A Game Theoretic Analysis of Intrusion Detection in Access Control Systems<sup>1</sup>-**Tansu Alpcan & Tamer Başar** 

> FEP3301 Computational Game Theory Paper Presentation

#### Kim Hammar

kimham@kth.se

Division of Network and Systems Engineering KTH Royal Institute of Technology

Oct 14, 2021

<sup>&</sup>lt;sup>1</sup>T. Alpcan and T. Basar. "A game theoretic analysis of intrusion detection in access control systems". In: 2004 43rd IEEE Conference on Decision and Control (CDC) (IEEE Cat. No.04CH37601). Vol. 2. 2004, 1568–1573 Vol.2. DOI: 10.1109/CDC.2004.1430267.

#### Outline

#### Use case: Intrusion Detection

#### System model

- A finite extensive-form model of the use case
   Elementary equilibrium analysis
- A static continuous-kernel game model of the use case
   Equilibrium analysis using Rosen's Theorem

## Use Case: Intrusion Detection

#### Defender = IDS:

 System Operator with task of detecting intrusions

#### ► IT Infrastructure:

- Consist of a set of components (also called "subsystems").
- The infrastructure is equipped with a network of sensors
- Sensors report anomalies/alerts to the defender
- Each component has a set of vulnerabilities

#### Attacker:

Can attack vulnerable components



## Use Case: Intrusion Detection

#### Defender = IDS:

 System Operator with task of detecting intrusions

#### IT Infrastructure:

- Consist of a set of components (also called "subsystems").
- The infrastructure is equipped with a network of sensors
- Sensors report anomalies/alerts to the defender
- Each component has a set of vulnerabilities

#### Attacker:

Can attack vulnerable components



## Use Case: Intrusion Detection

#### Defender = IDS:

 System Operator with task of detecting intrusions

#### IT Infrastructure:

- Consist of a set of components (also called "subsystems").
- The infrastructure is equipped with a network of sensors
- Sensors report anomalies/alerts to the defender
- Each component has a set of vulnerabilities

#### Attacker:

Can attack vulnerable components



## System Model

## ► Infrastructure components (Subsystems) ► T = {t<sub>1</sub>, t<sub>2</sub>,...t<sub>max</sub>}

- ▶ Vulnerabilities
   ▶ *I* = {*I*<sub>1</sub>, *I*<sub>2</sub>, ..., *I*<sub>max</sub>}
- Attacks
  - $\blacktriangleright \mathcal{A} = \mathcal{T} \times \mathcal{I}$

#### Sensors

- $\triangleright \ \mathcal{S} = \{s_1, s_2, \dots, s_{max}\}$
- Each sensor reports "alarm" or "no alarm".
- ► Each sensor s<sub>i</sub> ∈ S corresponds to a vulnerabilitiy I<sub>i</sub> ∈ I



## Example Infrastructure: One Subsystem, One Sensor and One Vulnerability

- Infrastructure components (Subsystems)
  - $\blacktriangleright \mathcal{T} = \{t_1\}$
- Vulnerabilities
   *I* = {*I*<sub>1</sub>}
- Attacks
   A = {a = (t₁, l₁)}
- Sensors
  - $\blacktriangleright \ \mathcal{S} = \{s_1\}$





• **Players**:  $N = \{$ Attacker, Defender, Nature $\} = \{1, 2, 3\}$ 

#### Action sets:

- ▶  $A_1 = \{ \text{Attack}, \text{Continue} \}, A_2 = \{ \text{Alert} \rightarrow \text{Defend}, \text{Alert} \rightarrow \text{Continue}, \text{No Alert} \rightarrow \text{Defender}, \text{No Alert} \rightarrow \text{Continue}, \}, A_3 = \{ \text{Alert}, \text{No Alert} \}$
- Nature's pre-defined strategy:
  - $f(\text{Alert}|\text{Attack}) = \frac{2}{3}$ ,  $f(\text{Alert}|\text{Continue}) = \frac{1}{3}$









## Extensive & Strategic Form of the Finite Game Model Extensive Form:



Strategic Form:

$$\begin{array}{ccccc} & DC & DD & CD & CC \\ A & \left[ \begin{array}{cccc} 0, -\frac{5}{3} & -5, 5 & 5, -\frac{25}{3} & 10, -15 \\ \frac{5}{3}, -\frac{10}{3} & \frac{5}{3}, -\frac{20}{3} & 0, -\frac{10}{3} & 0, 0 \end{array} \right]$$
(1)

## Extensive & Strategic Form of the Finite Game Model Extensive Form:



Strategic Form:

$$\begin{bmatrix} DC & DD & CD & CC \\ A & \begin{bmatrix} 0, -\frac{5}{3} & -5, \frac{5}{2} & 5, -\frac{25}{3} & \underline{10}, -15 \\ \frac{5}{3}, -\frac{10}{3} & \frac{5}{3}, -\frac{20}{3} & 0, -\frac{10}{3} & 0, \underline{0} \end{bmatrix}$$
(2

 $\Rightarrow$  No pure Nash equilibrium

# Unique Mixed Nash Equilibrium Computed using Lemke-Howson's Algorithm



Mixed Nash equilibrium  $(s_1^*, s_2^*) \in \Delta(A_1) \times \Delta(A_2)$ :

• 
$$s_1^*(\text{Attack}) = \frac{1}{5}, \ s_1^*(\text{Continue}) = \frac{4}{5}$$

## Limitations of the Finite Game Model



#### Limitations:

- Hard to analyze for large systems (scalability)
- Hard to define all of the parameters in the model
- Equilibria depend greatly on the payoff-values, are they realistic?
- Does mixed equilibria make sense?

## Limitations of the Finite Game Model



#### Alternative model: continuous-kernel game

**Limitations**:

- Hard to analyze for large systems (scalability)
- Hard to define all of the parameters in the model

• Attacker Action Space:  $A_1 = \mathbb{R}^{A_{max}}_+$  instead of  $A_1 = \{a_1, \dots, A_{max}\}$ 

▶ Defender Action Space: A<sub>2</sub> = ℝ<sup>D<sub>max</sub><sub>+</sub> instead of A<sub>2</sub> = {a<sub>1</sub>,..., D<sub>max</sub>}</sup>

Nature/(Sensor Network) Action Space:

 $\blacktriangleright A_3 = \mathbb{R}^{A_{max} \times A_{max}}_+ \text{ instead of } A_3 = \{\text{alert}, \text{no alert}\} \times A_2 = \bar{\boldsymbol{P}}$ 

*P*<sub>ij</sub> represent the alert-weight that nature put on attack j when attack i occurred.

