TAS-S
ANNUAL REPORT 2020-2021

UKRI AUTONOMOUS SYSTEMS NODE IN SECURITY

Foreward

The UKRI Trustworthy Autonomous Systems Node in Security (TAS-S) constitutes an exciting opportunity for collaborative research that is additionally supported by our unique security and Autonomous Systems (AS) test beds. The project has extensive stakeholder support, both domestic and international, from academics to AS providers, AS users and AS regulators.

Further details on all of the sections in this report can be found on the TAS-S website.

I am delighted to present the first TAS-S annual report.

Despite the challenges posed by the COVID-19 pandemic, this has been a successful and busy 12 months for the project. Exciting results are already emerging from our three research strands and we are developing valuable opportunities to further expand our community of researchers and stakeholders from a wide range of industries. The latest updates on our research activities are included this report..

Professor Neeraj Suri,
Principal Investigator, TAS-S
Lancaster University

In addition, we are taking an innovative approach towards ethical considerations for research with our Ethical, Legal and Social Issues (ELSI) Framework for the Security of Autonomous Systems. This framework has been developed by Research Strand 3 and is being incorporated into different aspects of our work across all research strands.. We are also working with our stakeholders to explore the most effective ways to apply this to industry. Further details about ELSI can be found in the Research Strand 3 update.

This report details our activities from November 2020-November 2021 and I am very much looking forward to working with colleagues across the project to develop our activities further in 2022.

Contents

Introduction

Autonomous Systems (AS) can be broadly categorised as the ability to effectively conduct a mission with varied levels of “absence of human intervention” including completely unsupervised operations. Typical examples, spanning an ever growing diversity of civilian, industrial and military applications across terrestrial, aerial and aquatic environments include autonomous vehicles, industrial automation, assisted living and a variety of logistical support to complement and supplement societal needs.

As technologically complex and networked cyber-physical entities, an AS needs to ensure “safe and secure” mission functionality despite the occurrence of any encountered cyber-physical disruptions. As such, an AS is a highly-dynamic entity that needs to adapt to the vagaries of its operational environments and security profiles (including changing threats). Providing “predictable, scalable and composable” security (of the AS assets, of the AS operations and the AS usage environment) in “uncontrolled and dynamic” operational environments is the objective of TAS-S.

The TAS Security Node’s research is centred around a seamless collaboration between fundamental cross-disciplinary security research and autonomous systems research at Lancaster and Cranfield Universities. To accomplish this vision, TAS-S utilizes interlinked cross disciplinary Research Strands (RS) to address 3 core challenge areas in autonomous system (AS) security:

Research Strand 1:
Securing the AS "usage" environment

Research Strand 2:
Can we secure the AS "operations" environment?

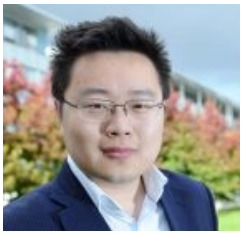
Research Strand 3:
Can we secure the AS "user" environment?

TAS-S Team

TAS-S assembles a cross-disciplinary team of internationally reputed security experts from Lancaster and Cranfield Universities who are based across a wide range of research areas including Distributed Systems, Controls, AI, Communications, Sociology and Law.

Research Strand Leads and Project Manager

Prof. Neeraj Suri,
PI, RS1 Lead
Lancaster University



Prof. Weisi Guo
Co-I, RS2 Lead
Cranfield University

Prof. Corinne May-Chahal
Co-I, RS3 Lead
Lancaster University

Pamela Forster
Project Manager
Lancaster University

Co-Is, Lancaster University

Prof. Plamen Angelov
Prof. David Hutchison
Dr. Daniel Prince
Dr. Joe Deville
Dr. Catherine Easton

Co-Is, Cranfield University

Prof, Gokhan Inalhan
Prof. Antonios Tsourdos
Dr. Lisa Dorn

Postdoctoral Researchers, Lancaster University

Dr. Zhengxin Yu
Dr. Andrew Sogokon
Dr. Luke Moffat
Eduardo Almeida Soares
Pierre Ciholas

Postdoctoral Researchers, Cranfield University

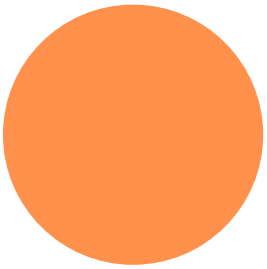
Dr. Burak Yuksek
Dr. Zhuangkun Wei
Dr. Oscar Gonzalez Villarreal

Governance

The node has an agile management structure to provide (a) efficient and responsive internal project management and (b) engagement with TAS Hub/nodes and external stakeholders. Further details can be found on the following pages or on the TAS-S website.

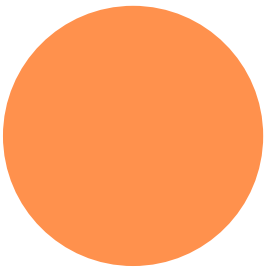
Project Management

The PI, Co-Is and Project Manager meet regularly as part of the Node's scheduled monthly "Research Activity Group (RAG)" and "Coordination Group (COG)" meetings. These groups have oversight of the day-to-day running of the project, plan engagement activities and events, and monitor progress with respect to the objectives, research outcomes and emerging risks.



Research Management

Each research strand has its own meeting structure, through which updates from each theme (2 or 3 per research strand) are discussed and opportunities for further collaborative work are identified. The Postdoctoral researchers from both Lancaster and Cranfield meet online every two weeks to present particular aspects of their research, suggest reading from the wider field, and discuss ways to work cross-research theme/cross-institution.



Governance (cont.)

Strategic Approach

The Advisory Group provides strategic advice and feedback on the project's research approaches, progress, quality and activity development. The Group has recently welcomed 2 additional members and now consists of 6 distinguished external stakeholders from academia and industry across the UK, Europe and the US. We are also in regular contact with the Hub Liaison Team to ensure that our Node works closely with the TAS Hub. Further details can be seen below:

Advisory Goup

Prof. Robin Bloomfield
Adelard

Prof. Phil Koopman
Carnegie Mellon University.

Dr. Hector Figueiredo
QinetiQ

Dr. Carl Segueira
Flarebright

Dr. Arthur van der Wees
Arthur's Legal

Prof. Carl Landwehr
Center for Democracy & Technology, University of Michigan

Hub Liaison

Prof. Luca Vigano, King's College London
Prof. Derek McAuley, University of Nottingham
Prof. Jose Such, King's College London

Testbeds

Lancaster and Cranfield Universities are home to specilalist testbed facilities, including a unique autonomous systems test facility for combined air-ground vehicles at Cranfield. Further details about our facilities can be found on the dedicated Testbed page on the TAS-S website.

Autonomous Systems Protocol Testbed

A testbed to replay, simulate, and benchmark autonomous systems communications supporting dynamic topologies, network state fluctuations, and highly scalable (2 to 10k instances).

Cyber Threat Laboratory

A partnership between Security Lancaster and Fujitsu Enterprise and Cyber Security, the lab is a collaborative platform that allows analysis of threats and behaviour to take place in a safe and controlled environment.

Lancashire Cyber Foundry

A series of multi-million pound secure digitalisation projects to help SMEs across Lancashire embrace innovations in digital and cyber technologies to defend, innovate and grow their business.

innovative Digital Infrastructure Defence (iDID)

iDID addresses pragmatic industrial requirements using applied research methods and focusses on cyber security and cyber threat intelligence for internet-enabled cyber physical systems.

