

Inherent Vulnerability of Demand Response Optimisation against False Data Injection Attacks in Smart Grids

Thusitha Dayaratne, Carsten Rudolph, Ariel Liebman, Mahsa Salehi

Faculty of Information Technology

Monash University

Melbourne, Australia

{Thusitha.Dayaratne, Carsten.Rudolph, Ariel.Liebman, Mahsa.Salehi}@monash.edu

Abstract—The transition of energy networks to so-called smart grids benefits from advancements in Internet of Things technology. Energy management systems enable efficient and effective demand response (DR) schemes optimising load distribution. The increased user involvements through such DR schemes creates a new vector for false data injection attacks (FDIA), where authentic users themselves inject false data. Unlike in most existing FDIAs, no breaches to communication or devices are needed to execute this type of FDIA. In this work, we depict that this new FDIA can impact any optimisation-based DR scheme. Further, we show that adversaries achieve financial benefits independently from the actual algorithm used for optimisation, as long as they are able to inject false demand predictions. Compared to traditional FDIAs, reliable security mechanisms such as proper authentication, security protocols, security controls or sealed/controlled devices cannot prevent this new type of FDIA. Additionally, we show that there is no straightforward solution and we highlight the need for highly reliable FDIA detection mechanisms to thwart this type of attacks.

Index Terms—Demand response, False data injection attack, Inherent vulnerabilities, Smart grids

I. INTRODUCTION

Serving peak demand with the ever increasing energy usage is extortionate since utility companies/energy providers need to operate reserve generation to maintain the demand-supply balance. The formation of Smart Grids and its descendant technologies such as Demand Response (DR) schemes try to shape these peak demands to ensure the operating efficiency and cost-effectiveness for utility companies while reducing emissions. Enhancing traditional power grids with the use of Information and Communication Technology enables a range of improved operational functionalities. These include more convenient and efficient energy management, compared to traditional approaches. Many countries are augmenting existing power systems with Advanced Metering Infrastructures (AMI) which enables a two-way communication between utility providers and consumers as well as increase the overall system visibility. For example, the United Kingdom is planning to replace existing meters with Smart Meters (SM) by 2020 [1]. In addition, the Australian state of Victoria has completed a state-wide AMI roll-out of over two million SMs in 2016.

Thus, the basis for smart energy networks is being laid. As soon as technology on the demand side in the form of energy management systems follows, demand response optimisation can be practically implemented in a large scale.

Utility companies will then provide financial benefits/incentives to keep their users actively engaged in their DR systems. Commonly considered financial incentives include Real-Time Pricing (RTP), Day Ahead Pricing (DAP) and Time of Use (ToU) [2], [3]. These programs allow consumers to make more informed decisions about their usages and reduce their costs through reducing their peak demands, which in turn reduces the overall peak demand, helping utility companies to better utilize the existing grid infrastructure, obtain more accurate demand forecasts and reduce the need for increasing the reserved generation capacity. Users need to make intelligent decision to achieve the expected results rather than being passive participants in most of the DR systems. However, manually responding to the periodically (possibly every half an hour) changing incentive/pricing signal(s) to actively participate in DR systems is highly inconvenient to a user. Therefore, researchers have proposed distributed DR systems which use the Smart Meters (SMs) and Home Energy Management Systems (HEMSs) to assist consumers with deciding the best time for consuming electricity. These systems obtain information from the utility companies and schedule devices automatically for consumers in a way that satisfies consumption requirements and preferences of all customers and reduces their overall costs and the total peak demand [4], [5], [6].

Rapid advances in IoT-related technologies make a significant impact on distributed/decentralized DR schemes which increased the number of DR schemes proposed in recent years. These DR schemes are capable of producing optimal or close to optimal solutions given the fact that these schemes receive 100% accurate and reliable data. Therefore, data integrity is vital in such distributed DR systems. Having complete, accurate and timely data can guarantee that all consumers will obtain the maximum benefits from these systems. Preserving data integrity is a valid assumption if we have a 100% secure system where a single entity (usually utility company) is

responsible for producing the optimal solution on behalf of all its user without depending on individual users' inputs. However, for demand prediction, individual households need to actively interact with the systems to set their preferences with a potential for a high penetration with distributed DR schemes in the future. Hence, we cannot guarantee that all users will set their preferences truthfully and control their consumption honestly. Some users may act strategically to gain financial benefits, which may be impossible if they act upon rules and regulations of these DR schemes. Since any distributed DR schemes calculate solutions in a distributed manner unlike in strictly centralized schemes, depending on the assumption that the inputs are always accurate and reliable will not be practical with next generation distributed DR schemes.

In this work, we aim to construct a generic model to prove that independently of how well an algorithm can optimise the load distribution, any optimisation algorithm for distributed DR scheme which depends on user preferences to construct day ahead price signal, is susceptible to False Data Injection Attacks (FDIAs) and adversaries can obtain benefits, if the bills are calculated based on the actual consumption. In particular we make the following contributions through this paper:

- Build a generic model that abstracts DR optimisation in smart grids.
- Depict the inherent vulnerability of DR schemes against a new type of FDIA.
- Emphasize the lack of proper security mechanisms to hinder the impact of the inherent vulnerability of DR schemes.

The remainder of the paper is organized as follows. In section 2 we discuss the current state of the art in FDIA with respect to DR schemes and SG. Section 3 discusses the system model of the DR scheme and how the optimisation works. In section 4 we present the attack model together with our proof. Section 5 opens up a discussion about the attack and we conclude the paper in section 6.

II. BACKGROUND

Data integrity is one of the critical factors that guarantees proper functioning of SGs and DR systems up to their promises. However, FDIAs inject false data into these systems, which can eventually affect costs or availability. Although there is a no widely accepted definition for FDIA, the process of injecting false measurements or manipulating existing measurements that lead a system to take defective decisions or actions, which do not represent an accurate system state can be defined as an FDIA. These attacks can induce incorrect or sub-optimal power flows, excessive or insufficient generation, financial losses, and inconvenience for the users.

