2019 American Control Conference (ACC)
Philadelphia, PA, USA, July 10-12, 2019

Learning-based Attack Schedule against Remote State Estimation in Cyber-Physical Systems

Xiaolin Wang, Jianping He, Shanying Zhu, Cailian Chen and Xinping Guan

***Abstract*—Malicious attack against remote state estimation in cyber-physical systems has attracted considerable attention. Nevertheless, most existing works assume that the attacker is powerful and has mastered the communication information when designing the attack strategy. In this paper, we consider that the attacker is energy-limited and has no prior knowledge of transmission pattern. To solve this problem, we introduce a learning-based method for the attacker, which consists of a learning phase and an attack phase, to achieve a smart attack. We first formulate an optimal attack schedule problem, aiming to maximize the estimation error while considering the trade-off between the learning accuracy and attack efficiency. Since it is hard to solve this problem directly, we split it into two subproblems: i) optimizing the attack pattern; ii) optimizing the eavesdropping times and attack times. Theoretically, we prove that the optimal attack pattern is that the learning phase precedes the grouped attack from the viewpoint of possibility. Furthermore, we propose an algorithm to design the rational learning times and attack times for the attacker. Numerical examples are used to demonstrate the effectiveness of the proposed method.**

***Index Terms*—Cyber-physical systems security, remote state estimation, denial-of-service, energy constraint, smart attack**

I. Introduction

Cyber-Physical Systems (CPS) integrate sensing, communication, computing and control technologies to seamlessly connect the information space with the physical world [1]–[3]. The introduction of wireless communication networks makes the CPS very vulnerable to malicious attacks. For example, the massive power blackouts in Brazil and the attack by the StuxNet computer worm on the Iranian nuclear facility both caused huge losses and harms [4]. Security issues of CPS have been investigated in numerous works from various perspectives. Some focus on the impact on the system performance caused by different attacks, e.g., Denial-of-Service (DoS) attacks [5] and deception attacks, and design the optimal attack strategies [6]. Others study the detection and prevention policies against different attacks [7], [8].

Considering the security problem on the remote state estimation of CPS, many efforts have been devoted to investigating the optimal attack strategy to degrade the system performance. For example, Guo et al. [9] propose an undetected linear deception attack strategy to successfully inject false data. [10] analyzes the impact of bad data injection attacks on the electricity market, and studies the strategy of both the attacker and the defender using a game-theoretic framework. Some works investigate the scheduling problem for the DoS attack. Zhang et al. [11] provide the optimal DoS attack schedule with energy constraint, in which the expected average estimation error is maximized. This work only concentrates on the case of a perfect single channel. Qin et al. [12] further study the optimal attack schedule over the packet-dropping network. In [13], the authors formulate a two-player zero-sum stochastic game framework to model interactive decision making process between the attacker and the sensor over multi-channel network, and present a Nash Q-learning algorithm to obtain the optimal schedule for both sides.

In practical applications, the attacker often knows little about system information. However, in [9], [10], the assumptions that the attacker is powerful and has mastered prior knowledge of communication pattern are often hard to meet. Meanwhile, in [11], [12], the process of information acquisition is ignored. Although in [13], the attacker obtains the system information through trial and error, it is energy-consuming. With energy constraint, the attacker expects to get better attack benefits from the beginning, the trial-and-error method is undesirable. Therefore, it is necessary to take the learning process into account when designing attack schedules. The work [14] considers the learning process to maximize the number of successful attacks to reduce network throughput. But it focuses on the communication problems and assumes the distribution of the transmission period to be learned.

In this paper, a more practical and smart attack is investigated, where the DoS attacker has limited energy budget and no prior knowledge of transmission pattern. Data transmission between the sensor and the remote estimator is achieved via the multi-path wireless communication networks. Owing to inadequate capacity, the attacker needs to eavesdrop the communication process to improve attack effectiveness. We consider the learning process and aim to obtain the optimal attack schedule under energy constraint. The main contributions are summarized as follows:

- To the best of our knowledge, this is the first time to consider the learning process and the scenario of multi-path routing transmission in the design of the optimal attack schedule against remote state estimation.
- We establish the optimal attack schedule problem that considers the trade-off between the learning accuracy and the attack efficiency. We prove that the optimal attack pattern is the learning phase precedes the grouped attack, in the sense of possibility degree. Besides, a novel algorithm is proposed to design the rational eavesdropping times and the attack times based on the relationship of the eavesdropping error and the eavesdropping times.

This paper is organized as follows. Problem formulation is reported in Section II. In Section III, some preliminaries are

978-1-5386-7926-5/$31.00 ©2019 AACC

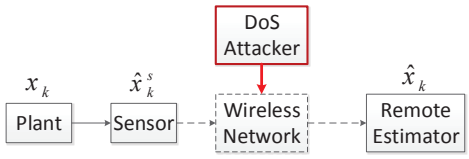

Fig. 1: System architecture.

provided. The optimal attack pattern in the sense of possibility degree is derived and an algorithm is proposed to design the eavesdropping times. Numerical examples are shown in Section IV. Section V concludes this paper.

