Applicability of Single and Multi-Metric Trust Management Frameworks in Underwater Autonomous Networks

DRAFT for 8pg submission to TrustCom 31 Mar

Andrew Bolster
Department of Electical Engineering and Electronics
University of Liverpool
Liverpool, UK
Email: bolster@liv.ac.uk

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

***Abstract*—In this paper, we demonstrate the need for multi-metric trust assessment in Underwater Autonomous Networks (UAN) and a performance comparison against a selection of current single metric Trust Management Framework (TMF) assessment processes. We investigate the operation of a selection of traditional MANET TMFs in this environment, highlighting the challenges in this, and demonstrating that a multi-metric approach to Trust greatly enhances the usefulness of Trust in these environments.**

I. Introduction

As mobile ad-hoc networks (MANETs) grow beyond the terrestrial arena, their operation and the protocols designed around them must be reviewed to assess their suitability to different communications environments to ensure their continued security, reliability, and performance.

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 [**?**], and maintaining throughput in the presence of malicious actors [**?**]

Most current TMFs use a single type of observed action to derive trust values, i.e. successfully forwarded packets. These observations then inform future decisions of individual nodes, for example, route selection [**?**].

Recent work has demonstrated use of a number of metrics to form a "vector" of trust. The Multi-parameter Trust Framework for MANETs (MTFM)[**?**], uses a range of physical metrics beyond packet delivery/loss rate (PLR) to form a vector of trust. This vectorized trust allows a system to detect and identify the tactics being used to undermine or subvert trust. To date this work has been limited to terrestrial, RF based networks. As autonomous underwater vehicles (AUVs) become more capable, and economical, they are being used in many applications requiring trust [**?**]. With this use being increasingly isolated from stable communications networks, the decentralised establishment of trust between nodes is essential for the reliability and stability of such teams. As such, the use of trust methods developed in the terrestrial MANET space must be re-appraised for application within the challenging underwater communications channel.

The paper is laid out as follows. In Section II we discuss Trust and Trust Management Frameworks, defining our terminology and reviewing the justifications for the use and development of Trust Management Frameworks in marine acoustic networks. In Section III, we review selected features of the underwater communications channel, highlighting particular challenges and differentials against terrestrial equivalents. In Section IV, we establish an experimental configuration for the marine space, and review the scenarios and results presented in [**?**]. In Section V, we present our findings in trust establishment and malicious behaviour detection, comparing with other current TMFs (OTMF and Beta) and analyse the use of this multi-parameter approach to detecting malicious and selfish behaviour in autonomous marine networks.

The contributions of this paper are a study on the comparative operation and performance between terrestrial and underwater MANETs using TMFs, and a review of metric suitability for TMFs in marine environments, informing future metric selection for experimenters and theorists. We also show that single metric trust systems are not directly suitable for the marine context in therms of the different threat and cost scenario in that environment.

II. Trust and Trust Management Frameworks

A. *Trust in 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. Taking some examples;

- *The Objective Trust Management Framework* takes a Bayesian Beta function to model per-link Packet Loss Rate (PLR) over time, combining "Trust" and "Confidence of Assessment" into a single value [**?**]. OTMF however does not appropriately combat multi-node-collusion in the network [**?**].
- *Trust-based Secure Routing*[**?**] demonstrated an extension to Dynamic Source Routing (DSR), incorporating a Hidden Markov Model of next-hop network, reducing the efficacy of Byzantine attacks such as black-hole routing.
- *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.
- *Fuzzy Trust-Based Filtering*; [**?**] presents the use of Fuzzy Inference to adapt to malicious recommenders using conditional similarity to classify performance with overlapping Fuzzy Set Membership, filtering assessments across a network.

These TMFs can be generalised as single-value probabilistic estimation, based around using a binary input state and generating an 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. In turn, this causes a super-linearly negative effect in 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.

B. OTMF and Beta Trust Assessment

Introdution to OTMF

This expectation value is beta$(p|\alpha,\beta) \rightarrow E(p) = \frac{\alpha}{\alpha+\beta}$ where α and β represent the number of successful and unsuccessful interactions respectively.

