***Index Terms*— Smart Home, Cybersecurity, Cyberattack, Advanced Metering Infrastructure, Electricity Pricing Manipulation**

I. INTRODUCTION

Smart home controls the end use of a power grid and it has become a critical component in the smart grid infrastructure. Smart home scheduling technique plays a key role in the smart home system as it controls the energy usage of the customers. It reduces the energy consumption during the time period when electricity price is high and shift it to the cheap time period. In this fashion, the average monetary cost of customers for energy consumption is reduced and the energy load of the power grid is balanced over the time horizon. From the utility perspective, a nicely designed electricity pricing curve will guide the energy usage of customers which can result in a balanced energy load in the community.

In fact, in the prevailing US electricity market, there are *two pricing models* which are popularly used together. The first one is the *real time pricing*. Based on the energy consumed in a past time window (e.g., one hour), a utility can charge the local community through billing to each customer who consumes the energy in that time window. Basically, the real time pricing determines the charge based on the energy consumption. The second model is called *guideline pricing*. Since a smart home scheduler is to determine when to launch each appliance in advance, it needs to know the future utility pricing. However, this cannot be known since it depends on the actual energy consumption in the future. Thus, the utility has to predict the future load, design a predictive pricing curve, and use it to guide the smart home schedulers on when the cheap time slots would be. This predictive pricing is called guideline pricing. Basically, the guideline pricing is not used to charge the customers while it is necessary for guiding them on energy scheduling. Refer to Figure 2 for the two pricing models provided from ComED Corporation. Typically, the predictive guideline pricing well matches the actual real time pricing, which is as expected.

Implementing smart home scheduling needs the support from the communication system associated with a smart grid, which is known as *advanced metering infrastructure* (AMI) [1], [2]. The AMI enables the two-way communications between utilities and the customers and facilitates smart home scheduling. In AMI, the utility transmits the pricing information to a central computer in the local community or substation through Internet. Subsequently, such pricing information is broadcast to smart meters through Internet or WiFi network or a combination of them, depending on the communication infrastructure of the local community. For example, in a hierarchical infrastructure where a community consists of multiple subcommunities, pricing information is forwarded to each subcommunity through Internet and is then broadcast inside the subcommunity through the WiFi network. In a WiFi network, there are some access points, which serve as the agents to receive the pricing information from the subcommunity and forward it to the smart meters inside the subcommunity. Refer to Figure 1.

The above popular infrastructure is vulnerable to at least three attacking strategies. First, one can directly hack the computer in the local community or substation and modify the pricing information there. Subsequently, the pricing information forwarded to the whole community could be a faking one. Second, one can block an access point in the WiFi network using the jamming attack (i.e., sending excessive requests to the access point), create a fake access point and send the faking pricing information to the smart meters covered by the fake access point. Third, one can hack the smart meter and modify the pricing it receives. Since the modern smart meters are microprocessor based and installed with advanced operating system, one can hack it through uploading computer virus. With the different difficulty levels in implementation, the cyberattacker can choose which one to use.

In this paper, we will focus on the impact of pricing cyberattacks to the customers as well as the local power system. We show that it is possible to design specific cyberattacks on modifying the guideline pricing curves for reducing the expense of the cyberattacker and increasing the peak energy usage in the local power system. After analyzing these impacts, we will design defense technology against the pricing cyberattacks. The key is to identify the pricing manipulation. For this, we first predict the guideline pricing curve from recent historical pricing curves using the support vector regression technique since the guideline pricing tends to be similar in a short term. The predicted guideline pricing will be then compared to the guideline pricing curve received at the smart meter. For this, one could compute the maximum difference between the two curves and signal an alert when it is larger than a threshold. The main problem with this idea is that it would be difficult to design a good threshold in practice. Note that the goal for setting a good threshold is to balance the false detection rate and the impact to the local power system and customers. For example, if one sets the threshold to be zero, then all pricing manipulation can be identified but there will be a lot of false detection. On the other hand, it is reasonable to ask, but difficult to answer, what the best threshold is if customers can tolerate up to 1% bill increase.

In contrast, our philosophy is that since anyway the goal for setting a good threshold is to limit the impact to local community and customers, why not directly using the impact as the threshold? To explore this idea, we will use a new concept called impact differences to quantify the bill and the peak to average ratio (PAR) increase comparing the predicted guideline pricing curve and the received guideline pricing curve. When the impact differences are larger than some preset maximum tolerable impact differences, a potential pricing attack is spotted and an alert signal will be sent to the utility to request for further check which could need some amount of human interaction. This work will develop such a framework.

These pricing manipulation cyberattacks can be viewed as integrity cyberattacks. In fact, there are some related works in the literature. In [3], an attacking strategy is proposed in which the attacker jams the communication network and prevents the customers from knowing the updated electricity price. In [4], the attacker introduces bias to the sensor measurement, which misconducts the bidding in the electricity market and impacts the electricity price. However, the above works use coarse models in analyzing the relationship between energy and pricing. They do not consider the impact due to smart home scheduling which is however important as the energy usage scheduling results could be significantly different with different electricity pricing. Countermeasure techniques for integrity cyberattcks have also been studied in literature. In [5], a likelihood ratio test based algorithm is proposed to detect the malicious data attack on smart grid state estimation. In [6], the jamming attack modeling and detection are proposed for time critical networks. The work [7] designs a detection technique at the package level. However, these works do not address the pricing cyberattacks in the smart home context.

In this work, the vulnerability of the electricity pricing transmission in AMI is assessed. Two closely related pricing cyberattacks which manipulate the guideline electricity prices received at smart meters are considered. They target for reducing the expense of the cyberattacker as well as increasing the peak energy usage in the local community. A countermeasure technique which uses support vector regression and impact difference for pricing anomaly detection is then proposed for the pricing cyberattacks. Our contributions are summarized as follows.