Motivated by the first work on FDIA on state estimation by Liu et al [7], an extensive amount of work has been carried out on the FDIA against state estimation throughout the years [8], [9], [10], [11], [12]. Tan et al. systematically

investigated the stability of an RTP system which they developed, based on control theory with respect to scaling and delay attacks [13]. Their simulation and experimental results showed that RTP algorithms are in a stable state unless the price signals scale down or more than half of the consumers get delayed price signals. Giraldo et al. extended and generalized that work further with the use of sensitivity analysis where they highlighted the importance of developing a robust algorithm to mitigate the impact of FDIA [14]. Barreto et al. modeled two strategic attackers in their work [15]. Modelled defraud attacker wants to obtain a lower price whereas the malicious attacker wants to disturb the stability of the system. Unlike in most other work, they considered both, a dynamic pricing model as well as direct load control (DLC) systems and concluded that dynamic pricing systems are more resilient towards data integrity attacks.

The impact of FDIAs on optimal power distribution [16], [17] in multistep electricity price schemes [18] has been also explored. Zhang et al and Yang et al showed that FDIAs can cause a significant impact on the welfare of the affected users on a DR scheme that they developed [19], [6]. Liu et al. showed that an adversary can increase the peak to average ratio of a DR system by 35.7% and gain about 34.3% of electricity bill reduction [20], [21]. A coordinated attack where the adversary manipulates both the price signal and the actual demand consumption values in a set of smart meters was explored in [22]. Weldehawaryat et al demonstrated that FDIAs can lead their decentralized demand management scheme to create a sub-optimal solution or force it not to converge [23].

However, with the existence of efficient encryption mechanisms, protecting price signals and measurement data can be effectively realised using standard components. Types of FDIAs, where adversaries require to manipulate communication channels or breach one or more entities in the system, can be hindered with these improved standard security controls. Nevertheless, as distributed DR systems getting popular, a significant portion of control of these distributed DR schemes is outside of the energy providers perimeter and depends on end users (e.g. the HEMSs or individual IoT devices). Hence, a novel FDIA type is emerging, where false data can be directly injected from these devices without the need to manipulate communication or hack into smart grid components. To the best of our knowledge, no work has analysed the impacts of this emerging type of FDIA or generalized the applicability of this FDIA for any distributed DR scheme in general. Therefore, this work addresses that gap by developing a model to analyse the impact of FDIAs on distributed DR schemes in general.

III. SYSTEM MODEL

A typical household contains both controllable and uncontrollable devices. Uncontrollable devices consume power throughout the day without being scheduled by energy management systems (e.g. refrigerator), while controllable devices consume power within given time-slots (e.g. washing machines, dishwashers, air conditioners). These controllable

devices can be optimised using distributed DR systems to reduce both the electricity generation cost for utility companies and the consumption cost for consumers. Scheduling devices optimally under these DR programs, such as the ones that use RTP, can effectively maximize the benefits that households and the utility provider get from a lower energy cost [24], [25]. A possible communication flow of a typical RTP-based device scheduling scheme that is commonly adopted in the distributed DR systems literature [26], [27], [5] is depicted in Figure 1.

For the following discussion, we assume a distributed DR scheme, which equally divides a day into P pricing slots. One day ahead of the actual energy consumption, the utility company/demand coordinator sends a guideline pricing signal for the 24 hours of the following day to all the SMs in the grid. The price signal consists of P price values where each value corresponds to the unit price for one pricing slot for some initially assumed demand. We index unit price values from 1 to p . Each household has a number of devices that need to be scheduled. The household owner configures the earliest start time, the latest finish time, preferred start time, device demand (power rate), running duration and the penalty factor for each device. The penalty factor models the inconvenience occurred to the owner if the device does not run at the preferred time. Setting up these configurations can be considered as a one-time operation unless the owner preferences change or they want to add new device. The HEMS locally schedules devices between the configured earliest start time and the latest finish time. If a device does not start at its preferred start time, there will be an associated cost (inconvenience/penalty) which is calculated based on the penalty factor. A device may run and finish within a single pricing slot or run for multiple slots. Table I shows the notation used throughout the paper.

Further we define the unit price of the p^{th} slot as a function of the aggregated demand of the same slot, where f is monotonically increasing.

$$f : ad_p \rightarrow up_p$$

The goal of optimising the demand distribution is to achieve the lowest overall costs. Thus, we can expect that the optimisation will remove peaks with high demand for particular time-slots, as long as flexibility of devices and penalty costs allow for a suitable re-distribution. For a generic discussion of this kind of optimization, we assume a perfect optimisation oracle that achieves an ideal distribution of demand.

For the formal definition of optimal demand distribution, we need to consider the following values.

The sum of the demand over all appliances scheduled to run during a pricing period p depicts the aggregated demand value of this pricing period (1).

$$ad_p = \sum_{h=1}^H \sum_{i=1}^{A_h} d_{h,i} \times S_{p(h,i)} \quad (1)$$

Total electricity bill of the community is calculated using (2) which is the summation of the product between each pricing

