Trusted Artificial Intelligence for Armaments in Uncertain Environments

Sami Saliba, Andrew Evans, Justin Abel
Mentored by Hunter Moore
University of Virginia Department of Systems and Information Engineering
August 20, 2024

Executive Summary

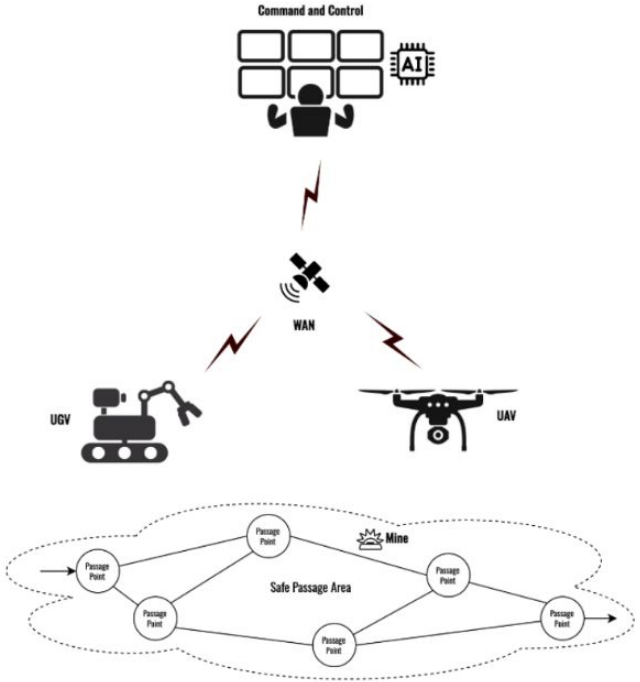
- Techniques to improve minefield traversal are introduced
  - Regression modeling to determine key factors in human and AI accuracy
  - Reinforcement learning methods to optimize UAV and UGV routing
  - Reliable system design to ensure warfighter trust
- Criteria for solution analysis are introduced and used to facilitate iterative design
  - These include human safety, traversal time, and accuracy variance
- Considerations for future project stages are explored

Challenge Introduction

- The integration of AI solutions is critical in maintaining high-performing, intelligent systems
- AI integration into military systems creates ethical challenges; human lives are on the line
- A reliable AI system functions as intended and performs well in high stress environments, developing trust amongst users