• Attack detection metric:  $dq(i) = \frac{P_{ii}}{\sum_{i=1}^{A_{max}} \vec{P}_{ii}}$ 

For notational convenience, define:

$$oldsymbol{P} = egin{cases} p_{i,j} = -ar{p}_{i,j} & ext{if } i=j \ p_{i,j} = ar{p}_{i,j} & ext{otherwise} \end{cases}$$

- ► Attacker Action Space: A<sub>1</sub> = ℝ<sup>A<sub>max</sub><sub>+</sub> instead of A<sub>1</sub> = {a<sub>1</sub>,..., A<sub>max</sub>}</sup>
- ▶ Defender Action Space: A<sub>2</sub> = ℝ<sup>D<sub>max</sub><sub>+</sub> instead of A<sub>2</sub> = {a<sub>1</sub>,..., D<sub>max</sub>}</sup>

Nature/(Sensor Network) Action Space:

- $\blacktriangleright A_3 = \mathbb{R}^{A_{max} \times A_{max}}_+ \text{ instead of } A_3 = \{\text{alert}, \text{no alert}\} \times A_2 = \bar{\boldsymbol{P}}$
- *P*<sub>ij</sub> represent the alert-weight that nature put on attack j when attack i occurred.

 $\blacktriangleright$   $\bar{P} = I$  if the sensors are perfect.

• Attack detection metric:  $dq(i) = \frac{\bar{P}_{ii}}{\sum_{i=1}^{A_{max}} \bar{P}_{ii}}$ 

For notational convenience, define:

$$oldsymbol{P} = egin{cases} p_{i,j} = -ar{p}_{i,j} & ext{if } i=j \ p_{i,j} = ar{p}_{i,j} & ext{otherwise} \end{cases}$$

- Attacker Action Space:  $A_1 = \mathbb{R}^{A_{max}}_+$  instead of  $A_1 = \{a_1, \dots, A_{max}\}$ , pure strategy  $a^A \in A_1$
- ▶ Defender Action Space:  $A_2 = \mathbb{R}^{D_{max}}_+$  instead of  $A_2 = \{a_1, \dots, D_{max}\}$ , pure strategy  $a^D \in A_2$
- Nature/(Sensor Network) Action Space:
  - ►  $A_3 = \mathbb{R}^{A_{max} \times A_{max}}_+$  instead of  $A_3 = \{ \text{alert}, \text{no alert} \} \times A_2 = \bar{P}$
  - *P*<sub>ij</sub> represent the alert-weight that nature put on attack j when attack i occurred.
  - $\bar{P} = I$  if the sensors are perfect.
  - Attack detection metric:  $dq(i) = \frac{P_{ii}}{\sum_{i}^{A_{max}} \bar{P}_{ij}}$
  - For notational convenience, define:

$$\boldsymbol{P} = \begin{cases} p_{i,j} = -\bar{p}_{i,j} & \text{if } i = j \\ p_{i,j} = \bar{p}_{i,j} & \text{otherwise} \end{cases}$$
(5)

- **•** Defender cost of being attacked:  $c^D \in \mathbb{R}^{A_{max}}$
- Attacker gain of successful attack:  $c^A \in \mathbb{R}^{A_{max}}$
- ▶ Vulnerability matrix:  $Q \in \mathbb{R}^{A_{max} \times A_{max}}$  diagonal matrix that models degree of vulnerability per attack
- ▶ Defense matrix:  $\bar{Q} \in \{0,1\}^{A_{max} \times D_{max}}$  where  $\bar{Q}_{i,j} = 1$  if defense detects the attack and 0 otherwise.
- **•** Cost of defender actions:  $\alpha \in \mathbb{R}^{D_{max}}_+$
- **Cost of attacker actions:**  $\beta \in \mathbb{R}^{A_{max}}_+$
- False alarm weight: γ, defines how much to penalize false alarms

- **>** Defender cost of being attacked:  $\boldsymbol{c}^{D} \in \mathbb{R}^{A_{max}}$
- Attacker gain of successful attack:  $c^A \in \mathbb{R}^{A_{max}}$
- ▶ Vulnerability matrix:  $Q \in \mathbb{R}^{A_{max} \times A_{max}}$  diagonal matrix that models degree of vulnerability per attack
- ▶ Defense matrix:  $\bar{Q} \in \{0, 1\}^{A_{max} \times D_{max}}$  where  $\bar{Q}_{i,j} = 1$  if defense detects the attack and 0 otherwise.
- **•** Cost of defender actions:  $\alpha \in \mathbb{R}^{D_{max}}_+$
- **Cost of attacker actions:**  $\beta \in \mathbb{R}^{A_{max}}_+$
- False alarm weight: γ, defines how much to penalize false alarms

- **>** Defender cost of being attacked:  $\boldsymbol{c}^{D} \in \mathbb{R}^{A_{max}}$
- Attacker gain of successful attack:  $c^A \in \mathbb{R}^{A_{max}}$
- ► Vulnerability matrix: Q ∈ ℝ<sup>A<sub>max</sub> × A<sub>max</sub> diagonal matrix that models degree of vulnerability per attack</sup>
- ▶ Defense matrix:  $\bar{Q} \in \{0, 1\}^{A_{max} \times D_{max}}$  where  $\bar{Q}_{i,j} = 1$  if defense detects the attack and 0 otherwise.
- **•** Cost of defender actions:  $\alpha \in \mathbb{R}^{D_{max}}_+$
- Cost of attacker actions:  $\beta \in \mathbb{R}^{A_{max}}_+$

False alarm weight: γ, defines how much to penalize false alarms

- **>** Defender cost of being attacked:  $\boldsymbol{c}^{D} \in \mathbb{R}^{A_{max}}$
- Attacker gain of successful attack:  $c^A \in \mathbb{R}^{A_{max}}$
- ► Vulnerability matrix: Q ∈ ℝ<sup>A<sub>max</sub> × A<sub>max</sub> diagonal matrix that models degree of vulnerability per attack</sup>
- ▶ Defense matrix:  $\bar{Q} \in \{0, 1\}^{A_{max} \times D_{max}}$  where  $\bar{Q}_{i,j} = 1$  if defense detects the attack and 0 otherwise.
- Cost of defender actions:  $\alpha \in \mathbb{R}^{D_{max}}_+$
- Cost of attacker actions:  $\beta \in \mathbb{R}^{A_{max}}_+$
- False alarm weight: γ, defines how much to penalize false alarms