- Two cyberattacks which manipulate the guideline electricity price received at the smart meters are considered. They target for reducing the expense of the cyberattacker as well as increasing the peak energy usage in the local community.
- A countermeasure technique based on the support vector regression and the impact difference is proposed to detect the pricing manipulation.
- The proposed cyberattacks explore the vulnerability of the power system due to the cyberattacks on the communication system, through studying the interdependance between the electricity pricing and the energy load. These are the first such cyberattacks in the smart home context.
- Our simulation results demonstrate that the pricing cyberattack can reduce the attacker's bill by 34.3% at the cost of the increase of others' bill by 7.9% on average. In addition, the pricing cyberattack can significantly unbalance the energy profile of the local power system as it increases the peak to average ratio by 35.7%.
- Our simulation results show that the proposed countermeasure technique can effectively detect the electricity pricing manipulation.

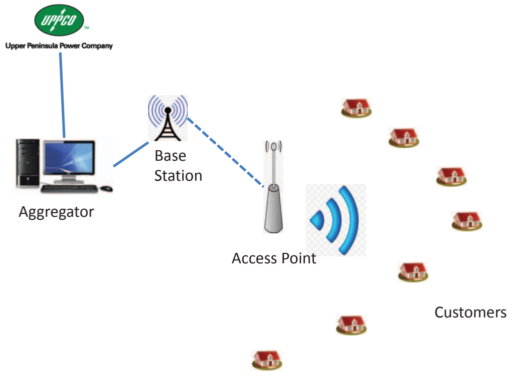

Fig. 1. System Model

II. Preliminaries

A. Smart Home System Model

A set of $\mathcal{N} = \{1, \ldots, N\}$ customers supplied by the same utility are considered in the system. Refer to Figure 1. Each customer schedules the energy consumption during the next 24 hours from the current moment, which is divided into H time slots such that $\mathcal{H} = \{1, 2, \ldots, H\}$. Each customer is equipped with a smart meter, which receives the electricity price from the utility and sends the measurement of real time energy consumption to the aggregator. In the following, we will give an overview of the game model and solution of smart home scheduling, which have been developed in our previous works [8], [9].

1) Energy Consumption: Each customer $n \in \mathcal{N}$ has a set of home appliances denoted by $\mathcal{A}_n$. At each time slot h, the home appliance $m \in \mathcal{A}_n$ works under power level $x_{m,h} \in \mathcal{X}_m$, where $\mathcal{X}_m$ is the set of available power levels for home appliance m. In general, there are two categories of home appliances, namely, manually controlled home appliances and automatically controlled ones. The category of manually controlled home appliances consists of TV set, computer, refrigerator etc.The category of automatically controlled home appliances consists of washing machine, cloth dryer, dish washer, electric vehicle (EV) etc. It is worth noting that HVAC (Heating, Ventilation and Air Conditioning) system could be classified into both categories. For example, the customer can adjust the working power level of the air conditioner manually according to the temperature in the room. However, there also exists situation that the customer needs the temperature to reach a certain level at a certain time point, in which the air conditioner works in the automatic mode. For each customer n, denote by $\mathcal{P}_n$ the set of manually controlled home appliances and denote by $\mathcal{Q}_n$ the set of automatically controlled home appliances. We have $\mathcal{A}_n = \mathcal{P}_n \bigcup \mathcal{Q}_n$. For customer n, the total energy consumption of the manually controlled home appliances at time slot h is denoted by $l^p_{n,h}$. We have

$$l^p_{n,h} = \sum_{m \in \mathcal{P}_n} x_{m,h} t_{m,h}, \tag{1}$$

where $t_{m,h}$ is the actual execution time of home appliance m at time slot h. At time slot h, the total energy consumption of automatically controlled home appliances is denoted by $l^q_{n,h}$. We have

$$l^q_{n,h} = \sum_{m \in \mathcal{Q}_n} x_{m,h} t_{m,h}. \tag{2}$$

For each automatically controlled home appliance $m \in \mathcal{Q}_n$, the power level is chosen subject to these constraints.

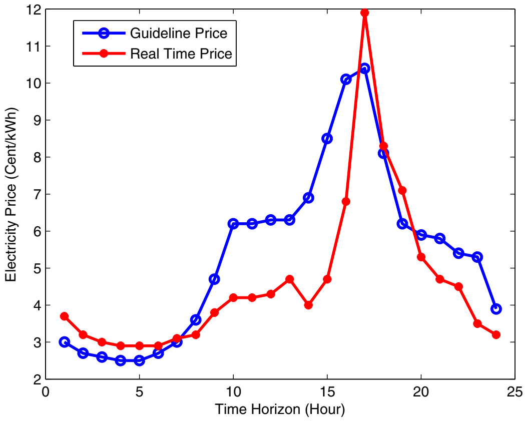

Fig. 2. Guideline and Real Time Electricity Price Provided by ComED [10].

(1) The total energy consumption of the home appliance m is equal to the required total energy consumption E_m. That is,

$$E_m = \sum_{h\in\mathcal{H}} x_{m,h} t_{m,h}. \tag{3}$$

(2) For a specified task, the home appliance m needs to be executed after the earliest start time α_m and before the deadline β_m such that

$$x_{m,h} = 0, \forall h < \alpha_m, h > \beta_m. \tag{4}$$

At each time slot h, the total energy load of the community is denoted by L_h, and thus

$$L_h = \sum_{n\in\mathcal{N}} (l^p_{n,h} + l^q_{n,h}). \tag{5}$$

2) Monetary Cost: As is mentioned before, there are two types of electricity price in the smart home system, which are guideline electricity price and real time electricity price respectively as depicted in Figure 2. The guideline electricity price is provided to the customers to facilitate smart home scheduling, while the real time electricity price is used in computing the bill.

