# Game Theory in Robotics: Patrolling

Tomáš Kroupa

Artificial Intelligence Center
Department of Computer Science
Faculty of Electrical Engineering
Czech Technical University in Prague

2021

# Patrolling in Mobile Robotics

*To patrol is to keep watch over an area*
*by regularly walking or travelling around it.*

- The mobile surveillance of an area in order
    - to detect an adversary and
    - to give some guarantees of doing so
- The agents are called the patroller (defender) and the intruder (attacker)

# Classification of Patrolling Models

## Area representation

① *graph*
  - open perimeter
  - closed perimeter
  - fully connected

② geometric
  - lines
  - polygons

## Number of patrollers

① *single agent*

② multiple agents

## Objective function

① non-adversarial
  - maximize repeated coverage
  - maximize worst idleness

② *adversarial*
  - the environment with targets of different values
  - a rational attacker tries to intrude into the targets

# Lecture Goals/Outline

*To understand how*

- 👉 *how a simple patrollling problem can be modeled with game-theoretic tools and*
- 👉 *that revealing patroller's mixed strategy might not be a disadvantage.*

① Motivation: Patrolling on a digraph with targets

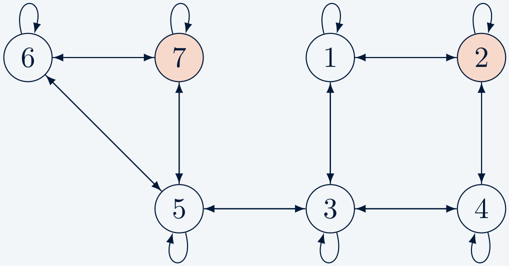
② Beyond zero-sum games: Various game forms and equilibria

③ Patrolling Security Games

# Motivation: Patrolling on a Digraph

# Example of the Environment

Labelled digraph



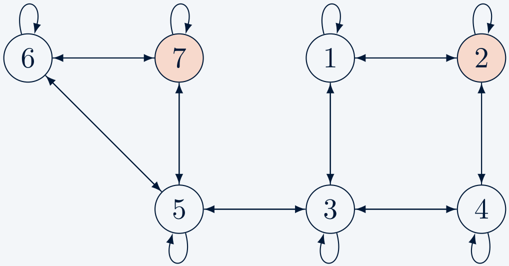
- Vertices represent the locations of the area
- Target vertices have values for the defender and the attacker
- The defender walks along the arcs to locate the attacker

## General Model of the Environment

$$\mathcal{G} = (V, E, T, v_d, v_a, \tau)$$

- $(V, E)$ is a directed graph
- $T \subseteq V$ is a nonempty set of targets
- $v_d : T \to \mathbb{R}^+$ is the value for the defender in case of successful protection
- $v_a : T \to \mathbb{R}^+$ is the value for the attacker in case of successful intrusion
- $\tau : T \to \mathbb{N}^+$ represents the time the attacker needs to spend on $t$ for getting $v_a(t)$

## Example



- $V = \{1, \ldots, 7\}$
- $T = \{2, 7\}$ with $v_d(2) = 40$, $v_d(7) = 60$ and $v_a(2) = v_a(7) = 50$
- $\tau(2) = 3$ and $\tau(7) = 2$

# The Patrolling Setting

👉 Adding the agents

### Defender

- moves along $\mathcal{G}$ spending one turn to cover one arc
- can sense only the area corresponding to the current vertex
- captures the attacker it they are at the same target $t$

### Attacker

- can wait indefinitely outside the environment
- observes the past defender's actions
- attacks a target $t$ by visiting it
- has to stay $\tau(t)$ turns in target $t$

# The Patrolling Setting

👉 Questions

- How to represent the turns/moves of agents?
- How to define the utility function of each agent?
- How to express the knowledge of attacker about defender's strategy?

