Метод обнаружения аномального поведения агентов в роевых робототехнических системах с использованием показателя локальной дезорганизации

П. Ю. Шамрай, Т. Ю. Мельников, Е. А. Домницкий, И. Ю. Попов, Д. А. Заколдаев

***Аннотация*— В ряде прикладных задач требуется сохранять периодическую пространственную структуру в рое робототехнических средств (РРТС). В данной работе введён локальный показатель дезорганизации, позволяющий выявлять аномалии в пространственной конфигурации роя на основе навигационной информации, в частности данных о расстояниях между агентами. На основе предложенного показателя разработан метод обнаружения и локализации нарушителя, использующий обмен информацией о значении показателя и идентификаторе агента вплоть до второго координатного слоя. Для разделения высокоуровневой информации и данных о геометрическом положении агентов введено понятие навигационного поля. Проведённые численные эксперименты подтвердили чувствительность разработанного метода обнаружения для рассматриваемой модели нарушителя. Ограничение метода связано с уязвимостью к византийским атакам, что может быть решено дальнейшим добавлением показателей репутации и доверия и использования распределенного консенсуса. Полученные результаты могут служить основой для дальнейших исследований в области обнаружения нарушений информационной безопасности РРТС.**

***Ключевые слова*— роевые робототехнические системы, аномальное поведение, информационная безопасность.**

I. Введение

Применение групп беспилотных воздушных судов (далее БВС) в структурированных, и в частности, в периодических формациях открывает новые возможности для решения сложных задач, требующих координированного взаимодействия множества агентов.

Поисково-спасательные операции, воздушное наблюдение и охрана могут быть выполнены группой БВС в линейных или решетчатых формациях для обеспечения равномерного покрытия целевой области. Нарушение однородности построения в такой формации приводит к появлению «слепых зон», которые могут привести к потере разыскиваемых людей или же могут быть использованы для незаметного проникновения на защищаемые объекты. Картографирование и обработка сельскохозяйственных угодий требуют точного соблюдения геометрической структуры формации для обеспечения равномерного охвата и предотвращения пропусков при опрыскивании или съемке. Отклонение отдельных агентов от заданной структуры ведет к неравномерной обработке посевов или появлению артефактов на создаваемых картах [1]. Так же высокая точность формации может требоваться при осуществлении группового целеуказания - при этом БВС выстраиваются в сенсорную сеть. В данных условиях отклонения в структуре группы будут приводить к увеличению ошибки локализации целей.

Таким образом, возможность контролировать качество пространственной структуры необходима в указанных выше задачах, так как от этого зависит исправность функционирования группы и корректность результатов ее работы. Поэтому, тема настоящего исследования является актуальным и требует внимательного изучения.

Ряд исследований посвящен выявлению неисправностей на уровне отдельных БВС. В [2] предложили метод обнаружения и изоляции отказов для формаций БВС с динамической структурой, использующую фильтры Калмана и χ^2-тесты для идентификации неисправных лидеров и ведомых агентов. В [3] представили гибридный модуль обнаружения аномалий, сочетающий LSTM-автоэнкодер с Isolation Forest для анализа телеметрии в режиме реального времени на встроенных платформах. В [4] предложили метод распределенных наблюдателей для одновременного обнаружения отказов и управления формацией. Исследователи в [5] предложили двухступенчатый подход, основанный на глубоком обучении для определения аномального поведения агентов из группы БВС с последующим установлением возможной причины отклонений для подозрительного агента. В работе [6] был представлен метод обнаружения аномалий в группе БВС с использованием подхода паритетного пространства, где каждое судно сравнивает значение энтропии Цаллиса своих соседей для определения неисправного члена группы.

Другие исследования посвящены противодействию злонамеренным воздействиям на БВС. В своей

Статья получена 29 января 2025.

