Preprints of the 23rd IFAC World Congress
Busan, Republic of Korea, August 23-28, 2026

Adaptive Filtering and Dual Compensation for Resilient Coverage Control against Coordinated Cyber Attacks

Yun Gao [*,**] **, Yanjing Huang** [*] **, Hao Gao** [*] **, Kaishun Wu** [***] **, and Yiding Ji** [*]

[*] *Robotics and Autonomous Systems Thrust, The Hong Kong University of Science and Technology (Guangzhou), Guangzhou, China (e-mail: y.gao@gaoyunailab.com; {yhuang403, hgao906}@connect.hkust-gz.edu.cn; jiyiding@hkust-gz.edu.cn.).*
[**] *Department of Electronic and Computer Engineering, The Hong Kong University of Science and Technology, Hong Kong SAR, China.*
[***] *Information Hub, The Hong Kong University of Science and Technology (Guangzhou), Guangzhou, China (e-mail: wuks@hkust-gz.edu.cn).*

Abstract: This paper studies resilient coverage control for multi-robot systems under coordinated cyber attacks (CCA). We first propose an adaptive safety-belt that screens exchanged neighbor information to identify compromised updates using increment-based consistency constraints. Additionally, a nonlinear attack observer is constructed, which reconstructs adversarial perturbations from the residual between observed and predicted neighbor motions. Based on these estimates, we design a dual-layer coverage controller for attack compensation, which corrects corrupted position vectors of the Voronoi computation at the state layer and mitigates residual attack-induced deviations at the control level. Then an input-to-state type practical stability analysis is established for the coverage error of the closed-loop system, which ensures that the robots converge to a bounded neighborhood of the nominal centroidal configuration under persistent coordinated attacks. Extensive simulations under various attacks further validate the resilience of the proposed method compared with several baseline methods.

Keywords:
Multi-robot systems; coverage control; coordinated attacks; adaptive filtering; observer design.

1. INTRODUCTION

Multi-robot systems (MRS) have become an enabling paradigm for large-scale autonomous exploration and environmental monitoring (Debie et al., 2023). By distributing sensing, computation, and control across many relatively simple robots, MRS can accomplish tasks that are difficult for a single robot, such as environmental mapping, target search, infrastructure inspection, and disaster response (Toumieh and Floreano, 2024). A central problem in these applications is coverage control, where robots must distribute themselves over a spatial domain to improve sensing quality, balance workload, and maintain stable coordination (Gao et al., 2025a,b, 2026). Unlike centralized deployment, distributed MRS coverage must reconcile local autonomy with global coordination, allowing each robot to respond to environmental variations and neighboring robots' motions (Chen et al., 2024).

A natural formulation of this problem is Voronoi-based coverage control, where the workspace is partitioned into dynamic Voronoi cells and each robot moves toward the centroid of its own cell (Liu et al., 2024). This geometric principle aligns sensing or actuation effort with spatial importance, enabling scalable and energy-efficient deployment (Lin et al., 2025). However, because each Voronoi cell is determined by neighboring robot states, coverage performance is highly sensitive to the consistency of exchanged geometric information. In complex and time-varying environments, maintaining such consistency requires reliable information exchange and state estimation, which are increasingly difficult to guarantee in networked robotic systems (Pratissoli et al., 2025).

This dependence on local communication makes Voronoi-based coverage vulnerable to cyber attacks. FDI, replay, and DoS attacks can corrupt exchanged state information or interrupt inter-robot communication (Gao et al., 2022). For coverage control, such attacks can distort the partition geometry, break centroidal consistency, and trigger cascading degradation of the coverage pattern. Consequently, robots may cluster in redundant regions while critical areas remain unobserved, resulting in fragmented coverage and degraded mission performance (Strobel et al., 2023).

Related work. Existing defenses against cyber attacks have mainly focused on multi-robot consensus (Zhou et al., 2021), where resilient filters, trust-weighted estimators, and topology reconfiguration methods have been developed to counter specific threats such as data injection, DoS, and spoofing (Yemini et al., 2021). However, most

⋆ This work is supported by National Natural Science Foundation of China under grants 62303389 and 62373289; Guangdong Province Scientific Research Platform and Project Scheme under grant 2024KTSCX039; Youth Talent Support Program of Guangdong Association for Science and Technology under grant SKXRC2025463 and Guangdong Provincial Key Lab of Integrated Communication, Sensing and Computation for Ubiquitous Internet of Things (No.2023B1212010007). *(Corresponding authors: Yiding Ji and Kaishun Wu.)*

©2026 the authors. Accepted by IFAC for publication under a Creative Commons Licence CC-BY-NC-ND

existing methods address isolated attack modes and rarely consider coordinated attacks that jointly affect sensing and communication channels. In contrast, coverage control under cyber attacks remains far less studied. Most Voronoi-based coverage methods are designed for attack-free or noise-limited settings and primarily target efficiency rather than resilience. Although several recent works have considered robust coverage through redundant sensing or robust control (Gong et al., 2024), they generally overlook coordinated cyber attacks (CCA). Therefore, security-aware coverage control remains absent in the literature.

