Agent-based trust management and prediction using D3-FRT

Funmi Onolaja

Rami Bahsoon

School of Computer Science

The University of Birmingham, UK.

Georgios Theodoropoulos

IBM Research, Dublin, Ireland.

International Conference on Computational Science, ICCS2012

UNIVERSITY OF BIRMINGHAM

Outline

- ✓ Introduction to trust computational models
- ✓ Issues and requirements
- ✓ Our approach (D3-FRT: Dynamic Data Driven Framework for Reputation and Trust management)
- ✓ Results of empirical study

Trust Computational Models

- ✓ **Motivation**
- ✓ **Trust and Reputation**

- ✓ **Trust management**

Trust computational models e.g. CORE, CONFIDANT, TrustGuard, Feedback forum on eBay etc

- ❖ Domains in WSNs, MANETs, P2P, Online communities etc
- ❖ Distinguish members
- ❖ Incentives / punishment
- ❖ Trust Values (TVs)

Trust management: issues

- ✓ Assumption that past behaviour is an indication of future behaviour.
- ✓ Not true with *intoxication.* Simple example, WSN for traffic management.
- ✓ Difficult to identify - sudden misbehaviour.
- ✓ Effect of past good behaviour outweighs the effect of current actions on reputation.
- ✓ *Collusion*: two or more domain entities deceive the system.
- ✓ Watchdog mechanism.

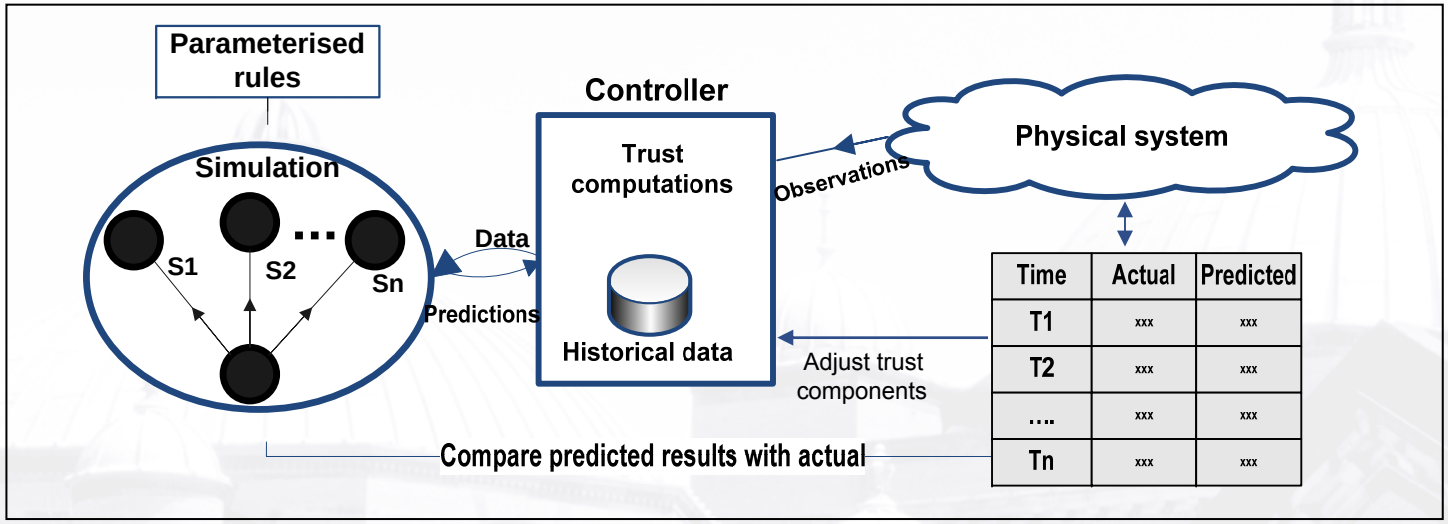
Trust management: dynamic nature

- ✓ Trust dynamics: Trust is not static but dynamic, computation of trust should be equally dynamic.
- ✓ Dynamic approach to identifying and isolating misbehaving (group of) members.

Requirements

- ✓ Dynamic online rating of members (data) to capture sudden changes.
- ✓ Anticipate events and predict TVs.

