2024 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)
October 14-18, 2024. Abu Dhabi, UAE

A Cooperative Recovery Framework for Resilient Multi-Robot Swarm Operations Under Loss of Localization in Unknown Environments

Paul J Bonczek and Nicola Bezzo

***Abstract*— Localization is one of the most important tasks for mobile robot operations. Without such capability, a robot may wander toward unsafe states and never complete a desired task. Such capability is even more important in multi-robot system (MRS) operations in which their motion is coordinated based on consensus schemes that leverage information from surrounding neighbors. Thus, in the event of compromised or malfunctioning on-board positioning sensing (e.g., due to cyber attacks or faults) on individual robots, the entire robotic system may be hijacked toward undesired states. In this work, we target this problem by proposing a decentralized framework where: i) robots with loss of localization capabilities detect the anomalous behavior then generate a notification signal within information exchanges to alert neighboring robots, and ii) neighboring robots leverage their mobility to aid in recovery allowing compromised robots to re-localize. Our framework is validated in simulations and lab experiments on proximity-based formations of homogeneous unmanned multi-robot swarms.**

I. Introduction

Localization sensing is one of the most critical information required to achieve intelligent and autonomous capabilities in unmanned systems like autonomous mobile robots. Global Positioning Systems (GPS) are the most common sources for outdoor positioning, but there exists other sensing methods to localize, such as: proximity sensors, radio frequency (RF), external cameras, and LiDAR [1]. However, security-related threats and faults to positioning sensors can compromise system operations, with examples demonstrating catastrophic consequences that include GPS spoofing to divert vessels off course [2] and GPS interference on unmanned aerial vehicles that cause undesirable control behavior leading to crashes [3].

The issue of undesirable on-board positioning sensing is exacerbated within multi-robot systems (MRSs). When left unchecked, compromised robots can negatively impact the entire system's operational performance by hijacking the MRS motion to unsafe regions. Resilient measures have been incorporated to MRSs to ensure continued safe operations in the presence of uncooperative robots within a swarm [4]–[6]. To accomplish this, most methods (e.g., [6]–[8]) make the uncompromised robots typically "ignore" any misbehaving robot to remove undesirable effects that could compromise the MRS mission. Thus, robots experiencing cyber attacks or faults are essentially discarded without a recovery method implemented. In turn, discarded agents can potentially enter undesirable/restricted regions within the environment and the asset will be more likely lost or damaged.

In this work, we propose a recovery framework that leverages MRS mobility to cooperatively recover compromised robots that experience cyber attacks or faults to on-board positioning sensors, as depicted in Fig. 1. Robots that detect compromised sensing send a notification signal within information broadcasts to create a detectable signature to alert neighboring robots. Upon detection of unreliable sensing capabilities, compromised robots perform checkpointing for state reconstruction and compute reachable sets to ensure safety while continuing to navigate in the environment. Detection of the notification signature triggers the neighboring uncompromised robots to perform a cooperative motion behavior to come within sensing/visual range (i.e. mobile landmarking) to aid in re-localizing the compromised robot. Our decentralized framework is designed to be robust within various environments, such as environments that may be absent of known landmarks that could be used for localization. We also assume that robots will nominally operate beyond distance/visual sensing range from each other. In other words, we assume that we cannot count on any landmark (static or mobile) to perform localization. We build upon our previous works where we leveraged residual-based detection techniques to discover anomalous behaving agents subject to stealthy cyber attacks [5] and to infer safety-critical information via hidden signatures within broadcast data exchanges [7]. We summarize our contributions as follows: i) we propose a detection scheme that leverages an expected randomness model of a signal to broadcast notification signatures alerting neighboring robots of unreliable on-board positioning sensors and ii) we present a decentralized control framework by using neighboring mobile robots as landmarks to cooperatively recover (re-localize) compromised robots.

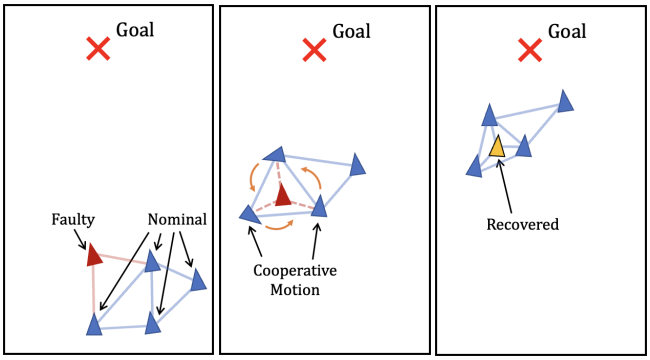

Fig. 1. A pictorial representation of the cooperative recovery problem we wish to solve. Uncompromised robots (in blue) perform (a) detection of the faulty robot, (b) cooperative behavior mode to act a mobile landmarks, and (c) aid in re-localization (recovery) of the compromised robot (yellow).

A. Related Work

The topic of resilient MRS operations has been extensively researched recently in the robotics community [9]. Within

Paul J Bonczek and Nicola Bezzo are with the Charles L. Brown Department of Electrical & Computer Engineering and members of the Link Lab at The University of Virginia, Charlottesville, VA, USA. Email: {pjb4xn,nb6be}@virginia.edu; Paul.Bonczek@jhuapl.edu

the cooperative localization literature, in an early work [10], robots in a swarm were leveraged as landmarks splitting the swarm into two subgroups differing in roles and motion from each other and a decentralized approach is proposed in [11] to alleviate scalability issues. Authors in [12] utilize the joint estimate of robot poses using both centralized and decentralized methods for robots to estimate their pose from the shared information between other robots. These methods however suffer from scalability issues and have strict requirements on swarm size prior to an operation. Other works have improved upon scalability issues by leveraging the Covariance Intersection Algorithm to perform belief updates of neighboring robots in a decentralized manner [13].

We also find works that cover approaches for MRS resiliency from various types of attacks/faults on: sensors, actuators, communications, and physical damage [9]. For example, authors in [14] propose a recovery framework on swarms to utilize LOS measurements when localization sensors fail, with the assumption that neighboring agents are within visual range at all times. In [6], authors propose a resilient flocking framework that leverages a consensus algorithm along with a hybrid control policy that maintains connectivity of the mobile robot team when uncooperative robots share incorrect information. In [15], the Cumulative Sum detector is used to discover spoofs to localization sensors of individual robots within MRSs to isolate and remove malicious agents. Different from these works, we assume that robots operate in unknown environments (i.e., lacking known landmarks) and are positioned beyond range sensing of other robots. When a cyber attack or fault occurs to positioning sensors, an impacted robot loses the ability to localize itself within the environment. In turn, it is necessary for neighboring robots to aid in recovery for re-localization.

