Towards Effective Swarm-Based GPS Spoofing Detection in Disadvantaged Platforms

Enguang Fan, Anfeng Peng, Matthew Caesar
Department of Computer Science
University of Illinois at Urbana-Champaign
Urbana, IL, USA
{enguang2,anfengp2,caesar}@illinois.edu

Jae Kim, Josh Eckhardt, Greg Kimberly, Denis Osipychev
Boeing Research and Technology
Bellevue, WA, USA
{jae.h.kim,josh.d.eckhardt,greg.kimberly,denis.osipychev}@boeing.com

***Abstract*— Modern battlefields are subject to spoofing of GPS signals. While large aircraft platforms can counter the effects of GPS spoofing via redundant and dissimilar sensors, the disadvantaged nodes with smaller platforms such as Air Launched Effects (ALE), with more limited capabilities, can be vulnerable. That said, other sensors on the platform may give clues to the drone about where it is located. In this paper, we investigate the ability of sensor fusion to remediate spoofing of GPS signals for ALE platforms. We first conduct performance comparison among several complementary techniques, including the use of inertial measurement units (IMUs), communication with nearby ALEs (to compare GPS readings), and received signal strength from networking connections (to estimate distance to neighboring ALEs, etc.) We then propose a novel architecture that performs sensor fusion to intelligently combine observations across multiple sensors so as to maximize the ability to detect GPS spoofing as well as to reconstruct coordinates with confidence levels. From a simulation study based on real-world mobility and sensor traces, we find that our approach can improve location estimates accuracy by multiple orders of magnitude as compared to simple baseline techniques, supplementing the ability for ALEs to navigate and execute missions in GPS-denied environments.**

I. Introduction

Modern battlefields are highly contested environments with an electronic and cyber threat that is continually advancing in capability. Due to the advent of software-defined radio capabilities as well as increasingly sophisticated hardware, the ability to "jam" or spoof communications in advanced forms is increasingly within reach of even moderately capable enemies. Unfortunately, ability of modern warfighters to conduct missions is ever more intertwined with satellite-hosted localization systems such as the Global Positioning System (GPS) [1]. Such systems are crucial for planning, logistics and routing of assets, as well as navigation capabilities for autonomous platforms such as UAVs. While large aircraft platforms may make use of technologies to counter the effects of GPS spoofing via redundant and dissimilar radio stacks, smaller platforms such as Air Launched Effects (ALEs), with more limited capabilities, can be very vulnerable. Disrupting location services for such platforms can incur dire effects, from termination of missions to capture of classified technologies.

While these challenges harm the ability of the warfighter to trust GPS, the architecture of ALEs may provide avenues towards a solution. In particular, ALEs often contain a broad spectrum of sensors of various sorts, from cameras to gyroscopes and LiDAR. Such sensors provide observational capabilities that may provide clues as to the ALE's location during an attack. For example, an ALE may be able to use its Inertial Measurement Unit (IMU) to estimate how far it has turned and traveled since an attack began; or a camera may be able to estimate altitude.

In our work, we investigate the question: is it possible to combine data from low-accuracy non-positional sensors of a drone to infer its position accurately in a GPS-denied environment? To achieve this, we propose a novel framework that integrates multiple sensor inputs, and derives the ALE's updated location. Our approach works by firstly developing analytical models that derive positional inferences from individual sensor inputs. These inferences are then fed into a learning pipeline, where data is intelligently cleaned, smoothed, and then fused leveraging an Error-State Extended Kalman-Filtering (ES-EKF) based approach. Our fusing approach comprises two stages, first combining sensors within a single ALE, then by leveraging observations from neighboring ALEs to further refine the result. The pipeline then outputs the combined positional inference along with a confidence value, which the warfighter (or software running on the ALE) can use to estimate location, as well as to determine how much they can trust their received GPS signal in contested battlefield environment scenarios. To verify our approach, we evaluate and cross-validate with both emulation (using real-world GPS and IMU data) and simulation (using an event-driven wireless simulation environment). Our results show the ES-EKF based fusion algorithm with low accuracy sensors inputs can achieve a substantially higher accuracy than single sensor alone. Specifically, the sensor fusion method for commercial grade RSSI receiver and IMU can achieve up to a two order-of-magnitude increase in positioning accuracy compared to dead reckoning with tactical grade IMU. Fusing outputs across multiple disparate sensors also greatly reduces variance in position uncertainty, and can leverage collaboration between neighboring drones to further improve performance.

The rest of this paper is structured as follows. Section II describes several works related to our undertakings. Section III introduces our design, including architectural and analytic contributions. We then present our evaluation results in Section IV

and conclude in Section V.