*We need to look beyond two-person zero-sum games.*

# Game Forms and Equilibria

# Two-Player Zero-Sum Game

☞ *Simultaneous* moves, the utility version

1. Players are the defender and the attacker
2. Strategy sets are $M$ and $N$
3. The matrix $\mathbf{U} = [u_{ij}]_{i \in M,\, j \in N}$ of utilities for the defender

However, the players may act *sequentially* with the defender revealing the strategy:

1. The defender commits to maxmin strategy $\bar{\mathbf{x}}$
2. The attacker plays minmax strategy $\bar{\mathbf{y}}$

The outcome is $\bar{\mathbf{x}}^\intercal \mathbf{U} \bar{\mathbf{y}} =$ the value of the game

# General-Sum Game

👉 *Simultaneous* moves, the utility version

① Players are the defender and the attacker

② Strategy sets are $M$ and $N$

③ The utility matrices $\mathbf{U}^d = [u^d_{ij}]_{i \in M,\, j \in N}$ and $\mathbf{U}^a = [u^a_{ij}]_{i \in M,\, j \in N}$

When agents use mixed strategies $\mathbf{x}$ and $\mathbf{y}$, the expected utilities are

$$\mathbf{x}^\intercal \mathbf{U}^d \mathbf{y} \qquad \text{and} \qquad \mathbf{x}^\intercal \mathbf{U}^a \mathbf{y}$$

*Which strategies will utility-maximizing agents seek?*

# Nash Equilibrium

☞ Nash, 1951

A Nash equilibrium is a pair of mixed strategies $(\bar{\mathbf{x}}, \bar{\mathbf{y}})$ such that

$$\bar{\mathbf{x}}^\intercal \mathbf{U}^d \bar{\mathbf{y}} \geq \mathbf{x}^\intercal \mathbf{U}^d \bar{\mathbf{y}} \qquad \text{and} \qquad \bar{\mathbf{x}}^\intercal \mathbf{U}^a \bar{\mathbf{y}} \geq \bar{\mathbf{x}}^\intercal \mathbf{U}^a \mathbf{y} \qquad \forall \mathbf{x} \in \Delta_M, \forall \mathbf{y} \in \Delta_N$$

- 👍 Every general-sum game has at least one NE
- 👎 Computing NE is a notoriously difficult problem

# Example

☞ $3 \times 3$ bimatrix game

| | | Attacker | | |
|---|---|---|---|---|
| | | 1 | 2 | 3 |
| Defender | 1 | $6, 2$ | $0, 6$ | $4, 4$ |
| | 2 | $2, 12$ | $4, 3$ | $2, 5$ |
| | 3 | $0, 6$ | $10, 0$ | $2, 2$ |

# Example

$3 \times 3$ bimatrix game

|   | 1 | 2 | 3 |
|---|---|---|---|
| 1 | $6, 2$ | $0, 6$ | $4, 4$ |
| 2 | $2, 12$ | $4, 3$ | $2, 5$ |
| 3 | $0, 6$ | $10, 0$ | $2, 2$ |

- The second row strategy is *strictly dominated* by $\frac{1}{2}(\mathbf{e}_1 + \mathbf{e}_3)$
- So we can omit it from the defender’s pure strategies

|   | 1 | 2 | 3 |
|---|---|---|---|
| 1 | $6, 2$ | $0, 6$ | $4, 4$ |
| 3 | $0, 6$ | $10, 0$ | $2, 2$ |

# Example

$3 \times 3$ bimatrix game

|  | 1 | 2 | 3 |
|---|---|---|---|
| 1 | $6, 2$ | $0, 6$ | $4, 4$ |
| 3 | $0, 6$ | $10, 0$ | $2, 2$ |

- The third column strategy is also *strictly dominated*
- So we can omit it from the attacker's pure strategies

|  | 1 | 2 |
|---|---|---|
| 1 | $6, 2$ | $0, 6$ |
| 3 | $0, 6$ | $10, 0$ |

$$\bar{\mathbf{x}} = (\tfrac{3}{5}, \tfrac{2}{5})$$

$$\bar{\mathbf{y}} = (\tfrac{5}{8}, \tfrac{3}{8})$$

utilities $(\frac{15}{4}, \frac{18}{5})$

## Nash Equilibrium: Properties

For general-sum games

In contrast to zero-sum games:

- Computing NE is PPAD-complete
- Equilibrium points may have different values
- It may not be clear which equilibrium will a rational player select
- Revealing the used equilibrium strategy may reduce the player's payoff

|  | $1$ | $2$ |
|---|---|---|
| $1$ | $0, 0$ | $2, 1$ |
| $2$ | $3, 2$ | $1, 2$ |