In real time pricing, at each time slot the monetary cost of energy consumption depends on the total energy load of the grid. In this paper, the quadratic cost function is used to compute the total monetary cost of all the customers, which is a popular pricing model used in literatures [1]. Thus, the total monetary cost at time slot C_h is given as

$$C_h = a_h L_h^2 \tag{6}$$

where a_h is the pricing parameter which models the relationship between energy consumption and monetary cost. At each time slot, the monetary cost is distributed to each customer according to energy usage of the customer. For customer n, the monetary cost is denoted by $C_{n,h}$ where

$$C_{n,h} = (l^p_{n,h} + l^q_{n,h})\frac{C_h}{L_h} = a_h(l^p_{n,h} + l^q_{n,h})L_h. \tag{7}$$

B. Smart Home Scheduling

Given the above constraints, one can formulate a game where each customer n aims to minimize the individual monetary cost through solving the problem **P1** [8], [9].

$$\begin{aligned}
\textbf{P1} \quad & \underset{x_{m,h}\in\mathcal{X}_m,\forall h\in\mathcal{H},m\in\mathcal{Q}_n}{\textbf{minimize}} \sum_{h\in\mathcal{H}} a_h(l^p_{n,h} + l^q_{n,h})L_h \\
& \textbf{subject to} \quad l^p_{n,h} = \sum_{m\in\mathcal{P}_n} x_{m,h} t_{m,h} \\
& \qquad l^q_{n,h} = \sum_{m\in\mathcal{Q}_n} x_{m,h} t_{m,h} \\
& \qquad E_m = \sum_{h\in\mathcal{H}} x_{m,h} t_{m,h} \\
& \qquad L_h = \sum_{n\in\mathcal{N}} (l^p_{n,h} + l^q_{n,h}) \\
& \qquad x_{m,h} = 0, \forall h < \alpha_m, h > \beta_m
\end{aligned}$$

Since the monetary cost of each customer n depends on the total energy load of the community including all other customers, the scheduling of one customer has an impact on all other customers. This naturally leads to a game. The monetary cost of each customer n can be divided into two parts. We have

$$C_{n,h} = a_h(l^p_{n,h} + l^q_{n,h})(l^p_{n,h} + l^q_{n,h}) + a_h(l^p_{n,h} + l^q_{n,h})l_{-n,h}, \tag{8}$$

where $l_{-n,h}$ is the community wide energy load excluding the energy consumption of customer n at time slot h and

$$l_{-n,h} = \sum_{i\in\mathcal{N}, i\neq n} (l^p_{i,h} + l^q_{i,h}). \tag{9}$$

The game can be then formulated as follows [8].

Game Model:

- **Players:** All the customers in the system.
- **Payoff Function:** $P(l^q_{n,h}|l_{-n,h}) = -C_{n,h} = -a_h(l^p_{n,h} + l^q_{n,h})(l^p_{n,h} + l^q_{n,h}) - a_h(l^p_{n,h} + l^q_{n,h})l_{-n,h}$.
- **Shared Information:** $l_{-n,h}$.
- **Problem formulation:**

$$\begin{aligned}
\textbf{P1} \quad & \underset{x_{m,h}\in\mathcal{X}_m,\forall h\in\mathcal{H},m\in\mathcal{Q}_n}{\textbf{minimize}} \sum_{h\in\mathcal{H}} a_h(l^p_{n,h} + l^q_{n,h})L_h \\
& \textbf{subject to} \quad l^p_{n,h} = \sum_{m\in\mathcal{P}_n} x_{m,h} t_{m,h} \\
& \qquad l^q_{n,h} = \sum_{m\in\mathcal{Q}_n} x_{m,h} t_{m,h} \\
& \qquad E_m = \sum_{h\in\mathcal{H}} x_{m,h} t_{m,h} \\
& \qquad L_h = \sum_{n\in\mathcal{N}} (l^p_{n,h} + l^q_{n,h}) \\
& \qquad x_{m,h} = 0, \forall h < \alpha_m, h > \beta_m
\end{aligned}$$

In the game, each customer aims to minimize the total payment through scheduling the energy consumption of the automatically controlled home appliances. Nash Equilibrium is achieved when no one can further reduce his/her own monetary cost without changing the energy consumption of any other customer [11]. To solve this game, a decentralized technique is adopted from our previous work [9]. This is an iterative algorithm. In each iteration, each customer solves the Problem **P1** using the dynamic programming based algorithm while assuming the energy consumption of all the others, i.e., $l_{-n,h}$, is fixed [8], [9]. After each iteration, each customer obtains the new energy consumption information from all others according to their new scheduling solutions. With the new energy consumption of each customer, $l_{-n,h}$ is updated and each customer uses dynamic programming to schedule their home appliances with the updated

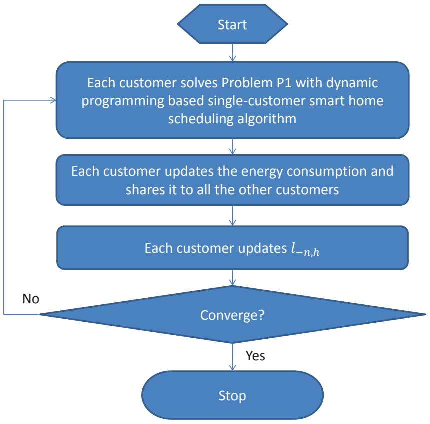
Start
Each customer solves Problem P1 with dynamic programming based single-customer smart home scheduling algorithm
Each customer updates the energy consumption and shares it to all the other customers
Each customer updates $l_{-n,h}$
No
Converge?
Yes
Stop

Fig. 3. Algorithmic Flow of Multiple Customer Smart Home Scheduling [8], [9].

$l_{-n,h}$. This process is iterated until convergence. Refer to Figure 3 for the algorithmic flow [8], [9].

III. Pricing Cyberattacks

Consider the communication infrastructure of the AMI in Figure 1. First, the utility transmits the pricing information to a central computer in the local community (substation) through Internet. Second, such pricing information is broadcast to smart meters through Internet or WiFi network or a combination of them, depending on the communication infrastructure of the local community. For example, in a hierarchical infrastructure where a community consists of multiple subcommunities, pricing information is forwarded to each subcommunity through Internet and is then broadcast inside the subcommunity through the WiFi network. In a WiFi network, there are some access points, which serve as the agents to receive the pricing information from the subcommunity and forward it to the smart meters inside the subcommunity.