II. Related Work

Localization involves determining the location of a device or object in a given environment. Numerous approaches have been proposed to address this challenge. GPS [1] is widely used for outdoor localization. Current NAVSTAR GPS receivers receive signals from satellites and use triangulation to calculate the device's position. GPS suffers from limited accuracy from GPS-denied technologies and atmospheric conditions. Another widely used technology is range-based localization, including time of arrival (ToA), time difference of arrival (TDoA), angle of arrival (AoA), or received signal strength indicator (RSSI). E-911 [2] is a system used to locate a phone during emergency calls. It determines the phone's location by intersecting the lines formed by AoA from each cell tower. SLAM creates a real-time map using visual sensors such as camera or non-visual sensor such as LiDAR and track the position or movement of the device. Work from Alonso et al [3] demonstrates that visual SLAM updates the device's position and orientation on a digital map for localization purpose. However, large-scale environments, motion distortion, and computing resource requirement restricts the scenario where visual SLAM can be used [4]. Non-visual SLAM can also generate maps for localization purposes. Tam et al. [5] use LiDAR to get 3D pointclouds registration and combine it with an offline map for localization. However, environmental effects and costs associated with advanced sensors impose restrictions on its practical usage. There has also been prior work on using individual sensors to estimate location of an asset. Angelino et al. [6] utilizes IMUs to gather information on the UAV's attitude changes, but does not attempt to handle the issue of error amplification over time. Some works [7] use visual-based sensors for height estimation but face challenges in scenarios with occlusions in the field of view. Received Signal Strength Indicator (RSSI) may also be used to estimate distance between a single pair of nodes, but such approaches face challenges in environments which subject to interference and suffer from varying accuracy over different ranges. Zhang et al's work [8] leverages LiDAR and UWB and Marquez's work [9] fuses UWB and IMU for an indoor scenario over sensors co-located within a small and controlled area. In our work, we explore the ability of sensor fusion technologies to overcome shortcomings of individual sensor inferences, as well as to explore the tradeoffs of using different sensors in varied environments. We leverage these insights to build a multi-sensor framework to protect ALEs from GPS-based spoofing and jamming attacks that may arise in battlefield scenarios.

III. System Design

The goal of our system is to provide a way to efficiently combine and leverage non-positional sensor outputs to form intuition on the drone's current location and heading. Then, by comparing between the estimated location of the drone and its associated confidence level with the corresponding GPS coordinate, it becomes possible to ascertain spoofed GPS signal. To do this, we need some way to translate from varying kinds of sensors into positional data, as well as some way to combine those various inferences together into a single inference on position.

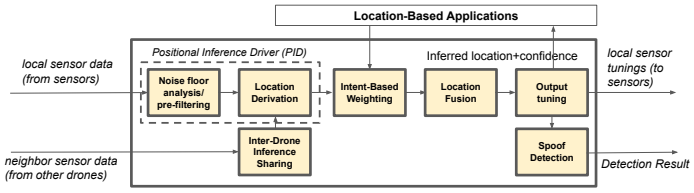

Fig. 1. Overview of system architecture.

To do this, we propose a location inference system, as shown in Figure 1. There are two key algorithmic steps that underlie our approach, which are described as follows. First, we need some way to convert sensor data into location information. To do this, our design leverages the concept of a *positional inference driver* (PID). Each sensor is equipped with a PID; these PIDs acquire outputs from individual sensors and reconstruct position information. PIDs differ for each sensor type, as the manner in which we infer location information from a camera is likely to be different from the way we infer from an IMU. PIDs are run independently for each sensor, as each PID maintains state (which it leverages to improve its inferences) which is specific to that specific sensor. Each PID acquires data at a certain frequency (determined by the sensor type and measurement constraints). Each PID then performs a pre-filtering stage, to clean the data, downsampling to remove biases, as well as smoothing noise to improve signal. The use of PIDs enables a modular architecture; different ALE platforms may have different sensor types, PIDs allow us to interface with the sensors that are available in a modular and extensible fashion.

Second, once location information is inferred from each of the sensors, we need to decide which sensors to actually listen to. To do this, we leverage *sensor fusion*, to combine perceived location information in an effective manner. In practice, we may want to listen to some sensors more than others (e.g., IMU inferences may become less accurate over time due to accumulated error, but may be very accurate soon after GPS signal loss). To achieve this, we adapt the Error State Extended Kalman Filter (ES-EKF) [10] [11]. Our Kalman Filter-based model takes data from multiple individual sensors, then compares the data from each sensor to compensate the measurement noise from each individual sensor. If desired, sensor inputs received from other drones may be fused as well in this step. The fusion model then outputs the estimated location of the drone with a covariance matrix representing the confidence level of the estimation. Our fusion model provides the capability to incorporate human guidance through the acceptance of weights, as humans may have domain or scenario-specific information (e.g., if the drone is entering a space of known signal interference, it may be desired to reduce reliance on RSSI measurements in advance of entering

the area). System applications running on the drone can then use these measurements to both determine location as well as estimate confidence in location measurements over time as they perform their missions.

A. Positional Inference Drivers

To construct a positional inference driver, we need a way to translate location "hints" collected from the sensor into a continuous stream of inferred locations. The manner in which this is done depends on the sensor type. For example:

1) An IMU outputs the angular speed and acceleration on three axes, which allows us to interpolate location information of the drone after proper integration operations [6]. A unique advantage of IMU is its robustness, as an intrinsic sensor, IMU does not rely on extrinsic information from its surroundings which make it difficult to be jammed or spoofed.
2) A camera provides a sequence of digital images, where we can use visual odometry [4] [7] [12] to estimate the displacement between frames and hence infer the position and attitude changes (image features are extracted and tracked across continuous images).
3) LiDAR, consisting of multiple laser transmitters and receivers, can precisely estimate the distance to an object, thereby allowing the drone to construct a map for its surroundings. While camera and LiDAR performance may be more subject to environmental conditions, they provide a method to continuously update location over time.
4) Signal strength (RSSI) to neighboring drones or ground stations may provide estimates about distance [13]. Triangulation may be used to narrow down the set of possible locations of the drone [14].

