Stealthy False Data Injection Attack on Unmanned Aerial Vehicles with Partial Knowledge

Umair Ahmad Mughal*, Muhammad Ismail*, and Syed Ali Asad Rizvi†
*Department of Computer Science, Tennessee Technological University, Cookeville, TN, USA
†Department of Electrical and Computer Engineering, Tennessee Technological University, Cookeville, TN, USA
Emails: {uamughal42, mismail, srizvi}@tntech.edu

2023 IEEE Conference on Communications and Network Security (CNS) | 979-8-3503-3945-1/23/$31.00 ©2023 IEEE | DOI: 10.1109/CNS59707.2023.10289001

***Abstract*—Unmanned aerial vehicles (UAVs) are widely used for military and civilian applications. Despite their wide adoption, they are vulnerable to cyber-attacks, which could lead to serious consequences that cause the UAVs to crash or be directed to undesired locations. In this context, several works have investigated false data injection (FDI) attacks. However, existing studies often assume that the attacker knows all the system and control parameters, which may not always be the case. In this paper, we propose a strategy for stealthy FDI attacks that does not require full knowledge of the system and control parameters. This approach eavesdrops and injects false commands and data into two main communication channels, namely, the feedback channel (from the UAV to the ground controller) and the forward channel (from the ground controller to the UAV). The attacker incorporates two separate Kalman filters that process inputs from both channels and subsequently generate an estimate of the UAV's current state. Then, stealthy false data and commands are injected while ensuring UAV's stability. We tested the proposed attack in simulation and validated it through experiments on an actual UAV. Our experimental results demonstrate that the proposed FDI attack causes the UAV to deviate from its original path while remaining stealthy. We also demonstrate that our attack causes a large deviation in the UAV's path compared to other attacks that assume full knowledge of the system parameters. Through this study, we aim to shed some light that helps to develop robust defense mechanisms against such attacks.**

***Index Terms*—UAV, stealthy false data injection, validation.**

I. Introduction

Unmanned aerial vehicles (UAVs) are cyber-physical systems that depend on real-time communication, onboard sensor measurements, and control system to carry out flight missions. Their several attractive features have led to their wide adoption in both civilian and military applications, such as the delivery of goods [1], search and rescue operations [2], disaster management [3], and surveillance [4]. These applications rely on the state estimation process, which plays a significant role in UAV's operation. It involves the use of mathematical models and sensor data to estimate the system's state variables, such as position, velocity, and orientation. However, the state estimation process is susceptible to cyber-attacks over the communication channels connecting the UAV to the ground controller [5]. Among various types of cyberattacks, false data injection (FDI) attacks have been the focus of several research works [6]. In these attacks, false command and sensor data are injected into the communication channels, leading to incorrect state estimates that can result in severe accidents and potentially irreversible damages [7], [8].

A particular type of FDI attack that has been gaining attention is the stealthy FDI attack. This attack involves the manipulation of the system's state estimation in such a way that it goes undetected by anomaly detectors, potentially leading the UAV to a malicious destination, causing a crash, or even hijacking. Most of the existing research mainly focuses on the detection of stealthy FDI attacks, with only a few studies exploring the practical design of such stealthy attacks. Furthermore, the few existing works that explore attack design involve unrealistic assumptions, such as the attacker having full access to the system and control parameters. This motivates us to design an FDI attack that does not require the attacker's full knowledge of the system and control parameters. Furthermore, we aim to validate this attack on an actual UAV, as all existing studies are based on simulations. In what follows, we summarize the related works and highlight their limitations to motivate our research and contributions.

A. Related Works

We first summarize the existing works on the detection of FDI attacks. Then, we cover the closely related works on the design of stealthy FDI attacks.

1) FDI Attack Detection: Extensive research has been conducted to ensure the secure operation of the UAV, considering vulnerabilities of the communication network, control system, sensor measurements, and feedback channel to develop effective methods for detecting cyber-attacks. In this context, several studies have developed strategies to detect FDI attacks. Xiao *et al.* [9] proposed an estimation-based modified sliding innovation sequence method to detect FDI attacks on UAV sensors and actuators. Ye *et al.* [10] proposed a summation of the innovation method to detect FDI attacks on UAVs. Zhou *et al.* [11] designed a matrix-based framework that gives weights to the different threat levels for attack detection. Miao *et al.* [12] proposed a coding-based detection scheme to detect FDI attacks. Zhao *et al.* [13] designed a data-driven attack and proposed a detection method based on the subspace matrix

This publication was made possible by NPRP13S-0127-200182 from the Qatar National Research Fund (a member of the Qatar Foundation). The statements made herein are solely the responsibility of the authors.

979-8-3503-3945-1/23/$31.00 ©2023 IEEE

observability and coding theory. While these works studied the detection of FDI attacks, not many of them considered stealthy attacks that can bypass the anomaly detector.

2) FDI Attack Design: Several works investigated FDI attack design in networked cyber-physical industrial control systems. These designs include optimal strategies and conditions subject to undetectability [14], Kullback–Leibler (KL) probability divergence [15], residual distributions [16], optimal switching conditions [17], and game theory [18], [19]. Furthermore, the FDI attack on quantized network control was studied in [20]. In [21], a control signal attack was proposed on the linear quadratic Gaussian control system where the optimal trade-off between the stealthiness of the attack and damage caused was studied. In [22], the statistic error cost was used to find the optimal attack strategy. However, the prior attacker's knowledge of the system matrices is required to model and design the attack strategy. Similarly, [23] presented an FDI attack formulated as an optimization problem where the attack sequence is subject to the upper bound bias introduced in the detection statistics. However, the work assumes full attacker's knowledge of the statistical properties of the system, feedback controller, and actuator.

In a closely related work, Kwon *et al.* [24] developed a stealthy FDI attack on the UAV's state estimation. The author investigates the conditions to predict the effect of the attack and then designed the optimal FDI attack by solving a constrained optimization problem. However, this work assumed that the attacker knows the physical system matrices and controller parameters. Dash *et al.* [25] introduced malware-based FDI attacks considering the vulnerabilities of the Proportional-Integral-Derivative (PID) control system. The work in [25] proposed an FDI attack assuming that the adversary already knows the PID control parameters and that the parameters are constant. Chen *et al.* [26] proposed an FDI attack on a magnetometer. Su *et al.* [27] developed a spoofing strategy to manipulate the GPS, leading the UAV to any arbitrary destination following a constrained optimization problem.