Objective: Minimize costs rather than maximize utilities

**Defender Cost Function:** 

$$J^{D}(\boldsymbol{a}^{A}, \boldsymbol{a}^{D}, \boldsymbol{P}) =$$

$$\underbrace{\gamma(\boldsymbol{a}^{A})^{T} \boldsymbol{P} \bar{\boldsymbol{Q}} \boldsymbol{a}^{D}}_{\text{false alarm}} + \underbrace{(\boldsymbol{a}^{D})^{T} diag(\alpha) \boldsymbol{a}^{D}}_{\text{cost of defense}} + \underbrace{\boldsymbol{c}^{D}(\boldsymbol{Q} \boldsymbol{a}^{A} - \bar{\boldsymbol{Q}} \boldsymbol{a}^{D})}_{\text{cost of attack}}$$
(6)

Attacker Cost Function:

$$J^{A}(\boldsymbol{a}^{A}, \boldsymbol{a}^{D}, \boldsymbol{P}) =$$

$$-\underbrace{\gamma(\boldsymbol{a}^{A})^{T} \boldsymbol{P} \boldsymbol{Q} \boldsymbol{a}^{D}}_{\text{detected}} + \underbrace{(\boldsymbol{a}^{A})^{T} \text{diag}(\beta) \boldsymbol{a}^{A}}_{\text{cost of attack}} + \underbrace{\boldsymbol{c}^{A}(\boldsymbol{\bar{Q}} \boldsymbol{a}^{D} - \boldsymbol{Q} \boldsymbol{a}^{A})}_{\text{gain of attack}}$$
(7)

Summary: Continuous-kernel general-sum game with strictly convex cost functions.

- Objective: Minimize costs rather than maximize utilities
- Defender Cost Function:

$$J^{D}(\boldsymbol{a}^{A}, \boldsymbol{a}^{D}, \boldsymbol{P}) =$$

$$\underbrace{\gamma(\boldsymbol{a}^{A})^{T} \boldsymbol{P} \bar{\boldsymbol{Q}} \boldsymbol{a}^{D}}_{\text{false alarm}} + \underbrace{(\boldsymbol{a}^{D})^{T} diag(\alpha) \boldsymbol{a}^{D}}_{\text{cost of defense}} + \underbrace{\boldsymbol{c}^{D}(\boldsymbol{Q} \boldsymbol{a}^{A} - \bar{\boldsymbol{Q}} \boldsymbol{a}^{D})}_{\text{cost of attack}}$$
(8)

Attacker Cost Function:

$$J^{A}(\boldsymbol{a}^{A}, \boldsymbol{a}^{D}, \boldsymbol{P}) =$$

$$-\underbrace{\gamma(\boldsymbol{a}^{A})^{T} \boldsymbol{P} \boldsymbol{Q} \boldsymbol{a}^{D}}_{\text{detected}} + \underbrace{(\boldsymbol{a}^{A})^{T} \text{diag}(\beta) \boldsymbol{a}^{A}}_{\text{cost of attack}} + \underbrace{\boldsymbol{c}^{A}(\boldsymbol{\bar{Q}} \boldsymbol{a}^{D} - \boldsymbol{Q} \boldsymbol{a}^{A})}_{\text{gain of attack}}$$
(9)

Summary: Continuous-kernel general-sum game with strictly convex cost functions.

- Objective: Minimize costs rather than maximize utilities
- Defender Cost Function:

$$J^{D}(\boldsymbol{a}^{A}, \boldsymbol{a}^{D}, \boldsymbol{P}) =$$

$$\underbrace{\gamma(\boldsymbol{a}^{A})^{T} \boldsymbol{P} \boldsymbol{\bar{Q}} \boldsymbol{a}^{D}}_{\text{false alarm}} + \underbrace{(\boldsymbol{a}^{D})^{T} diag(\alpha) \boldsymbol{a}^{D}}_{\text{cost of defense}} + \underbrace{\boldsymbol{c}^{D}(\boldsymbol{Q} \boldsymbol{a}^{A} - \boldsymbol{\bar{Q}} \boldsymbol{a}^{D})}_{\text{cost of attack}}$$
(10)

Attacker Cost Function:

$$J^{A}(\boldsymbol{a}^{A}, \boldsymbol{a}^{D}, \boldsymbol{P}) =$$

$$-\underbrace{\gamma(\boldsymbol{a}^{A})^{T} \boldsymbol{P} \bar{\boldsymbol{Q}} \boldsymbol{a}^{D}}_{\text{detected}} + \underbrace{(\boldsymbol{a}^{A})^{T} \text{diag}(\beta) \boldsymbol{a}^{A}}_{\text{cost of attack}} + \underbrace{\boldsymbol{c}^{A}(\bar{\boldsymbol{Q}} \boldsymbol{a}^{D} - \boldsymbol{Q} \boldsymbol{a}^{A})}_{\text{gain of attack}}$$
(11)

Summary: Continuous-kernel general-sum game with strictly convex cost functions.

Objective: Minimize costs rather than maximize utilities

**Defender Cost Function:** 

$$J^{D}(\boldsymbol{a}^{A}, \boldsymbol{a}^{D}, \boldsymbol{P}) =$$

$$\underbrace{\gamma(\boldsymbol{a}^{A})^{T} \boldsymbol{P} \bar{\boldsymbol{Q}} \boldsymbol{a}^{D}}_{\text{false alarm}} + \underbrace{(\boldsymbol{a}^{D})^{T} diag(\alpha) \boldsymbol{a}^{D}}_{\text{cost of defense}} + \underbrace{\boldsymbol{c}^{D}(\boldsymbol{Q} \boldsymbol{a}^{A} - \bar{\boldsymbol{Q}} \boldsymbol{a}^{D})}_{\text{cost of attack}}$$
(12)

Attacker Cost Function:

$$J^{A}(\boldsymbol{a}^{A}, \boldsymbol{a}^{D}, \boldsymbol{P}) =$$

$$-\underbrace{\gamma(\boldsymbol{a}^{A})^{T} \boldsymbol{P} \bar{\boldsymbol{Q}} \boldsymbol{a}^{D}}_{\text{detected}} + \underbrace{(\boldsymbol{a}^{A})^{T} \text{diag}(\beta) \boldsymbol{a}^{A}}_{\text{cost of attack}} + \underbrace{\boldsymbol{c}^{A}(\bar{\boldsymbol{Q}} \boldsymbol{a}^{D} - \boldsymbol{Q} \boldsymbol{a}^{A})}_{\text{gain of attack}}$$
(13)

Summary: Continuous-kernel general-sum game with convex cost functions.

