Situation Coverage Based Safety Analysis of an Autonomous Aerial Drone in a Mine Environment

Nawshin Mannan Proma
Computer Science
University of York
York, United Kingdom
nawshinmannan.proma@york.ac.uk

Victoria J Hodge
Computer Science
University of York
York, United Kingdom
victoria.hodge@york.ac.uk

Rob Alexander
Computer Science
University of York
York, United Kingdom
rob.alexander@york.ac.uk

***Abstract*—If we want to integrate autonomous aerial drones into safety-critical contexts, particularly in dynamic and hazardous environments like mining operations, we need to rigorously assure their safety. Despite significant technological advancements in drone technology over the past decade, this remains a challenge. The current safety engineering methods employed in drones cannot demonstrate convincingly that AI techniques can effectively mitigate unsafe situations with a specified level of confidence and reliability. In this paper, we present a brief study of various approaches, with particular focus on the situation coverage-based approach. A key challenge lies in identifying a finite set of representative situations for testing from the infinite possibilities that could occur in real-world scenarios. This research contributes to advancing our understanding of situation coverage based safety assessment methodologies and coverage criteria.**

***Index Terms*—drone, safety, testing, situation, coverage, mine**

I. Introduction

Autonomous Aerial Drones (AAD) have become popular across various domains, including military operations, environmental monitoring, and agricultural activities. Establishing a detailed safety assessment process for AAD, especially in settings with humans like mines, is crucial during the design phase [1]. Such a process aims to identify potential failure scenarios during AAD operation, assess their consequences, and define mitigation measures to minimize risks. To do this, it imperative that the process considers sufficient situations. A novel approach for this is situation coverage based safety testing [2], [3]. In the subsequent discussion, we will delve into the taxonomy of situation coverage based safety testing, exploring its foundational principles and applications derived from existing literature.

Recent literature focuses on finding representative situation for approving Autonomous Vehicles (AVs) using a situation-based approach [4], [5], [6] . The diverse literature uncovers various strategies contributing to different aspects of this approach.To address this, we developed a taxonomy (figure 1) to organize and understand the different stages to situation coverage based safety testing .

There are various situation sources like expert knowledge, standards, and driving data, which can be gathered from field tests or accident records. Recently, more organizations are making their driving data publicly accessible [7], [8], expanding the available resources. [9] propose a new method using drones to capture traffic data, which offers advantages like lower costs and less disruption, although it is limited to shorter sections of roads. However, there are still challenges, such as capturing highway scenarios effectively [9].

When generating scenarios, we can use either knowledge-based methods, relying on expert knowledge stored in ontologies, or data-driven approaches, often employing machine learning techniques like clustering. [10] proposed a fundamental ontology for AV guidance, which serves as a foundation for many subsequent studies. [11] utilized ontologies to create scenarios specifically for German highways, incorporating all layers of their model. For data-driven scenario generation, various methods exist, such as unsupervised clustering techniques by [12] [13], mixed similarity measures by [14], and Bayesian learning methods by [15]. These approaches aim to extract concrete scenarios from real driving data, which can be classified into logical scenario categories. Additionally, techniques like Kernel Density Estimation [16] and particle filters [17] are used to estimate and simulate scenario parameters from field data, ensuring the safety of autonomous vehicles.

Situation coverage(SC) measures can be approached from a Macro- or Micro- perspective [2]. Macro-SC means looking at the situation overall and saying whether certain things are covered.Micro-SC means that we watch the system run and see what small-scale situations are encountered.