In this subsection, we describe PID designs for two important sensor types: Inertial Measurement Units (IMUs), and Received Signal Strength Indication (RSSI).

1) IMU-based positional inference: A typical IMU is a 6 degree of freedom sensor that consists of a gyroscope and an accelerometer. The gyroscope and accelerometer provide the angular speed and translational acceleration on three axes, both under the body frame [6] [10] [11] [15].

$$\begin{aligned} \mathbf{R}_{t+1} &= \mathbf{R}_t \exp\left((\boldsymbol{\omega}_t dt)_\times\right), \\ \mathbf{v}_{t+1} &= \mathbf{v}_t + \left(\mathbf{R}_t \mathbf{a}_t - \mathbf{g}\right) dt, \\ \mathbf{p}_{t+1} &= \mathbf{p}_t + \mathbf{v}_t dt + \frac{1}{2}\left(\mathbf{R}_t \mathbf{a}_t - \mathbf{g}\right) dt^2. \end{aligned} \tag{1}$$

Given known initial location and attitude, Equation 1 transforms the data from the IMU into the current pose and speed information by observing rotation and acceleration across time. We employ the direction cosine matrix $\mathbf{R}_t$ to represent the attitude of the drone. Additionally, we use $\mathbf{v}_t$ and $\mathbf{p}_t$ to represent the velocity and position of the drone at time t, respectively. We then use $\exp\left((\boldsymbol{\omega}_n dt)_\times\right)$, an exponential map of a cross product matrix, to represent the corresponding cosine rotation matrix for attitude changes from time t to $t+1$. The $\mathbf{R}_t \mathbf{a}_t$ transforms the acceleration from the body frame into the reference frame where we then compensate for gravity g.

The gyroscope and accelerometer noise models are given as follows:

$$\begin{aligned} \boldsymbol{\omega}_m &= \boldsymbol{\omega} + \omega_{\boldsymbol{b}} + \omega_{\boldsymbol{n}}, \\ \boldsymbol{a}_m &= \mathbf{a} + \mathbf{a}_{\boldsymbol{b}} + \mathbf{a}_{\boldsymbol{n}}, \end{aligned}$$

where $\boldsymbol{\omega}_m$ and $\boldsymbol{a}_m$ represent measured angular velocity and acceleration, $\boldsymbol{\omega_b}$ and $\mathbf{a_b}$ represent the bias, $\boldsymbol{\omega_n}$ and $\mathbf{a_n}$ represents zero-mean Gaussian noise. The non-static bias $\boldsymbol{\omega_b}$ and $\mathbf{a_b}$ are driven by a random walk process, where $\epsilon_{\boldsymbol{r}}^{\boldsymbol{a}}$, $\epsilon_{\boldsymbol{r}}^{\boldsymbol{\omega}}$ are Gaussian noises.

$$\begin{aligned} \delta\omega_{\boldsymbol{b}} &= \omega_{\boldsymbol{b}}^{\boldsymbol{t}} - \omega_{\boldsymbol{b}}^{\boldsymbol{t-1}} = \epsilon_{\boldsymbol{r}}^{\boldsymbol{\omega}} \\ \delta a_{\boldsymbol{b}} &= \mathbf{a}_{\boldsymbol{b}}^{\boldsymbol{t}} - \mathbf{a}_{\boldsymbol{b}}^{\boldsymbol{t-1}} = \epsilon_{\boldsymbol{r}}^{\boldsymbol{a}} \end{aligned}$$

2) RSSI-based multilateration positional inference: Received Signal Strength Indicator (RSSI) reflects the strength of a received signal and can be used to infer the distance between a signal transmitter and a receiver. Unlike Global Positioning System (GPS), which requires highly synchronized atomic clock to measure the time of arrival of a radio signal, RSSI measurement requires only basic radio receivers. The position of the target can then be estimated through the technique of multilateration (MLAT) once the RSSI measurements from neighboring drones or stations are obtained.

An empirical RSSI channel model is the lognormal shadowing path loss model [13] [14].

$$RSSI = A - 10 \cdot \eta \cdot \log_{10}(d) + \epsilon_{RSSI} \tag{2}$$

where A is the received signal strength at reference distance, η is the path loss exponent, d is the actual distance between receiver and transmitter, and ϵ_{RSSI} is a zero-mean Gaussian random variable following $\epsilon_{RSSI} \sim N(0, \sigma_{RSSI})$ that models the noise in the measurement. In a typical free-space environment, path loss exponent $\eta = 2$ [13] [14].