II. Preliminaries

Let us consider a multi-robot system of N_r homogeneous robots modeled as a directed graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$. We denote $\mathcal{V} = \{1, \ldots, N_\mathrm{r}\}$ as the robot set and the edge set $\mathcal{E} \subset \mathcal{V} \times \mathcal{V}$, where an edge $(i, j) \in \mathcal{E}$ indicates a connection from robot $i \in \mathcal{V}$ to robot $j \in \mathcal{V}$. We consider each robot i can be represented in a linear time-invariant (LTI) state space form

$$\mathbf{x}_i^{k+1} = \mathbf{A}\mathbf{x}_i^k + \mathbf{B}\mathbf{u}_i^k + \boldsymbol{\nu}_i^k \tag{1}$$

$$\mathbf{y}_i^k = \mathbf{C}\mathbf{x}_i^k + \boldsymbol{\eta}_i^k \tag{2}$$

with state $\mathbf{A}$ and input $\mathbf{B}$ matrices, state vector $\mathbf{x}_i^k \in \mathbb{R}^n$, control input $\mathbf{u}_i^k \in \mathbb{R}^m$, and $\boldsymbol{\nu}_i^k \in \mathbb{R}^n$ denoting zero-mean Gaussian noise. All N_r robots rely on noisy sensors represented by the output vector $\mathbf{y}_i^k \in \mathbb{R}^{N_\mathrm{s}}$ with the output matrix $\mathbf{C}$ and zero-mean Gaussian noise $\boldsymbol{\eta}_i^k \in \mathbb{R}^{N_\mathrm{s}}$ at discrete-time iterations $k \in \mathbb{N}$. However, our framework is agnostic to the dynamical model of the robots. Each robot i utilizes a Kalman Filter to provide a state estimate $\hat{\mathbf{x}}_i^k \in \mathbb{R}^n$ and range sensor(s) with a limited range $\delta_\mathrm{r} > 0$ for collision avoidance. We represent the edge set by $\mathcal{E}_\mathcal{U} = \big\{(i,j) \big| j \in \mathcal{S}_i, \forall i \in \mathcal{V}\big\}$ that is defined within our *control* graph $\mathcal{G}_\mathcal{U} = (\mathcal{V}, \mathcal{E}_\mathcal{U})$ of the robot set $\mathcal{V}$, where $\mathcal{S}_i \subset \mathcal{V}$ is a neighboring robot set that a robot i utilizes for proximity-based control. The communication model assumes that any robots i and j positioned within a maximum communication range $\delta_\mathrm{c} \gg \delta_\mathrm{r}$ from each other can communicate. Information $\mathcal{I}_i = \langle \hat{\mathbf{x}}_i^k, \mathbf{u}_i^k \rangle$ exchanged by any robot $i \in \mathcal{V}$ in broadcasts consists of its state estimate and control input.

In an effort to cooperatively maintain a desired proximity-based formation, the robots within the MRS exchange state information with each other. Each robot $i \in \mathcal{V}$ computes its control input $\mathbf{u}_i^k$ that obeys a control consensus $\mathcal{U}(\cdot, \cdot)$ to achieve the desired inter-robot proximity by:

$$\mathbf{u}_i^k = \mathcal{U}\big(\hat{\mathbf{x}}_i^k, \hat{\mathbf{x}}_j^k\big) \tag{3}$$

where $j \in \mathcal{S}_i$ and $i \neq j$.

A. Threat Model

We assume the multi-robot system is navigating within an adversarial environment $\mathcal{M} \subset \mathbb{R}^2$, where robots may be subject to malicious cyber attacks or faults to on-board positioning sensors. During a cyber attack, we assume that an attacker can persistently modify position measurements with false information in an attempt to intentionally hijack the MRS. Furthermore, position sensors may experience various faults, such as: drift, scaling, and hard faults [16]. Without loss of generality, we characterize both attacks and faults to sensor measurements on-board a robot i as

$$\tilde{\mathbf{y}}_i^k = \mathbf{y}_i^k + \boldsymbol{\xi}_i^k = \mathbf{C}\mathbf{x}_i^k + \boldsymbol{\eta}_i^k + \boldsymbol{\xi}_i^k \tag{4}$$

with $\boldsymbol{\xi}_i^k$ denoting the vector considering an attack or fault to position measurements.

B. Problem Formulation

We consider an MRS tasked to navigate within an environment with the assumption that they maintain communication connectivity, but operate beyond visual/sensing range from each other. Robots that experience cyber attacks or faults to positioning sensors can impede on the ability of all robots within the MRS to successfully execute the mission. The goal here is to find a strategy to cooperatively identify and recover robots that have lost localization capabilities.

Problem 1: (Cooperative Detection and Recovery) Consider an MRS tasked to navigate an unknown environment towards a goal at position $\mathbf{p}_g \in \mathcal{M}$. Create a decentralized policy for any robot $i \in \mathcal{V}$ to detect that a neighboring robot $j \in \mathcal{V}$ has lost localization capability (i.e., $\boldsymbol{\xi}_j^k \neq 0$) and switch to a cooperative recovery control mode $\breve{\mathbf{u}}_i^k$ to re-localize the compromised robot j such that:

$$\mathbb{E}[\mathbf{p}_j^k - \hat{\mathbf{p}}_j^k] = 0. \tag{5}$$

with $\mathbf{p}_j^k$ and $\hat{\mathbf{p}}_j^k$ denoting the true and estimated positions of a compromised robot $j \in \mathcal{V}$.

III. Cooperative Recovery Framework

The diagram in Fig. 2 highlights the overall architecture of our approach to allow for resilient MRS operations. Each robot i monitors its on-board positioning sensors for consistent behavior. Upon detection of anomalous sensor behavior, compromised robots perform state checkpointing and reachability analysis to maintain safety during the absence of localization capabilities. A compromised robot then generates a notification signal to notify nearby robots of its loss of localization capabilities. Detection of the alerting

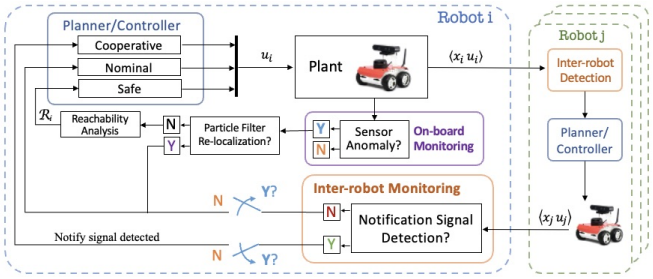

Fig. 2. The overall architecture of our framework followed by each robot.

signature by neighboring robots triggers a cooperative control mode to aid in recovery by coming within sensing range to provide mobile landmarks, allowing a compromised robot to leverage a particle filtering-based method to re-localize itself within the environment. Next, the on-board anomaly detection procedure performed by each robot is presented.