The goal of falsification approaches is to identify counterexamples that violate safety requirements during micro assessment. These approaches can either select existing concrete scenarios from a database or define logical scenarios with parameter ranges. Selection methods include using accident databases, modifying existing scenarios to increase criticality, or identifying critical scenarios within predefined parameter ranges. Several studies, such as those by [18], [19] and [20], utilize accident data to understand system requirements and simulate accident scenarios for system evaluation. However, solely relying on accident data may not adequately assess the safety of autonomous vehicles (AVs) beyond Level 3 of autonomy [21], as it only addresses past accidents rather than predicting future risks. Methods like the one presented in [22] efficiently determine the risk of real traffic situations to select critical scenarios for AV testing, focusing on the behavior of other road users. Other studies, like those by [23] and [24], develop frameworks to consider scenario complexity when selecting challenging scenarios for AV testing, which has shown to reveal more system errors. These approaches play a crucial role in enhancing the safety assessment of AVs by identifying and addressing potential risks in complex real-world scenarios.

The aim of testing-based approaches for scenario selection is to efficiently sample a subset of concrete scenarios for micro safety assessment, which can then be aggregated for macro assessment [25]. These approaches typically involve one of two sampling methods: sampling within parameter ranges or sampling from parameter distributions to incorporate scenario probability. N-wise sampling is often applied to simpler systems like Lane-Keeping Assistants [26], while other studies utilize techniques such as Design of Experiments (DoE) for scenario generation [26]. Additionally, some research focuses on generating road networks or modifying AV behavior using methods like Signal Temporal Logic (STL) monitoring or randomization of traffic vehicle parameters [27]. Accelerated sampling methods, including Extreme Value Theory and Importance Sampling Theory, are also used to predict system safety levels based on real data and criticality metrics, significantly reducing the need for extensive real-world testing [28]. These diverse approaches contribute to the development of comprehensive testing methodologies for assessing AV safety.

In the following sections, we will discuss our problem and motivation, our proposed solution, and my current research status.

II. Problem and Motivation

Our proposed research aims to develop a system-level validation approach for autonomous aerial drone (AAD) to ensure their safety and quality of service in mine environments. Existing safety assurance methods focus either on component-level approaches, which lack adaptability to the system level approaches. Therefore, the research will focus on deriving a comprehensive system-level validation method for autonomous vehicles [27]. The inspiration for our work comes from [27] vision paper on system-level safety testing. Their proposal serves as a foundational framework, which we intend to adapt initially and refine later if necessary to suit our test case. Our focus lies on applying their preliminary solution for autonomous aerial drones operating within mine settings.

Our research will address the following three key research questions:

RQ1: How can we define situation coverage of system-level test suites for AAD?

To achieve measurable guarantees through testing, it is essential to define coverage criteria tailored to the application domain. While various coverage criteria exist for software testing, there is a lack of well-defined situation coverage criteria specific to autonomous systems testing [2]. Addressing this gap, my research will formally define situation coverage and explore how to define situation coverage of system-level test suites for AAD.

RQ2: How can we evaluate situation coverage of existing system-level test suites for AAD?

One primary application of the newly defined situation coverage criteria will be the evaluation of existing test suites. By applying situation coverage measurements to these test suites, we can compare different test suite generation approaches based on their achieved safety assurance level.

RQ3: How can we generate relevant test situations and systematically drive simulation towards critical scenarios to justifiably increase situation coverage?

The defined situation coverage concept can guide test scenario generation towards scenarios that provide high coverage. Developing a novel test suite generation approach, we aim to define, derive, and simulate complex test scenarios efficiently to increase situation coverage.

III. Proposed Solution

A. Test Environment

In our research, we established a test environment in our lab (depicted in Figure 2a) where a drone was utilized to gather point cloud data suitable for simulations (Figures 2b and 2c) from a 'mine' reconstructed in the Lab. This process was iterated twice, resulting in the creation of two mock mines one for training navigation algorithms and one for testing. Thus, the physical mine in the Lab and depicted in Figure 2a served as the basis for the simulated mine shown in Figure 2b, designated as the ALOFT: Self-Adaptive Drone Controller testbed [29]. Our objective is to employ ALOFT for situation coverage-based safety testing of AAD.

