AFRIAT'S TEST FOR DETECTING MALICIOUS AGENT [1]

B.Tondi

Dept. of Information Engineering and Mathematical Sciences, *University of Siena*

1 Vikram Krishnamurthy, IEEE signal processing letters, vol.19, No.12, December 2012

Contents

- **The problem addressed**:

  Detecting the *presence* of malicious agents in networks

- **Formalization of the problem**:

  How malicious agents *behave*?

- **Possible way to solve the problem**:

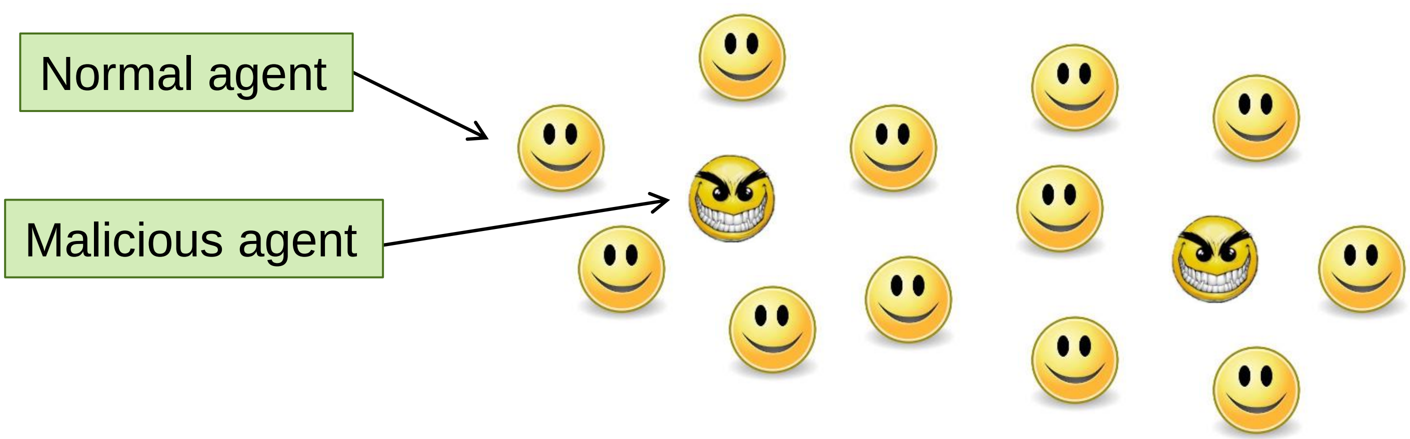
  Afriat's theorem (economic literature)

An overview of the problem

- **Why we are interested in detecing if an agent is malicious?**

  Only if a malicious agent is detected the system moves to an alert state and some countermisures are adoped (*resource saving*).

  - The system should be able to distinguish **malicious agents** from **normal agents** so to be able to reveal the *presence* of attackers

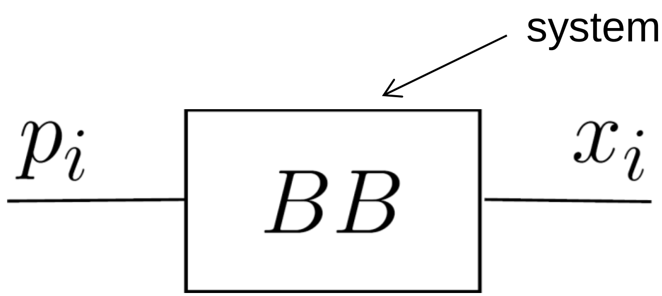


  - This is not possible by means of TdG.

Schematization of the problem

- At each time *i:*

$p_i \rightarrow$ probe vector
$x_i \rightarrow$ responce



$p_i, x_i \in \mathbb{R}_+^k$

- There exists an *utility function* that the black box is *maximizing* to generate its response x_i to probe input p_i ? (decision test)

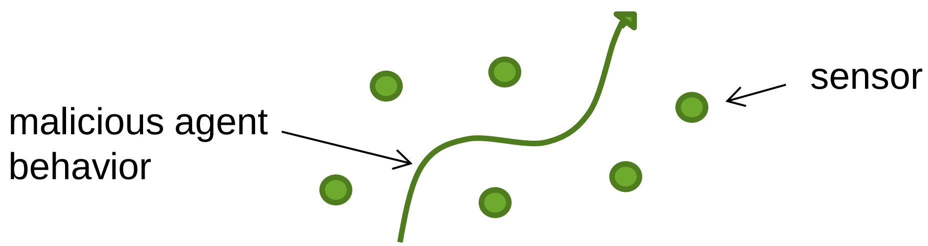
- In many practical scenarios:

Malicious agent are utility maximizer

 $\longleftrightarrow$ $\max u(x)$

Malicious agents: examples (1/2)