TABLE I
NOTATIONS

Symbol	Definition
H	Number of households, $H \in \mathbb{N}^+$
P	Number of pricing slots per day, $P \in \mathbb{N}^+$
p	Pricing slot index
A_h	Number of appliances in a household h
$a_{h,i}$	i^{th} appliance of household h
$es_{h,i}$	Earliest start time of i^{th} appliance of household h , $es_{h,i} \in \{1, 2, \dots, P\}$
$lf_{h,i}$	Latest finish time of i^{th} appliance of household h , $lf_{h,i} \in \{1, 2, \dots, P\}$
$ps_{h,i}$	Preferred start time of i^{th} appliance of household h , $ps_{h,i} \in \{1, 2, \dots, P\}$
$os_{h,i}$	Optimised start time of i^{th} appliance of household h , $os_{h,i} \in \{1, 2, \dots, P\}$
$rs_{h,i}$	Real start time of i^{th} appliance of household h , $rs_{h,i} \in \{1, 2, \dots, P\}$
$rt_{h,i}$	Number of pricing slots i^{th} appliance of household h runs, $rt_{h,i} \in \{1, 2, \dots, P\}$
$pf_{h,i}$	Penalty costs of i^{th} appliance of household; indicates the costs per pricing slot that an appliance is moved from the desired slot. h
$d_{h,i}$	Demand value of i^{th} appliance of household h
$S_{p(h,i)}$	Indicates, whether i^{th} appliance of household h is running on p^{th} pricing slot, $S_{p(h,i)} \in \{0, 1\}$
$ad_{p(\text{init})(X)}$	Initial (before optimising) aggregated demand value of p^{th} pricing slot in X scenario, $X \in \{\text{no-attack}, \text{with-attack}\}$
$ad_{p(\text{opt})(X)}$	Optimised aggregated demand value of p^{th} pricing slot in X scenario, $X \in \{\text{no-attack}, \text{with-attack}\}$
$ad_{p(\text{real})(X)}$	Real aggregated demand value of p^{th} pricing slot in X scenario, $X \in \{\text{no-attack}, \text{with-attack}\}$
$up_{p(\text{init})(X)}$	Initial unit price value of p^{th} pricing slot in X scenario, $X \in \{\text{no-attack}, \text{with-attack}\}$
$up_{p(\text{opt})(X)}$	Optimised unit price value of p^{th} pricing slot in X scenario, $X \in \{\text{no-attack}, \text{with-attack}\}$
$up_{p(\text{real})(X)}$	Real unit price value of p^{th} pricing slot in X scenario, $X \in \{\text{no-attack}, \text{with-attack}\}$

slot's aggregated demand ad_p and its correspondent unit price up_p .

$$Total\ bill = \sum_{p=1}^P ad_p \times up_p \quad (2)$$

The difference between preferred start time $ps_{h,i}$ and the optimised start time $os_{h,i}$ multiplied by the appliance's penalty costs $pf_{h,i}$ conceives the inconvenience of a given schedule for a given appliance to a user. Sum of all the inconvenience conceive the total penalty costs for the community which is (3).

$$Total\ penalty = \sum_{h=1}^H \sum_{i=1}^{A_h} |os_{h,i} - ps_{h,i}| \times pf_{h,i} \quad (3)$$

The optimisation will motivate the HEMSs to change $os_{h,i}$ of the devices with the ultimate goal to move the system towards minimizing the total cost (4).

$$min. Total\ cost = Total\ bill + Total\ penalty \quad (4)$$

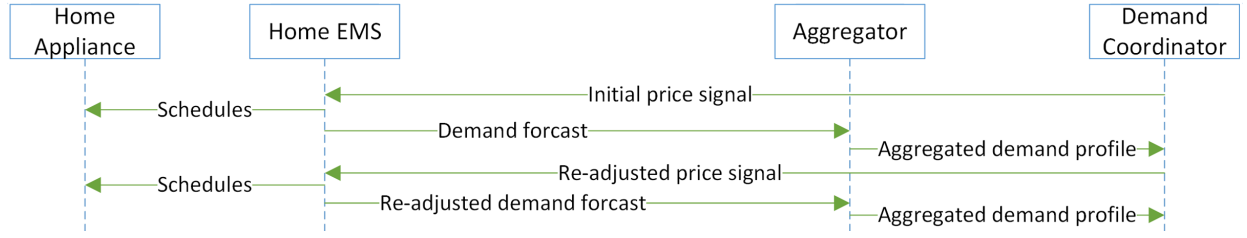


Fig. 1. Sequence of communication flow of a distributed DR scheme in general

Finding the minimum cost can be modeled as a convex optimisation problem and can be solved either using centralized or distributed approaches. For this discussion, the actual type of optimisation is not relevant. Instead, we assume a perfect optimisation oracle as depicted in Fig 2 which always provides the optimal demand distribution with respect to total costs for a given community. Input to the optimisation oracle consists of the initial demand distribution for P number of slots and the set of penalty costs. The oracle outputs optimal demand distribution for the P number of pricing slots.

We assume that the costs defined by the energy provider can be different for each pricing slot. However, for every individual pricing slot it will monotonically grow, either step-wise, linear or exponential.

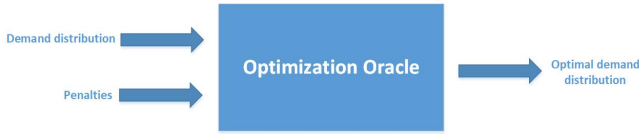


Fig. 2. optimisation oracle

A. Optimisation

This section explains in more detail the result of the demand schedule optimisation. For the simplicity, we restrict the explanation to the following case.

- 1) We identify 2 pricing slots p_i and p_j with $i \neq j$.
- 2) Some appliances may run during just one of the two pricing slots, for both pricing slots or not run at either of the two pricing slots.

We denote $ad_{i(init)}$, $ad_{j(init)}$, $ad_{i(opt)}$, and $ad_{j(opt)}$ as the initial aggregated demand of the i^{th} pricing slot, the initial aggregated demand of the j^{th} pricing slot, the optimised aggregated demand of the i^{th} pricing slot and the optimised aggregated demand of the j^{th} pricing slot respectively. Similarly, we denote respective unit price values as $up_{i(init)}$, $up_{j(init)}$, $up_{i(opt)}$, and $up_{j(opt)}$. Figure 3. depicts the initial unit price value variation with the aggregated demand.