In our test scenario, a human is present on the landing area so the AAD cannot land safely. The safety property also requires the ALOFT self-adaptive drone controller to stop upon detecting a human within 3 seconds, ensuring a safe landing. We check to make sure the drone only lands when there are no people, making sure it follows safety rules.

B. Overall Approach

1) Step 1 : Simulation — We use simulators or real test vehicles to observe how AAD behave in specific situations defined by test contexts.
2) Step 2: Qualitative Abstraction— We simplify the geospatial, causal, and temporal information from these situations using graph queries, creating situation graphs that represent relationships as labeled graphs. These graphs are maintained during simulation.
3) Step 3: Runtime Monitoring — We continuously monitor changes in these situation graphs using complex event processing techniques, which provide precise formal semantics [30].
4) Step 4: Situation Coverage— We measure how well our existing test scenarios cover different situations on an abstract level. This will be done by adapting metrics that account for model diversity and graph shapes [31],

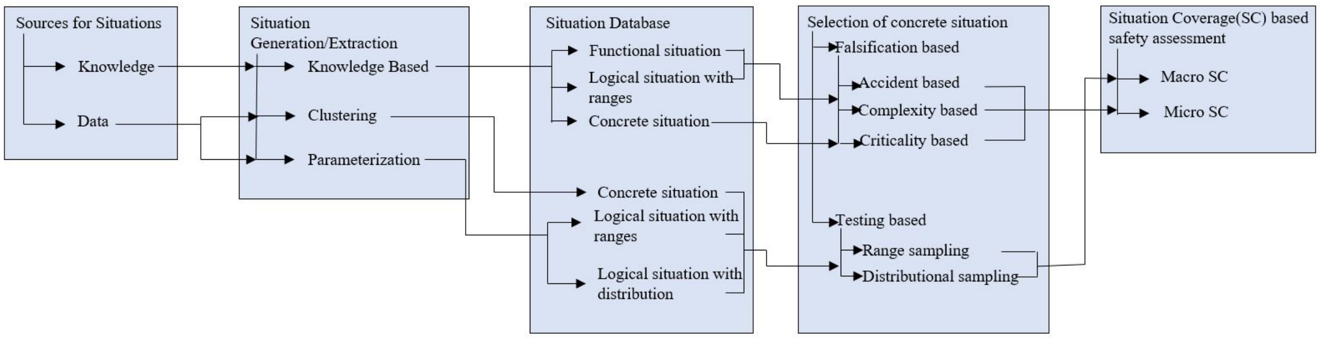

Fig. 1. Taxonomy of Situation Coverage Based Testing Approach.

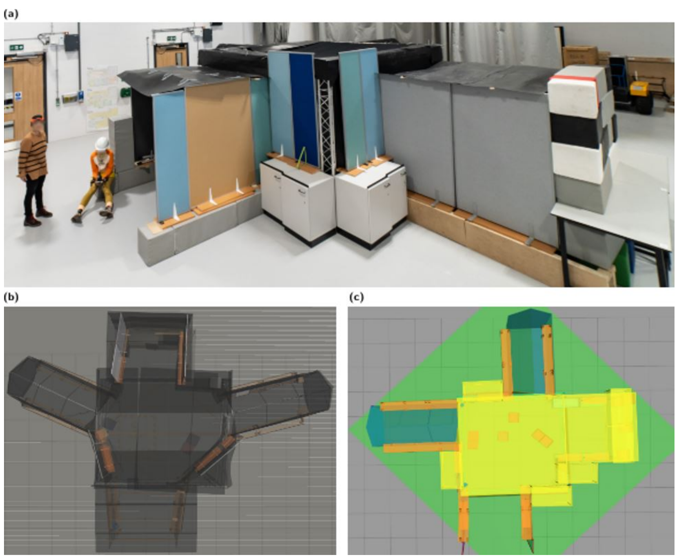

Fig. 2. ALOFT Testbed(from [29]).

ensuring a comprehensive understanding of covered scenarios.