## Equilibrium Analysis (1/3)

Since J<sup>A</sup>, J<sup>D</sup> are strictly convex, the best-response correspondences are obtained from the first order conditions:

$$\nabla_{\mathbf{a}^{D}}(J^{D}(\mathbf{a}^{A}, \mathbf{a}^{D}, \mathbf{P})) = 0 \qquad (14)$$

$$\nabla_{\mathbf{a}^{D}}(\gamma(\mathbf{a}^{A})^{T} \mathbf{P} \bar{\mathbf{Q}} \mathbf{a}^{D} + (\mathbf{a}^{D})^{T} diag(\alpha) \mathbf{a}^{D} + \mathbf{c}^{D}(\mathbf{Q} \mathbf{a}^{A} - \bar{\mathbf{Q}} \mathbf{a}^{D})) = 0$$

$$\gamma(\mathbf{a}^{A})^{T} \mathbf{P} \bar{\mathbf{Q}} + (\mathbf{a}^{D})^{T} (2diag(\alpha)) - \mathbf{c}^{D} \bar{\mathbf{Q}} = 0$$

$$\implies Br_{D}(\mathbf{a}^{A}, \mathbf{P}) = \{\mathbf{c}^{D} \bar{\mathbf{Q}}(2diag(\alpha))^{-1} - \gamma(2diag(\alpha))^{-1} \bar{\mathbf{Q}}^{T} \mathbf{P}^{T} \mathbf{a}^{A}\}$$

and, analogously for the attacker:

$$\nabla_{\boldsymbol{a}^{A}}(J^{D}(\boldsymbol{a}^{A},\boldsymbol{a}^{D},\boldsymbol{P})) = 0$$
(15)  
$$-\gamma \boldsymbol{P} \boldsymbol{\bar{Q}} \boldsymbol{a}^{D} + (\boldsymbol{a}^{A})^{T} (2 \operatorname{diag}(\beta)) - \boldsymbol{c}^{A} \boldsymbol{Q} = 0$$
  
$$\boldsymbol{a}^{A} = (2 \operatorname{diag}(\beta))^{-1} \boldsymbol{c}^{A} \boldsymbol{Q} + \gamma (2 \operatorname{diag}(\beta))^{-1} \boldsymbol{P} \boldsymbol{\bar{Q}} \boldsymbol{a}^{D}$$
$$\implies Br_{A}(\boldsymbol{a}^{D},\boldsymbol{P}) = \{(2 \operatorname{diag}(\beta))^{-1} \boldsymbol{c}^{A} \boldsymbol{Q} + \gamma (2 \operatorname{diag}(\beta))^{-1} \boldsymbol{P} \boldsymbol{\bar{Q}} \boldsymbol{a}^{D}\}$$

## Equilibrium Analysis (1/3)

Since J<sup>A</sup>, J<sup>D</sup> are strictly convex, the best-response correspondences are obtained from the first order conditions:

$$\nabla_{\boldsymbol{a}^{D}}(J^{D}(\boldsymbol{a}^{A}, \boldsymbol{a}^{D}, \boldsymbol{P})) = 0$$
(16)  
$$\nabla_{\boldsymbol{a}^{D}}(\gamma(\boldsymbol{a}^{A})^{T}\boldsymbol{P}\boldsymbol{\bar{Q}}\boldsymbol{a}^{D} + (\boldsymbol{a}^{D})^{T}diag(\alpha)\boldsymbol{a}^{D} + \boldsymbol{c}^{D}(\boldsymbol{Q}\boldsymbol{a}^{A} - \boldsymbol{\bar{Q}}\boldsymbol{a}^{D})) = 0$$
  
$$\gamma(\boldsymbol{a}^{A})^{T}\boldsymbol{P}\boldsymbol{\bar{Q}} + (\boldsymbol{a}^{D})^{T}(2diag(\alpha)) - \boldsymbol{c}^{D}\boldsymbol{\bar{Q}} = 0$$
  
$$\implies Br_{D}(\boldsymbol{a}^{A}, \boldsymbol{P}) = \{\boldsymbol{c}^{D}\boldsymbol{\bar{Q}}(2diag(\alpha))^{-1} - \gamma(2diag(\alpha))^{-1}\boldsymbol{\bar{Q}}^{T}\boldsymbol{P}^{T}\boldsymbol{a}^{A}\}$$

and, analogously for the attacker:

$$\nabla_{\boldsymbol{a}^{A}}(J^{D}(\boldsymbol{a}^{A}, \boldsymbol{a}^{D}, \boldsymbol{P})) = 0$$

$$-\gamma \boldsymbol{P} \boldsymbol{\bar{Q}} \boldsymbol{a}^{D} + (\boldsymbol{a}^{A})^{T} (2diag(\beta)) - \boldsymbol{c}^{A} \boldsymbol{Q} = 0$$

$$\boldsymbol{a}^{A} = (2diag(\beta))^{-1} \boldsymbol{c}^{A} \boldsymbol{Q} + \gamma (2diag(\beta))^{-1} \boldsymbol{P} \boldsymbol{\bar{Q}} \boldsymbol{a}^{D}$$

$$\implies Br_{A}(\boldsymbol{a}^{D}, \boldsymbol{P}) = \{(2diag(\beta))^{-1} \boldsymbol{c}^{A} \boldsymbol{Q} + \gamma (2diag(\beta))^{-1} \boldsymbol{P} \boldsymbol{\bar{Q}} \boldsymbol{a}^{D} \}$$