Testbeds (cont.)

Multiuser Environment for Autonomous Vehicle Innovation (MUEAVI)

This outdoor test facility supports the rapid development of on and off highway. ground and airborne autonomous vehicles. The facility includes sensors with 4G/5G connectivity sensors, along with Lidar object recognition/test drone tracking on both LOS/NLOS basis.

National Digital Aviation Research and Technology Centre (DARTeC)

DARTeC is a £65 million facility integrating research and practice. It includes a fully functional airport, digital control tower, and air space control to offer a unique research and development environment.

National Beyond Visual Line of Sight Experimentation Corridor (NBEC)

NBEC provides a safe, managed environment to test and develop concepts , principles and related technologies to enable flying unmanned aircraft systems beyond visual line of sight in non-segregated airspace.

Website and Social Media

TAS-S website

We have developed a dedicated website which details all the different aspects of the Node including:

- Node overview
- TAS-S Team
- Advisory Group/Hub Liaison
- Stakeholders
- Research Nodes
- Testbeds
- Publications

We have also taken this a step further by developing the website as a resource not just for our Node but for the whole TAS network and the wider field. Therefore, it is regularly updated with details regarding the following:

- Seminars from TAS-S, Lancaster University, Cranfield University, and the other Nodes.
- Events from the TAS Hub and the wider field.
- News which is of interest to the Autonomous Systems Community.
- Blogs from our Project Manager and Researchers regarding the Node's latest activities and research.

A small sample of our current news posts and blogs can be seen on the right-hand side of the page. Find out more at https://tas-security.lancs.ac.uk/

Meeting the TAS-S team!

Project Manager, Pam Forster, explains the behind-the-scenes work that went into making this first face-to-face event a success.

READ THE BLOG HERE

UKRI Trustworthy Autonomous Systems Hub

Join the TAS Hub!

Vacancy for a Project Manager to join the TAS Hub team in Southampton. Closing date, 18th February 2022.

Law Commission (26th Jan 2022)

The final report on Autonomous Vehicles is now available to read

Twitter and LinkedIn

Our social media accounts contain a huge range of information about upcoming seminars and events, latest job opportunities and news from across the TAS network and wider field.

Engagement

Despite the complications caused by the COVID-19 pandemic, the Node has organised and taken part in a range of online and hybrid events. These have included workshops and seminars, as well as engagement in the TAS Hub and other Nodes' initiatives. Further details can be found on the Node's website.

Event name/type	Description and impact
External Stakeholders' Group Workshop (ESG) 29th March 2021	• This full-day event provided the Node with a critical opportunity to network with around 50 TAS-S stakeholders, an international representation of prominent autonomous systems leaders from 30 academic, governmental and industrial organizations. • The second ESG workshop is scheduled for 1st March 2022.
ECR/Postdoctoral Researchers' Workshop 24th/25th November 2021	• This event provided an excellent opportunity for our postdoctoral researchers from Lancaster and Cranfield Universities to meet each other face-to-face for the first time to discuss their research and explore collaboration opportunities. • In addition, it enabled our researchers to network with PhD students and colleagues from Lancaster University's Security Institute. • The second Researchers' workshop will held in Cranfield in Spring 2022.
ELSI Workshop May 2021	• This ELSI (Ethical, Legal and Social Issues) workshop was facilitated by RS3 and was attended by around 20 external participants who explored potential ethical issues around developing drones.

Engagement (cont.)

Event name/type	Description and impact
Stakeholder Seminar Series (#TASSTalks) August-December 2021	• These online seminars are presented by our external stakeholders to capture their requirements, experiences and challenges. Around 25-30 external participants have attended the following talks: • 19/08/2021: "*Towards Safe, Trustworthy and Efficient Autonomous Vehicles*", Dr. Dezong Zhao, University of Glasgow. • 10/09/2021: '*Security by Design for IOT and Automotive*', Dr. Willibald Krenn, Austrian Institute of Technology • 24/09/2021: '*Secure-by-Design – the challenge of moving beyond Cyber Risk Management to Cyber Resiliency*', Dr. Alex Tarter, Thales • 22/10/2021: '*Safe Autonomous Systems: Challenges and Potential Solutions*', Wilfried Steiner, TTTech Labs • 12/11/2021: '*Security challenges for collaborative autonomous aircraft systems*', Dr. Cora-Lisa Perner, Airbus. • 19/11/2021: '*Trustworthy Autonomy*", Spirent. • 10/12/2021: '*Industrial Perspectives of Artificial Intelligence',* Prof. Nick Colosimo, BAE Systems.
'Secure, ethical digital technologies: What role for social science'? 15th July 2021	• This workshop was facilitated by RS3 and was was aimed at a younger audience, primarily applicants to Lancaster University to engage them with our research.
Stakeholder workshops 15th July 2021	• Engagement with multiple current TAS-S stakeholders and additional externals for the exchange of AS data and collaborations. An initiative with the Lancashire Police on vehicle forensics has also started.

Engagement (cont.)

Event name/type	Description and impact
'Living With AI'' podcast series. 24th March 2021	• Prof. Suri took part in a podcast discussing the 'Challenges of Living with AI'.
'TAS All Hands Meeting' 14th-16th September 2021	• Prof. Suri produced a research video and took part in the Fireside Chat. • Prof. Guo was a panel member for a Mentorship session discussing experiences of applying for funding and Fellowship opportunities. • Research Associate. Dr. Moffat was an invited speaker at the Programme Workshop 'Specifying for Trustworthiness'. • RS2 and RS3 produced posters on '*A Rapid Evidence Review of Methods: Autonomous Vehicle Security and Human Behaviour*' and '*Secure Operations of Trustworthy Autonomous Systems*' respectively.
'Thought Pieces' Workshop 9th December 2021	• Prof. Suri, Prof. Guo and Prof. May-Chahal discussed our approaches to security in automated systems with members of the TAS Hub and industrial partners. This will form part of a white paper being produced by the TAS Hub.
Collaborations with other Nodes	• RS1 Theme A is developing a cooperative technical approach to specifying and verifying security compositions relevant to AS's (with the Verifiability Node). • Prof. Suri presented a seminar for the Resilience node. • A workshop on the theme "Surveillance" was held on 29th September (with the Resilience node).

Research

The interconnectivity of our three research strands is illustrated in the image below. In addition, each strand is split into specialist themes focusing on a dedicated area of research. Further information about each of the research strands and themes can be found on the following pages and on the research pages of the TAS-S website.