A. On-board Sensor Anomaly Detection

During operations, each robot i computes the measurement residual to monitor for consistent behavior from its on-board position sensor measurements [5]. We define the measurement residual as $\mathbf{r}_i^k = \tilde{\mathbf{y}}_i^k - \mathbf{C}\hat{\mathbf{x}}_i^k$ which is a normally distributed vector $\mathbf{r}_i^k \sim \mathcal{N}(\mathbf{0}, \mathbf{\Sigma}_i)$. $\mathbf{\Sigma}_i$ is the expected measurement residual covariance matrix during nominal conditions. We employ the chi-square detection scheme where the measurement residual vector is reduced to a chi-squared scalar *test measure* defined by $z_i^k = \left(\mathbf{r}_i^k\right)^{\mathsf{T}} \mathbf{\Sigma}_i^{-1}\, \mathbf{r}_i^k$. The chi-square detector on-board a robot i triggers an alarm $\zeta_i = 1$ when $z_i^k > \tau$ is satisfied where $\tau \in \mathbb{R}_+$ is a user-defined threshold, otherwise $\zeta_i^k = 0$. When the estimated alarm rate $\hat{A}_i^k$ no longer follows the expected alarm rate $\mathbb{E}[A]$:

$$\hat{A}_i^k \neq \mathbb{E}[A] \quad \longrightarrow \quad \text{Anomaly Detection} \tag{6}$$

a robot i concludes that a cyber attack or fault has occurred to its positioning sensor then subsequently places itself in the compromised robot set $i \in \mathcal{V}_i^{\mathrm{C}}$, where $\mathcal{V}_i^{\mathrm{C}} \subset \mathcal{V}$ [5].

B. Checkpointing and Reachability for Safe Motion

With the positioning sensor deemed compromised, a robot can no longer rely on its position estimate while navigating. A compromised robot performs state reconstruction via checkpointing and reachability analysis along with a virtual physics-based scheme to maintain the system safety to avoid reaching undesired areas $\mathcal{A} \subset \mathcal{M}$ in the environment $\mathcal{M}$.

1) Checkpointing for State Reconstruction: Once a robot i detects a failure to its localization sensor, the robot leverages saved historical information of the robot's states and control inputs over a window of length $L \in \mathbb{N}$. To recover the system state, the robot "goes back" to a time when the state was considered safe to reconstruct an estimate of the failed state elements, by using a checkpointing and recovery procedure [17]. The robot then *rolls forward* the checkpointed state by utilizing the stored control inputs from the checkpointing time $k-L$ up until the present time k. From this roll-forward procedure [17], a reconstructed state $\hat{\hat{\mathbf{x}}}_i^k \in \mathbb{R}^n$ is computed by $\hat{\hat{\mathbf{x}}}_i^k = \mathbf{A}^L \hat{\mathbf{x}}_i^{(k-L)} + \sum_{l=1}^{L} \mathbf{A}^{l-1} \mathbf{B} \mathbf{u}_i^{(k-l)}$ that is then used for control purposes. We assume the window length L of historical data is chosen to contain enough information to detect position sensor anomalies [17].

2) Reachability Analysis: With the loss of localization capabilities, we want to ensure that compromised robots do not enter an undesired area $\mathcal{A} \subset \mathcal{M}$ within the environment. Compromised robots are tasked to compute a reachable set $\mathcal{R}_i \subset \mathcal{M}$ that is based on the dynamical and noise models to provide all possible positions that the robot's true state may be located given the localization uncertainties to help avoid navigating into any undesired areas. We utilize Monte Carlo simulations to generate reachable sets [18] where ellipsoid bounds $\varepsilon_i^{k^*}$ of the reachable set $\mathcal{R}_i$ are computed. This method encompasses all possible locations of an ith compromised robot, where k^* is the time since beginning the reachable set computation.

3) Safe Motion: For a compromised robot i to ensure safety while in motion, the goal is to maintain a safe distance away from any known undesired restricted area $\mathcal{A} \subset \mathcal{M}$. To do this, the robot includes a virtual spring from the minimum distance between the reachable set bounds $\varepsilon_i^{k^*}$ and the undesired region to provide a repulsive force modeled as:

$$\mathbf{u}_{i,\mathcal{A}}^k = \kappa_{\mathcal{A}}(l_{i\mathcal{A}} - l_o^0)\vec{\mathbf{d}}_{i\mathcal{A}}, \quad \text{if } l_{i\mathcal{A}} < l_o^0, \tag{7}$$

where $l_{i\mathcal{A}} = \min \left\| \varepsilon_i^{k^*} - \mathbf{p}_{\mathcal{A}} \right\|$ and $\mathbf{p}_{\mathcal{A}} \in \mathbb{R}^2$ are positions within the undesired region $\mathcal{A}$. The spring constant is represented by $\kappa_{\mathcal{A}}$, while $l_o^0 > 0$ is a safety distance, and $\vec{\mathbf{d}}_{i\mathcal{A}}$ indicates the unit vector of the repulsive force direction. The *safe* control input of compromised robot i follows:

$$\bar{\mathbf{u}}_i^k = \mathbf{u}_{i,\mathcal{A}}^k + \kappa_{ig} l_{ig} \vec{\mathbf{d}}_{ig} - \gamma_i \hat{\tilde{\mathbf{v}}}_i^k \tag{8}$$

with $l_{ig} = \|\mathbf{p}_g - \hat{\tilde{\mathbf{p}}}_i^k\|$ denoting the distance between the goal point and the reconstructed position estimate, κ_{ig} is the spring constant, and $\vec{\mathbf{d}}_{ig}$ is the unit vector for the force direction. $\gamma_i > 0$ and $\hat{\tilde{\mathbf{v}}}_i^k$ are damping coefficient and the reconstructed velocity estimate, respectively. This allows the compromised robot to safely continue navigating toward the goal point while a reliable position estimate is unavailable.