Based on the RSSI measurement, we have an empirical estimator of the distance. It is important to note that this is a biased estimator because the Gaussian noise ϵ_{RSSI} is scaled up by the power operator.

$$\hat{d} = 10^{\frac{A - RSSI}{10\eta}} \tag{3}$$

Suppose we have a set of n RSSI measurements $\hat{d}_i$ and corresponding known locations of the transmitters $p_i = (x_i, y_i, z_i)$ for $i = 1, \ldots, n$, and we aim to find the location of the target drone $p = (x, y, z)$. using multilateration with the least squares method. We can formalize this problem as:

$$\hat{p} = \arg\min_{p} \sum_{i=1}^{n} (\hat{d}_i - ||p - p_i||)^2 \tag{4}$$

where $\hat{p}$ is the estimated target position, p_i represents the position of the i^{th} neighboring drone, $\hat{d}_i$ represents the RSSI-based distance estimation for the i^{th} neighboring drone. $||\cdot||$ denotes the Euclidean distance between two points. Equation 4 indicates that we aim to find the value of p that minimizes the

sum of squared differences between the measured distances and the actual distances between the target location and the neighboring drones locations. Equation 4 can be solved by various non-linear least squares algorithms, specifically, we use the Levenberg-Marquardt algorithm.

B. Fusion Module

The fusion module is composed of three sub-modules, the Local Location Fusion (ES-EKF) sub-module, the Inter-Drone Inference Sharing sub-module and the Intent-based Weighting sub-module. For clarity, we describe the fusion process in the above order with IMU and RSSI as example sensor inputs.

1) Error State Extended Kalman Filtering (ES-EKF): Unlike RSSI multilateration which transforms the raw RSSI measurement to an estimated position, we directly input the raw RSSI measurement to the ES-EKF algorithm. This allows us to properly model Gaussian noise σ_{RSSI} instead of dealing with the non-linear positional error from RSSI multilateration, which is introduced by Equation 4. We use an Error State Extended Kalman Filter (ES-EKF) to fuse the sensor data from RSSI and IMU. The key idea behind ES-EKF is to estimate the bias term of sensors and the perturbation between nominal and true states of the drone. The estimation of the bias and perturbation would then be feedback to sensor and system models for compensation. The IMU works as the backbone of the ES-EKF sub-module with a high sampling frequency while RSSI receiver provides with complementary sensor input at a much lower sampling frequency. We define the kinematic state and the error state of a drone as following: $\mathbf{x} \triangleq [p, v, q]^T \in \mathbb{R}^{10}$, $\delta\mathbf{x} \triangleq [\delta p, \delta v, \delta\theta, a_b, \omega_b]^T \in \mathbb{R}^{15}$, where $p \in \mathbb{R}^3$ represents position, $v \in \mathbb{R}^3$ represents velocity, $q \in \mathbb{R}^4$ is a quaternion that represents the attitude. Correspondingly for error state $\delta\mathbf{x}$, we define perturbation for position and velocity as $\delta p = p - \hat{p}$, $\delta v = v - \hat{v}$. Additionally, $\delta\theta \in \mathbb{R}^3$ represents the infinitesimal rotation change between nominal and true attitude in Euler angle format, where $q = \delta q \otimes \hat{q} = e^{\frac{\delta\theta}{2}} \otimes \hat{q}$ [10]. We also define $u_m \triangleq [a_m, \omega_m]^T \in \mathbb{R}^6$ as the sensor input from IMU. The error state transition function for δx and the kinematic state transition function for x are formulated as

$$\delta x \leftarrow f(x, \delta x, u_m) = F_x(x, u_m) \cdot \delta x + G_x(x) \cdot w$$
$$x \leftarrow F(x, u_m)$$

where kinematic transition function F is implicitly defined by Equation 1. We then define $w \triangleq [a_n, \omega_n, \delta a_b, \delta\omega_b]^T \in \mathbb{R}^{12}$ to represent the measurement noise a_n, ω_n and biased random walk noise $\delta a_b, \delta\omega_b$, the covariance matrix of w is denoted by Q. We also define $R \triangleq R\{q\} \in \mathbb{S}O(3)$ as the directional cosine matrix associated with quaternion q, which indicates the transformation from the IMU body frame to the reference world frame.

We derive F_x and G_x following [10]:

$$F_x = \begin{bmatrix} I_3 & I_3\Delta t & 0 & 0 & 0 \\ 0 & I_3 & -[R(a_m - a_b)]_\times \Delta t & -R\Delta t & 0 \\ 0 & 0 & I_3 & 0 & -R\Delta t \\ 0 & 0 & 0 & I_3 & 0 \\ 0 & 0 & 0 & 0 & I_3 \end{bmatrix}$$

$$G_x = \begin{bmatrix} 0 & 0 & 0 & 0 \\ I_3\Delta t & 0 & 0 & 0 \\ 0 & I_3\Delta t & 0 & 0 \\ 0 & 0 & I_3 & 0 \\ 0 & 0 & 0 & I_3 \end{bmatrix}$$

Suppose we have $\mathbf{k}$ neighboring drones; the measurement equation is $y = h(x) + v$, where $y \in \mathbb{R}^k$ is the RSSI measurement vector, $\mathbf{h}$ is the RSSI measurement function and $v \in \mathbb{R}^k$, is a vectorized extension of ϵ_{RSSI}, denoting Gaussian noise vector with covariance V.