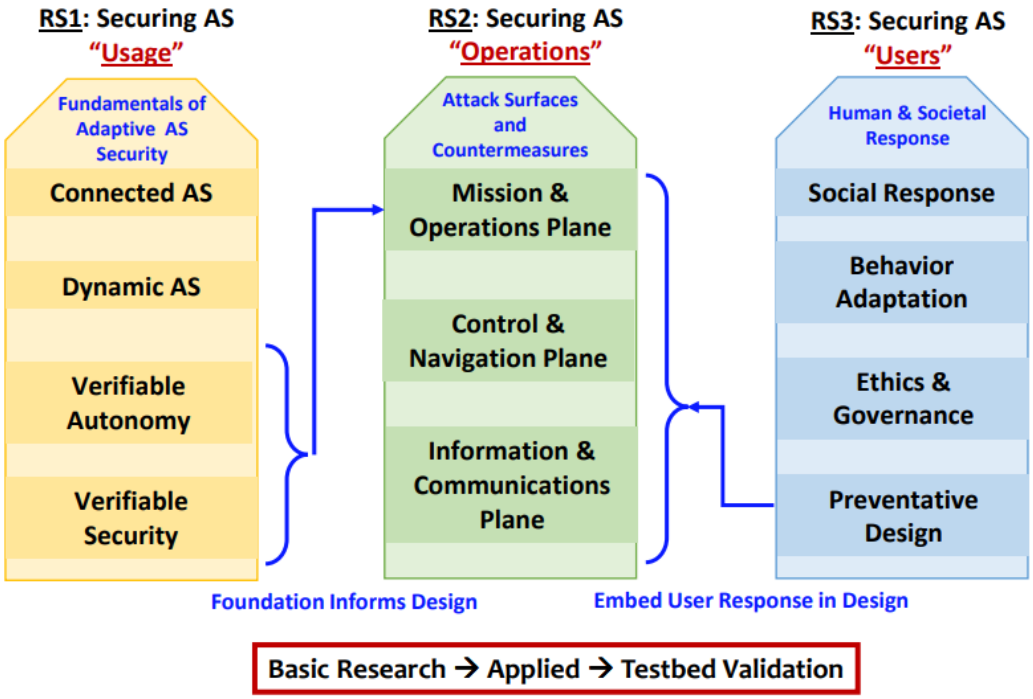


Fig. 1: The interconnectivity of the Research Nodes

RS1: Securing the AS "usage" environment.

To establish the fundamental AS "usage" framework for providing and assessing multi-layered, multi-dimensional adaptive AS security in dynamic mixed mode environments.

RS2: Can we secure the AS "Operations" environment?

To ascertain exposure (and their consequent mitigation) of AS "operations" to cyber-physical attacks by characterizing the attack surfaces (i.e. entry points and likelihoods) across the mission, control and information surfaces in a technology and mission-invariant manner.

RS3: Can we secure the AS "User" environment?

To ascertain the overall AS threats across multiple attacks, our approach tackles three interdependent AS surfaces (mission, control and communication), while the security foundations of RS1 and the human behaviour from RS3 are used to create holistic mitigation strategies.

RS1

RS1-Theme A: Dynamic and Compositional AS Security

Lead: N. Suri. Participants: A. Tsourdos, G. Inalhan, A. Sogokon.

The research addresses the fundamental challenges of specifying AS interfaces and the emergent security properties over compositions across AS and/or with the environment and especially adaptivity that characterizes AS operations.

Our intent is to develop a conceptual framework characterising the relationships across security attributes and the role of collaborative, disruptive and scalable security composition in AS, along with a run-time security policy framework for AS.

RS1-Theme B: Explainable and Verifiable Decision Making for AS Security

Lead: P. Angelov. Participants: N. Suri, W. Guo. G. Inalhan, Z. Yu, A. Sogokon, E. Almeida Soares.

Two research challenges will be addressed. First, the control behaviour of an AS is often non-deterministic as an AS adapts to changes in the operational environment, resources, sensory streams and objectives to yield an “optimal” solution. This nondeterminism makes verification of the security attributes unviable by classical testing and verification approaches that, typically, verify a given static property. This is a standalone challenge as autonomy, usually, results in non-deterministic outcomes unable to support reproducibility of scenarios and results. Second, AS operate on data streams from sensory inputs for analysing data related to the mission, situation awareness, the navigation, and control. This results in the use of data-driven reasoning techniques.

Our intent is to develop dynamic verification methodologies, explainable-by-design DL architectures that lend themselves to reasoning interpretation as well as to visualization, and symbolic surrogate models for DL-based automation reasoning techniques.

Research Activities (RS1A)

Dynamic and Compositional AS Security

Discrete AS's dynamically collate (join: aggregate & leave: disaggregate/ failures) into swarms of Networked AS's (NAS) to realize enhanced functionalities, e.g. the coordinated sensory coverage of a geo-distributed target or fleet capabilities in vehicles/drones. Consequently, the NAS structure and threats corresponding to the changing operational environment are also dynamic. The challenge is to understand how security attributes (e.g. CIA) compose as discrete AS's compose to form a NAS.

Research in Year 1

Our research aims to understand the fundaments of security composition. We focus on two aspects, namely, the "system of systems" NAS model and its security model. We are investigating two NAS models where the AS's are modeled and specified as either stateless or stateful artifacts. Following the specification of the discrete AS components, the interface linking them (protocol, communication) is specified to establish how the core AS functionality of PCD (Perception, Cognition & Decision) and corresponding security properties get sustained or disrupted. We are demarcating security into Liveness (Availability) and Safety (Access Control/Authorization and Integrity), where our initial focus is on Liveness.

Plan for Year 2

We plan to develop a formal model of "stateful" components. Unlike classical approaches of component-based software where, typically the complete specification of the software component functionality and their interconnections are assumed to be statically specifiable, the proposed model considers a) components to have hybrid cyber-physical state parameters such as time, location, movement in addition to the static functional specifications, and b) dynamic interactions where ASs can join, leave or get compromised by attacks. We plan to look at specification schema and corresponding run-time verification approaches specifically for incomplete specification NAS scenarios. A paper, tentatively entitled, "Security-Minded Verification of Cooperative Awareness Messages" is in development also involving members of the Verifiability node.

Research Activities (RS1B)

Explainable & Verifiable Decision Making for AS Security

Two research challenges will be addressed. Firstly, the control behaviour of an AS is often non-deterministic as an AS adapts to changes in the operational environment, resources, sensory streams and objectives to yield an “optimal” solution. This nondeterminism makes verification of the security attributes unviable by classical testing and verification approaches that, typically, verify a given static property. This is a standalone challenge as autonomy usually results in non-deterministic outcomes unable to support reproducibility of scenarios and results. Secondly, AS operate on data streams from sensory inputs for analyzing data related to the mission, situation awareness, the navigation, and control. This results in the use of data-driven reasoning techniques.

Research in Year 1

We studied methods which provide guarantees that can be established about the trustworthiness of the decision module of autonomous systems. In particular, formal guarantees of correctness related to safety or security are of interest and the methods that are being considered include Simplex Architecture for implementing the switching logic between vehicle controllers and in most cases does not come with formal guarantees, and vector Lyapunov functions in relation to neural networks. This work relates directly to proofs of stability properties which are important in cyber-physical systems (especially when dealing with large interconnected systems).

Plan for Year 2

A paper (tentatively) titled “Learning vector Lyapunov functions" to address safety-critical stability properties in autonomous systems using machine learning tools. Further development is proposed of the Simplex Architecture incorporating runtime monitoring and a decision module that offers formal correctness guarantees (resulting in trustworthy decisions) even in cases where the autonomous system relies on machine learned components for control and navigation. During the next year we will focus on the challenging problem of dynamic verification and validation of AI-driven control systems designed towards continual assurance. A part of this year’s effort will cover embedding features into the learning design process to ensure explainability of the learning enabled control designs.

Research Activities (RS1B)

Explainable & Verifiable Decision Making for AS Security

Plan for Year 2 (cont.)

Inter-linked with RS2A, we plan to propose an adaptive ML framework to enhance security and achieve resilience in AS. The proposed framework dynamically generates hierarchical loops (inner loops and outer loops) to learn ML models among ASs. Outer loops are used to build more generalizable ML systems. Inner loops are designed for dynamically matching new requirements from ASs and making predictions with the help of outer loops. This framework defragments ML processing architecture where hierarchical loops can have dynamic interactions when ML attacks occur. Even if some inner loops fail, other loops can still collaboratively train ML models. A paper entitled, “An Adaptive and Hierarchical Peer-to-Peer Federated Meta-Learning Framework” is in development.