Contributions. To address the vulnerability of distributed coverage control under CCA, this work introduces a resilience-enhanced control framework for MRS. The proposed method employs an adaptive safety-belt (ASB) to enforce increment-based consistency screening and a nonlinear attack observer (NLAO) to estimate adversarial perturbations from prediction residuals. These components are integrated into an attack-aware coverage controller that compensates for both geometric state deviations and control-input disturbances, allowing the MRS to achieve resilient coverage performance under CCA. The main contributions are summarized as follows:

(1) We develop an ASB mechanism for geometry-aware screening of exchanged neighbor states in Voronoi-based coverage control. Unlike fixed-threshold or redundancy-based filters, the proposed ASB constructs an online motion-consistency admissible set by combining local neighbor-increment envelopes with prediction-residual information, enabling selective rejection of corrupted updates while reducing false alarms caused by legitimate fast motions.
(2) We propose an NLAO tailored to Voronoi-driven coverage dynamics, which estimates adversarial perturbations from prediction–residual discrepancies without requiring redundancy or probabilistic attack models. Under bounded attack variation, the observer provides a provable bound on the attack-estimation error.
(3) We design a dual-layer resilient coverage controller that mitigates CCA-induced deviations at both the state and control levels. The state-layer correction recovers a locally consistent Voronoi tessellation, whereas the input-layer compensation attenuates residual attack-induced deviations in the robot motion. We establish an input-to-state (ISS) type practical stability result (Lakshmikantham et al., 1990), ensuring that the coverage error remains ultimately bounded and converges to a neighborhood of the nominal centroidal configuration under CCA.

Organization. Section 2 introduces the system model, attack model, and problem formulation. Section 3 presents the proposed resilient coverage framework and ISS practical stability analysis. Section 4 provides extensive simulation results, followed by concluding remarks in Section 5.

2. PRELIMINARIES

2.1 Multi-Robot Systems Model

Consider a team of $n \in \mathbb{Z}_+$ mobile robots operating in a compact and convex planar domain $\mathcal{Q} \subset \mathbb{R}^2$. At time k, the position of robot i is denoted by $p_i(k) \in \mathbb{R}^2$, with the robot index set $\mathcal{I} = \{1, \dots, n\}$. The motion of each robot follows the discrete-time first-order integrator dynamics:

$$p_i(k+1) = p_i(k) + T_s u_i(k), \tag{1}$$

where $T_s > 0$ is the sampling period and $u_i(k) \in \mathbb{R}^2$ denotes the control input, satisfying $\|u_i(k)\| \le v_{\max}$, with $v_{\max}$ denoting the maximum allowable speed.

Each robot is equipped with a wireless transceiver that enables communication with other robots located within a communication radius $R_c > 0$. Specifically, robots i and j can communicate at time k if $\|p_i(k) - p_j(k)\| \le R_c$. Whenever communication is available, robot j can transmit its position vector to robot i, and the position message received by robot i from robot j is denoted by $p_j^i(k)$. The set of robots satisfying the communication condition constitutes the neighbor set $\mathcal{N}_i(k) \subseteq \mathcal{I}$ of robot i. The overall communication topology at time k is captured through an undirected graph $\mathcal{G}(k) = (\mathcal{V}, \mathcal{E}(k))$, where $\mathcal{V} = \mathcal{I}$ is the node set and $(i,j) \in \mathcal{E}(k)$ represents an available bidirectional communication link. Accordingly, the adjacency matrix $A(k) = [a_{ij}(k)] \in \{0,1\}^{n\times n}$ is defined such that $a_{ij}(k) = 1$ if $(i,j) \in \mathcal{E}(k)$ and $a_{ij}(k) = 0$ otherwise. Since the graph is undirected, $A(k)$ is symmetric and satisfies $a_{ii}(k) = 0$ for all $i \in \mathcal{I}$. The neighbor set $\mathcal{N}_i(k)$ can thus be expressed as $\mathcal{N}_i(k) = \{j \in \mathcal{I} \setminus \{i\} : a_{ij}(k) = 1\}$.

In addition, each robot employs a virtual interaction function, defined in (2), over a circular region $\mathcal{C}(p_i(k), R_{vi}) = \{\, q \in \mathcal{Q} : \|q - p_i(k)\| \le R_{vi} \,\}$, where $R_{vi} > 0$ is the interaction radius. This function does not model physical sensing; instead, it represents a virtual marking or attenuation effect that locally reduces the information density, thereby generating the gradient needed for coverage computation.

$$f(q, p_i(k)) = \begin{cases} \beta e^{-\alpha \|q - p_i(k)\|^2} - \mu & q \in \mathcal{C}(p_i(k), R_{vi}), \\ 0 & q \notin \mathcal{C}(p_i(k), R_{vi}), \end{cases} \tag{2}$$

where $\beta \in (0,1)$ and $\alpha > 0$ shape the interaction intensity, and $\mu = \beta e^{-\alpha R_{vi}^2}$ makes the interaction vanish at the boundary $\partial\mathcal{C}(p_i(k), R_{vi})$.

2.2 Coordinated Cyber Attack Model

In this study, we consider an adversary whose objective is to compromise a subset of robots in the MRS or their communication links and manipulate the exchanged positional information so as to disrupt the coverage control process. Three canonical attack mechanisms, namely FDI, replay, and DoS, are incorporated into a CCA model.