C. Notification Signature Generation and Detection

Next, we formalize the notification signal that compromised robots emit to neighboring robots to trigger cooperative recovery. We begin by modeling an inter-robot residual that robots leverage to monitor for consistent behavior of neighboring agents [7]. Each robot $i \in \mathcal{V}$ exchanges its state estimate $\hat{\mathbf{x}}_j^k$ and control input $\mathbf{u}_j^k$ to enable neighboring robots to predict its motion. A robot i is then able to make state predictions $\hat{\mathbf{x}}_{ij}^{k+1} \in \mathbb{R}^n$ of any nearby robot $j \in \mathcal{V}$ by computing $\hat{\mathbf{x}}_{ij}^{k+1} = \mathbf{A}\hat{\mathbf{x}}_j^k + \mathbf{B}\mathbf{u}_j^k$. A robot i then computes an *inter-robot residual*

$$\mathbf{r}_{ij}^k = \hat{\mathbf{x}}_j^k - \hat{\mathbf{x}}_{ij}^k \tag{9}$$

to monitor for nominal behavior of a robot j. The expectation of the inter-robot residual vector follows $\mathbf{r}_{ij}^k = \mathcal{N}(\mathbf{0}, \mathbf{\Sigma}_{ij}) \in \mathbb{R}^n$, with each element $q = \{1, \ldots, n\}$ is characterized by

$$\mathbb{E}[\mathrm{r}_{ij,q}] = 0, \quad \mathrm{Var}[\mathrm{r}_{ij,q}] = \sigma_{ij,q}^2 = \sum_{s=1}^{N_{\mathrm{s}}} \left(\mathrm{K}_{(q,s)} \sigma_{j,s} \right)^2 \tag{10}$$

where $\mathrm{K}_{(q,s)}$ is the element in the qth row and sth column of the Kalman gain $\mathbf{K}$ [7]. Next, we formalize a procedure to generate a notification signal to alert nearby robots.

1) Notification Signal Generation: Once a compromised robot $i \in \mathcal{V}_i^{\mathrm{C}}$ detects anomalous sensor behavior (6), it desires to produce a detectable signature within the broadcast information, all while the inter-robot residual (9) continues to follow the expected distribution. The compromised robot leverages the reconstructed state $\hat{\mathbf{x}}_i^k$ and safe control input (8) to compute the next reconstructed state estimate to time $k+1$ by integrating forward system dynamics, as follows:

$$\hat{\mathbf{x}}_i^{k+1} = \mathbf{A}\hat{\mathbf{x}}_i^k + \mathbf{B}\bar{\mathbf{u}}_i^k. \tag{11}$$

Additionally, the compromised robot i overlays a notification signal $\mathbf{w}_i^k \in \mathbb{R}^n$ to the reconstructed state by

$$\hat{\mathbf{x}}_{i,\mathcal{N}}^k = \hat{\mathbf{x}}_i^k + \mathbf{w}_i^k \tag{12}$$

which is an additive Gaussian distributed vector that emulates the stochastic state estimate behavior from the viewpoint of neighboring robots. The updated state (12) containing the overlaid signal is then broadcast to the nearby robots. In the following Lemma, we show how a compromised robot $j \in \mathcal{V}_j^{\mathrm{C}}$ constructs the notification signal $\mathbf{w}_j^k$ such that the inter-robot residual (9) maintains the expected distribution as in nominal conditions from the perspective of a robot i.

Lemma 1: A compromised robot $j \in \mathcal{V}_j^{\mathrm{C}}$ covertly disguises the notification signal $\mathbf{w}_j^k$ such that the inter-robot residual distribution from the perspective of neighboring robots emulates nominal behavior if each element $q = \{1, \ldots, n\}$ of the notification signal vector follows $\mathbb{E}[\mathrm{w}_{j,q}] = 0$ and $\mathrm{Var}[\mathrm{w}_{j,q}] = \frac{\sum_{s=1}^{N_{\mathrm{s}}} (\mathrm{K}_{(q,s)}\sigma_{j,s})^2}{2}$.

Proof: We first consider the forward dynamics in (11) and inter-robot prediction $\hat{\mathbf{x}}_{ij}^{k+1}$ where both leverage the same state space dynamics and control input but compute one step ahead projections with differing states. The compromised robot j projects forward with the deterministic dynamics the state $\hat{\mathbf{x}}_i^k$ while the neighboring robot i uses the state $\hat{\mathbf{x}}_{i,\mathcal{N}}^k$ from (12) which contains the notification signal. A neighboring robot i monitoring inter-robot residuals of robot j, is in effect monitoring the difference between two consecutive Gaussian notification signals overlaid in the state information at times k and $k-1$, denoted by the difference vector $\mathbf{d}_{ij}^k = \mathbf{w}_j^k - \mathbf{w}_j^{k-1}$. Given this, the expected distribution of each qth difference vector element follows:

$$\begin{aligned} \mathbb{E}[\mathrm{d}_{ij,q}^k] &= \mathbb{E}[\mathrm{w}_{j,q}^k] - \mathbb{E}[\mathrm{w}_{j,q}^{k-1}] = 0, \\ \mathrm{Var}[\mathrm{d}_{ij,q}^k] &= \mathrm{Var}[\mathrm{w}_{j,q}^k] + \mathrm{Var}[\mathrm{w}_{j,q}^{k-1}] = 2\mathrm{Var}[\mathrm{w}_{j,q}]. \end{aligned} \tag{13}$$

Setting the difference vector $\mathbf{d}_{ij}^k \sim \mathcal{N}(\mathbf{0}, 2\mathrm{Var}[\mathbf{w}_j])$ equal to the expectation of the inter-robot residual (9), we establish that the notification signal must contain properties:

$$\mathbb{E}[\mathbf{w}_j] = \mathbb{E}[\mathbf{r}_{ij}], \quad \mathrm{Var}[\mathbf{w}_j] = \frac{\boldsymbol{\Sigma}_{ij}}{2}, \tag{14}$$

thus each element of the notification signal $\mathrm{w}_{j,q}$ follows $\mathbb{E}[\mathrm{w}_{j,q}] = 0$ and $\mathrm{Var}[\mathrm{w}_{j,q}] = \frac{\sum_{s=1}^{N_{\mathrm{s}}} (\mathrm{K}_{(q,s)}\sigma_{j,s})^2}{2}$ thereby, concluding the proof. ∎

By employing the deterministic forward dynamics (11) with the notification signal (12), the observation of the inter-robot residual from the perspective of a neighboring robot remains unchanged (i.e., preserved Gaussian distributed properties). However, the inter-robot residual contains a differing sign randomness signature from the nominal operating conditions. Next, we describe the conditions for nearby robots to implement in order to observe a change in inter-robot residual sign switching rate (i.e., randomness properties).

2) Notification Signal Detection: To detect a randomness-based notification signal (12), a robot i monitors an inter-robot residual sign switching rate $\boldsymbol{\psi}_{ij} = \left[\psi_{ij,1}, \ldots, \psi_{ij,n}\right]^{\mathsf{T}}$ of nearby robots $j \in \mathcal{V}$. The following procedure triggers an alarm (i.e., $\zeta_{ij,q}^k = 1$) when a sign switch occurs on any qth inter-robot residual element at a time k by:

$$\zeta_{ij,q}^k = \begin{cases} 1, & \text{if } \mathrm{sgn}\big(\mathrm{r}_{ij,q}^k\big) = -\mathrm{sgn}\big(\mathrm{r}_{ij,q}^{k-1}\big) \\ 0, & \text{otherwise.} \end{cases} \tag{15}$$

The sign switching alarm $\zeta_{ij,q}^k \in \{0, 1\}$ is then sent to a runtime alarm rate estimation algorithm [7] to provide an updated estimate of the sign switching rate $\hat{\psi}_{ij,q}^k \in [0, 1]$.