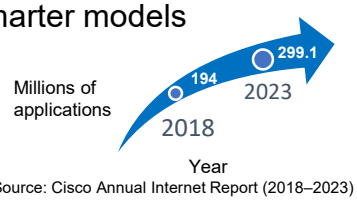
An Adaptive Federated Meta-Learning Framework for Autonomous Systems

Zhengxin Yu, Neeraj Suri (PI), Lancaster University

Background

Data is born at the edge

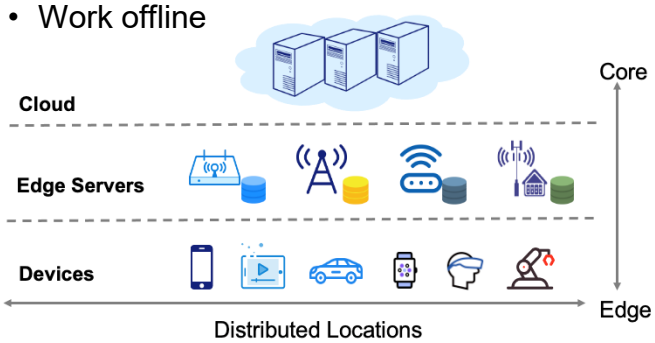
- Billions of smart devices generate data
- Data enables smarter models



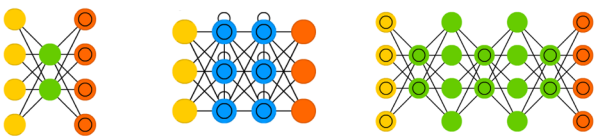
Source: Cisco Annual Internet Report (2018–2023)

Data processing is moving on edges

- Improved service latency
- Work offline



Emerging machine learning technologies

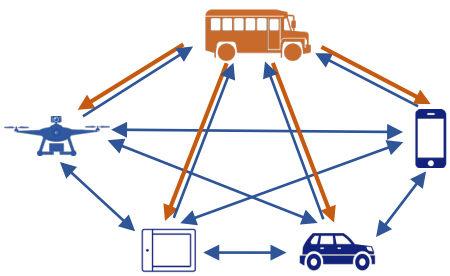


Challenges

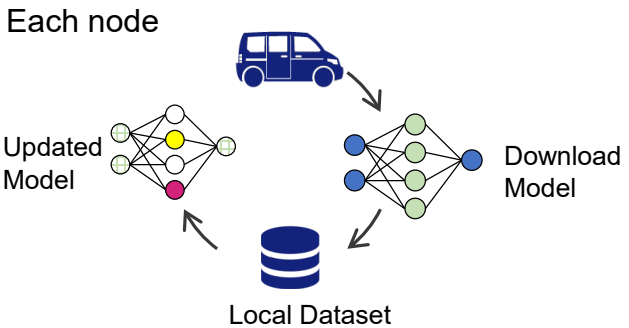
- High mobility of ASs
- Diversity of the AS environment
- Non-IID data distribution among ASs
- Security and privacy risks

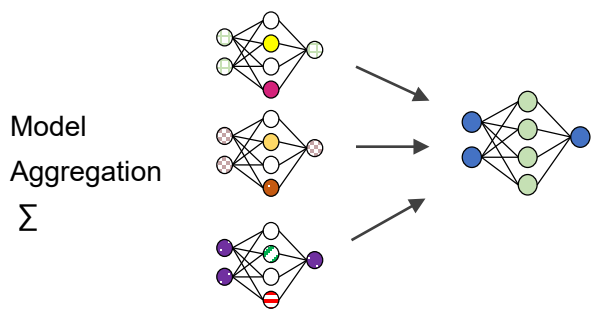
Methods

- **An adaptive meta-learning architecture** is proposed to adapt to new environment.
- A **peer-to-peer FL framework** is developed to reduce privacy risks and support mobility of ASs.
- An initial shared model and personalized models for ASs are trained in the framework.

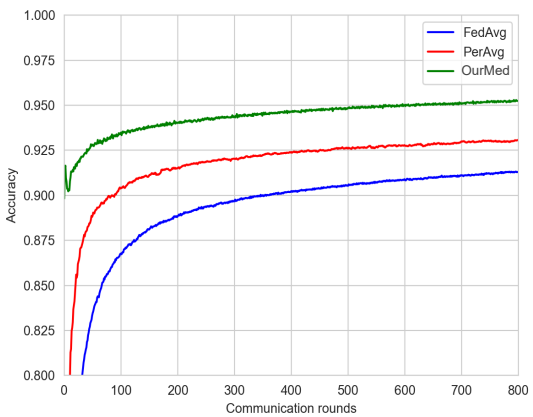


Model Training





Experimental Results



Acknowledgements

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

Lancaster University

Engineering and Physical Sciences Research Council

Cranfield University

TAS-S

RS2

RS2-Theme A: Security in the Mission and Operational Surface

Lead: P. Angelov. Participants: A. Tsourdos, Z. Yu, E. Alemida Soares, O. Gonzalez Villiarreal.

AS pose specific requirements and challenges to the detection and mitigation of cyber security risks and attacks due to their complexity and dynamic characteristics combined with the limited and unreliable network connectivity. The mission surface is the core, where the decisions and execution take place; it is dynamic and sensitive by its definition and verifiable security is of critical importance. This complicates the traditional approach that involves continual monitoring and update with patches, which links closely to the control surface below. We will develop methods and algorithms that reduce the risks and costs associated with these challenges and in turn, improve the reliability and resilience of AS.

RS2-Theme B: Securing the Control Surface

Lead: G. Inalhan. Participants: P. Angleov, A. Tsourdos, B. Yuksek.

AS relies on the ability to conduct run time adaptations of control decisions over attacks which can result from information and dynamic environment uncertainties. Specifically, in the context of learning enabled AS, it is crucial for the control system to exhibit self-aware learning in which the boundaries of “safe” state-space and the control space are tracked through their evolution. This is particularly challenging when the system is undertaking dynamic decisions within the AS mission surface.

RS2-Theme C: Securing the Cross-Layer Networking

Lead: W. Guo. Participants: D. Hutchison, D. Prince, P. Ciholas, Z. Yu, Z. Wei.

The research is divided between physical level attacks (Cranfield) and network level attacks (Lancaster). At the Physical (PHY) level, we know digital security can be derived from both antenna beamforming (codeless defence) [2C-1] and deriving distributed keys from channel state information (code-based defence) [2C-2]. The latter is particularly of interest as it can produce secure cipher keys without a common key pool or sharing keys. Yet, it must observe 3 conditions in the PHY channel, namely: (1) reciprocal to allow decentralised synchronous key generation, (2) dynamic to defence against brute force attacks, and (3) unique to avoid correlated attacks. The challenge is that the idealised conditions are often not met for ASs especially in open static spaces and airborne spaces.

Research Activities (RS2A)

Security in the Mission and Operational Surface

An Adaptive & Hierarchical Peer-to-Peer Federated Meta-Learning Framework.

A growing number of applications (i.e., autonomous vehicles, distributed sensing, etc) are supported by distributed systems where the spread of distributed users request varied data from disparate edge nodes. Such data has inherent characteristics, including non-IID distribution, heterogeneous with patterns of data, and limited data size that add considerable challenges to conduct accurate and efficient Machine Learning (ML) models. Meta-learning has been considered a promising approach to efficiently learn new tasks with few samples. However, in real-world, the high mobility of edge nodes and dynamic data pattern further increase the complexity of meta-learning. Moreover, gathering training data in a centralised way raises privacy and security risks.