In the FDI mode, the adversary injects perturbations into the transmitted position information of the robots. The falsified position $p_j^{\mathrm{FDI}}(k)$ of robot j at time k is given by:

$$p_j^{\mathrm{FDI}}(k) = p_j(k) + \theta_j(k), \tag{3}$$

where $\theta_j(k) \in \mathbb{R}^2$ is an adversarial perturbation. In the replay mode, the adversary records legitimate transmitted information and later substitutes the current packet with an outdated one. Let $\tau_j \in \mathbb{Z}_+$ be the replay window of robot j. The compromised transmission is modeled as:

$$p_j^{\mathrm{REP}}(k) = p_j(k - \tau_j), \tag{4}$$

effectively freezing the information flow and preventing the coverage gradient from evolving correctly. In the DoS mode, the adversary blocks the communication channel by preventing successful packet delivery, which is modeled by

setting $a_{ij}(k) = 0$. In this case, robot i receives no fresh information from robot j and uses the most recently stored position, which is abstracted as:

$$p_j^{\mathrm{DOS}}(k) = p_j(k-1). \tag{5}$$

Let $p_j^i(k)$ denote the position message of robot j received by i at time k. To model CCA, we introduce activation indicators $r_0, r_1, r_2, r_3 \in \{0,1\}$ with $r_0 + r_1 + r_2 + r_3 = 1$, thus exactly one of the four modes is active at each transmission. The received message is then defined as:

$$p_j^i(k) = r_0 p_j(k) + r_1\, p_j^{\mathrm{FDI}}(k) + r_2\, p_j^{\mathrm{REP}}(k) + r_3\, p_j^{\mathrm{DOS}}(k). \tag{6}$$

Furthermore, we define a unified attack vector that provides a consistent representation of position deviations in situations where the attack status is uncertain:

$$\delta_j^i(k) = p_j^i(k) - p_j(k). \tag{7}$$

We assume that the attack deviation has bounded amplitude and bounded variation, *i.e.*, there exist constants $z > 0$ and $\varepsilon > 0$ such that $\|\delta_j^i(k)\| \le z$ and $\|\delta_j^i(k+1) - \delta_j^i(k)\| \le \varepsilon$. Furthermore, the collection of all attack vectors received by robot i from its neighbors at time k can be represented compactly as $\delta^i(k) = [\delta_1^i(k), \dots, \delta_j^i(k), \dots]^\top$. Aggregating these terms over all robots yields the global attack matrix, denoted by $\delta(k) = [\delta^1(k), \dots, \delta^i(k), \dots]$.

2.3 Problem Formulation

Consider $m \in \mathbb{Z}_+$ targets distributed in Q, and let $s_l \in \mathbb{R}^2$ denote the location of target l. Each target acts as an information source described by the Gaussian function $g(q, s_l) = \eta_l \exp(-\|q - s_l\|^2/(2\sigma_l^2))$, where η_l is the peak intensity and σ_l determines the attenuation rate. The environment information density is:

$$\phi(q,k) = \sum_{l=1}^{m} g(q, s_l). \tag{8}$$

The environment information field evolves under the influence of (2), which models how robot i virtually marks the environment for coverage computation. The information field update is given by:

$$\phi(q,k) = (1 - f(q, p_i(k)))\phi(q, k-1). \tag{9}$$

Since $f(q, p_i(k))\phi(q,k)$ represents the instantaneous reduction of information at point q, the robots aim to maximize the global reduction across the domain. This motivates the coverage objective:

$$\mathcal{H}(P(k), V(k), k) = \sum_{i=1}^{n} \int_{V_i(k)} f(q, p_i(k))\phi(q,k)dq, \tag{10}$$

where $V_i(k)$ is the Voronoi cell of robot i, defined as:

$$V_i(k) = \{\, q \in \mathcal{Q} : \|q - p_i(k)\| \le \|q - p_j(k)\|,\ \forall j \ne i\}. \tag{11}$$

Following standard coverage control, the gradient of $\mathcal{H}$ with respect to $p_i(k)$ yields a centroid-seeking motion. Therefore, in the absence of CCA, the control law becomes:

$$u_i(k) = -2\kappa M_i(k)(p_i(k) - C_i(k)), \tag{12}$$

where $M_i(k)$ and $C_i(k)$ are the corresponding mass and centroid of Voronoi cell i. They are computed as follows:

$$M_i(k) = \int_{V_i(k)} 2\beta\alpha e^{-\alpha\|q - p_i(k)\|^2} \phi(q,k)dq, \tag{13}$$

$$C_i(k) = \frac{1}{M_i(k)} \int_{V_i(k)} 2\beta\alpha e^{-\alpha\|q - p_i(k)\|^2} \phi(q,k) q dq. \tag{14}$$

Remark 1. *The values of $M_i(k)$ and $C_i(k)$ depend implicitly on the neighboring positions that determine the Voronoi cell of robot i. To emphasize the dependency, we write $M_i(k, \{p_j(k)\}_{j\in\mathcal{N}_i(k)})$ and $C_i(k, \{p_j(k)\}_{j\in\mathcal{N}_i(k)})$. For simplicity, $C(k, \mathbf{p}(k))$ denotes the stacked centroid vector.*