II. Problem Formulation

A. System Model

Consider the following plant modeled by a linear time-invariant system:

$$x_{k+1} = Ax_k + w_k, y_k = Cx_k + v_k, \tag{1}$$

where $k \in \mathbb{N}$, $x_k \in \mathbb{R}^n$ is the system state at time k, $y_k \in \mathbb{R}^m$ is the sensor measurement, $w_k \in \mathbb{R}^n$ and $v_k \in \mathbb{R}^m$ are zero-mean i.i.d. Gaussian random noises with $\mathbb{E}[w_k w_j'] = \delta_{kj} Q(Q \geq 0)$, $\mathbb{E}[v_k v_j'] = \delta_{kj} R(R > 0)$, and $\mathbb{E}[w_k v_j'] = 0 \ \forall j, k \in \mathbb{N}$. The pair (A, C) is assumed to be detectable and $(A, Q^{1/2})$ is controllable. The sensor is capable of computing data after monitoring the plant, and runs a Kalman filter to estimate the system state x_k. Define the minimum mean squared error (MMSE) estimate as

$$\hat{x}^s_k = \mathbb{E}[x_k | y_1, y_2, ..., y_k],$$

the corresponding error covariance is

$$P^s_k = \mathbb{E}\left[(x_k - \hat{x}^s_k)(x_k - \hat{x}^s_k)' | y_1, y_2, ..., y_k\right].$$

The sensor sends the local estimate $\hat{x}^s_k$ to the remote estimator via wireless communication networks. The DoS attacker launches attacks on wireless networks to block the data transmission, as shown in Fig. 1.

B. Communication Model

The sensor transmits the state estimate $\hat{x}^s_k$ through the wireless mesh networks, where each node acts as a router to communicate with other nodes. Considering from the network layer, there are d routing paths for the sensor to transmit the data to the remote estimator, see Fig. 2. Based on some metrics (for instance trust level, Quality of Service request or energy-hops metrics [15]), only a single routing path is selected to forward the packet at each time instant k. Assuming that the wireless networks' routing structure is deterministic, while each routing path will be selected probabilistically for data transmission. Suppose that the transmission probability of each routing path is $p_i, i = 1, 2, ..., d$.

We assume that the communication routing path is perfect in the absence of attack, i.e., data can be successfully transmitted to the remote estimator without any attack [11].

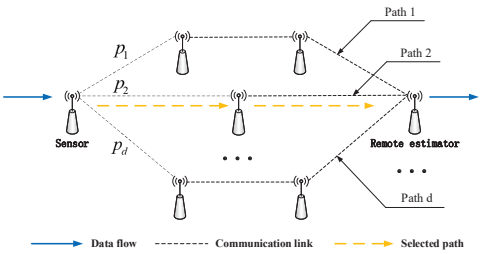

Fig. 2: Routing path.

C. Attacker Model

The attacker's goal is to obtain the maximum attack benefits. Due to energy constraint, it attacks one routing path at each time instant. There are two operational modes for the attacker, eavesdropping phase and attack phase.

- Learning-based eavesdropping. The attacker eavesdrops the communication process between the sensor and the remote estimator to learn the transmission probability of d routing paths simultaneously. Then, the attacker obtains the corresponding estimated transmission probability of each routing path, $\hat{p}_1, \hat{p}_2, ..., \hat{p}_d$.
- Attack the path with the largest transmission probability. After acquiring the estimated probability, the attacker attacks the routing path with the largest estimated probability $\hat{p}_{sel}$, which is the estimation of p_{sel}.

Let $\theta_k \in \{0, 1\}$ represent whether the attacker attacks the routing path at time step k. $\theta_k = 1$ denotes the attacker conducts attacking. $\theta_k = 0$ denotes the attacker is in the eavesdropping phase. Suppose that once the attacker successfully intercepts the packet, the packet will be lost. Therefore, after eavesdropping times $\tilde{n}_0$, the data drop probability completely depends on the transmission probability of the chosen routing path $p_{sel}(\tilde{n}_0)$.

Let E_t denote the attacker's energy budget. The energy consumption of eavesdropping and the energy consumption of attack at every time instant is defined as E_l and E_a, respectively. The total eavesdropping times is defined as $n_0 \in \mathbb{N}^*$ and the total attack times is represented by $\sum \theta_k = n \in \mathbb{N}^*$. We derive the mathematical expression of energy constraint,

$$E_t = E_l \cdot n_0 + E_a \cdot n. \tag{2}$$

D. Problem Formulation

From the viewpoint of the attacker, our object is to design the optimal attack schedule. It's obvious that the more times spending on eavesdropping, the more precise for the attacker to block the routing path with the largest probability, which causes more damage. However, under energy constraint, longer eavesdropping times will decrease the attack opportunities, resulting in a decline in the attack benefits. Thus, there is a trade-off between the eavesdropping times and attack times.