П.Ю. Шамрай, факультет безопасности информационных технологий, Университет ИТМО (email: pavel.shamray@mail.ru)

Т. Ю. Мельников, факультет безопасности информационных технологий, Университет ИТМО (email: tim_melnikov@mail.ru)

Е. А. Домницкий, факультет безопасности информационных технологий, Университет ИТМО (email: egor.dom0923@gmail.com)

И. Ю. Попов, факультет безопасности информационных технологий, Университет ИТМО (email: ilyapopov27@gmail.com)

Д. А. Заколдаев, факультет безопасности информационных технологий, Университет ИТМО (email: d.zakoldaev@mail.ru)

диссертации [7] автор проанализировал архитектуру сетей 5-6G и предложил систему обнаружения вторжений на основе машинного обучения для выявления злонамеренных воздействий на БВС. В работе [8] исследователи предложили фреймворк для развертывания систем обнаружения сетевых вторжений на гетерогенной группе БВС с учетом ресурсных ограничений. В исследовании [9] представили CoDetect - схему кооперативного обнаружения аномалий, включающую механизм аутентификации и использующую алгоритм консенсуса для решения задачи о византийских генералах в группе БВС.

Несмотря на обширные исследования в области обнаружения аномального поведения отдельных БВС и системных аномалий, задача целенаправленного выявления факта нарушения однородности пространственной структуры формации остается недостаточно изученной. Большинство существующих подходов фокусируются на обнаружении и обеспечении устойчивости при отказах полезной нагрузки или поиске аномалий в поведении отдельных агентов. При этом часто рассматривают группу в целом с использованием глобальной информации. В свою очередь, отклонения в геометрической конфигурации группы остаются за рамками рассмотрения. Вместе с этим упускается тот факт, что каждый агент может располагать ограниченной, локальной информацией о ближайших соседях. Это создает необходимость в разработке специализированного метода, обеспечивающего возможность обнаруживать нарушения заданной пространственной структуры, что и является предметом настоящего исследования. Необходимо гарантировать обнаружение неоднородности в отсутствие высокоуровневого информационного сообщения и при наличии только локальной сенсорной информации.

Таким образом, цель работы - обеспечить возможность локального обнаружения неоднородности пространственной структуры формации группы РРТС. Для этого предложен показатель локальной дезорганизации группы, позволяющий децентрализованно, на каждом агенте, определить нарушение пространственной однородности, а также предложен подход к идентификации аномального агента на основе обмена информацией, содержащей этот показатель и идентификатор агента.

II. Модель и ограничения

A. Введение понятия навигационного поля

Пусть взаимодействие и обмен данными между агентами роевой робототехнической системы (РРТС) осуществляется в информационном поле $\mathbb{I}$. В информационное поле входит любая информация передаваемая в сети РРТС независимо от способа передачи. Это могут быть, в том числе, данные о деталях миссии, принимаемые в реальном времени для достижения консенсуса между агентами, а также данные о состоянии других агентов, например, для принятия решении о дальнейшем движении или возвращении на базу агентов с малым энергетическим запасом. Так, в информационном поле $\mathbb{I}$ можно выделить навигационное поле $\mathbb{N} \subset \mathbb{I}$, в которое входят:

$$\mathbb{N} \supset \mathbb{A} \cup \mathbb{P} \cup \mathbb{M} \cup \mathbb{E} \cup \mathbb{S} \cup \mathbb{S}_{nb},$$

где $\mathbb{A}$ – алгоритмы взаимодействия агентов, выполнения элементов миссии (например, движение к точке, облет препятствий, движение по прямой), $\mathbb{P}$ - параметры алгоритмов (могут меняться во время выполнения миссии), $\mathbb{M}$ – миссии (с точки зрения навигации, то есть информация о маршруте, целевых координатах, скоростях, без описания действий, не связанных с движением), $\mathbb{E}$ – информация об окружающей среде (препятствия, агенты не входящие в группу, погодные условия), влияющая на движение, $\mathbb{S}$ – состояния агентов (координата, скорость, ускорения, угловые скорости, магнитный курс, показания иных датчиков), $\mathbb{S}_{nb}$ – состояние соседей агентов из группы (по аналогии с собственным состоянием), а также показатели взаимодействия с ними (такие как дальность, углы, относительная скорость до соседей).