B. Limitations

The existing attack strategies assume that the attacker has full knowledge of the system's physical matrices and control parameters. While assuming that the attacker knows the system matrices is possible [25], assuming that the attacker has full knowledge of the system and control parameters is not realistic. There is a need for an FDI attack strategy that relaxes this strict assumption and considers cases where the attacker has partial knowledge of the system and control parameters of the UAV to launch the stealthy FDI attack. Also, studying the effect of this partial knowledge of the system and control parameters at the attacker on the effectiveness of the attack requires further investigations. Moreover, one of the key motivations and contributions of our study is the practical implementation of the attack on an actual UAV, moving beyond theoretical models and simulations that have been conducted in the existing research.

C. Contributions

To address the aforementioned limitations and fill the research gap, we have carried out the following:

- We developed a technique for a stealthy FDI attack that does not require the attacker's full knowledge of the system and control parameters. This approach targets two communication channels, namely, the feedback channel (from the UAVs to the ground controller) and the forward channel (from the ground controller to the UAV). We incorporate two separate Kalman filters at the attacker that process inputs eavesdropped from both channels and subsequently generate an estimate of the UAV's current state. The attacker then targets the feedback channel to manipulate the sensors' readings, leading to incorrect state estimation. The attack on the forward channel aims to keep the system in a stable state so that the anomaly detector does not raise an alarm.
- To validate the proposed attack, we have developed a testbed that involves an actual UAV, ground control station, and attacker. The experimental validation involved collecting sensor measurements and control signals from the UAV and injecting false commands and data following our proposed attack strategy, which ultimately caused the UAV to deviate from its original path.
- We have evaluated the proposed strategy via simulations and experimental validation. Our simulations demonstrate that the proposed attack strategy caused a large state estimation error (e.g., large deviations in the UAV's path) compared with existing strategies while remaining stealthy. Also, our experimental results demonstrated that the residual errors after the attack remain close to zero, indicating the stealthiness of the attack. We have also reported the corresponding deviations in the UAV's path.

The remainder of this paper is as follows. In Section II, we introduce the system model including the state estimation, anomaly detector, and attack model. In Section III, we present the proposed attack strategy. Section IV presents the simulation results, a description of the testbed and the practical implementation of the attack, and the experimental results. Finally, Section VI concludes the paper and discusses potential directions for future research.

II. System Model

We consider the system model shown in Fig. 1. The system consists of a UAV, a ground controller, and an attacker. The UAV and ground controller communicate over two channels: (a) the feedback channel carries the sensor measurements from the UAV to the ground control station. The ground station uses the sensor measurements to estimate the state of the UAV (e.g., position, speed, etc.) and determines the control signal; (b) the forward channel carries the control signals from the ground control station to the UAV to control the physical actions of the UAV (e.g., movement). The ground control station employs

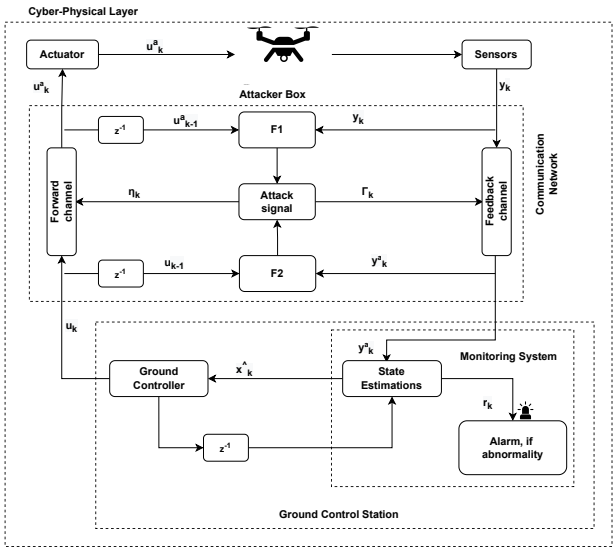

Fig. 1. UAV system architecture under proposed FDI attacks

a residual-based anomaly detector. While we consider herein a Chi-square detector, the attack strategy described herein works against any residual-based anomaly detector. As shown in Fig. 1, an attacker is located in the middle between the UAV and the ground control station. The attacker eavesdrops on the feedback and forward channels between the UAV and the ground station and uses the information in two Kalman filters (F_1 and F_2 in Fig. 1) to inject stealthy false commands and data into these channels while bypassing the Chi-square anomaly detector at the ground control station. These two filters eliminate the need for the attacker's knowledge about the exact control parameters.

Next, we discuss the system state estimation, the residual-based anomaly Chi-Square detector, and the attack model.

A. State Estimation and Control at Ground Station

Consider a UAV system modeled as a linear time-invariant system [6], [9], [10] given by

$$\begin{aligned} x_{k+1} &= Ax_k + Bu_k + w_k, \\ y_k &= Cx_k + v_k, \end{aligned} \tag{1}$$

where $A, B,$ and C represent the state transition, control, and measurement matrices, respectively, which are known as system matrices, $x_k \in \mathbb{R}^n, u_k \in \mathbb{R}^m$ and $y_k \in \mathbb{R}^p$ represent the system states, control inputs, and sensor measurement outputs, respectively, $w(k) \in \mathbb{R}^n$ and $v(k) \in \mathbb{R}^p$ represent the process noise and measurement noise that are independent and identical distribution (i.i.d.) Gaussian random vectors with zero mean and covariance matrices Q and R, respectively, i.e., $w_k \sim \mathcal{N}(0, Q)$, and $v_k \sim \mathcal{N}(0, R)$.

