An Investigation into Trust and Reputation Frameworks for Collaborative Teams of Autonomous Underwater Vehicles

Andrew Bolster

University of Liverpool

andrew.bolster@liv.ac.uk

June 4, 2014

Research Context

- Project launched at QUB ECIT in 2011 under the DSTL/DGA Anglo French Defence Research Group PhD Programme
- What lessons from the Mobile Ad Hoc Network (MANET) space can be transferred to the marine environment?
- Teams of 3 - 16 Autonomous Underwater Vehicles (AUVs) Mine countermeasures, Hydrography, and Patrol Capabilities (MHPC)
- Defence focus, assumption of highly capable enemy attempting to compromise communications / operations
- Primary Simulation/Analysis work done in 12/13
- Moved to UoL Oct 13 after 2 mth placement @ DSTL PDW Naval Systems / Information Systems departments.
- CDE Project on Precision Timing for Positioning with NPL/Plextek

Trust in Ad-Hoc Systems and the context of this document

- Particularly interested in the application of Trust in Decentralised (P2P) Autonomous Systems of Systems, Autonomous Underwater Vehicles (AUVs) for example

Trust in Ad-Hoc Systems and the context of this document

- Particularly interested in the application of Trust in Decentralised (P2P) Autonomous Systems of Systems, Autonomous Underwater Vehicles (AUVs) for example
- Trust: *The expectation of an actor performing a certain task or range of tasks within a certain confidence or probability*

Trust in Ad-Hoc Systems and the context of this document

- Particularly interested in the application of Trust in Decentralised (P2P) Autonomous Systems of Systems, Autonomous Underwater Vehicles (AUVs) for example
- Trust: *The expectation of an actor performing a certain task or range of tasks within a certain confidence or probability*
- Full System Views of Trust
    - Design Trust - that a system of systems will perform as spec'd / designed in operation

Trust in Ad-Hoc Systems and the context of this document

- Particularly interested in the application of Trust in Decentralised (P2P) Autonomous Systems of Systems, Autonomous Underwater Vehicles (AUVs) for example
- Trust: *The expectation of an actor performing a certain task or range of tasks within a certain confidence or probability*
- Full System Views of Trust
  - Design Trust - that a system of systems will perform as spec'd / designed in operation
  - Operational Trust - the systems within a larger system will perform as designed in field ✓

Trust Management Frameworks

- Provide information regarding the estimated future states and operations of nodes within networks

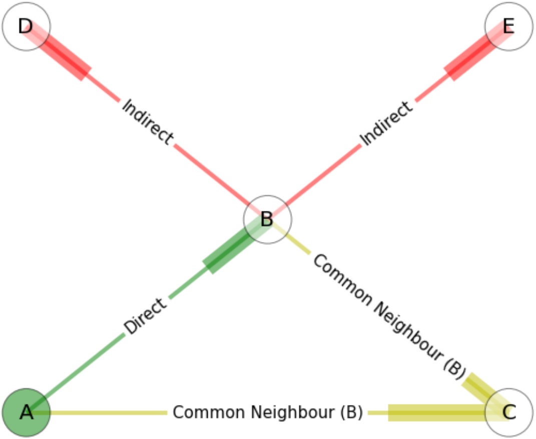
Trust Management Frameworks

- Provide information regarding the estimated future states and operations of nodes within networks
- "[. . .]collecting the information necessary to establish a trust relationship and dynamically monitoring and adjusting the existing trust relationship" - [4]

Trust Management Frameworks

- Provide information regarding the estimated future states and operations of nodes within networks
- "[. . .]collecting the information necessary to establish a trust relationship and dynamically monitoring and adjusting the existing trust relationship" - [4]
- Enables nodes to form collaborative *opinions* on their cohort nodes based on
  - Direct Observation of Communications Behaviour (eg Successfully Forwarded Packets)
  - Common-Neighbour Recommendation
  - Indirect Reputation