Problem 1 (Resilient coverage control)**.** *Consider a team of n robots governed by* (1) *and operating over the information field* (8) *under the CCA model* (6)*. The objective is to design a distributed attack-resilient controller using only each robot's own position, locally received neighbor messages, and local memory, such that the resulting coverage error $e_i(k) = p_i(k) - C_i(k)$ remains practically stable in the presence of persistent attacks, namely, there exists a finite residual radius $\ell_{\mathrm{res}} > 0$ such that:*

$$\limsup_{k\to\infty} \|e_i(k)\| \le \ell_{\mathrm{res}}. \tag{15}$$

Moreover, ℓ_{res} should decrease as the attack-estimation residual and uncompensated attack magnitude decrease. In the attack-free case, the controller should recover the nominal centroid-seeking coverage behavior.

3. RESILIENT COVERAGE CONTROL

3.1 Control Scheme Overview

This section develops a resilient coverage control framework that integrates adaptive state screening, nonlinear attack estimation, and dual-layer compensation. Each robot first employs an adaptive safety belt (ASB) filter to assess the motion consistency of its neighbors and identify suspicious state updates. When inconsistency is detected, a nonlinear attack observer (NLAO) is activated to reconstruct adversarial perturbations from prediction residuals. The resulting attack estimates are then incorporated into a two-layer compensation mechanism. At the state level, they are used to correct corrupted neighbor positions and recover a consistent Voronoi tessellation for centroid computation. At the control level, they contribute to compensating the induced disturbance in the coverage input, preserving the nominal descent behavior of the closed-loop coverage dynamics. In this way, our architecture simultaneously restores geometric consistency and mitigates attack-induced dynamical deviations under CCA.

3.2 Adaptive Safety-Belt Filter

Conventional hard-threshold filtering rejects any state deviation beyond a fixed bound. However, this criterion is ill-suited to Voronoi-based coverage control, in which the local geometric configuration evolves continuously and is coupled to neighbor motion, target-density variation, and virtual interaction effects. Under the dynamics, abrupt yet legitimate state changes arise naturally, making fixed-threshold screening prone to false alarms. To overcome this limitation, we introduce an ASB filter that adjusts its admissible range following the local coverage evolution.

To model local motion consistency among neighboring robots and develop robust screening against corrupted position vectors, we first define the observed increment:

$$\Delta p_j^i(k) = p_j^i(k) - p_j^i(k-1), \tag{16}$$

which measures the magnitude of the most recent state variation received by robot i from its neighbor j. Based on

the observed increments from all neighbors, the minimum and maximum increment norms are defined as:

$$\underline{\pi_j^i(k)} = \min_{j\in\mathcal{N}_i(k)} \|\Delta p_j^i(k)\|, \tag{17}$$

$$\overline{\pi_j^i(k)} = \max_{j\in\mathcal{N}_i(k)} \|\Delta p_j^i(k)\|, \tag{18}$$

which provide a local envelope of plausible motion magnitudes under normal dynamics. To filter out outliers caused by CCA, we introduce an ASB for each robot i:

$$\mathcal{B}_i(k) = [(1-\iota\|d_j^i(k)\|)\underline{\pi_j^i(k)}, (1+\iota\|d_j^i(k)\|)\overline{\pi_j^i(k)}], \tag{19}$$

where $d_j^i(k) = p_j^i(k) - \hat{p}_j^i(k)$ is a predicted position error determined by the predicted neighbor position $\hat{p}_j^i(k)$, and $\iota > 0$ is a sensitivity constant. The term $\iota\|d_j^i(k)\|$ quantifies the tolerable motion deviation induced by $\hat{p}_j^i(k)$ and thereby determines the width of the ASB.

Theorem 1 (Local State Consistency)**.** *Consider the case where no attacks exist and all robots follow the kinematic model* (1) *with bounded control inputs* $u_i(k)$*. For every robot* i *and every neighbor* $j \in \mathcal{N}_i(k)$*, the observed increment norm of the true neighbor-state update satisfies:*

$$\|\Delta p_j^i(k)\| \in \mathcal{B}_i(k). \tag{20}$$

Theorem 1 states that, under normal Voronoi-induced motion, the observed neighbor increments are bounded within the corresponding ASB. The ASB gate is given as:

$$\varpi_j^i(k) = \begin{cases} 0, & \text{if } \|\Delta p_j^i(k)\| \in \mathcal{B}_i(k), \\ 1, & \text{if } \|\Delta p_j^i(k)\| \notin \mathcal{B}_i(k). \end{cases} \tag{21}$$

