An Investigation into Physical and Communications Trust Frameworks for Collaborative Teams of Autonomous Underwater Vehicles

Andrew Bolster, Prof. Alan Marshall, Prof. Jean-Guy Fontaine

Advanced Networks Research Group, University of Liverpool, UK

Project Background and Outputs

- Attendance at UComms 2012 (Sestri Levante, Italy)
- Poster Presentations in 2012 (Kassam, Oxford) and 2013 (Heathrow, London and Bagneaux, Paris)
- Summer Research Placement with DSTL (Software Systems and Dependability for Autonomous Teams)(2013, Portsdown West, Portsmouth)
- Short Paper Presentation to the Association for the Advancement of Artificial Intelligence (AAAI) on “A Multi Vector Trust Framework for Autonomous Systems” (2014, Stanford, CA)[1]
- Technical Report for the UK/US/CAN/AUS/NZ Technical Cooperation Programme ((2014). Analysis of Trust Interfaces in Autonomous and Semi-Autonomous Collaborative MHPC Operations. The Technical Cooperation Program, Technical Report TR-C3I-06-2014) (June 13 - April 14)[2]
- DSTL CDE Collaboration with NPL and Plextek Ltd. on “Precision Timing and Navigation, Challenge 1: Resilient Time and Location Estimation for Networked Assets” (CDE 33135) (Oct 13-May 14)

Introduction

Aim of project: To combine physical and communications observations to assess and maintain trust within mobile, marine, ad-hoc networks

Small fleets of AUVs (*Autonomous Underwater Vehicles*) will be expected to operate in isolated environments.

This requires an auditable sense of trust within the remote intra-fleet communications networks, incorporating

- Communications Activity
- Mission Suitability/Capability
- Behavioural Monitoring

The use of centrally coordinated trust models presents a single point of failure.

Secure communication in marine environments is expensive and time consuming; adopting a decentralised form of trust assurance will reduce these costs by localising the per-node security environment.

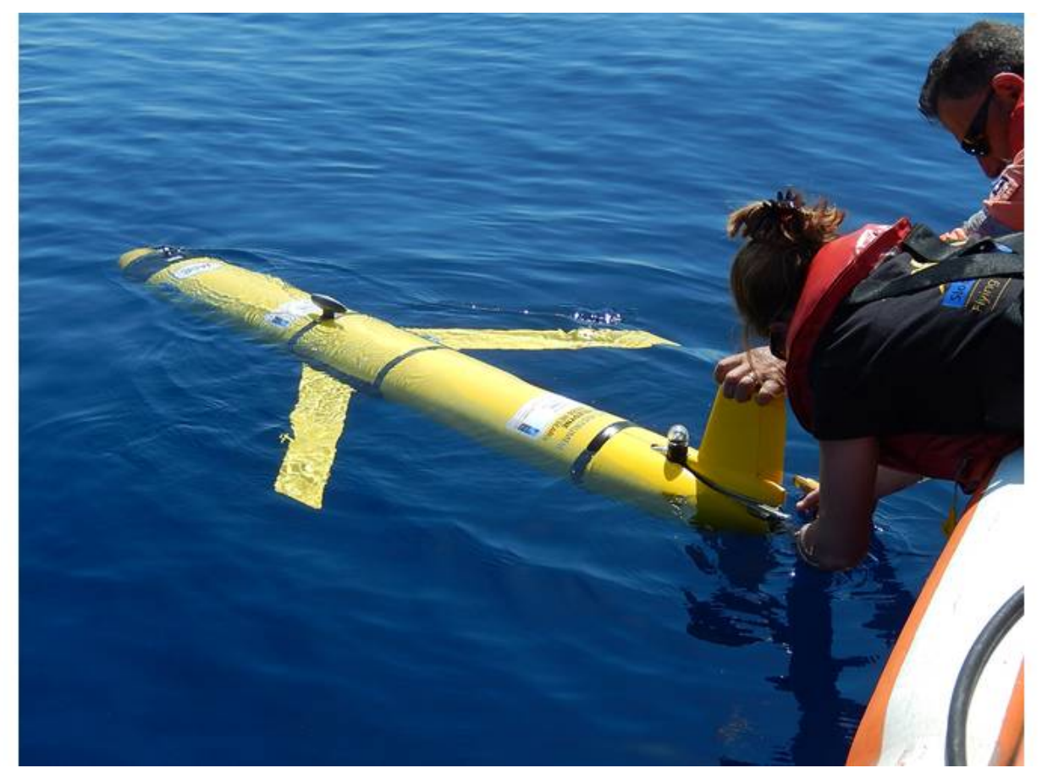


Figure 1: REMUS 100 AUV, as deployed at CMRE, a potential target platform for this work