The above popular infrastructure is vulnerable to at least three attacking strategies. First, one can directly hack the computer in the substation and modify the pricing information there. Subsequently, the pricing information forwarded to the whole community could be a faking one. Second, one can block an access point in the WiFi network using the jamming attack (i.e., sending excessive requests to the access point), create a fake access point and send the faking pricing information to the smart meters covered by the fake access point. Third, one can hack the smart meter and modify the pricing it receives. As is indicated from [12], "commercial smart devices including smart meters are often designed and manufactured utilizing off-the-shelf components and/or solutions. Therefore, security protection methods can only be applied at the application/network level and can hardly cover the hardware infrastructure. As a result, attackers can easily bypass firmware verification and install malicious OS kernel in the device to remotely control the smart device [13]". For example, an attacker can remotely manipulate the guideline pricing received at a smart meter without being realized. With the different difficulty levels in implementation, the cyberattacker can choose which one to use in practice.

A. Cyberattack for Bill Reduction

The first possible cyberattack is to fake the guideline pricing curve such that the utility bill of the cyberattacker can be reduced at the cost of bill increase of others in the community. Consider the following scenario. The guideline electricity price in the early morning such as 1:00am to 8:00am are usually not high due to limited amount of human activities. However, if a cyberattacker schedules a large load during this period, it could still be expensive. Therefore, if the cyberattacker fakes the guideline pricing curve such that the electricity price during 1:00am and 8:00am is very high, then almost no customers in the community will schedule energy during this period. Subsequently, the cyberattacker can schedule his/her own large load there, resulting in the significant reduction of his/her own bill. Of course, such a reduction comes from the increase of the bill of other customers. The procedure for the cyberattack for bill reduction using pricing manipulation is as follows.

(1) Determine the starting time t_s and end time t_e for the hacker to schedule his/her own energy load.

(2) Create the manipulated guideline price with a high price from t_s to t_e and low price at other time slots.

(3) Manipulate the guideline electricity prices received at the target smart meters.

(4) Schedule his/her own energy load from t_s to t_e.

When the guideline electricity price is high from t_s to t_e, the customers tend not to schedule the energy consumption there according to the smart home scheduling. This reduces the energy load during these time slots, and results in the decrease of the real time electricity price there. Subsequently, the cyberattack could schedule the energy consumption from t_s to t_e, and makes profit through saving his/her own bill at the cost of increasing the bill of other customers.

B. Cyberattack for Forming the Peak Energy Load

The second possible attack is to fake the guideline pricing curve such that a peak energy usage can be formed. Consider the following scenario. The guideline electricity price at 8:00pm is usually expensive since the utility discourages the excessive energy usage during this period which is typically occupied with various human activities (e.g., watching TV). If a cyberattacker creates a faking guideline pricing curve with very low price at this slot, significant amount of energy (e.g., laundry load) will be accumulated there. This will form a peak in the energy usage, which could significantly impact the power system stability. The procedure for the cyberattack for forming a peak energy load using pricing manipulation is as follows.

(1) Determine the starting time t_s and end time t_e of peak energy usage hours.

(2) Create the manipulated guideline electricity price such that it is very low from t_s to t_e.

(3) Manipulate the guideline electricity prices received at the target smart meters.

(4) A peak energy load will be formed from t_s to t_e.

If the guideline electricity price is very low from t_s to t_e, the customers tend to schedule large energy load there due to smart home scheduling. This increases the energy load during this time period and could potentially form a peak in energy consumption.

IV. Defense Technology

The above two pricing cyberattacks need to significantly perturb the guideline pricing curves. The key to design the countermeasure against them is to identify the guideline pricing manipulation. Machine learning and statical data analysis techniques would be natural choices since the typical guideline pricing curves should be similar to each other in a short term (such as several weeks). For example, as is shown in Figure 4 the electricity prices are similar in the three consecutive days according to a realistic industrial pricing study in [14]. Given the guideline electricity prices in the last few days, various techniques such as Supported Vector Regression (SVR), Hidden Markov Model and Neural Networks can be applied to predict the current guideline electricity price. Among them, we choose the popular SVR technique since it tends to be more robust than other techniques [15].

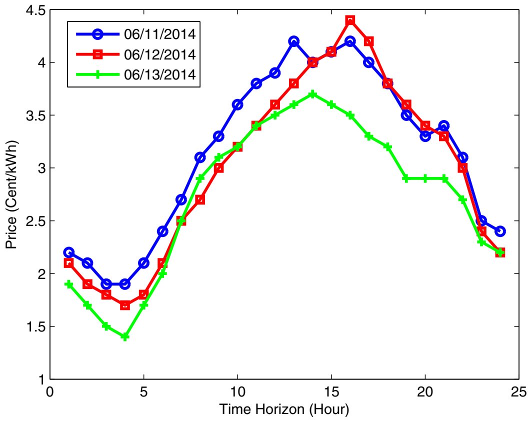

Fig. 4. Real Time Electricity Price from 06/11/2014 to 06/13/2014 Provided by Ameren Illinois [14].

After computing the predicted guideline price curve, one can compare it with the guideline pricing curve received at the smart meter. For comparison, one could compute the maximum difference between the two curves and signal an alert when it is larger than a threshold. However, the main problem with this idea is that it would be difficult to design a good threshold in practice. Note that the goal for setting a good threshold is to balance the false detection rate and the impact to the local power system and customers. For example, if one sets the threshold to be zero, then all pricing manipulation can be identified but there will be a lot of false detection. On the other hand, it is reasonable to ask what the best threshold is if customers can tolerate up to 1% bill increase? However, this question is difficult to answer.

In contrast, our philosophy is that since anyway the goal for setting a good threshold is to limit the impact to local community and customers, why not directly using the impact as the threshold? To explore this idea, we define a set of thresholds, called the maximum tolerable impact differences. An example maximum tolerable impact differences could be up to 2% increase in peak to average ratio (PAR) and up to 1% increase in bill. Given the predicted guideline pricing curve and the received guideline pricing curve, one can perform smart home scheduling simulations to compute the average bill and PAR for each of the two guideline pricing curves, and then compute the differences between them, called the actual impact differences. When the actual impact differences are larger than the maximum tolerable impact differences, a potential pricing attack is spotted and an alert signal will be sent to the utility to request for further check which could need some amount of human interaction. This work will develop such a framework.