3.3 Nonlinear Attack Observer

To compensate for the position deviations induced by CCA, we develop an NLAO. Since robot i cannot access the true control input $u_j(k)$ of its neighbor j, it must rely exclusively on previously received neighbor-state reports to infer the neighbor's motion. Using the two most recent observations $p_j^i(k-2)$ and $p_j^i(k-1)$, robot i forms a finite-difference estimate of its neighbor's control input at $k-1$:

$$\hat{u}_j(k-1) = \frac{p_j^i(k-1) - p_j^i(k-2)}{T_s}, \tag{22}$$

which provides the best available prediction of the neighbor's motion based solely on received position updates. Then, the predicted position of neighbor j is obtained using the attack-free kinematic model:

$$\hat{p}_j^i(k) = 2p_j^i(k-1) - p_j^i(k-2). \tag{23}$$

Define the predicted increment and residual as:

$$\Delta\hat{p}_j^i(k) = \hat{p}_j^i(k) - \hat{p}_j^i(k-1), \tag{24}$$

$$D_j^i(k) = \Delta p_j^i(k) - \Delta\hat{p}_j^i(k). \tag{25}$$

Based on these quantities, the NLAO is given by:

$$\hat{\delta}_j^i(k+1) = \hat{\delta}_j^i(k) + \varpi_j^i(k)\Gamma\varphi(d_j^i(k)), \tag{26}$$

where $\Gamma \in \mathbb{R}^{2\times 2}$ is a positive definite gain matrix and $\varphi(\cdot)$ is a nonlinear saturation operator defined by:

$$\varphi(x) = \frac{x}{\max\{1, \|x\|/h\}}, \tag{27}$$

where $h > 0$ denotes the saturation radius. Under this design, if $\|\Delta p_j^i(k)\| \in \mathcal{B}_i(k)$, then the incremental residual satisfies $\|D_j^i(k)\| = 0$, implying $\varpi_j^i(k) = 0$ and the NLAO remains inactive. Conversely, when the observed motion departs from the ASB $\mathcal{B}_i(k)$, the resulting nonzero incremental residual $\|D_j^i(k)\| \neq 0$ activates the NLAO, which employs $d_j^i(k)$ to track the true attack vector $\delta_j^i(k)$.

Theorem 2 (Boundedness of the NLAO)**.** *Consider the attack model in* (6)*, and suppose that the attack variation satisfies* $\|\delta_j^i(k+1) - \delta_j^i(k)\| \le \varepsilon$*. Then, for the NLAO* (26)*, the prediction error is bounded as:*

$$\|\delta_j^i(k) - \hat{\delta}_j^i(k)\| \le \frac{\varepsilon}{\lambda_{\min}(\Gamma)} + h, \tag{28}$$

where $\lambda_{\min}(\Gamma)$ *is the smallest eigenvalue of* Γ*.*

3.4 Double-Layer Compensation Control Mechanism

This section presents a control law equipped with a dual compensation mechanism that operates at both the state and control levels. For state compensation, the NLAO facilitates the robot i to maintain an online estimate $\hat{\delta}_j^i(k)$ of the adversarial deviation embedded in the received position $p_j^i(k)$ for each neighbor j. To prevent corrupted data from interrupting the Voronoi computation, we define:

$$\tilde{p}_j^i(k) = p_j^i(k) - \hat{\delta}_j^i(k). \tag{29}$$

Robot i uses the corrected neighbor positions to construct a locally consistent Voronoi partition. Specifically, $V_i(k)$ is computed from $\{\tilde{p}_j^i(k) : j \in \mathcal{N}_i(k)\}$ rather than the corrupted messages $\{p_j^i(k) : j \in \mathcal{N}_i(k)\}$. Thus, the correction removes attack-induced geometric bias and affects the coverage law through the resulting partition.

Second, for control-level compensation, the resilient control law is given by:

$$\tilde{u}_i(k) = -2\kappa M_i(k)(p_i(k) - C_i(k)) - K_i \sum_{j\in\mathcal{N}_i(k)} \varpi_j^i(k)\hat{\delta}_j^i(k), \tag{30}$$

where $K_i > 0$ is the compensation gain. The term $-K_i\sum_j \varpi_j^i(k)\hat{\delta}_j^i(k)$ compensates the residual disturbance injected by the attacker that affects the robot's motion dynamics. The proposed law achieves resilience to CCA through a unified incremental-distortion modeling and compensation framework. The ASB filter prevents corrupted neighbor states from entering the Voronoi computation, while the NLAO reconstructs and compensates residual adversarial perturbations online. Consequently, the closed-loop dynamics remain bounded and converge to a coverage-consistent neighborhood even under CCA.

3.5 Practical Stability Analysis

We now analyze the closed-loop stability of the coverage controller under CCA, and introduce three assumptions:

Assumption 1. *Under the control law* (12)*, there exists a constant* $\rho \in (0,1)$ *such that for all* $k \ge 0$*:*

$$\|e(k+1)\| \le \rho\|e(k)\|, \tag{31}$$

whenever the MRS evolves in the absence of CCA, where $e(k) = [e_1(k), \ldots, e_n(k)]^\top$*.*

Assumption 2. *The predicted attack vector residual* $\chi_j^i(k) = \delta_j^i(k) - \hat{\delta}_j^i(k)$ *is globally bounded. That is, there exists a constant* $\xi > 0$ *such that* $\|\chi_j^i(k)\| \le \xi$*.*

Assumption 3. *There exists a constant* $\zeta > 0$ *such that, for any time* k *and any global attack matrix* $\delta(k)$ *of the neighbor states, the corresponding variation of the centroid vector* $C(k, \{p_j(k)\})$ *satisfies:*

$$\|C(k, \{p_j(k)\}) - C(k, \{p_j(k) + \delta_j^i(k)\})\| \le \zeta\|\delta(k)\|. \tag{32}$$

We are now ready to state the main stability result.

