PAP Report 2015

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

Andrew Bolster,
Supervised by
Prof. Alan Marshall and Prof. Simon Maskell

May 29, 2015

1 Overview

- **Started:** October 2011 @ Queen's University Belfast
- **Transferred:** October 2013 @ University of Liverpool
- **Target Submission:** November 2015

The project is focused developing a multi-vector trust management framework (TMF) for collaborative mobile autonomous networks (CMANs).

Specifically, this project is looking at the relationships between physical behaviour and communications behaviour within teams of autonomous underwater vehicles (AUVs) for uses related to mine counter measures and port protection for defence, as well as persistent survey behaviour for environmental and petrochemical applications.

This work is undertaken as part of The UK-Fr PhD Programme which is jointly managed by Direction Générale de l'Armement (DGA) and Defence Science and Technology Laboratory (DSTL). It was agreed at the 2010 Anglo-French Summit as one of the ten priorities in 2011 for the Anglo French Defence Research Group (AFDRG). In 2011, five PhDs were funded (two from the UK and three from France), In 2012, the programme grew to nine PhDs (five from the UK and four from France). In 2013 a further ten PhDs were funded (five from the UK and five from France). These PhDs are currently investigating a variety of topics including meta-materials, synthetic biology, sensors, vehicle armour, and human and social sciences as well as this project's work in autonomous underwater vehicles.

2 Summary of Current Outputs

Since the launch of the project, the majority of time has been spent developing a bespoke simulation framework based on Python, developing a variety of "normal", "abnormal" and "malicious" physical and communicative behaviours, as well as developing a range of analytical techniques to detect and identify these behaviours to an extremely high degree of statistical accuracy.

Major outputs and research interactions of the project are:

- Attendance at UComms 2012 (Sestri Levante, Italy)
- Poster Presentations in 2012 (Kassam, Oxford) and 2013 (Heathrow, London)
- 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)
- Rejected submission to The 14th International Conference on Ad Hoc Networks and Wireless (AdHocNow)
- "Single and Multi-Metric Trust Management Frameworks for use in Underwater Autonomous Networks" submitted to The 14th IEEE International Conference on Trust, Security and Privacy in Computing and Communications (IEEE TrustCom-15) (Pending)

3 Field Background

3.1 What is 'Trust'?

"Trust" is a word that gets used a lot in many different ways. Mirriam Webster's Dictionary defines trust as "assured reliance on the character, ability, strength, or truth of someone or something".

This rather broad definition is very attractive to distributed network design as this "trustworthiness" can be used to inform autonomous actors to the "best" courses and paths to action, making them more efficient and resilient in their operation. Within this context, we define trust as "The expectation of an actor performing a certain task or range of tasks within a certain confidence or probability".

In the real world use-case of deployable autonomous systems this trust can take on two real forms; [2]

- Design Trust, where there is an expectation that a system of systems will perform as specified or designed in operation, and
- Operational Trust that the individual systems within a larger system will and are performing as designed in the field. It is this area with which we are particularly concerned.

3.2 Trust Management Frameworks (TMFs)

This desire for operational trust in mobile ad-hoc networks has lead to the development of several Trust Management Frameworks (TMFs), where such frameworks provide information regarding the estimated future states and operations of nodes within such a network. As such, the operation of such frameworks has been summarised as "collecting the information necessary to establish a trust relationship and dynamically monitoring and adjusting the existing trust relationship" [3].

Almost all of the work currently applied to establishing trust in mobile adhoc networks (MANETs) relies either on shared key exchanges with a centralised trust repository (PKI) or with exclusive assessment based on communications behaviour alone, and in that case the vast majority only measure one value e.g. packet loss rate (PLR). Within MANETs, the requirement for distributed trust comes from the decentralised and dynamic communication paths used; if a node moves or an environment changes, the network topology can completely change with paths containing potentially malicious nodes that can disrupt, modify, or reject communications coming from or going to a given node. The motivation for that application of TMFs to MANETs is that by taking historical performance information into accounts, malicious or inefficient actors can be at least detected and routed isolated, preventing further compromise to the distributed network.

These single metric TMFs provide malicious actors with a significant advantage if their activity is undetectable by that one assessed metric, especially if the attacker knows the metric in advance. The objective of operating a TMF is to increase the confidence in, and efficiency of, a system by reducing the amount of undetectable negative operations an attacker can perform. This space of potential attacks can be described as the "Threat Surface". In the case where the attacker can subvert the TMF, the metric under assessment by that TMF does not cover the threat mounted by the attacker. In turn, this causes a highly detrimental effect in the efficiency of the network.

The TMF is assumed to have reduced the threat surface when in fact it has simply made it more advantageous to attack a different part of it. [4] also raised the need for a more expanded view of trust but did so with a domain-partitioning approach rather than combining trust assessments from multiple domains within networks.

3.3 Multi-Parametric Trust Management Frameworks

Guo developed a form of vectorised trust that allowed the combination and cross-correlation of a range of communications observations (Packet loss rate, power and noise characteristics, Delay, Throughput, etc) not only as a combined singular value in the form of a Grey Interval, but also presented the ability to use this vectorised trust assessment to detect and identify malicious or abnormal behaviours through weight assessment perturbations [5].

Guo[5] demonstrated the ability of Grey Relational Analysis (GRA)[6] to normalise and combine disparate traits of a communications link such as instantaneous throughput, received signal strength, etc. into a Grey Relational Coefficient, or a "trust vector".