## Equilibrium Analysis (1/3)

Since J<sup>A</sup>, J<sup>D</sup> are strictly convex, the best-response correspondences are obtained from the first order conditions:

$$\nabla_{\boldsymbol{a}^{D}}(J^{D}(\boldsymbol{a}^{A}, \boldsymbol{a}^{D}, \boldsymbol{P})) = 0$$
(18)  
$$\nabla_{\boldsymbol{a}^{D}}(\gamma(\boldsymbol{a}^{A})^{T}\boldsymbol{P}\boldsymbol{\bar{Q}}\boldsymbol{a}^{D} + (\boldsymbol{a}^{D})^{T}diag(\alpha)\boldsymbol{a}^{D} + \boldsymbol{c}^{D}(\boldsymbol{Q}\boldsymbol{a}^{A} - \boldsymbol{\bar{Q}}\boldsymbol{a}^{D})) = 0$$
  
$$\gamma(\boldsymbol{a}^{A})^{T}\boldsymbol{P}\boldsymbol{\bar{Q}} + (\boldsymbol{a}^{D})^{T}(2diag(\alpha)) - \boldsymbol{c}^{D}\boldsymbol{\bar{Q}} = 0$$
  
$$\implies Br_{D}(\boldsymbol{a}^{A}, \boldsymbol{P}) = \{\boldsymbol{c}^{D}\boldsymbol{\bar{Q}}(2diag(\alpha))^{-1} - \gamma(2diag(\alpha))^{-1}\boldsymbol{\bar{Q}}^{T}\boldsymbol{P}^{T}\boldsymbol{a}^{A}\}$$

and, analogously for the attacker:

$$\nabla_{\mathbf{a}^{A}}(J^{D}(\mathbf{a}^{A}, \mathbf{a}^{D}, \mathbf{P})) = 0$$
(19)  

$$-\gamma \mathbf{P}\bar{\mathbf{Q}}\mathbf{a}^{D} + (\mathbf{a}^{A})^{T}(2diag(\beta)) - \mathbf{c}^{A}\mathbf{Q} = 0$$
  

$$\mathbf{a}^{A} = (2diag(\beta))^{-1}\mathbf{c}^{A}\mathbf{Q} + \gamma(2diag(\beta))^{-1}\mathbf{P}\bar{\mathbf{Q}}\mathbf{a}^{D}$$
  

$$\implies Br_{A}(\mathbf{a}^{D}, \mathbf{P}) = \{(2diag(\beta))^{-1}\mathbf{c}^{A}\mathbf{Q} + \gamma(2diag(\beta))^{-1}\mathbf{P}\bar{\mathbf{Q}}\mathbf{a}^{D}\}$$

## Equilibrium Analysis (2/3)

For notational convenience, define

$$\theta^{D}(\boldsymbol{c}^{D}\bar{\boldsymbol{Q}},\alpha) \triangleq [(\boldsymbol{c}^{D}\bar{\boldsymbol{Q}})_{1}/(2\alpha_{1}),\ldots,(\boldsymbol{c}^{D}\bar{\boldsymbol{Q}})_{D_{max}}/(2\alpha_{D_{max}})]$$
$$\theta^{A}(\boldsymbol{c}^{A}\boldsymbol{Q},\beta) \triangleq [(\boldsymbol{c}^{A}\boldsymbol{Q})_{1}/(2\beta_{1}),\ldots,(\boldsymbol{c}^{A}\boldsymbol{Q})_{A_{max}}/(2\beta_{A_{max}})]$$

Then we can write the best response functions as:

$$Br_D(\boldsymbol{a}^A, \boldsymbol{P}) = [\theta^D - \gamma(diag(2\alpha))^{-1} \bar{\boldsymbol{Q}}^T \boldsymbol{P}^T \boldsymbol{a}^A]^+$$
  
$$Br_A(\boldsymbol{a}^D, \boldsymbol{P}) = [\theta^A + \gamma(diag(2\beta))^{-1} \boldsymbol{P} \bar{\boldsymbol{Q}} \boldsymbol{a}^D]^+$$

The set of Nash equilibria is

 $\{(\boldsymbol{a}^{A}, \boldsymbol{a}^{D}) \mid \boldsymbol{a}^{A} \in Br_{A}(\boldsymbol{a}^{D}, \boldsymbol{P}), \boldsymbol{a}^{D} \in Br_{D}(\boldsymbol{a}^{A}, \boldsymbol{P})\}$ (20)

How large is this set? What do the elements of this set look like?

## Equilibrium Analysis (3/3)

► For notational convenience, define

 $\theta^{D}(\boldsymbol{c}^{D}\bar{\boldsymbol{Q}},\alpha) \triangleq [(\boldsymbol{c}^{D}\bar{\boldsymbol{Q}})_{1}/(2\alpha_{1}),\ldots,(\boldsymbol{c}^{D}\bar{\boldsymbol{Q}})_{D_{max}}/(2\alpha_{D_{max}})]$  $\theta^{A}(\boldsymbol{c}^{A}\boldsymbol{Q},\beta) \triangleq [(\boldsymbol{c}^{A}\boldsymbol{Q})_{1}/(2\beta_{1}),\ldots,(\boldsymbol{c}^{A}\boldsymbol{Q})_{A_{max}}/(2\beta_{A_{max}})]$ 

▶ Then we can write the best response functions as:

$$Br_D(\boldsymbol{a}^A, \boldsymbol{P}) = [\theta^D - \gamma(diag(2\alpha))^{-1} \bar{\boldsymbol{Q}}^T \boldsymbol{P}^T \boldsymbol{a}^A]^+$$
  
$$Br_A(\boldsymbol{a}^D, \boldsymbol{P}) = [\theta^A + \gamma(diag(2\beta))^{-1} \boldsymbol{P} \bar{\boldsymbol{Q}} \boldsymbol{a}^D]^+$$

#### The set of Nash equilibria is

$$\{(\boldsymbol{a}^{A}, \boldsymbol{a}^{D}) \mid \boldsymbol{a}^{A} \in Br_{A}(\boldsymbol{a}^{D}, \boldsymbol{P}), \boldsymbol{a}^{D} \in Br_{D}(\boldsymbol{a}^{A}, \boldsymbol{P})\}$$
(21)

How large is this set? What do the elements of this set look like?