Kalman filter is used at the ground control station to estimate the state x_k using y_k as follows

$$\begin{aligned} \hat{x}_{k_0} &= A\hat{x}_{k-1} + Bu_{k-1}, \\ P_{k_0} &= AP_{k-1}A^T + Q, \\ K_k &= P_{k_0}C^T\left(CP_{k_0}C^T + R\right)^{-1}, \\ \hat{x}_k &= \hat{x}_{k_0} + K_k\left(y_k - C\hat{x}_{k_0}\right), \\ P_k &= \left(I - K_kC\right)P_{k_0}, \end{aligned} \tag{2}$$

where $\hat{x}_{k_0}$ represents the initial estimate of state x_k, P_{k_0} is a corresponding error covariance matrix having $P_{k_0} = \mathbb{E}[(x_{k_0} - \hat{x}_{k_0})(x_{k_0} - \hat{x}_{k_0})^T]$, $K_k \in \mathbb{R}^{m\times p}$ represents the Kalman's filter gain at time instant k, and I is the identity matrix. The Kalman filter converges to a constant K provided with any initial condition [28]. Therefore, the steady-state error covariance becomes $P \triangleq \lim_{k\to\infty} P_k$, where P is a positive semi-definite matrix with the proper dimensions. Hence, the Kalman filter has a fixed gain as follows

$$K = PC^T\left(CPC^T + R\right)^{-1}. \tag{3}$$

Based on the current system estimate $\hat{x}_k$, the feedback control can be designed as [27], [29]

$$u_k = -L\hat{x}_k, \tag{4}$$

where L is the controller's gain, which determines the strength or influence of the control action. This feedback control is designed to regulate the system's behavior and ensure it follows a desired path or maintains a specific state. The negative sign in front of $L\hat{x}_k$ indicates that the control action is applied in the opposite direction to the estimated state. This is a characteristic of negative feedback control, which is commonly used to stabilize systems. In simple terms, the controller estimates the current state of the system, compares it with the desired state, and applies a control action proportional to the difference between the two. The control action is applied in such a way as to drive the system toward the desired state.

Define the estimation error under normal operation as

$$e_k \triangleq x_k - \hat{x}_k. \tag{5}$$

From (2), we can write the next estimate as follows

$$\begin{aligned} \hat{x}_{k+1} =& A\hat{x}_k + Bu_k \\ &+ K(y_{k+1} - C(A\hat{x}_k + Bu_k)). \end{aligned} \tag{6}$$

Using (1) and (6), we have

$$\begin{aligned} e_{k+1} =& x_{k+1} - \hat{x}_{k+1} \\ =& (I - KC)Ae_k + (I - KC)w_k \\ &- Kv_{k+1}. \end{aligned} \tag{7}$$

Using (4) in (1) and merging with (7), we have the system evolves as follows

$$\begin{aligned} \begin{bmatrix} x_{k+1} \\ e_{k+1} \end{bmatrix} =& \begin{bmatrix} A - BL & BL \\ 0 & (I - KC)A \end{bmatrix} \begin{bmatrix} x_k \\ e_k \end{bmatrix} \\ &+ \begin{bmatrix} I \\ I - KC \end{bmatrix} w_k - \begin{bmatrix} 0 \\ K \end{bmatrix} v_{k+1}. \end{aligned} \tag{8}$$

For the stability of the system, there are two closed-loop gains: L (controller's gain) and K (Kalman's gain). The system in (8) is stable only if the eigenvalues of the matrices $(A-BL)$ and $(I-KC)A$ are in the unit circle. This is possible by designing the controller and filter gains (i.e., L and K) [24], [27], whereas (A, B) is controllable and (A, C) is observable.

B. Chi-Square Detector

To ensure the normal operation of the UAV, the ground control station is equipped with an anomaly detector that monitors the system operation and detects anomalies such as false data using the statistical attributes of the acquired innovations. We consider a χ^2-detector [30], which is a residual-based detector that is used to detect anomalies in the system. The detector makes the decision if there is an anomaly based on the summation of the squared innovations normalized by covariance P. The criteria for detection [30] is given by

$$\sum_{i=k-\tau+1}^{k} r_k^T P^{-1} r_k \underset{H_1}{\overset{H_0}{\lessgtr}} \rho(\phi), \tag{9}$$

where τ is the size of the detection window, r_k is a residual of the system with i.i.d. Gaussian process with zero mean and covariance P, i.e., $r_k \sim \mathcal{N}(0, P)$ with $P = CP_kC^T + R$, ρ is the threshold, and null hypothesis H_0 is true when the system is under normal condition, whereas, the hypothesis H_1 represents the presence of an anomaly. Finally, $\phi = \chi^2_{1-\mu}$, where μ is the detector false alarm rate.

C. Attack Model

In UAVs, the actuators and sensors are connected to the forward and feedback communication channels, respectively [31], [24]. The FDI attack is launched as follows

$$y_k^a = y_k + \Gamma_k, \tag{10}$$

$$u_k^a = u_k + \eta_k, \tag{11}$$

where Γ_k and η_k are the attack vectors against the feedback and forward communication channels, respectively, and y_k^a and u_k^a represent the compromised sensor's output and controller's output due to the feedback and forward channel attacks (10) and (11), respectively, as shown in Fig. 1. The attacker's objective is to design Γ_k and η_k such that large estimation errors are attained at the controller, causing deviations to the UAV's trajectory, while remaining stealthy (i.e., bypassing the H_1 hypothesis of the χ^2-detector).

III. Sealthy FDI Attack Strategy

In this section, we discuss the design strategy of the proposed FDI attack.

A. FDI Attack

The system in (1) can be written as follows under attack

$$\begin{aligned} x_{k+1}^a &= Ax_k^a + Bu^a{}_k + w_k^a, \\ y_k^a &= Cx_k^a + v_k^a, \end{aligned} \tag{12}$$

where x_{k+1}^a, $u^a{}_k$, y_k^a represent the system state, control, and measurement under attack. In order to simplify our notation, the superscript a representing 'attack' has been omitted, since all variables from here on will be in the context of a system under attack. The term a will be incorporated only in the controller's and Kalman's gains (L^a and K^a) to differentiate them from the actual gains (L and K) adopted at the ground control station. In the following, we consider that the attacker has a partial system and control knowledge. Hence, we assume that the attacker only knows the system matrices, i.e., (A, B, C). However, the attacker does not know K and L. This relaxation of the assumption of full attacker's knowledge is practical as the attacker can infer the system matrices (A, B, C) [25] as we will demonstrate in the implementation section. We will demonstrate a strategy where the attacker does not need to use the exact K and L adopted at the ground control station for stealthy FDI, which is unlike the existing literature that requires the attacker's knowledge of system matrices (A, B, C), Q, R, and parameters K and L.

