Machine Learning-Based Delay-Aware UAV Detection and Operation Mode Identification Over Encrypted Wi-Fi Traffic

Amir Alipour-Fanid, Student Member, IEEE, Monireh Dabaghchian, Member, IEEE, Ning Wang, Member, IEEE, Pu Wang, Student Member, IEEE, Liang Zhao, Member, IEEE, and Kai Zeng, Member, IEEE.
investigators to restore the course of the incidents, which could be used as court evidences in a legal process and help law enforcement officials to improve countermeasures or responses to various possible UAV incidents.

Many physical detection mechanisms, such as radar [6], [7], acoustic [8], [9], and vision [10]–[12], have been proposed for UAV detection. When using only one of these sensors for detection, these methods may get less effective in some practical scenarios, especially in a crowded urban environment. The radar signals may get blocked by walls, buildings, and other obstacles, which are very common in a civilian environment. The vision detection technique cannot detect the UAV in non-line-of-sight scenarios and dark. The acoustic detection can be interfered by the environment noises which may overwhelm the relative small sound produced by tiny rotor-craft or gliding fixed-wing UAV.

Aiming to complement the above conventional physical detection mechanisms, we propose to explore machine learning-based Wi-Fi traffic identification approaches to achieve fast UAV detection and operation mode identification. It is motivated by the observation that many existing consumer UAVs are equipped with Wi-Fi interfaces and communicate with a user handheld device (e.g., smartphone) for command control or video streaming. Detecting UAVs through wireless traffic identification brings us several advantages over existing mechanisms. First, Wi-Fi signal sensing and packet capturing are less affected by obstacles, other flying objects, acoustic noise, or light conditions that could affect physical detection mechanisms. Second, Wi-Fi data traffic provides cyber information about UAVs’ type and their operation mode, which can be very useful for forensics investigation.

Challenges: At the same time, UAV detection through Wi-Fi traffic identification introduces unique challenges that separate it from traditional traffic identification [13]–[15] and sensing tasks as follows:

1) UAV traffic can be encrypted. Therefore, existing network monitoring and intrusion detection mechanisms that are based on packet header examination or port filtering are not applicable to encrypted UAV traffic. For example, Wi-Fi controlled UAVs (such as DJI and Bebop drones) use WPA2 to secure the wireless communication. Although SSID in the MAC frame may reveal information about the type or vendor of the drone, it can be easily changed through drone control apps. 2) Existing machine learning methods cannot be directly used to identify UAV traffic in a timely manner. For real-time applications, we need to identify the UAV as soon as it is appearing in or approaching to a restricted area. From learning and classification perspective, traditional machine learning methods [13], [14] that only aim at minimizing detection error cannot be directly applied. Detection delay introduced by the computations on feature generation and future packet arrival time should also be considered. 3) Traditional time series early detection strategies [16] cannot be applied to UAV traffic. The inter-packet arrival time of UAV traffic is random, so the traditional time series early detection method which is based on fixed time intervals cannot be directly applied.

To address the above challenges, we propose a delay-aware machine learning-based UAV detection framework to strike a tunable balance between UAV detection accuracy and delay. Our classification framework treats the encrypted data flow as a time series and extracts statistical features only based on the packet size and inter-arrival time. By considering the computation time among different features, our framework adopts a re-weighted $\ell_1$-norm regularization and integrates feature selection and performance optimization in one objective function. To tackle the packet inter-arrival time uncertainty when estimating the delay cost function, we use maximum likelihood estimation (MLE) method to estimate the packet inter-arrival time. Finally, expected total cost function integrates misclassification/misdetection and delay cost which are updated online when a new packet arrives and an optimal detection decision is made to minimize the expected total cost function.

Our main contributions are summarized as follows:

- We propose a machine learning-based framework to achieve delay-aware UAV detection and operation mode identification over encrypted Wi-Fi traffic. This framework extracts features derived only from information of packet size and inter-arrival time. This framework can be applied to other types of encrypted traffic, such as cellular traffic or proprietary protocol traffic as long as the packet size and interval can be measured.
- In order to reduce the model prediction time for fast UAV detection, our framework adopts $\ell_1$-norm regularization and integrates feature selection and accuracy optimization in one objective function, which considers the feature importance and difference of computation time among different features.
- We propose to use model-based MLE method to estimate the packet inter-arrival time. Then using the mean square error (MSE) as a well-known metric, we evaluate the performance of the estimation on the collected real-world dataset.
- Other than detecting and identifying different types of UAVs, our proposed method further identifies the UAV’s operation mode, such as standby, hovering, flying, etc.
- We collect a large amount of real-world encrypted Wi-Fi data traffic of non-UAV and eight types of consumer UAVs, and conduct extensive evaluations on the performance of the proposed methods.

Through comprehensive study, we obtain the following findings:

- The UAV traffic presents different patterns from non-UAV traffic. Therefore, machine learning based methods work well to differentiate UAV traffic from a wide range of non-UAV traffic.
- Due to vendor specific implementation of UAV command control and video streaming protocols, different types of UAVs present different traffic patterns which can be used to classify UAVs from different vendors.
- The UAV Wi-Fi traffic presents different patterns under different UAV operation modes. This finding implies a strong correlation or coupling between cyber information (data traffic) and physical information (operation mode) of UAVs. This finding is expected to motivate new cyber-physical defense and forensics mechanisms
that leverage this cyber-physical coupling. We believe this methodology can be applied to other cyber-physical systems (CPS) and motivate more in-depth study on cyber-physical attack co-detection or co-defense for many Internet of Things (IoT) applications, such as connected cars, smart home, smart healthcare, and industrial control systems.

The rest of the paper is organized as follows. Section II discusses the related works. Problem setup is described in Section III. Section IV presents our proposed method on delay-aware UAV early detection and operation mode identification. Collection and analysis of real-world dataset is described in Section V. Extensive performance evaluation is conducted in Section VI. Section VII provides more detailed discussion on practical aspects of the proposed methodology and possible future work in this area. Section VIII concludes the paper.

II. RELATED WORK

A. UAV Detection Mechanisms

In this section, we briefly describe the existing UAV detection methods.

1) UAV Detection Through Physical Sensing: Existing UAV detection mechanisms mainly focus on physical sensing through various means, including radar, vision, and acoustic.

Radar system is one of the well-known and oldest techniques in aircraft detection dating back to World War II. In order to adapt the detection of small size UAVs, X-band radar systems were proposed [6], [17]. However, in the metropolitan areas (e.g., a city) radar based detection may become less effective due to its line-of-sight requirement [18]. The vision-based UAV detection based on video cameras [19] has the same weakness as radar based techniques, as it also requires line-of-sight between the camera and UAV. However, if the cost is not a much concern, building a detection system based on multiple radars and cameras fusion to cover the targeted area would be a reliable and promising UAV detection system.

The acoustic signal-based UAV detection is a method that can solve the out-of-sight problem [8], [9], [20]. However, this method has its own drawbacks as well. First, the acoustic signal coming from the UAV can be quite noisy due to the noise generated at the motors of electric-powered rotorcraft with fixed wings [9]. Second, other similar acoustic signal generating devices such as electric weed whackers can generate sound signals quite similar to UAV’s. In order to overcome the drawbacks of individual techniques, hybrid solutions have been proposed by combining the acoustic sensor and video camera [21]. Another hybrid solution incorporates the radar sensor as well [22].