Framework: components



Physical system

- ✓ Historical and recent online rating with weightings
- ✓ Qualitative data to quantitative value (Controller); Fuzzy like

Simulation

- ✓ Parameterised / behavioural rules
- ✓ Probabilities (Bayesian theory) of collaboration / misbehaviour
- ✓ Scenario-based prediction of future ratings
- ✓ Compare with the real system and adjust weights (feedback)
- ✓ Ties with DDDAS

Empirical study

✓ Used agent based simulation approach (Repast – repast.sourceforge.net) with P2P file sharing network as a case study.

- ❖ Peers in the network remain from the start of the simulation to the end.
- ❖ Exchanges between peers are not necessarily symmetric; for example, a peer A may request a file from peer B whereas peer B might not request a file from A.
- ❖ We assume that all peers in a cluster are mutually connected and do not abort transfers; all interactions are completed.
- ❖ The peers are self-contained as they are uniquely identifiable with a set of attributes: historical, online and predicted TVs.

Empirical study

✓ The peers exhibit different behaviours which include: intoxication, collusion, active file upload and download.

✓ Trust value is computed for each peer per transaction:

- TV = $\mu_h tv_h^R + \mu_o tv_o^R + \mu_f tv_f^S$,

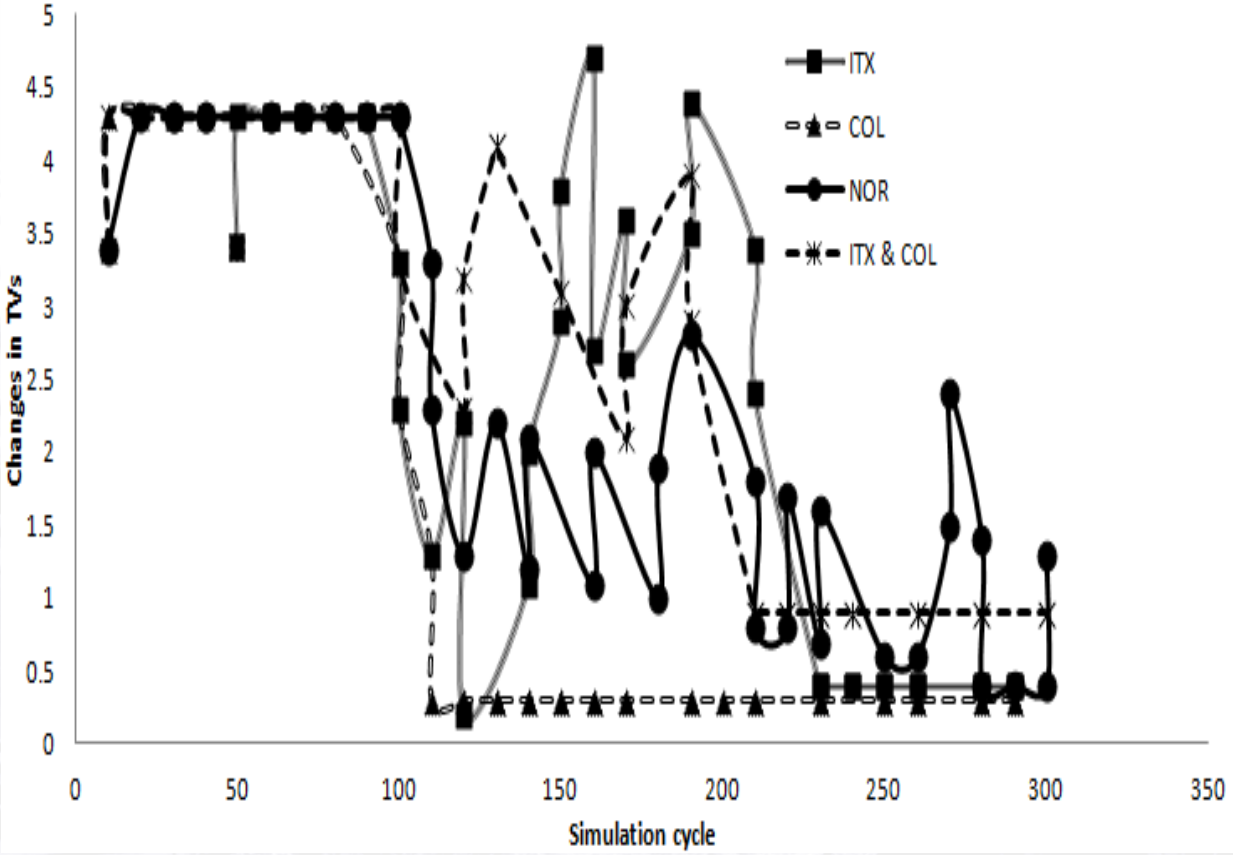
$$tv_f^S = \frac{(tv_o^S)\,S_1 + (tv_o^S)\,S_2 + (tv_o^S)\,S_3 + \cdots + (tv_o^S)\,S_n}{N}$$

