Machine Learning Supported Metric Selection and Characterisation for Multi-Domain Trust in Autonomous Underwater MANETs

(Invited Paper for Underwater Communications Security session) Pages: 4, Deadline: 1/4

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

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

***Abstract*—The increasing sensing and communications capabilities of Autonomous Underwater Vehicles (AUVs) presents an opportunity to assess and establish trust based on an increasing range of available metrics. This increased information availability has been demonstrated to be beneficial in the detection and identification of single-node misbehaviour in Underwater Acoustic Networks (MANs) [1]. We present an assistive Machine Learning based methodology to assess and characterise near-optimal trust metric sets and the weights of those metrics for use in Multiparmeter Trust Framework for MANETs (MTFM) [2]**

I. Introduction

Trust in terrestrial MANETs has been an area of active research for many years; with implicitly limited resources in terms of power, mobility, and communications range, the elimination or isolation of malicious nodes within the decentralised MANET is an obvious driver for increasing overall network efficiency. However, as these decentralised networks expand beyond the terrestrial arena into aerial and underwater domains, these terrestrially-born efficiencies and security protocols must be assessed for their suitability to these new regions of operation.

With respect to the comparatively stable terrestrial RF environment, the underwater acoustic communications environment is unforgiving; generally static or stable assumptions about propagation delay and paths, frequency-based attenuation and minimal refraction simply don't apply to many marine communications channels. This, coupled with the highly location, depth, weather, and flora and fauna dependent variability of these fundamental channel parameters make it difficult to make strong assumptions about if instantenous network performance is the fault of a misbehaving node or of a whale passing between nodes for instance.

Trust Management Frameworks (TMFs) provide information to assist the estimation of future states and actions of nodes within networks. This information is used to optimise the performance of a network against malicious, selfish or defective misbehaviour by one or more nodes. Existing research has demonstrated the advantages of implementing TMFs to 802.11 based MANETs in terms of preventing selfish operation and maintinaing throughput in the presence of malicious nodes [3], [4].

In this paper we build upon previous work [1] that demonstrated the use of a naieve Random Forest regression [5] to assess the relative importance of Communications Metrics in a simulated MAN, and extend the presented methodology with novel optimisation target functions and applying a range of Machine Learning techniques to investigate their suitability in Metric Assessment across both Physical and Communications Metrics.

This paper is laid out as follows: Sec. II outlines the operation and parameters of MTFM and summarises the comparison of MTFM and other classically terrestrial MANET TMFs.

II. MTFM and TMF suitability in the Marine Communications Environment

A. Terrestrial Trust Management Frameworks

1) Single Metric TMFs: 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* 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 [6]. *Objective Trust Management Framework* (OTMF) builds upon Hermes and distributes node observations across the network [7], however does not appropriately combat multi-node-collusion in the network [8]. *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 [9]. *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 [4]. *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 [10].

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.

2) Multiparameter Trust for MANETs: Guo et al. [2] demonstrated the ability of grey relational analysis (GRA) [11] 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.

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 abstract trust values for decision support without requiring per-environment calibration or characterisation. This presents a stark difference between the Grey and Probabilistic approaches. Grey assessments are relative in both fairly and unfairly operating networks; all nodes 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.

The grey relational vector is given as

$$\theta^t_{k,j} = \frac{\min_k |a^t_{k,j} - g^t_j| + \rho \max_k |a^t_{k,j} - g^t_j|}{|a^t_{k,j} - g^t_j| + \rho \max_k |a^t_{k,j} - g^t_j|} \qquad (1)$$
$$\phi^t_{k,j} = \frac{\min_k |a^t_{k,j} - b^t_j| + \rho \max_k |a^t_{k,j} - b^t_j|}{|a^t_{k,j} - b^t_j| + \rho \max_k |a^t_{k,j} - b^t_j|}$$

where $a^t_{k,j}$ is the value of an observed metric x_j for a given node k at time t, ρ is a distinguishing coefficient set to 0.5, g and b are respectively the "good" and "bad" reference metric sequences from $\{a^t_{k,j} k = 1, 2 \dots K\}$, i.e. $g_j = \max_k(a^t_{k,j})$, $b_j = \min_k(a^t_{k,j})$ (where each metric is selected to be monotonically positive for trust assessment, e.g. higher throughput is presumed to be always better).