Before the optimisation, *Total cost* is equal to the *total bill*, as no penalty needs to be applied. All appliances run at the initially scheduled time.

$$Total\ cost = Total\ bill = \sum_{p=1}^P ad_p \times up_p$$

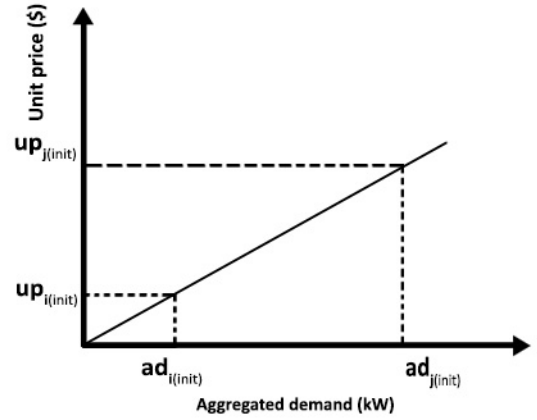


Fig. 3. Unit price variation with aggregated demand before the optimisation

The goal of the optimisation is to minimize the *Total cost*. Note that the total demand of the community remains the same, since the DR scheme is not curtailing any load from the system but rather adjusting the starting time of the devices. Now, as an example assume that the optimisation only moves load between pricing slots p_i and p_j . Then,

$$ad_{i(init)} + ad_{j(init)} = ad_{i(opt)} + ad_{j(opt)} \quad (5)$$

The oracle is responsible of calculating the aggregated demand values $ad_{i(opt)}$ and $ad_{j(opt)}$ such that *Total cost* is minimized. We assume that the oracle can find values for $ad_{i(opt)}$ and $ad_{j(opt)}$ by moving α : ($\alpha > 0$) amount of demand from the second pricing slot p_j to the first pricing slot during the optimisation as depicted in Figure 4

Then, the followings hold for the optimisation.

$$ad_{j(opt)} = ad_{j(init)} - \alpha$$

$$ad_{i(opt)} = ad_{i(init)} + \alpha \quad \because (5)$$

$$\therefore ad_{i(opt)} > ad_{i(init)} \text{ and } ad_{j(opt)} < ad_{j(init)}$$

Due to the monotonic growth of the price, we can conclude the following for the unit prices.

$$up_{i(opt)} > up_{i(init)} \text{ and } up_{j(opt)} < up_{j(init)}$$

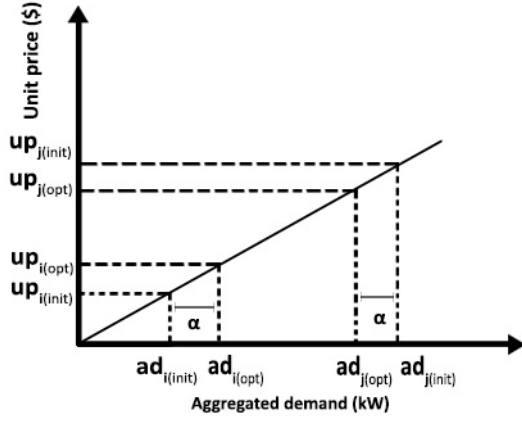


Fig. 4. Unit price variation with aggregated demand after the optimisation

Now let's consider the influence of this change to the total costs, as expressed by the following expression.

$$ad_i(init) \times up_i(init) + ad_j(init) \times up_j(init) - (ad_i(opt) \times up_i(opt) + ad_j(opt) \times up_j(opt) + Total\ penalty)$$

Replacing the optimised demand with the initial plus/minus the change α leads to this equivalent expression.

$$\Rightarrow ad_i(init) \times up_i(init) + ad_j(init) \times up_j(init) - ((ad_i(init) + \alpha) \times up_i(opt) + (ad_j(init) - \alpha) \times up_j(opt) + Total\ penalty)$$

This is equal to:

$$\Rightarrow ad_i(init) \times (up_i(init) - up_i(opt)) + ad_j(init) \times (up_j(init) - up_j(opt)) + \alpha \times (up_j(opt) - up_i(opt)) - Total\ penalty$$

The initial assumption was that the optimisation oracle will lead to the perfect result. Thus, it needs to provide the optimal values for $ad_i(opt)$, $ad_j(opt)$ and α for given unit prices $up_i(opt)$, $up_j(opt)$ so that the following statement remains true.

$$ad_i(init) \times (up_i(init) - up_i(opt)) + ad_j(init) \times (up_j(init) - up_j(opt)) + \alpha \times (up_j(opt) - up_i(opt)) - Total\ penalty > 0$$

Likewise, for the general case, we assume that the Oracle is capable of finding the optimal demand distribution for any given P pricing slots, which results as follows.

$$\sum_{p=1}^P ad_p(init) \times up_p(init) - (\sum_{p=1}^P ad_p(opt) \times up_p(opt) + Total\ penalty) > 0$$

s.t

$$\sum_{p=1}^P ad_p(init) = \sum_{p=1}^P ad_p(opt)$$

IV. ATTACK MODEL

In this section, we compare two scenarios, where the DR scheme executes with and without attack from an adversary's perspective. Let's assume that we have a DR scheme with a perfect optimisation oracle as we discussed in the previous section and all the communications between the nodes are reliable. Further, let's assume that all data is securely encrypted throughout the network to ensure that adversaries cannot manipulate or intercept the data. There can be adversaries who want to gain some financial advantages in this system. However, since the communication is perfectly secure, adversaries cannot intercept (and alter) the price signals or other demand profiles. Therefore, traditional FDIA approaches would not impact such a DR scheme. Nevertheless, as adversaries are also participants in the system, they can falsify data that they input into the DR scheme via smart devices or HEMSs. These can be either their own systems within their direct control or compromised systems located at another household.