In the case of the terrestrial communications network used in [5], the observed metric set $X = x_1, \dots, x_M$ representing the measurements taken by each

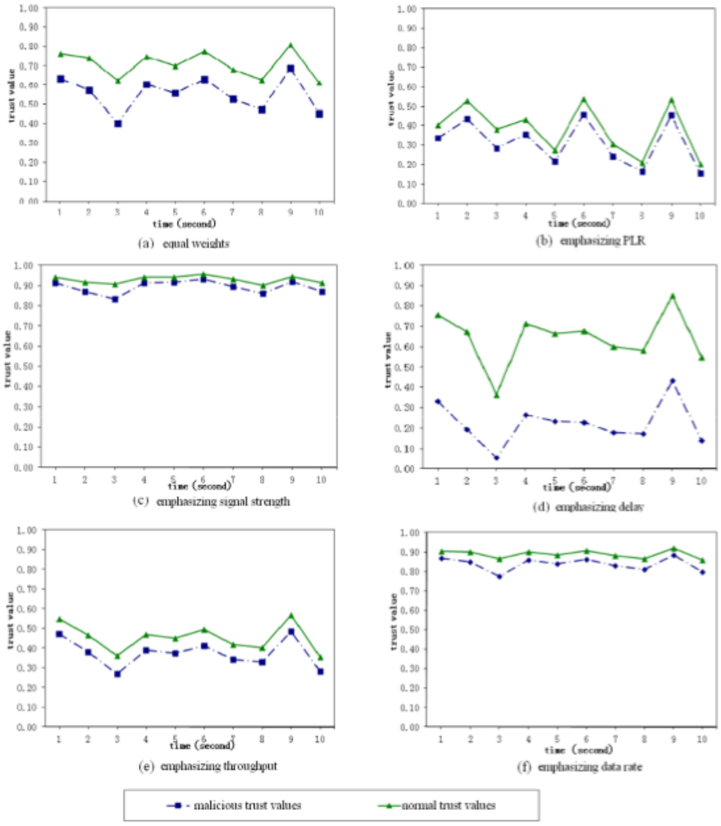


Figure 1: Example of weight-based attack classification taken from (Guo, Marshall & Zhou 2011)

node of its neighbours at least interval, is defined as X = [packet loss rate, signal strength, data rate, delay, throughput]. The trust vector is given as

$$\theta^t_{k,j} = \frac{\min_k |a^t_{k,j} - g^t_j| + \rho \max_k |a^t_{k,j} - g^t_j|}{|a^t_{k,j} - g^t_j| + \rho \max_k |a^t_{k,j} - g^t_j|} \tag{1}$$

$$\phi^t_{k,j} = \frac{\min_k |a^t_{k,j} - b^t_j| + \rho \max_k |a^t_{k,j} - b^t_j|}{|a^t_{k,j} - b^t_j| + \rho \max_k |a^t_{k,j} - b^t_j|}$$

where $a^t_{k,j}$ is the value of a 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 $\{a^t_{k,j}k = 1, 2 \ldots K\}$, e.g. $g_j = \max_k(a^t_{k,j})$, $b_j = \min_k(a^t_{k,j})$ (where each metric is selected to be monotonically positive for trust assessment, e.g. higher throughput is always better).

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

Where $H = [h_0 \dots h_M]$ is a metric weighting vector such that $\sum h_j = 1$, and in unweighted case, $H = [\frac{1}{M}, \frac{1}{M} \dots \frac{1}{M}]$. θ and ϕ are then scaled to $[0, 1]$ using the mapping $y = 1.5x - 0.5$. To minimise the uncertainties of belonging to either best (g) or worst (b) sequences in (1) the $[\theta, \phi]$ values are reduced into a scalar trust value by $T^t_k = (1 + (\phi^t_k)^2/(\theta^t_k)^2)^{-1}$ [**Hong2010**].

For applications involving low fidelity, temporally sparse metrics with unknown statistical distributions, this Grey Relational Analysis is a more stable comparative analysis, providing an interval of potential trust values rather than fuzzy-logic or the Bayesian-Beta distributions found in current TMFs [7].

However, this still ignores the usefulness of physical behaviours in the development of trust assessments, which became the basis for this project.

4 Project Objectives and Novelty

The aims of this project are to extend this "vectorised" methodology to not only combine individual trust metrics (observations) into trust vectors, but to combine such vectorised assessments of trust across "domains" (i.e. combining trust assessments based on observations in the communications domain with trust assessments based on observations in the physical domain).

The use of metrics from other domains is particularly relevant in CMANs where several autonomous systems interact and affect each-others behaviour, making classical constraint based testing or monitoring inappropriate.

In single metric trust, such as the use of packet loss rate in Hermes[8], and Optimal Trust Management Framework(OTMF)[9], this only monitors a relatively small area of the potential attack space. Combining several domain specific metrics, in this case in communication, provides not only the capability to detect a much wider range of attack types, but also to discriminate between them, potentially disclosing of the tactics of an attacker.

Our hope is that by using multiple domains together, we can provide a higher level, strategic point of view on an attacker or attackers, enabling the generation of preventative and reactive strategies to defend against them. This would leave even the most knowledgeable attacker with no option but to behave properly to avoid detections, eliminating any potential reward they could achieve without detection.

