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

Andrew Bolster

University of Liverpool

andrew.bolster@liv.ac.uk

July 1, 2015

Research Context

- Project launched at QUB ECIT in 2011 under the DSTL/DGA Anglo French Defence Research Group PhD Programme under Profs. Alan Marshall and Jean-Guy Fontaine
- What lessons from the Mobile Ad Hoc Network (MANET) space can be transferred to the marine environment?
- Teams of 3 - 16 Autonomous Underwater Vehicles (AUVs) Mine countermeasures, Hydrography, and Patrol Capabilities (MHPC)
- Defence focus, assumption of highly capable enemy attempting to compromise communications / operations
- Primary Simulation/Analysis work done in 12/13
- Moved to UoL Oct 13 after 2 mth placement @ DSTL PDW Naval Systems / Information Systems departments.
- Communications Analysis work done in 13/14

Research Collaborations

- DSTL
  - Visits and Placements (Summer '13) at DSTL Porton Down and Portsdown West
  - CDE Exhibition, London, (Spring '12)
  - PhD National Conferences, Oxford, London and Paris
- DGA/UPMC
  - DGA Conference (Autumn, '12)
  - Visits fo CRIIF (Autumn, '12)
- NATO/CMRE
  - UComms'12
  - Visits & Ongoing data sharing with CMRE(NURC) in La Spezzia
- NPL/Plextek
  - CDE Project on Precision Timing for Positioning with NPL/Plextek

Trust in Ad-Hoc Systems

- Particularly interested in the application of Trust in Decentralised (P2P) Autonomous Systems of Systems, Autonomous Underwater Vehicles (AUVs) for example

Trust in Ad-Hoc Systems

- Particularly interested in the application of Trust in Decentralised (P2P) Autonomous Systems of Systems, Autonomous Underwater Vehicles (AUVs) for example
- Trust: *The expectation of an actor performing a certain task or range of tasks within a certain confidence or probability*

Trust Management Frameworks

- Provide information regarding the estimated future states and operations of nodes within networks

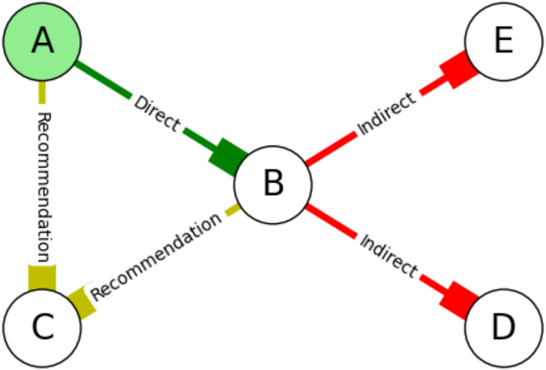
Trust Management Frameworks

- 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" - [1]

Trust Management Frameworks

- 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" - [1]
- 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

Transitivity in Trust Networks



TMFs in Ad Hoc Autonomous Systems

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

TMFs in Ad Hoc Autonomous Systems

- Multiple transitive relationships can be maintained over time, providing trust resilience with dynamic network topology
- Enable trust establishment from partial-strangers via indirect trust and direct observation

TMFs in Ad Hoc Autonomous Systems

- Multiple transitive relationships can be maintained over time, providing trust resilience with dynamic network topology
- Enable trust establishment from partial-strangers via indirect trust and direct observation
- Enables nodes to inform internal processes for global efficiency given observed network behaviour / 'wellness', similar to those found in human social networks eg
  - Update routing table based on 'safest' node chains (Phone Tree)
  - Manoeuvre away from misbehaving nodes (Shunning)
  - Inform as to 'trustworthiness' of forwarded information (Healthy sense of Skepticism)
  - Historic Distrust/Trust decaying over time (Forgiveness/Relationship Decay)

Reason for using TMFs in MANETs

- Provide Risk Mitigation against many classical MANET attacks
  - Black/Grayhole
  - Routing Loop
  - Selective misbehaviour / selfishness
- Generally; to constrain potential malicious behaviour that can operate without detection

Trust in Autonomous Systems

- Public Key Infrastructure - Requires Centralised Control and pre-shared keys
- Resurrecting Duckling - Uses in-action keying with a trusted source
- Evidence Based Trust - Uses shared keys
- Reputation Based Trust - Uses Packet forwarding success rate for prediction of future actions
  - CONFIDANT - Trust-based router implementation using packet forwarding rate
  - Hermes - Bayesian based estimation of trust from successful interactions
  - OTMF - Trust including transitive information from other nodes