Equilibria: $(1, 2)$ and $(2, 1)$

# Stackelberg Game

👉 Sequential moves

① Players are the defender and the attacker

② Strategy sets are $M$ and $N$

③ The utility matrices $\mathbf{U}^d = [u^d_{ij}]_{i \in M,\, j \in N}$ and $\mathbf{U}^a = [u^a_{ij}]_{i \in M,\, j \in N}$

- The defender commits to $\mathbf{x} \in \Delta_M$
- The attacker plays a pure strategy $\mathbf{e}_j$ maximizing $\mathbf{x}^\intercal \mathbf{U}^a \mathbf{e}_j$
- Typically there are more such best responses $j \in \mathrm{BR}(\mathbf{x})$

*The defender wants to maximize $\mathbf{x}^\intercal \mathbf{U}^d \mathbf{e}_j$, but which $j$ should be considered?*

# Optimal Strategy of Defender

👉 aka Strong Stackelberg equilibrium

*We assume that the attacker picks the best action for the defender.*

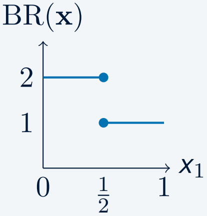
An optimal strategy of defender is a mixed strategy $\bar{\mathbf{x}} \in \Delta_M$ such that

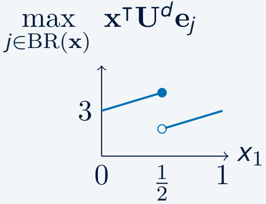
$$\max_{j \in \mathrm{BR}(\bar{\mathbf{x}})} \bar{\mathbf{x}}^\intercal \mathbf{U}^d \mathbf{e}_j = \max_{\mathbf{x} \in \Delta_M} \max_{j \in \mathrm{BR}(\mathbf{x})} \mathbf{x}^\intercal \mathbf{U}^d \mathbf{e}_j$$

# Optimal Strategy of Defender

Example

|  | 1 | 2 |
|---|---|---|
| 1 | $2, 1$ | $4, 0$ |
| 2 | $1, 0$ | $3, 1$ |





- The optimal strategy of defender is $\bar{\mathbf{x}} = (\frac{1}{2}, \frac{1}{2})$ and payoff $3.5$
- The best response of attacker is $j = 1$ or $j = 2$ with equal payoffs $0.5$

# How to Find the Optimal Strategy?

☞ Conitzer, Sandholm (2006)

① For every attacker's strategy $j \in N$, solve this LP$_j$:

$$\begin{aligned} \text{Maximize} \quad & \mathbf{x}^\intercal \mathbf{U}^d \mathbf{e}_j \\ \text{subject to} \quad & \mathbf{x} \in \Delta_M \\ & \mathbf{x}^\intercal \mathbf{U}^a \mathbf{e}_j \geq \mathbf{x}^\intercal \mathbf{U}^a \mathbf{e}_{j'} \quad \forall j' \in N \end{aligned}$$

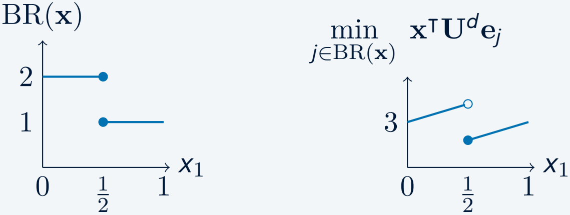
② The optimal strategy $\bar{\mathbf{x}}$ is the optimal solution of LP$_j$ with the maximal value

# The Worst-Case Optimality

aka Weak Stackelberg equilibrium

*Let's assume now that the attacker picks the worst action for the defender.*

$$\begin{array}{c|c|c|} & 1 & 2 \\ \hline 1 & 2,1 & 4,0 \\ \hline 2 & 1,0 & 3,1 \\ \hline \end{array}$$

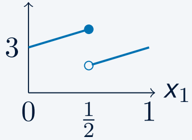


- The function on the right has no maximum
- So the optimal defender's strategy in this sense fails to exist!