2) RF-Fingerprinting Based UAV Detection: Recently, Zhao et al. [23] proposed a new method of RF signal fingerprinting in order to detect and identify the type of the UAVs. To do that, they propose to use Auxiliary Classifier Wasserstein Generative Adversarial Networks (AC-WGANs) based on the wireless signals collected from various types of UAVs. According to their results, this method can detect the UAVs in indoor and outdoor environments with average accuracy of 95% and 80%, respectively. In the other recent work [23], Bisio et al. [24] proposed a Wi-Fi statistical fingerprint-based amateur UAV detection method by applying existing multiclass classification machine learning algorithms. In this work the detection delay is not a concern, and thus, the main goal is to train a machine learning model to detect the intruding UAV based on the predefined and fixed number of statistical features which are computed in every fixed window size. Our proposed method considers detection delay and strike a tunable balance between detection accuracy and delay as well as feature computation time. Ezuma et al. [25] proposed a new detection and classification of micro-UAVs using RF fingerprints of the signals transmitted from the controller to the micro-UAV. In their technique, they utilized wavelet domain analysis to remove the bias in the signals which also helped in the processing data size reduction. For the classification purposes, a naive Bayes approach has been applied to distinguish the UAV signal frames from the non-UAV classes. In the testing phase, a signal energy level detection is also integrated to improve the detection performance. In average, the micro-UAV detection accuracy of 96.3% is achieved under various signal-to-noise ratio (SNR) levels on the channel. We believe integrating our work with the RF fingerprinting method proposed by Ezuma et al. [25] would result in a delay-aware, more robust, and accurate UAV detection system as each method could be a very suitable complement to the other.

A recent work [26] has been proposed to detect the approaching of a UAV within a short distance through the observation of received signal strength (RSS) changes of Wi-Fi signals. However, an intruder UAV may just launch inside a restricted area and hover, or standby on a neighboring roof to spy on someone. In these scenarios, this method will not work. Moreover, a changing RSS is not necessarily introduced by UAVs, but could be other moving objects with Wi-Fi interfaces, such as mobile users carrying smartphones or a driving car equipped with Wi-Fi connections. So the application scenario of this proposed technique is limited.

In another work [27], the authors propose a new RF-based drone detection method based on the physical characteristics of the drone, such as body vibration and body shifting, which impact the wireless signal transmitted by the drone during the communication. This method is not useful when the UAV is in the standby mode. Moreover, both [26] and [27] require line-of-sight connection between the RF signal monitoring system and UAV. Our proposed method based on Wi-Fi traffic identification relaxes this strong assumption.

Recently, Sciancalepore et al. [28] proposed to detect drones’ status, flying or laying on the ground, using encrypted network traffic identification. Three standard binary classification algorithms, Trees-J48 (J48), random forest (RF), and neural networks (NN) are applied to a 3DR SOLO UAV traffic dataset to identify the drone status. Different from this work, in our work we identify eight different operation modes as well as we propose a machine learning framework which addresses the delay-aware UAV detection problem. Our detection system is also tested in the presence of non-UAV data traffic to detect the UAVs used in the experiment.
B. Data Traffic Classification/Identification

Classical approaches such as port-based, payload-based and deep packet inspection can be used to identify the type of the non-encrypted network data traffic. However, nowadays many application data traffic are encrypted for security purposes, and our work is closely related to encrypted data traffic classification/identification. There are several works for identifying the encrypted data flow based on protocol data fingerprinting in wired and wireless networks, where commonly a combination of statistical and machine learning approaches have been used [13]–[15], [29], [30].

In [13], a new support vector machine (SVM) based method is proposed to identify three types of traffic, HTTP, File Transfer Protocol (FTP), and Email. One of the pioneering works in this area applies classification techniques to classify traffic in a wired network into classes of bulk transfer, small transactions, and multiple transactions [14]. Bar Yanai et al. [15] show that it is possible to distinguish the behavior of an application from the observation of the size and the direction of the first few packets of the Transmission Control Protocol (TCP) connection. In this work, three classical clustering algorithms, K-Means, Gaussian Mixture Model and spectral clustering are applied on the dataset to identify the flow. However, this method requires packet header traces analysis, and initial TCP connection packets. Xie et al. [29] proposes a new method called subspace clustering technique (SubFlow), which learns the intrinsic statistical features of each application to classify and identify the flow.

However, our work is different from the existing traffic identification works in the following aspects: 1) Our model provides packet-by-packet analysis, hence the decision is made in a timely manner as packets enter the detection system. 2) Our model adaptively finds the optimal number of the packets that are needed for optimal identification with high accuracy, while considering time cost (or delay). 3) When training the model, feature generation time is also considered and critical features are selected. Therefore, in the prediction/detection, useless features are not generated, and thus the detection delay is reduced.

Compared to our earlier work [31], in this paper, operation mode of the detected UAV is identified. To do so, we collect a large amount of real-world operation mode data traffic of four UAVs and apply multiclass classification machine learning algorithms to identify the modes. We also extend our delay-aware UAV detection test from four to eight commonly used consumer UAVs and conduct extensive evaluations on the performance of the proposed methods. Moreover, we provide performance evaluation for the packet inter-arrival time estimation using MLE. The results indicate that mean square error (MSE) of estimation is reduced when the information of a large number of packets are available in the detection system.

Note that our detection system is applicable to detect the intruding UAVs controlled by a user handheld device (e.g., smartphone). In other words, communication links between the UAV and controller should be established in order to monitor the traffic and detect the UAV by the proposed method. On the other hand, our method will be ineffective if the intruding UAV is equipped with the advanced autonomous systems such as Autonomous Guidance, Navigation and Control (GN&C) where no ground control station is required for command and control.

III. Problem Setup

A. System Setup

There is a Wi-Fi signal sensing and packet capturing system that can collect all the Wi-Fi traffic within a physical sensing range in real-time. There can be multiple Wi-Fi users in the sensing range and multiple UAVs or non-UAV devices. The sensing system may capture non-encrypted packets, for which we assume the system can tell the application types of the corresponding flow by examining the packet headers or contents. Since these non-encrypted packets can be easily identified by existing methods, we will only focus on encrypted Wi-Fi traffic in this paper. For an encrypted Wi-Fi frame, the information we can obtain about the frame is its source and destination MAC addresses, transmitter and receiver MAC addresses, packet size, and packet arrival time together with other MAC header information, such as frame type (control, management, or data), sequence number, and duration/connection ID.

In this paper, we mainly focus on two sequential tasks. 1) Wi-Fi controlled UAV detection: There are a large body of such kind of UAVs on the market, such as DJI, Bobop, DBPower drones, and etc. We assume the Wi-Fi communication between the drone and controller (e.g., smartphone) is encrypted using security protocols, such as WPA2. We assume a drone restricted area (i.e., No Drone Zone) which no UAVs are allowed to enter and operate (see Fig. 1). Our detection system can be implemented in the center of UAV restricted zone to monitor the area and detect any approaching UAVs as quickly as possible with a high accuracy. 2) UAV operation mode identification: For any detected UAV types in the first step, further identify its operation mode. Operation mode consists of standby, hover, forward, backward, and etc. It is noted that, in case of a malicious spying UAV just get turned on or launched inside the restricted zone and stay on in the standby or hovering mode to accomplish the spying mission, our detection system not only can detect the presence of the spying UAV, but it also can identify in which mode the spying UAV is operating.