Problem

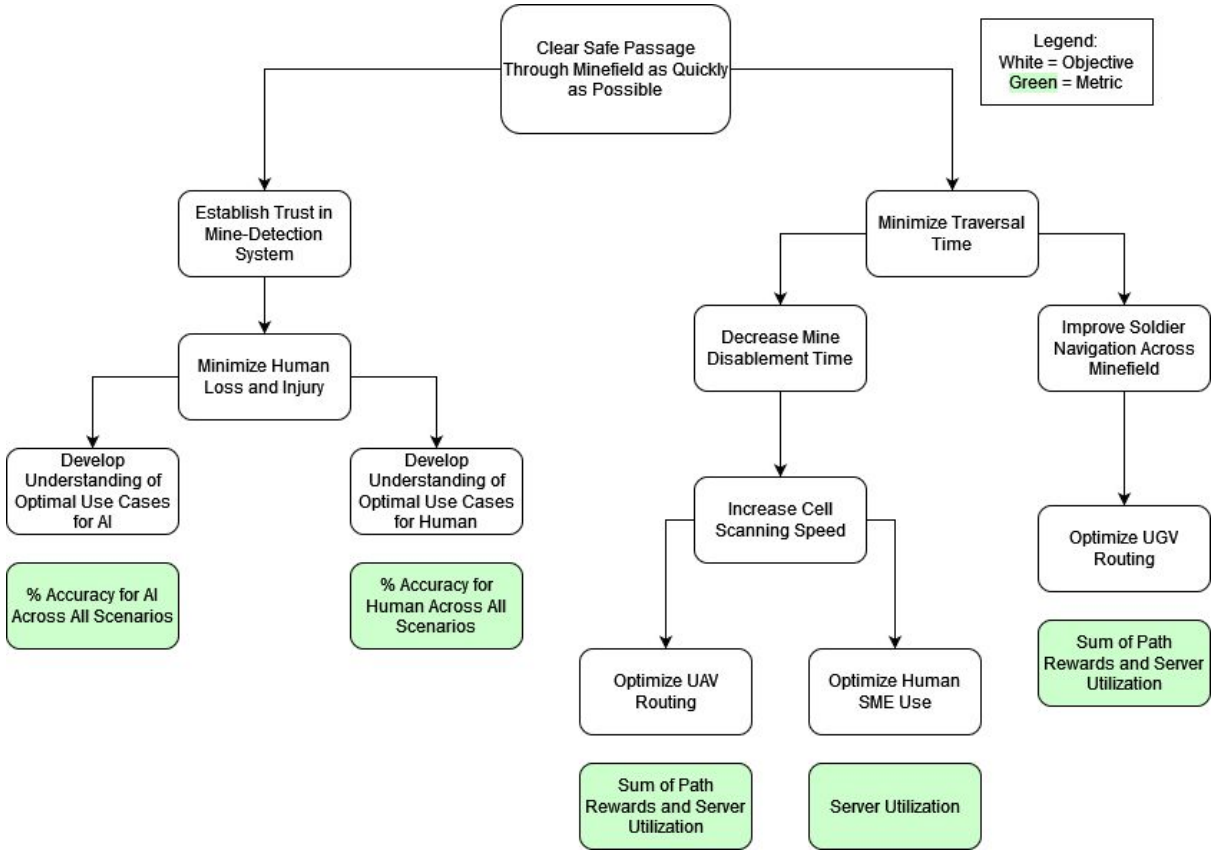
- How can safe passage through a minefield be achieved as quickly as possible while depending on unreliable subsystems?
- How can human trust in the AI mine-identification system be built?
    - Warfighter use of the system provides a competitive advantage on the battlefield



System Goals

- Clear safe passage through minefields as quickly as possible
  - Minimize Human Loss and Injury
  - Minimize Traversal Time
- Protect area assets and mission security
- Ensure trust and reliability in the AI mine-detection system
- Establish an overarching system that is more resilient than its subsystems

Objectives and Metrics for System Analysis



Criteria for Candidate Ranking

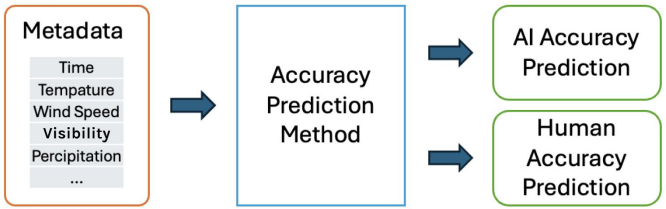
- Trust and reliability
- Traversal time
- Variability in solution performance across environmental conditions
  - Variation in confidence levels of detection accuracy
  - Performance in novel situations
- Resource utilization
  - What are the utilization rates of UAV, AI, human, UGV
    - Server utilization
    - Average number of processes occurring concurrently
- Monetary and time costs

Methods

- Accuracy Determination
  - Identifies environmental factors that are most influential in AI and Human mine-identification accuracy
  - Regression analysis, decision trees, random forests, neural networks
- UAV Routing
  - Optimizes UAV scanning path through full and partial observation of minefield
  - Formulation of a Markov Decision Process for use in a Reinforcement Learning model, Deep Q-Networks, Actor-Critic Methods
- UGV Routing
  - Optimizes UGV routing to minimize distance travelled and time spent
  - Routed based on path of UAV, Q-Learning, shortest path algorithms