Область навигационного поля может быть уменьшена до локального навигационного поля $\mathbb{N}_{loc}$, включающее в себя только ту навигационную информацию, которую агент может получить от своих датчиков и вычислительных устройств, но не которую он получил в результате обмена информационными сообщениями (посредством высокоуровневых протоколов) с другими участниками группы. Локальность навигационного поля агента обусловлена в первую очередь тем, что датчики агента как правило имеют ограничения по дальности действия. Так, сигналы ГНСС уже не входят в $\mathbb{N}_{loc}$.

Среди особенностей локального навигационного поля можно выделить:

- динамическая топология, предварительный расчет которой невозможен;
- ограниченность систем связи агента по дальности;
- невозможность получение агентом глобальной информации о группе в отсутствие обмена высокоуровневым информационными сообщениями;
- изначальное отсутствие возможности оценить непротиворечивость получаемой информации.

Мотивацией к выделению отдельного навигационного поля является недостаточное внимание к вопросам информационной безопасности в условиях радиомолчания и отсутствия связи между агентами в группе и внешними акторами.

B. Модель нормального функционирования РРТС

Каждый агент РРТС обладает:

- набором аппаратно-программных средств (работающих в НП), позволяющим ему определять взаимное расположение судов в группе без использования обмена высокоуровневыми сообщениями в ИП;
- информацией о выполняемой миссии $\mathbb{M}$, содержащую последовательность абстрактных примитивов (примитивами называются промежуточные путевые точки, их последовательности, линии заданного пути и проч.).

Задача РРТС завершить миссию с минимальными затратами энергии (с минимальными временными задержками) и минимальными отклонениями от примитивов миссии (все примитивы должны быть пройдены)

C. Модель нарушителя

Нарушителем является агент:
- программно- и аппаратно-идентичный агентам роя;
- обладающий или не обладающий информацией об актуальной полетной миссии $\mathbb{M}$ (внешний нарушитель обладает только устаревшей полетной миссией, тогда как внутренний может внести в актуальную миссию злонамеренные изменения);
- обладающий алгоритмами пространственного взаимодействия $\mathbb{A}$;
- обладающий знанием о параметрах алгоритмов пространственного взаимодействия $\mathbb{P}$ (внешний нарушитель имеет неактуальные параметры, внутренний вносит злонамеренные изменения);

Ожидаемые воздействия нарушителя на функционирование РРТС:
- Предоставлять искаженную информацию о своем состоянии $\mathbb{S}$.
- Двигаться в противодействие остальным участникам роя (с риском скомпрометировать себя), тем самым, например, уводя рой с маршрута, либо внося возмущение в структуру и снижая энергетический запас агентов.

Эти два типа воздействия являются идентичными с точки зрения агентов РРТС, так как они будут формировать команды управления только на основе навигационной информации.

Агент может быть скомпрометирован как посредством внутреннего саботажа, так и в случае перехвата и последующего внедрения третьей стороной. В первом случае оператор или техник, а во втором, злоумышленник третьей стороны, имеющие доступ к аппаратной и программной части, могут изменить алгоритмы взаимодействия агентов $\mathbb{A}$, параметры алгоритмов $\mathbb{P}$ и детали миссии $\mathbb{M}$.

III. Метод обнаружения нарушителя

Для оценки отклонения от периодической формации роя введен показатель дезорганизации пространственной структуры:

$$C_d = 1 - \frac{1}{N_{nbr}} \sum_{i=1..N_{nbr}} e^{-\frac{|r_i - r_0|}{\alpha r_0}} \qquad (1)$$

где N_{nbr} – количество соседей, r_i – расстояние до i-го агента, r_0 – целевое расстояние между агентами, α – коэффициент чувствительности. Значение $C_d = 0$ – идеальная периодическая структура, $C_d = 1$ – полная дезорганизация.

Такая форма показателя выбрана с учетом имеющийся информации только о ближайших соседях и периодической пространственной структуры.

