Trusted Big Data Sharing

Researching alliances and infrastructure models across multiple autonomous organizations

SC’ 16 November 14-17th 2016
Salt Lake City, Utah.

Leon Gommans, Ameneh Deljoo, Ralph Koning, Ben de Graaff, Tristan Suerink, Gerben van Malenstein, Axel Berg, Erik Huizer, Rob Meijer, Tom van Engers, Cees de Laat
Data representing value in airline context

- Passenger flow handling
- Integrated Vehicle Health Management & Predictive Maintenance
- Cargo load optimization & scheduling
- Passenger experience
- Cybersecurity (NWO-COMMIT/SARNET project)
Big Data Sharing in commercial Enterprise environments
Sharing Big Data assets needs:

- Clearly defined and agreed common benefit
- Established common rules governing use, access AND benefit sharing.
- Organizing trust amongst group members as means to reduce risk
  Research Topic 1: Cybersecurity Alliance context
- Infrastructure supporting implementation of trust
  Research Topic 2: Trusted Big Data Sharing Infrastructure context
Organizing trust

Business Level
Governing common benefit and risk

IT Level
Enabling Trusted Access and Use

Rulemaking
Judicial body
Group Rules
Executive
Administration
Enforcement
Infrastructure
Using the Service Provider Group framework describing a way trust could be organized, the main question became how to quantify trust, ie:

What trustworthiness estimators expresses alliance member behavior influencing risk and benefits for the group of alliance members?
Agent Based Models that simulate alliance member behavior trusting the alliance as a whole, based on uncertainty of its environment:

1. How to model an alliance context using trustworthiness estimates?
2. How accurate, robust and reliable are particular sets of trustworthiness estimates?
3. Are trustworthiness estimates influenced when agents have limited knowledge about its environment?
4. ...

Contributing to general questions:

Are Agent Based models useful to model trust within organizational networks in open environments, in particular in the cyber security domain?

and

Is it possible to develop a social computational trust model?
Agent Model evaluating Trust

- Trustworthiness
- Observation
- Expectation
- Trust Assessment
- Risk
- Outcome
- Impact
- Trust Assessment & Risk taking
First step: Evolutionary Prisoners Dilemma using ABM Simulation

Agents choose from different strategies:
- Collaborate
- Defect
- During simulation: Agents predict next behavior of neighboring agents learned from observing past behavior.

Simulation observes tendency to maximize individual welfare instead of helping the group.

This type of simulation will be base to simulate more complex collaborations of autonomous organizations.

Research performed by Ameneh Deljoo, PhD candidate University of Amsterdam.
ABM Simulation

Evolutionary Prisoner's Dilemma

Game Type  Rules  Strategies

Main Display

Player Type
- Very 'Nice': 0.36 %
- Mostly Cooperative: 2.88 %
- Cooperative: 8.44 %
- Balanced ('nice'): 27.16 %
- Balanced ('nasty'): 34.88 %
- Tendency to Defect: 17.8 %
- Mostly Defects: 7.72 %
- Very 'Nasty': 0.76 %

Population Fitness Stats
- Minimum Payoff: 0.1925
- Maximum Payoff: 4.105
- Average Payoff: 2.24

Use an Evolutionary Algorithm
- Use a Genetic Algorithm

Start  Stop

View Fittest Individual  View Weakest Individual

Graph

1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 31 32 33 34 35 36 37 38 39 40 41 42 43 44 45 46 47 48 49 50

Generation
Alliances supported by a Digital Market Place - Research

• Digital Market Place (DMP) is a member organization as independent legal entity.
• Goal of the DMP is to organize trust between members wanting to gain a particular common benefit no single member can gain on its own.
• Members of the DMP can be a supplier or consumer of data or both.
• All members have equal rights within a DMP
• DMP is governed by a board of members in which all members participate
• DMP establishes a regulation consisting of market rules and the admission requirements
• DMP appoints a market master in charge of market operations
• DMP establishes a regulation for conflict settlement
• DMP appoints an adjudication committee
• Members can obtain rights (licenses) from the DMP within the framework of the DMP regulation to act in a particular defined market role.

• What elements of the DMP can be digitized?
1. Given an agreed benefit to share data within a group of autonomous organizations:

   How can trusted sharing of (big-) data assets be securely implemented in an infrastructure?

2. Given future, software definable Internet capabilities provides virtually unlimited amounts of dedicated and secure bandwidth:

   What infrastructure models are best suited to perform (big-) data analyses?
Traditional Hub Sharing Model

Domain D

Domain A

Domain B

Domain C

Analyses
Public Cloud Sharing Model

Domain A

Domain B

Domain C

Analyses
Encryption in public cloud.

Analyses

Domain A

Domain B

Domain C
Proces local, share results
Turntable model*

* Demo at SC 2005: Seamless Live Migration of Virtual Machines over the MAN/WAN: Franco Travostino, et. al.
Searching behavioral patterns

Domain A

Domain B

Domain C
Separating Data from Compute using High Performance Network links

E.g. a 100 Gb/s link is potentially **20/80x faster** when compared with a local SSD / HDD performance.
(Science-) Data Hub

Domain A

Domain B

Domain C
Meta Data Hub with Peering

Domain D

Domain A

Domain B

Domain C

Ana-lyses
Acces control based value sharing

Domain D

Domain A

Domain B

Domain C

Aanalyses

blockchain
Using an Open Lightpath Exchange infrastructure.
Digital Market Model

Domain A  
Domain B  
Domain C

In memory analyses  
e-Market  
Blockchain
Imagine a globally con-federated digital market system testbed..
Participation in testbeds

ExoGENI Testbed

Link using ESnet DTN technology

PRP Participants include:
- Univ. of Hawaii System
- Montana State Univ.
- Northwestern Univ.
- NCAR
- MREN
- StarLight
- UIUC
- Chameleon
- UVA
- AARNet
- KISTI/KREONet
- Univ. of Tokyo
- NCSA
- Clemson Univ.
Thank you