The problem of deriving the optimal attack schedule that maximizes attack benefits $F = Tr[J_A(\theta)]$ with energy con-

straint is formulated by

$$\mathbf{P_0}: \quad \max_{\theta\in\Theta} Tr[J_A(\theta)]$$
$$s.t. \quad \sum_{k=1}^{T}\theta_k = n$$
$$E_t = E_l\cdot n_0 + E_a\cdot n,$$

where $J_A(\theta) = \frac{1}{T}\sum_{k=1}^{T}\mathbb{E}[P_k(\theta)]$ is the average expected state estimation error covariance, $\Theta = \{\theta|\theta_k \in \{0,1\}\}$ means the attack schedule space, and $T = n_0 + n$.

Different from the traditional attacker strategy, our problem takes the learning process into consideration. It is interesting and worth exploring to design the attack schedule under energy constraint, considering the trade-off between the learning accuracy and attack efficiency. However, due to the attacker's inadequate capacity, the transmission probability is unknown, which makes $\mathbf{P_0}$ difficult to solve.

III. Main results

In this section, we first give some preliminaries. We next analyze the impact of the introduction of the learning process, and split $\mathbf{P_0}$ into two subproblems. Then, we obtain the optimal attack pattern in the sense of possibility degree. Finally, we propose an algorithm to design the eavesdropping times and attack times.

A. Preliminaries

In this subsection, we introduce some properties and preliminaries.

Definition 1 ([16]). *For arbitrary intervals $a = [a^L, a^U]$ and $b = [b^L, b^U]$, define $P(a \geq b)$ as*

$$P(a\geq b) = \min\{1, \max(\frac{a^U - b^L}{a^U - a^L + b^U - b^L}, 0)\}. \quad (3)$$

We call $P(a \geq b)$ possibility degree of $a \geq b$. Meanwhile, we have $P(a \geq b) + P(b \geq a) = 1$.

Lemma 1 ([17]). *If $P(a \geq b) > P(b \geq a)$, we have that in the sense of possibility degree, $a \geq b$.*

At the sensor side, the MMSE estimate is actually the Kalman filter. As the Kalman filter converges from any initial condition exponentially fast [18], we assume that the error covariance has converged to a steady-state value $P_k^s = \bar{P},\ \forall k \in \mathbb{N}$, which is calculated by $\tilde{g}\circ h(X) = X$. Then define the Lyapunov and Riccati operators h and $\tilde{g}: \mathbb{S}_+^n \to \mathbb{S}_+^n$ as $h(X) \triangleq AXA' + Q,\ \tilde{g}(X) \triangleq X - XC'(CXC' + R)^{-1}CX$. And $h^i(X) \triangleq \underbrace{h\circ h\circ\cdots\circ h(X)}_{\text{i times}}$. The property of function h is $\bar{P} \leq h(\bar{P}) \leq h(\bar{P})^2 \leq \cdots \leq h(\bar{P})^i \leq \cdots, \forall i \in \mathbb{Z}^+$ [18].

B. Transform into Two Subproblems

Since the eavesdropping process makes our problem more complicated, we first give a lemma to have an insight into the learning phase.

Lemma 2. *After a period of eavesdropping times $\tilde{n}_0$, given the number of routing paths d and the confidence level β, the relationship between p_{max} and $p_{sel}(\tilde{n}_0)$ is*

$$p_{max} - p_{sel}(\tilde{n}_0) \leq 2\varepsilon \leq 2\sqrt{\ln(\frac{1-\beta}{2d})/-2\tilde{n}_0}, \quad (4)$$

where $p_{max} = \max[p_1, ..., p_d]$ is the largest transmission probability, ε denotes the deviation error.

Proof: We use A_{im} to represent whether the packet is transmitted in path i and $A_{im} \in \{0,1\}, m = 1, ..., \tilde{n}_0$ obeys Bernoulli distribution (i.e. $Pr(A_{im} = 1) = p_i, Pr(A_{im} = 0) = 1 - p_i$). According to the Hoeffding's inequality [19], for a single path i we have

$$Pr(|\hat{p}_i - p_i| > \varepsilon) \leq 2\exp(-2\tilde{n}_0\varepsilon^2). \quad (5)$$

We generalize the inequality to d routing paths and derive

$$Pr(\forall \text{ path } i,\ |\hat{p}_i - p_i| \leq \varepsilon) \geq 1 - 2d\exp(-2\tilde{n}_0\varepsilon^2). \quad (6)$$