$$h(x) = A - 10 \cdot \eta \cdot \log_{10}(||p - p_i||)$$

Note that the Jacobian matrix of H is with respect to the error state δx because we are estimating the error state instead of the kinematic state. The Jacobian matrix of H can be derived by chain rule and calculated by H_x and $X_{\delta x}$.

$$H \triangleq \frac{\partial h}{\partial \delta x} = \frac{\partial h}{\partial x}\frac{\partial x}{\partial \delta x} = H_x \cdot X_{\delta x} = \begin{bmatrix} \frac{\partial}{\partial x} RSSI_1 \\ \frac{\partial}{\partial x} RSSI_2 \\ \vdots \\ \frac{\partial}{\partial x} RSSI_k \end{bmatrix} \cdot X_{\delta x} \quad (5)$$

$$\frac{\partial}{\partial x} RSSI_i = [\frac{\partial}{\partial p_x} RSSI_i, \frac{\partial}{\partial p_y} RSSI_i, \frac{\partial}{\partial p_z} RSSI_i, 0, \ldots, 0]^T$$

where $RSSI_i$ represents the RSSI measurement from the i^{th} neighboring drone. $\frac{\partial}{\partial x} RSSI_i$ represents the Jacobian of $RSSI_i$ to kinematic state x, in which only related to position p. The derivation of $X_{\delta x}$ is trivial except for the quaternion $q = [q_x, q_y, q_z, q_w]^T$ [10].

$$X_{\delta x} = \frac{\partial x}{\partial \delta x} = \begin{bmatrix} I_6 & 0 & 0 \\ 0 & Q_{\delta\theta} & 0 \\ 0 & 0 & I_6 \end{bmatrix} \quad Q_{\delta\theta} = \frac{1}{2}\begin{bmatrix} -q_x & -q_y & -q_z \\ q_w & q_z & -q_y \\ -q_z & q_w & q_x \\ q_y & -q_x & q_w \end{bmatrix}$$

Algorithm 1 ES-EKF Algorithm

Input: $x_{initial}$, $\delta x_{initial}$, $P_{initial}$, u_m, V, Q
Output: $\hat{x}$, $\hat{\delta x}$, P

loop
  $\hat{u} = CORRECTION_{bias}(u_m, \hat{\delta x})$
  $\hat{x} \leftarrow F(\hat{x}, u)$
  $P \leftarrow F_x P {F_x}^T + G_x Q {G_x}^T$
  if *RSSI measurement available* **then**
    $H = H_x \cdot X_{\delta x}$
    $K \leftarrow PH^T(HPH^T + V)^{-1}$
    $P \leftarrow (I - KH)P$
    $\hat{\delta x} \leftarrow K(y - h(\hat{x}))$
    $\hat{x} = CORRECTION_{perturbation}(\hat{x}, \hat{\delta x})$
  end
end loop

For multiple sensors, we introduce the usage of multi-staged sensor fusion; each stage of sensor fusion only takes inputs from two individual sensors and produce a positional output.

The output from the previous stage of fusion would then be in turn taken as a pseudo-sensor input for the next stage of sensor fusion, together with another actual sensor input. By using a multi-staged sensor fusion framework, we effectively leverage inputs from multiple sensors.

2) Inter-Drone Inference Sharing: To further improve accuracy of our inferences, our system allows nearby drones to contribute findings to the localization algorithm. We construct a simple mesh network between neighboring drones, allowing them to share their findings within a single hop (in our evaluation section, we will also consider performance benefits from sharing across multiple ALE hops).

In our approach, drones periodically exchange information relevant to estimating their current location. In particular, each drone sends its GPS location (which may or may not be correct, e.g., if the drone is undergoing a spoofing attack), as well as the individual outputs from each of its PIDs (location estimation and confidence value) and relevant metadata (such as time since attack, if known). Each drone receives this information from its neighbors, then incorporates it into its fusion process. The operator may wish to configure how much a drone should trust or rely on its peers based on mission-specific knowledge. To support this, our approach incorporates a weight metric vector that is applied to external ALE PID values before they are incorporated into the fusion process on the local drone. Neighboring UAVs transmit their GPS readings (which may or may not be spoofed) as well as sensor inferences; attacked UAVs then perform localization leveraging other UAVs' findings by incorporating this information into their fusion processes.

3) Intent-based Weighting Fusion: The warfighter may want to input manual configurations for the system. For example, the warfighter may be aware of the mission parameters, sensor resilience and reliability that may influence the extent we could trust different sensors. To support this, our design incorporates configurable weights to provide control over the fusion process. In particular, the warfighter modifies an inverse scaling factor $\alpha \in (\mathbf{0}, \mathbf{1}]$ for each sensor's noise covariance, inducing the sensor fusion process to adapt its weights, and consequently, the confidence levels associated with various sensors' inputs.

$$\sigma_{weighted} = \frac{\sigma_{raw}}{\alpha} \in [\sigma_{raw}, \infty) \tag{6}$$

IV. Evaluation

To evaluate feasibility of our approach, and to investigate its benefits and shortcomings, we undertook an experimental evaluation. In this section, we first describe our emulation- and simulation-based framework. We then describe several key performance results of our design.