Trust

Trust is:

- *The expectation of an actor performing a certain task or range of tasks within a certain confidence or probability*
- a belief on the reliability of an entity
- based on both direct and indirect historical experience

Individual trust opinions are shared within the network concerning a range of activities:

- Transmission Relaying (Local and/or Backhaul)
- Position Relaying
- Reporting Accuracy

These Trust opinions also apply to extra-fleet entities, such as surface platforms, submarine comms links, and coastal stations, allowing the fleet to collaboratively form an opinion of these actors.

Trust Management Frameworks (TMFs)

TMFs are protocols designed to provide information regarding the estimated future states and operations of nodes within networks

“[. . .]collecting the information necessary to establish a trust relationship and dynamically monitoring and adjusting the existing trust relationship” - [3]

Enables nodes to form collaborative *opinions* on their cohort nodes based on

- Direct Observation of Communications Behaviour (eg Successfully Forwarded Packets)
- Common-Neighbour Recommendation
- Indirect Reputation

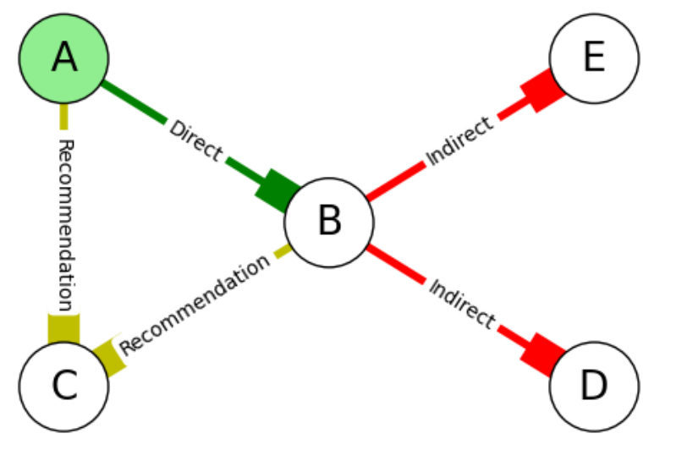


Figure 2: Direct, Recommendation, and Indirect trust relationships

Multiple transitive relationships can be maintained over time, providing trust resilience with dynamic network topologies

The Need for Multi-Domain Trust in Autonomous Systems

Communications not the only target for an attacker (or failure);

- Following to restricted area
- Masquerading
- Hardware Degradation
- Resource attack via propulsive power

Physical observation presents opportunity to further reduce the available threat surface while also discriminating between ’True’ attacks and mechanical failure.

Potential attacks exist within a multi-domain threat surface, and as further metrics and domains of trust are included in a TMF, attackers are increasingly restricted in their behaviour until the only way to avoid detection is to behave correctly.

