An Investigation into Trust and Reputation Frameworks for Autonomous Underwater Vehicles

Author: Andrew Bolster

Supervisors: Prof. Alan Marshall, Prof. Simon Maskell (UoL)
Prof. Jean-Guy Fontaine (UPMC)

University of Liverpool

Thursday 6th October 2016

Structure of this presentation

Structure

- As short a summary of the work as can be managed
- Fundamental Errata (hint: there haven't been any)
- Discussion of new research that has entered the field since submission
- Quick walk through of the main findings of each chapter for review
- Straight into Defence of Work Discussion
- Wealth of supporting slides for discussion

Summary of Contributions

Primary

- 1st comparative application of Trust in UAN
- 1st application of direct Physical/Mobility based Trust Metrics in any context
- 1st automatic, behaviour based optimisation of MTFM weighting

Secondary

- Reactive agent based Simulation system
- Review of Trust in the marine defence context

Publications

- Single and Multi-metric Trust Management Frameworks for Use in Underwater Autonomous Networks. IEEE TrustCom 2015
- Analytical Metric Weight Generation for Multi-Domain Trust in Autonomous Underwater MANETs. IEEE UComms 2016
- Analysis of Trust Interfaces in Autonomous and Semi-Autonomous Collaborative MHPC Operations, The Technical Cooperation Program, Portsmouth, UK 2014.
- A Multi-Vector Trust Framework for Autonomous Systems, AAAI 2014.

Erratta

Erratta

- Many small typographic issues corrected
- Out-of-order paras in 4.2.5 (Top should be bottom)

State-of-the-field

Trust

- Interesting general move towards decentralised trust[**Korzun2015**]
- Ditto cohort based relative trust assessment [**Singh2016**]
- Increasing use of ML techniques to assess contextual trust dynamically [**Rishwaraj2017**]
- Human Factors emerging as a increasingly vital area of research [**Saeidi2009**, **Matthews2016**, **Lahijanian2016**]
- Novel/Updated techniques for generalised TMF assessment emerging [**Janiszewski2016**]

State-of-the-field

Acomms

- Assumptions of Gaussian noise naive for real applications [**Mahmood2016**, **Deane2016**]
- The Beaufort Sea has fundamentally changed it's characteristics in 20 years and highlights fundamental flaws in channel modelling assumptions [**Schmidt2016**]
- Higher-Stack level functionality problems remain open(i.e. MAC+Route+ID+Interop) [**Diamant2016** , **Petroccia2016a**, **Petroccia2016b**, **Anjangi2016**]
- Assumptions on increasing passive localisation proving accurate [**Vio2016**, **Ferreira2016**, **Das2016**]

UNIVERSITY OF LIVERPOOL

Expression of terms and context

Focus On

- Trust
- Autonomy
- Decentralised networks
- Harsh Environments

Stated deficiencies in

- Single Metric Trust
- Systemic Trust
- Lack of modelling of Trust in Harsh environments

Deep background

Focus On

- Network/Graph concepts
- Routing
- Trust Perspectives and Models
- Trust Relationships
- Multi-Party Trust
- Trusted Threats
- Autonomy and Design constraints of Autonomous Systems
- Current Trust Management Frameworks

Key Outcomes

- Definition of Trust
- Constraints of Autonomy
- Threats to Trust
- Threats to MANETs
- Need for Trust in Autonomous Systems

Deep background

Focus On

- Marine Acoustics
- AComms Modelling
- AUV Operations
- Need for Trust in AUV AComms

Key Outcomes

- Channel Emulation Models
- Selection of characteristic constraints
- Operational Threat Surface
- Operational / Mechanical constraints

Original Work

Focus On

- Comparative factors between UAN/WLAN
- Application of TMF to each environment (terre/aqua)
- Relevant Metric Selection re AComms
- MTFM weight variation assessment and regression

Key Findings

- Modelled optimal performance range @ $\approx$0.015-0.025pps/100-300m node separations
- MTFM outperforms single metric TMFs for selected misbehaviours
- MTFM dimensional weighting further improves performance and tolerance
- Long collection times due to sparsity can impact trust assessment relevance