Transitivity in Trust Networks



TMFs in Ad Hoc Autonomous Systems

- Multiple transitive relationships can be maintained over time, providing trust resilience with dynamic network topology

TMFs in Ad Hoc Autonomous Systems

- Multiple transitive relationships can be maintained over time, providing trust resilience with dynamic network topology
- Enable trust establishment from partial-strangers via indirect trust and direct observation

TMFs in Ad Hoc Autonomous Systems

- Multiple transitive relationships can be maintained over time, providing trust resilience with dynamic network topology
- Enable trust establishment from partial-strangers via indirect trust and direct observation
- Enables nodes to inform internal processes for global efficiency given observed network behaviour / 'wellness', similar to those found in human social networks eg
  - Update routing table based on 'safest' node chains (Phone Tree)
  - Maneuver away from misbehaving nodes (Shunning)
  - Inform as to 'trustworthiness' of forwarded information (Healthy sense of Skepticism)
  - Historic Distrust/Trust decaying over time (Forgiveness/Relationship Decay)

Reason for using TMFs in MANETs

- Provide Risk Mitigation against many classical MANET attacks
  - Black/Grayhole
  - Routing Loop
  - Selective misbehaviour / selfishness
- Generally; to constrain potential malicious behaviour that can operate without detection

Trust in Autonomous Systems

- Public Key Infrastructure - Requires Centralised Control and pre-shared keys
- Resurrecting Duckling - Uses in-action keying with a trusted source
- Evidence Based Trust - Uses shared keys
- Reputation Based Trust - Uses Packet forwarding success rate for prediction of future actions
  - CONFIDANT - Trust-based router implementation using packet forwarding rate
  - OTMF - Trust including transitive information from other nodes

- . . . and there are plenty more along the same lines
- Predominantly use single metrics or only communications metrics

Trust in Autonomous Systems

- Public Key Infrastructure - Requires Centralised Control and pre-shared keys
- Resurrecting Duckling - Uses in-action keying with a trusted source
- Evidence Based Trust - Uses shared keys
- Reputation Based Trust - Uses Packet forwarding success rate for prediction of future actions
  - CONFIDANT - Trust-based router implementation using packet forwarding rate
  - OTMF - Trust including transitive information from other nodes
  - MPTM - Relationships and Multiple Metrics combined with Gray Interval assessment
- . . . and there are plenty more along the same lines
- Predominantly use single metrics or only communications metrics

Vectorised Trust

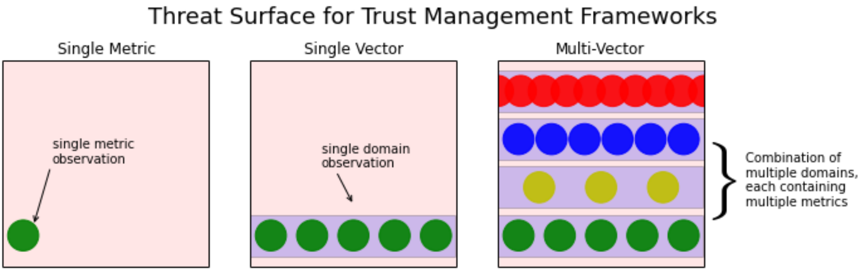
- Application of several individual metrics for the construction of a single trust measurement
- For example:
  - $X = \{packet\ loss, signal\ strength, datarate, delay, throughput\}$
- This multi-parameter trust prevents 'smart' attackers; leveraging a known trust metric to subvert a TMF without detection
- Normally expressed as a vector, but can be condensed into an abstracted or weighted form for comparison [3]