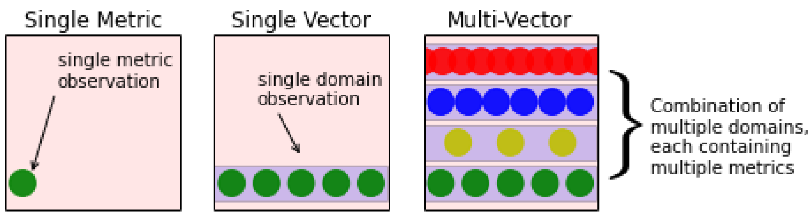


Figure 2: With the addition of further metrics, the threat surface available to an attacker is greatly reduced

Our main goal in this project has been to investigate if the methodologies that Guo applied to communications in the conventional, terrestrial, MANET space, can equally be applied to physical as well as communications behaviour in te Underwater CMAN environment. Our approach to this has been to develop an agent based simulation platform built in Python, emulating both the physical and communicative environments for teams of collaborating nodes performing some task, such as mine countermeasure (MCM) survey, port protection, mothership protection and others.

This interaction in the physical world opens up a range of metrics that can be used for trust assessment, in an effort to further restrict undetected malicious behaviour on the exposed threat surface of a network.

In summary, the novelties of the work accomplished so far are:

- A fused Trust Assessment using Physical Behaviour in collaborative mobile autonomous networks (CMAN)
- An analytical comparison of Marine and Terrestrial Communicative trust environments
- Exploration and Demonstration of the advantages of multiple metric trust over single metric trust (in underwater acoustic communications)
- A defined protocol for quantitative assessment of Metric suitability through comparative regression analysis

Additional novelties to be investigated prior to submission:

- Definition of asynchronous trust assessment with back-propogation (required to deal with both multi-domain and time delay factors in reporting)
- Demonstration of advantage between single domain (MTFM) and multi domain trust.

5 Current Progress and Results

5.1 Physical Behaviours

Through collaboration NATO's Centre for Maritime Research and Experimentation (CMRE) based in Italy, and the UK's Defence Science and Technology Laboratory (DSTL), we arrived upon a range of observable metrics based on the positions and attitudes of other nodes in the group. These were;

- The Inter Node Heading Deviation (INHD), i.e. the deviation in heading from a local group average,
- The Inter Node Distance Deviation (INDD), i.e. the deviation of a given node's position with respect to the average node spacing across the rest of the local group,
- and, The Node's Absolute Speed.

Along with metric selection, we also arrived at a collection of both malicious and non-malicious misbehaviours which were simulated in the framework. These were:

- **The Shadow**, where a node is following the fleet without appropriate mission knowledge such as the waypoints in a patrol path, modeling an 'infected' or masquerading node in the fleet
- **The Spy**, where a node consistently or intermittently rises to the top of the fleet, potentially surfacing to relay mission information to an unauthorised third party via a back-channel communications channel such as RF
- **The Sloth**, where a node is selfishly conserving energy by not making complete patrol paths, consistently taking a minimal path around the tolerances of the waypoint path or some other selfish behaviour.
- **The Stalker**, where a node in a multi-node system preferentially 'tails' a given node above others
- **The Scoundrel**, where a node falsely reports it's position or velocity with the intention to corrupt any collaborative positioning system
- **The Slow Coach**, where a node is operating correctly but has a defect in it's power train causing reduced maneuverability, i.e. a runtime defect in an otherwise good node
- **The Spin Doctor**, likewise is operating correctly but it's Inertial Navigation System is damaged and consistently bears left/right by some level.

Of these "Seven CMAN Sins" (S^7), we've directly considered Shadow and Slow Coach behaviours, as this pairing presents a useful side effect of trust based on physical behaviour; the ability for a network to self-test its health by discerning the difference between broken and malicious behaviour.

In a simple port protection scenario where nodes are patrolling around a series of waypoints around an area we can demonstrate this detective and selective power.

Each vertical of Fig. 5.1 shows a different behaviour; with the baseline waypointing behaviour in the middle, flanked on the left by the malicious "Shadow" behaviour and on the right by the benign but sub-optimal "Slowcoach" behaviour. The per-node metrics as described are shown in the horizontal, with Internode Heading Deviation at the top, followed by Absolute Speed and Distance Deviation in the middle and finally a combined trust 'vectorised' (dis)trust assessment in the bottom, based on a decaying windowed average of instantaneous trust assessments across observation types.

With figure chart it's shown that the Slowcoach behaviour manifests itself in the Speed metric more so than the Shadow behaviour. Using this information we can build a trust weight vector to discriminate between these two ostensibly similar behaviours.

With our current analysis we can do so with an average 96% positive identification rate with a 2% false positive rate, e.g. instances where a Slowcoach is confidently detected as a Shadow and visa versa. It should be noted that in this analysis, there were no cases where the malicious behaviour (Shadow) was misdirected as the purely abnormal/faulty behaviour (Slow Coach).