The probability represents the confidence level β of the confidence interval $p_i \pm \varepsilon$. Suppose β and $\tilde{n}_0$ is fixed, we obtain

$$\varepsilon \leq \sqrt{\ln(\frac{1-\beta}{2d})/-2\tilde{n}_0}. \quad (7)$$

Based on equation (6), we have $p_{max} \leq \hat{p}_{max} + \varepsilon \leq \hat{p}_{sel} + \varepsilon \leq p_{sel}(\tilde{n}_0) + 2\varepsilon$, the proof is completed. ■

Let $\varepsilon^*(\tilde{n}_0) = \sqrt{\ln(\frac{1-\beta}{2d})/-2\tilde{n}_0}$, as $\tilde{n}_0$ increases, $\varepsilon^*(\tilde{n}_0)$ decreases. Thus, the learning accuracy is improved, which affects the attack benefits. Meanwhile, the attacker's decision at each time instant, i.e., whether the attack pattern is decentralized or aggregated, also affects the attack benefits. Therefore, problem $\mathbf{P_0}$ needs to decide the optimal attack pattern and design the optimal eavesdropping times. Because these two decisions are coupled together and we can only obtain the bound of the learning deviation error, $\mathbf{P_0}$ is difficult to solve directly. Thus we split it into two subproblems, $\mathbf{P_1}$ and $\mathbf{P_2}$. Specifically, the first subproblem $\mathbf{P_1}$ is to determine the optimal attack pattern under the given but unknown eavesdropping times n_0 and attack times n, i.e.,

$$\mathbf{P_1}: \quad \max\ Tr[J_A(\theta_k)]. \quad (8)$$

Then, based on the optimal solution of $\mathbf{P_1}$, the second subproblem $\mathbf{P_2}$ is to design the optimal eavesdropping times that maximizes the attack benefits under energy constraint:

$$\mathbf{P_2}: \quad \max_{\mathbf{n_0}\in N^*} Tr[J_A(n_0)]$$
$$s.t. \quad E_t = E_l\cdot n_0 + E_a\cdot n. \quad (9)$$

By solving the above two subproblems, we derive the optimal attack pattern and design the eavesdropping times and attack times for the problem $\mathbf{P_0}$.

C. Optimal Attack Pattern

In this subsection, the optimal attack pattern in the sense of possibility degree is obtained to solve $\mathbf{P_1}$. For the given but unknown eavesdropping times n_0 and attack times n, we have the following theorem.

Theorem 1. *From the perspective of possibility degree, the following attack pattern is optimal:*

$$(\underbrace{0,0,\cdots,0,0}_{n_0 \text{ times}}\underbrace{1,1,\cdots,1,1}_{n \text{ times}}).$$

Under this optimal attack pattern, the expression of $J_A(\theta)$ is

$$J_A(\theta) = \bar{P} + \frac{1}{T}\sum_{z=1}^{n}(n-z+1)p_{sel}^{z}(n_0)[h^z(\bar{P}) - h^{z-1}(\bar{P})].$$

Proof: Firstly, we need to obtain the specific expression of $J_A(\theta)$ under the general case for sequential analysis. Let $\lambda_k \in \{0,1\}$ denote whether the packet drops at time instant k, i.e., $\lambda_k = 1$ if the packet drops, $\lambda_k = 0$ otherwise. Assuming that λ_k's are i.i.d. Bernoulli random variables. At the remote estimator side, the state estimation $\hat{x}_k$ and corresponding error covariance P_k satisfy the recursion:

$$\hat{x}_k = \begin{cases} A\hat{x}_{k-1}, & \text{if } \lambda_k = 1 \text{ and } \theta_k = 1, \\ \hat{x}_k^s, & \text{otherwise,} \end{cases} \tag{10}$$

$$P_k = \begin{cases} h(P_{k-1}), & \text{if } \lambda_k = 1 \text{ and } \theta_k = 1, \\ \bar{P}, & \text{otherwise.} \end{cases} \tag{11}$$

We consider the following attack strategy, the attack times n is divided into a number of consecutive attack sequences: c_1, c_2, $\cdots$, c_r, and corresponding eavesdropping sequences are l_1, l_2, $\cdots$, l_r, and $\sum_{j=1}^{r} c_j = n$, $\sum_{j=1}^{r} l_j = n_0$, i.e.,

$$(\underbrace{0,\cdots,0}_{l_1 \text{ times}}\underbrace{1,\cdots,1}_{c_1 \text{ times}}\underbrace{0,\cdots,0}_{l_2 \text{ times}}\underbrace{1,\cdots,1}_{c_2 \text{ times}}\underbrace{0,\cdots,0}_{l_r \text{ times}}\underbrace{1,\cdots,1}_{c_r \text{ times}}).$$