Chapter 5: Use of Physical Behaviours for Trust Assessment

Original Work

Focus On

- Physical Misbehaviours and Metrics
- "Failure" vs "Misbehaviour" vs "Malice"
- AUV Kinematics
- Metric variability in collaborative collision avoidance (flocking)
- Metric based classifier (Q-test based, not "Trust")

Key Findings

- First physical misbehaviour detection system
- Identified clear differentiating observations in different composite metrics
- Highly accurate blind behaviour classifier ($\approx$ 0% FP, $\gtrsim$ 90 % TP)

Original Work

Focus On

- Combination of Comms. & Phys. Metrics
- Domain Specific Behaviour in Cross Domain Metric Space
- Random Forest based metric significance correlation to build H weighting vector for MTFM
- Relative significance between behaviour domain and metric domain grouping $(\Delta T, \Delta T^{-})$
- Generation and Appraisal of alternate/targeted "domains"

Key Findings

- Metric Domains and Behaviour Responses not "naturally" coupled
- Inherent redundancy (eg INDD/RSSI) allows differential behaviour to be detected
- Application level selfishness (STS) very difficult to identify
- Extended C4 behaviour based optimisation of MTFM to dynamically select relevant metrics inclusion

Done

Σ

- UWA Multi Metric/Domain Trust
- Detection of non-comms misbehaviours/fouling even just using comms metrics
- Methodology for exploring / training / metric relevance
- Minimal performance specification
- UWA Trust is Hard & it's mostly the channels' fault
- Single-Metric Trust is unstable in such an environments
- Multi-Metric Trust works & can discriminate behaviours
- Not all metrics are equally useful
- Simple classifiers can be V good in some behaviours (MPC)
- - can be not so good for others (STS)

Left/Next

$\notin$

- Smarter Detection Classifier
- Cooperative / Periodic / Variable attack profiles
- Commonality of detection filters across Multiple-base scenarios
- Heterogenous Node capabilities
- Real experiments and Cross sim-implementations

References I

Thank You / QnA?

Say Hello!

- bolster@liv.ac.uk
- me@bolster.online
- @bolster
- bolster.online
- andrewbolster
- "Andrew Bolster"
- bolster

Fig 1.1 Multi-Domain Threat Surface

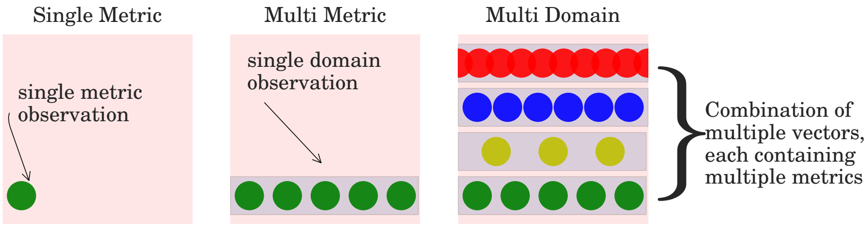


Fig. 1: Multi-Domain Threat Surface

Tab 2.3 Definitions of Trust

Table 1: Selected Definitions of Trust

Definition	Source
Assured reliance on the character, ability, strength, or truth of someone or something.	Merriam-Webster
Firm belief in the reliability, truth, or ability of someone or something	OED
The willingness of a party to be vulnerable to the actions of another party based on the expectation that the other will perform a articular action important to the trustor, irrespective of the ability to monitor or control that other party	**Mayer1995**
An expectancy held by and individual or a group that the word, promise, verbal or written statement of another individual or group can be relied upon	**Rotter1967**