For the UAV detection and operation mode identification, we only use data frames. We divide the encrypted Wi-Fi traffic into individual flows according to the pair of source
and destination MAC addresses. A unique flow includes the packets between a pair of nodes. The traffic in a flow can be bi-directional or unidirectional. In a real-time scenario, these flows usually interleave with each other in time. The goal of this paper is to identify the UAV data flows when frames are captured and decide the UAV type and its operation mode in a quick manner with high accuracy.

B. Delay-Aware UAV Detection Problem Formulation

The UAV detection over encrypted Wi-Fi traffic can be formulated as a machine learning classification problem. Let’s assume that we can obtain a large training dataset with \( m \) flow traces with each trace having \( n \) consecutive packets. The traces contain UAV and non-UAV flows which are labeled with their corresponding flow types \( y_i \in \mathcal{Y} \), where \( \mathcal{Y} = \{ \text{UAV}_1, \text{UAV}_2, \ldots, \text{UAV}_{v-1}, \text{non-UAV} \} \) and \( v = |\mathcal{Y}| \) denotes the number of class types in set \( \mathcal{Y} \). UAV\(_j\) for \( j \in [1, \ldots, v-1] \) denotes the UAV type \( j \).

Packet size and packet inter-arrival time are two key attributes we extract from these traces for UAV detection and operation mode identification. The sequences of packet size and packet inter-arrival time for the \( i \)-th trace are denoted by \( x_i \) and \( t_i \), respectively. Now, let \( x_i = (x_{i,1}, \ldots, x_{i,n}) \), where \( x_{i,j} \) for \( i = 1, \ldots, m \) and \( j = 1, \ldots, n \) indicates the size of the \( j \)-th packet in the \( i \)-th trace. Similarly, let \( t_i = (t_{i,1}, \ldots, t_{i,n}) \), where \( t_{i,j} \) denotes the inter-arrival time between the \( j \)-th and \( (j+1) \)-th packets (\( j < n \) in the \( i \)-th trace. Define a finite set \( S = \{(x_i, t_i), y_i\}_{i \in [1, \ldots, m]} \) where the pair \( (x_i, t_i) \) represents the packet size and inter-arrival time of the \( i \)-th trace in set \( S \), respectively.

Let \( \hat{x}(t_k) \) denotes the received incoming traffic up to its \( k \)-th packet arrived at time \( t_k \). Assume a set of multiclass classifiers \( \mathcal{H} = \{(h^y_j)_{j \in [1, \ldots, m]} \} \) are trained to classify the incoming traffic flow \( \hat{x}(t_k) \), where \( y \in \mathcal{Y} \). When the Wi-Fi sensing system receives a new packet of the incoming traffic flow \( \hat{x}(t_k) \), its new features are extracted and incorporated in the prediction system. Intuitively, as more packets arrive, more accurate information about the traffic can be gained. On the other hand, collecting more packets introduces longer identification delay. Therefore, there is a trade-off between detection accuracy and delay in the detection process.

Let \( C_1(\hat{y}, \hat{y}) : \mathcal{Y} \times \mathcal{Y} \rightarrow \mathbb{R} \) denotes the test misclassification cost function where \( (\hat{y}, \hat{y}) \in \mathcal{Y} \), and \( \hat{y} = h^y_j(\hat{x}(t_j)) \) is the predicted class label, while the true class label of the incoming flow is \( \hat{y} \). Let \( C_2(t_p) \in \mathbb{R}, p > k \), be the time cost function which indicates the time cost value if UAV detection is postponed up to time instant \( t_p \). Thus, the estimated total cost function is given by

\[
J(\hat{x}(t_k)) = C_1(h^y_j(\hat{x}(t_k)), \hat{y}) + C_2(t_k). \tag{1}
\]

In order to find an optimal trade-off between the detection accuracy and delay, the total cost function \( J \) needs to be minimized. Hence, we formulate the delay-aware UAV detection optimization problem as follows:

\[
p^* = \arg\min_{p \in [k, \ldots, n]} J(\hat{x}(t_p)), \tag{2}
\]

where \( p^* \) indicates the optimal number of the packets that needs to be received from the incoming flow before performing the UAV detection decision, and \( t_p^* \) denotes the \( p^* \)-th packet arrival time. When \( p^* = k \) is the solution of the optimization problem in (2), there is no need to collect more packets because \( J(\hat{x}(t_k)) < J(\hat{x}(t_q)) \) for \( q = k+1, \ldots, n \). Therefore, the detection is performed instantly at time \( t_k = t_{p^*} \). However, when \( p^* \neq k \) is the solution of (2), it means that the total cost is minimized for \( k+1 \leq p^* \leq n \), thus the detection process is deferred to collect more packets from incoming flow. For the notation purposes, let the indicator function \( \mathbb{I}(p^*) = 1 \) when \( p^* = k \); and otherwise \( \mathbb{I}(p^*) = 0 \), where \( p^* = k \) corresponds to the UAV detection at time \( t_{p^*} \).

C. UAV Operation Mode Identification Problem

In the second task, for the detected UAV, we further identify the UAV’s operation mode. Eight UAV operation modes are labeled as \( Z = \{ \text{“Standby”, “Hover”, “Forward”, “Backward”, “Up”, “Down”, “Right”, “Left”} \} \). In this problem, using existing machine learning algorithms a set of multiclass classifiers are trained packet-by-packet on a real-world dataset to minimize the total operation mode misidentification cost.

IV. DELAY-AWARE UAV DETECTION AND OPERATION MODE IDENTIFICATION

In this section, we propose a delay-aware learning-based predictive model in order to solve the problem formulated in (1) and (2). Then, we extend our work to further identify the operation mode of the detected UAV. Fig. 2 illustrates the main workflow of the proposed method.

A. Learning-Based Model Design

Based on the definition of traffic flow dataset \( S \) in Section III-B, we further define the data of incomplete traffic

Authorized licensed use limited to: George Mason University. Downloaded on June 16,2020 at 18:22:09 UTC from IEEE Xplore. Restrictions apply.
flow, where only the first \( j \leq n \) packets are available in the dataset, denoted as \( S_j = \{(x_i^j, \tau_i^j, y_i)\}_{i \in \{1, \ldots, n\}, j \in \{1, \ldots, m\}} \) where \( x_i^j = (x_{i,1}, x_{i,2}, \ldots, x_{i,j}) \) and \( \tau_i^j = (\tau_{i,1}, \tau_{i,2}, \ldots, \tau_{i,j}) \) indicate the sequence of packet size and inter-arrival time of the \( i \)th trace in the \( j \)th subset, respectively. Next, for each dataset \( S_j \), we generate a design matrix of \( X_j = [X_j^1, X_j^2, \ldots, X_j^m]^T \in \mathbb{R}^{m \times 2l} \), where \( X_j^i \in \mathbb{R}^{2l} \) is a row vector as \( X_j^i = \left[ V_1(x_i^j), \ldots, V_l(x_i^j), V_1(\tau_i^j), \ldots, V_l(\tau_i^j) \right] \) where \( V_1(\cdot), \ldots, V_l(\cdot) \) are functions which compute the statistical features of the input samples (i.e., \( x_i^j, \tau_i^j \)). \( l \) denotes the number of features. A list of statistical feature functions with their associated computation formula is shown in Table I.