Важно отметить, что расстояние учитывается не только между агентом и его ближайшими соседями, но и непосредственно между ближайшими соседями. Соседи могут определяться как дальностью действия сенсоров, так и искусственным ограничением на дальность. Во втором случае можно ограничиваться первым координатным слоем – то есть теми агентами, которые располагаются на расстоянии $r_i < 1{,}6r_0$ (Рисунок 1). Соседи второго координатного слоя располагаются на расстоянии $1{,}6r_0 < r_i < 3{,}2r_0$

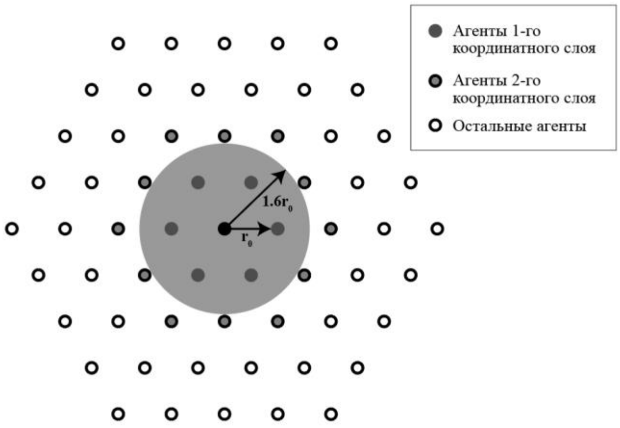

Рисунок 1. Изображение периодической пространственной структуры РРТС.

Фактом обнаружения внутреннего нарушителя ИБ РРТС является переход его состояния из множества соседей $\mathbb{S}_{nb}$ в множество объектов окружающей среды $\mathbb{E}$. Локальное обнаружение конкретным агентом или группой агентов локально не позволит идентифицировать нарушителя. Для этого необходимо обменяться информацией, то есть выйти за пределы локального навигационного поля $\mathbb{N}_{loc}$. Для этого на каждой итерации управления каждый i-ый агент:

1. Вычисляет $C_d{}^{(i)}$с ближайшими соседями N_{nbr_i}.
2. Передает $C_d{}^{(i)}$ и множество идентификаторов N_{nbr_i} всем ближайшим соседям и их соседям (агентам из второго слоя).

3. Получает $\left\{(C_d{}^{(j)}, N_{nbr\_j})\right\}_{j=1}^{N_{nbr\_i}}$ и для каждого j из N_{nbr_i} $\left\{(C_d{}^{(k)}, N_{nbr\_k})\right\}_{k=1}^{N_{nbr\_j}}$

4. Вычисляет для j из N_{nbr_i} мгновенный вклад в «дезорганизацию» $\overline{D^{(j)}} = \frac{1}{N_{peer}^{(j)}} \sum_{k=1}^{N_{nbr\_j}} (1 - C_d{}^{(k)})$, где $N_{peer}^{(j)}$- число агентов, для который j – ближайший сосед и участвует в вычислении их локальной организационной метрики.

5. Накапливает в течение некоторого количества итераций T кумулятивную *метрику нарушителя* для соседей j из N_{nbr_i} : $M_{intruder\_j} = \sum_{it}^{T} \overline{D^{(j)}}$ и применяет min-max масштабирование для вектора метрик ближайших соседей размером N_{nbr_i} .

Так как для расчета метрики нарушителя $M_{intruder\_j}$ происходит накопление информации о локальной дезорганизации, то это сглаживает временные задержки между отправкой, приемом и обработкой этой информации. Степень влияния этого фактора на выявление нарушителя может зависеть от скорости движения нарушителя. Так как обмен осуществляется максимум с агентами из второго координатного слоя, то в данном методе не учитываются возможные задержки при обмене информацией.

IV. Результаты численных экспериментов

В качестве модели роевого управления использовался метод искусственных потенциалов с тремя зонами (притяжения, нейтральной и отталкивания) [10]. Аномальный агент начинал движение в сторону виртуальной цели, расположенной в координатах (-1000,-1000). При этом осуществляя взаимодействие с остальными агентами группы (Рисунок 2).