Lemma 2: Given a inter-robot residual (9) during attack-free conditions, the expected value and variance of the sign switching rate to signify nominal random behavior follows $\mathbb{E}[\psi] = \frac{1}{2}$ and $\mathrm{Var}[\psi] = \frac{1}{4(2\ell-1)}$, respectively.

Proof: See [7] for a similarly designed proof. ∎

The overlaid notification signal $\mathbf{w}_j^k$ by a robot j transforms the stochastic inter-robot residual variable to contain serial randomness characteristics [19] while observed by a robot i. The following Lemma provides the required observed sign switching behavior by a robot i to determine that a neighboring robot j is emitting the notification signal.

Lemma 3: A robot $i \in \mathcal{V}$ detects the notification signal $\mathbf{w}_j^k$ from a compromised robot j when the observed inter-robot residual (9) sign switching rate for all $q = 1, \ldots, n$ elements satisfy $\hat{\psi}_{ij,q}^k \in [\Psi'_-, \Psi'_+]$.

Proof: See [7] for a similarly designed proof. ∎

After a robot i identifies that a robot j is compromised, robot j is included into the compromised robot set $j \to \mathcal{V}_i^{\mathrm{C}}$.

Corollary 1: Given a user-defined window length $\ell_\psi > 0$ for a robot i to estimate inter-robot sign switching rate elements $\hat{\psi}_{ij,q}$, $q = 1, \ldots, n$ of a robot j, the probability of false detection (FD) is described as $\Pr(FD) = \beta$.

Proof: Under the assumption that each qth inter-robot residual element is independent, the probability β that a qth sign switching alarm rate $\hat{\psi}_{ij,q}$ travels above the lower bound for notification signal detection (i.e., $\hat{\psi}_{ij,q} > \Psi'_-$) follows:

$$\begin{aligned} \beta &\approx 1 - \int_{-\infty}^{\Psi'_-} f\big(X \mid \mathbb{E}[\psi], \mathrm{Var}[\psi]\big), \\ &\approx 1 - \int_{-\infty}^{\Psi'_-} \frac{1}{\sqrt{2\pi\mathrm{Var}[\psi]}} \exp\left\{-\frac{1}{2}\left(\frac{X - \mathbb{E}[\psi]}{\sqrt{\mathrm{Var}[\psi]}}\right)^2\right\} \end{aligned} \tag{16}$$

where $f\big(X \mid \mathbb{E}[\psi], \mathrm{Var}[\psi]\big)$ denotes the probability density function of the sign switching alarm rate under nominal conditions. Then, the probability for at least $n_s \in \{1, \ldots, n\}$ sign switching rate elements to travel above the lower detection bound (during nominal conditions) is found by:

$$\Pr(FD) = \sum_{n_s=1}^{n} \binom{n}{n_s} \beta^{n_s}(1-\beta)^{n-n_s} \tag{17}$$

such that when $n_s = n$ (i.e., a false detection) results in $\Pr(FD) = \beta^n$, thus concluding the proof. ∎

D. Cooperative Recovery

In this subsection, we describe the decentralized cooperative behavior of a robot i that detects a notification signature from a compromised robot $j \in \mathcal{V}_i^{\mathrm{C}}$ that also belongs to its control neighbor set $j \in \mathcal{S}_i$. A robot i switches to a cooperative control mode to move within sensing range of the compromised agent to provide itself as a mobile landmark for recovery.

Neighboring robots employ cooperative motion to aid in recovery of compromised robots via a commonly-used "spiral in" pattern (i.e., encircling, rotating, and converging motion) used in search and recovery [20]. First, each cooperative robot i estimates who the other cooperative neighbors h of the compromised robot $j \in \mathcal{V}_i^{\mathrm{C}}$ are for cooperative recovery by leveraging Gabriel Graph (GG) rule [21], a connected graph with no crossing control edges [5]. A robot i estimates the cooperative set $\hat{\mathcal{C}}_{j,i}$ of robot j by:

$$\hat{\mathcal{C}}_{j,i} = \left\{ h' \in \mathcal{V} \;\middle|\; \widehat{jhh'} \leq \pi/2,\ h, h' \in \mathcal{V} \setminus \mathcal{V}_i^{\mathrm{C}} \right\} \tag{18}$$

where $\widehat{jhh'}$, $j \neq h \neq h'$ is the interior angle of the three robot position configuration and positions of robots $h, h' \in \mathcal{V} \setminus \mathcal{V}_i^{\mathrm{C}}$ are received from information broadcasts. Each cooperative robot i utilizes the estimated set $\hat{\mathcal{C}}_{j,i}$ to maneuver into locations to surround the compromised robot j in equal angular intervals. The desired angle between the other cooperative robots follows $\hat{\theta}_i^{\mathrm{des}} = \frac{2\pi}{|\hat{\mathcal{C}}_{j,i}|}$. To maintain the desired intervals between its cooperative neighbors while encircling around the compromised robot j, each robot i includes a control input to generate the tangential force:

$$\mathbf{u}_{i,\mathrm{E}}^k = \left[\kappa_{\mathrm{E}}(\hat{\theta}_i^{\mathrm{des}} - \widehat{ijh_i^{\mathrm{L}}}) - \kappa_{\mathrm{E}}(\hat{\theta}_i^{\mathrm{des}} - \widehat{ijh_i^{\mathrm{R}}}) \right] \vec{\mathbf{d}}_{ij}(\perp) \tag{19}$$

where κ_{E} is a user-defined control gain, while h_i^{L} and h_i^{R} are the nearest cooperative neighbors to the left and right from the perspective of robot i, respectively. Additionally, $\widehat{ijh_i^{\mathrm{L}}}$ and $\widehat{ijh_i^{\mathrm{R}}}$ denote inner angles between the nearest left and right cooperative neighbors and $\vec{\mathbf{d}}_{ij}(\perp)$ denotes the direction of the tangential force from the vector of robots i to j.