- ...and there are plenty more along the same lines
- Predominantly use **single metrics** or **only communications metrics**

Trust in Autonomous Systems

- Public Key Infrastructure - Requires Centralised Control and pre-shared keys
- Resurrecting Duckling - Uses in-action keying with a trusted source
- Evidence Based Trust - Uses shared keys
- Reputation Based Trust - Uses Packet forwarding success rate for prediction of future actions
    - CONFIDANT - Trust-based router implementation using packet forwarding rate
    - Hermes - Bayesian based estimation of trust from successful interactions
    - OTMF - Trust including transitive information from other nodes
    - MTFM - Relationships and Multiple Metrics combined with Gray Interval assessment
- . . . and there are plenty more along the same lines
- Predominantly use **single metrics** or **only communications metrics**

Trust in Constrained Networks

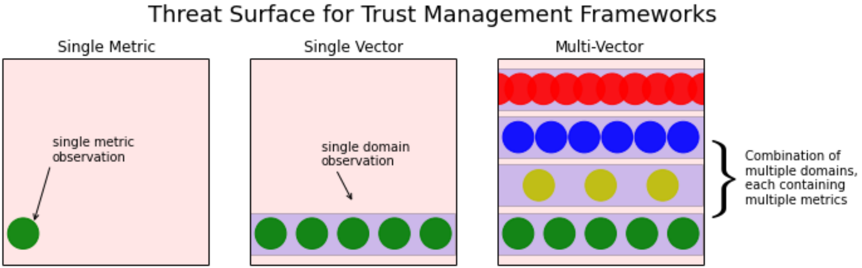
- Most TMF research assumes healthy, stable, low latency channel.
- Interesting work from Chen[2] on DTNs, particularly in optimisation methodologies.
- Usually restricted to state spaces / Markov processes rather than continuous metric assessment.
- Also work from Caiti[3] specifically on UANs, however does rely on PKI-like architecture and does not make Join/Leave protocols clear

Trust in Constrained Networks is still and active and vibrant research area, while still making this work novel in terms of distributed and multi-domain nature.

The Need for Multi-Vector Trust Assessment

- 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.
- Also could provide additional 'handshake' protocols for 'friendly' fleets/teams through reactionary behaviours

Multi-Vector Trust and the Threat Surface



Potential attacks exist across a multi-domain threat surface

Multi-Parameter Trust Assessment for MANETS (MTFM)

- Application of several individual metrics for the construction of a single trust measurement
- For example:
  - $X = \{packet\ loss, signal\ strength, datarate, delay, throughput\}$
- This multi-parameter trust prevents 'smart' attackers; leveraging a known trust metric to subvert a TMF without detection
- Normally expressed as a vector, but can be condensed into an abstracted or weighted form for comparison [4]

Gray (MTFM) Trust Assessment

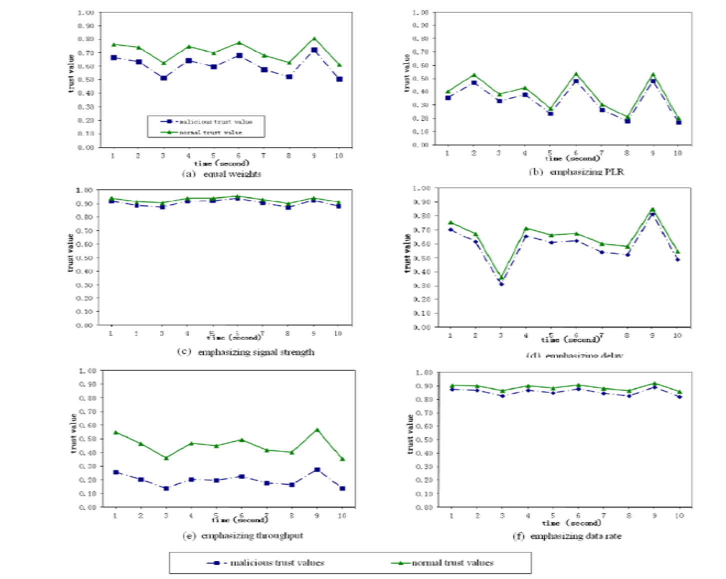
$$[\theta_{k,j}^t, \phi_{k,j}^t] = \frac{\min_k |a_{k,j}^t - r_j^t| + \rho \max_k |a_{k,j}^t - r_j^t|}{|a_{k,j}^t - r_j^t| + \rho \max_k |a_{k,j}^t - r_j^t|}, r \in [g, b] \quad (1)$$

