An Investigation into Trust and Reputation Frameworks for Autonomous Underwater Vehicles

Research Update and Plan Detail

Andrew Bolster

[1] Institute of Electronics, Communication, and Information Technology
Queen's University Belfast

Outline

Outline

Current Context of work: Use of physical behaviours to assess trust without wasting communications time / energy.

Motivations

- Direct Secure communications are expensive and time consuming
- Centralised security mediation is a single point of failure
- Timing technology has reached a point at which accurate (secure) localisation within a trusted fleet is relativly simple (CSACs)

Outline

Trust could be required for:

- Alien[1] Node joining Secure Fleet
- Known node (re)joining Fleet
- bi-directional Fleet authentication with USV
- Node 'disappearance'

It is also expected that a trust metric for a particular fleet should be stable outside these activities

[1] In this sense, Alien does not necessarily imply unauthorized, but simply that it has not interacted with the fleet previously

A composite Trust metric (consisting of a weighted collection of individual metric observations over time) must hold the following qualities

- Associativity - Fleet will settle to same value given same weightings regardless of initial configuration
- Reflexitivity - Fleet Trust should be stable to temporary additions/subtractions:

$$\begin{cases} T_0 = T(Fleet) \\ T_1 = T(Fleet - x) \neq T_0 \\ T_2 = T((Fleet - x) + x) = T_0 \end{cases} \tag{1}$$

A composite trust metric **may** also:

- exhibit distributivity - in the case where a fleet is temporarily minus a fleet member, it may be the case that the difference in trust behaviours is observable enough to allow a rogue entity to sufficiently emulate the missing member to fool the fleet into trusting it's behaviour
- indicate/identify technical failure - in the case where a trust metric is impacted by some technical fault, it may be possible to not only recogise the difference between a technical trust failure and a rogue operation, but to identify the node and even the subsystem that is failing

All in all, a composite metric should follow the properties of a vectorized binomial opinion aka beta distribution in subjective logics.

This would incorporate belief, uncertainty, expectation, and allow the a-priori establishment of trust within a formal system.

A single observational trust metric should be:

- Atomic - one measurement for one factor, so as to allow for accurate and causitive statistical analysis
- Orthogonal to others - Again, metrics should not indirectly measure the same physical factors

Outline

Bespoke Simulation framework consisting of three modules:

- `Aietes` - Ancient King of Ephyra in the Illiad
- `Bounos` - Recieved the kingdom of Ephyra from Aietes
- `Ephyra` - An ancient kingdom near modern day Parga in Greece

- `Aietes` : the original base behavioural simulator, performing agent-based modeling of the motions of AUV's within an environment
- `Bounos` : a collection of data processing and collation functions
- `Ephyra` : a global visualisation (and later, control) system for both Aietes and Bounos

These in total currently consist of nearly 2600 lines of python runtime code, including animation and GUI navigation capabilities.

- Arbitrary node configurations (both in terms of physical and communications capabilities)
- Generically Based on the REMUS 100 configuration and physical model
- Support for runtime and a-postori statistical analysis with numpy/scipy/pandas
- Componentized Behaviour network, with a 3-rule flocking base

Flocking:

$$v_{t+1} = v_t + \sum_{\forall b \in B} b_v(p_t, v_t) \cdot b_f \tag{2}$$

where: $b_v(p_t, v_t)$ is the individual force vector exerted by a given behaviour, and b_f is the user-controlled weight of that behaviour

Cruising Behaviour

$$p_{t+1} = p_t + \begin{cases} v_{t+1} & v_{t+1} \leq v_{cruising} \\ \frac{v_{t+1}}{|v_{t+1}|} \cdot \frac{1}{e^{-(|v_{t+1}| - v_{cruising})}} & v_{t+1} > v_{cruising} \end{cases} \tag{3}$$

- Attraction to a point

$$F_A(p, p_A, d_\infty) = (\widehat{p - p_A}) \cdot \frac{|p - p_A|}{d_\infty} \tag{4}$$

- Repulsion from a point

$$F_R(p, p_R, d_\infty) = (\widehat{p_R - p}) \cdot \frac{d_\infty}{|p - p_R|} \tag{5}$$

- in both cases, the d_∞ variable is set to be a limiting distance i.e. objects within the radius d_∞ produce attractive factor > 1

The urge to attract to the center of gravity of the fleet

$$F_{j,C} = F_A\left(p_j, \frac{1}{N}\sum_{\forall i \neq j}^{N} p_i, d_{collision}\right) \quad (6)$$

The urge to avoid local collisions

$$F_{j,H} = \sum_{\forall i \neq j}^{N} F_R\left(p_j, p_i, d_{collision}\right) \Big| d_{collision} > \|p_i - p_j\|\Big) \tag{7}$$