# Stackelberg vs. Nash

|   | 1 | 2 |
|---|---|---|
| 1 | $2, 1$ | $4, 0$ |
| 2 | $1, 0$ | $3, 1$ |

$$\max_{j \in \mathrm{BR}(\mathbf{x})} \mathbf{x}^\intercal \mathbf{U}^d \mathbf{e}_j$$



Stackelberg

$\bar{\mathbf{x}} = (\frac{1}{2}, \frac{1}{2})$ with payoff $3.5$

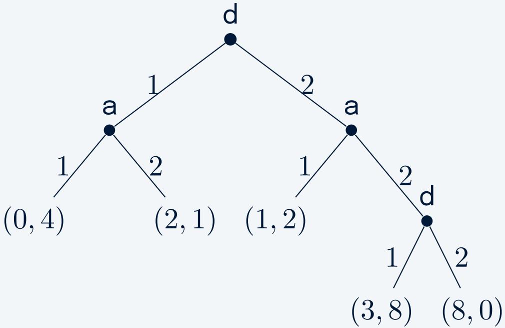
Nash

$\bar{\mathbf{x}} = (1, 0)$ with payoff $2$

*In fact, the optimal payoff of defender in the Stackelberg game is always $\geq$ the optimal payoff in any Nash equilibrium.*

## Extensive-Form Games

👉 The most general game representation



- Moves are explicitly modeled
- Actions are represented by branching
- Outcomes are at the leaves

👉 Strategies depend on the history

👉 The history of past actions may not be observable by the other player

# Patrolling Security Games

# Patrolling Security Game

👉 The environment $\mathcal{G} = (V, E, T, v_d, v_a, \tau)$

A two-player multi-stage game with imperfect information and infinite horizon:

- The agents act *simultaneously* in each move
- The game ends when a single target is captured or no attack is carried out
- The attacker can derive the defender’s strategy from observing past actions

*We need to define actions, outcomes, and utility functions.*

# Actions

👉 At turn $k$ of the game

**Defender**

① `move`$(i)$ means that the defender visits adjacent vertex $i \in V$ at turn $k + 1$ and checks it for the intruder's presence

**Attacker**

① `wait` means that the attacker makes no attempt at intrusion

② `enter`$(t)$ represents the intrusion into target $t \in T$ and blocks the attacker for the next $\tau(t)$ moves

## Outcomes of the Game

`no-attack`
The attacker plays `wait` at every turn $k$

`intruder-capture`
The attacker plays $\texttt{enter}(t)$ at turn $k$ and the patroller visits $t$ in the interval

$$\{k, \ldots, k + \tau(t) - 1\}$$

$\texttt{penetration-}t$
The attacker plays $\texttt{enter}(t)$ at turn $k$ and the patroller does not visit $t$ in that interval

# Utility Functions

👉 Outcome $x$

## Defender

$$u_d(x) = \begin{cases} \sum_{i \in T} v_d(i) & x = \texttt{intruder-capture} \quad \texttt{or} \quad \texttt{no-attack} \\ \sum_{i \in T \setminus \{t\}} v_d(i) & x = \texttt{penetration-}t \end{cases}$$

## Attacker

$$u_a(x) = \begin{cases} 0 & x = \texttt{no-attack} \\ v_a(t) & x = \texttt{penetration-}t \\ -\epsilon & x = \texttt{intruder-capture} \end{cases}$$

where $\epsilon > 0$ is a penalty

## Strategies

Let $H$ denote the set of all histories of defender's actions

Defender

$$\sigma_d : H \to \Delta_V$$

Attacker

$$\sigma_a : H \to \Delta_{T \cup \{\texttt{wait}\}}$$

*Computing defender's optimal strategy is tractable in several special cases.*

# References

- Basilico, Nicola, Nicola Gatti, and Francesco Amigoni. Patrolling Security Games: Definition and Algorithms for Solving Large Instances with Single Patroller and Single Intruder. *Artificial Intelligence* 184–185: 78–123, 2012.
- Robin, Cyril, and Simon Lacroix. Multi-Robot Target Detection and Tracking: Taxonomy and Survey. *Autonomous Robots* 40 (4): 729–60, 2016.