SYSTEMS ENGINEERING RESEARCH CENTER

UNIVERSITY of VIRGINIA

Human-Machine Team (HMT) Concepts for Resilient Autonomous Systems

Sponsor: DASD(SE)

By

Drs. Inki Kim, Barry Horowitz, Peter Beling, Stephen Adams

10th Annual SERC Sponsor Research Review

November 8, 2018

FHI 360 CONFERENCE CENTER

1825 Connecticut Avenue NW, 8th Floor

Washington, DC 20009

www.sercuarc.org

10 YEARS of SERC

Why Human-Machine Team (HMT) Concepts Matter for Cyberattack Resilient Systems?

- To date, engineering of Unmanned Aerial Vehicle (UAV) systems has not actively incorporated major concerns of cyber security [1]:
  - Cyberattacks may directly inflict permanent damages to the operational capacity (OC) of system components
  - Cyberattacks may indirectly impact the system via the attacks propagating in coupled subsystems [2]
  - Cyberattacks can significantly disrupt the mission, as well as the system resources
- Need to develop a resilient HMT concept for robust UAV systems
  - In general, a resilient solution is less prone to the possible violations of the assumptions, requirements, or rules that could influence its design and deployment [3]

Prior Findings by UVA Research Team

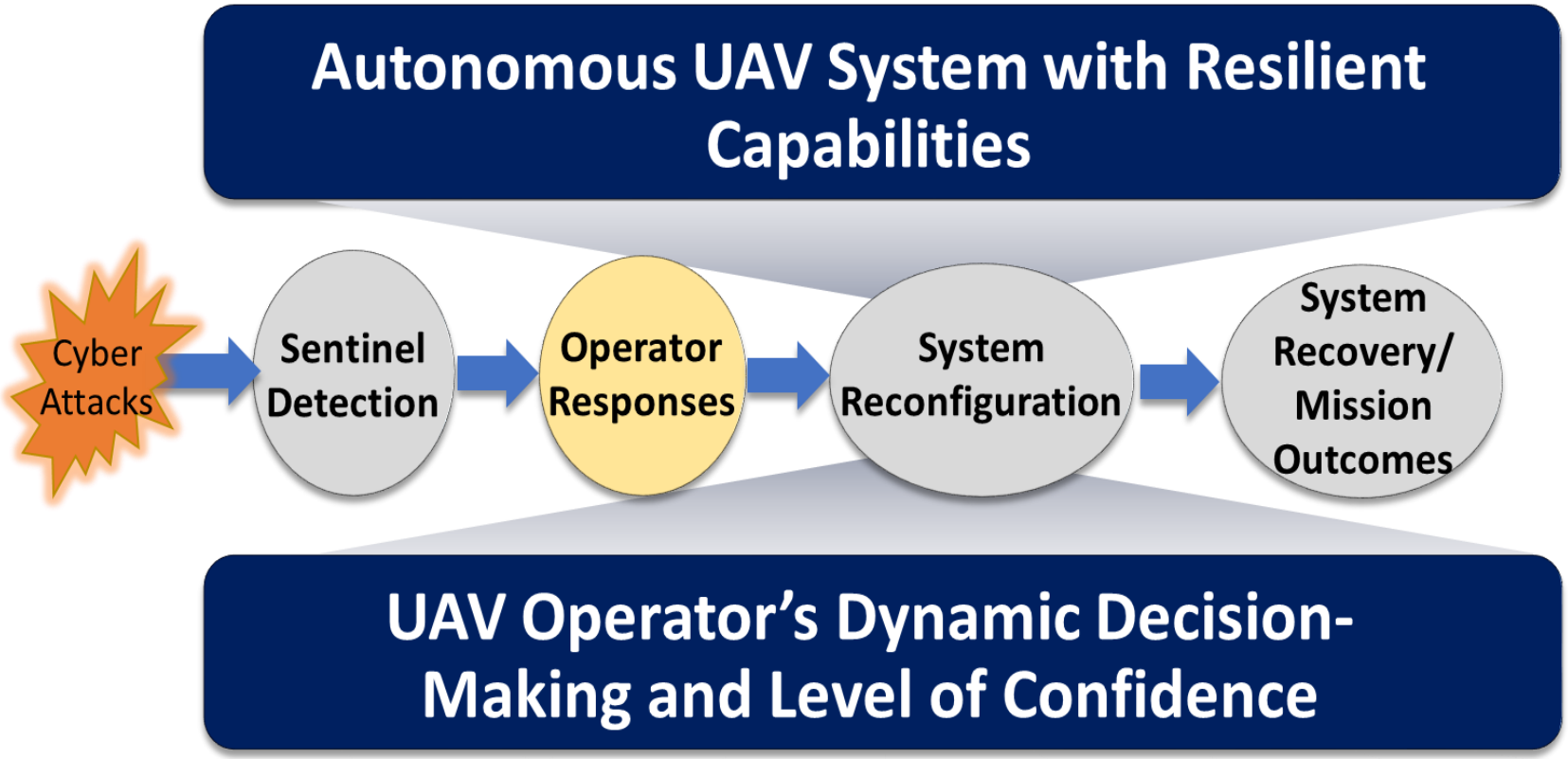
- Both technology and human factors are crucial to providing cyberattack resilient systems.
  - —Resiliency solutions require adjusting both the technical and operational configuration of the attacked system, thereby requiring operators "in-the-loop".
  - —It is necessary to investigate the human dimensions of decision-making under the uncertainty of cyberattacks and subsequent recovery.
- Without dedicated orientation, operators of physical systems are not capable of responding in a timely manner, or dependably choose resilience solutions upon a detected cyberattack [4].
  - —Operator's non-familiarity with:
    - cyberattacks and their potential consequences
    - the technologies for providing resilience and their expected performance
    - the possibilities for cyber attacking adversaries to respond in near-real time to resilience solutions

