TAS Node in Security (TAS-S)

Observations on AV Models & Realities

Prof. Neeraj Suri, Project PI

July 2022

This work is supported by the Engineering and Physical Sciences Research Council [grant number EP/V026763/1]

Safety/Security

1. Specify a system model
2. Specify a threats model
3. Specify associated assumptions
4. Assert a consequent property (as a compromise of the models and/or assumptions)

Assumptions for Validity:

- That our models are accurate and complete
- That our assumptions are accurate and complete
- That the environment and attackers behave as modeled!
- That the assertions are accurate and complete!

Expectations: Correct, Reproducible, Verifiable, …

Models: Approximations of Reality

- System Model (Computing/Communication)
  - Structured relationships across the elements
  - Delineated control and data flows
- Threat Model (over a characterized attack surface)
  - Specific scenarios: type, #, duration, collusion ...
  - Worst case (UU) scenarios modeled via weak assumptions

Needs: Accuracy, Completeness

Approach: Refine & validate model post-deployment

Expectancy: Deliver outcomes that are predictable, reproducible & can be validated

AV Models

Context: Multi-AV + Full Autonomy Scenarios

AI based PCD (**P**erception, **C**ognition, **D**ecisions) Chain

- System Model
    1. Computing and Communications Model (underpins PCD ops)
        - VVV/Unstructured: Dynamic/variable SoS relationships
    2. AI/ML Model (underpins CD ops)
        - Dynamic/variable SoS data streams
    3. Socio-technical Users/Usage Model
        - Users as active elements: behavior/response/ethics/liability models
        - Environment model (urban, terrestrial, aerial…)
- Threat Model (Variable attack surfaces)
    - TM for computing infrastructure
    - TM for AI/ML infrastructure
    - Complex inter-linked System/ML TM's over SoS AV Environment
        - Limited separation of concerns across system/ML control & data flows
        - Infinite cases over unanticipated socio-technical scenarios
        - Worst cases approach of weak assumptions work in a limited manner

Issues: AV Specification and Models

- CD critically depends on accuracy & completeness of models
  - CD (Cognition and Decision) functionality developed on (incomplete) pre-deployment training models… limited validation
  - Reality: Form, adapt and solve SoS AS models at run time – validation?
  - AI/ML is accuracy and performance driven → discards rare/outlier cases. Rare cases are THE AV cases to worry about. Run-time determination?
- Safety-critical CD decisions based (a) on run-time models, (b) on dynamic & incomplete data-streams, (c) using non-deterministic AI/ML technology to deliver deterministic safety/security outcomes in unstructured AV space?

Security-by-design without a complete, accurate, stable model? Engg solutions are nice but …