A. Experimental Setup

To investigate larger-scale scenarios, we created a scalable and deterministic event-driven simulator along with a collection of models for mobility and wireless. On simulation startup, the drones are randomly deployed on a map with random initial and target states including acceleration and speed. The simulator supports realistic constraints on drones mobility (acceleration, max speed). Each drone is equipped with an IMU and a communication unit, providing sensor data to the system at different frequencies. Additive Gaussian noise is included as measurement noise. To evaluate the performance under varying hardware platforms, we consider IMUs with three levels of accuracy: low (AHRS380SA), mid (IMU381), and high (ADA581016). We also consider pre-filtered readings provided to the localization algorithm at a rate of 100 Hz. To further improve realism, we implemented an emulation framework, which allows the replay of real traces into our design. Here, we leveraged traces collected from real GPS and IMU sensor modules.

B. Metrics

We use root-mean-square error (RMSE) to evaluate the accuracy of different localization algorithms.

$$\text{RMSE} = \sqrt{\frac{1}{n}\sum_{i=1}^{n} ||p_i - \hat{p}_i||^2} \tag{7}$$

where p_i and $\hat{p}_i$ represent the i^{th} actual and predicted position at the same time, and n is the total number of observations. The RMSE estimates the translational deviation between the actual and predicted trajectory. Note that rotational deviation is not evaluated in this context because RSSI multilateration does not provide information about the orientation or rotational aspects of the drones.

C. Single Sensor Efficiency

When the attacked drone needs to determine its own position, one approach would be for it to leverage just a single sensor at a time to perform its location inferences. To help us study how much benefit comes from individual sensor types, we configured our system to leverage one PID at a time. Below, we present results for our approach leveraging just RSSI, and just IMU data. We evaluate the accuracy of single-sensor based localization algorithms.

1) RSSI: Figure 2a shows the error in location estimation for varying radio ranges, using RSSI to correct the drone's position information.

We firstly demonstrate the accuracy of RSSI-based multilateration. As the radio range increases, the attacked drone can communicate with more nearby drones to obtain RSSI values and their positions. This gradually reduces the error in RSSI-based localization algorithms.

We proceed to examine how the swarm density affects the accuracy of multilateration. To ensure consistency, we predefine a set of 20 neighboring drones with known locations. The range of available neighboring drones varies from a minimum of 4, to a maximum of 20. During each iteration, we randomly sample K drones, where $4 <= K <= 20$, from the original set of neighbors. We then execute equation 4 using these sampled drones to estimate the target location. By repeating this process across multiple iterations, we assess the accuracy of localization, the result is shown in Figure 2c.

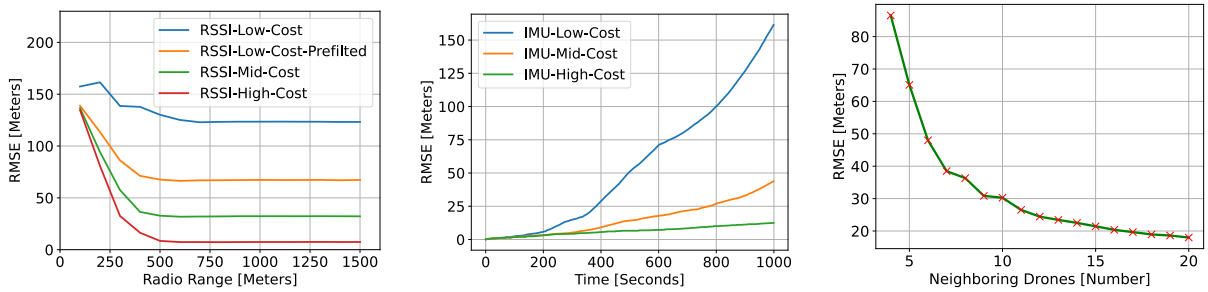

Fig. 2. (a) Position error with RSSI (b) Position error with IMU (c) Effect of swarm density on RMSE.

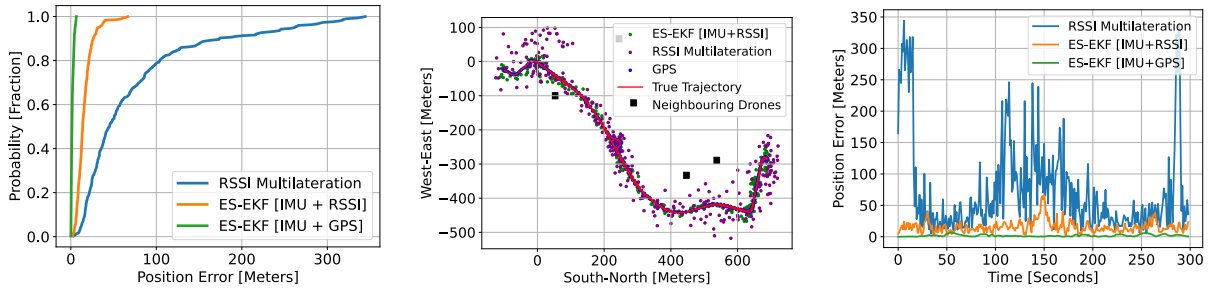

Fig. 3. (a) Distribution of positional error (b) True vs. computed trajectory from trace data (c) Time-series variation in position error.