Research in Year 1 and Plan for Year 2

We propose a layered ML approach, namely Peer-to-Peer Federated Meta-learning (PPFM) where the defragmented ML architecture adaptively matches the characteristics of the distributed heterogeneous data to enhance the efficiency and accuracy of ML models in a distributed environment, while reducing privacy and security risks. In the PPFM, hierarchical outer loops and inner loops are dynamically created based on diverse data features to adapt and generalise to unseen learning tasks or environments with limited exposure. Furthermore, the proposed peer-to-peer learning framework can ease reliance on the fixed central server and support such high-mobility scenarios. Figure 2 illustrates this approach.

Specifically, 1) A peer-to-peer federated meta-learning framework is proposed to adapt to the dynamic scenarios in a distributed environment. A moving user acts as an FL central server to learn meta parameters, rather than a fixed central server. It also mitigates the need to share raw data, so security risks can be largely reduced. 2) Hierarchical federated meta-learning approach is designed. Based on the data features, similar data features are clustered. In each cluster, hierarchical outer loops and inner loops are dynamically generated. Outer loops are used to build more generalisable ML systems, which stays mostly invariant. In contrast, their corresponding inner loops are designed for dynamically matching data features to ML models to train a more personalised model for each user. 3) Comprehensive experiments are conducted to evaluate the performance of PPFM under various distribution environments and compare the PPFM with state-of-the-art FL methods with respect to the accuracy. Experimental results demonstrate that our proposed method outperforms baseline FL approaches.

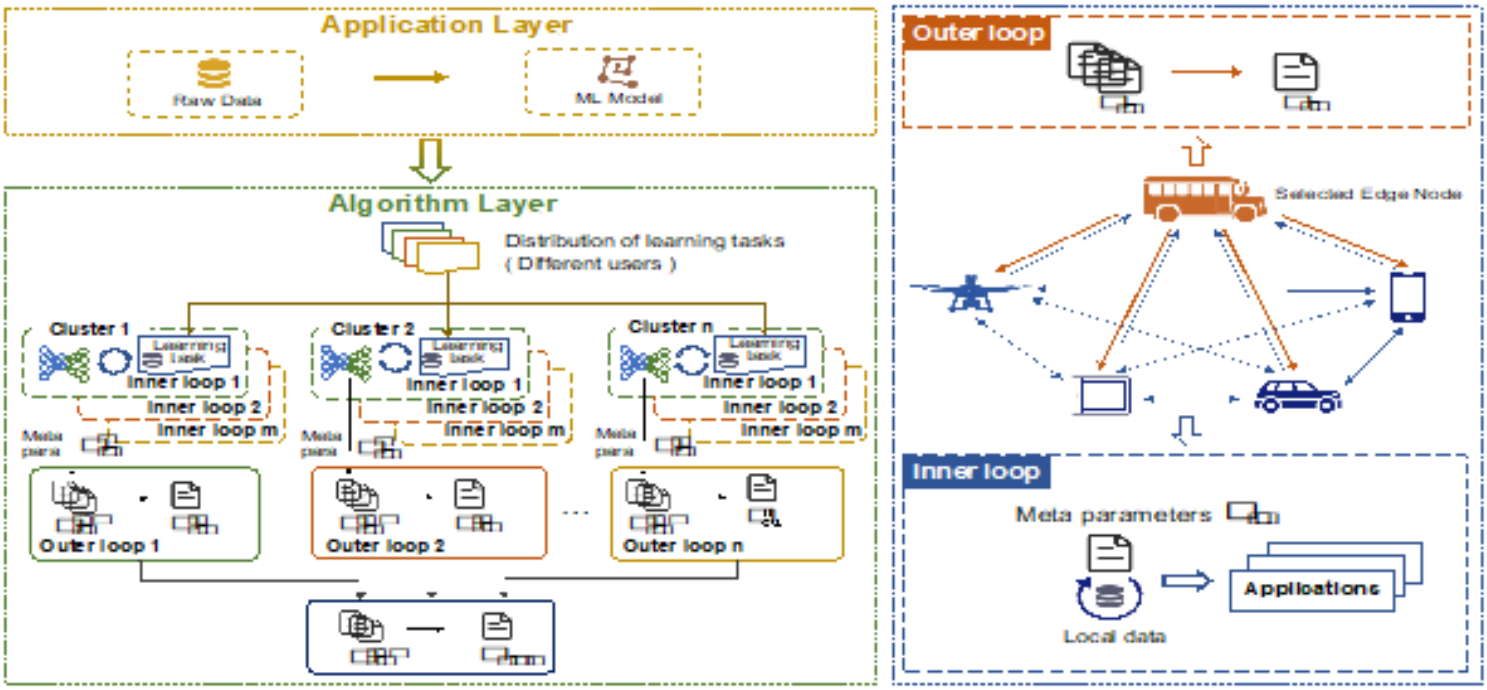


Fig. 2: The proposed peer-to-peer learning framework from RS2A

Research Activities (RS2A)

Security in the Mission and Operational Surface

Adversarial Attacks to Autonomous Systems that utilise Deep Learning.

During the last decade, deep neural networks have achieved tremendous success as they could achieve high accuracy on different complex applications as computer vision, and natural language processing. However, recent findings have shown that deep learning models have several vulnerabilities to adversarial attacks. Deep learning tends to make wrongly overconfident predictions on modified data (Fig. 3). Furthermore, their “black-box" nature makes it extremely difficult to audit their decisions.

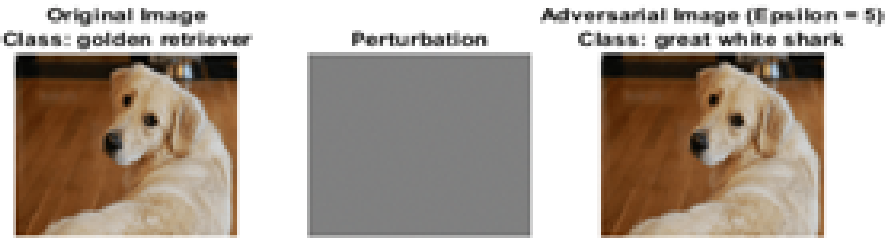
Security aspects of machine learning are extremely important especially on high stake applications as autonomous cars, and medical applications. In particular, robustness to adversarially-chosen inputs is becoming a crucial design goal as recent work shows that an adversary is often able to manipulate the input so that the model produces an incorrect output.