$$[\theta_k^t, \phi_k^t] = \left[\sum_{j=0}^{M} h_j \theta_{k,j}^t, \sum_{j=0}^{M} h_j \phi_{k,j}^t\right] \quad (2)$$

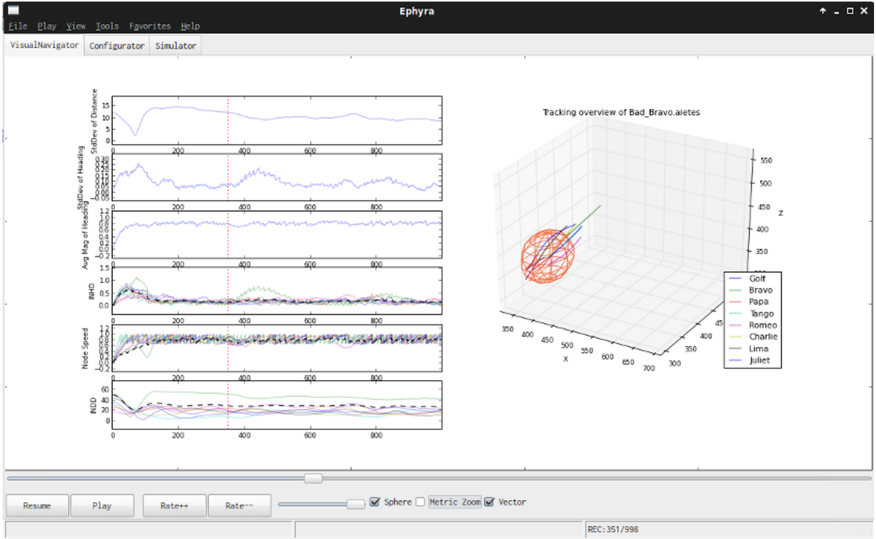
$$T_k^t = (1 + (\phi_k^t)^2/(\theta_k^t)^2)^{-1} \quad (3)$$

where $a_{k,j}^t$ 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. $g_j = \max_k(a_{k,j}^t)$, $b_j = \min_k(a_{k,j}^t)$
These metric coefficients are then accumulated (2) and combined to present a singular trust value for analysis (3).

Malicious Behaviour Discrimination



Agent Based Behaviour Simulator



Operational Mission Profiles

- Flocking with Intent: MCM, Port Protection, Survey, Protection Detail, etc.

Operational Mission Profiles

- Flocking with Intent: MCM, Port Protection, Survey, Protection Detail, etc.
- Metric Selection in collaboration CMRE/DSTL
  - Inter Node Heading Deviation
  - Inter Node Distance Deviation
  - Node Speed