## Main Contribution of the Paper

#### Theorem

There exists a unique NE. Further, if:

$$\gamma < (22)$$

$$\min \left[ \frac{\min_{i} \theta^{D}}{\max_{i} (diag(2\alpha))^{-1} \bar{\boldsymbol{Q}}^{T} \boldsymbol{P}^{T} \theta^{A}}, \frac{\max_{i} \theta^{A}}{\max_{i} (diag(2\beta))^{-1} (-\boldsymbol{P}) \bar{\boldsymbol{Q}} \theta^{D}} \right]$$

Then the unique NE  $(\mathbf{a}^{D*}, \mathbf{a}^{A*})$  satisfy  $\mathbf{a}^{D,*} > 0$  and  $\mathbf{a}^{A,*} > 0$  and is given by:

$$\boldsymbol{a}^{A*} = (I+Z)^{-1} [\theta^A + \gamma (diag(2\beta))^{-1} \boldsymbol{P} \bar{\boldsymbol{Q}} \theta^D]$$
(23)

$$\boldsymbol{a}^{D*} = (\boldsymbol{I} + \bar{\boldsymbol{Z}})^{-1} [\theta^D - \gamma(\operatorname{diag}(2\alpha))^{-1} \bar{\boldsymbol{Q}}^T \boldsymbol{P}^T \theta^A] \qquad (24)$$

where  $Z \triangleq \gamma^2 (\operatorname{diag}(2\beta))^{-1} \boldsymbol{P} \bar{\boldsymbol{Q}} (\operatorname{diag}(2\alpha))^{-1} \bar{\boldsymbol{Q}}^T \boldsymbol{P}^T$  and  $\bar{Z} \triangleq \gamma^2 (\operatorname{diag}(2\alpha))^{-1} \bar{\boldsymbol{Q}}^T \boldsymbol{P}^T (\operatorname{diag}(2\beta))^{-1} \boldsymbol{P} \bar{\boldsymbol{Q}}.$ 

## Rosen's Existence Theorem<sup>23</sup>

## Theorem (Pure Nash Equilibrium Existence for Continuous-Kernel Games)

For each player  $i \in N$ , let  $A_i$  be a compact and convex subset of a finite-dimensional Euclidean space, and the cost functional  $J^i : A_1 \times \ldots \times A_N \to \mathbb{R}$  be jointly continuous in all its arguments and strictly convex in  $a_i$  for every  $a_j \in A_j, j \in N, j \neq i$ . Then, the associated N-person nonzero-sum game admits a Nash equilibrium in pure strategies.

<sup>&</sup>lt;sup>2</sup>T. Başar and G.J. Olsder. *Dynamic Noncooperative Game Theory*. Classics in Applied Mathematics. Society for Industrial and Applied Mathematics, 1999. ISBN: 9780898714296. URL: https://books.google.se/books?id=k1oF5AxmJ1YC.

<sup>&</sup>lt;sup>3</sup>J. B. Rosen. "Existence and Uniqueness of Equilibrium Points for Concave N-Person Games". In: *Econometrica* 33.3 (1965), pp. 520–534. ISSN: 00129682, 14680262. URL: http://www.jstor.org/stable/1911749.

## Proof of Theorem 1, Existence

#### **Existence of Pure NE**:

- $A^A, A^D$  are convex subsets of a Euclidean space
- ► J<sup>A</sup>, J<sup>D</sup> are jointly continuous in all their arguments and strictly convex in a<sup>A</sup>, a<sup>D</sup> respectively,

#### $\blacktriangleright$ $A^A, A^D$ are not compact.

- However,  $J^D(a^A, a^D, P)$  and  $J^A(a^A, a^D, P)$  grow unbounded as  $|a| \rightarrow \infty$
- $\blacktriangleright \implies$  by Rosen's existence theorem, the game has a pure Nash equilibrium.

## Proof of Theorem 1, Existence

#### **Existence of Pure NE**:

- $A^A, A^D$  are convex subsets of a Euclidean space
- ► J<sup>A</sup>, J<sup>D</sup> are jointly continuous in all their arguments and strictly convex in a<sup>A</sup>, a<sup>D</sup> respectively,

#### • $A^A, A^D$ are not compact.

- However,  $J^D(a^A, a^D, P)$  and  $J^A(a^A, a^D, P)$  grow unbounded as  $|a| \rightarrow \infty$
- $\blacktriangleright \implies$  by Rosen's existence theorem, the game has a pure Nash equilibrium.

## Proof of Theorem 1, Existence

#### **Existence of Pure NE**:

- $A^A, A^D$  are convex subsets of a Euclidean space
- ► J<sup>A</sup>, J<sup>D</sup> are jointly continuous in all their arguments and strictly convex in a<sup>A</sup>, a<sup>D</sup> respectively,
- $A^A, A^D$  are not compact.
- However,  $J^D(a^A, a^D, P)$  and  $J^A(a^A, a^D, P)$  grow unbounded as  $|a| \to \infty$
- $\blacktriangleright \implies$  by Rosen's existence theorem, the game has a pure Nash equilibrium.

Rosen's Uniqueness Theorem (1/3) - Pseudo-gradient<sup>4</sup>

Definition (Pseudo-Gradient g(a) and Pseudo-Gradient Operator  $\overline{\nabla}$ )

Let  $\overline{\nabla}$  be the pseudo-gradient operator, defined through its application on the cost vector J as:

$$\bar{\nabla}J \triangleq \begin{bmatrix} \frac{\partial J_{1}(a)}{\partial a_{1}} \\ \frac{\partial J_{2}(a)}{\partial a_{2}} \\ \vdots \\ \frac{\partial J_{|N|}(a)}{\partial a_{|N|}} \end{bmatrix} = g(a)$$
(25)

<sup>&</sup>lt;sup>4</sup>J. B. Rosen. "Existence and Uniqueness of Equilibrium Points for Concave N-Person Games". In: *Econometrica* 33.3 (1965), pp. 520–534. ISSN: 00129682, 14680262. URL: http://www.jstor.org/stable/1911749.

## Proof Preliminaries (2/3) - Pseudo-Hessian<sup>5</sup>

Definition (Pseudo-Hessian)

Let G(a) be the Jacobian of the pseudo-gradient g(a) with respect to a (also called pseudo-Hessian):

$$G(a) \triangleq \begin{bmatrix} \frac{\partial^2 J_1(a)}{\partial a_1^2} & \frac{\partial^2 J_1(a)}{\partial a_1 \partial a_2} & \cdots & \frac{\partial^2 J_1(a)}{\partial a_1 \partial a_{|N|}} \\ \vdots & & \ddots & \vdots \\ \frac{\partial^2 J_{|N|}(a)}{\partial a_{|N|} \partial a_1} & \frac{\partial^2 J_{|N|}(a)}{\partial a_{|N|} \partial a_2} & \cdots & \frac{\partial^2 J_{|N|}(a)}{\partial a_{|N|}^2} \end{bmatrix}$$
(26)

#### Definition (Symmetrized Pseudo-Hessian)

Let G(a) be the Jacobian of the pseudo-gradient g(a) with respect to a, i.e. the pseudo-Hessian, then the symmetrized pseudo-hessian is defined as:

$$\mathcal{G}(a) \triangleq G(a) + G(a)^T$$
 (27)

## Proof Preliminaries (3/3) - Rosen's Uniqueness Theorem<sup>6</sup>

Theorem (Unique Pure Nash Equilibrium Existence for Continuous-Kernel Games)

If the symmetrized pseudo-Hessian  $\mathcal{G}(a)$  is positive definite, the pure equilibrium of a continuous-kernel game with strictly convex cost functions is unique.