- Weights μ_o, μ_h, μ_f: scaling factors for the online, historical and predicted TVs.
- $[\mu_h, \mu_o]>0$ and $\mu_o>\mu_h$, more emphasis on recent behaviour.
- tv_h^R , tv_O^R, tv_f^S and tv_o^S : historical, online TVs in the network, and the predicted and online TVs in simulation.
- Time is measured in simulation *ticks* – a compression of time.
- *N* is number of scenarios and *(S1, S2, . . . Sn)* are the scenarios

Trust dynamics I

- ✓ Purpose: Test the dynamism of D3-FRT.
- ✓ Parameters: 50 peers, 40% misbehaving.
- ✓ Observation: Changes in online TVs as simulation progresses.
- ✓ Lesson learnt: D3-FRT observes and computes ratings online.

Changes in online TVs as simulation progresses



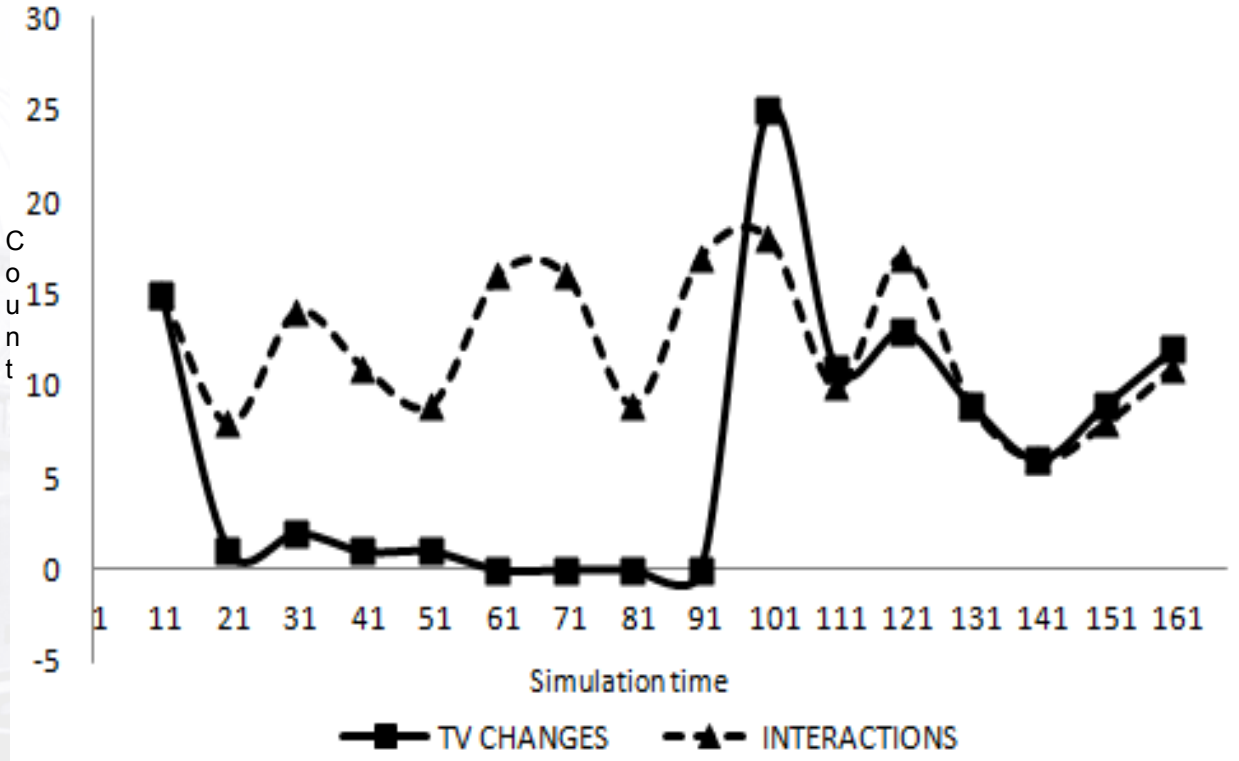
ITX- Intoxication, COL – Collusion, NOR – Cooperating in exchange of files