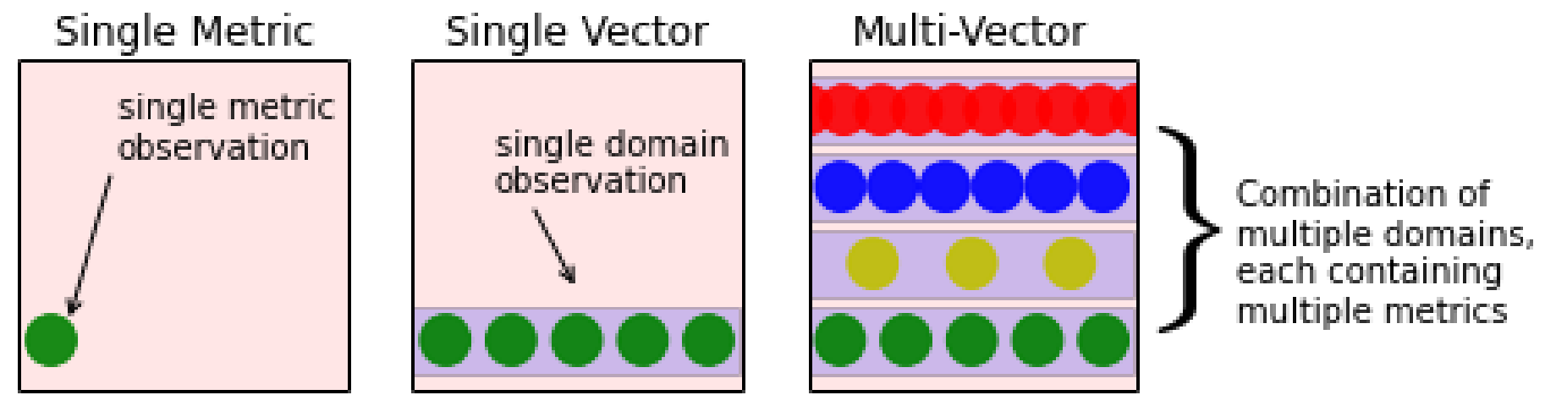

Single Metric
single metric observation
Single Vector
single domain observation
Multi-Vector
Combination of multiple domains, each containing multiple metrics

Figure 3: Threat Surface for Trust Management Frameworks

Multi-Metric Trust Assessment

Most TMFs can be generalised as single-metric estimators based on a binary input stats (Packet Loss/Delivery Rates), which provides malicious actors advantage if their activity does not affect that metric. MTFM[4] analyses more than PLR to make it’s assessment, including Received and Transmitted signal strength, delay, and throughput as well as taking account of dynamic network topology to inform assessment.

$$[\theta^t_{k,j}, \phi^t_{k,j}] = \frac{\min_k |a^t_{k,j} - r^t_j| + \rho \max_k |a^t_{k,j} - r^t_j|}{|a^t_{k,j} - r^t_j| + \rho \max_k |a^t_{k,j} - r^t_j|}, r \in [g, b] \tag{1}$$

$$[\theta^t_k, \phi^t_k] = \left[\sum_{j=0}^{M} h_j \theta^t_{k,j}, \sum_{j=0}^{M} h_j \phi^t_{k,j}\right] \tag{2}$$

$$T^t_k = (1 + (\phi^t_k)^2/(\theta^t_k)^2)^{-1} \tag{3}$$

where $a^t_{k,j}$ is the value of an observed metric x_j for a given node k at time t, ρ is a distinguishing coefficient set to 0.5, g and b are respectively the “good” and “bad” reference metric sequences from a, i.e. $g_j = \max_k(a^t_{k,j})$, $b_j = \min_k(a^t_{k,j})$ These metric coefficients are then accumulated (2) and combined to present a singular trust value for analysis (3). The weights used in (2) can be used to interrogate the trust value space, putting more emphasis on one or more metrics to identify and better characterise a misbehaviour.

Operational Mission Profiles

Generic behaviours currently under investigation include:

- *Waypointing* - Attraction to a point or a chain of points, providing pre-described patrol networks
- *Surveying* - Fleets can be tasked to provide one-shot, or persistent coverage of an area of the environment utilising a dynamic lawnmower pattern.
- *Dynamic Constraint* - Repulsion from a series of points, analogous to sea-borders or shipping lanes.

Potentially Exploitable Behaviours not yet developed include:

- *Capacity Based Homing* - Where a node leaves and later returns to the fleet, for instance for refuelling or resupply.
- *Dynamic Communications Maintenance* - The fleet can adjust to changing communications environments.