There are also situations where the observed metrics will include significant noise and occur at irregular, sparse, intervals. Conventional approaches such as probabilistic estimation do not produce trust values that reflect the underlying reality and context of the metrics available, as they require a-priori assumption that the trust value under exploration has an expected distribution, that distribution is mono-modal, and the input metrics are binary. In scenarios with variable, sparse, noisy metrics, estimating the distribution is difficult to accomplish a-priori.

C. Grey Theory and MTFM

Grey Theory performs cohort based normalization of metrics at runtime, providing a "grade" of trust compared to other observed nodes in that interval, while maintaining the ability to reduce trust values down to a stable assessment range for decision support without requiring every environment entered into to be characterised. This presents a stark difference between the Grey and 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 no-one else "bad" to compare against, and variations in assessment will be primarily driven by topological and environmental factors.

Guo et al.[**?**] demonstrated the ability of Grey Relational Analysis (GRA)[**?**] 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 [**?**], the observed metric set $X = x_1, \ldots, x_M$ representing the measurements taken by each node of its neighbours at least interval, is defined as $X =$ [packet loss rate, signal strength, data rate, delay, throughput]. The grey relational 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).

Weighting can be applied before generating a scalar value (2) allowing the detection and classification of misbehaviours.

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

Where $H = [h_0 \ldots h_M]$ is a metric weighting vector such that $\sum h_j = 1$, and in the basic case, $H = [\frac{1}{M}, \frac{1}{M} \ldots \frac{1}{M}]$ to treat all metrics evenly. 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 $T^t_k = (1 + (\phi^t_k)^2/(\theta^t_k)^2)^{-1}$.

Cite LMS reduction from Guo

MTFM combines this GRA with a topology-aware weighting scheme(3) and a fuzzy whitenization model(4). There are three classes of topological trust relationship used; Direct, Recommendation, and Indirect. Where an observing node, n_i, assesses the trust of another, target, node, n_j; the Direct relationship is n_i's own observations n_j's behaviour. In the

Recommendation case, a node n_k, which shares Direct relationships with both n_i and n_j, gives its assessment of n_j to n_i. In the Indirect case, similar to the Recommendation case, the recommender n_k, does not have a direct link with the observer n_i but n_k has a Direct link with the target node, n_j. These relationships give us node sets, N_R and N_I containing the nodes that have recommendation or indirect, relationships to the observing node respectively.

$$\begin{aligned} T_{i,j}^{MTFM} =& \frac{1}{2} \cdot \max_s \{f_s(T_{i,j})\} T_{i,j} \\ &+ \frac{1}{2} \frac{2|N_R|}{2|N_R| + |N_I|} \sum_{n \in N_R} \max_s \{f_s(T_{i,n})\} T_{i,n} \\ &+ \frac{1}{2} \frac{|N_I|}{2|N_R| + |N_I|} \sum_{n \in N_I} \max_s \{f_s(T_{i,n})\} T_{i,n} \end{aligned} \tag{3}$$

Where $T_{i,n}$ is the subjective trust assessment of n_i by n_n, and $f_s = [f_1, f_2, f_3]$ given as:

$$\begin{aligned} f_1(x) &= -x + 1 \\ f_2(x) &= \begin{cases} 2x & \text{if } x \leq 0.5 \\ -2x + 2 & \text{if } x > 0.5 \end{cases} \\ f_3(x) &= x \end{aligned} \tag{4}$$

III. Marine Acoustic Networks

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 it's interfaces with the seabed / air. The resultant challenges include; long delays due to propagation, 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, multipath 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 [**?**].