The urge to maintain a globally average heading

$$F_{j,CA} = \frac{1}{N} \cdot \left(\sum_{\forall i \neq j}^{N} \hat{v}_i \right) \tag{8}$$

The urge to head towards a goal / waypoint

$$F_{j,W} = F_A(p_j, p_w, \frac{d_\phi}{2} \tag{9}$$

where d_ϕ the satisfaction distance of a waypoint, i.e. the success distance from a positional waypoint

- The configuration of Aietes is quite delicate and complex and needs an overhaul
- Aietes doesn't make it easy to extract the **causes** of decisions after they are made due to its agent-based nature
- Application of vector-weights (see next slide) has been more fraught than expected

All three of these are being quickly solved in parallel development with Bounos

One proposed method of securing a fleets behaviour is to apply a vector rather than scalar relation to it’s behaviour weights, i.e.

Different Nodes behave differently within the fleet; prefer different limiting distancs, weight repulsion more than attraction, etc.

This would drastically increase the complexity of any observer deriving these values.

Ephyra is used to process the simulated path files and perform a-postori analysis of the fleets behaviour, in the same way an observer would.[2]

- A wireframe sphere representing the gravity of the fleet (colured based on the standard deviation from the centre of the fleet) can be overlaid
- Weighted Vectors can also be overlaid showing individual node headings
- The LHS panel also shows the time series responses of a selection of metrics

[2] i.e. these analyses are not subject to the same 'fudging factors' that are applied to the simulated nodes obeservations of eachother

The current defaults for this metric displays are:

- Positional Standard Deviation from the Fleet CoG
- Average Inter-Node distances
- Standard Deviation of headings from fleet average
- Average Speed of the fleet (based on average heading)
- Min, Mean, and Max speeds of nodes within the fleet

overview.png

Analysis
Dynamic Trail

dynamic_tail.png

Demo: If Remote Desktop is working. . .

Outline

Establish Behaviour Trust metric within perfect network[3]:

1. Assess observability of behaviour factors in Normal Mission Profiles (port protection, shadowing, minesweeping, survey)
2. Assess behaviour of principal observable components with NMP with selective node failures (total, instant failures)
3. As above with fresh authorised nodes being introduced
4. As above with alien nodes being introduced (non authenticated, but normal behaviour)
5. As above with rogue nodes introduced (bad behaviour; eg falling behind, pushing ahead, false-heading, etc)

[3] No communications loss, perfect realtime knowledge of locations and headings of nodes

Proposed Experiments
Within Next Six Months

1. Same situations but with non-fleet static and dynamic obstacles
2. Same situations but with communications-based information
3. Same situations with communications based information and information warfare (i.e. lying)

This proposed metric will be assessed for resiliance, and accuracy at each stage and should be publishable, at least in the journal of the Marine Technology Society, if not IEEE Trans. Comms

This will provide a bed upon which to build a transactional protocol for marine trust that could be integrated with communications trust systems (i.e. those that use communications artefacts as their base metrics)

Outline

The communications segments will use the SUNSET framework, which has been heavily verified.

The use of SUNSET as a communicative and control layer also allow for easy implementation of any developed framework onto hardware availabile at CMRE and other facilties.

Expected outcomes

- Compound Metric definition for behavioural trust in any communications environment.
- Identification schema for fleets/nodes based on behaviour factors

Summary

- Behaviour of a fleet of flocking individuals is more interesting that I thought
- The biggest blocker so far is lack of practical data
- This isn't guaranteed to work, but if it provably can't work, that's still very positive
- Outlook / Immediate Action points
  - Generation of stable behavioural vectors is not solved
  - Real-Time interaction with simulations is not solved

For Further Reading I

Chiara Petrioli and Roberto Petroccia,

SUNSET: Simulation, Emulation and Real-life Testing of Underwater Wireless Sensor Networks,

Proceedings of IEEE UComms 2012, (Sestri Levante, Italy), IEEE Computer Society.

Karim Konate and Abdourahime Gaye,

Attacks Analysis in Mobile Ad Hoc Networks: Modeling and Simulation.

2011 Second International Conference on Intelligent Systems, Modelling and Simulation

Andrea Caiti

Cooperative distributed behaviours of an AUV network for asset protection with communication constraints

OCEANS, 2011 IEEE-Spain

Qiuling Jia, Guangwen Li

Formation Control and Obstacle Avoidance Algorithm of Multiple Autonomous Underwater Vehicles(AUVs) Based on Potential Function and Behavior Rules.

2007 IEEE International Conference on Automation and Logistics