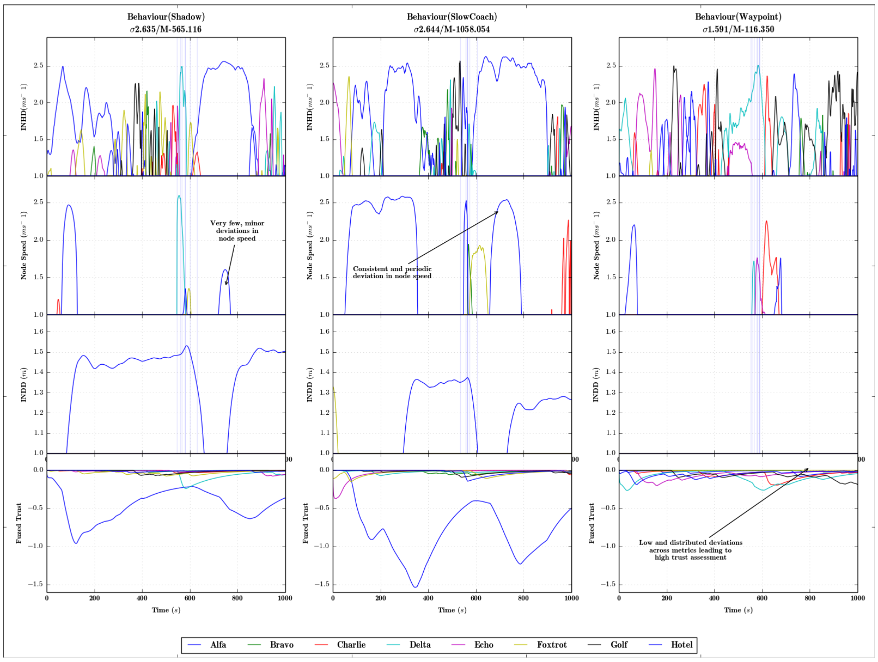


Figure 3: Charts showing deviation activation curves for three different simulated behaviours, with a fused trust value in the bottom row.

5.2 Communications Behaviours

The major amount of recent work has been concerned with the development and testing of simulated acoustic communications channels. Supposedly feature-complete implementations such as [10] were found to be incomplete and not totally tested, and this put back progress by several months.

However, we were able to incorporate our experiences under the DSTL/NPL/Plextek collaboration project (CDE 33135) to reimplement and improve both the routing and physical modelling used in the AUVNetSim framework. These implementations were validated against [11] and [12].

Having then implemented Multi-parameter Trust Framework for MANETS (MTFM)[5], Hermes[8], and Optimal Trust Management Framework(OTMF)[9], exploratory simulations were run to establish a comparable operation range for the simulated underwater context.

Given the differences in delay and propagation between RF and marine networks, (Table 2) it would not be expected that the same application rates (e.g. packet emission rates or throughput) and node separations are equally stable in this environment.

To establish a safe operating zone in terms of packet rate and node separation, four mobility scenarios were implemented on a six-node CMAN and repeatedly simulated to assess the total network throughput and end-to-end delay characteristics of the networks as a whole. These results are summarised in Figs 4 and 5, showing the throughput-delay ratio across the packet rate - node separation parameter space.

As the separation is increased, the emission rate at which the network be-

Table 1: Identification results from 384 simulated runs (i.e. ++ indicating correct and confident identification, and -+ meaning an incorrect confident identification)

Scenario	Correct ++	Correct +-	Incorrect -+	Incorrect −
Shadow	124	4	0	0
SlowCoach	111	5	3	9
Waypoint	NA	NA	8	120

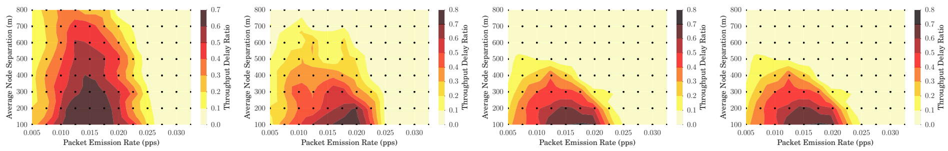
Table 2: Comparison of system model constraints as applied between Terrestrial and Marine communications

Parameter	Unit	Terrestrial	Marine
Simulated Duration	s	300	18000
Trust Sampling Period	s	1	600
Simulated Area	km^2	0.7	0.7-4
Transmission Range	km	0.25	1.5
Physical Layer		RF(802.11)	Acoustic
Propagation Speed	m/s	3×10^8	1490
Center Frequency	Hz	2.6×10^9	2×10^4
Bandwidth	Hz	22×10^6	1×10^4
MAC Type		CSMA/DCF	CSMA/CA
Routing Protocol		DSDV	FBR
Max Speed	ms^{-1}	5	1.5
Max Data Rate	bps	5×10^6	≈ 240
Packet Size	bits	4096	9600
Single Transmission Duration	s	10	32
Single Transmission Size	bits	10^7	9600

comes saturated decreases, reducing overall throughput. This throughput degradation is tightly coupled with the mobility, as increasing mobility leads to increasing delays as routes are constantly broken, re-advertised and re-established. For instance, where all nodes are static, we do not see significant drops in saturation rates until node separation approaches 800m, nearly double the initial estimate. When all nodes are randomly walking the saturation point collapses from 0.025pps at 300m to 0.015pps at 400m.

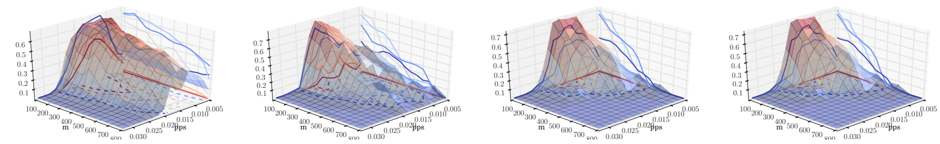
Our results indicate that the best area to continue operating in for a range of node separations is at 0.015pps, and that a reasonable position scaling is from 100m to 300m, beyond which communication becomes increasingly unstable, especially in terms of end-to-end delay. These results are similar to work performed in [10], and are expected in such a sparse, noisy, and contentious environment.

