Multi-Domain Trust Frameworks for Harsh Environments

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—**With the increasing application of autonomy in cyber-physical systems, Trust Management Frameworks (TMFs) are being used to improve the efficiency, security, and reliability of decentralized and distributed autonomous systems. In such systems, subtle misbehaviors can significantly impact the operation and performance of the system as a whole. Techniques have been developed for high-speed, uncontended environments such as terrestrial 802.11 MANETs. However, these do not perform well in sparse / harsh environments such as those found in Underwater Acoustic Networks (UANs), where network nodes experience significant and variable delays, comparatively low data rates, large contention periods, and considerable multi-path artifacts. In such sparse networks, trust establishment based on statistical observations of success/failure events become unstable and ineffective in detecting or identifying misbehaviors. Additionally, these methodologies focus solely on the communications actions of entities and do not incorporate the physical domain. 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 UANs as an exemplar application. 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.**

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

1) Trust in Networked Systems:

Insert Generic Discussion of Trust Here: pg

A. 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 [**?**]. Objective Trust Management Framework (OTMF) builds upon Hermes and distributes node observations across the network [**?**], 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 sub-networks, 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 uses 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 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 [**?**] 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. [**?**] 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 (GRC), or a “trust vector” in this instance.

¡Really not sure how much MTFM detail to go in to here¿

III. 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.