Research Challenge – 1

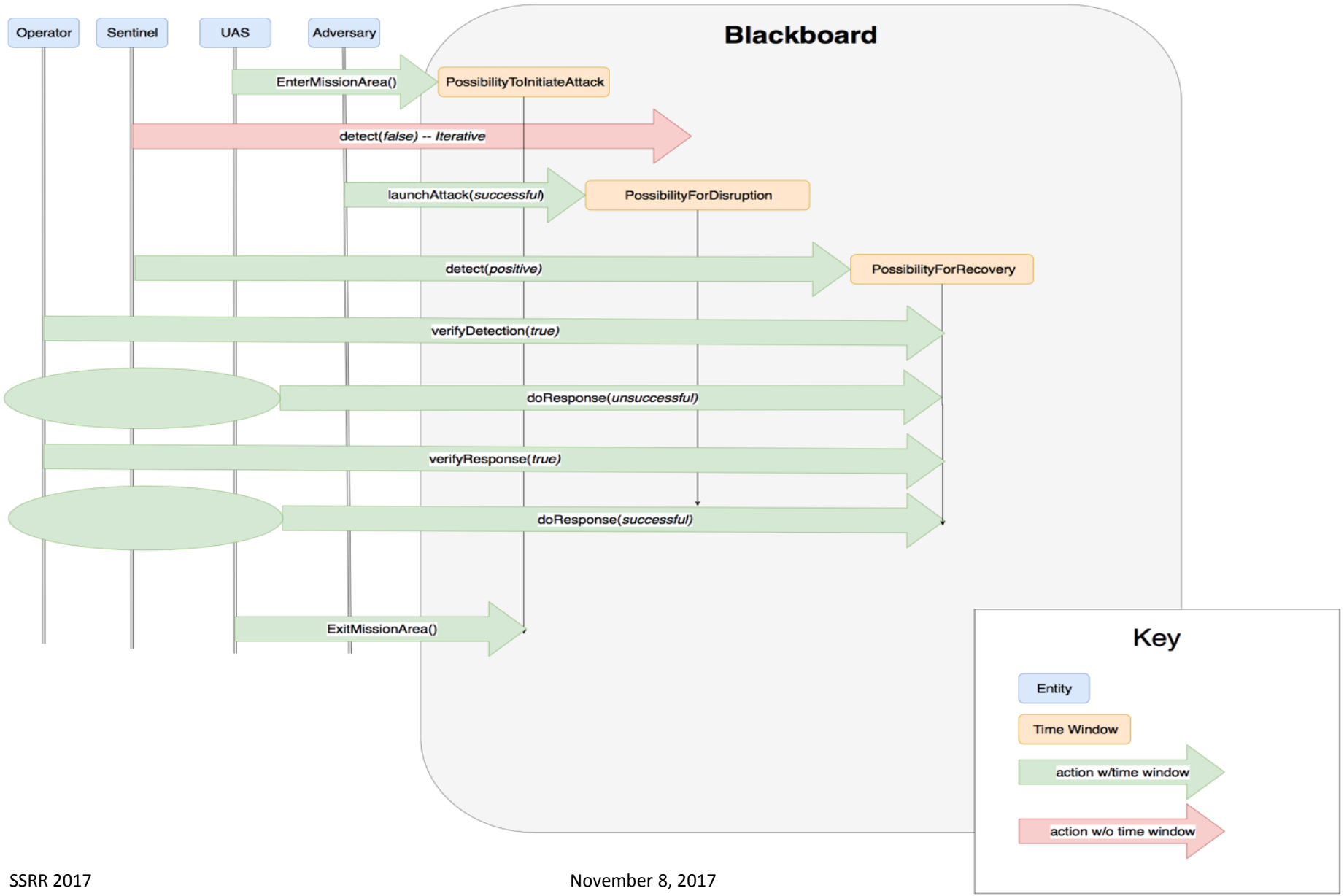
- **Representation and evaluation of HMT performance for cyberattack resilient systems**