Define the variable $C_j = \sum_{i=1}^{j} c_i, L_j = \sum_{i=1}^{j} l_i$, and $C_0 = 0, L_0 = 0$. Note that $p_{sel}(L_j), j = 1,2,...,r$ is the transmission probability of the routing path that selected to attack after L_j eavesdropping times. Considering the attacker launches c_1 consecutive attacks at the time period $[a_1+1, a_1+c_1]$, $a_j = L_j + C_{j-1}$ is the attack start sequence. Assuming that at time a_1 the data is received at the remote estimator, i.e., $P_{a_1} = \bar{P}$. Based on [11] and above analysis, the distribution of P_{a_1+k}, $k \in [a_1+1, a_1+c_1]$ is

$$Pr\{P_{a_1+k} = h^j(\bar{P})\} = \begin{cases} p_{sel}^{k}(L_1), & j = k; \\ p_{sel}^{j}(L_1) - p_{sel}^{j+1}(L_1), & j \le k-1. \end{cases}$$

Now we derive the specific expression of $J_A(\theta)$,

$$\begin{aligned} J_A(\theta) &= \frac{1}{T}\sum_{j=1}^{r}\sum_{m=1}^{c_j}\mathbb{E}[P_{a_j+m}] + \frac{T-n}{T}\bar{P} \\ &= \bar{P} + \frac{1}{T}\sum_{j=1}^{r}\sum_{m=1}^{c_j}\sum_{z=1}^{m} p_{sel}^{z}(L_j)\left[h^z(\bar{P}) - h^{z-1}(\bar{P})\right]. \end{aligned} \tag{12}$$

Secondly, based on Lemma 2, we obtain the value range of $p_{sel}(L_j)$ as $p_{sel}(L_j) = [p_{max} - 2\varepsilon^*(L_j), p_{max}]$. From (7), we have $\varepsilon^*(L_1) > \varepsilon^*(L_2) > ... > \varepsilon^*(L_r) = \varepsilon^*(n_0)$. Thus, the lower bound of the value range of $p_{sel}(L_j)$ is changed.

Both $p_{sel}(L_j)$ and the attack pattern have effects on $J_A(\theta)$, for the convenience of demonstration, we firstly fix $p_{sel}(L_j)$ to a random value $p_{sel}(\tilde{n}_0)$. Define $\tilde{J}_A(\theta) = \bar{P} + \frac{1}{T}\sum_{j=1}^{r}\sum_{m=1}^{c_j}\sum_{z=1}^{m} p_{sel}^{z}(\tilde{n}_0)\left[h^z(\bar{P}) - h^{z-1}(\bar{P})\right]$, then we have

$$\begin{aligned} \tilde{J}_A(\theta) &\le \bar{P} + \frac{1}{T}\sum_{j=1}^{r}\sum_{m=1}^{c_j}\sum_{z=1}^{m+C_{j-1}} p_{sel}^{z}(\tilde{n}_0)\left[h^z(\bar{P}) - h^{z-1}(\bar{P})\right] \\ &= \bar{P} + \frac{1}{T}\sum_{z=1}^{n}(n-z+1)p_{sel}^{z}(\tilde{n}_0)[h^z(\bar{P}) - h^{z-1}(\bar{P})]. \end{aligned} \tag{13}$$

Based on Lemma 1, from the perspective of possibility degree we have

$$\begin{aligned} &\bar{P} + \frac{1}{T}\sum_{z=1}^{n}(n-z+1)p_{sel}^{z}(\tilde{n}_0)[h^z(\bar{P}) - h^{z-1}(\bar{P})] \\ &\le \bar{P} + \frac{1}{T}\sum_{z=1}^{n}(n-z+1)p_{sel}^{z}(n_0)[h^z(\bar{P}) - h^{z-1}(\bar{P})]. \end{aligned}$$

Similarly, in the sense of possibility degree, for $J_A(\theta)$ we have

$$J_A(\theta) \le \bar{P} + \frac{1}{T}\sum_{j=1}^{r}\sum_{m=1}^{c_j}\sum_{z=1}^{m} p_{sel}^{z}(n_0)\left[h^z(\bar{P}) - h^{z-1}(\bar{P})\right].$$

From (13), we also have

$$\begin{aligned} &\bar{P} + \frac{1}{T}\sum_{j=1}^{r}\sum_{m=1}^{c_j}\sum_{z=1}^{m} p_{sel}^{z}(n_0)\left[h^z(\bar{P}) - h^{z-1}(\bar{P})\right] \\ &\le \bar{P} + \frac{1}{T}\sum_{z=1}^{n}(n-z+1)p_{sel}^{z}(n_0)[h^z(\bar{P}) - h^{z-1}(\bar{P})]. \end{aligned} \tag{14}$$

Therefore, the optimal attack pattern is to use the first n_0 times for eavesdropping and the last n times for attack, in the sense of possibility degree. ■

In the above analysis, we prove the optimal attack pattern in the sense of possibility degree which solved $\mathbf{P_1}$.