Trust dynamics II

- ✓ Purpose: Test the dynamism of D3-FRT.
- ✓ Parameters: 50 peers, 40% misbehaving peers.
- ✓ Observation and Lesson learnt: As the simulation progresses and peers collaborate, corresponding trust values are computed per peer and the number of interactions are captured.

This implies that the framework dynamically computes new ratings after each peer interaction.

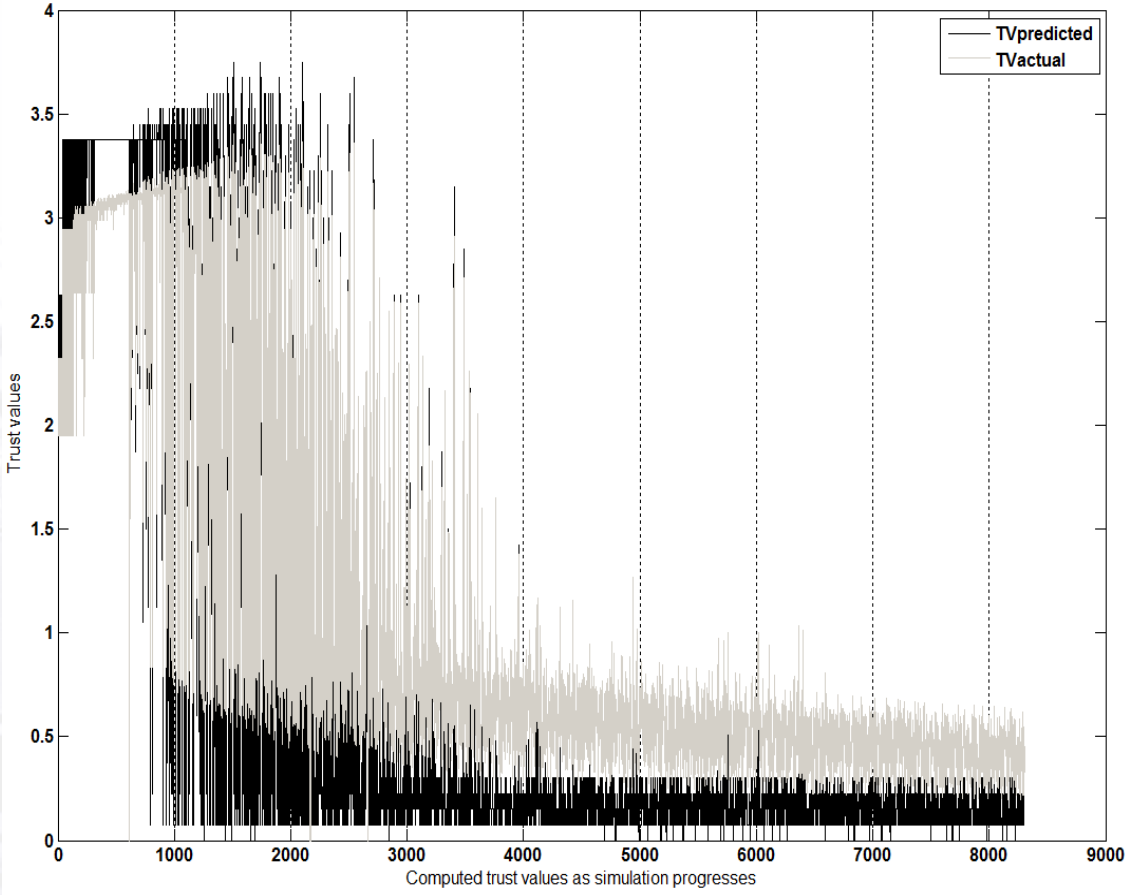
Changes in online TVs as simulation progresses