- For resilient HMT solutions, a reconfigurable system architecture needs be flexibly distributed between operators and autonomous agents in response to the mission context [5].
  - —This "adjustable" autonomy necessitates a sequence of processes to represent, measure, distribute, and evaluate performance in human-machine team [6].
  - —HMT performance must consider the effectiveness and efficiency of collaborative responses of both human and machine under specific mission contexts
  - —In classic psychology, a situation and its constraints are not considered an essential element of cognition as they are in Information-Processing theories [7].

Scope of Research



An Illustrative Framework of HMT Interaction in a UAV System



Research Challenge – 2

- **Improvement of the military personnel's readiness to manage resiliency against potential cyberattacks**

- Lack of knowledge about the sources of variability in operators' decision making under dynamically-changing environments when rare instance of cyberattacks occur.
  - A vast body of literature has attempted to explain this variability, with emphases on behaviors [8], cognition [9], [10], perception [11], or adaptation [12].
  - Development of orientation methods, training programs, and education curriculum are much needed for the next-generation workforce development, as supervisors of autonomous systems.

Research Challenge – 3

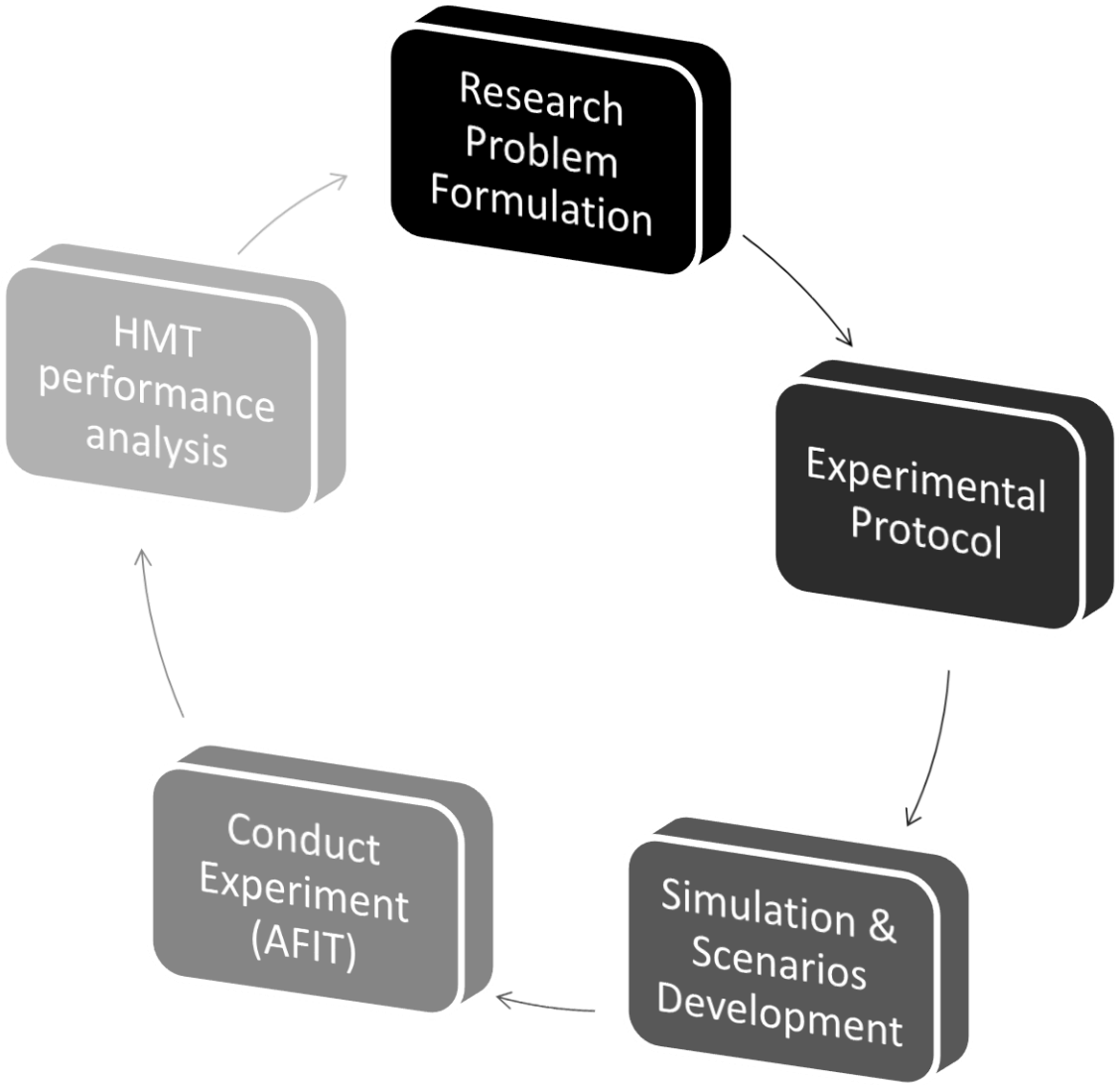
- **A formal decision model and experimental methodology that help explain dynamic tradeoffs among quality, time cost, and confidence in the context of the mission being sustained**
  - —A decision response is considered an outcome of interdependence among human, machine, and situational contexts, rather than human alone.
  - —A resiliency solution to disruptive events can be represented on a problem space, in terms of time (i.e., how long it is expected to take for the solution to handle the problem) and performance (i.e., the expected quality of the system after recovery, compared to the one in normal operations).
  - — In settling down for a solution, the pilot and automation agent should collaboratively handle the uncertainty associated with the solution.

Tradeoffs in Problem Space

Research Objectives

- This research aims to examine resilient human-machine cooperation under suboptimal systems operation due to malicious cyberattacks on military assets.
- **Objective 1**: Propose a framework to represent and evaluate the human-machine team (HMT) performance for cyberattack resilient systems
- **Objective 2**: Provide a training guideline to improve the military personnel's readiness to manage resiliency against potential cyberattacks
- **Objective 3**: Develop a balanced decision model that considers dynamic tradeoffs among quality, time, cost, and confidence in the context of the mission being sustained.