Corollary 1. *If n_0 is long enough, $\varepsilon(n_0)$ approaches 0. According to Lemma 2, $p_{sel}(n_0) = p_{max}$, the attacker could be able to attack the one with the largest transmission probability. Thus we have $p_{sel}(n_0) > p_{sel}(L_j), j = 1,2,...,r-1$. The optimal attacking pattern is learning before grouped attack in the deterministic sense.*

D. Design Eavesdropping Times and Attack Times

In this subsection, we solve the second subproblem $\mathbf{P_2}$ under the optimal attack pattern of $\mathbf{P_1}$. We present the main idea of designing an iterative algorithm to transform the problem $\mathbf{P_2}$ of optimizing the eavesdropping times into finding the eavesdropping times in the worst case. In the expression of J_A given in Theorem 1, the transmission probability is unknown, thus $\mathbf{P_2}$ cannot be solved directly. We approximate the optimal solution by reducing the difference between the attack benefits

Algorithm 1 Design Eavesdropping Times

Input: N, β, d, η

Initialize: $j = 3$.
for $i = 1$ to N **do**
  $n_0 = j$, $n = N - n_0/3$, $\varepsilon_1^* = \sqrt{\frac{\ln(\frac{1-\beta}{2d})}{-2n_0}}$;
  compute $\Delta F_1 = \min\{F(p_{max} - 2\varepsilon_1^*) - F(p_{max})\}$;
  $n_0 = j + 3$, $n = N - n_0/3$, $\varepsilon_2^* = \sqrt{\frac{\ln(\frac{1-\beta}{2d})}{-2n_0}}$;
  compute $\Delta F_2 = \min\{F(p_{max} - 2\varepsilon_2^*) - F(p_{max})\}$;
  $j = j + 3$;
  if $\Delta F_2/\Delta F_1 > \eta$ **then**
    $n_0^* = n_0$;
  end if
end for

Output: eavesdropping times n_0^*.

generated by attacking the optimal path with p_{max}, and the attack benefits generated by attacking the selected path with $\hat{p}_{sel}$. In this way, the eavesdropping times is obtained.

Firstly, we formulate the problem of maximizing the difference of attack benefits $\Delta F = F(p_{max}) - F(p_{sel})$. Based on the previous analysis, it's obvious that if $p_1 > p_2$, we have $F(p_1) > F(p_2)$. Therefore, let $p_{max} - p_{sel} = 2\varepsilon^*$ in order to obtain the maximal ΔF considering the worst case. Then, as eavesdropping times increase, the worst-case ΔF will decrease until an inflection point n_0^* occurs, after which the decrease of ΔF is very small. And this inflection point is the eavesdropping times we are looking for. However, ΔF is hard to be solved since p_{max} is unknown. But the range of p_{max} is available. We can find a p_{max} that maximizing ΔF by solving the following optimization problem

$$\begin{aligned} &\max \{F(p_{max}) - F(p_{max} - 2\varepsilon^*)\} \\ &s.t. \quad p_{max} \in [\frac{1}{d} + \varepsilon^*, \frac{1 + 2\varepsilon^*}{2}]. \end{aligned} \tag{15}$$

Remark 1. *The range of p_{max} is discussed below. When the transmission probability of other routing paths reaches the maximum value (not greater than $\frac{1}{d}$), p_{max} takes the minimum value. Similarly, when the transmission probability of other routing path is closed to 0, p_{max} takes the maximum value.*

The specific process of obtaining the eavesdropping times is shown in Algorithm 1.

Remark 2. *According to [14], the energy consumption of attack is larger than eavesdropping, i.e., $E_a = 3E_l$. Therefore, the total energy quantization times is $N = \frac{E_t}{E_a} = \frac{n_0}{3} + n$. As eavesdropping times increase, ΔF will reduce as the learning accuracy is improved. However, due to the energy constraint, the attack times will decrease and the reduction of ΔF will also decrease. Given a threshold $\eta = 90\%$, increase the eavesdropping times n_0 until $\frac{\Delta F_2}{\Delta F_1} > \eta$, which indicates that the gradient of the difference of attack benefits is small. Then there is no need to increase the eavesdropping times, because the improvement in learning accuracy has less impact on attack benefits. Thus n_0^* is obtained through Algorithm 1.*

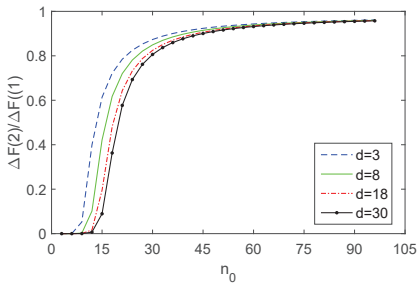

Fig. 3: The effect of different number of routing paths.