**TABLE I**

<table>
<thead>
<tr>
<th>Feature</th>
<th>Description</th>
<th>Compt. time</th>
</tr>
</thead>
<tbody>
<tr>
<td>( V_1(x) ): mean</td>
<td>( \mu = \frac{1}{N} \sum_{i=1}^{N} x(i) )</td>
<td>0.672 ( \mu s )</td>
</tr>
<tr>
<td>( V_2(x) ): median</td>
<td>The higher half value of a data sample.</td>
<td>4.365 ( \mu s )</td>
</tr>
<tr>
<td>( V_3(x) ): MAD(^2)</td>
<td>( \text{MAD} = \text{median}(</td>
<td>x(i) - \text{median}(x)</td>
</tr>
<tr>
<td>( V_4(x) ): STD(^2)</td>
<td>( \sigma = \sqrt{\frac{1}{N-1} \sum_{i=1}^{N} (x(i) - \text{mean}(x))^2} )</td>
<td>1.608 ( \mu s )</td>
</tr>
<tr>
<td>( V_5(x) ): Skewness</td>
<td>( \gamma = \frac{1}{N \sigma^3} \sum_{i=1}^{N} (x(i) - \text{mean}(x))^3 )</td>
<td>14.917 ( \mu s )</td>
</tr>
<tr>
<td>( V_6(x) ): Kurtosis</td>
<td>( \beta = \frac{1}{N \sigma^4} \sum_{i=1}^{N} (x(i) - \text{mean}(x))^4 )</td>
<td>14.995 ( \mu s )</td>
</tr>
<tr>
<td>( V_7(x) ): MAX</td>
<td>( H = (\max(x(i)))_{i=1}^{N} )</td>
<td>0.464 ( \mu s )</td>
</tr>
<tr>
<td>( V_8(x) ): MIN</td>
<td>( L = (\min(x(i)))_{i=1}^{N} )</td>
<td>0.652 ( \mu s )</td>
</tr>
<tr>
<td>( V_9(x) ): Mean Square</td>
<td>( MS = \frac{1}{N} \sum_{i=1}^{N} (x(i))^2 )</td>
<td>1.147 ( \mu s )</td>
</tr>
<tr>
<td>( V_{10}(x) ): RMS</td>
<td>( RMS = \sqrt{\text{mean}(x)} )</td>
<td>1.275 ( \mu s )</td>
</tr>
<tr>
<td>( V_{11}(x) ): FS(^3)</td>
<td>( 3(\text{mean}(x) - \text{median}(x)) )</td>
<td>8.911 ( \mu s )</td>
</tr>
<tr>
<td>( V_{12}(x) ): MAD(^2)</td>
<td>( \text{MAD} = \sum_{i=1}^{N}</td>
<td>x(i) - \text{mean}(x)</td>
</tr>
</tbody>
</table>

\(^1\text{MAD}: \text{median absolute deviation}\)

\(^2\text{STD}: \text{standard deviation}\)

\(^3\text{PS}: \text{Pearson skewness}\)

\(^4\text{MAD}: \text{mean absolute deviation}\)

**Algorithm 1** Training Phase Framework for UAV Identification

**Input:** Wi-Fi traffic trace dataset \( \{(x_i(t_n), y_i), (y_i, \gamma) \in \mathcal{Y}, \forall i, j \in \{1, \ldots, n\}\};\)

**Output:** \( E(X_j^i, W_j^i), j \text{ set of classifiers } \mathcal{H} = \{h_j^i\}\);  

**step 1:** Extract packet size and inter-arrival time of encrypted Wi-Fi traffic traces and create dataset \( S = \{(x_i, \tau_i, y_i)\};\)

**step 2:** Define subsets \( S_j = \{(x_i^j, \tau_i^j, y_i)\}, \) where \( x_i^j = (x_{i,1}, x_{i,2}, \ldots, x_{i,j}) \) and \( \tau_i^j = (\tau_{i,1}, \tau_{i,2}, \ldots, \tau_{i,j}) \);

**step 3:** Determine design matrices \( X_j = [X_j^1, X_j^2, \ldots, X_j^m]^T, \) where \( X_j^i = [V_1(x_i^j), \ldots, V_l(x_i^j), V_1(\tau_i^j), \ldots, V_l(\tau_i^j)] \in \mathbb{R}^{2l}; \)

**step 4:** Train a set of classifiers \( \mathcal{H} = \{h_j^i\} \) by solving (3):  

\[
\min_{W_j^i} \mathcal{L}(y, h_j^i(X_j^i, W_j^i)) + \sum_{i=1}^{2l} \lambda_i^i |W_j^i|, \tag{3}
\]

where the strength of penalty \( \lambda_i^i \) for \( i \)th feature in \( j \)th subset is proportional to the computational time for this feature. Therefore, the objective function in (3) will minimize the misclassification error and enforce some \( W_j^i \)'s to be zeros, especially those that consume more computation time. Since for computing of expected misclassification cost function, we shall need probabilistic output of the classifier, then we choose \( h_j^i(\cdot) \) to be one-vs-all logistic regression function [33]. One-versus-all logistic regression is a generalized version of the logistic regression into multiclass classification.

Next, we compute the training expected misclassification cost function on each trace \( i \) for every subset \( j \) as follows:

\[
E_j^i(X_j^i) = \sum_{y_i \in \mathcal{Y}} P(y_i|X_j^i) \sum_{\tilde{y} \in \mathcal{Y}} P_j(\tilde{y}|X_j^i; W_j) C_j^i(\tilde{y}|y_i), \tag{4}
\]

where \( P(y_i|X_j^i) = 1 \) if \( \tilde{y} = y_i \) and 0 otherwise. Then, a set of one-vs-all logistic regression classifiers \( \mathcal{H} = \{h_j^i\} \) for \( j = 1, \ldots, n \) and \( \gamma \in \{\text{UAV1, UAV2, \ldots, UAV}_{n-1}, \text{non-UAV}\} \) are trained. Based on the probabilistic output of one-vs-all logistic function, we can compute \( P_j(\tilde{y} = y_i|X_j^i; W_j^i) = h_j^i(X_j^i, W_j^i)^{C_j^i(\tilde{y}|y_i)} \) denotes the misclassification cost function of training dataset. \( C_j^i(\tilde{y}|y_i) = 1 \) if \( \tilde{y} = y_i \) and 0 otherwise, and \( \tilde{y} = \max h_j^i(X_j^i, W_j^i) \). We summarize the model training phase for UAV identification in Algorithm 1.
### B. Delay-Aware Predictive Model