The attenuation that occurs in an underwater acoustic channel over a distance d for a signal about frequency f in linear power is given as $A_{\text{aco}}(d, f) = A_0 d^k a(f)^d$ and in dB form as (**??**)

$$10 \log A_{\text{aco}}(d, f)/A_0 = k \cdot 10 \log d + d \cdot 10 \log a(f) \tag{5}$$

where A_0 is a unit-normalising constant, k is a spreading factor (commonly taken as 1.5

cite

), and $a(f)$ is the absorption coefficient, expressed empirically using Thorp's formula (6) from [**?**]

$$10 \log a(f) = \frac{0.11 \cdot f^2}{1 + f^2} + \frac{44 \cdot f^2}{4100 + f^2} + 2.75 \times 10^{-4} f^2 + 0.003 \tag{6}$$

Refractive lensing and the multipath nature of the medium result in supposedly line of sight propagation being extremely unreliable for estimating distances to targets. The first arriving beam has as the very least bent 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.

CULL:Is this relevant?

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

Comparing $A_{aco}(d, f)$ with the RF Free-Space Path Loss model ($A_{\text{RF}}(d, f) \approx \left(\frac{4\pi df}{c}\right)^2$), the impact of range on signal power is exponential underwater, rather than quadratic in terrestrial RF ($A_{\text{aco}} \propto f^{2d}$ vs $A_{\text{RF}} \propto (df)^2$). While both frequency dependant factors are quadratic, approximating the factors in (6), $f \propto A_{\text{aco}}$ is at least 4 orders of magnitude higher than $f \propto A_{\text{RF}}$

A. Trust in Marine Networks

With demand for smaller, more decentralised marine survey and monitoring systems, and a drive towards lower per-unit cost, TMFs are expected to be increasingly applied to the marine space, as the benefits they present are significant

cite

. Beyond the constraints of the communications environment, knock on pressures are emerging in battery capacity, on-board processing, and locomotion..

rephrase

These pressures simultaneously present opportunities and incentives for malicious or selfish actors to appear to cooperate while not reciprocating, in order to conserve power for instance These multiple aspects of potential incentives, trust, and fairness do not directly fall under the scope of single metric trusts discussed above, and this context indicates that a multi-metric approach may be more appropriate.

IV. Initial System Model Characterization

A. Mobility, Topology, and Communications

Four mobility scenarios were used in [**?**] to explore trust behaviour; all nodes static, a central node n_1 performing a random walk with other nodes remaining static, all nodes but the central node (n_1) randomly walking, and all nodes randomly walking. From these we select the all static and all mobile cases for presentation. The reason for this is that giving a malicious node special privilege or capabilities will skew the results of trust assessment, as the behaviours of the static and mobile nodes will be significantly different regardless of misbehaviour.

The six nodes are initially arranged as per Fig. 1, with each node on average 100m from each other, as per [**?**]. The use of six nodes and the particular layout enables the investigation of the three trust relationships based on minimum path topologies

Possibly cite Guo paper on team sizes or find an alternative

, such that the node generating the trust assessments, n_0 has Direct, Recommendation, and Indirect trust assessments of n_1 available to it from itself, $[n_2, n_3]$, and $[n_4, n_5]$ respectively.

In all of the scenarios, each link from $n_i \rightarrow n_j$ periodically sent 10 second bursts of Constant Bit Rate (CBR) style traffic. Guo et al. demonstrated that when compared against OTMF and Beta trust assessment, MTFM provided increased variation in trust assessment over time, providing more information about the nodes behaviour than packet delivery probability. By weighting the metrics used in MTFM, it was shown that the trust assessments could be used to identify the style of misbehaviour being performed within the network and by who. We present a corollary method to investigate and apply this work to the Marine MANET field.

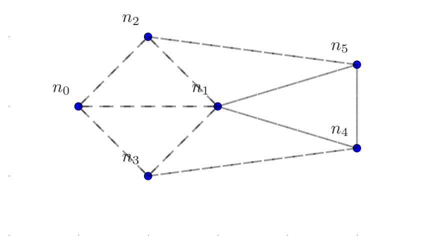


Fig. 1. Initial layout with nodes spaced an average of 100m apart

CULL: Not sure if this is necessary

B. Simulation Background

Simulations were conducted using a Python based simulation framework, SimPy[?], with a network stack built upon AUVNetSim[?], with transmission parameters (Table I) taken from and validated against [?] and [?].