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

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

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 $T^t_k = (1 + (\phi^t_k)^2/(\theta^t_k)^2)^{-1}$ [12]. 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 node sets, N_R and N_I containing the nodes that have recommendation or indirect, relationships to the observing node respectively.

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

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:

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

3) Summary of Previous Work: In [1], Hermes trust establishment and OTMF were selected indicative single-metric TMFs to compare against MTFM, as Hermes captures the core operation of a pure single metric assessment methodology and OTMF provides a comparison that combines assessments from across nodes to develop trust opinions.

To fairly transition from Terrestrial to Marine spaces, a series of simulations were performed to establish a suitable optimal scaling rage for a sample scenario with six nodes performing a persistent survey operation. Both communications rate and physical separations were scaled to optimise throughput and minimise network saturation across a range of mobility models. Subsequently, MTFM, Hermes and OTMF were applied to a range of control simulations in this new scaled range to assess their performance in detecting and identifying two modes of misbehaviour; Malicious Power Control (MPC) where an attacker raises it's transmission power for all nodes *except* a particular target node, making that target node appear to be selfishly conserving energy to

the rest of the network while the attacker itself appears to be operating fairly; and Selfish Target Selection (STS) where an "attacker" preferentially communicates with nodes near to itself to minimise energy use at the cost of network fairness and throughput.

This comparison demonstrated that while the performance of MTFM, Hermes and OTMF were all severely reduced in the marine environment compared to that in the terrestrial, Unweighted MTFM showed a consistent if small (10%) deviation in misbehaviour-cases and no such deviation in the Fair case. When the individual metrics in MTFM were preferentially weighted, it was demonstrated that a "Relevance Signature" could be generated for detectable behaviours based on a Random Forest regression of multiple weighted assessments, but this method of characterisation was incomplete and did not produce generalised metric weights for use in real-time detection.

III. Some clever section title

The identification and establishment of weight-based "filters" for MTFM has classically been accomplished through the use of variations on the Analytic Hierarchy Process (AHP)[13]. AHP attempts to systematically capture the relative metric performance in detecting alternative behaviours through the trust impact comparison of iteratively emphasising individual metrics. However, similar applications of AHP has been criticised in the past [14] for making potentially misleading or outright incorrect assumptions about the relative scaling and relationships between individual metrics, and this correctness being highly sensitive to the designer-imposed hierarchies used to generate "optimal" alternatives.

As such, we propose a purely computational methodology for filter-weight generation.

A. The one where you talk about Mean T Deltas

Before embarking on any optimisation, it is important to establish what analytical behaviour is being targeted; in this case this target is the identifiable exposure of a "low" trust value for a misbehaving node that is as distinct from other "fair" nodes as possible across a full (or multiple) simulation runs of a particular behaviour.

We characterise this "objective function" as $\Delta\overline{T}_{ix}$

$$\Delta\overline{T}_{ix} = \frac{\sum_{j \neq x} \left(\overline{T_{i,j}}^{\forall t} \right)}{N-1} \tag{5}$$

Where $T_{i,j}$ is the resultant weighted trust value from 3 of node j from the perspective of node i and N is the number of nodes in the current cohort. The larger this $\Delta\overline{T}_{ix}$ value, the further that particular weight shows the misbehaving node as a "Untrustworthy".

Initially, we mirror the methodology of MTFM (i.e. emphasising single metrics independently) but take it to an extreme by instead fully weighting one metric at a time ($H_3 = 0, 0, 0, 1, 0, 0$ for example).