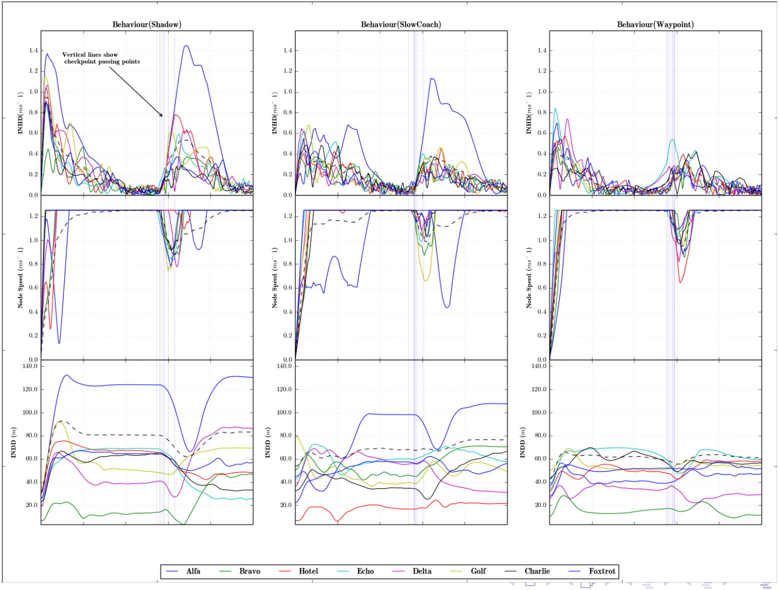
Operational Mission Profiles

- Flocking with Intent: MCM, Port Protection, Survey, Protection Detail, etc.
- Metric Selection in collaboration CMRE/DSTL
  - Inter Node Heading Deviation
  - Inter Node Distance Deviation
  - Node Speed
- Behaviour selection for testing
  - Shadow
  - Spy
  - Sloth
  - Stalker
  - Scoundrel

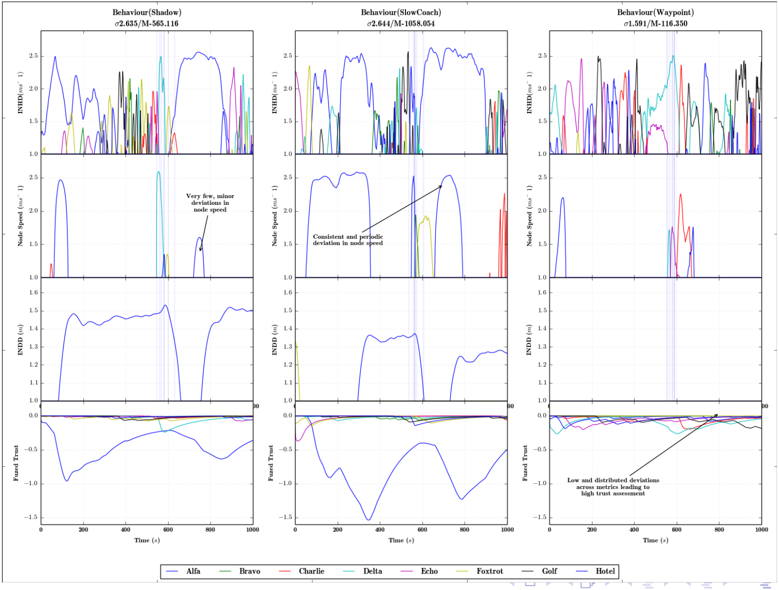
Operational Mission Profiles

- Flocking with Intent: MCM, Port Protection, Survey, Protection Detail, etc.
- Metric Selection in collaboration CMRE/DSTL
  - Inter Node Heading Deviation
  - Inter Node Distance Deviation
  - Node Speed
- Behaviour selection for testing
  - Shadow
  - Spy
  - Sloth
  - Stalker
  - Scoundrel
  - Slow Coach (non-malicious)
  - Spin Doctor (non-malicious)

Raw Behavioural Metric Assessment in AUVs



Behavioural Trust Assessment in AUVs



Behavioural Trust Assessment in AUVs

- Detection and identification based on basic weight-assessment classifier against windowed history of observations, with confidence based on a Grey Theoretic weight
- Currently >96% statistical accuracy of detection and confidence, but this needs more rigorous analysis

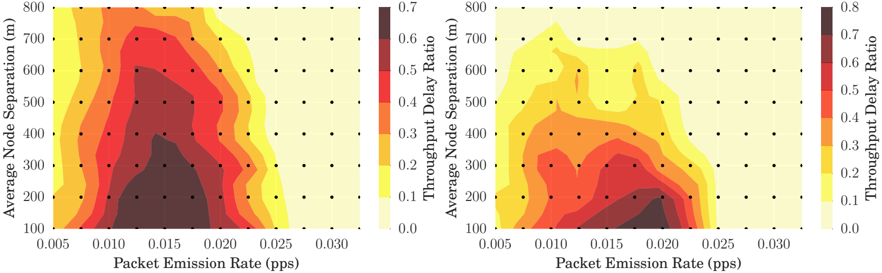
Marine Comms. TMF Performance Assessment

- Acoustic Network based on AUVNetSim [5] and validated against [6].
- Aim to investigate use of 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 (n_1) inflates it's power to all nodes except a target node (n_0) making it appear selfish
- **Selfish Target Selection**(STS) where n_1 preferentially communicates with nodes that are physically near-by, reducing its own power consumption.