IV. Performance Evaluation

In this section, we illustrate the results by using some numerical examples. Consider the system (1) with the following parameters:

$$A = \begin{bmatrix} 1.05 & 0.5 \\ 0 & 0.85 \end{bmatrix}, \quad C = \begin{bmatrix} 1.1 & 0.7 \end{bmatrix},$$
$$Q = \begin{bmatrix} 0.3 & 0 \\ 0 & 0.4 \end{bmatrix}, \quad R = 0.4.$$

A. Impact of the Number of Routing Paths

The effect of different number of routing paths is illustrated in Fig. 3. We plot four curves of $\Delta F_2/\Delta F_1$ as a function of n_0 when the number of routing paths $d = 3, 8, 18, 30$, respectively. Given the energy budget $N = 150$, for a fixed routing path number, the initial value of $\Delta F_2/\Delta F_1$ is very small, which illustrates that when n_0 is too short, the learning phase's impact on the attack benefits is small. When $n_0 > 10$, the rapid increase of $\Delta F_2/\Delta F_1$ reveals the improvement in learning accuracy. As n_0 increases, $\Delta F_2/\Delta F_1$ is getting smaller due to the reduction of attack times. According to Algorithm 1, given $\eta = 90\%$ we obtain n_0^* so that $\Delta F_2/\Delta F_1 > \eta$. Moreover, as the number of routing paths increases, the eavesdropping times that we designed also increases. Namely, the attacker needs to spend more eavesdropping times to find the routing path with the largest transmission probability.

B. Impact of Eavesdropping Times

Fig. 4 shows the influence of different eavesdropping times n_0. We simulate the transmission of 4 routing paths. Considering scenario 1, where the sensor selects the routing path with higher QoS quality or fewer hops, the transmission probability of each path varies greatly. So we give the following routing transmission probabilities: $p_1 = 0.62, p_2 = 0.05, p_3 = 0.28, p_4 = 0.05$. In scenario 2, each routing path has a similar transmission probability for energy balance. Thus the transmission probabilities are given by $p_1 = 0.28, p_2 = 0.21, p_3 = 0.34, p_4 = 0.17$. Given $N = 150, d = 4, \eta = 90\%, \beta = 90\%$, from Algorithm 1 we have $n_0^* = 39$. As shown in Fig. 4, in scenario 1, the attacker can find the routing path with the maximum transmission probability by eavesdropping several times. In scenario 2, the attacker needs more times to find out the path with the maximum transmission probability. However, in either case, eavesdropping times do not exceed n_0^*. Although n_0^* is not optimal, the attack benefits it brings are not much different from the optimal attack benefits.

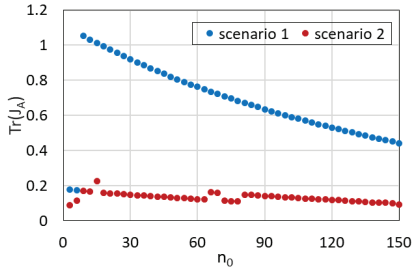

Fig. 4: The effect of different eavesdropping times.

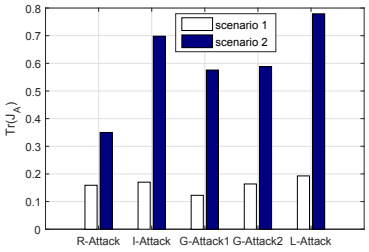

Fig. 5: Comparison results between learning-based attack and other attack strategies.

C. Comparison with Other Attack Strategies

We compare the attack benefits of different attack strategies in two representative scenarios, as shown in Fig. 5. We use L-Attack to represent our learning-based attack strategy. I-Attack represents the general attack strategy that eavesdropping and attacking are performed crosswise. G-Attack1 and G-Attack2 are two kinds of grouped attacks [11], but attack times are not in the end of the time series. G-Attack1 eavesdrops a short times before attack and G-Attack2 eavesdrops a long times before attack. R-Attack denotes the random attack strategy which randomly selects the routing path and uses all the energy for the aggregate attack. To have a fair comparison, we also set $n_0 = 39$ for all attack strategies except R-Attack. The experiment is repeated 10 times.

It shows that L-attack strategy performs better than other attack strategies in both scenarios. This reveals the superiority of our attack strategy. In scenario 1, the R-Attack strategy performs worst. Because it's unlikely to attack the optimal routing path without the learning process. In scenario 2, the G-Attack1 strategy performs worst, because the eavesdropping times is too short to find out the optimal routing path when the transmission probability of each routing path is similar. This reveals that the learning phase has significant influence on attack benefits. It's important to take the learning process into consideration, when the attacker has inadequate capacity.