The Need for Multi-Domain Trust Assessment

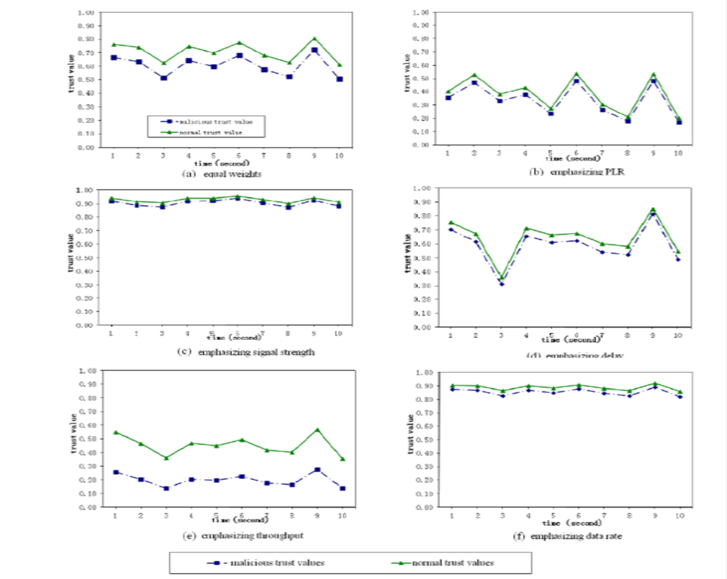
- Communications not the only target for an attacker (or failure);
  - Following to restricted area
  - Masquerading
  - Hardware Degradation
  - Resource attack via propulsive power
- Physical observation presents opportunity to further reduce the available threat surface while also discriminating between 'True' attacks and mechanical failure.
- Also could provide additional 'handshake' protocols for 'friendly' fleets/teams through reactionary behaviours

Multi-Vector Trust and the Threat Surface

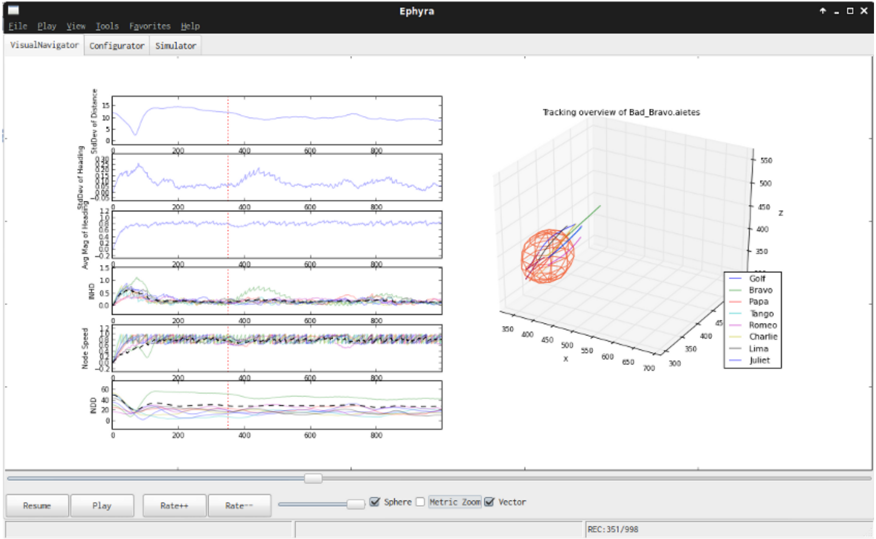
Potential attacks exist across a multi-domain threat surface



Malicious Behaviour Discrimination



Agent Based Behaviour Simulator



Trust in Mobile Autonomous Underwater Vehicles

- Flocking with Intent: MCM, Port Protection, Survey, Protection Detail, etc.

