# Multi-Domain Trust Frameworks for Harsh Environments

## Bringing Physical and Communications observations together for effective trust management

Andrew Bolster
Department of Electrical Engineering and Electronics
University of Liverpool
Liverpool, UK
andrew.bolster@liv.ac.uk

Alan Marshall
Department of Electrical Engineering and Electronics
University of Liverpool
Liverpool, UK
alan.marshall@liv.ac.uk

***Abstract*—In this paper we demonstrate the use and operation of a multi-domain trust management framework (MD-TMF) for collaborative mobile autonomous networks (CMANs), using simulated underwater autonomous networks (UANs) as an exemplar application of a resource constrained, sparsely communicating, delay-tolerant, cyber-physical system. We also present a machine learning methodology for assessing the relative and collective performance of varying metric sets and subsets in detection and differentiation of a range of communications and physical misbehaviors.**

***Keywords—component; formatting; style; styling; insert*** *(key words)*

## I. INTRODUCTION

With the increasing application of autonomy in cyber-physical systems, Trust Management Frameworks (TMFs) are increasingly being applied to assist the efficiency, security, and reliability of decentralised and distributed autonomous systems, from highway-bound autonomous vehicles to aerial battlefield drones.

Classical applications of trust management in Mobile Ad-Hoc Networks (MANETs) have focused solely on observations from the communications domain to make trust assessments. However, these methods are not as effective in applications exhibiting sparse, delayed, or otherwise challenged communications environments\cite{Pavan2015}. MD-TMF expands this paradigm to include relevant physical factors and movements to increase the threat area covered the trust framework. An example area of application is the underwater marine environment, where extreme challenges to communications present themselves (propagation delays, frequency dependent attenuation, fast and slow fading, refractive multi-path distortion, etc.).

In addition to the communications challenges, other considerations such as command and control isolation, as well as power and locomotive limitations, drive towards the use of teams of smaller and cheaper autonomous underwater vehicles (AUVs). These increasingly decentralised applications present unique threats against trust management. In underwater environments, communications is both sparse and noisy.

Therefore the observations about the communications processes that are used to generate the trust metrics, occur much less frequently, with much greater error (noise) and delay than is experienced in terrestrial RF MANETS.

Trust Management Frameworks (TMFs) provide information to assist the estimation of future states and actions of nodes within networks. This information is used to optimize the performance of a network against malicious, selfish, or defective misbehaviour by one or more nodes. Previous research has established the advantages of implementing TMFs in 802.11 based MANETs, particularly in terms of preventing selfish operation in collaborative systems \cite{Li2007}, and maintaining throughput in the presence of malicious actors \cite{Buchegger2002}. Most current TMFs use a single type of observed action to derive trust values, typically successfully delivered or forwarded packets. These observations then inform future decisions of individual nodes, for example, route selection \cite{Li2008}.

Recent work has demonstrated the use of a number of metrics to form a ``vector" of trust. The Multi-parameter Trust Framework for MANETs (MTFM) \cite{Guo11}, uses a range of communications metrics beyond packet delivery/loss rate (PLR) to assess trust. This vectorized trust also allows a system to detect and identify the tactics being used to undermine or subvert trust. The authors have previously applied this method to the marine space, comparing against a selection of existing communications TMFs \cite{Bolster2015b} showing that MTFM is more effective at detecting misbehaviours in sparse environments. This paper continues and extends that work to encompass physical as well as communications observations in the establishment of trust and the detection and classification of misbehaviours across both physical and communications domains.

This paper is laid out as follows; in section 2 we discuss Trust and TMFs, defining out terminology and reviewing the justifications for the use and development of TMFs in harsh environments such as UANs. In section 3 we review selected features of the underwater communications channel, highlighting particular challenges against terrestrial equivalents. In section 4 we review the findings of \cite{Bolster2015b} and establish experimental parameters and simulated behaviours under assessment. In section 5 we present our analysis pipeline for assessing misbehaviour using MTFM, and intermediate results of the independent detection of

physical and communications misbehaviours using single-domain observations. In section 6 we demonstrate results from multi-domain MTFM and discuss the significance of these findings in terms of detection and classification of cross-domain misbehaviour sets.