Behavioural Analysis

Three fleets are simluated performing a simple 8 hour duration patrol mission. The first includes a malicious node attempting to infiltrate the fleet (Shadow). The second has an impared or faulty but otherwise "good" node (SlowCoach). The third is a baseline fleet with all "Good" nodes

Physical Metrics under assessment:

- *INHD*: Inter Node Heading Deviation, or the per-node variation from the fleet-average direction
- *Node Speed*: The Magnitude of Velocity of each node compared to the fleet-average
- *INDD*: Inter Node Distance Deviation, or the variation between the inter-node distances between each node

In this initial model, there are no ’positive’ trust expression; all nodes are ’distrusted’ to some degree dependant on consistent or periodic deviation from fleet norms.

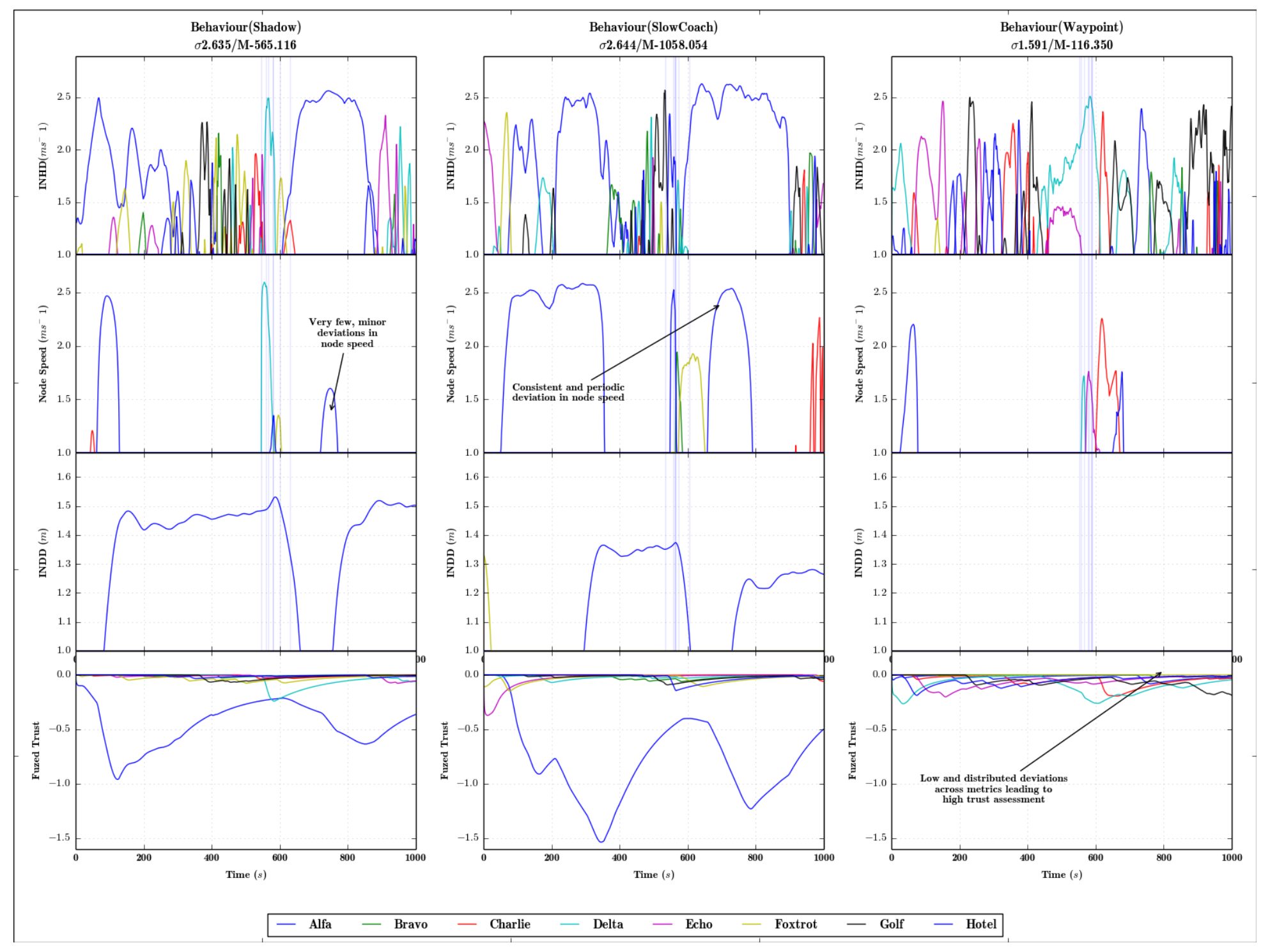


Figure 4: Per-Node deviations for each metric, with an additional row showing an EWMA based cross-metric trust assessment. Annotations highlight difference in ’Node Speed’ triggers between the malicious and impaired behaviours

