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 20.25 (Heathrow, London and Bagneaux, Paris)
- Summer Research Placement with DSTL (Software Systems and Dependability for Autonomous Teams)(20.25, PDW)
- Paper Presentation to the Association for the Advancement of Artificial Intelligence (AAAI) (Stanford, USA) [1]
- Technical Report for the UK/US/CAN/AUS/NZ Technical Cooperation Programme [2]
- DSTL CDE Collaboration with NPL and Plextek Ltd. on “Precision Timing and Navigation, Resilient Time and Location Estimation for Networked Assets” (CDE 33135)
- Paper Presentation to the IEEE International Symposium on Recent Advances of Trust, Security and Privacy in Computing and Communications (Accepted) (Helsinki, FI) [3]

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 environment, requiring an auditible sense of trust within the remote intra-fleet communications networks, incorporating

- Comms. Activity
- Mission Suitability/Capability
- Behaviour Monitoring

The use of centrally coordinated trust models presents a single point of failure, and secure communication in marine environments is expensive and time consuming.

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

Adopting a decentralised form of trust assurance will reduce these costs by localising the per-node security environment.

Trust & Trust Management Frameworks (TMFs)

Trust is the expectation of an actor performing a certain task or range of tasks within a certain confidence or probability
Individual trust opinions are shared within the network concerning a range of activities:

- Transmission Routing (Local and/or Back-haul)
- Position Reporting
- 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.
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” - [4]

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

- Direct Observation of Behaviour
- Common-Neighbour Recommendation
- Indirect Reputation

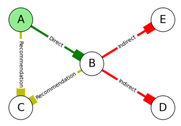


Fig. 2: Direct, Recommendation, & Indirect trust relationships

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

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.

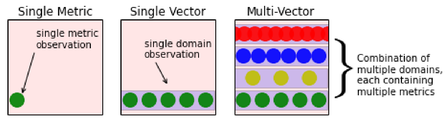


Fig. 3: Threat Surface for Trust Management Frameworks

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[5] 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.

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

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

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

where $\boldsymbol{a}^t_{k,j}$ is the value of an observed metric $\boldsymbol{x}_j$ for a given node $\boldsymbol{k}$ at time $\boldsymbol{t}$, ρ is a distinguishing coefficient set to **0.5**, $\boldsymbol{g}$ and $\boldsymbol{b}$ are respectively the “good” and “bad” reference metric sequences from $\boldsymbol{a}$, i.e. $\boldsymbol{g_j} = \max_k(\boldsymbol{a}^t_{k,j})$, $\boldsymbol{b_j} = \min_k(\boldsymbol{a}^t_{k,j})$
These metric coefficients are then accumulated (2) and combined to present a singular trust value for analysis (3).

Operational Mission Profiles

Generic mobility behaviours currently under investigation include:

- **Waypointing** - Pre-described patrol routes
- **Surveying** - Persistent coverage of an area.
- **Dynamic Constraint** - Repulsion from points (geofencing).

Potentially Exploitable Behaviours not yet developed include:

- **Capacity Based Homing** - e.g. periodic refuelling
- **Dynamic Communications** - Adjust to changing backhaul

A series of ’Malicious’ and ’Faulty’ misbehaviours have been designed:

- **Shadow** - Following the fleet without restricted mission data
- **Spy** - A node that intermittently rises in the fleet, potentially surfacing to relay mission information to a third party
- **Sloth** - Selfish energy conservation by taking minimal paths
- **Stalker** - Tailing a specific node, attempting to use this consistent history to poison the trust network
- **Scoundrel** - Falsely reporting its estimated position with the intention to corrupt collaborative localisation or induce a collision
- **Slow Coach** - Where a node’s power train is damaged, causing reduced manoeuvrability and performance
- **Spin Doctor** - Damage to control surfaces or to Inertial Navigation System

Behavioural Analysis

Three fleets are simulated performing a simple 8 hour patrol mission. The first included a malicious node attempting to infiltrate the fleet (Shadow). The second has a 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

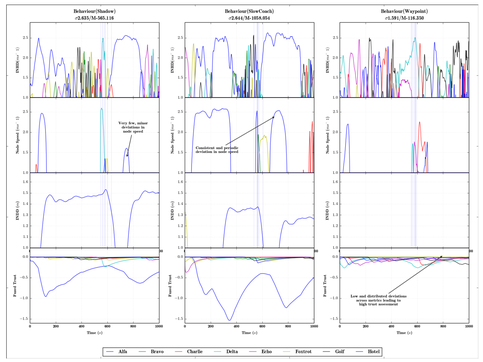


Fig. 4: Per-Node deviations for each metric, with an additional row showing an EWMA based cross-metric trust assessment

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.