Fig 2.5 Model of Trust

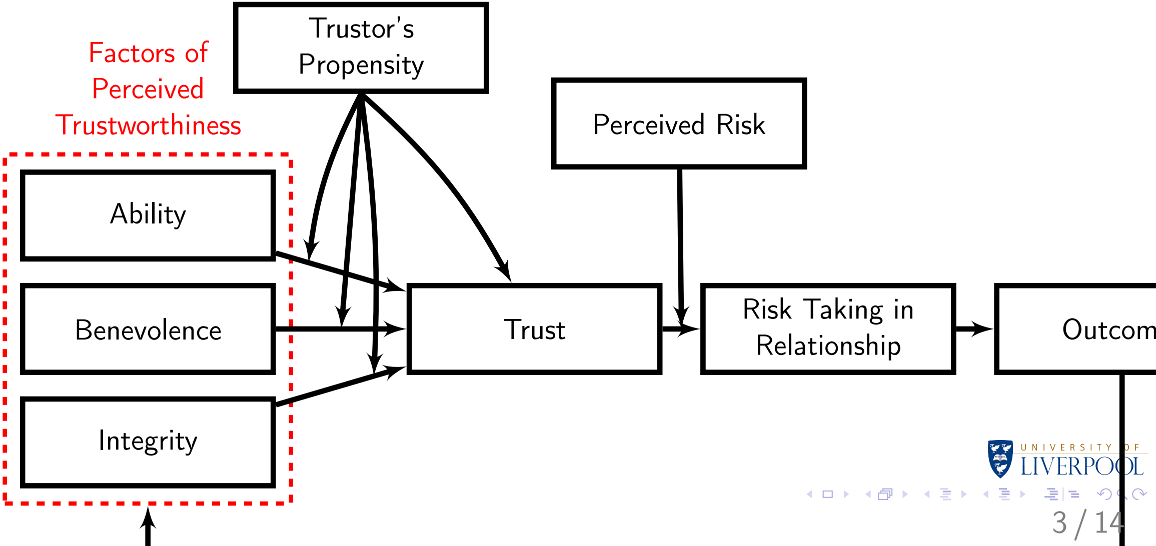


Fig 2.6 Trust Construct Relationships

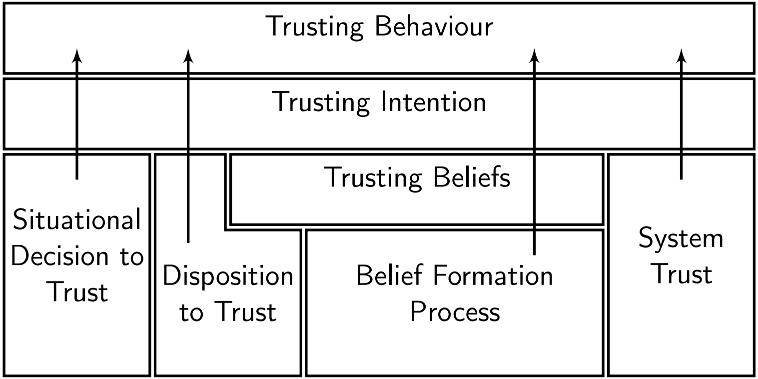
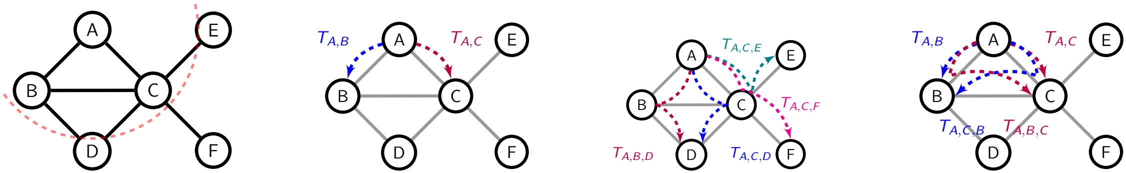


Fig. 3: Trust Construct Relationships (from **Liu2010**)

Fig 2.10 Trust Topologies



(a) Sample topology showing logical connections between nodes (Range of A shown in red dashed line)

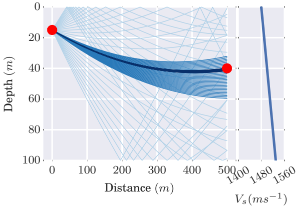
(b) Direct Relationships, the two possible trust assessments from A to its connected neighbours, B, C