From Fig. 4, *INDD* is a clear candidate for a suspicion ’trigger’, but looking at *INHD* values after several minutes of mission time; an anomaly is clearly being detected.

In addition, *Alfa* node (Blue) is clearly an outlier in terms of *INHD* and *Node Speed* in the earliest sections of the graph. This implies that a fusion of metrics would be more effective than a simple detection envelope on a single metric. Considering the baseline Waypointing results in Figure 4, it’s clear that these metrics are not infallible, as is demonstrated by the number of relatively short-lived false positives, demonstrating the need to use multiple metrics for reliable trust assessment.

Figure 4 demonstrates a windowed, weighted trust fusion, where deviations in individual metrics are combined to generate a Trust Value.

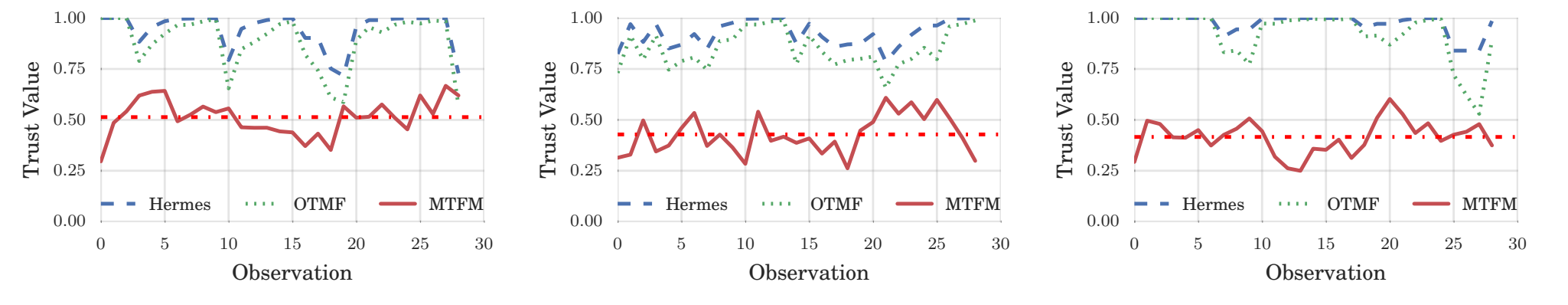
Single and Multi-Metric TMF operation in Marine Comms

Acoustic Network based on AUVNetSim [5] and validated against [6].

Aim to investigate use of Multi-Parameter Trust Framework for MANETS (MTFM), against current communications TMFs (Hermes/ Objective Trust Management Framework (OTMF)), which exclusively use Packet Loss Rate (PLR) as their assessment metric.

Two Communications Misehaviours were generated:

- **Malicious Power Control**(MPC) where a malicious node (n_1) inflates it’s power to all nodes except a target node (n_0) such that that node appears to be behaving selfishly.
- **Selfish Target Selection**(STS) where n_1 preferentially communicates with nodes that are physically near-by, reducing its own power consumption.



(a) Fair Scenario (b) MPC Scenario (c) STS Scenario

Figure 5: $T_{1,0}$ for Hermes, OTMF and MTFM assessment values for fair and malicious behaviours in the fully mobile scenario (mean of MTFM also shown)

From 5, in the challenging underwater environment, no assessment tool is able to appreciably differentiate between behaviours (while MTFM does display a 10% discriminating behaviour in the a-postori average assessment, shown as a red dashed line)
Metric suitability for trust assessment using a distributied weighting analysis using Random Forest Regression Techniques.