Research in Year 1

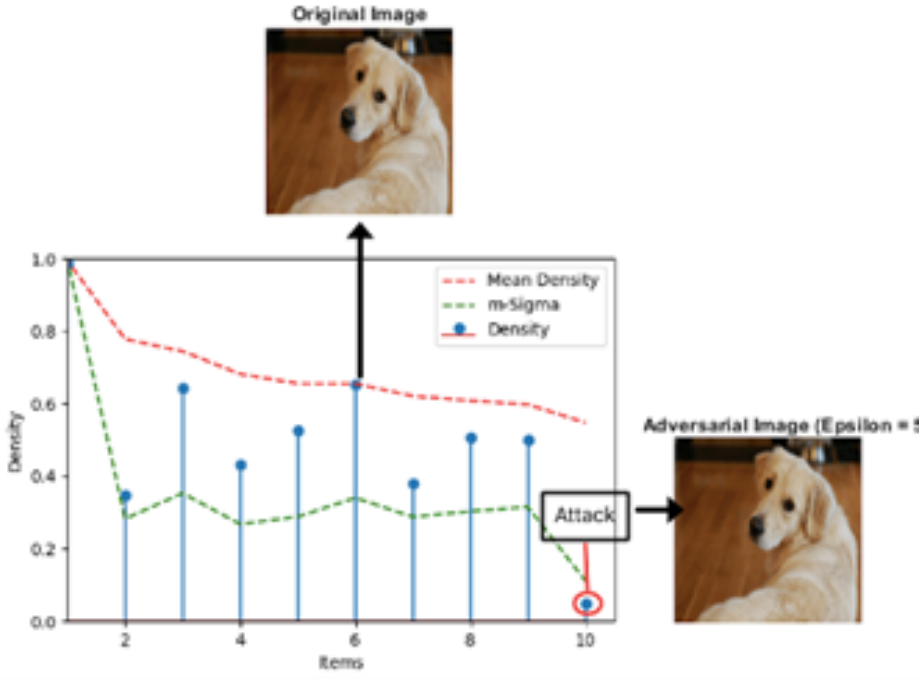
Defending against such attacks is an important research topic. In this sense, we are proposing a prototype-based method that is able to detect changes in the data patterns and detect imperceptible adversarial attacks on real time. Differently from traditional approaches, the proposed method does not require a specific training on adversarial data to improve its robustness. Initial results have shown its efficiency to detect imperceptible adversarial attacks of algorithms as PerC (Fig. 4). Further steps include experiments on different types of attacks and applications.

Plan for Year 2

We plan to develop an explainable-by-design and robust approach able to detect adversarial attacks without prior training. Differently from classical approaches, the proposed method is based on the concept of data density and similarity, therefore, it may not require previous training on adversarial data to be able to detect them. Any change on the concept of the data should be detected through the recursive density estimation mechanism. A paper is planned to formalize this concept, and its applications to different scenarios.



***Fi**g. 3 – Adversarial attack fooling a DenseNet network.*



***Fi**g. 4– Adversarial attack detection based on the drop of the density.*

Research Activities (RS2B)

Securing the Control Surface

AS rely on the ability to conduct run time adaptations of control decisions over attacks which can result from information and dynamic environment uncertainties. Specifically, in the context of learning enabled AS, it is crucial for the control system to exhibit self-aware learning in which the boundaries of “safe” state-space and the control space are tracked through their evolution. This is particularly challenging when the system is undertaking dynamic decisions within the AS mission surface

Research in Year 1

The theme has focused on designing novel AI enabled control methods to allow the necessary adaptive control needs in face of attacks on the control surface. In that respect we have designed a new reinforcement learning based closed loop adaptive control methodology and demonstrated the feasibility on advanced air mobility devices. The approach allows the control system to develop AI driven adaptation strategies in a well-known adaptive control scheme through reinforcement learning. One of the envisioned impacts of this approach is the natural pathway which can be further utilized towards certification as it is using an industrial control backbone. As such, the team has applied its core competence to works on safe reinforcement learning for autonomous airborne collision avoidance systems. In that sense, intelligent adaptation and safety has been at the forefront of this theme’s focus towards developing trustworthy autonomous systems. An overview of Year 1 progress in RS2 is shown next.

Plan for Year 2

The theme's focus will be on stability conditions and Lyapunov functions with the help of neural networks. A part of this year’s effort will cover embedding features into the learning design process to ensure explainability of the learning enabled control designs.

Secure Operations of Trustworthy Autonomous Systems

Cranfield University, Lancaster University

Investigators: Weisi Guo, Gokhan Inalhan, Plamen Angelov, Antonios Tsourdos, Vasileios Giotsas, David Hutchison
Research Fellows: Zhuangkun Wei, Oscar Villarreal, Burak Yuksek, Pierre Ciholas

Engineering and Physical Sciences Research Council

UKRI Trustworthy Autonomous Systems Hub

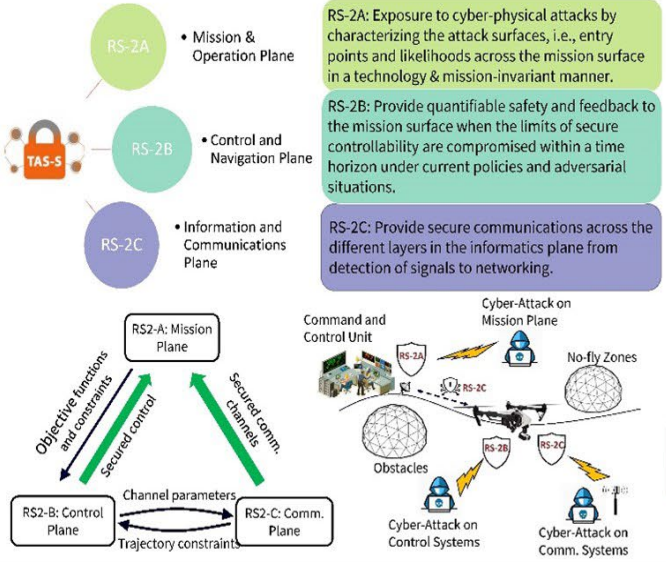
TAS-S

Secure Operation of Autonomous System

Autonomous systems face numerous challenges in their operation, due to the uncertain and dynamic multi-layer attack surfaces

TAS-RS2 aims to solve following challenges

- Modelling & addressing potential attacks in discrete mission, control and communication layers
- Study & address hybrid cascaded cross-layer threats in the dynamic AS space



RS-2A: Securing the Mission Surface

Mission Control for Secure Trustworthy Autonomous Systems requires flexible but reliable real-time optimal decision making and monitoring to handle a wide range of attacks

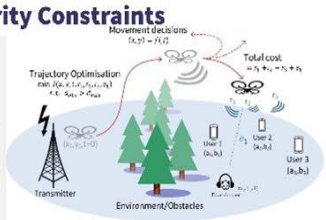
Critical Tasks:

- Handling Communication Errors and Security Constraints
- Assessing Control Faults and Performance Limitations
- Handling Environmental Limitations (Uncertainty/Dynamic Obstacles/No-Fly zones)
- Avoiding and Handling Electronic/Electro-magnetically induced attacks
- Achieving Deterministic/Real-Time Performance for the Optimal Decision Making
- Handling and Detecting Security Threats under Learning-based Scenarios

Key Challenges for Trustworthy Learning-based Mission Control under Security Constraints

Methods and Focus:

- Real-Time Non-Convex Trajectory Optimisation for Path Planning under Uncertainty, Power Consumption, Dynamic Obstacle Avoidance and Communication Security Constraints
- Adaptive and Fault-Tolerant Learning-based Design for Mission Control to improve reliability of safety critical systems
- Supervisory Control for Anomaly Detection and Isolation Systems
- Intelligent Resource Allocation for Multi-UAV Design under Security Threats
- Reliable Self-Assessment under Learning-based Scenarios



RS-2B: Securing the Control Surface

Autonomous Systems rely on the ability to conduct run time adaptations of control decisions over attacks or "perceived" attacks:

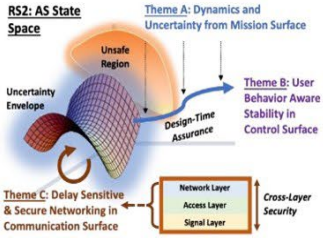
- Adversaries
- Environment uncertainties
- Degraded performance

How to do this in a "trustworthy" fashion?