2) IMU: Next, we explored the ability of IMUs to correct position information. Figure 2b measures RMSE versus time. IMU-based correction provides highly accurate positioning for a short period after attack. However, the mean error as well as variance increases substantially over time. For low-cost IMUs the mean error can reach up to 160 meters at the end of a 10-minute simulation period.

D. Effectiveness of Sensor Fusion

Using individual sensors can provide benefits, but each sensor type may provide varying benefits under different conditions. In this section we investigate whether we can improve performance by fusing locational inferences across sensor types, to overcome their individual limitation.

1) RSSI+IMU Weighted Average: To explore the best ways to combine multiple sensors together, we undertake a study where we average together the location information across them with different weightings (each plane in Figure 4 correspond to different weightings). Since RSSI and IMU are orthogonal methods, this fusion approach can enhance the accuracy of RSSI when the radio range is short and reduces the instability of IMU over time. The time-based fusion algorithm of RSSI and IMU improves results. Figure 5 compares the use of the first-order reciprocal and the 2.5th-order reciprocal of time, both of which effectively suppress the increasing error in IMU measurements over a long time. However, their differences are not significant in a 10-minute simulation.

2) Error-State Extended Kalman Filter Fusion: We then investigate the performance of ES-EKF algorithm over a real-world dataset, comprising GPS data collected at a sampling frequency of 1Hz and 6 DoF IMU sensor data obtained at sampling frequency of 100Hz, gathered on a vehicle in a urban environment with the duration of 300 seconds. To simulate

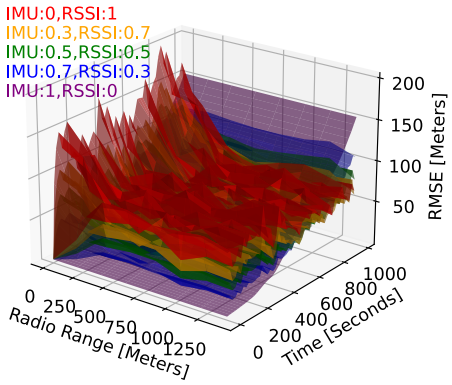

Fig. 4. Parameterized exploration of RSSI+IMU.

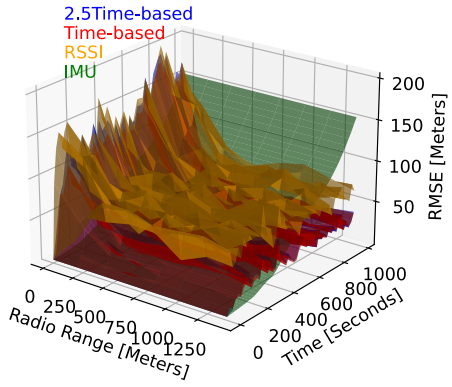

Fig. 5. Parameterized exploration with reciprocal-based combining.

RSSI measurements, we consider four neighboring drones and employ Equation 2 with $\sigma_{RSSI} = 0.5$. The positions and altitudes of four neighboring drones are randomly sampled

from the predefined region to verify algorithm effectiveness.

We compare the positioning accuracy between RSSI multilateration and ES-EKF based sensor fusion of RSSI and IMU. The RMSE from Equation 7 is used as the metric for positioning error. Figure 3a shows that the RSSI multilateration has a much higher positioning error, and a long-tailing effect, the worst 10 percent deviations of which are beyond 174 meters. The ES-EKF fusion of IMU and RSSI greatly alleviates the long-tailing effect and increases the accuracy of positioning, with the worst 10 percent of deviations are only beyond 26 meters. By leveraging sensor data from IMU and RSSI with ES-EKF fusion, we see a nearly 80 percent improvement in positioning accuracy over RSSI multilateration.

Figure 3b shows the disparity of positioning accuracy between RSSI multilateration and ES-EKF fusion between RSSI and IMU. It highlights a significant deviation of the RSSI multilateration results from the true trajectory, while the fusion results closely align with the true trajectory. Numerous outliers can be observed in the RSSI multilateration result.

Figure 3c demonstrates the time-based variation of different algorithms' positioning error, with ES-EKF of IMU and GPS serving as the baseline. We observe similar patterns between ES-EKF of IMU+RSSI and RSSI multilateration, some error peaks occur at the same time periods but at a much lower magnitude in the ES-EKF, implying that sensor fusion essentially attenuate the error from RSSI. Moreover, ES-EKF fusion greatly eliminates the variance in position uncertainty, achieving a two magnitude reduction on the variance of positioning error over RSSI multilateration.

The fundamental difference between RSSI multilateration and ES-EKF with RSSI+IMU is that the ES-EKF does not fuse the positional data produced by RSSI multilateration under Equation 4, instead, RSSI measurements from the k neighboring drones are directly used as a complementary input to the system by Equation 5. Moreover, this difference also imply that we can still perform ES-EKF algorithm when neighboring drones number are fewer than 4, in which case the RSSI multilateration would fail in a 3-dimensional scenario.