<sup>&</sup>lt;sup>6</sup>J. B. Rosen. "Existence and Uniqueness of Equilibrium Points for Concave N-Person Games". In: *Econometrica* 33.3 (1965), pp. 520–534. ISSN: 00129682, 14680262. URL: http://www.jstor.org/stable/1911749.

#### The pseudo-gradient is:

$$\bar{\nabla}J(\boldsymbol{a}) = \begin{bmatrix} (\gamma(\boldsymbol{a}^{A})^{T}\boldsymbol{P}\bar{\boldsymbol{Q}})_{1} + (\boldsymbol{a}^{D})^{T}(2\alpha_{1}) - (\boldsymbol{c}^{D}\bar{\boldsymbol{Q}})_{1} & \dots & (\gamma(\boldsymbol{a}^{A})^{T}\boldsymbol{P}\bar{\boldsymbol{Q}})_{D_{max}} + (\boldsymbol{a}^{D})^{T}(2\alpha_{D_{max}}) - (\boldsymbol{c}^{D}\bar{\boldsymbol{Q}})_{D_{max}} \\ -(\gamma\boldsymbol{P}\bar{\boldsymbol{Q}}\boldsymbol{a}^{D})_{1} + (\boldsymbol{a}^{A})^{T}(2\beta_{1}) - (\boldsymbol{c}^{A}\boldsymbol{Q})_{1} & \dots & -(\gamma\boldsymbol{P}\bar{\boldsymbol{Q}}\boldsymbol{a}^{D})_{A_{max}} + (\boldsymbol{a}^{A})^{T}(2\beta_{A_{max}}) - (\boldsymbol{c}^{A}\boldsymbol{Q})_{A_{max}} \end{bmatrix}$$

$$\tag{28}$$

#### The pseudo-gradient is:

$$\bar{\nabla}J(\boldsymbol{a}) = \begin{bmatrix} (\gamma(\boldsymbol{a}^{A})^{T}\boldsymbol{P}\bar{\boldsymbol{Q}})_{1} + (\boldsymbol{a}^{D})^{T}(2\alpha_{1}) - (\boldsymbol{c}^{D}\bar{\boldsymbol{Q}})_{1} & \dots & (\gamma(\boldsymbol{a}^{A})^{T}\boldsymbol{P}\bar{\boldsymbol{Q}})_{Dmax} + (\boldsymbol{a}^{D})^{T}(2\alpha_{Dmax}) - (\boldsymbol{c}^{D}\bar{\boldsymbol{Q}})_{Dmax} \\ -(\gamma\boldsymbol{P}\bar{\boldsymbol{Q}}\boldsymbol{a}^{D})_{1} + (\boldsymbol{a}^{A})^{T}(2\beta_{1}) - (\boldsymbol{c}^{A}\boldsymbol{Q})_{1} & \dots & -(\gamma\boldsymbol{P}\bar{\boldsymbol{Q}}\boldsymbol{a}^{D})_{A_{max}} + (\boldsymbol{a}^{A})^{T}(2\beta_{A_{max}}) - (\boldsymbol{c}^{A}\boldsymbol{Q})_{A_{max}} \end{bmatrix}$$

$$\tag{29}$$

The pseudo-hessian is:

$$G(\mathbf{a}) = \begin{bmatrix} 2\alpha_{1} & 0 & 0 & | & & \\ 0 & \ddots & 0 & | & \gamma P \bar{\mathbf{Q}} \\ 0 & 0 & 2\alpha_{D_{max}} & | & & \\ - & - & - & | & - & - \\ & & | & 2\beta_{1} & 0 & 0 \\ & -\gamma P \bar{\mathbf{Q}} & | & 0 & \ddots & 0 \\ & & | & 0 & 0 & 2\beta_{A_{max}} \end{bmatrix}$$
(30)

The pseudo-gradient is:

$$\bar{\nabla}J(\mathbf{a}) = \begin{bmatrix} (\gamma(\mathbf{a}^{A})^{T} P \bar{\mathbf{Q}})_{1} + (\mathbf{a}^{D})^{T}(2\alpha_{1}) - (\mathbf{c}^{D} \bar{\mathbf{Q}})_{1} & \dots & (\gamma(\mathbf{a}^{A})^{T} P \bar{\mathbf{Q}})_{D_{max}} + (\mathbf{a}^{D})^{T}(2\alpha_{D_{max}}) - (\mathbf{c}^{D} \bar{\mathbf{Q}})_{D_{max}} \\ -(\gamma P \bar{\mathbf{Q}} \mathbf{a}^{D})_{1} + (\mathbf{a}^{A})^{T}(2\beta_{1}) - (\mathbf{c}^{A} \mathbf{Q})_{1} & \dots & -(\gamma P \bar{\mathbf{Q}} \mathbf{a}^{D})_{A_{max}} + (\mathbf{a}^{A})^{T}(2\beta_{A_{max}}) - (\mathbf{c}^{A} \mathbf{Q})_{A_{max}} \end{bmatrix}$$
(31)

#### The pseudo-hessian is:

$$G(\boldsymbol{a}) = \begin{bmatrix} 2\alpha_{1} & 0 & 0 & | & & \\ 0 & \ddots & 0 & | & \gamma \boldsymbol{P} \bar{\boldsymbol{Q}} \\ 0 & 0 & 2\alpha_{D_{max}} & | & & \\ - & - & - & | & - & - \\ & & | & 2\beta_{1} & 0 & 0 \\ & & | & 0 & \ddots & 0 \\ & & & | & 0 & 0 & 2\beta_{A_{max}} \end{bmatrix}$$
(32)

Clearly,  $\mathcal{G}(\boldsymbol{a}) = G(\boldsymbol{a}) + G(\boldsymbol{a})^T = 4 diag([\alpha, \beta]^T)$ , which is positive definite.