Simultaneously, each cooperative robot i rotates around and converges toward the received position coordinate of the compromised robot j that are computed by:

$$\begin{aligned} \mathbf{u}_{i,\mathrm{R}}^k &= \left[\kappa_{\mathrm{R}}(l_{ij} - l_{v'}^0) \right] \vec{\mathbf{d}}_{ij}(\perp), \\ \mathbf{u}_{i,\mathrm{C}}^k &= \left[\kappa_{\mathrm{C}}(l_{ij} - l_{v'}^0) \right] \vec{\mathbf{d}}_{ij}, \end{aligned} \tag{20}$$

to generate the rotational $\mathbf{u}_{i,\mathrm{R}}^k$ and converging $\mathbf{u}_{i,\mathrm{C}}^k$ motion. Parameters that determine control behavior are the gains κ_{R} and κ_{C} for rotation and convergence. The desired distance $l_{v'}^0$ is reduced to $0 < l_{v'}^0 < \delta_{\mathrm{r}}$ to enable cooperative robots to come within sensing range δ_{r} of the compromised robot j. The position of robot j used for rotation and converging (i.e., utilized in $l_{ij} = \|\hat{\mathbf{p}}_i^k - \breve{\mathbf{p}}_j^k\|$) is determined by:

$$\breve{\mathbf{p}}_j^k = \begin{cases} \hat{\mathbf{p}}_{j,i}^k, & \text{if } \|\mathbf{p}_i^k - \mathbf{p}_j^k\| \leq \delta_{\mathrm{r}}, \\ \hat{\mathbf{p}}_{j,\mathcal{N}}^k, & \text{otherwise}, \end{cases} \tag{21}$$

where $\hat{\mathbf{p}}_{j,\mathcal{N}}^k$ is the received position (containing the notification signal) from compromised robot $j \in \mathcal{V}_i^{\mathrm{C}}$ and $\hat{\mathbf{p}}_{j,i}^k$ is the observed estimated position of robot j when within sensing range of robot i. The control input for each robot i during cooperative recovery follows

$$\breve{\mathbf{u}}_i^k = \mathbf{u}_{i,\mathrm{E}}^k + \mathbf{u}_{i,\mathrm{R}}^k + \mathbf{u}_{i,\mathrm{C}}^k + \mathbf{u}_{i,\mathrm{A}}^k + \kappa_{ig} l_{ig} \vec{\mathbf{d}}_{ig} - \gamma_i \mathbf{v}_i^k \tag{22}$$

enabling cooperative behavior to find the compromised robot j while still navigating towards the goal. The control input $\mathbf{u}_{i,\mathrm{A}}^k \in \mathbb{R}^m$ in (22) serves as a repulsive force for avoidance during cooperative recovery when separation distance $\mathrm{s}_{ih}^k = \|\mathbf{p}_i^k - \mathbf{p}_h^k\|$ between the robot i and any robot $h \in \mathcal{V} \setminus i$ is less than a minimum separation threshold $\delta_{\mathrm{s}} \in \mathbb{R}_+$ (i.e., when $\mathrm{s}_{ih}^k < \delta_{\mathrm{s}}$). Once a robot i and all cooperative neighbors have come within sensing range of the compromised robot, the robots return to the nominal formation controller (3) while maintaining reduced distances $l_{v'}^0$ between the cooperative and compromised robots. Pictured in Fig. 3 is an example of cooperative recovery motion from nearby robots (blue dots) to aid in re-localization of a compromised robot (red dot).

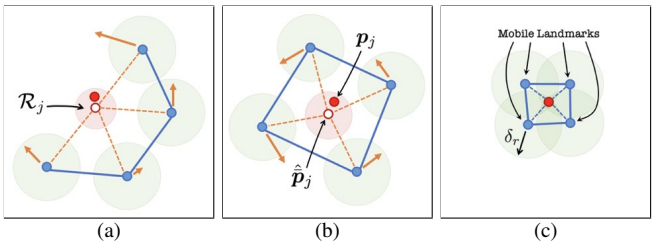

Fig. 3. A depiction of desired cooperative behavior of robots (blue disks) that aid in re-localization of a compromised robot (red). Orange arrows depict force vectors during cooperative recovery behavior and the larger red disk represents the reachable set $\mathcal{R}_j$ of robot j.

Remark 1 (Lost Robot): In a scenario where the compromised robot j is not found (i.e. unrecoverable), each cooperating robot i places robot j into its set of lost robots $j \in \mathcal{V}_i^{\mathrm{L}}$ to longer use it for control purposes. The cooperative robots then return to the nominal control mode (3).

E. Mobile Landmarks for Localization

Once one or more cooperative neighboring robots come within sensing range, a compromised robot $i \in \mathcal{V}_i^{\mathrm{C}}$ begins to re-localize itself within the environment. In this work, compromised robots use a particle filtering-based method for re-localization by leveraging the mobile landmarks (i.e., in-sensing range robots), which are obtained by measurements from on-board range sensors that are assumed to still be available and also utilizing the received position coordinates from nearby robots. Given the known reachable set $\mathcal{R}_i \subset \mathcal{M}$ and the sensing range δ_{r}, a robot i assumes that any robot j with position coordinates that satisfies:

$$\mathcal{P}_i = \begin{cases} j \in \mathcal{P}_i, & \text{if } \hat{\mathbf{p}}_j^k \in \mathcal{R}_i \oplus \delta_{\mathrm{r}}, \\ j \notin \mathcal{P}_i, & \text{otherwise}, \end{cases} \tag{23}$$

can potentially be the observed robot(s) within range sensing, denoted by the robot set $\mathcal{P}_i \subset \mathcal{V}$ where we represent $\mathcal{R}_i \oplus \delta_{\mathrm{r}}$ as the summation of areas from the computed reachable set $\mathcal{R}_i$ and the disk with sensing radius $\delta_{\mathrm{r}} > 0$. The output of our particle filter which utilizes the mobile landmarks provided by nearby robots $j \in \mathcal{P}_i$ (i.e., resulting position coordinates) acts as a replacement for position measurements, in place of the compromised on-board position sensor. This measurement is then used for state estimation and control.

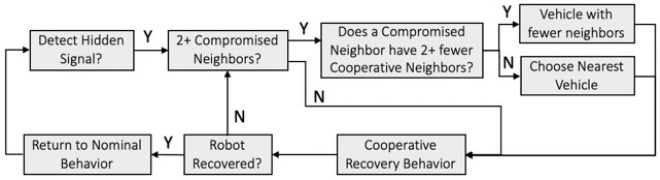

Fig. 4. Cooperative behavior logic of robots when more than one neighboring robot is emitting a notification signal.

F. Multiple Compromised Robots

Up to this point, we have characterized the cooperative recovery framework for robots when only a single control neighbor robot j is emitting a notification signal. However, a robot that encounters a scenario when two or more neighbors are emitting a notification alert signal must make a decision of which compromised neighbor to choose to aid in re-localization. We present a flow chart in Fig. 4 of the decision process made by any robot when this condition is present.

To summarize, a robot first infers the number of cooperative neighbors that each compromised robot has by using (18). If the number of cooperative neighbors for one compromised robot is at least 2 fewer than any other compromised robot, then the robot chooses to aid the robot with the minimum number of encircling neighbors. If the difference in cooperative neighbors does not satisfy the previous conditions, then the robot will choose the nearest robot (i.e., in terms of euclidean distance). At every time instant k, the robot infers the number of cooperative neighbors per compromised robot, allowing the robot to switch to another compromised robot if network topology conditions change.