1) *Expected Missclassification Cost Function C₁*: In the prediction phase of the incoming flow $\tilde{x}(t)$, in order to compute the expected missclassification cost function $C₁$, $E(X_i^j)$ is weighted based on the incoming traffic’s Euclidean distance from every trace in the training dataset. Consider $\tilde{x}(t_k)$ be the incoming flow and $\tilde{X}^k \in \mathbb{R}^d$ its corresponding feature values. The weight function is defined as a normalized sigmoid function by $f_{w_i}^k = \frac{x_i^k}{\sum_l x_l^k}$ where $x_i^k = 1/1 + exp^{-\eta \Delta_i^k}$, and $\eta$ is some positive constant, and $\Delta_i^k = \tilde{D}_i - d_i^k / \tilde{D}_i$ is the normalized average distances between $\tilde{X}^k$ and all the traces in the training dataset [16]. $d_i^k = ||\tilde{X}^k - X_i^k||_2$ indicates the Euclidean distance of the incoming flow from $i$th trace in the dataset. In fact, the weight function $f_{w_i}^k$ plays the role of a *similarity function* which measures how close the incoming traffic flow is to each of the traces in the training dataset. Hence, the expected missclassification cost function for $\tilde{x}(t_k)$ is defined as follows:

$$C_1(h_k^i(\tilde{x}(t_k), \tilde{y})) = \sum_{i=1}^{m} f_{w_i}^k E(X_i^j).$$

The above equation indicates that more weights are multiplied to the training expected missclassification value of the $i$th trace if its distance from the incoming flow is larger and vice versa.

2) *Estimated Time Cost Function C₂*: For the incoming flow $\tilde{x}(t_k)$, future packet arrival times are unknown and random. This uncertainty in packet arrival times introduces difficulties in constructing a delay-aware UAV identification algorithm. In order to tackle this challenge, we propose to estimate the incoming flow’s future packet inter-arrival time according to the exponential distribution with parameter $\tilde{u}_i$ for $i = 1, \ldots, \nu$ achieved by MLE method [34]. In Table II, we present the goodness-of-fit statistics to show that exponential distribution provides a good approximation for packet inter-arrival time estimation. For the illustration purposes, we also graphically show the goodness-of-fit for Bebop 2 and DJI Spark in Fig. 3.

#### Exponential Distribution Parameter Estimation Using MLE:

Let $\{T_j^n\}$ be a sequence of $n$ independent and identically distributed (i.i.d.) exponential random variables. Thus, $T_j^n \sim \text{Exp}($μ$)$ has a probability density function (pdf) of $\lambda(t_j^n) = \mu \exp(-\mu t_j^n)$ for $t_j^n \geq 0$ with parameter $\mu$, where $j = 1, \ldots, n$, $i = 1, \ldots, \nu$ and $\nu = |\mathcal{Y}|$. Given the data sequence $\{T_j^n\}$ our goal is to estimate the average packet inter-arrival time (i.e., $\mu$). Since $T_j^i$ for $i = 1, \ldots, \nu$ and $j = 1, \ldots, n$ are assumed to be i.i.d., then the likelihood function is given by

$$\mathcal{L}(\mu; t_1^n, \ldots, t_n^n) = \prod_{j=1}^{n} f_{T_j^n}(t_j^n; \mu) = \mu^n \exp\left(-\mu \sum_{j=1}^{n} t_j^n\right).$$

By taking logarithm of both sides in (6), we obtain the log-likelihood function as

$$l(\mu; t_1^n, t_2^n, \ldots, t_n^n) = n \ln(\mu) - \mu \sum_{j=1}^{n} t_j^n.$$  

Then, maximum log-likelihood estimation of $\mu$ is achieved by solving the first order maximization problem of

$$\hat{\mu}_i = \text{argmax}_\mu l(\mu; t_1^i, t_2^i, \ldots, t_n^i),$$

as $\frac{d}{d\mu}l(\mu; t_1^i, t_2^i, \ldots, t_n^i) = 0$, which results in

$$\hat{\mu}_i = \frac{n}{\sum_{j=1}^{n} t_j^i} \text{ for } i = 1, \ldots, \nu.$$

Next, we estimate the packet inter-arrival time of the incoming traffic flow $\tilde{x}(t_k)$, through the following steps: First, the Euclidean distance between $\tilde{x}(t_k)$ and each trace in the training set is computed. Second, the class label of the trace which has a minimum distance from the incoming flow is
identified. Third, the average inter-arrival time of the identified class is selected from (9) to estimate the packet inter-arrival time of $\tilde{x}(t_k)$ using the exponential distribution.

Now, let $\tilde{t}_{i+1} = t_{i+1} - t_i$ for $i = k, \ldots, n$ be the packet inter-arrival time of the $\tilde{x}(t_k)$ estimated by the above steps. Then, the estimated time cost function is obtained as

$$C_2(t_p) = \sum_{i=k}^{n} \tilde{t}_{i+1} \text{ for } p = k, \ldots, n \quad (10)$$

where $C_2(t_p)$ is a strictly increasing function.

3) MLE Performance Metric: We use MSE metric to measure the performance of the parameters estimated by the MLE. Considering an incoming traffic flow $\tilde{x}(t_k)$ and letting $\tilde{t}_{i+1}$ for $i = k, \ldots, n$ be the true packet inter-arrival time of $\tilde{x}(t_k)$, we have

$$MSE_p = \frac{1}{n-p} \sum_{i=p}^{n} (\tilde{t}_{i+1} - \tilde{t}_{i+1})^2 \text{ for } p = k, \ldots, n \quad (11)$$

where $MSE_p$ denotes the MSE estimation of packet inter-arrival time of $\tilde{x}(t_k)$ when $p$th packet arrives.

4) Estimated Expected Total Cost Function $J$: According to (1), the total cost function $J$ is defined based on $C_1$ and $C_2$ which can be computed using (5) and (10), respectively. Algorithm 2 summarizes the total cost function estimation and incoming traffic flow’s identification phase.

C. UAV Operation Mode Identification

We identify eight common operation modes for consumer UAVs in the market which are labeled as $\mathcal{Z} = \{ \text{"Standby", "Hover", "Forward", "Backward", "Up", "Down", "Right", "Left"} \}$. UAVs’ operation mode is based on the type of the command they receive from the controller. Each UAV operation mode produces a distinct traffic pattern in the Wi-Fi network. This pattern depends on the type of the command issued by the controller which governs different packet size and inter-arrival time in the trace. Therefore, a multiclass classification model trained on a suitable dataset can identify a UAV’s operation mode. Given a dataset which contains a specific UAV’s Wi-Fi traffic traces labeled with the operation modes mentioned in set $\mathcal{Z}$, two well-recognized multiclass classification algorithms, SVM and random forest (RF) are applied to create the discriminative model. Then, the incoming traffic flow $\tilde{x}(t_k)$ is provided as an input to the corresponding multiclass classifier to identify the operation mode of the detected UAV.