- Safe, Secure and Reliable

Attack Definitions

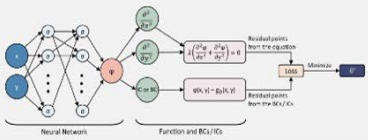
- Sensing and Communication Errors
- Loss of an actuator
- Environmental conditions
- Electronic attacks
- Electromagnetic deception
- Injecting false pattern into data



Key Solution Cornerstones in Learning-Enabled Context

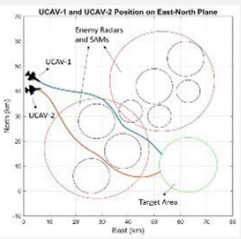
Interpretability => Explainable and Trustworthy AI

- Physics Informed Deep Learning
- Ability to identify system behaviour
- Generalization capability
- Anomaly detection/classification



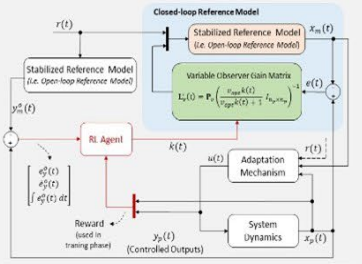
Continual Assurance

- Detect and avoid
- Learning enabled context



Adaptive Security Strategies

- Deep Reinforcement Learning Based Adaptive Controls



RS-2C: Physical Layer Security for AS

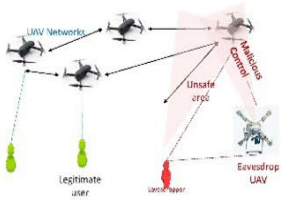
Secure and real-time communications serve as the fundamentals for Autonomous Systems to achieve reliable control and mission delivery.

Attack vectors:

- Key intercept
- Active interference (jamming)
- Passive eavesdropping
- Erode secrecy rate

Traditional Cryptography method:

- Complex key generation/distribution
- High computational complexity
- High latency
- No secrecy guaranteed by brute force



Physical Layer Security: using RECIPRO radio environment

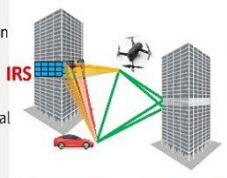
Innovation: exploit unique, dynamic, reciprocal CSIs between entities due to radio propagation nature

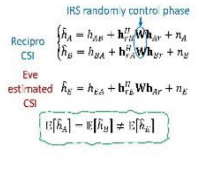
Advantages: low latency & complexity, using only physical channel properties:

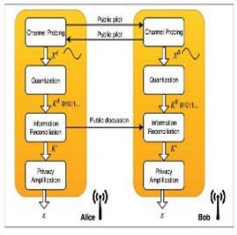
- randomness of wireless channel
- Superiority of legitimate over wiretap channels

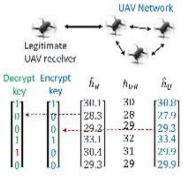
Steps:

- Generate randomness using intelligent reflecting surface (IRS)
- Channel Probing between legitimate users
- Generate cipher keys using recipro CSIs









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

Research Activities (RS2C)

Securing the Cross-Layer Networking (Communication, Sensors)

The focus of RS2C is on the communication and sensory planes of ASs. Here, our research is divided between physical level attacks (Cranfield) and network level attacks (Lancaster).

Physical (PHY) Level

At the Physical (PHY) level, we know digital security can be derived from both antenna beamforming (codeless defence) [2C-1] and deriving distributed keys from channel state information (code-based defence) [2C-2]. The latter is particularly of interest as it can produce secure cipher keys without a common key pool or sharing keys. Yet, it must observe 3 conditions in the PHY channel, namely: (1) reciprocal to allow decentralised synchronous key generation, (2) dynamic to defence against brute force attacks, and (3) unique to avoid correlated attacks. The challenge is that the idealised conditions are often not met for ASs especially in open static spaces and airborne spaces.

Research in Year 1

Our research focused on creating the necessary conditions for code-based physical layer security (PLS) by inducing channel randomness via an intelligent reflecting surface (IRS). Here, subtle random phase changes induce a scatter rich dynamic channel that enables desynchronized cipher code generation [2C-3]. A further challenge is that cooperative eavesdroppers can still attack the pilot channel needed for key generation and we show that a random matrix-based pilot channel that preserve singular values between legitimate users but not to eavesdroppers [2C-4].

Plan for Year 2

Our research will continue to focus on attack and defence theoretical limits for different eavesdropper and channel characteristics, expanding to: (1) collaborative eavesdropping with channel probing and augmentation capabilities, (2) impact of channel security and latency on (2) control and mission planes, and (3) federated autonomy.

[2C-1] “Transmit Beamforming for Layered Physical Layer Security,” W. Zhang et al., IEEE Trans. Vehicular Tech., 2019.
[2C-2] “PLS for IoT: Authentication and Key Generation,” J. Zhang et al., IEEE Wireless Communication, 2019. [2C-3] “Intelligent Reflecting Surface – Induced Randomness for mm-Wave Key Generation,” S. Yang H. Han, Y. Liu, W. Guo, L. Zhang, submitted to IEEE Int. Conf. on Communications (ICC), 2021
[2C-4] “Random Matrix based PLS Key Generation in Static Channels,” Z. Wei, W. Guo, submitted to IEEE Trans. Signal Processing, 2021

Research Activities (RS2C)

Securing the Cross-Layer Networking (Communication, Sensors)

Network Level

The Layers above PHYS are responsible for multi-hop service interconnection. While PHYS attacks fundamentally undermine service interaction, disruptions at the higher communication layers afford more complex opportunities to disrupt the overall composition of the security requirements for a dynamic, mobile, adaptive, ad hoc system with various levels of autonomy. This tight quality of service and resilience requirements to provide autonomy in the system presents a new set of necessary properties for the network layers above PHYS. Importantly, the interaction between the communications and the other elements of the autonomous system (mission layer etc.) require alternative approaches to transparency such that decisions for mission operation can be made in a trusted way, while under attack, providing differing and novel levels of degradation

Research in Year 1

The research has focused on developing a suitable testbed to explore the implications of attacks against the network stack within Autonomous Systems. Our objectives for the testbed were: 1) Dynamic topology (e.g., multiples network paths changing in real time), 2) Control over the simulated links characteristics (e.g., introduce sudden latency spikes, packet loss, disconnections, route changes, etc…), 3) Support for arbitrary protocols (since autonomous systems often use specific, possibly proprietary protocols), 4) High scalability (to simulate small networks or networks with thousands of nodes), and 5) Possibility of using tools for logging, benchmarking, and data analysis. After evaluating numerous platforms, the Testground software was selected and provides a platform for testing, benchmarking, and simulating distributed and peer-to-peer systems at scale. It's designed to be multi-lingual and runtime-agnostic, scaling gracefully from 2 to 10k instances when needed. Testground allows for tests to be written natively to generate behaviours or to inject existing traffic data sets, which was done with Bitswap/IPFS dataset and a Cranfield University autonomous vehicle dataset. The full platform will be made available to the research partners at the start of 2022.

Plan for Year 2

Our research will utilise the developed testbed to explore the core research focus of the cross-layer security issues. We will focus on understand the novel attack surface which the combination of autonomy and mobility provide and the necessary defences to provide assurances for correct functioning under differing levels of attack.

RS3

RS3-Theme A: Behaviour Adaptation as a Basis of Security by Design

Lead: L. Dorn. Participant: J. Deville.