Estimation error

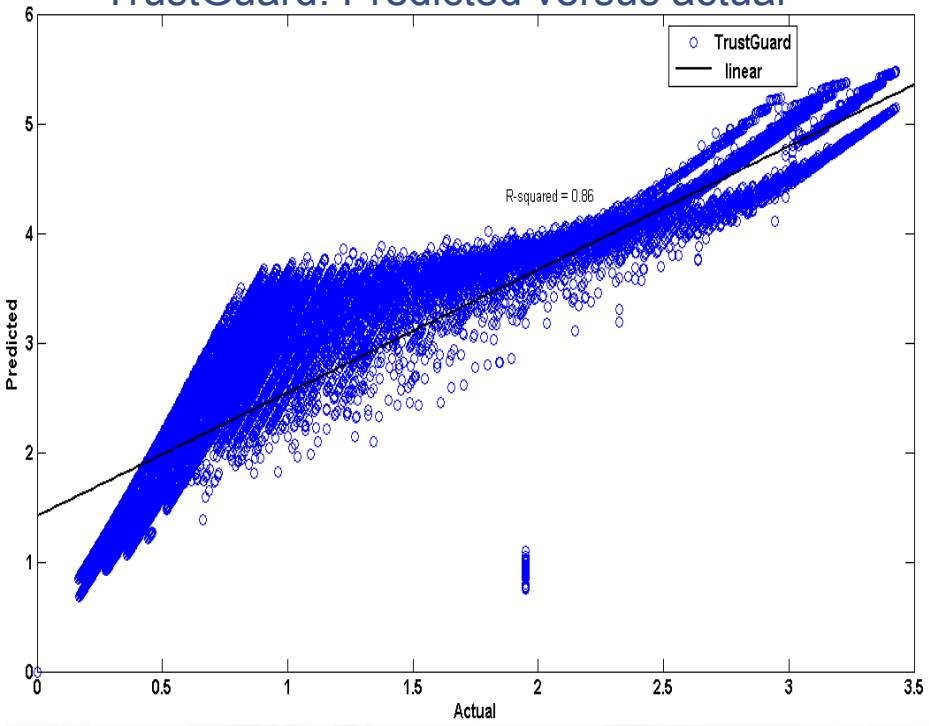
- ✓ Purpose: To test the reliability of D3-FRT with different scaling factors.
- ✓ Parameters: 50 with 12% and 10% of the peers exhibiting collusive and intoxicating behaviours.
- ✓ Observation: 5.1% estimation error rate.
- ✓ Lesson learnt: The best value for the scaling factors are $\mu_h = 0.3$; $\mu_o = 0.5$ as the best results were obtained with these values and in addition reducing the possibility of intoxication attacks.