Given the differences in delay and propagation between RF and marine networks, it is natural that the same application rates (e.g. packet emission rates or throughput) and node separations should not be assumed to be equally viable. Therefore, we characterise an operational zone of performance within which the network can operate stably.

TABLE I. COMPARISON OF SYSTEM MODEL CONSTRAINTS AS APPLIED BETWEEN TERRESTRIAL AND MARINE COMMUNICATIONS

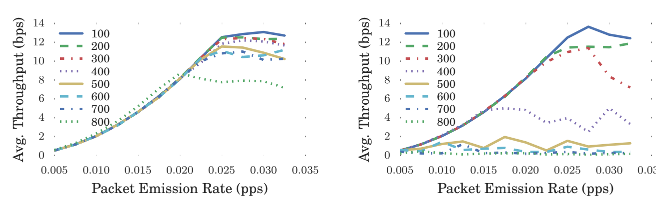
Parameter	Unit	Terrestrial	Marine
Simulated Duration	s	300	18000
Trust Sampling Period	s	1	600
Simulated Area	km^2	0.7	0.7-4
Transmission Range	km	0.25	1.5
Physical Layer		RF(802.11)	Acoustic
Propagation Speed	m/s	3×10^8	1490
Center Frequency	Hz	2.6×10^9	2×10^4
Bandwidth	Hz	22×10^6	1×10^4
MAC Type		CSMA/DCF	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

C. Scaling Considerations between Terrestrial and Underwater Environments

In this section we characterise the simulated communications environment, establishing an optimal packet emission rate for comparison against [?].

We establish a appropriate safe operating zone for marine communications by looking at the communications rate and physical distribution factors across the two selected mobility scenarios. In scaling the physical distribution of the nodes, we also scale the environment in which the nodes are restricted to, which has a significant impact on the number of potential runtime topologies, with nodes getting increasingly isolated as the environment space increases. This leads to increasing delays as routes are constantly broken, re-advertised and re-established. From Table I, the operating transmission range of this model of acoustic communications is ≈ 6 times further than that of 802.11, indicating that a suitable operating environment will have an area $\approx \sqrt{6}$ times the area of the 802.11 case. However, it was recognised in Section III that the relationship between attenuation and distance is exponential underwater, so this would represent an upper bound of performance, where nodes begin approximately 400m apart.

As the separation is increased, the emission rate at which the network becomes saturated decreases, reducing overall throughput (Fig. 2). This throughput degradation is tightly coupled with the mobility. For instance, in Fig. 2a, where all nodes are static, we do not see significant drops in saturation rates until we approach 800m, nearly double our initial estimate. However, in Fig. 2b, where all the nodes are randomly walking, the saturation point collapses from 0.025pps at 300m to 0.015pps at 400m. These 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 on average, beyond which communication becomes increasingly unstable, especially in terms of end to end delay (not shown)



(a) Static

(b) Mobile

Fig. 2. Throughput Characteristics for varying node separations across increasing packet emission rates

D. Introduced Misbehaviours

Guo et al. introduce a range of malicious behaviours, including modification of the packet loss rate of routing nodes and limiting throughput on a per-link basis as well as a selection of combined misbehaviours.

Given that the established links are already heavily constrained, heavy handed attacks such as introducing selective PLR and adding to the already extreme and hugely variable delays would severely impact the general performance of the network beyond the scope of simple selfishness, effectively triggering saturation collapses in regions that the network should be stable. Therefore, we select two misbehaviours to investigate; a Malicious Power Control behaviour, where n_1 increases it's transmission power by 20% for all nodes *except* communications with n_0 in order to make n_0 appear "bad" to the rest of the team, and a Selfish Target Selection behaviour where nodes preferentially communicate, forward and advertise to nodes that are physically close to them in effort to reduce overall power consumption for the individual.

V. Trust

Having established a safe operating range for comparison, at 300m average separation and an emission rate of 0.015pps, we repeat the static and mobile scenarios presented in [?]. We select an assessment period of 10 mins for a 5 hour mission to scale in comparison to relative bitrates experienced (1Mbps vs $\approx$ 15bps).