Let's assume that our perfect Oracle uses either a linear or a quadratic unit price function. Therefore, the unit price value of a given pricing slot depends on the aggregated demand value of the particular pricing slot. When ad_p increases, up_p increases as well. It should be noted that a linear increase models a quite modest price increase. Demand peaks induce high costs for production and distribution as both need to be prepared to react quickly to demand changes and need to ensure that sufficient energy reserves are available. However, though the attack works for the linear, quadratic and step-wise cases, we only focus on the linear case where demand and price are proportional.

$$up_p \propto ad_p$$

For the optimisation as well as the attacks to work, we need a few more assumptions on the flexibility of demand (i.e. can sufficient demand be moved to other pricing slots) and on potential penalty costs (i.e. the gain in unit price is higher than the penalty). Note that these assumptions constitute some of the basic pre-conditions for DR-schemes to actually make a difference for the efficiency of energy distribution. More formally, we assume the following. Let α_{init} be the initial demand difference between two pricing slots i and j . We assume that there is a threshold demand value γ such that if $\alpha_{init} < \gamma$, the optimisation oracle will move demand such that the optimised difference α_{opt} is now smaller than γ . As the optimisation oracle minimises overall costs, this also expresses that the penalty for moving demand from the preferred pricing slot is not higher than the gain in energy costs through the lower unit prices.

1) Attack-free Scenario:

Let's denote $ad_p(init)(no-attack)$, $ad_p(opt)(no-attack)$ and $ad_{real}(no-attack)$ as the initial aggregated demand of the

p^{th} pricing slot before the optimisation, the optimised aggregated demand value for the p^{th} pricing slot after the optimisation and the real demand consumption of the p^{th} pricing slot respectively. Similarly, we denote unit price values $up_{p(\text{init})(\text{no-attack})}$, $up_{p(\text{opt})(\text{no-attack})}$ and $up_{\text{real}(\text{no-attack})}$. Further, $Total\ bill_{(\text{init})(\text{no-attack})}$, $Total\ bill_{(\text{opt})(\text{no-attack})}$ and $Total\ bill_{(\text{real})(\text{no-attack})}$ denote the initial total bill before the optimisation, optimised total bill after the optimisation and the real total bill respectively. It is worth noting that optimised price/demand/cost implies the values that are calculated via the optimisation oracle before the real occurrence, whereas real price/demand/cost implies the exact values that occur on the run.

Similar to the optimisation example, let assume that $P = \{2|i, j\}$ and the optimisation oracle moves α amount of demand from the j^{th} pricing slot to i^{th} pricing slot. Then, once the oracle converges to a solution, the optimised aggregated demand value of the j^{th} pricing slot would be computed as

$$ad_{j(\text{opt})(\text{attack-free})} = ad_{j(\text{init-attack-free})} - \alpha \quad (6)$$

If the system can preserve the integrity of data and all the users follow their predicted schedules, the optimised aggregated demand value and the unit price value of the j^{th} pricing slot that calculated by the Oracle, should be equal to the real demand caused by the consumption and the real unit price on the subsequent day respectively.

$$\begin{aligned} ad_{j(\text{opt})(\text{attack-free})} &= ad_{j(\text{real})(\text{attack-free})} \\ \Rightarrow up_{j(\text{opt})(\text{attack-free})} &= up_{j(\text{real})(\text{attack-free})} \end{aligned}$$

Similarly, $up_{i(\text{opt})(\text{attack-free})} = up_{i(\text{real})(\text{attack-free})}$ since $ad_{i(\text{opt})(\text{attack-free})} = ad_{i(\text{real})(\text{attack-free})}$. Hence, in an attack-free scenario users are able to receive the expected return.

2) Attack Scenario:

We assume that there is an adversary in the system who wants to run a particular appliance with high demand, during a selected time slot of the day and tries to trick the system into reducing the unit price for this time slot. As a result of the lower unit price, a successful attack can reduce the usage cost for an adversary. The adversary can execute the attack by injecting some fake demand into the system. This fake demand can be introduced by different avenues, based on the adversary's attack capacity. Some simplest options are

- 1) Add a (fake) appliance into their HEMS
- 2) Increase the demand for an existing appliance
- 3) Collaborate with a set of other users to inject the demand

Let's assume that in order to carry out the attack, the adversary injects some false demand into the j^{th} pricing slot such that

$$ad_{j(\text{init-with-attack})} = ad_{j(\text{init-attack-free})} + \phi \quad (7)$$

where ϕ ($\phi > 0$) is the amount of false demand, and everything else remains same as the attack-free scenario.

Then from (7)

$$ad_{j(\text{init-with-attack})} > ad_{j(\text{init-attack-free})} \because \phi > 0$$

$$\Rightarrow up_{j(\text{init-with-attack})} > up_{j(\text{init-attack-free})}$$

Since $ad_{j(\text{init-with-attack})} > ad_{j(\text{init-attack-free})}$,

$$\sum_{\forall P} ad_{p(\text{with-attack})} > \sum_{\forall P} ad_{p(\text{attack-free})}$$

$$\therefore Total\ bill_{(\text{opt})(\text{with-attack})} > Total\ bill_{(\text{opt})(\text{attack-free})}$$

However, the optimisation oracle tries to minimize the $Total\ bill_{(\text{opt})(\text{with-attack})}$ as much as possible. If the adversary keeps maintaining the false demand ϕ in the j^{th} slot, the optimisation must clearly move some of the appliances that are scheduled in the j^{th} pricing slot to slots with lower demand (this would be i^{th} slot according to our case) to reduce the $ad_{j(\text{opt})(\text{with-attack})} \times up_{j(\text{opt})(\text{with-attack})}$ which results a minimized $Total\ bill_{(\text{opt})(\text{with-attack})}$. Note that, as explained above, we assume the penalty for flexibility is smaller than the benefit in unit price. If this was not the case, the overall costs could not be used as a target for optimisation, as high penalties would remove the incentive to re-schedule demand.