Guo et al. introduced a range of misbehaviours, including modification of



(a) All Static (b) One Mobile (c) All but One Mobile (d) All Mobile

Figure 4: Performance exploration of a range of packet emission rates and node separations



(a) All Static (b) One Mobile (c) All but One Mobile (d) All Mobile

Figure 5: 3D versions of Fig. 4

the packet loss rate of routing nodes and limiting throughput on a per-link basis as well as a selection of combined misbehaviours.

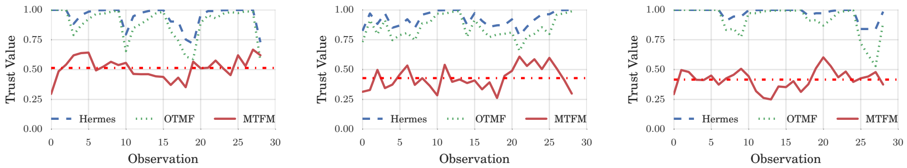
Building upon and extending these misbehaviours, we develop two comparative misbehaviours to initially investigate;

1. Malicious Power Control (MPC), where n_1 increases its transmit and forwarding power by 20% for all nodes *except* communications from n_0 in order to make n_0 appear to be selfishly conserving energy to the rest of the team, while n_1 itself appears to be performing very well.

2. Selfish Target Selection (STS), where n_1 preferentially communicates, forwards and advertises to nodes that are physically close to it in effort to reduce its own power consumption.

As per [5], "fair" scenarios were also performed with no malicious behaviour, applying OTMF and Hermes assessment as well as MTFM, providing like-for-like comparison of assessment. For simplicity of presentation, we only discuss the results from the fully-mobile scenario, as we are concerned with the establishment of trust in mobile networks

Fig. 6 shows a comparison between the unweighted response of MTFM compared to OTMF and Hermes assessment functions on the same data for the fair, malicious and selfish behaviours respectively. It is important to note a distinction between the expectations of MTFM compared to other TMFs; MTFM is primarily concerned with the identification of differences in the behaviours of nodes in a network, and is relative rather than absolute. That is to say that under MTFM, nodes are compared against the worst current performances across metrics of other observed nodes and graded against them, rather than the absolute (objective) approach taken by many TMFs. In these cases, particularly since the methods of attack were not directly related to PLR, OTMF and Hermes have not registered significant activity in either misbehaviour when compared to the fair scenario. The difference between the MTFM trust assessments under "fair" and "malicious" behaviour is lowered by $\approx 10\%$ in both



(a) Fair Scenario

(b) Malicious Power Control Scenario

(c) Selfish Target Selection Scenario

Figure 6: $T_{1,0}$ for Hermes, OTMF and MTFM assessment values for fair and malicious behaviours in the fully mobile scenario (mean of MTFM also shown)

cases, in terms of the mean values returned. At run time, similar results could be attained by an exponentially weighted moving average filter (EWMA).

On their own, neither OTMF, Hermes, nor unbiased MTFM are effective in detecting or identifying malicious behaviour in this environment, in fact OTMF and Hermes don't appear to differentiate between fair and selfish scenarios at all.

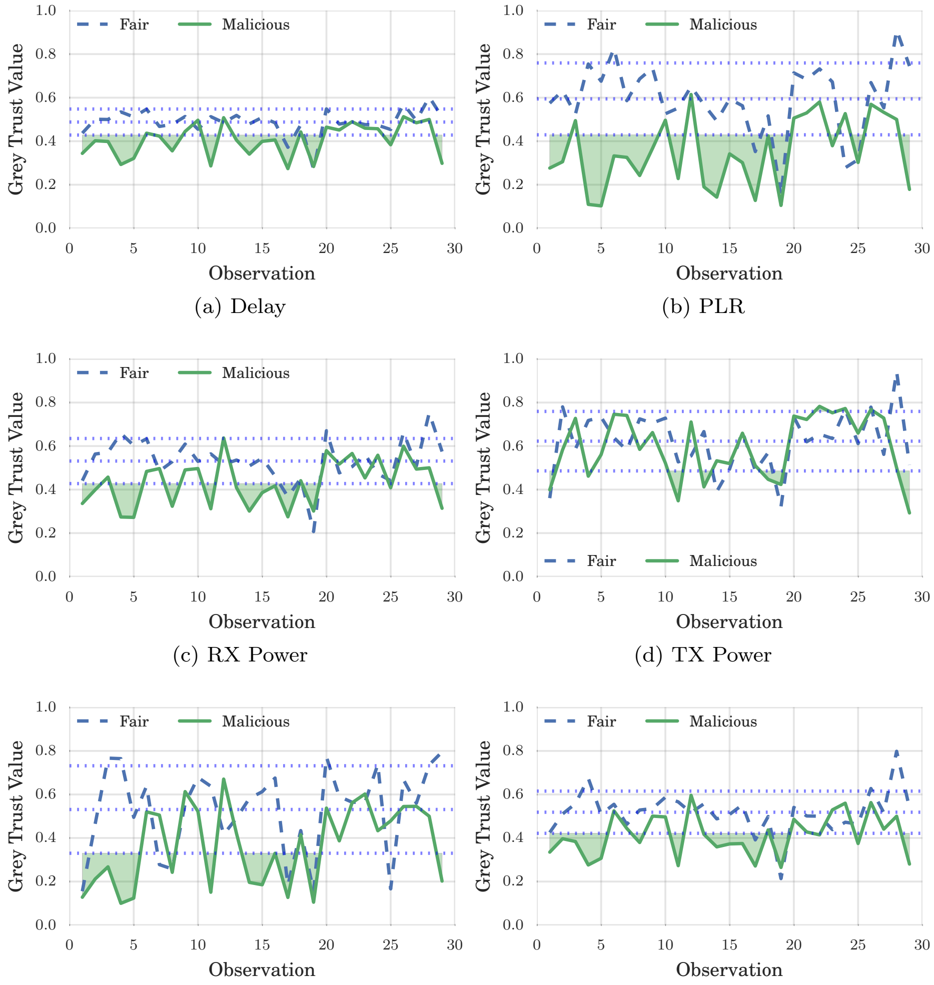
However, with MTFM we can apply a sequence of vectors that preferentially weight each metric in (2) to each of the three simulation runs. For a metric weight vector H, where the metric m_j is emphasised as being twice as important as the other metrics, we form an initial weighting vector $H' = [h_i...h_M]$ such that $h_i = 1 \forall i \neq j; h_j = 2$. We then scale that vector H' such that $\sum H = 1$ by $H = \frac{H'}{\sum H'}$. Using this process we can extract and highlight the primary aspects of an attack by comparing against the deviation from the "fair" result set.