Theorem 3 (Practical Stability of the Resilient Coverage Control)**.** *Consider the MRS model* (1)*, the CCA model* (6)*, the ASB* (19)*, the NLAO* (26)*, and the resilient coverage control law* (30)*. Suppose that Assumptions 1–3 hold and the attack signals $\delta_j^i(k)$ are bounded. Then there exists a positive constant $\gamma > 0$ such that the coverage error norm $\|\mathbf{e}(k)\|$ satisfies:*

$$\|\mathbf{e}(k+1)\| \le \rho\|\mathbf{e}(k)\| + \gamma. \tag{33}$$

Proof. Based on (1) and (12), the kinematics of the MRS in the attack-free case can be rewritten as:

$$p_i^{\mathrm{na}}(k+1) = p_i(k) - 2\kappa T_s M_i(k)\big(p_i(k) - C_i(k)\big). \tag{34}$$

Accordingly, under (30), the actual kinematics of the MRS under unknown attack conditions can be expressed as:

$$p_i(k+1) = p_i^{\mathrm{na}}(k+1) - T_s K_i \sum_{j\in\mathcal{N}_i(k)} \varpi_j^i(k)\hat{\delta}_j^i(k). \tag{35}$$

Let $e_i^{\mathrm{na}}(k+1) = p_i^{\mathrm{na}}(k+1) - C_i^{\mathrm{na}}(k+1)$ denote the attack-free coverage error. Quantities without the superscript na refer to the attacked closed-loop system. Then we obtain:

$$\begin{aligned} e_i(k+1) &= (p_i(k+1) - C_i^{\mathrm{na}}(k+1)) + (C_i^{\mathrm{na}}(k+1) - C_i(k+1)) \\ &= e_i^{\mathrm{na}}(k+1) - T_s K_i \sum_{j\in\mathcal{N}_i(k)} \varpi_j^i(k)\hat{\delta}_j^i(k) \\ &\quad + (C_i^{\mathrm{na}}(k+1) - C_i(k+1)). \end{aligned} \tag{36}$$

By the triangle inequality:

$$\begin{aligned} \|e_i(k+1)\| &\le \|e_i^{\mathrm{na}}(k+1)\| + T_s K_i \sum_{j\in\mathcal{N}_i(k)} \varpi_j^i(k)\|\hat{\delta}_j^i(k)\| \\ &\quad + \big\|C_i^{\mathrm{na}}(k+1) - C_i(k+1)\big\|. \end{aligned} \tag{37}$$

We now bound the three terms in (37). First, by Assumption 1, the error satisfies:

$$\|e_i^{\mathrm{na}}(k+1)\| \le \rho\|e_i^{\mathrm{na}}(k)\|. \tag{38}$$

Second, using Assumption 2, we have:

$$\|\hat{\delta}_j^i(k)\| \le \|\delta_j^i(k)\| + \|\chi_j^i(k)\| \le z + \xi. \tag{39}$$

Since $\varpi_j^i(k) \in \{0,1\}$ and the neighbor set is finite, there exists a constant $c_1 > 0$ such that:

$$T_s K_i \sum_{j\in\mathcal{N}_i(k)} \varpi_j^i(k)\|\hat{\delta}_j^i(k)\| \le c_1(z+\xi). \tag{40}$$

Third, by Assumption 3, we have:

$$\|C^{\mathrm{na}}(k+1) - C(k+1)\| \le \zeta\|\delta(k)\| \le c_2\zeta z, \tag{41}$$

where $c_2 > 0$ is a constant that depends on the number of attacked robots. Combining (37), (38), (39), and (41), and stacking all robots, we obtain:

$$\|\mathbf{e}(k+1)\| \le \rho\|\mathbf{e}(k)\| + c_1(z+\xi) + c_2\zeta z. \tag{42}$$

Let $\gamma = c_1(z+\xi) + c_2\zeta z$. Substituting this expression yields (33). This establishes an ISS-type bound for the closed-loop coverage dynamics, implying that $\|\mathbf{e}(k)\|$ converges exponentially to a neighborhood whose radius is proportional to the residual attack bound ξ. □

4. SIMULATION RESULTS

The simulation is conducted in a bounded workspace $\mathcal{Q} = \mathrm{conv}\{(0,0),(0,25),(25,25),(25,0)\}$, with all coordinates measured in meters. The domain is discretized by a uniform grid with resolution $\Delta x = \Delta y = 1\,\mathrm{m}$, and the system is simulated with $T_s = 0.1\,\mathrm{s}$ for 500 time steps.