Marine Comms. TMF Performance Assessment

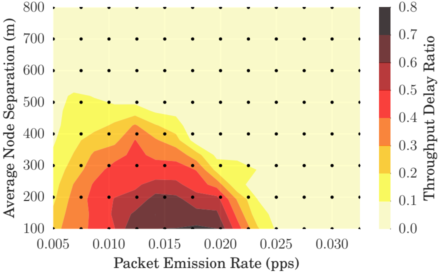


(a) All Static (b) One Mobile

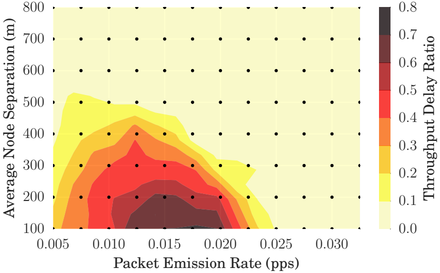
Fig. 1: Performance exploration of a range of packet emission rates and node separations

Selected 0.015pps emission / 300m separation for stable range

Marine Comms. TMF Performance Assessment



(a) All but One Mobile

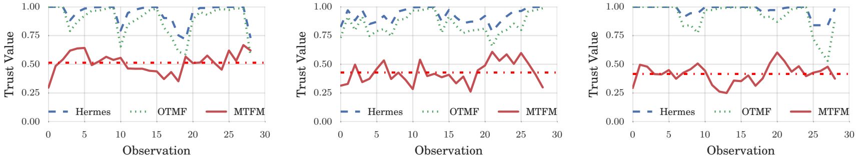


(b) All Mobile

Fig. 2: Performance exploration of a range of packet emission rates and node separations

Selected 0.015pps emission / 300m separation for stable range

Marine Comms. TMF Performance Assessment

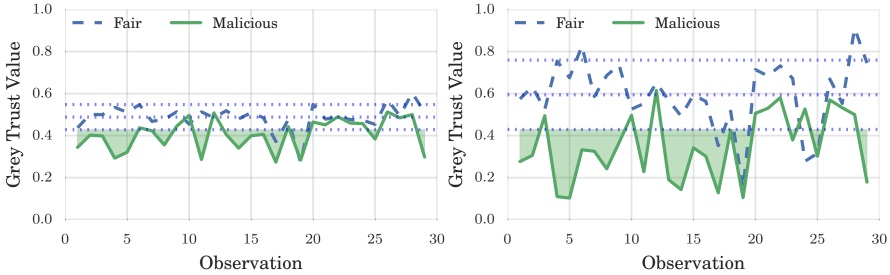


(a) Fair Scenario (b) MPC Scenario (c) STS Scenario

Fig. 3: $T_{1,0}$ for Hermes, OTMF and MTFM assessment values for fair and malicious behaviours in the fully mobile scenario (mean of MTFM also shown)

From 3, 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)

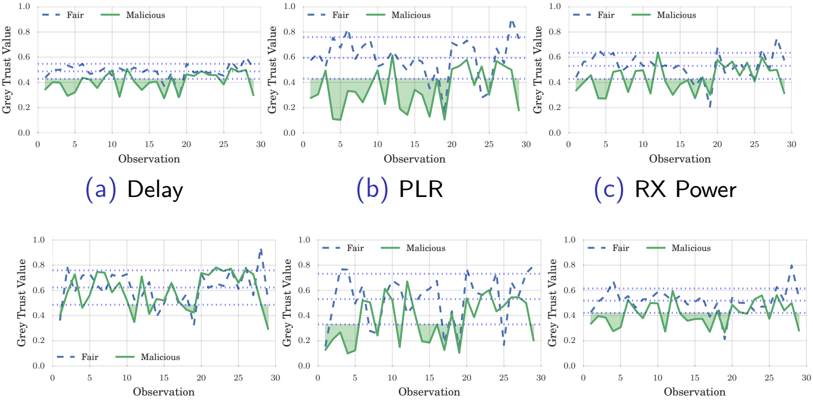
Metric Emphasis and Misbehaviour detectibility: MPC



(a) Delay (b) PLR

Fig. 4: $T_{1,MTFM}$ in the All Mobile case for the MPC behaviour, including dashed $\pm\sigma$ envelope about the fair scenario