In the malicious case (MPC) (Fig. 7) we can see that the malicious node is consistently outside the $\pm\sigma$ (one standard deviation above and below the mean) envelope of the fair scenario it's being compared to. This is particularly true for PLR, with smaller impacts on delay, received power and transmitted throughput. However, this weighted delta in received throughput is minimal to insignificant compared to the width of the detection envelope, occasionally breaching the envelope for a short period.

In the selfish case (STS) (Fig. 8) we observe much lower weighted delta in PLR and delay, with greatly increased impact on transmission power. In comparison to [5], these results are qualitatively similar, however here the differences between the fair case and the misbehaviours are less clear than in the comparable terrestrial space. Guo et al. show similar types of behaviour but report a weighted delta from ≈ 0.4 to ≈ 0.9 across the simulation period, compared to our maximum delta in TX Power of ≈ 0.3 for an inconsistent interval (Fig. 8d.)

To further investigate the relative importance of metric selection and weighting, we apply a random forest regression methodology. In this process, 729 different weighting vectors, variably emphasising all combinates of metrics and varibly weighting different metrics within each vector, are applied to the results set. As above, the area between the resultant weighted MTFM assessment outlying the equivalent fair case is selected as the target of the regression, i.e. we are interested in the metrics that provide the largest deviation between behaviours, allowing us to select and identify them.

Applying this target regression reveals that each of the three combina-



(a) Delay

(b) PLR

(c) RX Power

(d) TX Power

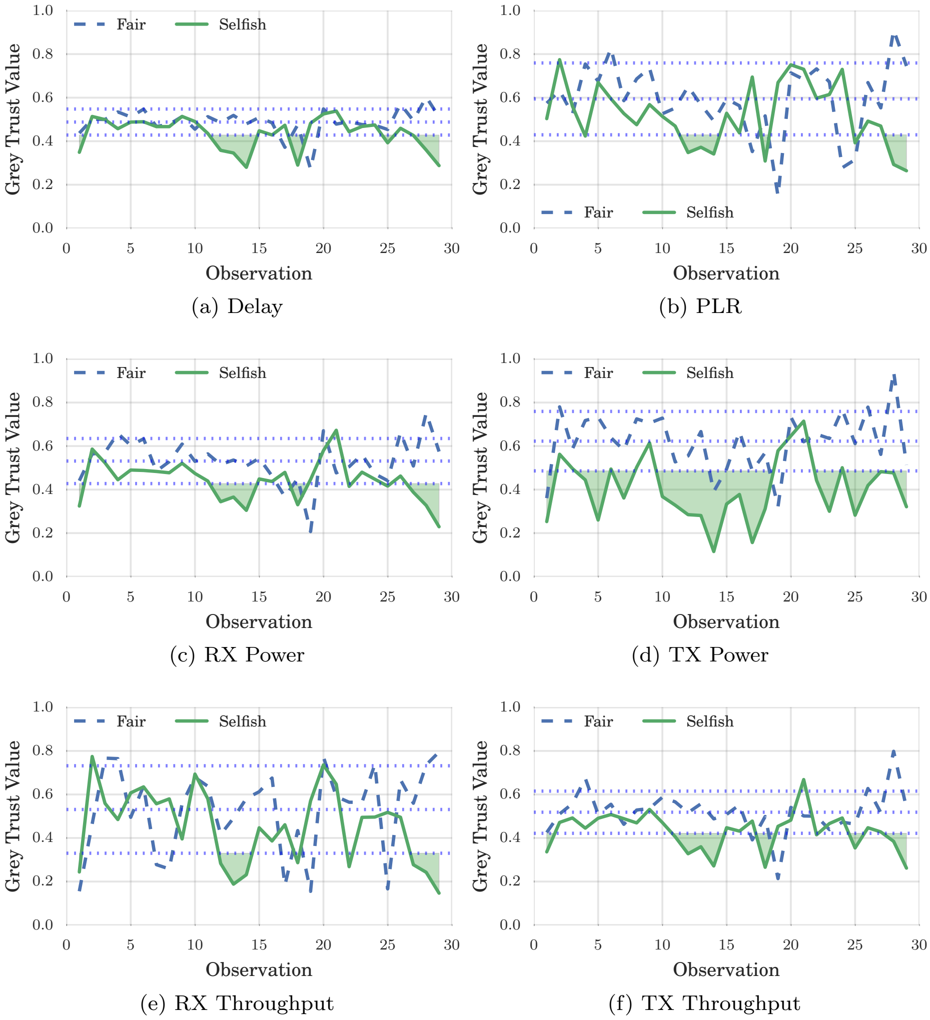
(e) RX Throughput

(f) TX Throughput

Figure 7: $T_{1,MTFM}$ in the All Mobile case for the MPC behaviour, including dashed $\pm\sigma$ envelope about the fair scenario