Trust in Mobile Autonomous Underwater Vehicles

- Flocking with Intent: MCM, Port Protection, Survey, Protection Detail, etc.
- Metric Selection in collaboration CMRE/DSTL
  - Inter Node Heading Deviation
  - Inter Node Distance Deviation
  - Node Speed

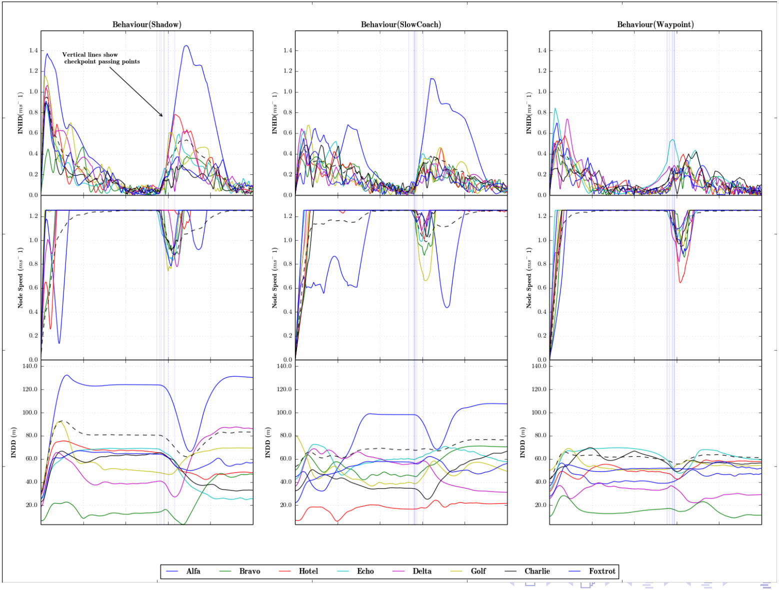
Trust in Mobile Autonomous Underwater Vehicles

- Flocking with Intent: MCM, Port Protection, Survey, Protection Detail, etc.
- Metric Selection in collaboration CMRE/DSTL
  - Inter Node Heading Deviation
  - Inter Node Distance Deviation
  - Node Speed
- Behaviour selection for testing
  - Shadow
  - Spy
  - Sloth
  - Stalker
  - Scoundrel

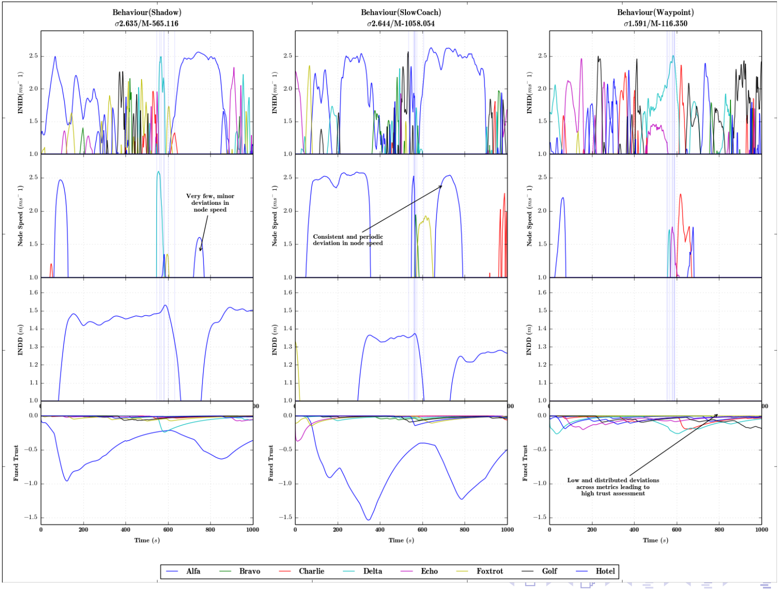
Trust in Mobile Autonomous Underwater Vehicles

- Flocking with Intent: MCM, Port Protection, Survey, Protection Detail, etc.
- Metric Selection in collaboration CMRE/DSTL
  - Inter Node Heading Deviation
  - Inter Node Distance Deviation
  - Node Speed
- Behaviour selection for testing
  - Shadow
  - Spy
  - Sloth
  - Stalker
  - Scoundrel
  - Slow Coach (non-malicious)
  - Spin Doctor (non-malicious)