(c) Indirect Relationships, showing the four possible trust assessments from A or the three disconnected leaf nodes, D, E, F

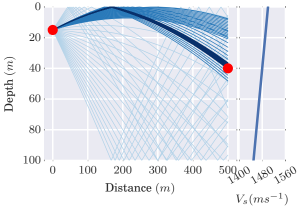
(d) Recommender Relationship, showing the two discrete paths trust assessments travel to A; $T^R_{A,B} = T_{A,C} \cdot T_{C,B}$ and $T^R_{A,C} = T_{A,B} \cdot T_{B,C}$

Fig. 4: Trust Topologies; Direct, Indirect, Recommender, etc. from the perspective of Node A

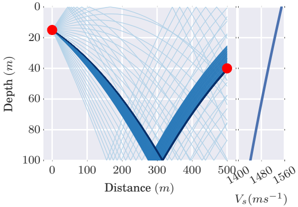
Fig 3.3: Bellhop Model



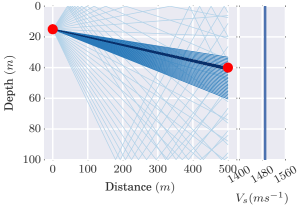
(a) Linear Increasing



(b) Linear Decreasing



(c) Quadratic



(d) Isovelocity

UNIVERSITY OF LIVERPOOL

Communications Channel Considerations

Key Characteristics of the Marine Acoustic Channel:**Urick1983a**, **Partan2006**, **Stojanovic2007**, **Stefanov2011**

- Slow propagation ($1400ms^{-1}$) incurring long delays
- Inter-symbol interference
- Doppler Spreading
- Non-Linear propagation due to refraction
- Fast & Slow fades from environmental factors (flora/fauna/surface and seabed conditions)
- Freq. dependant attenuation
- Significant destructive multipath effects

Attenuation in the Marine Acoustic Channel

The attenuation that occurs in an underwater acoustic channel over distance d about frequency f is given as $A_{\text{aco}}(d, f) = A_0 d^k a(f)^d$ or

$$10 \log A_{\text{aco}}(d, f)/A_0 = k \cdot 10 \log d + d \cdot 10 \log a(f) \tag{1}$$

where A_0 is a normalising constant, k is a spreading factor, and $a(f)$ is the absorption coefficient;**Stefanov2011**

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

Attenuation in the Marine Acoustic Channel

The attenuation that occurs in an underwater acoustic channel over distance d about frequency f is given as $A_{\text{aco}}(d, f) = A_0 d^k a(f)^d$ or

$$10 \log A_{\text{aco}}(d, f)/A_0 = k \cdot 10 \log d + d \cdot 10 \log a(f) \tag{1}$$

where A_0 is a normalising constant, k is a spreading factor, and $a(f)$ is the absorption coefficient;**Stefanov2011**

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

Compared to RF Free space PL: $\left(A_{\text{RF}}(d, f) \approx \left(\frac{4\pi df}{c}\right)^2\right)$

- Exponential in d: $A_{\text{aco}} \propto f^d$ vs $A_{\text{RF}} \propto (df)^2$
- f factor four orders higher in $f \propto A_{\text{aco}}$ vs $f \propto A_{\text{RF}}$

Multi-Metric TMF - Grey Grading

$$\theta^t_{k,j} = \frac{\min_k |a^t_{k,j} - g^t_j| + \rho \max_k |a^t_{k,j} - g^t_j|}{|a^t_{k,j} - g^t_j| + \rho \max_k |a^t_{k,j} - g^t_j|} \tag{3}$$

$$\phi^t_{k,j} = \frac{\min_k |a^t_{k,j} - b^t_j| + \rho \max_k |a^t_{k,j} - b^t_j|}{|a^t_{k,j} - b^t_j| + \rho \max_k |a^t_{k,j} - b^t_j|} \tag{4}$$