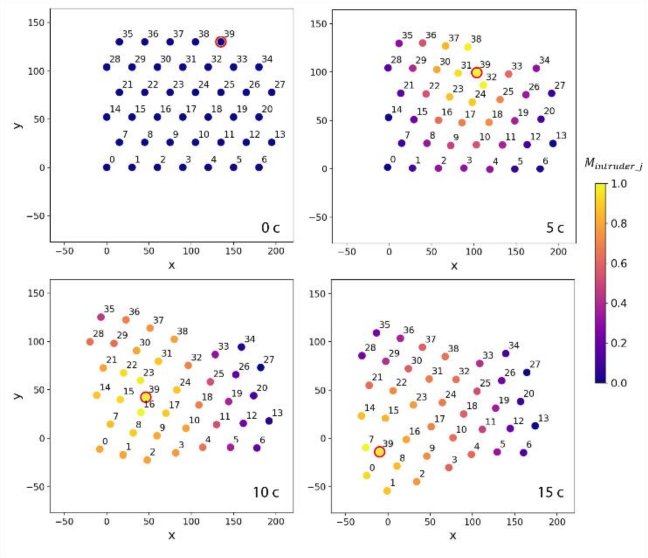

Рисунок 2. Положение агентов в РРТС при активном действии нарушителя (номер 39, выделен красной обводкой) в последовательные моменты симуляции. Цветом обозначена метрика нарушителя $M_{intruder\_j}$, рассчитанная на j-м агенте.

В ходе экспериментов варьировалось количество агентов группе от 10 до 40. С уменьшением количества агентов показатель становится менее чувствительным, потому что пространственная структура становится менее стабильной при внешнем воздействии.

Максимальное значение показателя локальной дезорганизации C_d и минимальное значение метрики нарушителя $M_{intruder\_j}$ уменьшаются при увеличении минимального за все время симуляции расстояния (Рисунок 3). Таким образом, можно говорить о возможности локализации нарушителя. Однако, стоит отметить, что при малой активности нарушителя, необходимо увеличивать чувствительность α, что снижает возможности по локализации из-за малой разницы между дисперсией показателя C_d с нарушителем и без него.

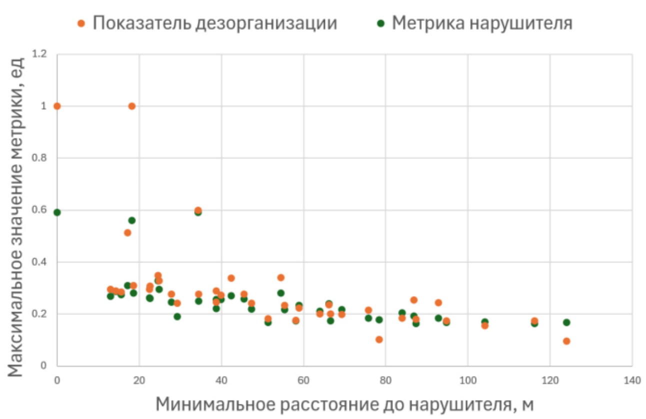

Рисунок 3. Зависимость минимального значения показателя локальной дезорганизации C_d и метрики нарушителя $M_{intruder\_j}$ от минимального расстояния до нарушителя для всех агентов.

V. Заключение

В работе введен показатель дезорганизации пространственной структуры в рое робототехнических средств, позволяющий идентифицировать аномалии в пространственной структуре. Важным отличием от существующих показателей является локальность и использование только навигационной информации при его расчете. В работе вводится понятие навигационного поля, необходимое для разделения высокоуровневой информации и информации о геометрическом местоположении агента. В качестве модели нарушителя был взят агент, преднамеренно или по неисправности осуществляющий движение к ложной цели, тем самым проходя через структуру группы. Было показано, что для агентов в рое, находящихся ближе к нарушителю за время его движения, показатель снижался сильнее, чем для более удаленных агентов. Также был предложен метод локализации нарушителя, однако требующий перехода из пространства локального навигационного поля в информационное.