## II. Trust and Trust Management Frameworks

### A. Trust in Networked Systems

<Insert Generic Discussion of Trust Here: ¼ pg>

### B. Trust Management in Conventional MANETS

The distributed and dynamic nature of MANETs mean that it is difficult to maintain a trusted third party (TTP) or evidence based trust system such as Certificate Authorities (CA) or Public Key Infrastructure (PKI). Distributed trust management frameworks aim to detect, identify, and mitigate the impacts of malicious actors by distributing per-node assessments and opinions to collectively police behaviour. Various models and algorithms for describing trust and developing trust management in distributed systems, P2P communities or wireless networks have been considered.

*Hermes Trust Establishment Framework* uses a Bayesian Beta function to model per-link Packet Loss Rate (PLR) over time, combining ``Trust" and ``Confidence of Assessment" into a single value \cite{Zouridaki2005}. *Objective Trust Management Framework* (OTMF) builds upon Hermes and distributes node observations across the network \cite{Li2008}, however does not appropriately combat multi-node-collusion in the network \cite{Cho2011}. *Trust-based Secure Routing* demonstrated an extension to Dynamic Source Routing (DSR), incorporating a Hidden Markov Model of sub-networks, reducing the efficacy of Byzantine attacks such as black-hole routing \cite{Moe2008a}. *CONFIDANT* presented an approach using a probabilistic estimation of PLR, similar to OTMF, also introducing a topology aware weighting scheme and also weighting trust assessments based on historical experience of the reporter \cite{Buchegger2002}. *Fuzzy Trust-Based Filtering* uses Fuzzy Inference to adapt to malicious recommenders using conditional similarity to classify performance with overlapping fuzzy set membership, filtering assessments across a network \cite{Luo2008}.

These TMFs can be generalised as single-value estimation based on a binary input state (success or failure of packet delivery) and generating a probabilistic estimation of the future states of that input. These single metric TMFs provide malicious actors with a significant advantage if their activity does not impact that metric.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. This causes a significant negative effect on the efficiency of the network, as the TMF is assumed to have reduced the possible set of attacks when it has actually made it more advantageous to attack a different part of the networks operation.

An example of such a situation would be in a TMF focused on PLR where an attacker selectively delays packets going through it, reducing overall throughput but not dropping any packets. Such behaviour would not be detected by the TMF.

*Multi-Parameter Trust Framework for MANETS* (MTFM) extends this single-parameter approach, applying Grey Relational Analysis \cite{Zuo1995} to provide cohort based normalization of a range of disparate metrics at runtime, providing a "grade" of trust compared to other observed nodes, while maintaining the ability to reduce trust valued down to a stable assessment range for decision support without requiring a-priori environmental or metric characterisation. This presents a stark difference between the previously discussed probabilistic approaches. Grey assessments are relative in both fairly and unfairly operating networks. All nodes will receive mid-range trust assessments if there are no malicious actors as there is nothing ``bad" to compare against, and variations in assessment will be primarily driven by topological and environmental factors.

Guo et al. \cite{Guo11} demonstrated the ability of grey relational analysis (GRA) \cite{Zuo1995} to normalise and combine disparate traits of a communications link such as instantaneous throughput, received signal strength, etc. into a grey relational coefficient (GRC), or a ``trust vector" in this instance.

<Really not sure how much MTFM detail to go in to here>

## III. Marine Acoustic Communications

The key challenges of underwater acoustic communications are centred around the impact of slow and differential propagation of energy (RF, Optical, Acoustic) through water, and its interfaces with the seabed / air.

The resultant challenges include; long propagation delays, significant inter-symbol interference and Doppler spreading, fast and slow fading due to environmental effects (aquatic flora/fauna, surface weather), carrier-frequency dependent signal attenuation, multi-path caused by reflective medium interfaces, variations in propagation speed due to depth dependant effects (salinity, temperature, and pressure), and subsequent refractive spreading and lensing due to that same propagation variation \cite{Partan2006}.