Research Process



Experimental Scenario Development – Mission Operation

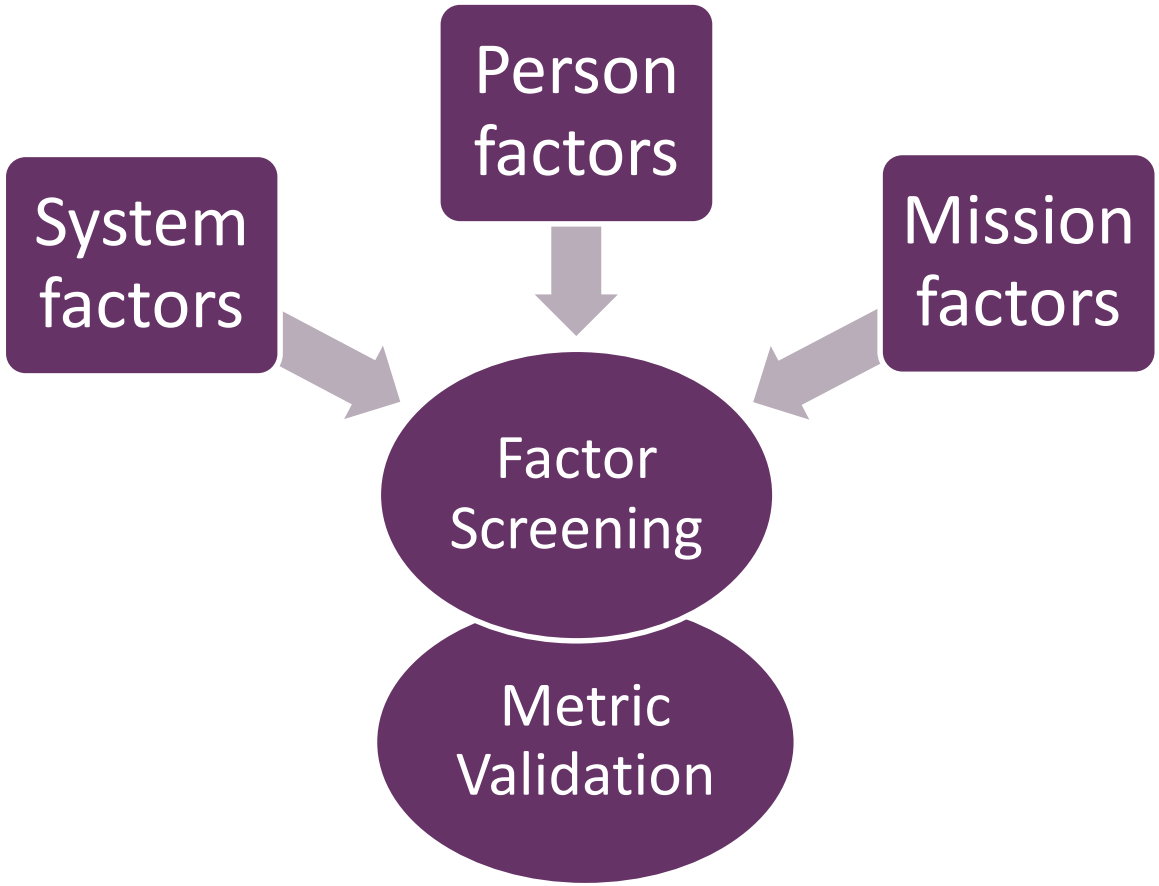
- A single pilot conducts a post-battle damage assessment surveillance mission over a broad area of enemy land where ground troops likely suffered significant injuries.
- Rapid mission completion is critical to identify the location of the casualties and call for medical support.
- The mission uses the group-two, fixed-wing UAVs (typically 20-50 pounds), with each vehicle operational for up to 1-hour surveillance mission.
- The pilot is responsible for planning, execution and modification of routes to inspect the region effectively and efficiently.

Experimental Scenario Development – Cyberattacks

- The adversary may have implanted Trojan horse into UAV control software, and the Trojan surreptitiously alters waypoints.
- Simultaneous attacks may be launched on the GCS display to conceal the fact that the UAV is going off course.
- In response to enemy cyberattacks, Sentinel detects the inconsistency of control data within the system, alerts system damages after the initial attack, and provides a set of available resiliency options for the pilot to choose from while launching an initial automated response.
- The system might autonomously recover itself to some degree, but it remains for the pilot to decide on the more complete follow-on response, which can include options beyond those suggested by the Sentinel.

Experimental Plan Overview

Phase-I Experiment



Phase-II Experiment

Identification of Person Factors

- These variables focus on individual differences in ability, awareness, and traits that are related to resilient behaviors.
  - (Mission/ Cyber) Situation awareness: measures the degree to which the subject is aware of current system capabilities and constraints within the ongoing mission context and the disrupting cyberattacks.
  - Resiliency scheme: measures the acquired ability to understand causal relations between a disruptive event, recovery options, and resiliency. The scheme can be learned, but also subject to individual traits.
  - Individual traits: measures self-confidence, attentional resources, mental workload, propensity to trust/ suspicion, and personality.

Identification of Mission Factors – Prior to Cyberattacks

- These factors constitute mission context for experimental scenarios.
  - **Mission location:** For a surveillance mission, the distribution of open vs. shielded geographic areas can impact the difficulty of mission execution.
  - **Home location**: For fixed-wing drones, a longer distance from home can limit mission time as well as slows the response option of calling-in a new vehicle.
  - **Target density**: Estimate of the total number of potentially injured target troops in the given search area.
  - **Mission progress (I)**: The percentage of scanned area out of the entire mission location, to be displayed on and after cyberattack detection.