Denoted by vectors $\mathbf{a}^p$ and $\mathbf{a}$ the predicted guideline electricity price and the received guideline electricity price, respectively. Denote by ΔB and ΔP the actual impact differences in bill and PAR, respectively. We also define two thresholds δ_B and δ_P which are the maximum tolerable impact differences in bill and PAR, respectively. If $\Delta B > \delta_B$ or $\Delta P > \delta_P$, the smart meter treats the guideline electricity price as being manipulated due to cyberattacks. Suppose that the guideline pricing of last T days, denoted by $\mathbf{A}$, are known. Adopt SVR to our problem context, the predicted guideline electricity pricing can be computed as follows.

- Denote the historical guideline electricity prices during the last T days by the matrix

$$\mathbf{A} = \begin{bmatrix} a_{1,1} & a_{1,2} & a_{1,3} & \dots & a_{1,H} \\ a_{2,1} & a_{2,2} & a_{2,3} & \dots & a_{2,H} \\ a_{3,1} & a_{3,2} & a_{3,3} & \dots & a_{3,H} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ a_{T,1} & a_{T,2} & a_{T,3} & \dots & a_{T,H} \end{bmatrix}. \tag{10}$$

  Denote by $\mathbf{a}_h = [a_{1,h}, a_{2,h}, \dots, a_{T,h}]$ the guideline electricity prices at time slot h during the last T days.
- We aim at computing the function $f(\mathbf{a}_h)$ which derives the predicted guideline electricity price at time slot h, denoted by a_h^p, from $\mathbf{a}_h$. That is,

$$a_h^p = f(\mathbf{a}_h). \tag{11}$$

- Given the historical guideline electricity price, the function $f(\mathbf{a})$ is computed through first solving the problem

$$\begin{aligned} \textbf{P2 maximize} & \sum_i^H u_i - \frac{1}{2} \sum_i^H \sum_j^H u_i u_j a_i a_j K(\mathbf{a}_i, \mathbf{a}_j) \\ \textbf{subject to} & \sum_i^H u_i a_i = 0 \\ & u_i \geq 0, \forall i \in \mathcal{H}, \end{aligned}$$

  where $K(\mathbf{a}_i, \mathbf{a}_j)$ is defined as kernel function [15]. In our implementation, radial basis kernel function is used such that

$$K(\mathbf{a}_i, \mathbf{a}_j) = \exp(-\gamma \times \|\mathbf{a}_i - \mathbf{a}_j\|^2). \tag{12}$$

  Obtaining u_i through solving Problem **P2**, the function $f(\mathbf{a})$ can be computed as

$$f(\mathbf{a}) = \sum_i^H u_i a_i K(\mathbf{a}_t, \mathbf{a}). \tag{13}$$

- After the function $f(\mathbf{a})$ is computed, the predicted guideline electricity price is derived according to Eqn. (11).

After computing the predicted guideline pricing curve $\mathbf{a}^p$, one can perform smart home scheduling simulations to compute the bill B_p and the PAR P_p due to using it. Similarly, one can compute the bill B and the PAR P due to using the received guideline pricing curve. Subsequently, the bill increase rate is computed as $\Delta B = \frac{B - B_p}{B_p}$ and the PAR increase is computed as $\Delta P = \frac{P - P_p}{P_p}$. If $\Delta B > \delta_B$ or $\Delta P > \delta_P$, an alert will be signaled. The proposed countermeasure technique for the guideline electricity pricing manipulation attack is summarized in Algorithm 1.

Algorithm 1 Electricity Pricing Manipulation Detection Algorithm

```
Obtain a set of historical guideline electricity prices $\mathbf{A}$.
Obtain the received guideline electricity price $\mathbf{a}$.
Compute the support vector regression model through solving Problem P2.
Compute the predicted guideline electricity price according to Eqn. (11).
Perform smart home scheduling simulations to compute the average bill $B_p$ and the PAR $P_p$ using the predicted guideline price.
Perform smart home scheduling simulations to compute the average bill $B$ and the PAR $P$ using the received guideline price.
Evaluate $\Delta B = \frac{B - B_p}{B_p}$ and $\Delta P = \frac{P - P_p}{P_p}$.
if $\Delta B > \delta_B$ or $\Delta P > \delta_P$ then
    Send an alert message to the utility.
end if
```

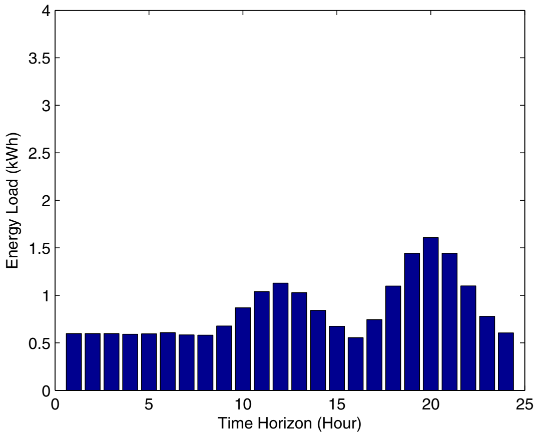

Fig. 5. Average Energy Load Created by Manually Controlled Home Appliances (Background Energy Load).

V. Simulation Results

We conduct simulations using MATLAB to analyze the impacts of electricity pricing manipulation cyberattacks. In the simulation setup, a community consisting of 500 customers is considered, and each customer is equipped with a smart meter connected with the AMI. For each customer, there are both manually controlled home appliances and automatically controlled home appliances. In our testcases, the average energy load due to manually controlled home appliances is shown in Figure 5. The range for the daily energy consumption and execution duration of the automatically controlled home appliances are shown in Table I, which is the same as our previous works [8]. The simulation duration is set to the next 24 hours divided into hourly time slots. The quadratic pricing model is used. Therefore, y axes in Figure 6(a), Figure 7(a), Figure 8(a) and Figure 10(a)(b) show the quadratic coefficients in pricing and one needs to multiple them by the energy load in the corresponding time slots to obtain unit price.