Scaling factors: $\mu_h = 0.3$, $\mu_o = 0.5$, $\mu_f = 0.17$

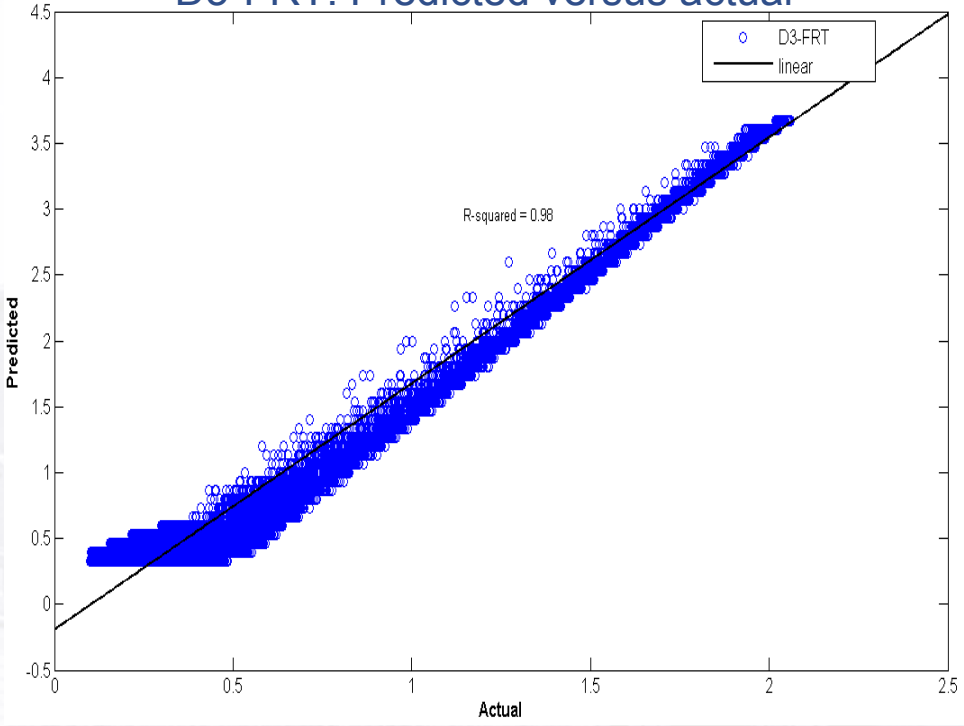


Versus TrustGuard: predictions

TrustGuard: Predicted versus actual



D3-FRT: Predicted versus actual



To test the predictive accuracy of both D3-FRT and the TrustGuard models.

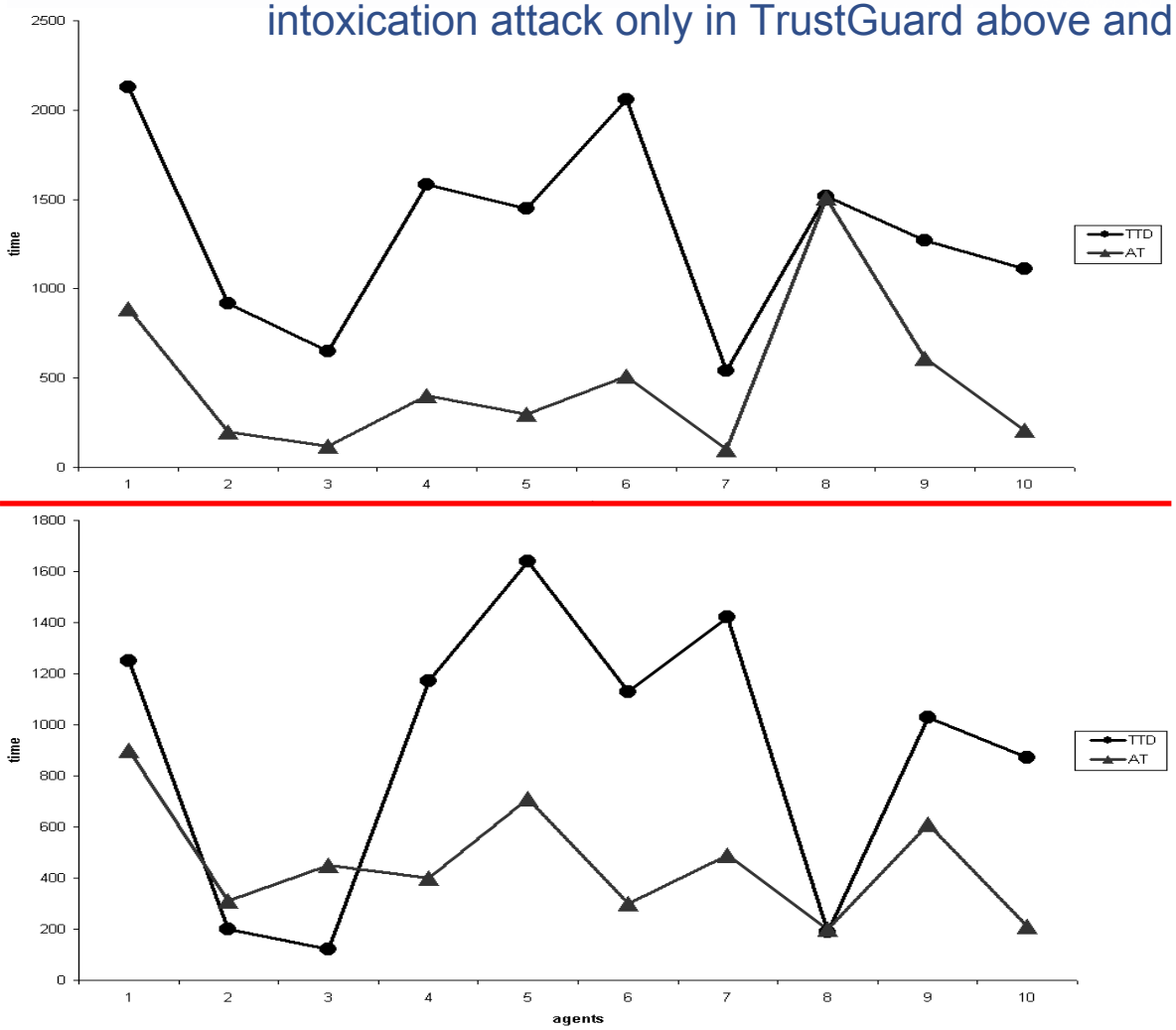
Similarities : a) flexibility by giving different trust components varying weights. b) sudden changes. c) degrees of centralisation. Differences: computing trust and predictions.