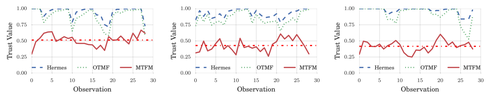
In addition, **Alfa** node (Blue) is clearly an outlier in terms of **INHD** and **Node Speed** in the earliest sections of the graph, and this is shown in the Fused Trust results, however the difference in expression in Node Speed enables not only detection but classification of misbehaviour with a high degree of selectivity (97%).

Single and Multi-Metric TMF operation in Marine Comms.

Acoustic Network based on AUVNetSim [6] and validated against [7].

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

Two Communications Misbehaviours were created: **Malicious Power Control**(MPC) where a malicious node ($\boldsymbol{n_1}$) inflates it’s power to all nodes except a target node ($\boldsymbol{n_0}$) making it appear selfish and **Selfish Target Selection**(STS) where $\boldsymbol{n_1}$ preferentially communicates with nodes that are physically near-by, reducing its own power consumption.



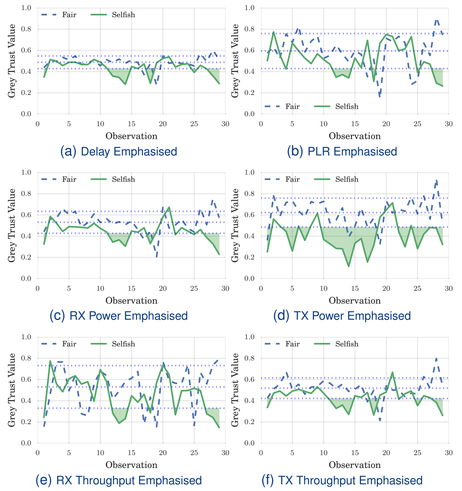
(a) Fair Scenario (b) MPC Scenario (c) STS Scenario

Fig. 5: $\boldsymbol{T_{1,0}}$ 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 Permutation and Classification

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. Using this process we can extract and highlight the primary aspects of an attack (MPC/STS) by comparing against the deviation from the “fair” result set.



(a) Delay Emphasised (b) PLR Emphasised
(c) RX Power Emphasised (d) TX Power Emphasised
(e) RX Throughput Emphasised (f) TX Throughput Emphasised

Fig. 6: $\boldsymbol{T_{1,MTFM}}$ in the All Mobile case for the STS behaviour, including dashed $\pm\sigma$ envelope about the fair scenario

Weight Significance Analysis for Behaviour Classification

Applying a Random Forest regression tree to 729 different weighting schemes for each of the three behaviours. We assess how important each metric is in differentiating between behaviours; clearly demonstrating that PLR is not an important metric in this regard, where as Received Throughput $\boldsymbol{T^P_{RX}}$ is a major discriminator between STS and MPC.

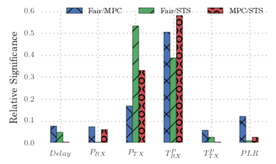


Fig. 7: Factor Analysis of Malicious (MPC), Selfish (STS) and Fair behaviours selectivity

Table 1: Correlation Coefficients between metric weights and behaviour detection targets

Correlation	*Delay*	P_{RX}	P_{TX}	T^P_{RX}	T^P_{TX}	*PLR*
Fair / MPC	0.199	0.159	-0.416	0.708	-0.238	-0.401
Fair / STS	0.179	-0.009	0.724	-0.697	-0.145	-0.052
MPC / STS	0.058	-0.134	0.146	-0.768	0.052	0.146

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?

Thesis Plan

- Trust and its applications to MANETs
  - Discussion on abstract analysis of trust networks
  - Discussion on the threat surface of Mobile Ad Hoc Networks and how that has been protected so far
  - Introduction to Trust Management Frameworks and their benefits
- Maritime Uses of Autonomous Systems
  - Discussion of current and future approaches to areas where autonomous systems can be used mainly focused on Mine counter measures, Hydrography and Patrol Capabilities (MHPC)
  - Discussion of the contextual human factors around integrating autonomous systems into existing human-based solutions, predominantly following on from [2], including development of representative malicious and abnormal behaviours
- Strategies for Multi-Domain Trust Assessment
  - Analytical establishment of Multi-Domain Trust, from an information theoretic standpoint.
- Modelling and Analysis of Collaborative Node Kinematic Behaviours in Underwater Acoustic MANETS
  - Focused on the mobility and assessment of mobility between nodes, including identification of suitable motive metrics and analyses of these motions to establish intent or abnormality
  - Incorporating collaboration with NPL/Plextek as supporting evidence.
- Comparative Analysis of Multi-Domain Trust Assessment in Collaborative Mobile Networks
- Investigation into the relative performance characteristics of multi-domain combination strategies in an exemplary context (AUV teams) against existing single and multi metric TMFs

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.

Andrew Bolster and Alan Marshall. “Single and Multi-Metric Trust Management Frameworks for use in Underwater Autonomous Networks”. In: *TrustCom2015*.

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.