- **Sensor networks** *(detecting intruders in a sensor field)*
  - *System goal:* to detect if an agent is avoiding being detected by the sensors
  - *Malicious agents behavior:* seek to evade detection by maximizing its associated distance to each sensor (based on the relative importance of the sensors)

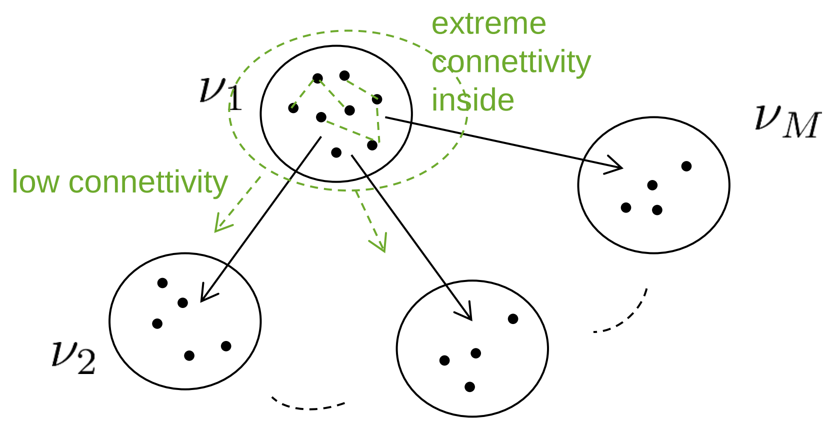


  - *Probe and responce model:* (p_i, x_i)

$$p_i = \text{importance parameter vector}$$
$$x_i = \text{distance between the agent and each sensor}$$

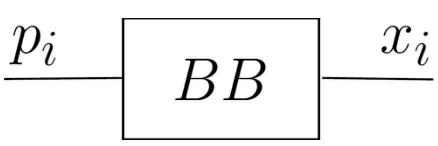
Malicious agents: examples (2/2)

- **Social networks** *(detecting tightly connected subgraphs)*
  - *Malicious agents behavior (e.g. hijackers):* maximize the connettivity to other nodes in their subgraph (*social group*) and minimize the connettivity to nodes outside.



ν_i = social group

  - *Probe and responce model:* (p_i, x_i)



p_i = QoS of the links between a node in ν_1 and a node in ν_k, $k = 1, .., M$ and viceversa

x_i = average amount of communication resources consumed by the nodes in ν_i

Then.....

- There are many *real scenarios* in which malicious agents behave as utility maximizer
- Given a system (BB)

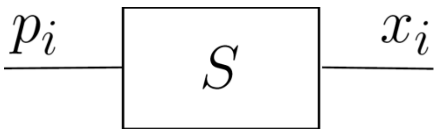
"Detecting the presence of malicious agents corresponds to determine if there exists an utility fuction that BB is maximizing"

- Challenging goal:

(p_i, x_i) *at each time* $i \longrightarrow$ *BB is an "utility maximizer" ?*

Afriat's test

Some terminology



$p_i, x_i \in \mathbb{R}^k_+$

- A system S is an ***utility maximizer*** if for every probe p_i, the chosen response x_i satisfies:

$$x_i = x^*(p_i) \in \arg\max_{p_i' x \leq 1} u(x)$$

  where $u(x)$ is a *nonsatiated* utility function.

  - Nonsatiated formally means that:

$$\forall \eta > 0, \exists x \quad \text{with} \quad ||x - x_i||_2 < \eta \quad s.t. \quad u(x) > u(x_i)$$

- We say that $u(\cdot)$ ***rationalizes*** the observed responses if and only if

$$u(x_i) = \max\{u(x) : p_i' x \leq 1\} \quad \forall i$$

Afriat's test (the original problem)

- Afriat's test [2] (1967) is a remarkable result in *Consumer Theory* concearned with 'how a rational consumer would make consumption decisions' (a widely studed topic in economic literature).
- ***Consumer problem (CP)***

$p \rightarrow$ price vector
$x \rightarrow$ purchased quantity vector
$w \rightarrow$ total consumer's wealth

$$\max_{x \in \mathbb{R}^k_+} u(x)$$

$$s.t. \quad p \cdot x \leq w$$

budget constraint

- Afriat answers the question of "*when a sequence of purchase decisions* (p_i, x_i) *is consistent with the purchaser maximizing a concave utility function* $u(\cdot)$" .