When this extremely simple

IV. Conclusion

The conclusion goes here.

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

[1] A. Bolster and A. Marshall, "Single and Multi-Metric Trust Management Frameworks for use in Underwater Autonomous Networks," in *Trust. 2015*, 2015.

[2] J. Guo, A. Marshall, and B. Zhou, "A new trust management framework for detecting malicious and selfish behaviour for mobile ad hoc networks," *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*, pp. 142–149, 2011. [Online]. Available: http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=6120813

[3] H. Li and M. Singhal, "Trust Management in Distributed Systems," *Computer (Long. Beach. Calif).*, vol. 40, no. 2, pp. 45–53, 2007. [Online]. Available: http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=4085622

[4] S. Buchegger and J.-Y. Le Boudec, "Performance analysis of the CONFIDANT protocol," in *Proc. 3rd ACM Int. Symp. Mob. ad hoc Netw. Comput. - MobiHoc '02*. ACM Press, 2002, pp. 226–236. [Online]. Available: http://dl.acm.org/citation.cfm?id=513800.513828

[5] L. Breiman, "Random forests," *Mach. Learn.*, pp. 5–32, 2001. [Online]. Available: http://link.springer.com/article/10.1023/A:1010933404324

[6] C. Zouridaki, B. L. Mark, M. Hejmo, and R. K. Thomas, "A quantitative trust establishment framework for reliable data packet delivery in MANETs," *Proc. 3rd ACM Work. Secur. ad hoc Sens. networks*, pp. 1–10, 2005.

[7] J. Li, R. Li, J. Kato, J. Li, P. Liu, and H.-H. Chen, "Future Trust Management Framework for Mobile Ad Hoc Networks," *IEEE Commun. Mag.*, vol. 46, no. 4, pp. 108–114, apr 2007. [Online]. Available: http://ieeexplore.ieee.org/xpls/abs{_}all.jsp?arnumber=4212452http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=4481349

[8] J.-h. Cho, A. Swami, and I.-r. Chen, "A survey on trust management for mobile ad hoc networks," *Commun. Surv. & Tutorials*, vol. 13, no. 4, pp. 562–583, 2011. [Online]. Available: http://ieeexplore.ieee.org/xpls/abs{_}all.jsp?arnumber=5604602

[9] M. E. G. Moe, B. E. Helvik, and S. J. Knapskog, "TSR: Trust-based secure MANET routing using HMMs," *... symposium QoS Secur. ...*, pp. 83–90, 2008. [Online]. Available: http://dl.acm.org/citation.cfm?id=1454602

[10] J. Luo, X. Liu, Y. Zhang, D. Ye, and Z. Xu, "Fuzzy trust recommendation based on collaborative filtering for mobile ad-hoc networks," *2008 33rd IEEE Conf. Local Comput. Networks*, pp. 305–311, 2008. [Online]. Available: http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=4664184

[11] F. Zuo, "Determining Method for Grey Relational Distinguished Coefficient," *SIGICE Bull.*, vol. 20, no. 3, pp. 22–28, jan 1995. [Online]. Available: http://doi.acm.org/10.1145/202081.202086

[12] L. H. L. Hong, W. C. W. Chen, L. G. L. Gao, G. Z. G. Zhang, and C. F. C. Fu, "Grey theory based reputation system for secure neighbor discovery in wireless ad hoc networks," *Futur. Comput. Commun. (ICFCC), 2010 2nd Int. Conf.*, vol. 2, 2010.

[13] T. L. Saaty, "Relative measurement and its generalization in decision making: why pair wise comparisons are central in mathematics for the measurement of intangible factors-The analytic hierarchy process," *Racsam*, vol. 102, no. 2, pp. 251–318, 2008.

[14] R. Whitaker, "Criticisms of the Analytic Hierarchy Process: Why they often make no sense," *Math. Comput. Model.*, vol. 46, no. 7-8, pp. 948–961, 2007.