TABLE I
Daily Energy Consumption and Execution Duration of Automatically Controlled Home Appliances [8]

Home Appliance	Daily Consumption	Execution Duration
Washing Machine	1.2kWh-2kWh	0.5h-1.5h
Dish Washer	1.2kWh-2kWh	0.5h-1h
Cloth Dryer	1.5kWh-3kWh	0.5h-1.5h
EV	9kWh-12kWh	4h-8h
Air Conditioner	2kWh-3kWh	1h-3h
Heater	2kWh-3kWh	1h-3h

A. *Cyberattack for Bill Reduction*

In this simulation, the smart home scheduling results with the unattacked guideline electricity price and the attacked guideline electricity price are compared. For the cyberattack, the hacker manipulates the guideline price and create a peak price from 1:00 am to 6:00 am and makes the rest flat. Refer to Figure 6 for the results without attack and Figure 7 for the results with attack, respectively. We make the following observations.

- From Figure 6(a), one sees that the guideline price well matches the real time price, which is as expected.
- From Figure 6(b), one sees that the energy load is well distributed over the time horizon.
- In contrast, Figure 7(a) shows that the guideline price and real time price are significantly different. The reason is that when a smart home scheduler sees a high guideline price from 1:00 am to 6:00 am, it tends not to schedule the load there, which can be clearly seen from Figure 7(b). Note that there are still scheduled appliances during that time period due to (1) the background energy such as refrigerator and (2) the appliances which are required to be scheduled there due to starting time and ending time constraints. Due to the reduced energy usage from 1:00 am to 6:00 am, the real time price at these time slots are lower than what it should be. Since the quadratic pricing model is used, the unit price is computed as the multiplication of quadratic coefficient (y axis) and the energy load. From 1:00 am to 6:00 am, the unit price is $0.0812 per kWh without cyberattack and $0.0528 per kWh with cyberattack on average, which is a 34.3% reduction. Thus, if the hacker schedules his/her own load during this time period, a significant reduction in his/her own bill can be achieved.
- This bill reduction of the hacker comes from the bill increases of other customers. Using the unattacked guideline electricity price, each customer pays $3.82 on average. However, using the attacked guideline price, the average bill increases to $4.12, which is 7.9% higher.
- As a byproduct, the cyberattack will also impact the energy load balancing. Comparing Figure 6(b) and Figure 7(b), the peak to average ratio (PAR) of the energy load from unattacked guideline electricity price is 1.107 while it becomes 1.358 with the attacked guideline electricity price.

B. *Cyberattack for Forming a Peak Energy Load*

In this simulation, the smart home scheduling results with the unattacked guideline electricity price and the attacked guideline electricity price are compared. For this cyberattack, the hacker manipulates the guideline price and create a dip from 7:00 pm to 9:00 pm. Refer to Figure 6 for the results without attack and Figure 8 for the results with attack, respectively. We make the following observations.

- Figure 8(a) shows that the guideline price and real time price are significantly different. The reason is that when a smart home scheduler sees a low guideline price from 7:00 pm to 9:00 pm, it tends to schedule a large amount of load there, which can be clearly seen from Figure 8(b). The PAR of the energy load from unattacked guideline electricity price is 1.107 while it becomes 1.502 with the attacked guideline electricity price which is a 35.7% increase in PAR. This means that the cyberattacks can significantly unbalance the local energy load.
- Due to the peak energy usage from 7:00 pm to 9:00 pm, the real time price at these time slots are higher than what it should be. From 7:00 pm to 9:00 pm, after converting quadratic pricing to unit price, one can obtain that the unit price without cyberattack is $0.160 per kWh and with cyberattack is $0.111 per kWh on average, which is a 43.9% increase. The average daily bill of each customer is $4.02, which is 5.24% more.

C. *Countermeasure Technique*

In this simulation, the performance of our proposed countermeasure technique is evaluated. Refer to Figure 9. Given a set of the historical guideline electricity prices of last 7 days, a predicted guideline electricity price is computed using the support vector regression on these data. We have tested on various pricing cyberattacks. To present our simulation results, we assume that the hacker chooses to use a bill reduction cyberattack through increasing the guideline electricity price during some time slots. Refer to Figure 10 for the simulation results. The maximum tolerable impact differences are set as $\delta_B = 5\%$ for average bill increase and $\delta_P = 2\%$ for PAR increase, respectively. We make the following observations.

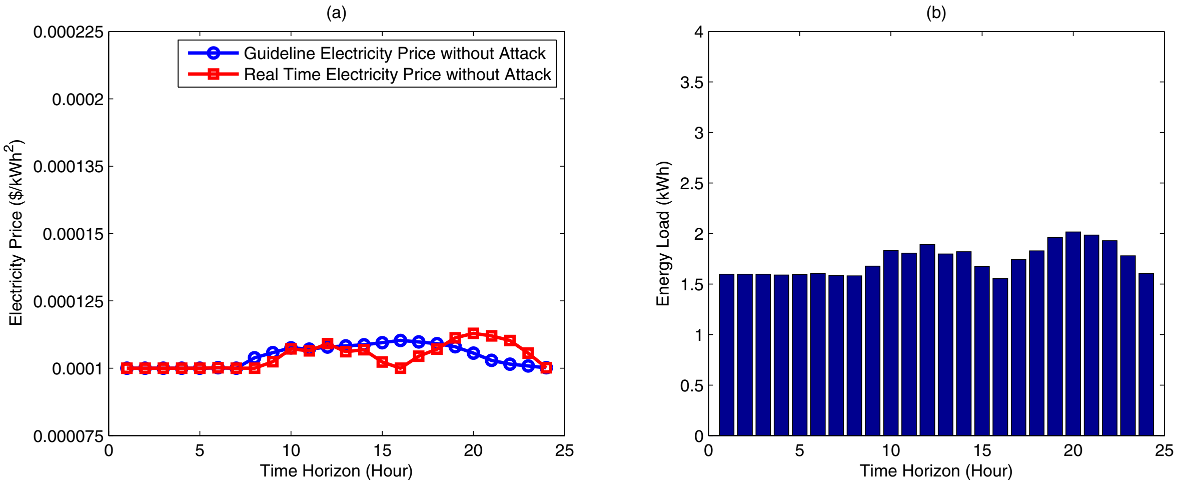

Fig. 6. Without Attack, (a) Guideline Electricity Price and Real Time Electricity Price, (b) Average Energy Load (PAR=1.107).