Given the system matrices, the attacker employs two filters to estimate the current state of the compromised UAV as shown in Fig. 1. Using the first filter F_1, the state estimation is given by

$$\begin{aligned} \hat{x}_{k+1} =& A\hat{x}_k + Bu_k \\ &+ K^a \left(y_{k+1} - C\left(A\hat{x}_k + Bu_k\right)\right), \end{aligned} \tag{13}$$

where $\hat{x}_k$ is the estimate of the current state based on the measurement y_k. A similar process is applied using the second filter F_2, such that

$$\begin{aligned} \hat{\hat{x}}_{k+1} =& A\hat{x}_k + Bu_k \\ &+ K\left(y_{k+1}^a - C\left(A\hat{x}_k + Bu_k\right)\right), \end{aligned} \tag{14}$$

where the second filter utilizes the measurement y_k^a to output the current estimate of the state. The attacker's filter gain K^a is used for both filters as it simplifies the design and maintains a consistent estimate of the system state. The error between the state estimation of both filters is computed as follows

$$e_k \triangleq \hat{x}_k - \hat{\hat{x}}_k. \tag{15}$$

Remark 1: From (8), the stability of (12) depends on the stability of $(A - BL^a)$ and $(I - K^aC)A$.

Remark 2: It is essential that the attack design still ensure the stability of the closed-loop system. Failure to ensure system stability can raise the anomaly detector's alarm, which negates the purpose of stealthiness.

B. Attack Strategy

The attacker specifies K^a such that $(I - K^aC)A$ is stable, i.e., the eigenvalues of $(I - K^aC)A$ are within the unit circle. As for L^a, this is computed to maintain the stability, and hence, stealthiness of the attack. Toward this goal, the attacker computes L^a that maximizes the controller's quadratic cost function, which is given by

$$J = \mathbb{E}\left\{\sum_{k=0}^{\infty}\left(x_k^\top W x_k + u_k^\top U u_k\right)\right\}, \tag{16}$$

where $W \in \mathbb{R}^{n\times n}$ and $U \in \mathbb{R}^{n\times n}$ are symmetric weighting positive semi-definite matrices such that W penalizes the states and U penalizes the control variables. We rewrite (16), from the attacker's perspective, using the attacker's control strategy $u_k = -L^a y_k$ such that

$$J = \mathbb{E}\left\{\sum_{k=0}^{\infty}\left(x_k^\top \left(W + C^\top L^{a\top} U L^a C\right) x_k\right)\right\}. \tag{17}$$

The attacker designs L^a to ensure the stability of the system while increasing the cost in (17) in a controlled manner so that the cost does not overshoot, and hence, is not caught by the anomaly detector. This is achieved by having the eigenvalues of the matrices $(A - BL^aC)$ less than 1, i.e., inside the unit circle, which can be expressed as

$$\mathcal{I} \triangleq \left\{L^a \in \mathbb{R}^{m\times p} : \lambda(A - BL^aC) < 1\right\}. \tag{18}$$

Hence, the attacker aims to compute L^a that maximizes the cost function (17) subject to the stability constraint in (18). To solve this optimization problem, we adopt the Lyapunov stability method. This method defines a set of differential equations to analyze the steady-state stability of dynamic systems. Lyapunov equations can be expressed as [32]–[34]

$$\begin{aligned} -x_k^\top \left(W + C^\top L^{a\top} U L^a C\right) x_k &= \frac{d}{dx}\left(x_k^\top S x_k\right), \\ &= \dot{x}_k^\top S x_k + x_k^\top S \dot{x}_k, \\ &= x_k^\top \left(A_c^\top S + S A_c\right) x_k, \end{aligned} \tag{19}$$

where $(A - BL^aC) = A_c$ and S is a semi-definite symmetric matrix. We have

$$\mathcal{V} \equiv A_c^\top S + S A_c + C^\top L^{a\top} U L^a C + W. \tag{20}$$

Then, the cost function J can be written in terms of matrix S as follows

$$J = \text{tr}(SX), \tag{21}$$

where $X = \mathbb{E}\left\{x_0 x_0^\top\right\}$ represents the initial condition. The Hamiltonian function H can be defined to incorporate the constraint (18) through a Lagrange multiplier Λ and the matrix $\mathcal{V}$ [32]–[34]. Hence, we have

$$H = \text{tr}(SX) + \text{tr}(\Lambda\mathcal{V}). \tag{22}$$

To solve the optimization problem, we set the partial derivatives of H with respect to Λ, S, and L^a to zero, i.e.,

$$0 = \frac{\partial H}{\partial \Lambda} = g = A_c^\top S + S A_c + C^\top L^{a\top} U L^a C + W, \tag{23a}$$

$$0 = \frac{\partial H}{\partial S} = A_c \Lambda + \Lambda A_c^{\mathrm{T}} + X, \tag{23b}$$

$$0 = \frac{\partial H}{\partial L^a} = U L^a C \Lambda C^{\mathrm{T}} - B^{\mathrm{T}} S \Lambda C^{\mathrm{T}}. \tag{23c}$$

The solution of these equations gives the controller gain L^a, which can be computed in terms of S and Λ, which is given by

$$L^a = U^{-1} B^{\mathrm{T}} S \Lambda C^{\mathrm{T}} \left(C \Lambda C^{\mathrm{T}}\right)^{-1}. \tag{24}$$

In summary, the optimal control strategy from the attacker's perspective tries to maximize the cost function while ensuring system stability. The approach involves defining the quadratic cost function, expressing this cost in terms of the attacker's control strategy, ensuring stability, simplifying the optimization problem, and finding the optimal solution using Lyapunov stability and Lagrange multipliers.

In order to derive the attack signal Γ_k, we define the residual r_k as the difference between the current measurement y_k and the previous estimate $\hat{x}_k$ as follows

$$\begin{aligned} r_k &= y_k - \hat{y}_{k-1}, \\ r_k &= y_k - C\hat{x}_{k-1}, \\ &= y_k - C(A\hat{x}_{k-1} + Bu_{k-1}). \end{aligned} \tag{25}$$

The difference between the compromised and normal residual can be defined as follows

$$r_k^d, \triangleq r_k^a - r_k, \tag{26}$$

where r_k^a represents the residual of the system under attack and r_k is the normal residual. To ensure that the attack is stealthy, the difference between the compromised and normal residuals must be zero, i.e.,

$$0 = r_k^a - r_k. \tag{27}$$

By substituting $r_k^a = y_k^a - C(A\hat{x}_k - Bu_k^a)$ and $r_k = y_k - C(A\hat{x}_k - Bu_k)$ in (27), the feedback channel attack signal Γ_k can be formulated as

$$\Gamma_{k+1} = C(Ae_k^a + B(L^a\eta_k)), \tag{28}$$

where $e_k^a = \hat{x}_k - \hat{\hat{x}}_k$ is the estimation error between filters F_1 and F_2 outputs as described in (15).