The observed high value of the coefficient of determination could be attributed to close relationship between actual and predicted ratings.

D3-FRT adapted and the error rate reduced with time. (500 peers, 10% and 15% intoxicating and collusive peers)

Versus TrustGuard: intoxication

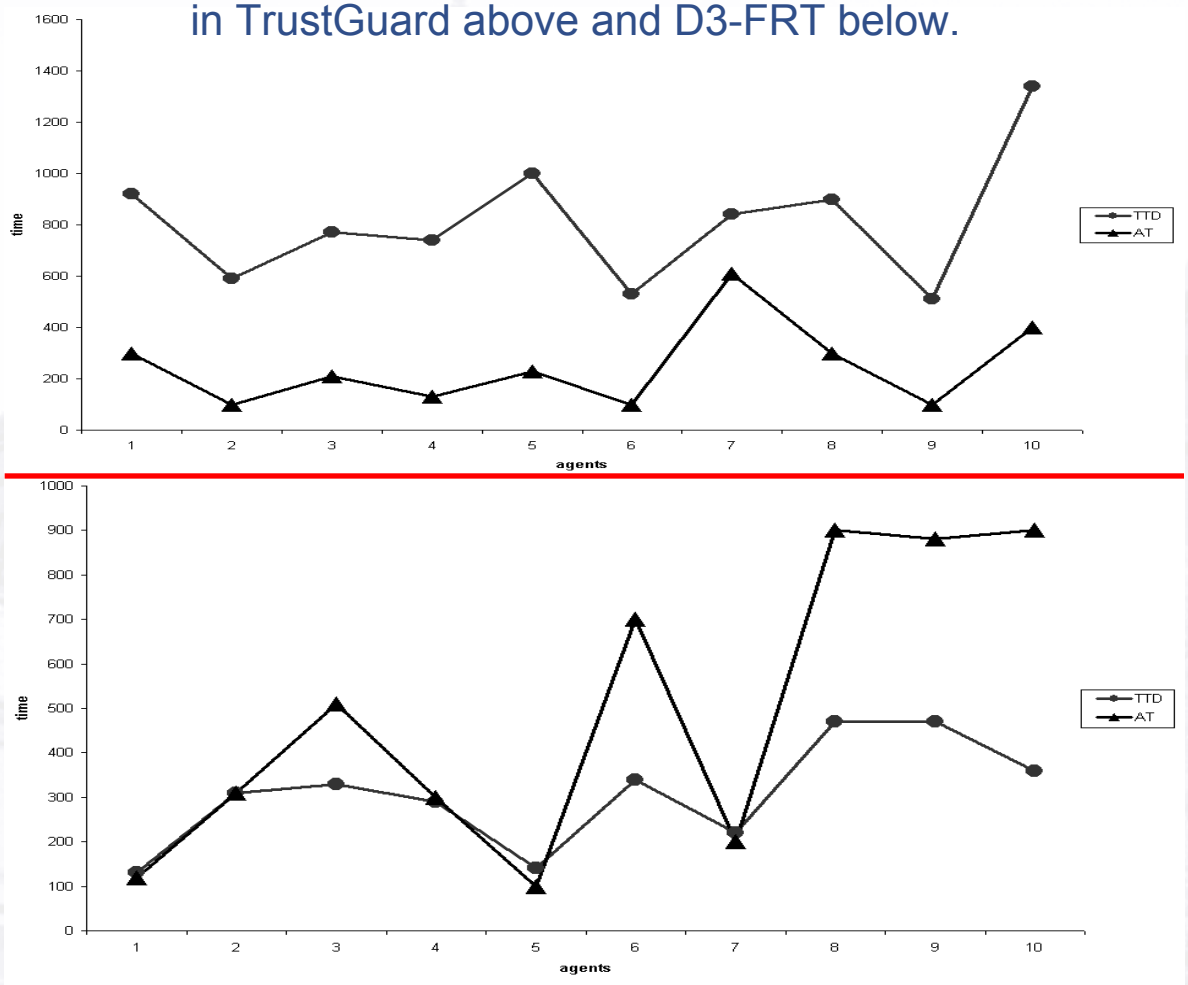
(Time-To-Detect) TTD versus the time of misbehaviour in the presence of intoxication attack only in TrustGuard above and D3-FRT below.