Future Applications

- Due to the high communications, motion, and computation costs, and lack of external location reporting (*e.g. GPS*), behavioural analysis in the marine environment is particularly difficult, but if successful, can be reliably applied in a wide variety of fields including but not limited to
  - ▷ Self-Driving Cars
  - ▷ Environmental Survey drones (terrestrial, marine, and aerial)
  - ▷ Satellite Communications Arrays
  - ▷ Internet Certificate Authority verification
  - ▷ Verifiable Distributed Computing

Conclusions

This research area presents a range of challenges and opportunities within both civil and defence operations; an auditable trust framework for automated marine craft would be a significant enabling factor to the roll-out of more low-maintenance or even “Fire and Forget” deployments for persistent patrol/monitoring tasks.

Open Hypotheses in this field that this project intends to answer are:

- How can optimality in trust assessment based on behaviour be defined win a distributed, dynamic network topology?
- Is there a quantifiable benefit to cross-domain comparison beyond single-vector trust? (i.e. 1-D vector vs cross domain comparison)
- Is there an optimal *generic* cross domain fusion methodology?

Development Plan

- Behaviour Detection (Q3 14) - Formal Analysis of Behavioural Trust Systems
  - ▷ ASON 2014 : Seventh Int. WS on Autonomous Self-Organizing Networks (Aug 14)
  - ▷ AHUC 2014 : The Fourth Int. WS on Ad Hoc and Ubiquitous Computing (Aug 14)
  - ▷ ICCAR 2015 : WASET Int. Conf. on Control, Automation and Robotics (Dec 14)
- MANET/Marine comparison (Q4 14) - Formal Comparison between Terrestrial MANET / Marine contexts
- Multi-Domain Trust Assessment (Q4 14) - Combination of Communicative and Physical Behaviour Trusts
  - ▷ IEEE Trans. on Communications / Dependable and Secure Computing / Intelligent Systems
- Reactionary/Perturbative Trust (Q1 15) - Exploration of reactionary behaviours for teams to ’shake down’ suspects
  - ▷ SASO15:Self-Adaptive and Self-Organizing Systems,
  - ▷ SEAMS15: Software Engineering for Adaptive and Self-Managing Systems

Bibliography

- Andrew Bolster and Alan Marshall. “A Multi-Vector Trust Framework for Autonomous Systems”. In: *2014 AAAI Spring Symp. Ser.* Stanford, CA, 2014, pp. 17–19. URL: http://www.aaai.org/ocs/index.php/SSS/SSS14/paper/viewFile/7697/7724.
- Andrew Bolster. *Analysis of Trust Interfaces in Autonomous and Semi-Autonomous Collaborative MHPC Operations*. Tech. rep. The Technical Cooperation Program, 2014.
- Huaizhi Li and Mukesh Singhal. “Trust Management in Distributed Systems”. In: *Computer (Long. Beach. Calif).* 40.2 (2007), pp. 45–53. ISSN: 00189162. DOI: 10.1109/MC.2007.76. URL: http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=4085622.
- Ji Guo, Alan Marshall, and Bosheng Zhou. “A new trust management framework for detecting malicious and selfish behaviour for mobile ad hoc networks”. In: *Proc. 10th IEEE Int. Conf. Trust. Secur. Priv. Comput. Commun. Trust. 2011, 8th IEEE Int. Conf. Embed. Softw. Syst. ICESS 2011, 6th Int. Conf. FCST 2011* (2011), pp. 142–149. DOI: 10.1109/TrustCom.2011.21. URL: http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=6120813.
- Josep Miquel and Jornet Montana. “AUVNetSim: A Simulator for Underwater Acoustic Networks”. In: *Program* (2008), pp. 1–13. URL: http://users.ece.gatech.edu/jmjm3/publications/auvnetsim.pdf.
- Andrej Stefanov and Milica Stojanovic. “Design and performance analysis of underwater acoustic networks”. In: *IEEE J. Sel. Areas Commun.* 29.10 (2011), pp. 2012–2021. ISSN: 07338716. DOI: 10.1109/JSAC.2011.111211.

A UK/Fr Joint Project funded by DSTL/DGA with assistance from UoL, CMRE, and UPMC

<andrew.bolster/alan.marshall>@liv.ac.uk