The FDI attack will cause the system to deviate from its normal course if the attack signal on the forward channel is divergent. Therefore, the attack signal on the forward channel can be designed as follows

$$\eta_{k+1} = \Omega\eta_k, \tag{29}$$

where Ω is a matrix that is designed to cause the system divergence. This is achieved by choosing Ω such that some of its eigenvalues are outside the unit circle. However, it is crucial

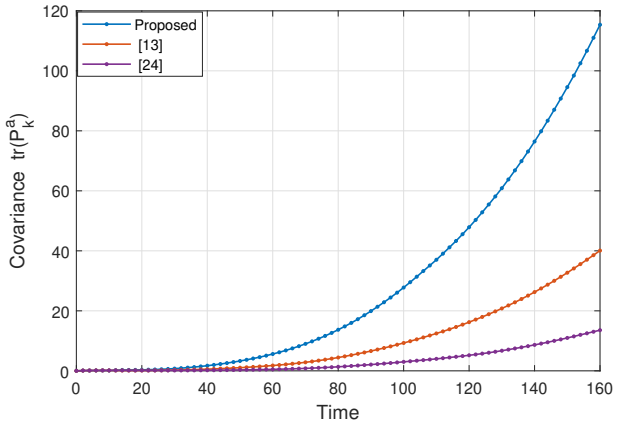

Fig. 2. Comparison of estimation error under different FDI attack strategies.

to ensure that while creating these disturbances, the overall system does not cross into a state of instability that would violate the Lyapunov stability criteria. Hence, Ω is designed to be just unstable enough to cause the intended divergence without destabilizing the overall system.

To evaluate the impact of the attack in the simulation results section, we compute the covariance error matrix P_k^a, when the system is under attack, which can be expressed as

$$P_k^a = \eta_k P_k \eta_k^T + \Gamma_k P_k \Gamma_k^T, \tag{30}$$

which is influenced by the attack signals η_k and Γ_k. The trace of the error covariance matrix P_k under attack can be written as

$$\text{tr}(P_k^a) = \text{tr}(\eta_k P_k \eta_k^T) + \text{tr}(\Gamma_k P_k \Gamma_k^T), \tag{31}$$

where $\text{tr}(P_k^a)$ shows the trace of the covariance P_k^a under the attack vector η_k and Γ_k.

Algorithm 1 summarizes the steps taken by the attacker to launch the proposed FDI attack. These include selecting a stable gain K^a, computing S and Λ using Lyapunov equations, computing the attacker's controller gain L^a, and deriving the feedback and forward attack signals Γ_k and η_k. Existing literature assumes that the attacker has complete knowledge of the system model's parameters (A, B, C, Q, R), as well as, the controller's parameters (K, L). Herein, we proposed a strategy where the attacker does not have knowledge of Q, R, K, L, yet, launches a stealthy attack. We will show in the experimental validation results how an attacker can deduce (A, B, C) matrices and follow Algorithm 1 to launch an effective yet stealthy attack. In summary, our proposed attack strategy effectively relaxes the strict requirements on the attacker's knowledge.

IV. Simulation and Experimental Results

In this section, we first simulate the proposed FDI attack and compare it with existing strategies. Then, we implement and validate the proposed attack on a testbed.

Algorithm 1: Proposed FDI Attack Strategy

Input: Collect: input u_k and output y_k data streams
Output: Γ_k, η_k, L^a // Attack signals and // Controller gain
begin
  Run Filter $F1$ to estimate x_k using y_k
  Run Filter $F2$ to estimate x_k using y_k^a
  Select the stable filter gain K^a // ensures $I - K^aC$ to be stable, i.e., eigenvalues within a unit circle
  Design the controller policy
    do (23a) to solve S // Lyapunov equation
    do (23b) to solve Λ // Lyapunov equation
      if *There is a feasible solution by solving (23a) and (23b) ; using Matlab Control System Toolbox* **then**
        Output: $L^a = U^{-1}B^{\mathrm{T}}S\Lambda C^{\mathrm{T}}\left(C\Lambda C^{\mathrm{T}}\right)^{-1}$ (24) // Optimum gain for the controller such that // $A - BL^aC$ asymptotically stable
      end if
      else
        Return to line 7
  Generate attack signals
    do (28) to attain Γ_k
    do (29) to attain η_k
  Breakdown

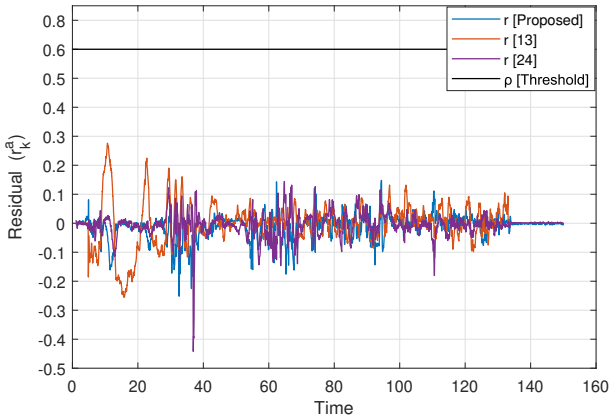

Fig. 3. Residual signals under different FDI attack strategies.

A. Simulation Results

The flight control system in [35] is considered herein with the following parameters

$$A = \begin{bmatrix} 0.9944 & -0.1203 & -0.4302 \\ 0.0017 & 0.9902 & -0.0747 \\ 0 & 0.8187 & 0 \end{bmatrix}, \quad B = \begin{bmatrix} 0.4252 \\ -0.0082 \\ 0.1813 \end{bmatrix}, \quad C = I_3, \tag{32}$$

where I_3 is a 3×3 identity matrix. The noise covariance matrices are $Q = 0.1I$ and $R = 0.1I$. The state definition $x \in \mathbb{R}^3$ denotes the pitch angle, pitch rate, and normal velocity. Using these parameters, the Kalman gain and controller gain

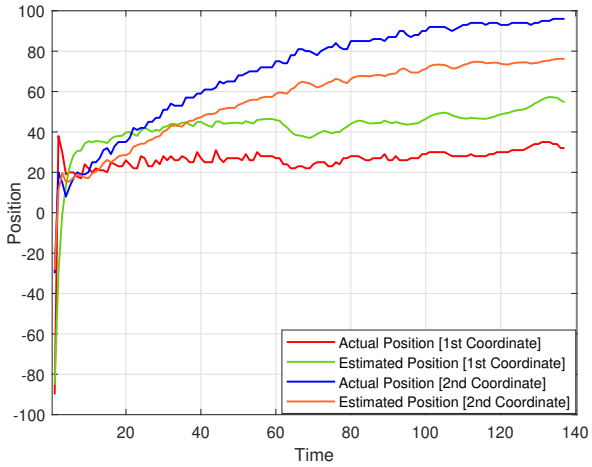

Fig. 4. UAV actual and estimated positions under the proposed FDI attack.