As AS design and application progress, how will people adapt their behaviour in relation to them? And how might behavioural adaptations weaken AS security? Little is known about how critical aspects of a security breach may go unnoticed when operators are out of the loop.

To start with, we are focusing on autonomous vehicles and identifying security issues that may apply to other AS. Previous studies to evaluate behavioural adaptation in responses to assisted and automated vehicles have shown how unintended consequences can mean that safety benefits are not realised and may even be put at risk. These studies have been short term in duration and lab-based with very few studies conducted in real-world fully autonomous vehicles. Longitudinal field-based studies across a range of platforms, with younger and older people, will investigate behavioural adaptation to inform interface design in order to capture the operators attention when the security of the system is compromised.

RS3-Theme B: Organizational Socio-Technical Mitigation

Lead: J. Deville. Participants: L. Dorn, L. Moffat.

As public-private collaborations become more prevalent, there is a need to clarify the liabilities and duties of private companies working in a public capacity because there are different legal and incentive frameworks between private and public organisations. While these collaborations open new possibilities, they also bring forth a range of ethical, legal and social issues which warrant careful consideration. In a collaborative information management setting, it is important to support and encourage reflection on such issues by making more visible the ethical and legal implications of outsourcing, subcontracting, and privatisation in general.

As the rate of technological innovation exponentially increases, the ways that organisations manage their data, business, and their ethics, must adapt. This is not simply a case of ‘keeping up’ with the technology, but of creating synergies, affordances, and spaces for response. Given the extent and diversity of contexts in which A/S do and will operate, organisation adaptation needs to happen ‘all the way through’, from policy and protocol, to everyday practice.

RS3 (cont.)

RS3-Theme C: Ethics and Governance of AS Security

Lead: C. Easton. Participants: L. Dorn, L. Moffat

ELSI

The Ethical, Legal and Social Issues (ELSI) framework is one of many cross-disciplinary approaches to technological innovation, which seeks to examine, address, and advise upon the wider implications of new technologies being implemented in society. In RS3, we draw on this framework to inform our research, our engagement with others within the TAS-S project, and our collaboration with external partners.

Ethical

While traditional ethical theories tend to focus on individual conduct and so here, on individual technological devices, the ELSI framework approaches ethics as an interconnected, complex process of negotiation, appraisal, and reflection. Who benefits from technologies being designed and used, and who is harmed as a result? Tools like Ethical Impact Assessment help developers in industry to audit their new technologies according to these values of benefit and harm.

Legal

Standardisation and best practice go hand in hand with a robust understanding of the legal landscape, as well as the capacity to change it. This requires opening channels between tech developers, operators, and policy advocates, so that legal practices can help forecast better AS futures, as well as responding to existing challenges.

Social

AS do not exist in a vacuum. They operate in, engage with, and respond to, pre-existing social structures and protocols. A key component of ELSI involves making space for communities to voice their thoughts, apprehensions, and desires for how AS work with and for them. Putting the ELSI framework into practice is not easy, nor should it be. But approaches like ELSI are essential for ensuring that new autonomous technologies are not only possible, but suitable for society. Future work in RS3 will also be using the Design Justice principles, in combination with ELSI, to look at ways of expanding the positives of AS according to community assets, and limiting harms imposed by techno-solutionism. You can find out more about Design Justice here

Research Activities (RS3A)

Behaviour Adaptation as a Basis of Security by Design.

Previous studies to evaluate behavioural adaptation (BA) have mostly considered short term effects and it is unclear how repeated longitudinal exposure to AS may impact individual response to security threats.

Research in Year 1

Research to investigate behaviour adaptation as a basis of security by design began in February 2021 with a scoping review of peer-reviewed, published primary studies to identify methods for investigating behavioural change in response to autonomous systems and how this may compromise security. Comprehensive searches were conducted using ESCBO databases and a Title and Abstract Sift (TAS) protocol was developed. Three rounds of calibration were conducted until a Cohen's Kappa score of 0.95 was reached.

In April 2021, a newly recruited PDRA working full time on the project undertook a Full Text Review (FTR) screening procedure of 880 papers resulting in 174 relevant papers selected for analysis. A Rapid Evidence Review was then conducted between June and August to investigate the mental models that guide how humans interact with Autonomous Systems and how specific behaviours change as humans adapt to autonomous systems. Summary findings were presented at the TAS All Hands Meeting in September, showing that previous studies are narrow and that no studies have been conducted on how human behaviour can compromise security for autonomous systems.

Plan for Year 2

The PDRA left Cranfield in October 2021 and it is hoped that a new PDRA will be recruited by Easter 2022. The research plans for 2022 include undertaking an in-depth analysis of a chosen set of 174 key papers to inform the design of a pilot study to investigate behavioural adaptation as a basis of security by design using collaborative robots. The research team also plans to develop a subjective instrument to measure behavioural adaptation that can be used alongside objective measures.

Research Activities (RS3B)

Organisational Socio-Technical Mitigation.

Research in Year 1

The team's focus in Year 1 has been on (a) developing new social scientific methods for enabling organisations to better deal with and anticipate security challenges related to the deployment of autonomous systems and (b) on engaging and building partnerships with organisations. This included the initial advisory board meeting, and in collaboration with RS3C, scoping the organisational challenges confronted a range of relevant stakeholders.

In this session, and in a follow up workshop organized in partnership with the National Cyber Security Centre, which brought together experts from industry, advocacy, and academia, the team explored the role of scenarios in enabling organisations to imagine their future engagements with autonomous systems.

Towards the end of Year 1, and again in collaboration with RS1C, the team also began the process of negotiating a research partnership with National Highways (formerly Highways England). It is anticipated this will form the basis of a 12-month collaboration, to start in Year 2 (see below). This work will provide the basis for a toolkit to be used by organisations seeking to engage more critically with the security challenges associated with in autonomous systems design and deployment.

Plan for Year 2

The collaboration with National Highways is the key priority this year. This will involve a sequence of workshops and 1:1 interviews with members of the organisation to understand (a) the security challenges associated with their use of autonomous systems, (b) the role of organisational content in shaping these challenges, and (c) the potential for innovative, creative / scenario-based methods to enable organisations facing similar challenges to National Highways to enable them to engage more critically with the security challenges associated with autonomous systems design and deployment.

Research Activities (RS3C)

Ethics and Governance of AS Security.

Research in Year 1

The team has focused on an analysis of the ethical issues underpinning the project's development. An analysis has been undertaken of on-going work in the discipline with a particular focus on security, an area that is lacking in the majority of ethics-focused studies. At the advisory board meeting, in collaboration with Theme B, legal and ethical issues were probed with a range of relevant stakeholders. Based on this, and again in collaboration with Theme B, National Highways was approached to collaborate in an intensive review of issues of ethics and security relating to the deployment of autonomous. At the end of Year 1, a formal collaboration proposal was sent to National Highways to review.

Plan for Year 2

Publication is a priority in this year, with plans developed for at least two articles in the area of ethics and governance. The aim is that these will be jointly written. Targeted research will be undertaken at a number of stakeholder workshops, likely to include workshops with National Highways.

Acknowledgements

We would like to thank our stakeholders for their continued support.

This work is supported, in part, by the Engineering and Physical Sciences Research Council [Grant number: EP/V026763/1]

Contact

TAS-S Node
Security Lancaster
InfoLab 21
Lancaster University
LA1 4WA

https://tas-security.lancs.ac.uk/
tas-s@lancaster.ac.uk
@TAS_Security