IV. Results

Our MRS approach is validated with MATLAB simulations and lab experiments using a swarm of Husarion ROSbot 2.0 robots. The MRSs are tasked to perform go-to-goal operations in unknown environments that lack known landmarks where agents are susceptible to cyber attacks and faults to vulnerable on-board position sensors. In all case studies, the agents in the MRS resiliently maintain a desired proximity-based formation with the utilization of a virtual spring-damper physics control model [5].

A. Simulations

We consider $N_{\mathrm{r}} = 10$ agents treated as double integrator point masses that are navigating in formation to a goal point within an environment $\mathcal{M}$ where desired distances between agents is 10m (i.e., virtual spring rest lengths), on-board sensing range is limited to $\delta_r = 3$m, and a sampling time $t_s = 0.05$s is used. A sequence of snapshots presented in Fig. 5 highlight our cooperative recovery framework to aid in re-localization of two agents $\{1, 7\}$ that experience malicious cyber attacks that falsify their position sensor measurements. Once the compromised agents (red dots) detect the cyber attacks, they emit the notification signature in $\hat{\hat{\boldsymbol{x}}}_{i,\mathcal{N}}$ to notify neighboring robots (blue dots), allowing them to perform cooperative recovery behavior as shown in Fig. 5(b). In Fig. 5(c), the cooperative agents come within sensing range of the compromised agents to aid in recovery by acting as a mobile landmark (recovered agents are depicted by yellow dots). Fig. 5(d) shows that all compromised agents have been recovered and the multi-robot system continues to navigate toward the desired goal point. From the perspective of robot 10, we show in Fig. 6 the observed inter-robot residual (9) sign switching rates of all neighboring agents. The compromised agents $\{1, 7\}$ have an observed increase in sign switching rates (shown only for position in x-direction), allowing agents to switch to the cooperative recovery control mode to aid in robot re-localization. Shown in Table I is the success rate of our cooperative recovery framework during varying MRS sizes and number of compromised robots. For each of the six scenarios, 100 simulations were ran to result in an observed success rate, which is defined as the percent of robots that were successfully re-localized. We observe as the number of compromised robots increases, the overall rate of success decreases since there are fewer uncompromised robots available to aid in recovery.

TABLE I
Cooperative Recovery Success Rate

Total Agents	10			25		
Compromised	1	3	5	3	5	10
Success (%)	100	91	56	100	96	64

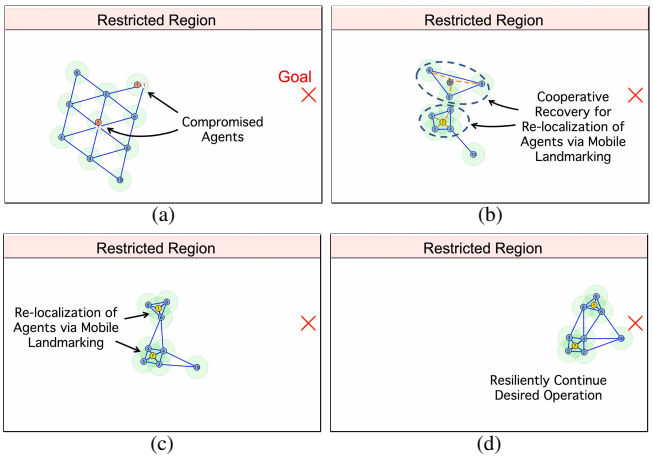

Fig. 5. A formation of $N_{\mathrm{r}} = 10$ agents resiliently navigating to a desired goal point (red 'X'). Robots (blue disks) perform cooperative recovery to aid in re-localization of the two compromised robots (red disks). Recovered robots are represented as yellow disks.

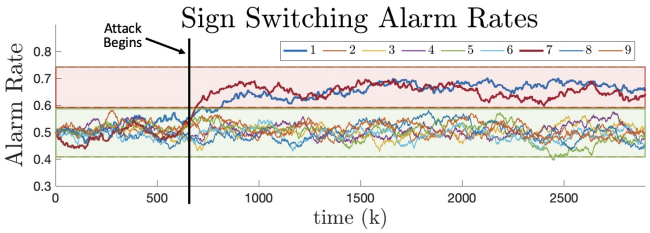

Fig. 6. Observed inter-robot residual sign switching rate (position in the x-direction) from the perspective of robot 10 throughout the simulation.

B. Experiments

In the experiment case study, we consider $N_{\mathrm{r}} = 6$ Husarion ROSbot 2.0 robots that navigate in a lab environment while resiliently performing go-to-goal operations. Depicted in Fig. 7, robot 5 is subject to attacks to its position sensor (Fig. 7(b)) in an attempt to hijack the robot to an undesirable (i.e., restricted) region. Upon detection of the

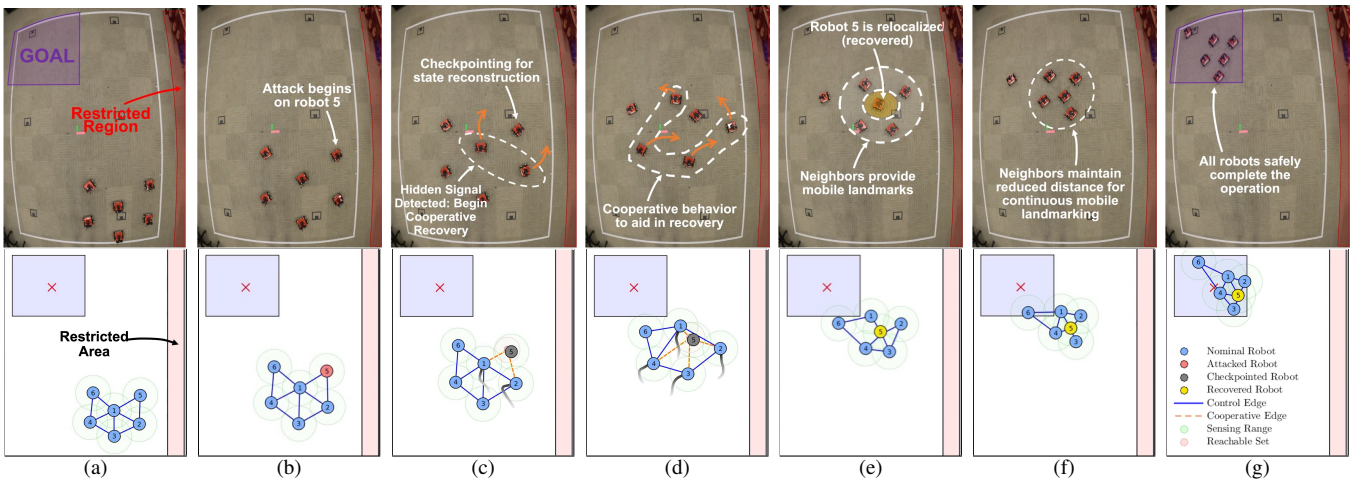

Fig. 7. An experiment showing $N_{\mathrm{r}} = 6$ robots navigating to a goal (purple region) with robot $i = 5$ experiencing a cyber attack to its position sensor. The neighboring robots detect the notification signal from robot 5, then cooperatively aid in re-localization via mobile landmarking.