tions (i.e. comparing fair to misbehaviours, and comparing the misbehaviours) present distinct patterns of significance in three primary metrics; received throughput, transmitted power, and PLR, with delay, received power and transmitted throughput playing a lesser role (Fig. 9)) Practically this means that in order to accurately distinguish between these scenarios, these primary metrics should be higher-weighted in the generation of $T_{1,MTFM}$ in (2).

It may initially appear odd that the relative significance of the received throughput is similar between all three scenario combinations, however a correlation analysis (Table. 3)shows that in the MPC attack; the received throughput is positively correlated with successful classification against the fair case ($R = +0.71, p \approx 10^{-100}$), while the inverse is the case for the STS attack ($R = -0.70, p \approx 10^{-100}$). It is expected that Transmitted power should be the



(a) Delay

(b) PLR

(c) RX Power

(d) TX Power

(e) RX Throughput

(f) TX Throughput

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

defining characteristic of STS ($R = +0.72, p < 10^{-100}$) as the node is acting fairly from a protocol perspective but is acting unfairly at a higher (incentive) level; it is performing fairly in terms of it's communications with other nodes, however it is preferring to communicate with nodes that it can expend less energy communicating with.

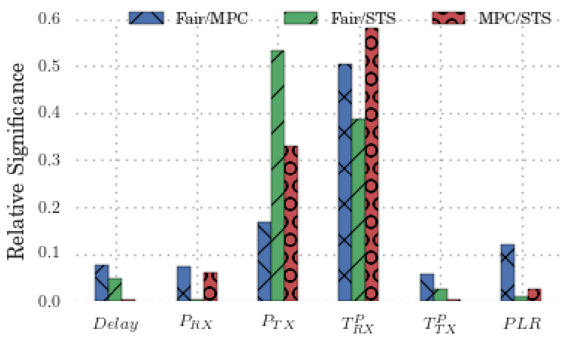


Figure 9: Random Forest Factor Analysis of Malicious (MPC), Selfish (STS) and Fair behaviours compared against each-other

Table 3: 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

5.3 Discussion of Current Results

From these results we are able to conclude several things;

- Current MANET trust assessment methods are not immediately applicable to the marine environment.
- While PLR has some behaviour discriminating performance, it is unmatched compared to other metrics, and it is much more effective to use several metrics together in the noisy and sparse communications environment.
- Behavioural metrics, like communications metrics, can be used to detect and characterise abnormal behaviours.

6 Changes to Assumptions

Over the course of the project, initial assumptions made about the nature, operation, and context of the project have changed wildly. This is largely due to the increased access to current and planned operational data from the Royal Navy while at DSTL Portsdown West, working with both the Naval Systems and Information Systems teams specifically on the area of maritime autonomy and the use of increasingly decentralised and autonomous systems for marine survey and observation. This work also contributed to a transnational collaboration looking at the establishment of systematic trust in autonomous systems.

Likewise, our work with Plextek and NPL on collaborative positioning using high precision timing was extremely helpful in developing my awareness not just of positioning techniques used by current platforms but also increased knowledge and time to generate simulation extensions to more accurately model not only the errors in these positioning systems but also algorithms for estimating the propagation time of acoustic comms.

Initial Assumption	**Correction**
AUV operation in open water	AUVs (currently) mainly operate in coastal/littoral/riverbed areas with highly dynamic SSP (Speed of Sound Profile) and environmental variations. Shallow environmental profile
AUV sensing dynamically adjusts to the environment and motion	SSS (Side Scanning Sonar) is extremely sensitive to not only height but strafe drift and can corrupt entire missions by being out of alignment more than 8-10m horizontally and 2-4m vertically (depending on initial configuration) SAR (Synthetic Aperture Sonar), which builds up acoustic profiles over several “pings” is slightly more tolerant but any variation in path massively increases computational complexity of imagery (Possibly a “Moore’s Law” Problem)
AUV positioning is equally bad in all dimensions due to loss of GNSS, relying on INS, with cumulative errors on the order of tens of metres per hour of operation	Depth sensing is extremely accurate INS errors are directionally cumulative based on a electro/gyroscopic bias in the device. This means that if you travel two hours in one direction and two hours back in the opposite direction, your cumulative INS (Gyroscopic) error will be minimal Additionally, the use of bottom tracking DVL (Doppler Velocity Log) provides quite good instantaneous velocity (through water) tracking but with normal errors. These errors are relative to the height above floor and cumulative with time. All in all for most cases, drift can be characterised at around half the expected rate (5/hr) Simulation Model Characterisations have been generated for both Gyro and DVL error statistics Further, collaborative measures can be taken to normalise these errors and consistently improve position accuracy by up to 40%. This collaborative system has also been implemented into the simulation framework, and directly leads into the Scoundrel malicious behaviour model.