Metrics used for Grey assessment are transmitted and received throughput and power, delay, and packet loss rate as calculated by aborted and unacknowledged, transmissions. Compared to [?], this metric set lacks a data rate quantity as the network is not dynamically adjusting bandwidth. In context of Grey Relational Coefficient generation (1), the best sequence g was selected using the lowest PLR, delay, and powers, and the highest throughputs, with the worst sequence, b the inverse of these metrics, reflecting the observations made in Section III-A.

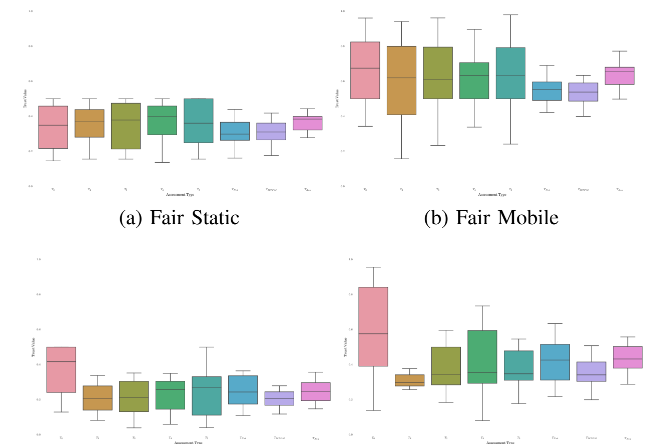
The particular factors under discussion are the relative performance of MTFM against OTMF and Beta with respect to statistical stability across mobilities and in responsiveness to changing network behaviour. We establish a similar result set by initially tracking the resultant trust values established by MTFM in the pair of mobility scenarios, shown in Fig.3. As per Guo et al. we are primarily concerned with the observational trust relationship between n_0 and n_1, i.e. n_0's assessment of the trustworthiness of n_1, or $T_{1,0}$. We are also concerned with the opinions of n_1 provided to n_0 by other nodes, where $[T_{1,2}, T_{1,3}]$ and $[T_{1,4}, T_{1,5}]$ denote the sets of recommendation and indirect trust assessment respectively.

We also include aggregate assessments; $T_{1,\text{Avg}}$, the flat average of direct trust assessments of n_1, $T_{1,\text{Net}}$, that weights assessments according to the network topology from (3), without the whitenization factor f_s, and $T_{1,\text{MTFM}}$, the final MTFM trust assessment value based on both network topology and whitenization from (4).

It's possible this paragraph is surplus to requirements and the relevant sections of the boxplot could be removed to simplify things

The variability in assessment is coupled to mobility; in the static case (Fig. 3c), we see that the nodes close to n_1 ($[n_0, n_2, n_3]$) have reasonably consistent distributions, and as the range increases out to $[n_4, n_5]$, this variability increases. In the full mobility case, shown in Fig. 3d, this subjective variability is greatly increased. As the topology is highly dynamic, delays due to re-establishing routes can be very large, perturbing the trust value. The aggregate trust values using topology information ($T_{1,\text{Net}}, T_{1,\text{MTFM}}$) display a decreased variation than those of the individual subjective observations in both cases.

In comparison to [?], these results are qualitatively similar, however in this case the weighted deltas are significantly less clear than in the comparable terrestrial space, where Guo et al. show the same type of malicious behaviour and demonstrates 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.



(a) Fair Static (b) Fair Mobile

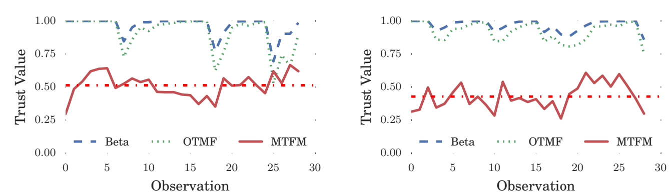
(c) Selfish Static (d) Selfish Mobile

Fig. 3. MTFM Trust assessments of n_1 ($T_{1,X}$)[1]