Consider $n = 25$ mobile robots initialized uniformly in $\mathcal{Q}$. They follow (1) with $\|u_i(k)\| \le v_{\max} = 0.3\,\mathrm{m/s}$. The communication and virtual interaction radii are set to $R_c = 5.5\,\mathrm{m}$ and $R_{vi} = 3.0\,\mathrm{m}$, respectively. The coverage parameters are $\alpha = 0.7$, $\beta = 0.8$, and $\kappa = 0.6$. For ASB and NLAO, we set $\iota = 0.5$, $\Gamma = \mathrm{diag}(0.8, 0.8)$, and $h = 0.6$. The compensation gain in (30) is $K_i = 0.35$ for all i.

We first compare three scenarios: the attack-free case, the attacked case without defense, and the attacked case with the proposed framework (see Fig. 1). The adversarial scenario involves five compromised nodes under CCA, with the attacks activated over $k \in [100, 400]$. From the randomized initial configuration in Fig. 1(a), the attack-free system reaches an approximate centroidal Voronoi tessellation in Fig. 1(b). Under CCA without defense, the nominal controller loses this centroidal structure, causing robot clustering and under-observed regions, as shown in Fig. 1(c). In contrast, Fig. 1(d) shows that our method preserves a more balanced distribution consistent with the underlying information density under persistent CCA.

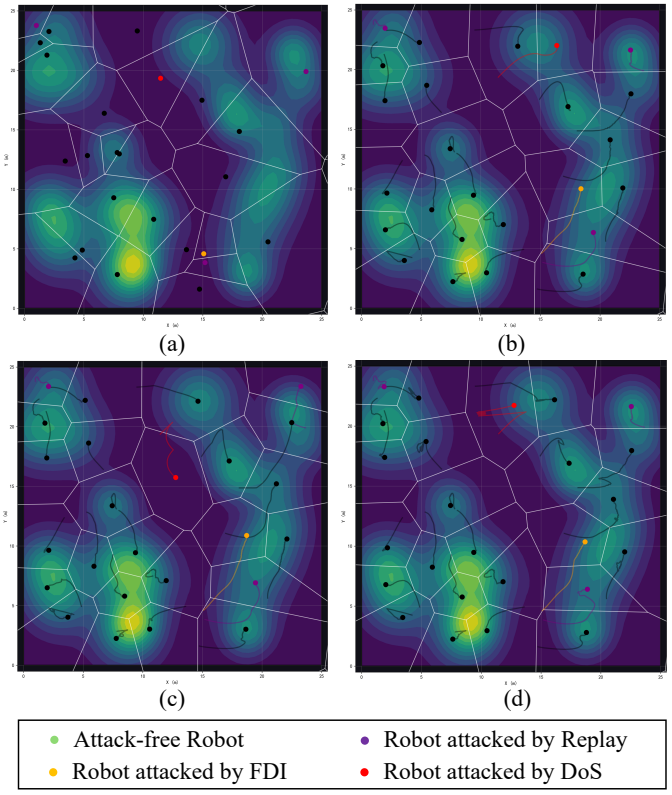
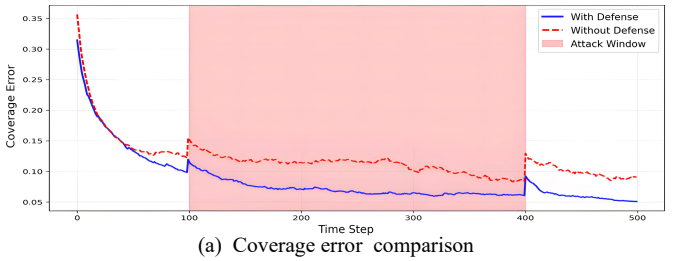


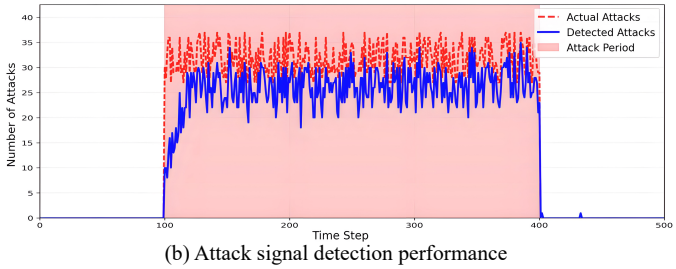
Fig. 1. Evolution of the MRS under different conditions: (a) initial robot configuration; (b) final attack-free configuration; (c) final configuration under CCA and without defense; (d) final configuration under CCA using our proposed method.

Fig. 2 illustrates the evolution of the global coverage error and the performance of the anomaly detection. The left panel compares the system performance with and without the proposed defense strategy. Upon the initiation of the combined FDI, replay, and DoS attacks at step $k = 100$, the undefended system exhibits an immediate performance degradation, evidenced by a sharp spike in the coverage error. Throughout the attack window $k \in [100, 400]$, the baseline controller fails to converge, as the persistent corruption of neighbor positions prevents the nominal cover-

age law from recovering centroidal consistency. In contrast, the proposed resilient framework effectively mitigates the adversarial influence, keeping the coverage error bounded and driving it to a significantly lower steady-state value, comparable to nominal operation. The right panel validates the efficacy of the anomaly detection module. The detected attack frequency closely tracks the actual number of malicious interactions, successfully identifying the majority of attacks while producing few false detections during non-attack periods. These results demonstrate that the proposed framework can effectively detect and mitigate CCA while maintaining resilient coverage performance.