V. DATA COLLECTION AND PREPARATION

A. UAV Detection Dataset

We collect traffic flows from eight types of consumer UAVs shown in Fig. 4: Parrot Bebop 1 Quadcopter Drone (Bebop 1), Parrot Bebop 2 Quadcopter Drone (Bebop 2), DJI Spark (Spark), DBPOWER UDI U842 Predator FPV (UDI), DBPOWER Discovery FPV (Discovery), DJI Tello (Tello), Tenergy TDR Phoenix Mini RC Quadcopter Drone (TDR), and Wingsland Mini Racing Drone (Wingsland). We use a DELL Latitude laptop embedded with a wireless network interface card (NIC), Intel Corporation Wireless 8260, operating in promiscuous mode to monitor and collect the Wi-Fi network traffic. For each UAV type, we collect the UAV traffic while they are flying and streaming video to the controller. To do so, we set the channel frequency of the monitoring sensor in the same channel as the UAV’s operating channel, then run Wireshark version 2.4.11 to capture the Wi-Fi traffic data. Each UAV type dataset contains 3,000 traffic traces with each trace having $n = 200$ consecutive packets.

After collecting the data and identifying the UAVs’ traffic flows, we clean the data and prepare it for the training and testing dataset. In the data cleaning phase, we remove all the broadcast packets (e.g., 802.11 beacon frames), damaged packets and packets with only receiving address (e.g., 802.11 ACK frames). The remaining packets include video streaming, control commands, UAV’s response to the control commands, and UAV status updates such as direction, velocity, height and GPS information. In Fig. 5, we show the packet size distribution of the UAV types used in our experiment.

Note that the data cleaning is performed in order to train and test the classification model offline. However, when testing in the real scenario the incoming traffic may consist of broadcast packets, or packets with only ACK frames as well which all are easily discarded by the predefined filtering option on the packet capturing/monitoring software (i.e., Wireshark) before entering to the delay-aware UAV detection system.
B. Non-UA V Dataset

In an effort to make a diverse non-UA V dataset, we create a dataset which consists of two main sub-dataset: First, we use the Wi-Fi data traffic available online from CRAWDAD database [35]. We choose this dataset because of the following reasons. 1) This dataset consists of live and non-live video streaming traffic captured from commonly seen popular applications such as Google Hangouts, ooVoo, Skype, TED and Youtube. 2) The traffic data are collected from a smartphone app where the user makes a diverse set of mobility patterns. Second, we have also captured encrypted Wi-Fi traffic on a university campus Wi-Fi network where a mixed multiple traffic types such as video streaming, social network apps, VoIP, email, web browsing applications are usually running. If the UAV identification system is set up on the campus, our method should be able to differentiate UAV traffic from these non-UA V traffic. The non-UA V dataset (Google Hangouts, ooVoo, Skype, TED, Youtube, and Campus traffic) also contains 3,000 traffic traces with $n = 200$ consecutive packets.

C. UA V Operation Mode Dataset

The following steps are taken for an operation mode data traffic collection of a specific UA V type: 1) Wi-Fi connection is established between the UAV and controller. 2) A specific operation mode command (e.g., “Forward”) is given via controller to the UAV and is held. 3) Wi-Fi medium monitoring sensor is activated to monitor the wireless channel traffic. 4) Wireshark is run on the promiscuous mode to capture the packets. 5) Before releasing the command in the controller, first, Wireshark is stopped, and then the collected traffic is saved and labeled according to the commanded operation mode. 7) This process is repeated for all the operation modes until enough data traffic is collected.

VI. PERFORMANCE EVALUATION

A. Learning-Based Model Performance Evaluation

By randomly sampling the dataset, we split the whole dataset into training and testing datasets with the ratio of 70% and 30%, respectively. We create subset $S_j$ for $j = 5, \ldots, 200$ by adding packet-by-packet information to each subset $j$ according to the step 2 in Algorithm 1. Then, we form the design matrix $X_j$ for $j = 5, \ldots, 200$ by extracting $2^l = 24$ statistical feature values listed in Table I. One-vs-all logistic regression multiclass classification algorithm with re-weighted $l_1$-norm technique proposed in (3) is run over each design matrix $X_j$.

Fig. 6(a) illustrates the accuracy of the classification algorithm in training and testing on each design matrix $X_j$. The shaded areas denote the regions surrounded by one standard deviation above and below the mean accuracy. The results show that a mean testing accuracy of higher than 88% is achieved when the information of fifty or more packets ($j > 50$) are available in the subset. The model learning process also confirms the intuition that as more consecutive packets are available in the subset, the mean accuracy of predictive model increases. F-measure (weighted harmonic mean of precision and recall) for different UA V types and non-UA V is shown in Fig. 6(b). Average F-measure of higher than 86% is achieved on the test data for $n = 200$. Fig. 7 also shows the corresponding precision and recall which indicate an acceptable discriminative power of the trained classifiers.

As a closely similar and related multiclass classification algorithm to one-vs-all logistic regression, we apply linear discriminant analysis (LDA) statistical method on the dataset [33]. This method is a generalized version of statistical Fisher’s LDA. The LDA method finds a linear combination of the features to distinguish different classes in the
dataset. The output of this analysis is shown in Fig. 6(c) for \( n = 200 \). Various types of linear feature combination for different classes are shown in this figure with arrows followed by its associated feature index \((X_1, X_2, \ldots, X_{24})\). Successive discriminant function in the LDA analysis provides four proportions \( LD_1 = 0.70, LD_2 = 0.1448, LD_3 = 0.0480, \) and \( LD_4 = 0.0232 \), which describes the proportion of between-class variances. It is well visualized in this figure that how UAV types and non-UAVs are distinguished on \( LD_2 \) versus \( LD_1 \) as a result of the features’ linear combination.

### B. Feature Selection and Computation Time

Our objective function in (3) jointly minimizes the misclassification error and runtime by discarding useless features. Fig. 8(a) shows the set of selected features for each model trained on \( j \)th subset for \( j = 5, \ldots, 200 \). As it is shown in the figure, when the sample size is small (e.g., \( n < 30 \)), all the features are selected by the model. This is because, on the one hand, small sample size does not provide enough information to the classifier to distinguish different classes with high accuracy, and on the other hand, it consumes less amount of time to compute the feature values. However, as the sample size increases, misclassification error is reduced and the feature computation time increases which results in the smaller set of selected features by the algorithm.

In order to evaluate the impact of feature selection method on the prediction time, we select 1,000 traces uniformly at random from the UAV dataset and consider them as incoming flows (i.e., \( X(t_k) \)). Then, we compute the feature generation time of the flows for \( k = 5, \ldots, 200 \). Fig. 8(b) illustrates the mean total feature generation time of the flows versus the number of the packets in the trace with and without feature selection method. The shaded area denotes one standard deviation above and below the mean total computation time. The results show that as the number of the packets increases, with feature selection, the mean total feature generation time oscillates in a non-increasing trend depending on the number of selected features. However, without feature selection, the total computation time increases when the number of packets increases. Therefore, the proposed feature selection method reduces the prediction runtime despite the fact that the sample size is increasing.

### C. MLE Performance Evaluation

For each UAV type, we select 1,000 traces uniformly at random from the UAV dataset. We consider the selected traces as incoming traffic flows. Then, we follow the step 2 in Algorithm 2 to estimate the packet inter-arrival time of each flow. Using (11), we evaluate the MLE-based estimation performance. Fig. 8(c) shows the mean MSE between the true and estimated packet inter-arrival time with shaded area of one standard deviation. The results show that as more packets arrive, the mean MSE decreases. This means that as more packets are captured, the cost function \( C_2 \) estimation improves as the estimation accuracy of inter-arrival time enhances. This results in achieving high quality delay-aware UAV identification.