Discuss differential of n0 ratings, also this is impossible to read

A. Comparison to OTMF and Beta

As per [?], parallel simulations were performed where there was no malicious behaviour, the "fair" scenario utilising OTMF and Beta assessment as well as MTFM, providing like-for-like comparison of assessment. The use of Forward Beam



(a) Fair Scenario (b) Selective Power Control Scenario

Fig. 4. $T_{1,0}$ for Beta, OTMF and MTFM assessment values for fair and selfish behaviours in the fully mobile scenario (mean of MTFM also shown)

Routing and a CSMA/CA MAC scheme from AUVNetSim[?] in our simulation mitigates a significant number of packet losses through collision avoidance, and contention handling, leading to the situation that the only genuinely lost packets occur when a node moves completely out of range of any other node and time out in route discovery rather than transmission. As such, confirmed packet losses are extremely rare and in a delaying network like this, it is difficult to set a differentiating

[1] Box plots centres indicate the median, bounds indicate the 25%-75% range, and whiskers represent the points within $\pm 2\sigma$

time-out between packets that are in the network but queued, and packets that are actually "lost". This renders OTMF and Beta assessment at best uninformative and at worst misleading; consistently providing nodes a high trust assessment as they have very little information to extract trust from.

The single metric TMFs used in Terrestrial MANETs require regular and constant streams of positive and negative validation to shape and adjust their evaluations, which for a network with significant delays such as this, is not practical.

Fig. 4 shows a comparison between the unweighted response of MTFM compared to OTMF and Beta assessment functions on the same data for the fair 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, agents are compared against the worst current performances across metrics of other nodes and graded against them, rather than the absolute (objective) approach taken by many TMFs. This relative versus absolute difference is particularly clear when comparing mobility models. In this case, particularly since the method of attack was not directly related to PLR, OTMF and Beta have not registered significant activity in the correct behaviour. Without comparing against any other known quantity; the difference between the MTFM trust assessments under "fair" and "malicious" behaviour is appreciably affected, particularly in looking at the average values returned.

B. Comparison under dynamic metric weighting

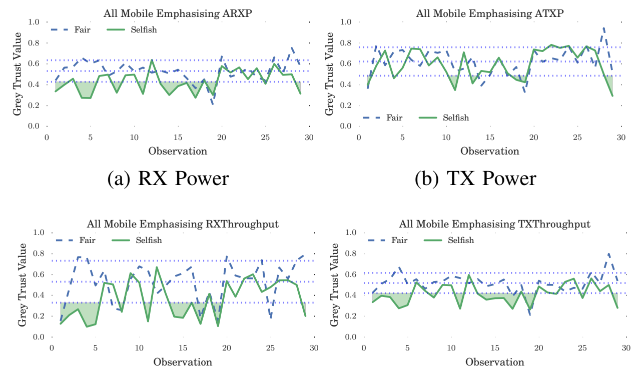
We apply a sequence of metric vectors that preferentially weight each metric during (2) to each of the two simulation runs. For an arbitrary metric weight vector H, where the metric m_j is emphasised as being twice as important as the other metrics, we form an initial weighting vector $H' = [h_i...h_M]$ 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.

(a) Delay (b) TX Power

(c) RX Throughput (d) TX Throughput

Fig. 5. $T_{1,MTFM}$ in the All Static case for the Selective Power Control behaviour, emphasising selected metrics and showing the mean and $\pm\sigma$ of $T_{1,MTFM}$ in the same 'fair' scenario

From Fig. 5 we can see that the selfish node is consistently outside the $\pm\sigma$ envelope of the fair node it's being compared



(a) RX Power (b) TX Power

(c) RX Throughput (d) TX Throughput

Fig. 6. $T_{1,MTFM}$ in the All Mobile case for the Selective Power Control behaviour

to, particularly TX Power, with smaller impacts on RX/TX Throughput, as would be expected for a power related selfish behaviour. However, the impact on delay is minimal to insignificant, occasionally breaching the envelope for a short period. This was to be expected in a contention-based medium access network operating close to its saturation point; it can be observed that the delay deviance appears to increase as simulation time progresses. This indicates that the variation in delay could be caused not by a malicious behaviour but simple congestion. In the mobile case (Fig. 6) we observe a similar pattern, however it should be noted that the deviation envelope is greatly increased compared to the static case due to the underlying variations in topology and configuration in this scenario.