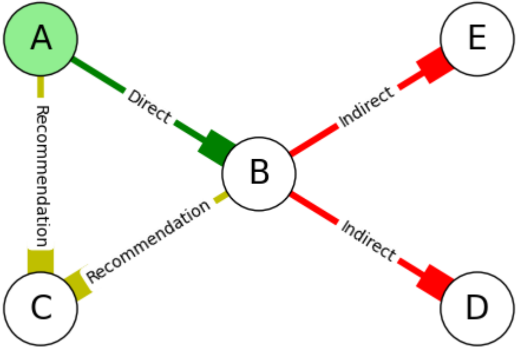
$$[\theta^t_k, \phi^t_k] = \left[\sum_{j=0}^{M} h_j \theta^t_{k,j}, \sum_{j=0}^{M} h_j \phi^t_{k,j}\right] \tag{5}$$

$$T^t_k = (1 + (\phi^t_k)^2 / (\theta^t_k)^2)^{-1} \tag{6}$$

Where $a^t_{k,j}$ is the value of an observed metric x_j for a given node k at time t, g and b are respectively the "good" and "bad" reference metric sequences from $\{a^t_{k,j} k = 1, 2 \ldots K\}$, $H = [h_0 \ldots h_M]$ is a metric weighting vector such that $\sum h_j = 1$

Includes shared assessments from other nodes weighted based on their relative topology to provide a final value[1]

$$T_{i,j}^{MTFM}$$



1 Details

$$
\begin{aligned}
T_{i,j}^{MTFM} =& \frac{1}{2} \cdot \max_s \{f_s(T_{i,j})\} T_{i,j} \\
&+ \frac{1}{2} \frac{2|N_R|}{2|N_R| + |N_I|} \sum_{n \in N_R} \max_s \{f_s(T_{i,n})\} T_{i,n} \\
&+ \frac{1}{2} \frac{|N_I|}{2|N_R| + |N_I|} \sum_{n \in N_I} \max_s \{f_s(T_{i,n})\} T_{i,n}
\end{aligned} \tag{7}
$$

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...

Grey Trust Equs II

$$
\begin{aligned}
f_1(x) &= -x + 1 \\
f_2(x) &= \begin{cases} 2x & \text{if } x \leq 0.5 \\ -2x + 2 & \text{if } x > 0.5 \end{cases} \\
f_3(x) &= x
\end{aligned}
\tag{8}
$$

Back

UNIVERSITY OF LIVERPOOL

System Model Constraints

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

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

UNIVERSITY OF LIVERPOOL

Metric Selection/Weighting

Table 3: ΔT_{ix} behaviour detection performance across meta-domains, including selected metrics

	Domain	Behaviour ΔT_{ix} MPC	STS	Shadow	SlowCoach	Mean	Metrics in Domain *Delay*	P_{RX}	P_{TX}	S	G	*PLR*	*INDD*	*INHD*	*Speed*
Basic	Full	0.81	-0.03	0.42	0.60	0.45	✓	✓	✓	✓	✓	✓	✓	✓	✓
	Comms	0.85	0.04	0.19	0.26	0.34	✓	✓	✓	✓	✓	✓			
	Phys	0.04	0.00	0.39	0.69	0.28							✓	✓	✓
Alternate	Comms alt.	0.85	0.03	0.38	0.45	0.43				✓	✓	✓	✓		
	Phys alt.	0.48	0.03	0.42	0.63	0.39	✓	✓					✓	✓	✓
Synthetic	MPC	0.89	0.01	0.35	0.54	0.45	✓	✓	✓					✓	
	STS	0.86	0.06	0.37	0.49	0.45	✓		✓	✓		✓	✓		
	Shadow	0.49	-0.00	0.44	0.66	0.40		✓					✓	✓	✓
	SlowCoach	0.47	0.00	0.37	0.72	0.39	✓	✓		✓					✓
	Mean	0.88	0.03	0.42	0.69	0.50		✓	✓		✓		✓		✓