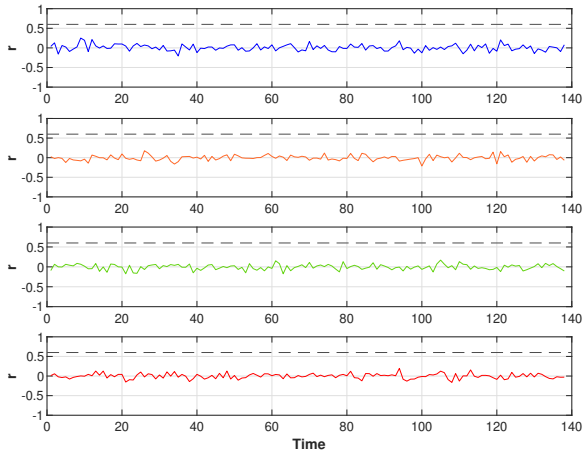

Fig. 5. Residual error of actual state and estimated state under the proposed FDI attack. The same color code as in Fig. 4 is used herein.

matrices from the ground control station perspective can be calculated as follows

$$K = \begin{bmatrix} 0.9173 & -0.0005 & -0.0006 \\ -0.0005 & 0.9156 & 0.0054 \\ -0.0006 & 0.0054 & 0.9135 \end{bmatrix}, \quad L = \begin{bmatrix} 0.7981 & -0.8586 & -0.2925 \end{bmatrix}. \tag{33}$$

In our strategy, the attacker does not need to follow the parameters in (33). Instead, the attacker parameters are calculated using Algorithm 1 and are found to be as follows

$$K^a = \begin{bmatrix} 0.1945 & -0.0001 & -0.0001 \\ -0.0011 & 0.1941 & 0.0011 \\ -0.0021 & 0.0011 & 0.1937 \end{bmatrix}, \quad L^a = \begin{bmatrix} 0.1596 & 0.1717 & -0.0585 \end{bmatrix}. \tag{34}$$

The corresponding eigenvalues of $(A-BL^a)$ and $(I-K^aC)A$ are $[0.7868, 0.7584, 0.0528]$ and $[0.0647, 0.9306, 0.8306]$, respectively. These eigenvalues are all within the unit circle, i.e., K^a and L^a ensure the system stability.

Fig. 2 shows the performance of the flight system [35] under the attack. We compared our attack strategy (using the parameters in (34)) with the attack strategies in [13] and [24] that require the attacker's knowledge of the parameters in (33). The trace of the estimation error covariance tr$(P^a{}_k)$ in Fig. 2 shows that the proposed attack caused a large estimation error covariance tr$(P^a{}_k)$ compared with [13] and [24]. Also, Fig. 3 shows the residuals of these attack strategies and ρ indicates the χ^2 detector's threshold. It can be seen that the residuals are below the detection threshold, which indicates the stealthiness of the attacks. Hence, in conclusion, our proposed attack strategy achieves high estimation error (hence causing the UAV to deviate from its planned path) without being detected by the χ^2 detector (i.e., being stealthy).

B. Experimental Validation

First, we will discuss the testbed and attack implementation. Then, we present and discuss the experimental results.

1) Tesbed: To validate our proposed attack, we have set up a testbed with the following components: a DJI Tello EDU drone [36], ALFA AWUS036ACH network adapter (antenna) [37], and two computers. The first computer acts as a ground control station where the UAV is connected to and receives the control commands over the wireless channels. This computer also has a χ^2 anomaly detection system as explained in Section II-B. The second computer acts as an attacker and runs Kali Linux, Aircrack-ng, Wireshark, and tcpdump, and is equipped with the ALFA adapter. This attacker eavesdrops on the communications between the Tello EDU drone and the ground control station, runs Algorithm 1, and injects the false command and measurements into the drone.

2) Attack Implementation: The following steps were taken in order to launch the proposed FDI attack on the DJI drone:

- *Dataset Collection:* A DJI flight map (3×3 m in size) is used with 2D Cartesian positional coordinates. The DJI Tello drone uses a vision positioning system that includes a camera and two 3D infrared sensors. The drone uses its positioning system to detect the flight map and sends the corresponding positional coordinates. The drone sends other telemetry data such as roll, pitch, yaw, time-of-flight, etc. In this validation, our FDI attack targets the 2D positional coordinates to deviate the drone from its original path. We record the 2D positional coordinates before and after the attack and report the deviation as the impact of the attack. We set the destination coordinate to be $(50, 80)$ and perform several flight missions, and collect the control input signal u_k and the output signal y_k representing the 2D positional coordinates of the drone.

- *Calculation of State Matrices:* First, part of the collected dataset was used in the Matlab System Identification toolbox [38] to calculate the system matrices (A, B, C) as in [25]. Then, (3) and (6) are used to compute the estimated system output. Then, a comparison is made between the estimated output and the actual output (real-time movement of the UAV as collected by the attacker). As some discrepancy exists between the estimated and actual output due to noise, we used an iterative approach to adjust the system matrices (A, B, C) for better accuracy. With each iteration, the system matrices are adjusted using collected data and the alignment between the estimated and actual output is determined. The adjustments of (A, B, C) continue until the estimated outputs converge toward the actual outputs.
- *Launching the FDI Attack:* To remain stealthy, the attacker injects attack vectors Γ_k and η_k such that deviations caused by the attack vectors are within the detection threshold. The actual threshold value is not necessarily to be the same as the attacker's assumed threshold. Here, the attacker assumes a low ρ value. The attacker runs Algorithm 1 as follows. First, the attacker specifies a K^a value that maintains $(I - K^a C)A$ stable, i.e., with eigenvalues inside the unit circle. Then, the attacker solves (23a) and (23b) to find S and Λ, which are then used in (24) to find L^a. Finally, the attack vectors Γ_k and η_k are found using (28) and (29), respectively, which are injected into the feedback and forward channels.