Разработанный метод может быть применен в роевых робототехнических системах как для обнаружения аномального поведения агента, так и злоумышленника. Ограничением такого геометрического показателя является уязвимость к византийским атакам без применения дополнительных мер, например, использование распределенного консенсуса и вычисления показателей репутации и доверия. Дальнейшим шагом может быть использование показателя дезорганизации для выявления, а также идентификации нарушителя, не проявляющего активных действий. Более того, необходимо более четко исследовать вопрос влияния задержек при информационном обмене на эффективность выявления нарушителя.

Благодарности

Работа выполнена в Университете ИТМО при финансовой поддержке Министерства науки и высшего образования Российской Федерации в рамках проекта № 70-2024-001354 «Разработка технологий и демонстратора комплексной системы группового управления, взаимодействия и организации поведения группы БВС при выполнении целевых задач».

Библиография

[1] Y. Liu, K. Han, и W. Rasdorf, «Assessment and Prediction of Impact of Flight Configuration Factors on UAS-Based Photogrammetric Survey Accuracy», *Remote Sensing*, т. 14, вып. 16, с. 4119, янв. 2022, doi: 10.3390/rs14164119.

[2] M. Zaeri Amirani, N. Bigdeli, и M. Haeri, «Distributed fault detection and isolation in time-varying formation tracking UAV multi-agent systems», *Asian Journal of Control*, т. 25, вып. 1, сс. 604–622, 2023, doi: 10.1002/asjc.2821.

[3] M. J. C. S. Reis и A. J. D. Reis, «Edge-Based Real-Time Fault Detection in UAV Systems via B-Spline Telemetry Reconstruction and Lightweight Hybrid AI», *Sensors*, т. 25, вып. 16, с. 4944, янв. 2025, doi: 10.3390/s25164944.

[4] X. Su, M. Hao, R. Yang, K. Yue, и Z. Feng, «Distributed Observer-Based Simultaneous Fault Detection and Formation Control of Multi-UAVs», в *2023 42nd Chinese Control Conference (CCC)*, июл. 2023, сс. 4968–4973. doi: 10.23919/CCC58697.2023.10240561.

[5] H. Ahn и S. Chung, «Deep learning-based anomaly detection for individual drone vehicles performing swarm missions», *Expert Systems with Applications*, т. 244, с. 122869, июн. 2024, doi: 10.1016/j.eswa.2023.122869.

[6] H. E. Sevil, «Anomaly Detection using Parity Space Approach in Team of UAVs with Entropy based Distributed Behavior», в *AIAA Scitech 2020 Forum*, в AIAA SciTech Forum. , American Institute of Aeronautics and Astronautics, 2020. doi: 10.2514/6.2020-1625.

[7] F. Alrefaei, «Machine Learning for Intrusion Detection into Unmanned Aerial System 6G Networks», *Doctoral Dissertations and Master's Theses*, май 2024, [Онлайн]. Доступно на: https://commons.erau.edu/edt/815

[8] V. Lannurien *и др.*, «A retrospective on DISPEED -- Leveraging heterogeneity in a drone swarm for IDS execution», 13 июнь 2025 г., *arXiv*: arXiv:2506.11800. doi: 10.48550/arXiv.2506.11800.

[9] T. Li, W. Lin, R. Ma, Z. Ma, Y. Shen, и J. Ma, «CoDetect: cooperative anomaly detection with privacy protection towards UAV swarm», *Sci. China Inf. Sci.*, т. 67, вып. 5, с. 159103, апр. 2024, doi: 10.1007/s11432-023-3984-7.

[10] A. Boyko и R. Girgidov, «Key features of a swarm assemlby algorythm for autonomous unmanned aerial vehicles (UAVs) in absence of GNSS and stable radio communication», *Rob. and Tech. Cyb. J.*, т. 10, вып. 1, сс. 25–31, мар. 2022, doi: 10.31776/RTCJ.10103.