Parameters: 10 peers with 20% and 30% collusive and intoxicating respectively for a period of 2000 ticks.

Versus TrustGuard: collusion

TTD versus the time of misbehaviour in the presence of collusion attack only in TrustGuard above and D3-FRT below.



✓ Observations: Generally, D3-FRT consistently provide timely trust computations compared to TrustGuard but both models performed better in presence of collusion only compared to that of intoxication only.

D3-FRT suffered less from this attack by having lower TTD values overall as can be seen from the time variation in the vertical axis.

Versus TrustGuard: estimate error

✓ Purpose: From a series of experiments, we compare the models using a more symmetric measure with their respective average Mean Variation from Estimates (MVREs).

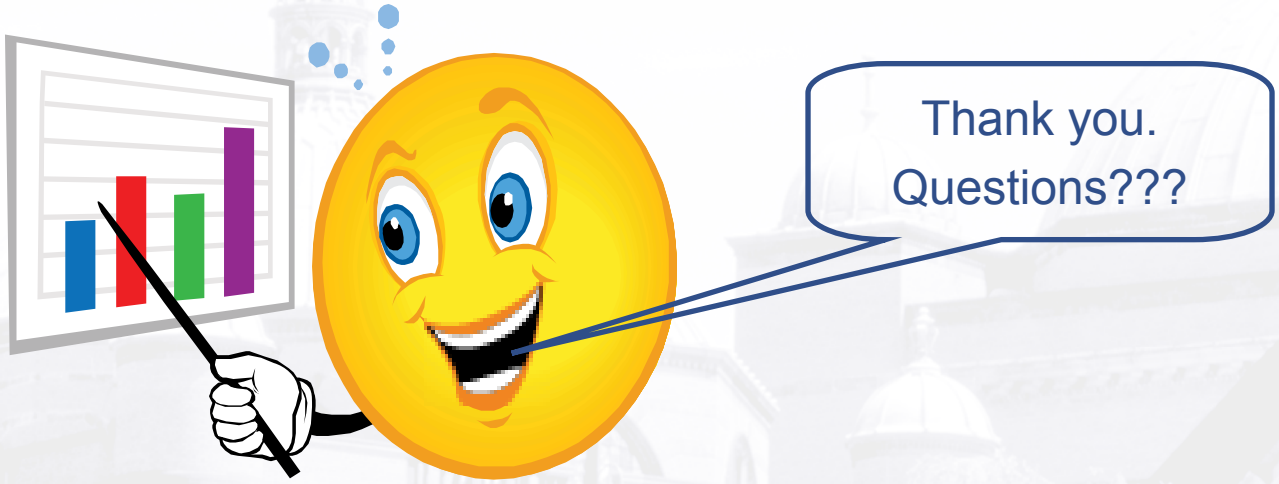
✓ Observations: The table shows that based on our implementation of the models, D3-FRT has a lower MVRE compared to that of TrustGuard.

The mean variation of estimate of D3-FRT is consistently less than that of TrustGuard.

Trial #	TrustGuard	Our approach
1	0.65	0.47
2	0.68	0.4
3	0.68	0.5
4	0.68	0.5
5	067	0.48
Average MVRE	0.67	0.49

Summary

- ✓ D3-FRT has the potential of providing a high level of dynamism for trust management allowing for more realistic analysis of the system and enabling predictions.
- ✓ Results suggest that hypothesis is most likely correct based on the scenarios considered.
- ✓ Preventive mechanisms for reducing the effect collusion and avoiding intoxication.
- ✓ Future direction
  - ❖ Model validation.
  - ❖ Emerging attacks.
  - ❖ Scalability.
  - ❖ Cost of simulation.



Funmi Onolaja
o.o.onolaja@cs.bham.ac.uk