Before the optimisation

$$ad_{p(\text{init-attack-free})} < ad_{p(\text{init-with-attack})}$$

Lets assume the oracle moves β amount of demand from j^{th} pricing slot to i^{th} slot. Therefore once the oracle optimised the distribution

$$ad_{j(\text{opt})(\text{with-attack})} = ad_{j(\text{init-with-attack})} - \beta \quad (8)$$

From 7 and 8

$$ad_{j(\text{opt})(\text{with-attack})} = ad_{j(\text{init-attack-free})} + \phi - \beta \quad (9)$$

Case 1: When $ad_{j(\text{opt})(\text{with-attack})} < ad_{j(\text{opt})(\text{attack-free})}$

If we consider the real consumption, ϕ amount of demand does not occur in the real time. Therefore

$$ad_{j(\text{real})(\text{with-attack})} = ad_{j(\text{opt})(\text{with-attack})} - \phi$$

$$\Rightarrow ad_{j(\text{real})(\text{with-attack})} + \phi < ad_{j(\text{opt})(\text{attack-free})}$$

$$ad_{j(\text{real})(\text{with-attack})} < ad_{j(\text{opt})(\text{attack-free})} \because \phi > 0$$

$$\Rightarrow ad_{j(\text{real})(\text{with-attack})} < ad_{j(\text{real})(\text{attack-free})}$$

$$\therefore ad_{j(\text{opt})(\text{attack-free})} = ad_{j(\text{real})(\text{attack-free})}$$

$$\therefore up_{j(real)(with-attack)} < up_{j(real)(attack-free)}$$

This implies that the adversary can receive a cheaper unit price for the j^{th} pricing slot compared to the attack-free scenario, by adding some false demand into the system.

Case 2: When $ad_{j(opt)(with-attack)} = ad_{j(opt)(attack-free)}$
Similar to the previous case,

$$ad_{j(real)(with-attack)} + \phi = ad_{j(opt)(attack-free)}$$

$$\therefore ad_{j(real)(with-attack)} < ad_{j(opt)(attack-free)}$$

$$\Rightarrow ad_{j(real)(with-attack)} < ad_{j(real)(attack-free)} \\ \therefore ad_{j(opt)(attack-free)} = ad_{j(real)(attack-free)}$$

$$\therefore up_{j(real)(with-attack)} < up_{j(real)(attack-free)}$$

Case 3: When $ad_{j(opt)(with-attack)} > ad_{j(opt)(attack-free)}$

$$ad_{j(real)(with-attack)} + \phi > ad_{j(opt)(attack-free)}$$

Clearly, in this case the attacker would not get an advantage. However, the attacker can increase the amount of ϕ to be bigger than the assumed threshold γ , such that the oracle will move more appliances away from the j^{th} slot (note that we have assumed perfect optimisation and sufficient flexible devices and low penalties). Therefore, above a certain threshold value of ϕ , we will be back to case one and $ad_{j(real)(with-attack)} < ad_{j(opt)(attack-free)}$

$$\Rightarrow ad_{j(real)(with-attack)} < ad_{j(real)(attack-free)} \\ \therefore ad_{j(opt)(attack-free)} = ad_{j(real)(attack-free)}$$

$$\therefore up_{j(real)(with-attack)} < up_{j(real)(attack-free)}$$

V. DISCUSSION

Smart grids and its descendent concepts such as DR schemes are yet to be fully realized by both utility providers and the consumers. However, if they are, financial incentives can drive misbehaviour within these highly distributed infrastructures. Further, boundaries of these systems cannot be rigidly defined or controlled. Outside of the control of energy providers and grid perimeter, consumers have the total control over their own home automation and energy management systems. These system play a significant role in DR schemes. Consequently, consumer preferences and demand forecasts which get inputted into the DR schemes via these HEMSs are entirely up to the consumers and the utility providers only has indirect influence via the price signal.

The FDIA strategy described here does not involve any tampering or compromise of the communication channel. Hence, the attack can be executed by any user in the community without much effort. The optimisation oracle always tries to produce an optimal demand distribution by optimising the overall costs. Due to the higher unit price for peak demand, optimising overall costs will flatten the peaks in the demand distribution as much as possible. However, the impact of optimisation highly depends on how flexible the demands are in the system. When an adversary injects false demand into the system, the injected false demand causes an increase in the aggregated demand of the attacked time-slot(s). When the aggregated demand of a time-slot increases, a higher unit price is applied for that slot. Hence, the costs of the households who have scheduled devices in the attacked time-slot(s) increases. The increase in their costs triggers their HEMSs to move some of the devices that have been scheduled in that time-slot(s) to other slots to reduce the total cost. However, the real demand consumption of the attacked time-slot(s) will always be lower than the expected optimised demand value, since the Oracle moves some of the devices away from the attacked time-slot(s). Therefore, if an adversary is capable of injecting false demand into the DR system and the existing demand are flexible enough, the adversary is capable of shifting some of the existing demand into other time-slots, independent from the optimisation algorithm the Oracle uses. Accordingly, the real unit price(s) for the attacked time-slot(s) will be lower than those in the convergent solution. Hence, the adversary can gain a lower usage cost and less inconvenience by running their devices in the attacked time-slot(s).

In our previous work [28] we have demonstrated how a strategic adversary could execute an FDIA on RTP based DR scheme to obtain cost reduction for their usage without any technical skills. Our experimental results showed that a small percentage of overall false demand injection can result a significant cost reduction for the adversary, orthogonal to the number of households in the community. Adversary's benefits can further increases with the penetration of smart house concept where more appliances can be used as controllable appliances in the future. As the proof shows, this attack is totally independent of the actual optimisation algorithm used. Therefore, additional controls are required to prevent these attacks or at least to induce a proper reaction. Obviously, utility companies can use reliable security mechanisms such as proper authentication, security protocols, security controls or sealed/ controlled devices to remediate the impact of traditional FDIAs where attackers need to tamper the nodes or broadcast (demand, price) signals. However, these security controls do not influence the type of new FDIAs where the false data originates from the systems which are outside the control of the energy provider. Even if all data from HEMSs is properly authenticated towards the DR scheme, it can still be fake demand, as the attack exploits attacker's access to the data source. The assumption that proper authentication, security protocols, security controls or sealed/ controlled devices can provide security against FDIAs represents a common miscon-

ception concerning the security of smart energy networks.