A method for detecting anomalous behavior of agents in swarm robotic systems using the local disorganization metric

P.Y.Shamray, T.Y. Melnikov, E.A. Domnitsky, I.Y. Popov, D.A. Zakoldaev

***Abstract* — In a number of applied problems, it is necessary to preserve a periodic spatial structure within a swarm of robotic agents. This paper introduces a local disorganization metric that enables the detection of anomalies in the spatial configuration of the swarm based on navigation information, in particular data on inter-agent distances. Based on the proposed metric, a method for detecting and localizing an intruder is developed. The method relies on the exchange of information about the metric value and the agent identifier up to the second coordination layer. To separate high-level information from data describing the geometric positions of agents, the concept of a navigation field is introduced. Numerical simulations have confirmed the sensitivity of the proposed detection method for the considered adversary model. A limitation of the method is its vulnerability to Byzantine attacks, which can be addressed by further incorporating reputation and trust metrics and by employing distributed consensus mechanisms. The obtained results may serve as a basis for further research in the field of information security intruder detection in robotic swarms.**

***Keywords* — swarm robotics, anomaly behavior, information security.**

REFERENCES

[1] Y. Liu, K. Han, and W. Rasdorf, "Assessment and Prediction of Impact of Flight Configuration Factors on UAS-Based Photogrammetric Survey Accuracy," *Remote Sensing*, vol. 14, no. 16, p. 4119, Jan. 2022, doi: 10.3390/rs14164119.

[2] M. Zaeri Amirani, N. Bigdeli, and M. Haeri, "Distributed fault detection and isolation in time-varying formation tracking UAV multi-agent systems," *Asian Journal of Control*, vol. 25, no. 1, pp. 604–622, 2023, doi: 10.1002/asjc.2821.

[3] M. J. C. S. Reis and A. J. D. Reis, "Edge-Based Real-Time Fault Detection in UAV Systems via B-Spline Telemetry Reconstruction and Lightweight Hybrid AI," *Sensors*, vol. 25, no. 16, p. 4944, Jan. 2025, doi: 10.3390/s25164944.

[4] X. Su, M. Hao, R. Yang, K. Yue, and Z. Feng, "Distributed Observer-Based Simultaneous Fault Detection and Formation Control of Multi-UAVs," in *2023 42nd Chinese Control Conference (CCC)*, Jul. 2023, pp. 4968–4973. doi: 10.23919/CCC58697.2023.10240561.

[5] H. Ahn and S. Chung, "Deep learning-based anomaly detection for individual drone vehicles performing swarm missions," *Expert Systems with Applications*, vol. 244, p. 122869, Jun. 2024, doi: 10.1016/j.eswa.2023.122869.

[6] H. E. Sevil, "Anomaly Detection using Parity Space Approach in Team of UAVs with Entropy based Distributed Behavior," in *AIAA Scitech 2020 Forum*, in AIAA SciTech Forum. , American Institute of Aeronautics and Astronautics, 2020. doi: 10.2514/6.2020-1625.

[7] F. Alrefaei, "Machine Learning for Intrusion Detection into Unmanned Aerial System 6G Networks," *Doctoral Dissertations and Master's Theses*, May 2024, [Online]. Available: https://commons.erau.edu/edt/815

[8] V. Lannurien *et al.*, "A retrospective on DISPEED -- Leveraging heterogeneity in a drone swarm for IDS execution," Jun. 13, 2025, *arXiv*: arXiv:2506.11800. doi: 10.48550/arXiv.2506.11800.

[9] T. Li, W. Lin, R. Ma, Z. Ma, Y. Shen, and J. Ma, "CoDetect: cooperative anomaly detection with privacy protection towards UAV swarm," *Sci. China Inf. Sci.*, vol. 67, no. 5, p. 159103, Apr. 2024, doi: 10.1007/s11432-023-3984-7.

[10] A. Boyko and R. Girgidov, "Key features of a swarm assemlby algorythm for autonomous unmanned aerial vehicles (UAVs) in absence of GNSS and stable radio communication," *Rob. and Tech. Cyb. J.*, vol. 10, no. 1, pp. 25–31, Mar. 2022, doi: 10.31776/RTCJ.10103.