Raw Behavioural Metric Assessment in AUVs



Behavioural Trust Assessment in AUVs



Behavioural Trust Assessment in AUVs

- Detection and identification based on basic weight-assessment classifier against windowed history of observations, with confidence based on a Grey Theoretic weight
- Currently >96% statistical accuracy of detection and confidence, but this needs much more rigorous analysis

Challenges in Multi-vector Trust

- How to define optimality in trust assessment when dealing with multiple vectors and transitive trust?
- Is there a quantifiable benefit to cross-domain comparison beyond single vector Trust?
- Is there an optimal generic cross-domain comparator?

Current Publications

- A Multi-Vector Trust Framework for Autonomous Systems [2]
    - Symposium paper to the Association for the Advancement of Artificial Intelligence on the current state of work, presenting our progress towards multi-vector trust
- Analysis of Trust Interfaces in Autonomous and Semi-Autonomous Collaborative MHPC Operations [1]
    - Part of a Five-Eyes defence strategy programme (TTCP) for assuring C3I capabilities as part of FF2020

Development Plan

1. Behaviour Detection (Q3 14) - Formal Analysis of Behavioural Trust Systems
   - ASON 2014 : Seventh Int. WS on Autonomous Self-Organizing Networks (Aug 14)
   - AHUC 2014 : The Fourth Int. WS on Ad Hoc and Ubiquitous Computing (Aug 14)
   - ICCAR 2015 : WASET Int. Conf. on Control, Automation and Robotics (Dec 14)
2. MANET/Marine comparison (Q4 14) - Formal Comparison between Terrestrial MANET / Marine contexts
3. Multi-Domain Trust Assessment (Q4 14) - Combination of Communicative and Physical Behaviour Trusts
   - IEEE Trans. on Communications / Dependable and Secure Computing / Intelligent Systems
4. Reactionary/Perturbative Trust (Q1 15) - Exploration of reactionary behaviours for teams to 'shake down' suspects
   - SASO15:Self-Adaptive and Self-Organizing Systems,
   - SEAMS15: Software Engineering for Adaptive and Self-Managing Systems

Thesis plan

- Abstract, Acknowledgements, Introduction,
- Background Information on Trust and it's applications to MANETs
- Background Information on Maritime Uses of Autonomous Systems
- Trust in Autonomous Systems of Systems for Maritime Defence Applications
- Strategies for Multi-Domain Trust Assessment
- Modelling and Analysis of Collaborative Node Kinematic Behaviours in Underwater Acoustic MANETS
- Comparative Analysis of Multi-Domain Trust Assessment in Collaborative Mobile Networks
- Reactionary Behaviours to increase decentralised trust in isolated environments
- Conclusions, Bibliography

References I

Andrew Bolster. *Analysis of Trust Interfaces in Autonomous and Semi-Autonomous Collaborative MHPC Operations*. Tech. rep. The Technical Cooperation Program, 2014.

Andrew Bolster and Alan Marshall. "A Multi-Vector Trust Framework for Autonomous Systems". In: *2014 AAAI Spring Symposium Series*. Stanford, CA, 2014, pp. 17–19. URL: http://www.aaai.org/ocs/index.php/SSS/SSS14/paper/viewFile/7697/7724.

References II

- Ji Guo, Alan Marshall, and Bosheng Zhou. "A New Trust Management Framework for Detecting Malicious and Selfish Behaviour for Mobile Ad Hoc Networks". In: *2011IEEE 10th International Conference on Trust Security and Privacy in Computing and Communications* (2011), pp. 142–149. DOI: 10.1109/TrustCom.2011.21. URL: http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=6120813.
- Huaizhi Li and Mukesh Singhal. "Trust Management in Distributed Systems". In: *Computer* 40.2 (2007), pp. 45–53. ISSN: 00189162. DOI: 10.1109/MC.2007.76. URL: http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=4085622.

The End