A significant factor of trust assessment in such a constrained environment is that there may be long periods where two edge nodes (for instance, $n_0 \rightarrow n_5$) may not interact at all. This can be due to a range of factors beyond potential 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). One potential solution to this would be to move from a stepping-window of trust periods to a continuous trust log, updated on packet reception rather than waiting for a number of packets to arrive.

Possibly move this section

For a more quantative assessment of the viability of this multi-metric trust assessment method, we take the qualatative analysis above and apply a Random Forest regression

cite

to quantify the relative importance of the selected metrics on relative detectability of malicious behaviour. The target function for this regression is given in (7), and captures the relative difference between the fair and misbehaviour curves as demonstrated in Figs. 5 and 6. Visually, this is the total area between the T_{MTFM} in the misbehaving case that is outside the standard deviation of T_{MTFM} of the fair case.

Applying this target function to 729 different metric weight vectors emphasis combinations (H) and applying the regression demonstrates that while PLR is the primary factor in differentiating the fair and malicious behaviours ($R = 0.75, p \approx 10^{-100}$), the recorded transmission strength is also significant ($R = -0.54, p \approx 10^{-50}$).

"While it is not feasible to perform this breadth of calculations at run time to detect and classify unknown or unexpected behaviours, a more nuaced multi-dimensional optimisation approach could be applied, providing additional resiliace to attack"

The selfish behaviour where the node preferentially select nodes to communicate with based on proximity is a more subtle attack on the network when compared to the outright power-flooding of the malicious behaviour. However, it still impacts the efficiency and utility of the network by creating artificial asymmetries in the networks information distribution. Applying the same approach to this behaviour, we find that the ensemble MTFM results appear similar to the malicious case, but OTMF and Beta don't appear to differentiate between fair and selfish scenarios at all.

need to add graphics for this

This is natural as these TMFs do not take protocol or application level behaviour into account in their assessment of fairness. Through the same regression as above, we find that this intuition is validated in feature extraction, with observed transmission power dominating ($R = 0.88, p < 10^{-100}$) followed by PLR ($R = -0.45, p \approx 10^{-35}$) and throughput ($R = -0.33, p \approx 10^{-20}$)

$$Y(H) = \int T_{mal}(H) - (T_{gd}(H) \pm \sigma_{T_{gd}(H)}) \, \mathrm{d}H. \qquad (7)$$

VI. Conclusions and Future Work

We have demonstrated that existing MANET Trust Management Frameworks are not directly suitable to the contentious and dynamic underwater medium. We presented a comparison between trust establishment in Terrestrial MANET and in the underwater space, 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 computationally, MTFM is significantly more intensive than the relatively simple Beta / OTMF algorithm, and the repeated metric matrix re-weighting required for real time behaviour detection is an area that requires optimization.

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 misbehaviour can be discovered. Another difference is that computationally, MTFM is significantly more intensive than the relatively simple Beta / OTMF algorithm, and the repeated metric matrix re-weighting required for real time behaviour detection is an area that requires optimization. As such, a hybrid system could be implemented, that used OTMF as a 'trigger' to detect potentially selfish or malicious behaviour, and allow MTMF weight matrix execution to be triggered at less regular intervals. 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.

However, with significant delays (order 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 of the channel that are useful for assessing the behaviour of nodes in the network. This indicates that there is a good case, particularly within constrained MANETs such 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.

Future work will investigate the stability of GRA under multi-node collusion, the development of real-time outlier detection and filtering for metrics (e.g differentiating between a very long delay that was an 'accident' and a malicious router), and the introduction of physical metrics and sensing capabilities into the trust management context.

Acknowledgment

The Authors would like to thank the DSTL/DGA UK/FR PhD Programme for their support during this project.