### D. Delay-Aware UAV Identification Test

We consider eight types of incoming UAV traffic flows each belonging to a specific UAV type, and run the delay-aware UAV detection algorithm on them. Due to the space limitation, we only show the test results for Bebop 1 in three steps Fig. 9(a), (b), (c), and summarize the outcome in the far right table in Fig. 9(d).

In Fig. 9(a), \( k = 10 \)th packet arrives at time \( t_k = 20.56 \text{ ms} \) and based on the received traffic flow till then, total cost function \( J \) is estimated. In this case, it is estimated that the minimum total cost function will occur when \( p^* = 71 \)th packet arrives at \( t_{p^*} = 139.41 \text{ ms} \). Therefore, the decision for the flow detection is deferred. The far right column of the table in Fig. 9(d) indicates that if the UAV identification is performed in \( k = 10 \), then the detection probability will be 42.54% (\( Pr = 0.4254 \)). In Fig. 9(b), \( k = 40 \)th packet arrives at time \( t_k = 84.75 \text{ ms} \). In this case, the algorithm estimates that the minimum expected total cost function will occur when \( p^* = 74 \)th packet arrives at \( t_{p^*} = 142.94 \text{ ms} \). If the identification is performed in \( k = 40 \), then with the probability of \( Pr = 0.6891 \) the flow will be detected as a Bebop 1 traffic flow. This process is continued until the arrival...
Fig. 9. Delay-aware UAV identification using algorithm 2.

<table>
<thead>
<tr>
<th>Traffic</th>
<th>$E[p^*]$ (Packet)</th>
<th>$E[t_{p^*}]$ (ms)</th>
<th>Accuracy (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bebop 1</td>
<td>87 (±8)</td>
<td>160.43 (±10.01)</td>
<td>87.84 (±1.20)</td>
</tr>
<tr>
<td>Bebop 2</td>
<td>95 (±13)</td>
<td>151.91 (±18.82)</td>
<td>90.75 (±1.74)</td>
</tr>
<tr>
<td>Spark</td>
<td>93 (±11)</td>
<td>142.80 (±15.57)</td>
<td>95.23 (±0.69)</td>
</tr>
<tr>
<td>UDI</td>
<td>141 (±21)</td>
<td>350.79 (±23.41)</td>
<td>85.76 (±2.38)</td>
</tr>
<tr>
<td>Discovery</td>
<td>94 (±3)</td>
<td>131.11 (±10.42)</td>
<td>92.52 (±0.85)</td>
</tr>
<tr>
<td>Tello</td>
<td>72 (±7)</td>
<td>121.65 (±31.30)</td>
<td>93.68 (±1.01)</td>
</tr>
<tr>
<td>TDR</td>
<td>68 (±13)</td>
<td>101.77 (±12.75)</td>
<td>89.66 (±2.10)</td>
</tr>
<tr>
<td>Wingsland</td>
<td>75 (±18)</td>
<td>92.46 (±21.22)</td>
<td>94.39 (±1.83)</td>
</tr>
</tbody>
</table>

Table III: Tested UAVs’ Identification Performance

This will be our next challenging problem to tackle the UAV detection using traffic identification when the traffic suffers from packet loss.

F. UAV Operation Mode Identification Evaluation

Consumer UAVs’ operation mode capabilities maybe different from each other depending on the vendor specifications and manufacturing model. Here, for the UAV types, Bebop 1, Bebop 2 and DJI we identify eight common and popular operation modes as $Z = \{"Standby", "Hover", "Forward", "Backward", "Up", "Down", "Right", "Left"\}$. However, FPV does not support the “Hover” mode, so we exclude this mode from set $Z$ for this type of UAV.

In order to identify the operation mode of these UAVs, we apply SVM and RF multiclass classifiers on the collected real-world data traffic. For each UAV type, we train the SVM and RF predictive model packet-by-packet for $n = 10, \ldots, 300$ by tuning the best model parameters for each subset. For the SVM classification method, we utilize radial basis function (RBF) kernel and tune the best model parameters. For Bebop 1, Bebop 2, DJI, and FPV the total number of operation mode traffic traces in the training dataset is equal to 9600, 9600, 9600, and 8400, respectively. By randomly sampling each dataset, we split the whole dataset into training, cross validation and test datasets with the ratio of 60%, 20% and 20%, respectively. Using 10-fold cross validation repeated three times the best model parameters are tuned. For example, for Bebop 1’s operation mode_identification when $n = 300$, the best model tuned parameters are $C = 64$ and $\epsilon = 0.15$ with the number of support vector machines of $\{173, 139, 142, 131, 126, 174, 150, 142\}$ for each operation mode in set $Z$, respectively.
Fig. 11 illustrates the accuracy of the classification in training and testing for the four UAV types when SVM and RF are utilized for operation mode identification. The gray (darker) and green (lighter) lines denote the mean accuracy of training and testing with the shaded area of one standard deviation, respectively. Since the SVM and RF models are trained and cross validated for parameter tuning on different UAV types operation mode dataset, training and testing accuracy varies to some extent for each UAV type. However, both SVM and RF methods show an acceptable accuracy in effectively distinguishing the tested UAVs’ operation modes.

Fig. 12 illustrates the results of the feature importance analysis for various number of packets in the set. We can see that for different UAVs, the most important feature sets can be different. For example, for Bebop 1 operation mode identification, $V_5(x)$, $V_{12}(x)$ (skewness and MAD of packet size), $V_{17}(r)$, and $V_{24}(r)$ (skewness and MAD of packet inter-arrival time) are indicating high importance value. However, for Bebop 2, $V_4(x)$, $V_{11}(x)$ (STD and PS of packet size), $V_{21}(r)$, $V_{22}(r)$, $V_{24}(r)$ (Mean Square, RMS and MAD of packet inter-arrival time) are showing high importance value. This indicates that: 1) For any given UAV type the data traffic patterns for various operation modes are different. 2) The operation modes of each UAV type follows a different data traffic pattern than the other UAV types. 3) In order to train an effective model for the UAV operation mode identification, it is safe to consider all of statistical features so that the model can freely choose an effective set of important features which provides higher discriminative power.

Confusion matrices for operation mode identification of four tested UAVs when $n = 300$ is shown in Fig. 13. This figure indicates the overall performance of the SVM and RF multiclass classification algorithms. In each confusion matrix, the diagonal and off-diagonal cells correspond to operation modes that are correctly and incorrectly identified, respectively. The right most column of the matrix indicates the precision (positive predictive value) and false discovery rate, as the top and bottom values of each cell, respectively. Similarly, the bottom row of the matrix shows the recall (true positive rate) and false negative rate, in top and bottom part of each cell, respectively. Lastly, the cell in the most bottom right of the matrix, indicates the overall operation modes.
identification accuracy and error, respectively. As the results show, the operation modes of the UAVs can be accurately identified with high accuracy of $88.5-98.2\%$ through wireless traffic fingerprinting.