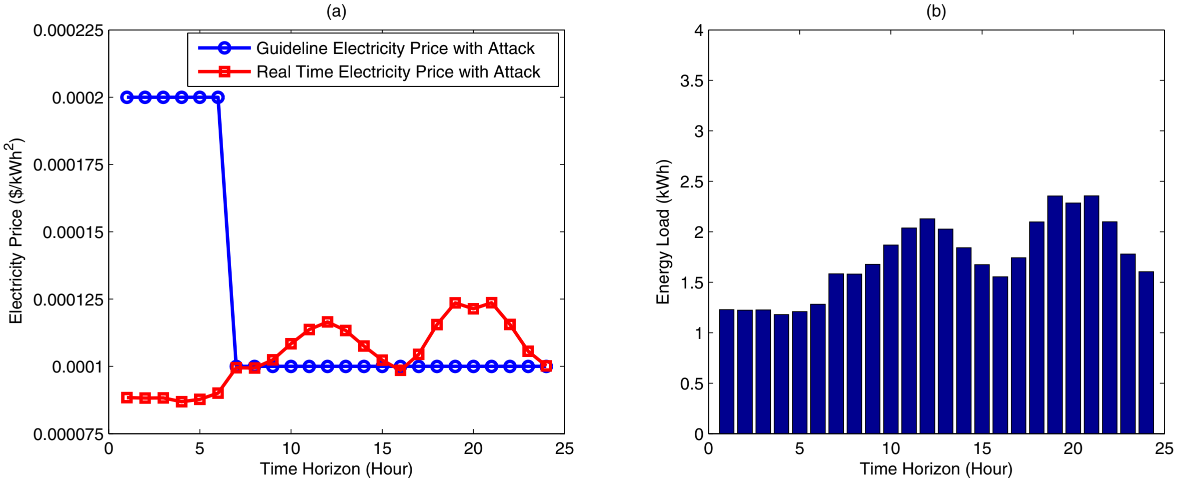

Fig. 7. With Pricing Cyberattack for Bill Reduction, (a) Guideline Electricity Price and Real Time Electricity Price, (b) Average Energy Load (PAR=1.358).

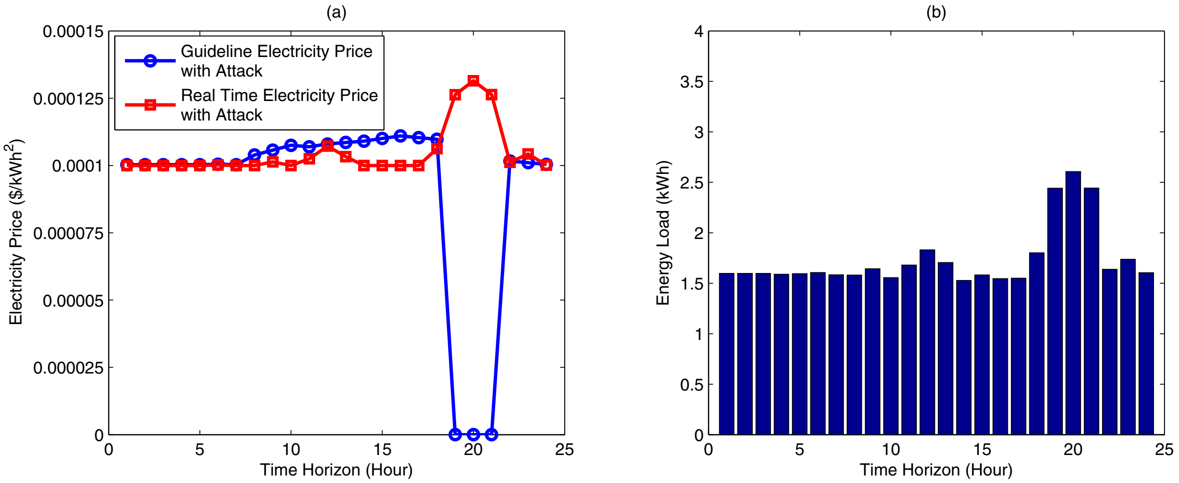

Fig. 8. With Pricing Cyberattack for Forming Peak Energy, (a) Guideline Electricity Price and Real Time Electricity Price, (b) Average Energy Load (PAR=1.502).

- When there is no pricing cyberattack, Figure 10(a) shows the received guideline electricity price and the predicted guideline electricity price. Comparing with the predicted electricity price, the average bill pay increase is -0.26%, smaller than δ_B and the PAR increase is -1.45%, smaller than δ_P. Thus, the received guideline electricity is regarded as normal.