2 S. N. Afriat, 'The construction of utility functions from expenditure data', *International Economic Review*, vol. 8, no 1, February, 1967

Afriat's theorem

- Given a dataset $D = \{(p_i, x_i) : i \in N = \{1, 2, ..., n\}\}$ with $p_i, x_i \in \mathbb{R}^k_+$, the following statements are equivalent:
  i. There exists a *non-satiated utility function that rationalizes the data*;
  ii. The data satisfies *GARP (Generalized Axiom of Revealed Preference)*, namely

  $$p_j \cdot x_{j+1} \leq p_j \cdot x_j, \quad \forall j \leq n-1 \quad \Rightarrow p_n \cdot x_1 \geq p_n \cdot x_n$$

  iii. There exist numbers $U_1, ..., U_n$ and $\lambda_1, ..., \lambda_n$ satisfying the *Afriat's inequalities*

  $$U_j - U_i - \lambda_i p_i (x_j - x_i) \leq 0, \quad \text{for all } i, j \in N$$

  iv. There exists a *non-satiated, concave, monotonic, continuous utility function that rationalizes the data*.

Remarkable consequence

Afriat's theorem gives ***necessary* and *sufficient* conditions** for a system to be a utility maximizer based **only on the *input-output* *responce***

- The remarkable feature of Afriat's Theorem is that the utility function $u(\cdot)$ *does not need to be known*.

- Afriat's test is viewed as a *blind test:* it detect utility maximizing behavior without knowledge of the utility function.

- This result is particularly useful in **detecting malicious agents** since the precise nature of the utility function that is being maximized is not known to the system (BB).

Testing utility maximization

- The price vectors p_i and the observed quantity vectors x_i can be checked for consistency with maximization of a non-satiated utility function $u(\cdot)$ in several ways ((ii.) or (iii.)) :
  1. checking whether or not the data satisfy GARP;
  2. using *linear programming methods* to check for the existence of a solution to Afriat's inequality, e.g.[3]

$$\min S_T \qquad\qquad S_T = \text{largest violation of the Afriat inequalities}$$

subject to

$$\begin{aligned}
&U_j - U_i - \lambda_i p_i (x_j - x_i) \leq S_T \quad \text{for all } i, j \in N \\
&\lambda_j > 0, \quad \text{for all } j \in N \\
&S_T \geq 0
\end{aligned}$$

3 simplified formulation of Fleissing and Whitney (2005): 'Testing for thee significance of violations of Afriat's inequalities', *Journal of Business and Economic Statistics*, 23,p 355-362

Afriat's Test in practical settings

- The responces x_i are measured via *noisy observations* y_i :

$$y_i = x_i + w_i, \quad x_i \in \mathbb{R}^k_+, \quad w_i \text{ i.i.d. noise vector}$$

Hp) additive
noise model

- Given a dataset $D_{noisy} = \{(p_i, y_i) : i \in N\}$ the question is: '*how can Afriat's Theorem be generalized to detect a utility maximizer?'*
- Jones and Edgerton[4] give a *decision test* to detect a utility maximizer using the noisy dataset D_{noisy} (statistical N-P test):
  - The test has a guaranteed upper bound on *Type-I errors* in detecting malicious agents

4 Barry E. Jones and David L. Edgerton 'Testing utility maximization with measurement errors in the data' *Advances iin Econometrics*, 2009, Vol.24, p 199-236

Statistical test for 'malicious' behavior (1/2)

- The noisy dataset: $D_{noisy} = \{(p_i, y_i) : i \in N\}$
- Based on Afriat's Theorem, we want to solve the *hypothesis test:*

  H0: the clean dataset D *satisfies utility maximization;*

  H1: the clean dataset D *does not satisfy utility maximization;*

  Errors: Type I $\longrightarrow$ *accept H1 when H0 holds (Type II* $\longrightarrow$ *accept H0 when H1 holds);*

- Jones And Edgerton (2009) consider the statistical test

significance level

$$\int_{\Phi^*(y)}^{\infty} f_M(\beta) d\beta \underset{H1}{\overset{H0}{\gtrless}} \alpha, \qquad (1)$$

test statistic $y \in N$

where: $M \equiv \max_{i,j}[p_i(w_i - w_j)]$ and $\Phi^*(y)$ is the solution of the *constrained optimization problem:*

$$\begin{aligned} \min \quad & \Phi \\ s.t. \quad & U_j - U_i - \lambda_i p_i(x_j - x_i) - \lambda_i \Phi \leq 0 \\ & \lambda_i > 0, U_i > 0, \Phi > 0 \quad \text{for } i, j \in \{1, 2, ..., n\} \end{aligned}$$

Statistical test for 'malicious' behavior (2/2)

- Jones And Edgerton (2009) prove the following theorem:

Theorem *(Statistical test for agent that seek to maximize utility)*

'Given the noisy dataset D_{noisy} , the probability that the statistical test (1) yields a Type-I error (reject H0 when true) is less than α '.

- The theorem guarantees that *the Type-I error probability is less than* α for the decision test (1). Through the optimization of the probe signal p_i it is possible *to reduce (minimize) the Type-II error probability*.