5) Step 5 : Situation Generation— We automatically create new challenging situations as abstract test cases using diverse graph generation techniques. This helps us expand the range of scenarios covered by our tests [32].
6) Step 6: Context Generation— Finally, we aim to turn abstract situations into concrete test contexts. This step increases the practical coverage and robustness of our test suite, making it more effective in real-world scenarios.

IV. Planning

In this part, we describe our research plan based on the proposed solution of section III.

1) Data Acquisition from ROS and Gazebo Simulation [33], [34]:
   - Use ALOFT [29] to access kinematic data from the simulation environment.
   - Take relevant information including drone position, speed, and any human presence on the simulated landing area.
2) Metamodel Creation using Eclipse [35]:
   - Use Eclipse to make a metamodel that captures the essential elements and relationships of the landing scenario.
   - Define entities such as drone, landing area, human presence, and their respective attributes.
3) Situation Modeling with VIATRA [36]:
   - Derive the relations in the qualitative abstraction chain, we utilize graph queries employing the VIATRA syntax. This syntax allows us to express the situation model derived from the metamodel effectively.
   - Define graph patterns and rules to specify the conditions and behaviors within the landing scenario.
4) Continuous Scene Monitoring with Runtime Monitoring Algorithm [30]:
   - Implement a runtime monitoring algorithm using VIATRA to continuously monitor the simulated scene against safety constraints.
   - Define safety constraints such as the drone's response time to human presence and its speed reduction.

V. Current Research status

As a first-year PhD student, my goal by the end of 2024 is to integrate the landing scenario's metamodel into VIATRA. Then, I aim to create a basic set of tests for AAD, run them, and collect simple safety data, like whether a simulated scenario results in a collision. Once this initial integration is done, I'll focus on RQ1 and RQ2. By mid-2025, I'll define "situation coverage" formally and test it using existing test suites for AAD. Next, I'll tackle RQ3 by creating a method to generate scenarios that maximize situation coverage. I'll test this method experimentally, like I did for RQ2. Based on my current timeline, I anticipate finishing my PhD by the end of 2026.

References

[1] V. J. Hodge, R. Hawkins, and R. Alexander, "Deep reinforcement learning for drone navigation using sensor data," *Neural Computing and Applications*, vol. 33, no. 6, pp. 2015–2033, 2021.

[2] R. Alexander, H. R. Hawkins, and A. J. Rae, "Situation coverage–a coverage criterion for testing autonomous robots," tech. rep., Department of Computer Science, University of York, 2015.

[3] H. Hawkins and R. Alexander, "Situation coverage testing for a simulated autonomous car–an initial case study," *arXiv preprint arXiv:1911.06501*, 2019.

[4] Z. Tahir and R. Alexander, "Intersection focused situation coverage-based verification and validation framework for autonomous vehicles implemented in carla," in *Modelling and Simulation for Autonomous Systems: 8th International Conference, MESAS 2021, Virtual Event, October 13–14, 2021, Revised Selected Papers*, pp. 191–212, Springer, 2022.

[5] Z. Tahir and R. Alexander, "Coverage based testing for v&v and safety assurance of self-driving autonomous vehicles: A systematic literature review," in *2020 IEEE International Conference On Artificial Intelligence Testing (AITest)*, pp. 23–30, 2020.

[6] N. M. Proma and R. Alexander, "Systematic situation coverage versus random situation coverage for safety testing in an autonomous car simulation," in *Proceedings of the 12th Latin-American Symposium on Dependable and Secure Computing*, LADC '23, (New York, NY, USA), p. 208–213, Association for Computing Machinery, 2023.

[7] J. Guo, U. Kurup, and M. Shah, "Is it safe to drive? an overview of factors, metrics, and datasets for driveability assessment in autonomous driving," *IEEE Transactions on Intelligent Transportation Systems*, vol. 21, no. 8, pp. 3135–3151, 2019.