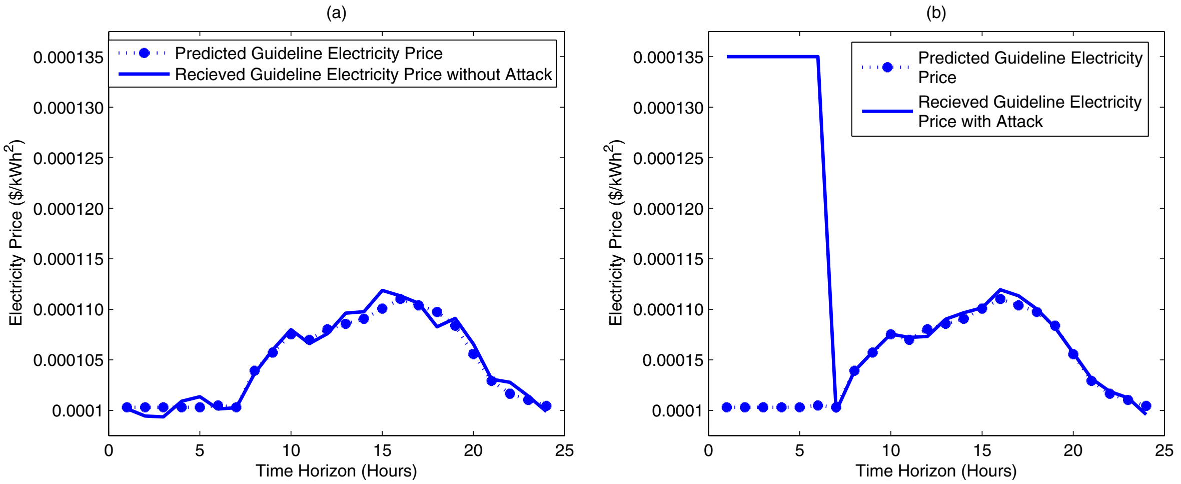

Fig. 10. (a) Without Cyberattack. For received guideline pricing, Bill=$3.83, PAR=1.170. and for predicted guideline pricing, Bill=$3.82, PAR=1.153, $\Delta B = -0.26\%$ and $\Delta P = -1.45\%$. (b) With Cyberattack. For received guideline pricing, Bill=$3.83, PAR=1.170. and for predicted guideline pricing, Bill=$4.09, PAR=1.203, $\Delta B = 6.79\%$ and $\Delta P = 2.82\%$.

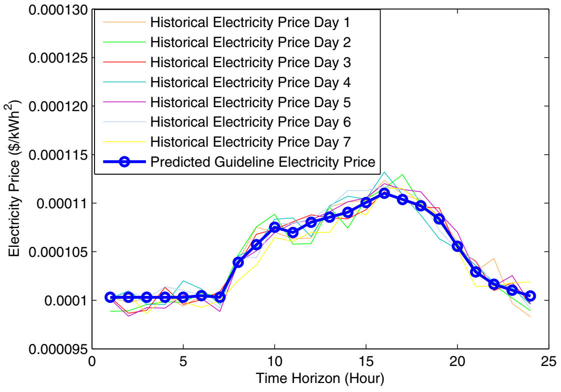

Fig. 9. Historical electricity price of the last 7 days and predicted guideline electricity price.

- When there is pricing cyberattack, Figure 10(b) shows the received guideline electricity price and the predicted guideline electricity price. Comparing with the predicted electricity price, the average bill increase is 6.79%, larger than δ_B and the PAR increase is 2.82%, larger than δ_P. In this case, the electricity pricing manipulation is detected.

VI. Conclusion

In this paper, the electricity pricing cyberattacks in the smart home system are considered. These cyberattacks aim at reducing the monetary expense of the cyberattacks and increasing the peak energy usage in the local power system. After analyzing them, a countermeasure technique based on support vector regression and maximum tolerable impact difference is proposed. Our simulation results demonstrate that the pricing cyberattack can reduce the attacker's bill by 34.3% at the cost of the increase of others' bill by 7.9% on average. In addition, the pricing cyberattack can significantly unbalance the energy profile of the local power system as it increases the peak to average ratio by 35.7%. Furthermore, our simulation results show that the proposed countermeasure technique can effectively detect the electricity pricing manipulation. The future work seeks to improve the pricing anomaly detection algorithm using more advanced machine learning techniques.

References

[1] A. Mohsenian-Rad, V. Wong, J. Jatskevich, R. Schober, and A. Leon-Garcia, "Autonomous demand-side management based on game-theoretic energy consumption scheduling for the future smart grid," *IEEE Transactions on Smart Grid*, vol. 1, no. 3, pp. 320–331, 2010.

[2] R. Khan and J. Khan, "A heterogeneous wimax-wlan network for ami communications in the smart grid," in *Proceedings of IEEE International Conference on Smart Grid Communications*, 2012, pp. 710–715.

[3] H. Li and Z. Han, "Manipulating the electricity power market via jamming the price signaling in smart grid," in *Proceedings of IEEE GLOBECOM Workshops*, 2011, pp. 1168–1172.

[4] L. Xie, Y. Mo, and B. Sinopoli, "False data injection attacks in electricity markets," in *Proceedings of IEEE International Conference on Smart Grid Communications*, 2010, pp. 226–231.

[5] O. Kosut, L. Jia, R. Thomas, and L. Tong, "Malicious data attacks on smart grid state estimation: Attack strategies and countermeasures," in *Proceedings of IEEE International Conference on Smart Grid Communications*, 2010, pp. 220–225.

[6] Z. Lu, W. Wang, and C. Wang, "From jammer to gambler modeling and detection of jamming attacks against time-critical traffic," in *Proceedings of IEEE INFOCOM*, 2011, pp. 1871–1879.

[7] H. Kim, R. Chitti, and J. Song, "Novel defense mechanism against data flooding attacks in wireless ad hoc networks," *IEEE Transaction on Consummer Electronics*, pp. 579–582, 2010.

[8] Y. Liu, S. Hu, and Z. Tian, "Analyze electricity market integrated with smart home scheduling," *Manuscript*.

[9] L. Liu, Y. Zhou, Y. Liu, and S. Hu, "Dynamic programming based game theoretic algorithm for economical multi-user smart home scheduling," in *Proceedings of IEEE Midwest Symposium on Circuits and Systems*, 2014.

[10] [Online]. Available: http://www.comed.com/

[11] J. Rosen, "Existence and uniqueness of equilibrium points for concave n-person games," *Econometrica*, vol. 33, no. 3, pp. 520–534, 1965.

[12] Y. Jin, "Personal communication."

[13] G. Hernandez, O. Arias, D. Buentello, and Y. Jin, "Smart nest thermostat: A smart spy in your home," in *Black Hat USA*, 2014.

[14] [Online]. Available: http://www.powersmartpricing.org/

[15] C.-C. Chang and C.-J. Lin, "LIBSVM: A library for support vector machines," *ACM Transactions on Intelligent Systems and Technology*, vol. 2, pp. 27:1–27:27, 2011.