Using sensor fusion demonstrates a significant performance improvement over relying solely on RSSI measurements. Note that the empirical estimator in Equation 3 is a biased estimator and hence the RSSI multilateration is in turn biased with a non-linear positional error. Whereas the ES-EKF directly uses the RSSI measurement with no further transformation, which prevents introduction of non-linear error.

3) Inter-ALE Fusion: Finally, we investigate if performance can be improved further by incorporating sensor readings from neighboring drones. We find that performance depends greatly on flight patterns. If drones fly in groups or are otherwise similarly located, locations are likely to be similar, and their inferences can provide a good estimate of the attacked drone's location. However, there are also circumstances where communication can harm result. In this experiment, drones move according to the random waypoint mobility model. Here, we find that correcting based solely on IMU or RSSI outperforms correcting based on neighbors' inferences.

V. Conclusion

GPS signals are crucial to battlefield communication, and hence an attractive target for adversaries to attack. In this paper, we propose a novel framework for the disadvantaged nodes with smaller platforms (e.g., drones) to enhance their location information by leveraging data from their other sensors as well as sensor data from neighboring drones to verify and correct their location information. Principled techniques such as Error State Extended Kalman Filtering is integrated into our framework to improve accuracy of location estimates. Overall, we find that our approach can substantially reduce the adversary's capability to disrupt a drone's positional estimates. For future work, it may be interesting to investigate whether positional error and error variance could be improved by leveraging machine learning techniques to adaptively change weights for different sensor inputs. This work was supported by The Boeing Company under the University Innovation Program Agreement BRT-Z1121-5030.

References

[1] T. Moore, "Global positioning system: Theory and practice," *The Journal of Navigation*, vol. 54, no. 3, p. 481–481, 2001.

[2] S. Swales, J. Maloney, and J. Stevenson, "Locating mobile phones and the us wireless e-911 mandate," 1999.

[3] I. Parra Alonso, D. F. Fernández Llorca, M. Gavilan, S. Álvarez Pardo, M. Garcia-Garrido, L. Vlacic, and M. Sotelo, "Accurate global localization using visual odometry and digital maps on urban environments," *IEEE Transactions on Intelligent Transportation Systems*, vol. 13, no. 4, pp. 1535–1545, 2012.

[4] I. Abaspur Kazerouni, L. Fitzgerald, G. Dooly, and D. Toal, "A survey of state-of-the-art on visual slam," *Expert Systems with Applications*, vol. 205, p. 117734, 2022. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0957417422010156

[5] G. K. L. Tam, Z.-Q. Cheng, Y.-K. Lai, F. C. Langbein, Y. Liu, A. D. Marshall, R. R. Martin, X. Sun, and P. L. Rosin, "Registration of 3d point clouds and meshes: A survey from rigid to nonrigid," *IEEE Transactions on Visualization and Computer Graphics*, vol. 19, pp. 1199–1217, 2013.

[6] C. V. Angelino, V. R. Baraniello, and L. Cicala, "UAV position and attitude estimation using IMU, GNSS and camera," *2012 15th International Conference on Information Fusion*, pp. 735–742, 2012.

[7] Y. Lu, Z. Xue, G.-S. Xia, and L. Zhang, "A survey on vision-based uav navigation," *Geo-spatial Information Science*, vol. 21, no. 1, pp. 21–32, 2018. [Online]. Available: https://doi.org/10.1080/10095020.2017.1420509

[8] C. Zhang, X. Ma, and P. Qin, "Lidar-imu-uwb-based collaborative localization," *World Electric Vehicle Journal*, vol. 13, p. 32, 02 2022.

[9] A. Marquez, B. Tank, S. K. Meghani, S. Ahmed, and K. Tepe, "Accurate UWB and IMU based indoor localization for autonomous robots," in *2017 IEEE 30th Canadian Conference on Electrical and Computer Engineering (CCECE)*, 2017, pp. 1–4.

[10] J. Solà, "Quaternion kinematics for the error-state kalman filter," *CoRR*, vol. abs/1711.02508, 2017. [Online]. Available: http://arxiv.org/abs/1711.02508

[11] A. I. Mourikis and S. I. Roumeliotis, "A multi-state constraint kalman filter for vision-aided inertial navigation," in *Proceedings 2007 IEEE International Conference on Robotics and Automation*, 2007, pp. 3565–3572.

[12] D. Nistér, O. Naroditsky, and J. Bergen, "Visual odometry," in *CVPR*, 2004.

[13] T. S. Rappaport, *Wireless communications. Principles and and practice*, 2nd ed. Upper Saddle River, NJ: Prentice Hall, 2009.

[14] J. Du, J.-F. Diouris, and Y. Wang, "A rssi-based parameter tracking strategy for constrained position localization," *EURASIP journal on advances in signal processing*, vol. 2017, no. 1, pp. 1–10, 2017.

[15] M. Brossard, A. Barrau, and S. Bonnabel, "AI-IMU dead-reckoning," *IEEE Transactions on Intelligent Vehicles*, vol. 5, no. 4, pp. 585–595, 2020.