As cryptography, secure transmission, digital signatures, perimeter protection etc. cannot prevent this attack, a straightforward approach would be to use anomaly detection to identify the false demand. Though there are many works which use machine learning techniques to identify FDIAs in the literature, no work focus on using machine learning techniques to detect false demand injection. Hence, we did some experiments to infer the most effective machine learning technique(s) to identify FDIA. We have extracted the energy consumption data of 168 houses in Austin, Texas from the Pecanstreet dataset [29]. The extracted data contains minute-wise energy consumption of different appliances running on houses for three months period between June to August (summer) 2017. We considered an electric car, washer, combined washer dryer, electric dryer, gas dryer, dishwasher, jacuzzi, pool 1, pool 2, pool pump and pump as flexible devices and all other appliances as inflexible. Several supervised (neural network, k-nearest-neighbor classifier, proximity forest) and unsupervised techniques (k-means clustering, local outlier factor, k-nearest-neighbor) have been applied to the data-sets in the first set of experiments. The detection accuracy was close to 99% when the false injection percentage is high. However, detection accuracy decreases when the false data injection percentage gets lower. Hence, straightforward application of existing detection algorithms might not be reliable in this type of FDIAs. Further experiments are needed, but the initial results show that there is no guarantee that user preferences remain consistent due to the inherently chaotic nature of energy usages. These chaotic consumption patterns makes anomaly detection arduous. Consumer preferences (ultimately become demand patterns) get affected not only due to the seasonality but also by unpredictable circumstances such as special events (birthday parties, special celebrations), their social status and sudden requirements (e.g. wash a heavy load after a trip). Furthermore, a considerable cost advantage can potentially be achieved with rather low percentage of fake demand. Therefore, anomaly detection might need to take into account the probability that a demand pattern is part of a particular attack strategy. Future research needs to explore options for smart anomaly detection in the smart grid environment.

Additionally, the price reduction that a typical user receives via an effective and efficient DR scheme is much higher compared to the increase resulting from a single attack. Thus, the probability of consumers noticing the attack impact can be considered relatively low. However, the attacker can gain a significant amount of bill reduction collectively, if the attacker is capable of executing the attack for several days without being detected.

Finally, another reaction could be to introduce penalties if there is a high deviation of real demand from the predicted demand. If the DR scheme provides the predicted demand to the energy provider, this deviation can be easily detected. Such a penalty could de-incentivise any FDIAs targeting unit prices. Thus, in principle, such a penalty should effectively prevent such attacks. However, smart devices and home energy

management systems sit in private home networks and are often not very secure. If an adversary can exploit some vulnerabilities in a HEMS or in appliances, false data can be directed on behalf of another home. The attack would still work, but the attribution of the attack to the attacker is no longer possible. Then the adversary is still able to receive the benefit of the attack while the compromised users have to pay the penalties. Nevertheless, isolating and restricting data from individual HEMSs, which causing anomalous inputs, can be used as primitive protection against FDIAs, despite, it will not be effective if the adversary is capable of distributing false demand among several HEMSs, where the injection from a single HEMS is minimum.

VI. CONCLUSION

The FDIA discussed in this paper is probably not the most sophisticated way to attack smart grids. Nevertheless, as soon as there are financial incentives, people will be motivated to attack these systems on various levels. This attack should be seen as examples for various types of similar attacks. If FDIAs are applied on the level of energy markets, effects can be much worse and go beyond the electricity bill of a few households.

While this particular FDIA obviously should be considered in DR implementations, in the long run we should take this example to create a wider view on security issues in distributed infrastructures. We need to understand business processes, markets and financial consequences of algorithmic behaviour in order to properly evaluate if false data injected or created somewhere in the distributed system of systems can disrupt these processes or lead to harm or financial losses.

REFERENCES

- [1] Foreign & Commonwealth Office and Sir David King, "The UK aims to have smart meters in place by 2020" - GOV.UK, 2016. [Online]. Available: <https://www.gov.uk/government/speeches/the-uk-aims-to-have-smart-meters-in-place-by-2020>
- [2] M. H. Albadi and E. F. El-Saadany, "Demand response in electricity markets: An overview," pp. 1–5, Jun. 2007.
- [3] C. Eid, E. Koliou, M. Valles, J. Reneses, and R. Hakvoort, "Time-based pricing and electricity demand response: Existing barriers and next steps," *Utilities Policy*, vol. 40, pp. 15 – 25, 2016.
- [4] S. He, A. Liebman, A. Rendl, M. Wallace, and C. Wilson, "Modelling RTP-based residential load scheduling for demand response in smart grids," in *International Workshop on Constraint Modelling and Reformulation*, 2014, pp. 36–51.
- [5] S. Mhanna, A. C. Chapman, and G. Verbic, "A fast distributed algorithm for large-scale demand response aggregation," *CoRR*, vol. abs/1603.00149, 2016.
- [6] X. Zhang, X. Yang, J. Lin, G. Xu, and W. Yu, "On Data Integrity Attacks Against Real-Time Pricing in Energy-Based Cyber-Physical Systems," *IEEE Transactions on Parallel and Distributed Systems*, vol. 28, no. 1, pp. 170–187, 2017.
- [7] Y. Liu, P. Ning, and M. K. Reiter, "False data injection attacks against state estimation in electric power grids," *Ccs*, vol. 14, no. 1, pp. 1–33, 2009.
- [8] S. Amini, F. Pasqualetti, and H. Mohsenian-Rad, "Dynamic Load Altering Attacks Against Power System Stability: Attack Models and Protection Schemes," *IEEE Transactions on Smart Grid*, pp. 1–1, 2016.
- [9] L. Jia, R. J. Thomas, and L. Tong, "Impacts of malicious data on real-time price of electricity market operations," in *Proceedings of the Annual Hawaii International Conference on System Sciences*, 2012, pp. 1907–1914.