Identification of Mission Factors – After Cyberattacks

- These factors also constitute mission context for experimental scenarios.
  - **Initial history report of cyberattack:** On detection, Sentinel will report the number of historic cases of waypoint manipulating cyberattacks. Although incomplete, this report may help the pilot gauge enemy doctrine and the risk of the imminent attack.
  - **Post-attack system function assessment**: Sentinel will identify subsystem(s) that were attacked
  - **Occurrence of cyberattacks:** In order to include scenarios with Sentinel's false alarm, a certain portion of scenarios will involve no cyberattacks but will include Sentinel false alerts.
  - **Occurrence of Sentinel detection**: In order to generate scenarios with Sentinel's missed detections, a certain portion of scenarios will involve no Sentinel alert.

Identification of System Factors

- **Battery power/ time-of-flight**: For the pilots, mission planning and response selection heavily depend on remaining battery level
- **Number of UAVs engaged in the mission**: The surveillance mission can be conducted by a single UAV operation, or a two-UAVs in a swarm.

Formulating Resiliency Solution

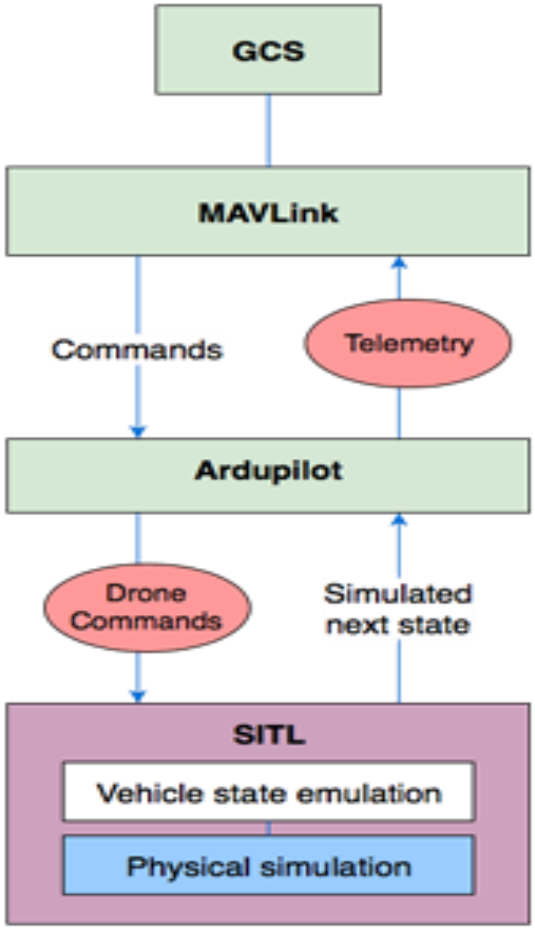
- The pilot may choose to reconfigure only the damaged sub-system(s), or to exploit diverse redundancy in a broader scope.
- To some response options, Sentinel will provide computational prediction of quality and time-to-recovery
  - —for quality, the percentage of system functions to be fully capable after recovery.
  - —for time, the predicted time to recover in minutes.
  - —In addition, prediction errors will follow the prediction value (±), to represent inherent uncertainty. This prediction error may impact the pilot's confidence in selecting a response.
- End-to-end test and verification actions: Test and verification of the resiliency solution(s) is not mandatory, but helps dispel doubts in the solution, and thus, the pilot can become more confident.

Scope of Resiliency Solution

System damage	Elementary Resiliency Options	Reconfiguration Scope
Navigation module	- Replace the navigation module - As an added precaution, replace the internal UAV communications to a backup communication channel between the navigation and flight control sub-system - (Automated Response) Reconfigure the control display to permit presentation of the updated trajectories	Damage-specific recovery options
Guidance module	- Recover waypoints (reset the navigation software and upload the original flight plan) - Use alternate waypoints (reset the navigation software and upload an alternate flight plan) - Switch to a manual guidance - (Automated Response) Reconfigure the control display to permit presentation of the updated trajectories	Damage-specific recovery options
Others	- Land the attacked UAV(s) - Recover the UAV that the Sentinel indicated as attacked - Call in a new UAV from home	Broader resiliency options
	- Continue mission while ignoring Sentinel alert	No resiliency options taken
	- Abort mission and go home with all UAVs	