Metric Emphasis and Misbehaviour detectibility: MPC

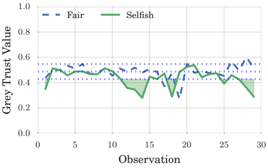


(a) Delay (b) PLR (c) RX Power

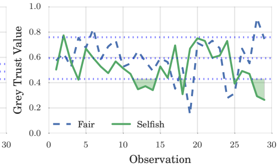
(d) TX Power (e) RX Throughput (f) TX Throughput

Fig. 5: $T_{1,MTFM}$ in the All Mobile case for the MPC behaviour, including dashed $\pm\sigma$ envelope about the fair scenario

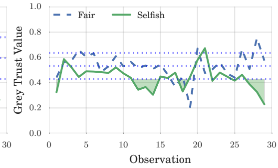
Metric Emphasis and Misbehaviour detectibility: STS



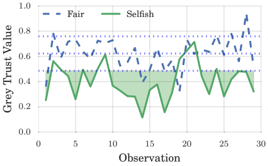
(a) Delay



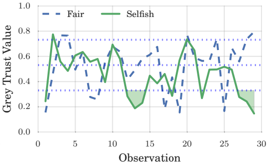
(b) PLR



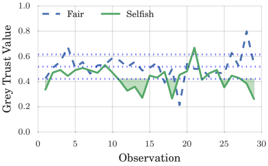
(c) RX Power



(d) TX Power



(e) RX Throughput



(f) TX Throughput

Fig. 6: $T_{1,MTFM}$ 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;

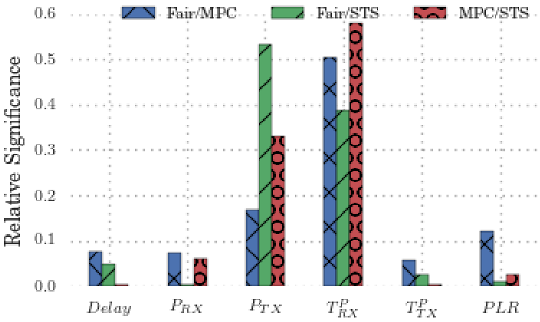


Fig. 7: Random Forest Factor Analysis of Malicious (MPC), Selfish (STS) and Fair behaviours compared against each-other

Weight Significance Analysis for Behaviour Classification

Applying a Random Forest regression tree to 729 different weighting schemes for each of the three behaviours;

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

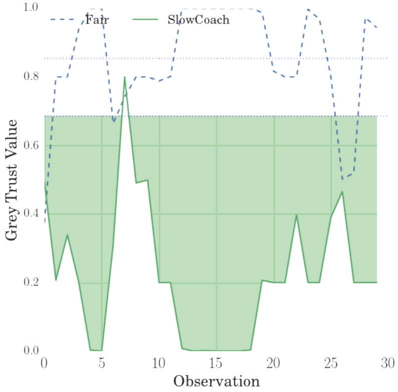
Marine TMF Performance Assessment

Two Physical Misbehaviours were created:

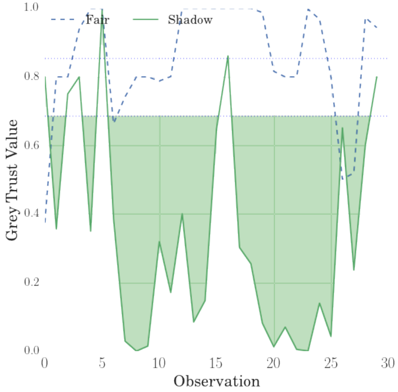
- **Slow Coach**(SC) where a node experiences simulated propulsion damage (lower acceleration)
- **Shadow**(Sh) where a node is following the fleet without appropriate mission knowledge such as the waypoints in a patrol path, modelling an "infected" or masquerading node in the fleet

Similar methodology as applied to comms for metric significance across three metrics (INHD, INDD, Speed)

Raw Grey Assessment Comparison



(a) Slow Coach



(b) Shadow

Fig. 8: $T_{1,MTFM}$ in the All Mobile case for the STS behaviour, including dashed $\pm\sigma$ envelope about the fair scenario