V. Conclusions and Further Discussions

In this paper, the attack schedule against remote state estimation under energy constraint is studied. The attacker is assumed to have no prior knowledge of the transmission pattern. Then, the learning phase is considered to eavesdrop the transmission probability of each routing path to improve the attack efficiency. The learning phase preceding the grouped attack is proved to be the optimal attack pattern in the sense of possibility degree. Furthermore, a novel algorithm is proposed to design the rational eavesdropping times. The simulation results illustrate that the learning-based attack strategy proposed in this paper preforms better than other attack strategies. In the future, we will investigate the corresponding defense strategy under this optimal attack strategy.

VI. Acknowledgment

This work is partially supported by The National Key R&D Program of China under the grant 2018YFB1702100, and by NSF of China under the grants 61521063, 61622307 and 61603251, and by NSF of Shanghai Municipality of China under the grant 18ZR1419900.

References

[1] P. Leitão, S. Karnouskos, L. Ribeiro, J. Lee, T. Strasser, and A. W. Colombo, "Smart agents in industrial cybercphysical systems," *Proceedings of the IEEE*, vol. 104, no. 5, pp. 1086–1101, 2016.

[2] L. Lyu, C. Chen, and et al, "Dynamics-aware and beamforming-assisted transmission for wireless control scheduling," *IEEE Transactions on Wireless Communications*, vol. 17, no. 11, pp. 7677–7690, 2018.

[3] J. He and L. Cai, "Hybrid content distribution framework for large-scale vehicular ad hoc networks," *ZTE COMMUNICATIONS*, 2016.

[4] S. Sridhar, A. Hahn, and M. Govindarasu, "Cybercphysical system security for the electric power grid," *Proceedings of the IEEE*, vol. 100, pp. 210–224, 2011.

[5] D. R. Raymond and S. F. Midkiff, "Denial-of-service in wireless sensor networks: Attacks and defenses," *IEEE Pervasive Computing*, vol. 7, pp. 74–81, 2008.

[6] H. Zhang, P. Cheng, L. Shi, and J. Chen, "Optimal dos attack scheduling in wireless networked control system," *IEEE Transactions on Control Systems Technology*, vol. 24, no. 3, pp. 843–852, 2016.

[7] F. Pasqualetti, F. Dörfler, and F. Bullo, "Attack detection and identification in cyber-physical systems," *IEEE Transactions on Automatic Control*, vol. 58, pp. 2715–2729, 2013.

[8] J. He, L. Cai, and X. Guan, "Optimal state estimation for distributed algorithm with noise adding mechanism," in *CDC*, 2017, pp. 4135–4140.

[9] Z. Guo, D. Shi, K. H. Johansson, and L. Shi, "Optimal linear cyber-attack on remote state estimation," *IEEE Transactions on Control of Network Systems*, vol. 4, no. 1, pp. 4–13, 2017.

[10] M. Esmalifalak, G. Shi, Z. Han, and L. Song, "Bad data injection attack and defense in electricity market using game theory study," *IEEE Transactions on Smart Grid*, vol. 4, no. 1, pp. 160–169, 2013.

[11] H. Zhang, P. Cheng, L. Shi, and J. Chen, "Optimal denial-of-service attack scheduling with energy constraint," *IEEE Transactions on Automatic Control*, vol. 60, no. 11, pp. 3023–3028, 2015.

[12] J. Qin, M. Li, L. Shi, and X. Yu, "Optimal denial-of-service attack scheduling with energy constraint over packet-dropping networks," *IEEE Transactions on Automatic Control*, vol. 63, no. 6, pp. 1648–1663, 2018.

[13] K. Ding, Y. Li, D. E. Quevedo, S. Dey, and L. Shi, "A multi-channel transmission schedule for remote state estimation under dos attacks," *Automatica*, vol. 78, pp. 194–201, 2017.

[14] Z. Yang, P. Cheng, and J. Chen, "Learning-based jamming attack against low-duty-cycle networks," *IEEE Transactions on Dependable and Secure Computing*, vol. 14, no. 6, pp. 650–663, 2017.

[15] K. Sha, J. Gehlot, and R. Greve, "Multipath routing techniques in wireless sensor networks: A survey," *Wireless Personal Communications*, vol. 70, no. 2, pp. 807–829, 2013.

[16] Y.Nakahara, M.Sasaki, and M.Gen, "On the linear programming problems with interval coefficients," *Computers & Industrial Engineering*, vol. 23, no. 1-4, pp. 301–304, 1992.

[17] A. Sengupta and T. K. Pal, *On Comparing Interval Numbers: A Study on Existing Ideas*, Springer Berlin Heidelberg, 2009.

[18] L. Shi, P. Cheng, and J. Chen, "Optimal periodic sensor scheduling with limited resources," *IEEE Transactions on Automatic Control*, vol. 56, no. 9, pp. 2190–2195, 2011.

[19] W. Hoeffding, "Probability inequalities for sums of bounded random variables," *Journal of the American Statistical Association*, vol. 58, no. 301, pp. 13–30, 1963.