(a) Coverage error comparison



(b) Attack signal detection performance

Fig. 2. Coverage-error and attack-detection performance under CCA. Left: evolution of the centroid tracking error. Right: comparison between the actual number of malicious interactions and the detected attacks.

Table 1 further compares our method with two representative resilient-consensus baselines, mean-subsequence-reduced (MSR) and dynamic rate-bound filtering (DRBF), also the no-defense case. Our framework achieves the lowest average error (0.076 ± 0.008) and final error (0.037), substantially better over the others. It also provides the highest detection rate (93.1%) and the lowest false-positive rate (3.4%) among all methods. In comparison, MSR reduces peak error but suffers from poor final convergence, while DRBF improves detection relative to MSR but still yields more false positives. These results indicate that our framework provides a more balanced defense against CCA.

Table 1. Quantitative comparison under CCA.

Method	Avg. Error	Max Error	Final Error	DR (%)	FPR (%)
ASB–NLAO	0.076±0.008	0.124	0.037	93.1	3.4
DRBF	0.093±0.012	0.151	0.078	71.6	14.2
MSR	0.105±0.015	0.130	0.081	54.8	21.0
No Defense	0.124±0.019	0.158	0.115	–	–

5. CONCLUSION

This paper proposed a Voronoi-based resilient coverage framework for MRS under CCA. By integrating an ASB, an NLAO, and dual-level compensation, our method identifies corrupted neighbor states, estimates adversarial deviations, and mitigates attacks at both the geometric and control levels without controller mode switching. We provide an ISS-type stability analysis for the coverage error bound, ensuring convergence to a neighborhood of the nominal centroidal configuration. Simulations with mixed attacks demonstrated that our method achieves balanced coverage, reduces coverage degradation, and outperforms baselines. Future work will extend our method to heterogeneous teams and learning based resilient coordination.

REFERENCES

Chen, J., Jayawardhana, B., and de Marina, H.G. (2024). Scaling up the formation of agents with heterogeneous sensing: mixed distance and bearing-only. *IEEE Trans. Autom. Control*, 69(11), 8011–8017.

Debie, E., Kasmarik, K., and Garratt, M. (2023). Swarm robotics: A survey from a multi-tasking perspective. *ACM Comput. Surv.*, 56(2), 1–38.

Gao, Y., Gao, H., Ji, Y., Zhou, J., and Shi, Y. (2025a). A dual calibration framework for exploring environments using heterogeneous robot swarms. In *51st Annual Conference of the IEEE Industrial Electronics Society*.

Gao, Y., Gao, H., Ma, W., Xiong, H., and Ji, Y. (2026). Coverage control with load consistency for dynamic targets in robot swarms using graph attention networks. In *IEEE International Conf. on Robotics and Automation*.

Gao, Y., Luo, K., Fang, C., and He, J. (2022). Fragility-aware stealthy attack strategy for multi-robot systems against multi-hop wireless networks. In *61st IEEE Conference on Decision and Control*, 4827–4832.

Gao, Y., Zhou, Z., Gao, H., Zhang, S., and Ji, Y. (2025b). Hybrid control for robot swarms to detect critical nodes in heterogeneous sensor networks. In *IEEE 21st Intl. Conf. on Automation Sci. and Engineering*, 3462–3467.

Gong, J., Kim, H., and Lee, S. (2024). Resilient multi-robot coverage path redistribution using boustrophedon decomposition for environmental monitoring. *Sensors*, 24(23), 7482.

Lakshmikantham, V., Leela, S., and Martynyuk, A.A. (1990). *Practical stability of nonlinear systems*. World Scientific.

Lin, R., Kim, S., and Egerstedt, M. (2025). Heterogeneous collaborative pursuit via coverage control driven by Fokker-Planck equations. *IEEE Trans. on Robotics*.

Liu, Q., Zhang, Z., Le, N.K., Qin, J., Liu, F., and Hirche, S. (2024). Distributed coverage control of constrained constant-speed unicycle multi-agent systems. *IEEE Trans. Automation Science Engineering*, 22, 2225–2240.

Pratissoli, F., Mantovani, M., Prorok, A., and Sabattini, L. (2025). Distributed coverage control for time-varying spatial processes. *IEEE Trans. Robot.*, 41, 1602–1617.

Strobel, V., Pacheco, A., and Dorigo, M. (2023). Robot swarms neutralize harmful byzantine robots using a blockchain-based token economy. *Science Robotics*, 8(79), eabm4636.

Toumieh, C. and Floreano, D. (2024). High-speed motion planning for aerial swarms in unknown and cluttered environments. *IEEE Trans. on Robotics*, 40, 3642–3656.

Yemini, M., Nedić, A., Goldsmith, A.J., and Gil, S. (2021). Characterizing trust and resilience in distributed consensus for cyberphysical systems. *IEEE Transactions on Robotics*, 38(1), 71–91.

Zhou, J., Lv, Y., Wen, G., and Yu, X. (2021). Resilient consensus of multiagent systems under malicious attacks: Appointed-time observer-based approach. *IEEE Trans. Cybern.*, 52(10), 10187–10199.