Weight Significance Analysis for Behaviour Classification

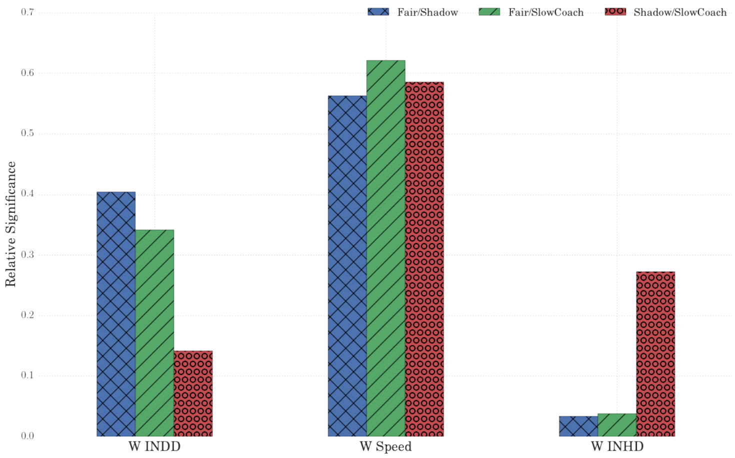


Fig. 9: Applying a Random Forest regression tree to 125 different weighting schemes for each of the three behaviours

Weight Significance Analysis for Behaviour Classification

Applying a Random Forest regression tree to 125 different weighting schemes for each of the three behaviours;

Table 2: Correlation Coefficients between metric weights and behaviour detection targets

Correlation	INDD	Speed	INHD
Fair / SlowCoach	-0.131	0.111	-0.057
Fair / Shadow	-0.194	0.280	-0.214
Shadow / SlowCoach	-0.018	-0.042	0.040

Like the communications results, these demonstrate the ability of Grey Assessment for the detection and differentiation of behaviours.

Remaining Work and Analysis

- Perform Cross and Inter-Domain analysis between Comms and Behavioural Environment and submit for publication to InfoComm 2016
- Extension of MTFM to be asynchronous and report-delay tolerant (back-propagation of delayed messages)

Challenges in Multi-vector Trust

- How to define optimality in trust assessment when dealing with multiple vectors and transitive trust?
- Is there a quantifiable benefit to cross-domain comparison beyond single vector Trust?
- Is there an optimal generic cross-domain comparator?

Current Publications

- A Multi-Vector Trust Framework for Autonomous Systems [7]
  - Symposium paper to the Association for the Advancement of Artificial Intelligence on the current state of work, presenting our progress towards multi-vector trust
- Analysis of Trust Interfaces in Autonomous and Semi-Autonomous Collaborative MHPC Operations [8]
  - Part of a Five-Eyes defence strategy programme (TTCP) for assuring C3I capabilities as part of FF2020
- Single and Multi-Metric Trust Management Frameworks for use in Underwater Autonomous Networks
  - Symposium paper to TrustCom15: RATSP [9]
- Multi-Domain Trust Management Framework for Underwater Autonomous Networks
  - Pending submission to InfoCom15

Thesis Plan I

1. Background Information on 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
2. Background Information on 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 work already accomplished under "Analysis of Trust Interfaces in Autonomous and Semi-Autonomous Collaborative MHPC Operations", including development of representative malicious and abnormal behaviours

Thesis Plan II

3. Strategies for Multi-Domain Trust Assessment
   - Analytical establishment of Multi-Domain Trust, from an information theoretic standpoint.
4. Modelling and Analysis of Collaborative Node Kinematic Behaviours in Underwater Acoustic MANETS
   - Touching on the development of the simulation platform but 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
   - Passing mention of work done in Drift analysis with NPL/Plextek as supporting evidence
5. Comparative Analysis of Multi-Domain Trust Assessment in Collaborative Mobile Networks
6. Investigation into the relative performance characteristics of multi-domain combination strategies in an exemplary context (AUV teams) against existing single and multi metric TMFs

References I

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.

Ing-Ray Chen et al. "Dynamic Trust Management for Delay Tolerant Networks and Its Application to Secure Routing". In: *IEEE Trans. Parallel Distrib. Syst.* 25.5 (2014), pp. 1200–1210. issn: 1045-9219. doi: 10.1109/TPDS.2013.116. url: http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=6497435.

Andrea Caiti et al. "Secure cooperation of autonomous mobile sensors using an underwater acoustic network". In: *Sensors* 12.2 (2012), pp. 1967–1989. issn: 14248220. doi: 10.3390/s120201967.

References II

- 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.

References III

- 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.
- 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*.

The End