VII. DISCUSSION

A. Significance of UAV Early Detection

Considering that a consumer UAV can fly at 50-70mph and some racing UAVs could even fly above 150mph, a delay of one second will translate to a flying distance of 22m to 66m, which can be significant in practice for incident responses and safety/privacy protection. Therefore, reducing the detection delay is paramount important in the UAV invasion detection application. Another improvement in detection time can be achieved using features’ computational dependencies property [36], [37]. Statistical features shown in Table I are computationally dependent, so new techniques proposed in [36] can be applied to reduce the detection time. If number of the features are large, then large-scale feature computational dependency graph proposed in [37] can be employed to reduce the detection time as much as possible.

B. Applicability to Other Communication Protocols

This work is based on the observation that many of consumer UAVs utilize Wi-Fi communication protocol for remote pilot control and video streaming. However, some type of consumer UAVs (e.g., DJI Phantom, 3DR Solo, Yuneec) may use other types of custom communication protocols such as Lightbridge, Sololink and Yuneec protocol. Our proposed framework is applicable to other types of consumer UAVs that use different communication protocols. Our framework works for encrypted wireless traffic and only needs packet size and inter-arrival time information (no need to use any packet content information). As long as we can obtain this information, our framework can be applied to not only detect the UAV using the proposed delay-aware mechanism, but also identify its operation mode.

In this paper, the hypothesis is that UAVs present unique traffic patterns that can be separated from other non-UAV traffic due to their use of a different set of communication protocols and physical operation. We believe smart IP camera and handheld smartphone gimbal using a different set of communication protocols will be separable from UAV traffic as well.

C. Recognizing New Types of UAVs

In this paper, we applied supervised learning frameworks which can classify the known classes (UAVs) appeared in the training set. It will be interesting to extend our work to recognize new types of UAVs (unseen classes). It belongs to the open set recognition problem which is still an open research problem in machine learning areas. Existing technologies including [38] could be explored to recognize new types of UAVs and at the same time reducing the model retraining overhead. Although only a limited number of types of UAVs are tested in this work, the proposed framework should be able to handle a large dataset of different UAV subtypes. Through experiments, this paper has demonstrated the discriminative power of the proposed classifier, which indicates the effectiveness of proposed methodologies. The users are free to adjust dataset to cover different applications on different UAV subtypes. Through experiments, this paper has demonstrated the discriminative power of the proposed classifier, which indicates the effectiveness of proposed methodologies. The users are free to adjust both UAV and non-UAV dataset to cover different application scenarios (e.g., university campus, government building, airport, etc).
D. More Sophisticated Scenarios

The framework developed in this paper could be extended to tackle more sophisticated scenarios, such as simultaneous detection of UAVs operating on multiple channels. A more powerful adversary could even hop among different channels to escape from detection. Some multi-channel network monitoring mechanisms [39] could be integrated in this scenario. An intelligent adversary could change its traffic pattern by injecting packets to avoid being detected. However, this kind of adversary could be limited by energy budget (i.e., limited number of packets can be injected due to limited battery capacity) and mission requirement (i.e., genuine command control packets and video streaming packets cannot be suppressed). Therefore, an enhanced machine learning model which can effectively test sub-traffic could still be effective when facing such an intelligent adversary. Furthermore, investigating the possibility of combining traffic information and physical layer information (such as RSS and modulation schemes) to enhance the identification performance and enable UAV localization is of great interest.

VIII. CONCLUSIONS

Detecting and identifying consumer UAVs is of utmost importance for regulation enforcement, forensics investigation, public security, and personal privacy protection. To complement existing physical detection mechanisms, we proposed a delay-aware machine learning-based UAV detection and operation mode identification framework over encrypted Wi-Fi UAV traffic. This framework extracts features from packet size and inter-arrival time and in the model training phase adopts re-weighted l_1-norm regularization with consideration of computation time among various features. Therefore, feature selection and performance optimization are integrated in one objective function. To deal with packet inter-arrival time uncertainty when estimating the cost function, we utilized model-based MLE method to estimate the packet inter-arrival times of the incoming flow. We collected a large amount of encrypted Wi-Fi traffic of eight types of consumer UAVs and conducted extensive evaluation on the performance of our proposed methods. Experimental results show that the proposed methods can detect and identify tested UAVs within 0.15 – 0.35s with the accuracy of 85.7 – 95.2%. The UAV detection range is within the physical sensing range of 70m and 40m in the line-of-sight (LoS) and non-line-of-sight (NLoS) scenarios, respectively. The operation modes of UAVs can also be well identified with accuracy in the range of 88.5 – 98.2%. The operation mode identification reveals the cyber-physical coupling property of UAVs. Based on this coupling, we can infer information on the physical status (operation mode) of UAVs given information on their cyber part (Wi-Fi traffic data).

Although this work uses Wi-Fi traffic to detect and identify consumer UAVs, we believe the proposed machine learning-based detection framework and methodology are general enough to be applied to other cyber-physical/IoT systems using different wireless communication technologies (e.g., Bluetooth and cellular). We hope this work to shed light on the cyber-physical attack co-detection or co-defense for many other CPS/IoT systems.

REFERENCES


Amir Alipour-Fanid received the B.S. degree in electrical engineering from the Islamic Azad University of Ardabil, Ardabil, Iran, in 2005, and the M.S. degree in electrical engineering from the University of Tabriz, Tabriz, Iran, in 2008. He is currently pursuing the Ph.D. degree with the Electrical and Computer Engineering Department, George Mason University, Fairfax, VA, USA. His research interests include machine learning applications in cybersecurity, cyber-physical systems (CPS) security, the Internet of Things (IoT) security, vehicle-to-vehicle (V2V) communication, and 5G wireless networks.

Ning Wang received the Ph.D. degree in information and communication engineering from the Beijing University of Posts and Telecommunication, Beijing, China, in 2017. He was an Engineer with Huaxin Post and Telecommunications Consulting Design Co., Ltd., Hangzhou, Zhejiang, China, from 2012 to 2013. He is currently a Post-Doctoral Scholar with the Electrical and Computer Engineering Department, George Mason University, Fairfax, VA, USA. His current research interests include physical layer security, machine learning, device identification, and RF fingerprinting.

Pu Wang received the B.S. degree in telecommunications engineering from Xidian University, Xi'an, China, in 2014, where he is currently pursuing the Ph.D. degree in cyber engineering. He was a Visiting Ph.D. Student with the Department of Electrical and Computer Engineering, George Mason University, from November 2017 until November 2019. His current research interests include backscatter communication, wireless information and power transfer, and physical layer security.

Liang Zhao received the B.S. and M.S. degrees from Northeastern University, China, and the Ph.D. degree from Virginia Tech, USA. He is an Assistant Professor with the Information Science and Technology Department, George Mason University. His research interests include spatial data mining, deep learning on graphs, sparse feature learning, interpretable machine learning, and nonconvex optimization.

Kai Zeng received the Ph.D. degree in electrical and computer engineering from the Worcester Polytechnic Institute (WPI) in 2008. He was a Post-Doctoral Scholar with the Department of Computer Science, University of California at Davis (UCD), from 2008 to 2011. He was an Assistant Professor with the Department of Computer and Information Science, University of Michigan, Dearborn, from 2011 to 2014. He is currently an Associate Professor with the Department of Electrical and Computer Engineering, Cyber Security Engineering, and the Department of Computer Science, George Mason University. His current research interests include spatial data mining, deep learning on graphs, sparse feature learning, interpretable machine learning, and nonconvex optimization.