cyber attack, robot 5 generates a notification signal in $\hat{\tilde{\boldsymbol{x}}}_{i,\mathcal{H}}$ to alert its neighbors. The snapshots in Fig. 7(c)-(d) show the cooperative motion of neighboring robots $\{1, 2, 3, 4\}$ to aid in recovery (re-localization) of robot 5 by acting as mobile landmarks. Once robot 5 is recovered (Fig. 7(e)), the entire swarm of robots is able to safely navigate to the desired goal region, as shown in Fig. 7(f)-(g).

V. Conclusions

In this work, we have proposed a decentralized multi-robot system framework to resiliently perform operations by utilizing cooperative motion to recover compromised robots that are subject to cyber attacks or faults to on-board positioning sensors. Compromised robots generate a notification signature to alert neighboring agents of the impending failure, allowing them to switch to a cooperative recovery control mode to aid in robot re-localization via mobile landmarking. In our future work, we plan to provide theoretical resiliency guarantees on the maximum number of robots that may be compromised in an MRS of N_{r} robots with our framework. We also plan to explore cooperative recovery within heterogeneous MRSs where we can exploit differing robot dynamics to optimally aid in recovery.

Acknowledgements

This work is based on research supported by National Science Foundation under grant numbers 1816591 and 1916760.

References

[1] G. A. Demetriou, "A survey of sensors for localization of unmanned ground vehicles (ugvs)." in *IC-AI*. Citeseer, 2006, pp. 659–668.

[2] J. Bhatti and T. E. Humphreys, "Hostile control of ships via false gps signals: Demonstration and detection," *Navigation*, vol. 64, no. 1, pp. 51–66, 2017.

[3] D. Hambling. (2020) Drone crash due to gps interference in u.k. raises safety questions. [Online]. Available: https://www.forbes.com

[4] V. Renganathan, K. Fathian, S. Safaoui, and T. Summers, "Spoof resilient coordination in distributed and robust robotic networks," *IEEE Transactions on Control Systems Technology*, vol. 30, no. 2, pp. 803–810, 2022.

[5] P. J. Bonczek, R. Peddi, S. Gao, and N. Bezzo, "Detection of nonrandom sign-based behavior for resilient coordination of robotic swarms," *IEEE Transactions on Robotics*, vol. 38, no. 1, pp. 92–109, 2022.

[6] K. Saulnier, D. Saldaña, A. Prorok, G. J. Pappas, *et al.*, "Resilient flocking for mobile robot teams," *IEEE Robotics and Automation Letters*, vol. 2, no. 2, pp. 1039–1046, 2017.

[7] P. J. Bonczek and N. Bezzo, "Detection and inference of randomness-based behavior for resilient multi-vehicle coordinated operations," in *2021 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, 2021, pp. 5844–5850.

[8] V. Renganathan and T. Summers, "Spoof resilient coordination for distributed multi-robot systems," in *2017 International Symposium on Multi-Robot and Multi-Agent Systems (MRS)*, 2017, pp. 135–141.

[9] D. Zhang, G. Feng, Y. Shi, and D. Srinivasan, "Physical safety and cyber security analysis of multi-agent systems: A survey of recent advances," *IEEE/CAA Journal of Automatica Sinica*, vol. 8, no. 2, pp. 319–333, 2021.

[10] R. Kurazume, S. Nagata, and S. Hirose, "Cooperative positioning with multiple robots," in *Proceedings of the 1994 IEEE International Conference on Robotics and Automation*, 1994, pp. 1250–1257 vol.2.

[11] A. G. Pires, D. G. Macharet, and L. Chaimowicz, "Towards cooperative localization in robotic swarms," in *Distributed Autonomous Robotic Systems*. Tokyo: Springer Japan, 2016, pp. 105–118.

[12] S. Roumeliotis and G. Bekey, "Collective localization: a distributed kalman filter approach to localization of groups of mobile robots," in *Proceedings 2000 ICRA. Millennium Conference. IEEE International Conference on Robotics and Automation. Symposia Proceedings (Cat. No.00CH37065)*, vol. 3, 2000, pp. 2958–2965 vol.3.

[13] L. C. Carrillo-Arce, E. D. Nerurkar, J. L. Gordillo, and S. I. Roumeliotis, "Decentralized multi-robot cooperative localization using covariance intersection," in *2013 IEEE/RSJ International Conference on Intelligent Robots and Systems*, 2013, pp. 1412–1417.

[14] R. Wang, J. Du, Z. Xiong, X. Chen, *et al.*, "Hierarchical collaborative navigation method for uav swarm," *Journal of Aerospace Engineering*, vol. 34, no. 1, pp. 1–14, 2021.

[15] S. Lee and B. Min, "Distributed direction of arrival estimation-aided cyberattack detection in networked multi-robot systems," in *2018 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, 2018, pp. 1–9.

[16] E. Balaban, A. Saxena, P. Bansal, K. F. Goebel, *et al.*, "Modeling, detection, and disambiguation of sensor faults for aerospace applications," *IEEE Sensors Journal*, vol. 9, no. 12, pp. 1907–1917, 2009.

[17] F. Kong, M. Xu, J. Weimer, O. Sokolsky, *et al.*, "Cyber-physical system checkpointing and recovery," in *2018 ACM/IEEE 9th International Conference on Cyber-Physical Systems (ICCPS)*, 2018, pp. 22–31.

[18] A. Devonport and M. Arcak, "Data-driven reachable set computation using adaptive gaussian process classification and monte carlo methods," in *2020 American Control Conference*, 2020, pp. 2629–2634.

[19] P. J. Bonczek and N. Bezzo, "Detection of hidden attacks on cyber-physical systems from serial magnitude and sign randomness inconsistencies," in *2021 American Control Conference*, 2021, p. 3281–3287.

[20] T. Carvalho, L. Ferreira, and D. C. Silva, "A study in search pattern efficiency using under-actuated aircraft," *International Journal of Engineering and Technology*, vol. 13, no. 3, pp. 24–30, 2021.

[21] K. R. Gabriel and R. R. Sokal, "A New Statistical Approach to Geographic Variation Analysis," *Systematic Biology*, vol. 18, no. 3, pp. 259–278, 09 1969.