[8] Y. Kang, H. Yin, and C. Berger, "Test your self-driving algorithm: An overview of publicly available driving datasets and virtual testing environments," *IEEE Transactions on Intelligent Vehicles*, vol. 4, no. 2, pp. 171–185, 2019.

[9] R. Krajewski, J. Bock, L. Kloeker, and L. Eckstein, "The highd dataset: A drone dataset of naturalistic vehicle trajectories on german highways for validation of highly automated driving systems," in *2018 21st international conference on intelligent transportation systems (ITSC)*, pp. 2118–2125, IEEE, 2018.

[10] S. Geyer, M. Baltzer, B. Franz, S. Hakuli, M. Kauer, M. Kienle, S. Meier, T. Weißgerber, K. Bengler, R. Bruder, *et al.*, "Concept and development of a unified ontology for generating test and use-case catalogues for assisted and automated vehicle guidance," *IET Intelligent Transport Systems*, vol. 8, no. 3, pp. 183–189, 2014.

[11] G. Bagschik, T. Menzel, and M. Maurer, "Ontology based scene creation for the development of automated vehicles," in *2018 IEEE Intelligent Vehicles Symposium (IV)*, pp. 1813–1820, IEEE, 2018.

[12] F. Kruber, J. Wurst, and M. Botsch, "An unsupervised random forest clustering technique for automatic traffic scenario categorization," in *2018 21st International conference on intelligent transportation systems (ITSC)*, pp. 2811–2818, IEEE, 2018.

[13] F. Kruber, J. Wurst, E. S. Morales, S. Chakraborty, and M. Botsch, "Unsupervised and supervised learning with the random forest algorithm for traffic scenario clustering and classification," in *2019 IEEE Intelligent Vehicles Symposium (IV)*, pp. 2463–2470, IEEE, 2019.

[14] H. Watanabe, T. Maly, J. Wallner, T. Dirndorfer, M. Mai, and G. Prokop, "Methodology of scenario clustering for predictive safety functions," in *9. Tagung Automatisiertes Fahren*, 2019.

[15] W. Wang and D. Zhao, "Extracting traffic primitives directly from naturalistically logged data for self-driving applications," *IEEE Robotics and Automation Letters*, vol. 3, no. 2, pp. 1223–1229, 2018.

[16] E. de Gelder and J.-P. Paardekooper, "Assessment of automated driving systems using real-life scenarios," in *2017 IEEE Intelligent Vehicles Symposium (IV)*, pp. 589–594, IEEE, 2017.

[17] M. R. Zofka, F. Kuhnt, R. Kohlhaas, C. Rist, T. Schamm, and J. M. Zöllner, "Data-driven simulation and parametrization of traffic scenarios for the development of advanced driver assistance systems," in *2015 18th international conference on information fusion (FUSION)*, pp. 1422–1428, IEEE, 2015.

[18] L. Stark, M. Düring, S. Schoenawa, J. E. Maschke, and C. M. Do, "Quantifying vision zero: Crash avoidance in rural and motorway accident scenarios by combination of acc, aeb, and lks projected to german accident occurrence," *Traffic injury prevention*, vol. 20, no. sup1, pp. S126–S132, 2019.

[19] L. Stark, S. Obst, S. Schoenawa, and M. Düring, "Towards vision zero: Addressing white spots by accident data based adas design and evaluation," in *2019 IEEE International Conference on Vehicular Electronics and Safety (ICVES)*, pp. 1–6, IEEE, 2019.

[20] F. Fahrenkrog, L. Wang, T. Platzer, A. Fries, F. Raisch, and K. Kompaß, "Prospective safety effectiveness assessment of automated driving functions—from the methods to the results," in *ESV 26th Conference Search*, vol. 11, p. 2019, 2019.