3) Experimental Results: Fig. 4 shows the effect of the attack on the drone's trajectory. The mission of the drone was to reach the destination at coordinate $(50, 80)$ with a tolerance of 5%, which means that it is acceptable for the drone to land within a 5% radius around the intended destination. The actual path of the drone is depicted by the red and blue lines for the 2D positional coordinates while the green and orange lines represent the drone's estimated position by the ground controller after the FDI attack. The final estimated position coordinates by the ground controller are $(54, 76)$ (i.e., within 5% of the intended destination) while the actual position coordinates (due to the FDI attack) are $(32, 97)$ as can be seen in Fig. 4. In our experiment, the drone initiated the landing procedure, under the wrong impression that it has reached its destination $(50, 80)$. In reality, however, the drone ends up at a different location, landing at the coordinate $(32, 97)$. This discrepancy between the actual and estimated positions indicates the impact of the proposed FDI attack, leading to a deviation in the drone's landing destination without being detected by the ground controller. This is confirmed by Fig. 5 which shows the residual error for the actual and estimated positions. As shown in Fig. 5, all the residual errors (actual and estimated) are around 0, i.e., below the detection threshold. This demonstrates the stealthiness of the attack. The black dotted line indicates the threshold adopted by the χ^2 anomaly detector. Also, Fig. 6 shows the associated injected false control signals for the 2D positions in the forward channel. As shown, all the control signals are smooth, and there are no abrupt changes in the control signals that can overshoot the quadratic cost of the controller.

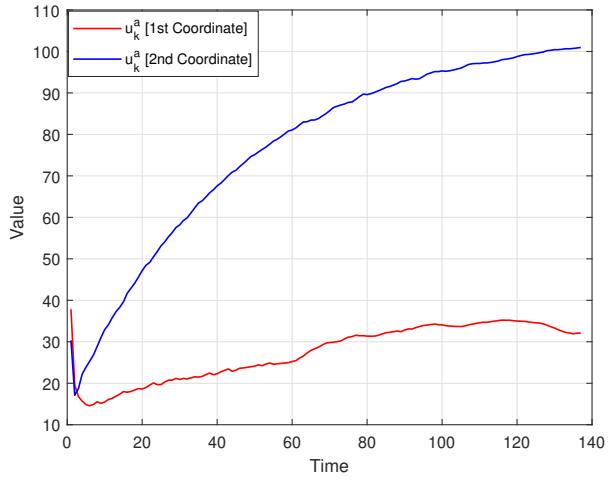

Fig. 6. Injected control signals in the forward channel for the position coordinates.

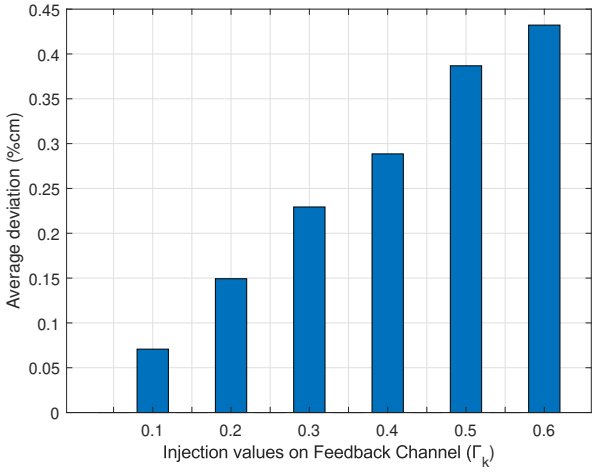

Fig. 7. Average deviation in position caused due to different injection values.

Fig. 7 shows the average deviation caused by the different injection values in the feedback channel (i.e., Γ_k). The x-axis depicts the range of optimal Γ_k values that have been calculated by the attacker at different time steps. These caused deviations by $7\% - 45\%$ in the planned path.

V. Conclusion

In this paper, we propose a stealthy false data injection attack that does not require the attacker's full knowledge of the system and control parameters. The proposed attack targets both the

feedback channel (from the UAV to the ground controller) and the forward channel (from the ground controller to the UAV). By setting up two estimation filters, the attacker can inject false commands and measurement data that cause a large estimation error at the ground controller while keeping the attack stealthy (i.e., bypassing the residual error detector). We examined the proposed strategy in simulation and also via experimental validation using an actual drone. Our simulation and experimental results confirmed the effectiveness and stealthiness of the proposed attack.

Our future work will explore design strategies for false data injection attacks in a multi-UAV network.

REFERENCES

[1] P. M. Kornatowski, M. Feroskhan, W. J. Stewart, and D. Floreano, "Downside up: Rethinking parcel position for aerial delivery," *IEEE Robotics and Automation Letters*, vol. 5, no. 3, pp. 4297–4304, 2020.

[2] A. Albanese, V. Sciancalepore, and X. Costa-Pérez, "SARDO: An automated search-and-rescue drone-based solution for victims localization," *IEEE Transactions on Mobile Computing*, vol. 21, no. 9, pp. 3312–3325, 2022.

[3] G. Aiello, F. Hopps, D. Santisi, and M. Venticinque, "The employment of unmanned aerial vehicles for analyzing and mitigating disaster risks in industrial sites," *IEEE Transactions on Engineering Management*, vol. 67, no. 3, pp. 519–530, 2020.

[4] B. Vergouw, H. Nagel, G. Bondt, and B. Custers, "Drone technology: Types, payloads, applications, frequency spectrum issues and future developments," in *The future of drone use*. Springer, 2016, pp. 21–45.

[5] A. Fotouhi, H. Qiang, M. Ding, M. Hassan, L. G. Giordano, A. Garcia-Rodriguez, and J. Yuan, "Survey on UAV cellular communications: Practical aspects, standardization advancements, regulation, and security challenges," *IEEE Communications surveys & tutorials*, vol. 21, no. 4, pp. 3417–3442, 2019.

[6] D. Ding, Q.-L. Han, X. Ge, and J. Wang, "Secure state estimation and control of cyber-physical systems: A survey," *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, vol. 51, no. 1, pp. 176–190, 2020.

[7] N. Shachtman and WIRED, "Computer virus hits us. drone fleet," *Available: https://edition.cnn.com/2011/10/10/tech/innovation/virus-hits-drone-fleet-wired/index.html*, 2011.

[8] K. Hartmann and C. Steup, "The vulnerability of UAVs to cyber attacks-an approach to the risk assessment," in *2013 5th international conference on cyber conflict (CYCON 2013)*. IEEE, 2013, pp. 1–23.