Refractive lensing and the multi-path nature of the medium result in line of sight propagation being extremely unreliable for estimating distances to targets.

The first arriving acoustic signal has as the very least curved in the medium, and commonly has reflected off the surface/seabed before arriving at a receiver, creating secondary paths that are sometimes many times longer than the first arrival path, generating symbol spreading over orders of seconds depending on the ranges and depths involved.

Forward Error Correction coding is used on such channels to minimise packet losses.

<I've dropped the equations from in here for discussion>

## IV. Existing Work

<summarise Bolster2015b, highlight major take-aways of:

1. Classical MANET Trust not directly suitable to harsh environment due to sparsity of observations leading to more statistical noise than signal in terms of detection of misbehaviour

2. Power of weight variation on the detection and characterisation of misbehaviours
3. Open Questions (not all of them will be answered here) >

We have demonstrated that existing MANET Trust Management Frameworks are not directly suitable to the sparse, noisy, and dynamic underwater medium.

We presented a comparison between trust establishment in MANETs in a simulated underwater environment, demonstrating that in order to have any reasonable expectation of performance, throughput and delay responses must be characterised before implementing trust in such environments. While the MTFM value does not display any immediate difference between the two behaviours, we have shown that by exploring the metric space by weight variation, the existence and nature of the malicious behaviour can be discovered. Another difference is that MTFM is significantly more computationally intensive than the relatively simple Hermes / OTMF algorithms. The repeated metric re-weighting required for real time behaviour detection is therefore an area that requires optimization.

We demonstrated initial, unfiltered Grey Trust assessment using all available metrics (transmitted and received throughput, delay, received signal strength, transmitted power, and packet loss rate), as well as the application of multiple weighting vectors to iteratively emphasise different aspects of trust operation to expose and identify misbehaviour on the network. With significant delays (from seconds to many minutes), in a fading, refractive medium with varying propagation characteristics, the environment is not as predictable or performant as classical MANET TMF deployment environments. We show that, without significant adaptation, single metric probabilistic estimation based TMFs are ineffective in such an environment.

We have shown that existing frameworks are overly optimistic about the nature and stability of the communications channel, and can overlook characteristics that are useful for assessing the behaviour of nodes in the network.

This indicates that there is a good case, particularly within constrained MANETs as this, for multi-vector, and even multi-domain trust assessment, where metrics about the communications network and topology would be brought together with information about the physical behaviours and operations of nodes to assess trust. Also, a significant factor of trust assessment in such a constrained environment, is that there may be long periods where two edge nodes (for instance, $n_0 \to n_5$) may not interact at all.

This can be due to a range of factors beyond malicious behaviour, including simple random scheduling coincidence and intermediate or neighbouring nodes collectively causing long back-off or contention periods.

This disconnection hinders trust assessment in two ways; assessing nodes that do not receive timely recommendations may make decisions based on very old data, and malicious nodes have a long dwelling time where they can operate under a reasonable certainty that the TMF will not detect it (especially if the node itself is behaving disruptively).

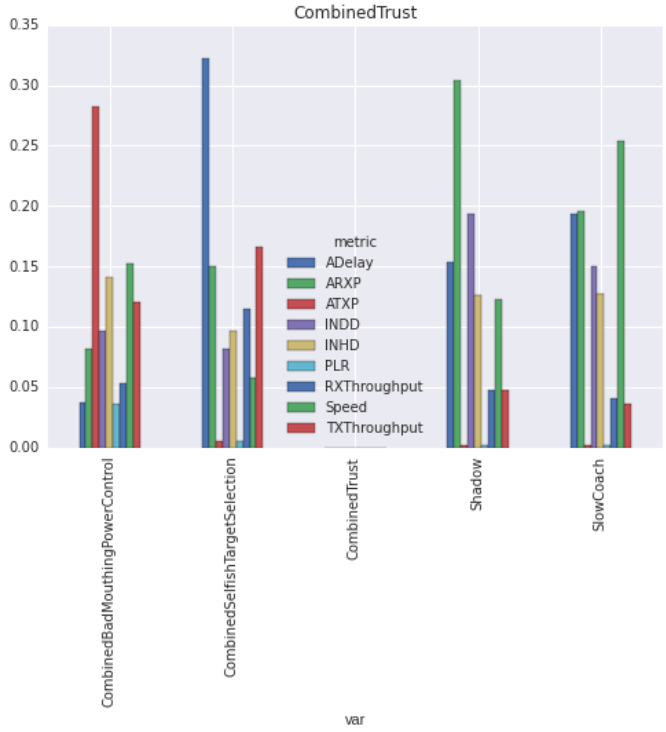
<Summarise Experimental setup and selected behaviours>