[21] L. Grayned, "Sae international: Levels of autonomous driving and sensor packages," *Oak Ridge Institute for Science and Education: Oak Ridge, TN, USA*, 2020.

[22] A. Pierson, W. Schwarting, S. Karaman, and D. Rus, "Learning risk level set parameters from data sets for safer driving," in *2019 IEEE Intelligent Vehicles Symposium (IV)*, pp. 273–280, IEEE, 2019.

[23] C. Zhang, Y. Liu, Q. Zhang, and L. Wang, "A graded offline evaluation framework for intelligent vehicle's cognitive ability," in *2018 IEEE Intelligent Vehicles Symposium (IV)*, pp. 320–325, IEEE, 2018.

[24] Y. Qi, Y. Luo, K. Li, W. Kong, and Y. Wang, "A trajectory-based method for scenario analysis and test effort reduction for highly automated vehicle," tech. rep., SAE Technical Paper, 2019.

[25] P. M. Junietz, "Microscopic and macroscopic risk metrics for the safety validation of automated driving," 2019.

[26] H. Beglerovic, A. Ravi, N. Wikström, H.-M. Koegeler, A. Leitner, and J. Holzinger, "Model-based safety validation of the automated driving functio highway pilot," in *8th International Munich Chassis Symposium 2017: chassis. tech plus*, pp. 309–329, Springer, 2017.

[27] I. Majzik, O. Semeráth, C. Hajdu, K. Marussy, Z. Szatmári, Z. Micskei, A. Vörös, A. A. Babikian, and D. Varró, "Towards system-level testing with coverage guarantees for autonomous vehicles," in *2019 ACM/IEEE 22nd International Conference on Model Driven Engineering Languages and Systems (MODELS)*, pp. 89–94, IEEE, 2019.

[28] D. Åsljung, J. Nilsson, and J. Fredriksson, "Using extreme value theory for vehicle level safety validation and implications for autonomous vehicles," *IEEE Transactions on Intelligent Vehicles*, vol. 2, no. 4, pp. 288–297, 2017.

[29] C. C. Imrie, R. Howard, D. Thuremella, N. M. Proma, T. Pandey, P. B. W. Lewinska, R. Cannizzaro, R. D. Hawkins, C. Paterson, L. Kunze, *et al.*, "Aloft: Self-adaptive drone controller testbed," in *SEAMS'24: Proceedings of the 19th Symposium on Software Engineering for Adaptive and Self-Managing Systems*, ACM, 2024.

[30] I. Dávid, I. Ráth, and D. Varró, "Streaming model transformations by complex event processing," in *Model-Driven Engineering Languages and Systems: 17th International Conference, MODELS 2014, Valencia, Spain, September 28–October 3, 2014. Proceedings 17*, pp. 68–83, Springer, 2014.

[31] M. Búr, G. Szilágyi, A. Vörös, and D. Varró, "Distributed graph queries for runtime monitoring of cyber-physical systems," in *Fundamental Approaches to Software Engineering: 21st International Conference, FASE 2018, Held as Part of the European Joint Conferences on Theory and Practice of Software, ETAPS 2018, Thessaloniki, Greece, April 14-20, 2018, Proceedings 21*, pp. 111–128, Springer International Publishing, 2018.

[32] O. Semeráth, A. S. Nagy, and D. Varró, "A graph solver for the automated generation of consistent domain-specific models," in *Proceedings of the 40th international conference on software engineering*, pp. 969–980, 2018.

[33] ROS, "Ros.org — powering the world's robots." https://www.ros.org/, 2020.

[34] OSRF, "Gazebo." http://gazebosim.org/, 2019.

[35] R. Gronback, "Eclipse modeling project — the eclipse foundation." http://www.eclipse.org/modeling/emf. Accessed:2024-04-22.

[36] "Viatra - scalable reactive model transformations." http://www.eclipse.org/viatra. Accessed:2024-04-22.