The pseudo-gradient is:

$$\bar{\nabla}J(a) = \begin{bmatrix} (\gamma(a^{A})^{T} P \bar{Q})_{1} + (a^{D})^{T} (2\alpha_{1}) - (c^{D} \bar{Q})_{1} & \dots & (\gamma(a^{A})^{T} P \bar{Q})_{Dmax} + (a^{D})^{T} (2\alpha_{Dmax}) - (c^{D} \bar{Q})_{Dmax} \\ -(\gamma P \bar{Q} a^{D})_{1} + (a^{A})^{T} (2\beta_{1}) - (c^{A} Q)_{1} & \dots & -(\gamma P \bar{Q} a^{D})_{Amax} + (a^{A})^{T} (2\beta_{Amax}) - (c^{A} Q)_{Amax} \end{bmatrix}$$
(33)

#### The pseudo-hessian is:

$$G(\mathbf{a}) = \begin{bmatrix} 2\alpha_{1} & 0 & 0 & | & & \\ 0 & \ddots & 0 & | & \gamma \mathbf{P}\bar{\mathbf{Q}} & \\ 0 & 0 & 2\alpha_{D_{max}} & | & & \\ - & - & - & | & - & - & - \\ & & | & 2\beta_{1} & 0 & 0 \\ & -\gamma \mathbf{P}\bar{\mathbf{Q}} & | & 0 & \ddots & 0 \\ z & & | & 0 & 0 & 2\beta_{A_{max}} \end{bmatrix}$$
(34)

Clearly,  $\mathcal{G}(\mathbf{a}) = G(\mathbf{a}) + G(\mathbf{a})^T = 4 \operatorname{diag}([\alpha, \beta]^T)$ , which is positive definite. Thus, by Rosen's Uniqueness theorem, the NE is unique

Proof of Theorem 1, Analytical Characterization of NE

**Recall the Best Response Functions:** 

$$Br_D(\boldsymbol{a}^A, \boldsymbol{P}) = [\theta^D - \gamma(diag(2\alpha))^{-1} \bar{\boldsymbol{Q}}^T \boldsymbol{P}^T \boldsymbol{a}^A]^+$$
  
$$Br_A(\boldsymbol{a}^D, \boldsymbol{P}) = [\theta^A + \gamma(diag(2\beta))^{-1} \boldsymbol{P} \bar{\boldsymbol{Q}} \boldsymbol{a}^D]^+$$

Proof of Theorem 1, Analytical Characterization of NE

**Recall the Best Response Functions:** 

$$Br_D(\boldsymbol{a}^A, \boldsymbol{P}) = [\theta^D - \gamma(diag(2\alpha))^{-1} \bar{\boldsymbol{Q}}^T \boldsymbol{P}^T \boldsymbol{a}^A]^+$$
  
$$Br_A(\boldsymbol{a}^D, \boldsymbol{P}) = [\theta^A + \gamma(diag(2\beta))^{-1} \boldsymbol{P} \bar{\boldsymbol{Q}} \boldsymbol{a}^D]^+$$

Substitute  $a^D$  in  $Br_A(a^D, P)$  with  $Br_D(a^A, P)$ , we then obtain the fixed point equation:

$$\begin{aligned} \mathbf{a}^{A*} &= Br_{A}(Br_{D}(\mathbf{a}^{A*}, \mathbf{P}), \mathbf{P}) \\ &= [\theta^{A} + \gamma(diag(2\beta))^{-1} \mathbf{P} \bar{\mathbf{Q}} [\theta^{D} - \gamma(diag(2\alpha))^{-1} \bar{\mathbf{Q}}^{T} \mathbf{P}^{T} \mathbf{a}^{A*}]^{+}]^{+} \end{aligned}$$

#### Proof of Theorem 1, Analytical Characterization of NE

**Recall the Best Response Functions:** 

$$Br_D(\boldsymbol{a}^A, \boldsymbol{P}) = [\theta^D - \gamma(diag(2\alpha))^{-1} \bar{\boldsymbol{Q}}^T \boldsymbol{P}^T \boldsymbol{a}^A]^+$$
  
$$Br_A(\boldsymbol{a}^D, \boldsymbol{P}) = [\theta^A + \gamma(diag(2\beta))^{-1} \boldsymbol{P} \bar{\boldsymbol{Q}} \boldsymbol{a}^D]^+$$

Substitute  $a^D$  in  $Br_A(a^D, P)$  with  $Br_D(a^A, P)$ , we then obtain the fixed point equation:

$$\begin{aligned} \mathbf{a}^{A*} &= Br_{A}(Br_{D}(\mathbf{a}^{A*}, \mathbf{P}), \mathbf{P}) \\ &= [\theta^{A} + \gamma(diag(2\beta))^{-1} \mathbf{P} \bar{\mathbf{Q}} [\theta^{D} - \gamma(diag(2\alpha))^{-1} \bar{\mathbf{Q}}^{T} \mathbf{P}^{T} \mathbf{a}^{A*}]^{+}]^{+} \end{aligned}$$

Solving for  $a^{A*}$  and similarly for  $a^{D*}$  yields:

$$\boldsymbol{a}^{A*} = (I+Z)^{-1} [\theta^A + \gamma (diag(2\beta))^{-1} \boldsymbol{P} \bar{\boldsymbol{Q}} \theta^D] \qquad (35)$$
$$\boldsymbol{a}^{D*} = (I+\bar{Z})^{-1} [\theta^D - \gamma (diag(2\alpha))^{-1} \bar{\boldsymbol{Q}}^T \boldsymbol{P}^T \theta^A] \qquad (36)$$

## Conclusion

#### Topic:

The paper provides a game theoretic analysis of intrusion detection

#### Contributions:

- A finite extensive form non-cooperative game model
- A infinite continuous-kernel strategic non-cooperative game model
- Existence and uniqueness proof of NE
- Repeated game simulation

#### Discussion

#### General questions/Comments?

Are there other existence/uniqueness theorems that could have been used?

#### Are cyber attacks continuous?

- The continuous-kernel model provide a richer analytical analysis
- But, does it make sense in practice?
- Which Model makes most sense:
  - Finite game model with NE in mixed strategies
  - Infinite game model with NE in pure strategies