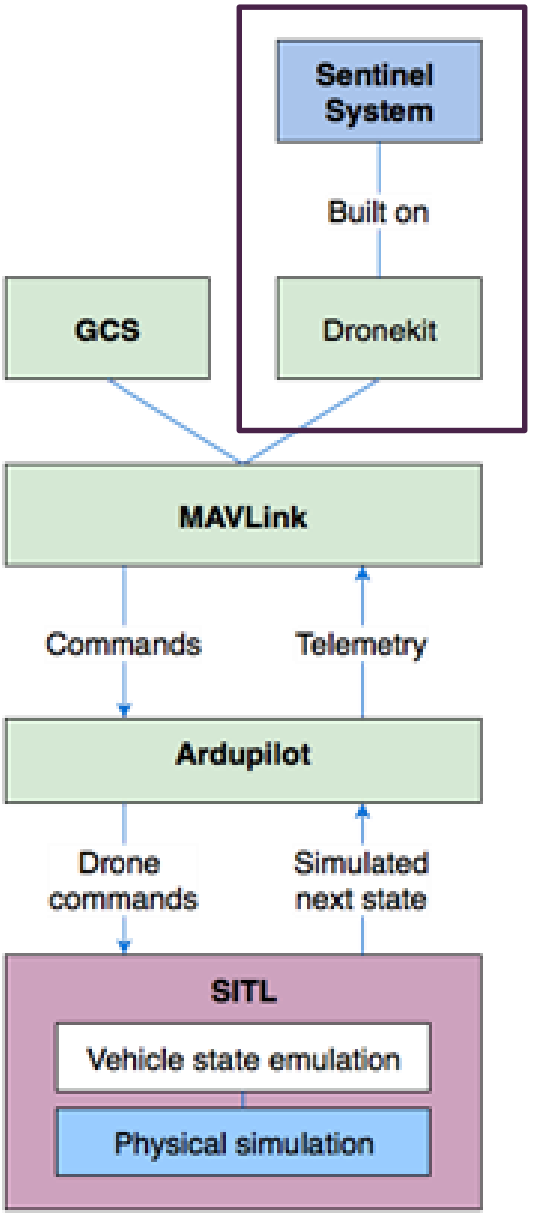
Simulation System Construction

- Software-in-the-loop Simulation Control Flow



Sentinel Design for Simulation

- For this project, a sentinel subsystem is assumed to be implemented on a separate computer system, separate from the regular ground control system.
  - —The sentinel system will be built on top of DroneKit, a drone interface tool built with the Python programming language.
  - —The sentinel system will determine the likelihood of an ongoing cyberattack from the telemetry and return its conclusion to the user in real time along with its recommended course of action.



UAV Ground Control System

- The overall idea of the simulation system is a cyber-physical view of the UAV system that accommodates interaction between cyber (command, control, and communication) and physical (sensor, actuators) components within the system.

Modes	Description
MANUAL	Manual control surface movement, pass-through
AUTO	Follow missions
LOITER	Circles point where mode switched
CIRCLE	Gently turns aircraft
GUIDED	Circles user defined point from GCS
Return to Launch (RTL)	Return to and circle home or rally point

UAV Ground Control Interface

Design of Experiment

- The phase-I experiment will adopt a 2-level Plackett-Burman design with 10 factors and 12 replicates (scenarios).

Scenario	Block	Mission location	Home location	Target density	Mission progress	History	Damage assessment	Cyberattack	Sentinel	UAVs	Battery power
1	1	open	near	dense	high	not provided	local damage	Attack	Alert	1	high
2		open	distant	sparse	low	provided	global damage	Attack	Alert	1	high
3		open	near	sparse	low	provided	local damage	No attack	No alert	1	low
4	2	open	distant	dense	low	not provided	local damage	No attack	No alert	2	high
5		open	distant	dense	high	provided	global damage	Attack	No alert	2	low
6		open	near	sparse	high	not provided	global damage	No attack	Alert	2	low
7	3	shielded	near	dense	high	provided	global damage	No attack	No alert	1	high
8		shielded	distant	sparse	high	not provided	local damage	Attack	No alert	1	low
9		shielded	distant	dense	low	not provided	global damage	No attack	Alert	1	low
10	4	shielded	near	dense	low	provided	local damage	Attack	Alert	2	low
11		shielded	distant	sparse	high	provided	local damage	No attack	Alert	2	high
12		shielded	near	sparse	low	not provided	global damage	Attack	No alert	2	high