## V. Analysis Design and Per-Domain Results

<Formal step-through from simulation results (i.e. position/packet logs) through trust assessment through metric weight exploration with appropriate visual aids. This will mostly be equations / algorithms with some explanations and graphs>

<Finish with presentation of Physical Only / Comms Only Metric/Behaviour results, along with maybe a small discussion on computational complexity of expanding the metric set and a discussion on strategies for mixing/pre-mixing intermediate trust results vs “throw em all in a bag”>

## VI. Cross-Domain Results and Discussion




<Centre behaviour is the “fair” case that everything else is being compared against, and is thus, zero, left are comms behaviours, right are physical behaviours. I’m working on better ways to present this but basically, each behaviour has a ‘signature’ that’s fairly clear that maximises the ‘outlier’ state. In reality, this is maximisation problem, so now that the ‘brute force’ case has been worked, there’s ample opportunity for putting together a greedy algo to work through these cases in much closer to real time. Depending on how successful my experiments are this weekend, I may slip that into this, or may leave for AAMAS in Nov>

## Acknowledgment

The Authors would like to thank the DSTL/DGA UK/FR PhD Programme for their support during this project, as well as NATO CMRE for their advice and assistance.

## REFERENCES

The template will number citations consecutively within brackets [1]. The sentence punctuation follows the bracket [2]. Refer simply to the reference number, as in [3]—do not use "Ref. [3]" or "reference [3]" except at the beginning of a sentence: "Reference [3] was the first ..."

Number footnotes separately in superscripts. Place the actual footnote at the bottom of the column in which it was cited. Do not put footnotes in the reference list. Use letters for table footnotes.

Unless there are six authors or more give all authors' names; do not use "et al.". Papers that have not been published, even if they have been submitted for publication, should be cited as "unpublished" [4]. Papers that have been accepted for publication should be cited as "in press" [5]. Capitalize only the first word in a paper title, except for proper nouns and element symbols.

For papers published in translation journals, please give the English citation first, followed by the original foreign-language citation [6].

[1] G. Eason, B. Noble, and I. N. Sneddon, "On certain integrals of Lipschitz-Hankel type involving products of Bessel functions," Phil. Trans. Roy. Soc. London, vol. A247, pp. 529–551, April 1955. *(references)*

[2] J. Clerk Maxwell, A Treatise on Electricity and Magnetism, 3rd ed., vol. 2. Oxford: Clarendon, 1892, pp.68–73.

[3] I. S. Jacobs and C. P. Bean, "Fine particles, thin films and exchange anisotropy," in Magnetism, vol. III, G. T. Rado and H. Suhl, Eds. New York: Academic, 1963, pp. 271–350.

[4] K. Elissa, "Title of paper if known," unpublished.

[5] R. Nicole, "Title of paper with only first word capitalized," J. Name Stand. Abbrev., in press.

[6] Y. Yorozu, M. Hirano, K. Oka, and Y. Tagawa, "Electron spectroscopy studies on magneto-optical media and plastic substrate interface," IEEE Transl. J. Magn. Japan, vol. 2, pp. 740–741, August 1987 [Digests 9th Annual Conf. Magnetics Japan, p. 301, 1982].

[7] M. Young, The Technical Writer's Handbook. Mill Valley, CA: University Science, 1989.