- [10] J. Kim, L. Tong, and R. J. Thomas, "Dynamic attacks on power systems economic dispatch," *2014 48th Asilomar Conference on Signals, Systems and Computers*, pp. 345–349, 2014.
- [11] O. Kosut, L. Jia, R. J. Thomas, and L. Tong, "Malicious data attacks on the smart grid," *IEEE Transactions on Smart Grid*, vol. 2, no. 4, pp. 645–658, 2011.
- [12] L. Xie, Y. Mo, and B. Sinopoli, "Integrity data attacks in power market operations," *IEEE Transactions on Smart Grid*, vol. 2, no. 4, pp. 659–666, 2011.
- [13] R. Tan, V. B. Krishna, D. K. Y. Yau, and Z. Kalbarczyk, "Integrity Attacks on Real-Time Pricing in Electric Power Grids," *ACM Transactions on Information and System Security*, vol. 18, no. 2, pp. 1–33, 2015.
- [14] J. Giraldo, A. Cardenas, and N. Quijano, "Integrity Attacks on Real-Time Pricing in Smart Grids: Impact and Countermeasures," *IEEE Transactions on Smart Grid*, vol. 8, no. 5, pp. 2249–2257, 2017.
- [15] C. Barreto, A. Cárdenas, N. Quijano, and E. Mojica-Nava, "CPS: Market analysis of attacks against demand response in the smart grid," in *ACM International Conference Proceeding Series*, vol. 2014-Decem, 2014.
- [16] Y. Abdallah, Z. Zheng, N. B. Shroff, H. E. Gamal, and T. M. El-Fouly, "The Impact of Stealthy Attacks on Smart Grid Performance: Tradeoffs and Implications," *IEEE Transactions on Control of Network Systems*, vol. 4, no. 4, pp. 886–898, 2017.
- [17] J. Lin, W. Yu, X. Yang, G. Xu, and W. Zhao, "On false data injection attacks against distributed energy routing in smart grid," in *Proceedings - 2012 IEEE/ACM 3rd International Conference on Cyber-Physical Systems, ICCPS 2012*, 2012, pp. 183–192.
- [18] J. Lin, W. Yu, and X. Yang, "On false data injection attack against multi-step electricity price in electricity market in smart grid," in *GLOBECOM - IEEE Global Telecommunications Conference*, 2013, pp. 760–765.
- [19] X. Yang, X. Zhang, J. Lin, W. Yu, X. Fu, and W. Zhao, "Data integrity attacks against the distributed real-time pricing in the smart grid," *2016 IEEE 35th International Performance Computing and Communications Conference (IPCCC)*, pp. 1–8, 2016.
- [20] Y. Liu, S. Hu, and T. Y. Ho, "Vulnerability assessment and defense technology for smart home cybersecurity considering pricing cyberattacks," in *IEEE/ACM International Conference on Computer-Aided Design, Digest of Technical Papers, ICCAD*, vol. 2015-Janua, no. January, 2015, pp. 183–190.
- [21] L. Yang, H. Shiyan, and H. T. Yi, "Leveraging Strategic Detection Techniques for Smart Home Pricing Cyberattacks," *IEEE Transactions on Dependable and Secure Computing*, vol. 13, no. 2, pp. 220–235, 2016.
- [22] Y. Liu, Y. Zhou, and S. Hu, "Combating Coordinated Pricing Cyberattack and Energy Theft in Smart Home Cyber-Physical Systems," *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, 2017.
- [23] G. K. Weldehawaryat, P. L. Ambassa, A. M. Marufu, S. D. Wolthusen, and A. V. Kayem, "Decentralised scheduling of power consumption in micro-grids: Optimisation and security," in *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, vol. 10166 LNCS, 2017, pp. 69–86.
- [24] C. Joe-wong, S. Member, S. Sen, and S. Ha, "Optimized Day-Ahead Pricing for Smart Grids with Device-Specific Scheduling Flexibility," *IEEE JOURNAL ON SELECTED AREAS IN COMMUNICATIONS*, vol. 30, no. 6, pp. 1075–1085, 2012.
- [25] Y. Ozturk, S. Member, D. Senthilkumar, S. Kumar, S. Member, G. Lee, and S. Member, "An Intelligent Home Energy Management System to," *IEEE Transactions on Smart Grid*, vol. 4, no. 2, pp. 694–701, 2013.
- [26] S. He, M. Wallace, G. Gange, A. Liebman, and C. Wilson, "A fast and scalable algorithm for scheduling large numbers of devices under real-time pricing," in *Principles and Practice of Constraint Programming*, J. Hooker, Ed. Cham: Springer International Publishing, 2018, pp. 649–666.
- [27] N. Li, L. Chen, and S. H. Low, "Optimal demand response based on utility maximization in power networks," in *Power and Energy Society General Meeting, 2011 IEEE*, Jul. 2011, pp. 1–8.
- [28] T. Dayaratne, C. Rudolph, A. Liebman, M. Salehi, and S. He, "High impact false data injection attack against real-time pricing in smart grids," in *2019 IEEE PES Innovative Smart Grid Technologies Europe (ISGT-Europe)*. IEEE, pp. 1–5.
- [29] "Pecan Street Inc. Dataport 2019." [Online]. Available: <https://dataport.cloud/>