Initial Assumption	Correction
— Nodes operate with a spherical distribution (i.e. vertically and horizontally distributed)	All of these together lead to the the improved assumption that nodes operate in a flat arrangement, maintaining a known altitude for tracking accuracy, DVL consistency, SSS/SAR resolution and SSP stability.
Acoustic Comms is terrible and curved	This assumption was validated and expanded upon, with the generation of a simplified “Bellhop” simulation model generated to reasonably estimate the time of flight characteristics (but not ISI) of acoustic comms using ray tracing. Current bitrates for <500m ranges are around the 100kbps scale, providing an upper limit on comms bandwidth for inter-node collaboration
AUV operations will be mostly isolated	Currently this is not the case, with AUV MHPC operations mainly based on either shore-based teams or Hunt/Sandown class based surface vessels in close proximity to operational area. However BLOS (beyond line of sight) AUV operation is part of the MoD’s Future Force 2020 programme so it is reasonable to continue operating as if that is the case.
AUVs will be “weaponless”	While there are direct-disposal tethered ROVs fitted with explosive payloads, this assumption has largely borne out; the biggest blockade to having weaponised AUVs is legal trust in the security of operation of such an autonomous system, making it extremely unlikely that decentralised AUVs will ever be legally weaponised.

7 Development, Publication Plans

7.1 Behaviour Detection

Current work is concentrated on the improvement of the analysis of behaviours, with the intention of submission for publication of a paper on the threat detection methodology. This was intended to be completed last year but has fallen by the way-side in favour of the communications comparison work package.

7.2 Multi-Domain Trust Assessment

All remaining focus is on the cross domain implementation of trust, which is, combining communicative and behavioural trust assessments.

Initially, It is expected that this work will be largely analytical, maintaining context-free analysis for multi-domain trust but utilising AUV MANETs as an exemplar implementation for critical assessment of proposed cross-domain combination strategies.

Ideally this work will go on as a Journal paper into potentially IEEE Comms, Dependable and Secure Computing, or Intelligent Systems Trans.

7.3 Reactionary/Perturbative Trust

It was hoped that some time be spent investigating reactionary behaviours, both communicative and physical, to allow a distributed, sparsely connected, decentralised team of nodes to dynamically “interrogate” or “test” the trustworthiness of nodes within the team. Given the compressed timeframe, this is unlikely.

8 Proposed Thesis Chapter Titles

1. Trust and its applications to MANETs
    - Discussion on abstract analysis of MANETS, trust networks and trust frameworks
    - Introduction to Trust Management Frameworks and their benefits
    - Discussion on the threat surface of Mobile Ad Hoc Networks and how that has been protected so far
2. Maritime Uses of Autonomous Systems
    - Review of the Underwater Communications Environment
    - 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)
    - (Minor) Discussion of the contextual human factors around integrating autonomous systems into existing human-based solutions.
3. Strategies for Multi-Domain Trust Assessment
    - Information Theory approach to single and multiple trust metrics including Bayesian and Grey backgrounds.
    - Analytical establishment of Multi-Domain Trust, from an information theoretic standpoint.
4. Trust from 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. Multi-Metric Trust Assessment of Communications in Collaborative Mobile Networks
    - Primarily focused on work submitted to INFOCOM
6. Performance analysis of Multi-Domain Trust Assessment in Marine context

- Comparative assessment of multi-domain attack and defense operational performance

7. Conclusions and Future Work

References

[1] 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.

[2] Andrew Bolster. *Analysis of Trust Interfaces in Autonomous and Semi-Autonomous Collaborative MHPC Operations.* Tech. rep. The Technical Cooperation Program, 2014.

[3] 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.

[4] Dijiang Huang, Xiaoyan Hong, and M Gerla. "Situation-aware trust architecture for vehicular networks". In: *Commun. Mag. IEEE* November (2010), pp. 128–135. URL: http://ieeexplore.ieee.org/xpls/abs%5C_all.jsp?arnumber=5621979.

[5] 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.

[6] Fengchao Zuo. "Determining Method for Grey Relational Distinguished Coefficient". In: *SIGICE Bull.* 20.3 (Jan. 1995), pp. 22–28. ISSN: 0893-2875. DOI: 10.1145/202081.202086. URL: http://doi.acm.org/10.1145/202081.202086.

[7] K J R Liu. "Information theoretic framework of trust modeling and evaluation for ad hoc networks". In: *IEEE J. Sel. Areas Commun.* 24.2 (2006), pp. 305–317. ISSN: 07338716. DOI: 10.1109/JSAC.2005.861389. URL: http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=1589110.

[8] Charikleia Zouridaki et al. "E-Hermes: A robust cooperative trust establishment scheme for mobile ad hoc networks". In: *Ad Hoc Networks* 7.6 (2009), pp. 1156–1168. ISSN: 15708705. DOI: 10.1016/j.adhoc.2008.10.003. URL: http://dx.doi.org/10.1016/j.adhoc.2008.10.003.

[9] Jie Li et al. "Future Trust Management Framework for Mobile Ad Hoc Networks". In: *IEEE Commun. Mag.* 46.4 (Apr. 2007), pp. 108–114. ISSN: 01636804. DOI: 10.1109/MCOM.2008.4481349. URL: http://ieeexplore.ieee.org/xpls/abs%5C_all.jsp?arnumber=4212452.

[10] 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.

[11] Milica Stojanovic. *On the relationship between capacity and distance in an underwater acoustic communication channel.* 2007. DOI: 10.1145/1347364.1347373. URL: http://www.mit.edu/~millitsa/resources/pdfs/bwdx.pdf.

[12] 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.