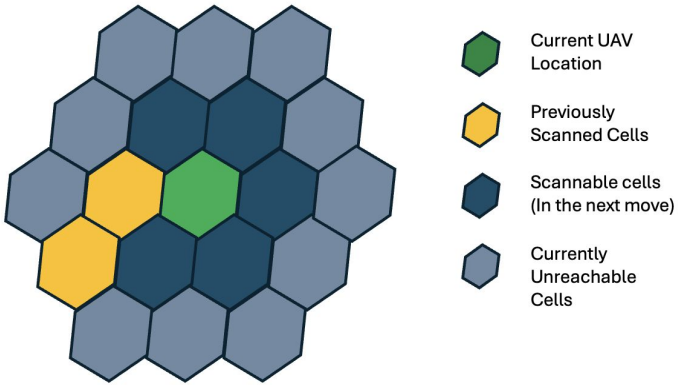
Accuracy Determination

- Analysis of environmental data to determine influential factors in Human and AI mine-detection accuracy
- Regression models can determine the marginal changes in accuracy due to changes in environmental factors
    - Struggles with non-linear relationships, high-dimensional data
- Decision trees, random forests, and neural networks can also identify influential factors
    - Perform better than regression models with large, complex datasets



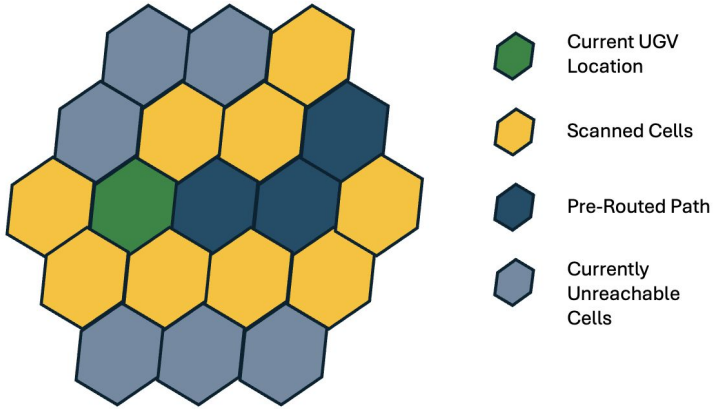
UAV Routing

- Optimization of UAV pathing to avoid unnecessary time costs and mine encounters
- A Markov Decision Process (MDP) can be developed for use in a Reinforcement Learning Model
  - MDP includes states, actions, and rewards
  - RL model finds optimal policy for MDP
- Partial and full map routing calculations can be synthesized
  - A full observation of the map allows for start to finish routing
  - A partial observation of the map allows for cell-to-cell routing based on adjacent cell conditions



UGV Routing

- Optimization of UGV pathing to travel along the shortest path
- UGV is treated as a secondary router; it follows the path determined by the UAV routing system
  - UGV could follow the path scanned by the UAV with the lowest probability of encountering a mine
- UGV utilizes a shortest-path algorithm such as A* to navigate along the shortest and lowest risk route



Future Considerations

- Scalability of system solutions to larger operations
  - Could future scenarios include multi-UAV or multi-UGV scenarios
  - Multiple warfighter battalions crossing the same minefield at once
    - Sharing of information between battalions, devices
- Identification of mines while troops/UGV have partially traversed a cell
  - Would it be faster to turn back around and take a different path?
- Physical limitations of UAV and UGV systems
  - In future stages, the UAV and UGV may not have infinite power supplies or ensured removal
- Effect on human and AI accuracy when images are sent over a WAN
  - A strategy to send some pictures back over a WAN before the UAV returns to the Command Center could improve traversal time

Future Plans and Actions

- Design regression and RL models
  - Additional data in the next stage will allow for more robust analysis of how system and AI behavior changes in different scenarios
- Perform testing and simulation to identify solution limitations
- Analyze how different algorithms perform on different subsets of data
- Test how solutions perform when the underlying situation changes