[9] J. Xiao and M. Feroskhan, "Cyber attack detection and isolation for a quadrotor UAV with modified sliding innovation sequences," *IEEE Transactions on Vehicular Technology*, 2022.

[10] D. Ye and T.-Y. Zhang, "Summation detector for false data-injection attack in cyber-physical systems," *IEEE transactions on cybernetics*, vol. 50, no. 6, pp. 2338–2345, 2019.

[11] Y. Zhou, K. G. Vamvoudakis, W. M. Haddad, and Z.-P. Jiang, "A secure control learning framework for cyber-physical systems under sensor and actuator attacks," *IEEE Transactions on Cybernetics*, vol. 51, no. 9, pp. 4648–4660, 2020.

[12] F. Miao, Q. Zhu, M. Pajic, and G. J. Pappas, "Coding schemes for securing cyber-physical systems against stealthy data injection attacks," *IEEE Transactions on Control of Network Systems*, vol. 4, no. 1, pp. 106–117, 2016.

[13] Z. Zhao, Y. Huang, Z. Zhen, and Y. Li, "Data-driven false data-injection attack design and detection in cyber-physical systems," *IEEE Transactions on Cybernetics*, vol. 51, no. 12, pp. 6179–6187, 2020.

[14] Y. Chen, S. Kar, and J. M. Moura, "Optimal attack strategies subject to detection constraints against cyber-physical systems," *IEEE Transactions on Control of Network Systems*, vol. 5, no. 3, pp. 1157–1168, 2017.

[15] Y. Mo, E. Garone, A. Casavola, and B. Sinopoli, "False data injection attacks against state estimation in wireless sensor networks," in *49th IEEE Conference on Decision and Control (CDC)*, 2010, pp. 5967–5972.

[16] Z.-H. Pang, G.-P. Liu, D. Zhou, F. Hou, and D. Sun, "Two-channel false data injection attacks against output tracking control of networked systems," *IEEE Transactions on Industrial Electronics*, vol. 63, no. 5, pp. 3242–3251, 2016.

[17] G. Wu, J. Sun, and J. Chen, "Optimal data injection attacks in cyber-physical systems," *IEEE Transactions on Cybernetics*, vol. 48, no. 12, pp. 3302–3312, 2018.

[18] R. Zhang and P. Venkitasubramaniam, "False data injection and detection in LQG systems: A game theoretic approach," *IEEE Transactions on Control of Network Systems*, vol. 7, no. 1, pp. 338–348, 2020.

[19] F. Li and Y. Tang, "False data injection attack for cyber-physical systems with resource constraint," *IEEE Transactions on Cybernetics*, vol. 50, no. 2, pp. 729–738, 2020.

[20] L. Liu and Z. Xi, "False data injection attack sequence design against quantized networked control systems," in *2019 IEEE International Conference on Unmanned Systems (ICUS)*. IEEE, 2019, pp. 542–547.

[21] R. Zhang and P. Venkitasubramaniam, "Stealthy control signal attacks in linear quadratic gaussian control systems: Detectability reward tradeoff," *IEEE Transactions on Information Forensics and Security*, vol. 12, no. 7, pp. 1555–1570, 2017.

[22] Y. Chen, S. Kar, and J. M. Moura, "Cyber-physical attacks with control objectives," *IEEE Transactions on Automatic Control*, vol. 63, no. 5, pp. 1418–1425, 2017.

[23] ——, "Optimal attack strategies subject to detection constraints against cyber-physical systems," *IEEE Transactions on Control of Network Systems*, vol. 5, no. 3, pp. 1157–1168, 2017.

[24] C. Kwon, W. Liu, and I. Hwang, "Analysis and design of stealthy cyber attacks on unmanned aerial systems," *Journal of Aerospace Information Systems*, vol. 11, no. 8, pp. 525–539, 2014.

[25] P. Dash, M. Karimibiuki, and K. Pattabiraman, "Stealthy attacks against robotic vehicles protected by control-based intrusion detection techniques," *Digital Threats: Research and Practice*, vol. 2, no. 1, pp. 1–25, 2021.

[26] W. Chen, Z. Duan, and Y. Dong, "False data injection on EKF-based navigation control," in *2017 International Conference on Unmanned Aircraft Systems (ICUAS)*. IEEE, 2017, pp. 1608–1617.

[27] J. Su, J. He, P. Cheng, and J. Chen, "A stealthy GPS spoofing strategy for manipulating the trajectory of an unmanned aerial vehicle," *IFAC-PapersOnLine*, vol. 49, no. 22, pp. 291–296, 2016.

[28] B. D. Anderson and J. B. Moore, *Optimal filtering*. Courier Corporation, 2012.

[29] A. J. Kerns, D. P. Shepard, J. A. Bhatti, and T. E. Humphreys, "Unmanned aircraft capture and control via GPS spoofing," *Journal of Field Robotics*, vol. 31, no. 4, pp. 617–636, 2014.

[30] R. Tunga, C. Murguia, and J. Ruths, "Tuning windowed chi-squared detectors for sensor attacks," in *2018 Annual American Control Conference (ACC)*. IEEE, 2018, pp. 1752–1757.

[31] Y.-L. Wang, C.-C. Lim, and P. Shi, "Adaptively adjusted event-triggering mechanism on fault detection for networked control systems," *IEEE Transactions on Cybernetics*, vol. 47, no. 8, pp. 2299–2311, 2017.

[32] B. Datta, *Numerical methods for linear control systems*. Academic Press, 2004, vol. 1.

[33] F. L. Lewis, D. Vrabie, and V. L. Syrmos, *Optimal control*. John Wiley & Sons, 2012.

[34] Y. Tang, Y. Zheng, and N. Li, "Analysis of the optimization landscape of linear quadratic gaussian (LQG) control," in *Learning for Dynamics and Control*. PMLR, 2021, pp. 599–610.

[35] L. Hu, Z. Wang, Q.-L. Han, and X. Liu, "State estimation under false data injection attacks: Security analysis and system protection," *Automatica*, vol. 87, pp. 176–183, 2018.

[36] DJI, "Tello edu." [Online]. Available: https://www.ryzerobotics.com/tello-edu

[37] ALFA Network, "Awus036ach." [Online]. Available: https://www.alfa.com.tw/products/awus036ach?variant=36473965871176

[38] MATLAB, "System Identification Toolbox," https://www.mathworks.com/products/sysid.html, n.d.