# Green Security Games along trails and the protection of National parks.

Mauricio Velasco*
Universidad Católica del Uruguay (UCU)

**JOINT work with:**
Nicolás Betancourt (Instituto de Recursos Biológicos Alexander von Humboldt, Colombia).

*IBERAMIA*
November, 2024
Montevideo, Uruguay

# Outline

1. Motivation: The ranger's problem.
2. Multi-armed bandits.
3. Results.

# Motivation:

Mitigating the negative impact of the current loss of biodiversity is one of the most significant challenges of the coming years.

# Motivation:

Mitigating the negative impact of the current loss of biodiversity is one of the most significant challenges of the coming years.

An essential mitigation mechanism is the *maintenance of large protected areas* as these serve as habitats for a wide variety of species as well as water reservoirs.

# Motivation:

Unfortunately, protected areas around the world are constantly under threat: **hunting, illegal logging and mining, and species trafficking**, among others.

# Motivation:

Unfortunately, protected areas around the world are constantly under threat: **hunting, illegal logging and mining, and species trafficking**, among others.

This is a challenge for those who care for these areas as they must **allocate the limited resources at their disposal to care for large areas. Caretakers (park rangers) are typically at a considerable disadvantage with respect to attackers.**

# Motivation:

Unfortunately, protected areas around the world are constantly under threat: **hunting, illegal logging and mining, and species trafficking**, among others.

This is a challenge for those who care for these areas as they must **allocate the limited resources at their disposal to care for large areas. Caretakers (park rangers) are typically at a considerable disadvantage with respect to attackers.**

**Question.** *Can we design systems that help rangers preserve protected areas?*

# Jama Coaque

The **Jama Coaque** ecological reserve in the coast of Ecuador is a protected area of $850HAs$ of *tropical rainforest* ($4\% \times$ MVD).

It is a **key resource** in the protection of biodiversity in the region.

# Jama Coaque

The **Jama Coaque** ecological reserve in the coast of Ecuador is a protected area of $850HAs$ of *tropical rainforest* ($4\% \times$ MVD).

It is a **key resource** in the protection of biodiversity in the region. From wikipedia: "The Jama-Coaque Ecological Reserve serves as habitat and key migratory channel for six endangered species of felines (jaguar, puma, ocelot, oncilla, margay, and jaguarundi) and two endangered species of primates (mantled howler monkey and white-fronted capuchin monkey). Other endangered mammals include the tayra, the three-toed sloth, the western agouti, and the spotted paca. In 2009, herpetologist Paul S. Hamilton discovered two new species of frog in the cloud forest of the Jama-Coaque Ecological Reserve".

# Jama Coaque:

# Jama Coaque:

# Jama Coaque:

- The reserve is protected by a team of rangers who patrol it. It is under constant threat, by **illegal logging** of balso wood and **illegal hunting** of zainos and deers.

# Jama Coaque:

- The reserve is protected by a team of rangers who patrol it. It is under constant threat, by **illegal logging** of balso wood and **illegal hunting** of zainos and deers.
- The terrain is so difficult that rangers (and poachers) move only along the paths in a fixed **trail graph** (with 128 edges).

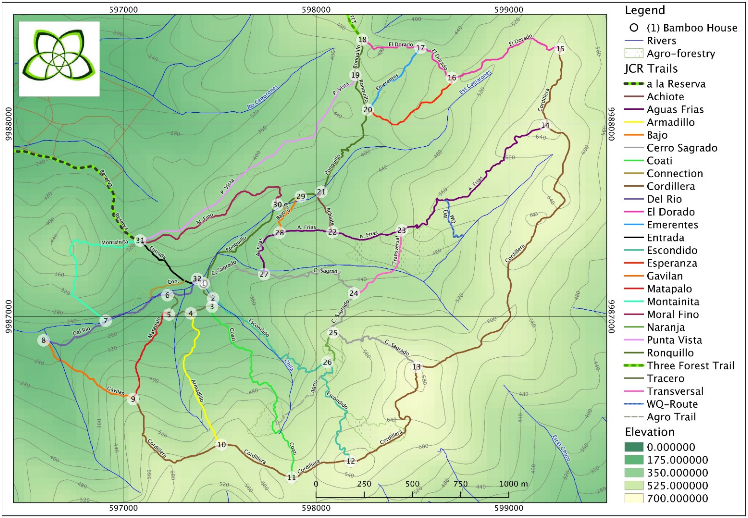
# Jama Coaque:

- The reserve is protected by a team of rangers who patrol it. It is under constant threat, by **illegal logging** of balso wood and **illegal hunting** of zainos and deers.
- The terrain is so difficult that rangers (and poachers) move only along the paths in a fixed **trail graph** (with 128 edges).
- Rangers patrol the reserve using **cycles** in this trail graph (there are more than 135.000 such cycles and more than 10.000 that go through the *bamboo house*)

# Jama Coaque:

- The reserve is protected by a team of rangers who patrol it. It is under constant threat, by **illegal logging** of balso wood and **illegal hunting** of zainos and deers.
- The terrain is so difficult that rangers (and poachers) move only along the paths in a fixed **trail graph** (with 128 edges).
- Rangers patrol the reserve using **cycles** in this trail graph (there are more than 135.000 such cycles and more than 10.000 that go through the *bamboo house*)
- The team of rangers has $B = 4$ people, who patrol it every day along distinct cyclical routes.

# Jama Coaque:


Legend
(1) Bamboo House
Rivers
Agro-forestry
JCR Trails
a la Reserva
Achiote
Aguas Frias
Armadillo
Bajo
Cerro Sagrado
Coati
Connection
Cordillera
Del Rio
El Dorado
Emerentes
Entrada
Escondido
Esperanza
Gavilan
Matapalo
Montainita
Moral Fino
Naranja
Punta Vista
Ronquillo
Three Forest Trail
Tracero
Transversal
WQ-Route
Agro Trail
Elevation
0.000000
175.000000
350.000000
525.000000
700.000000
597000
598000
599000
9988000
9987000
0
250
500
750
1000 m

# Main problem:

Every day the team of rangers needs to decide on a set of $B = 4$ cycles to patrol

# Main problem:

Every day the team of rangers needs to decide on a set of $B = 4$ cycles to patrol (choosing among around $\binom{10000}{4} \sim 10^{14}$ alternatives).

# Main problem:

Every day the team of rangers needs to decide on a set of $B = 4$ cycles to patrol (choosing among around $\binom{10000}{4} \sim 10^{14}$ alternatives).

1. *How should the set of patrol cycles be selected?*
2. *How should rangers use the available data to dynamically improve their routes?*

# Main problem:

Every day the team of rangers needs to decide on a set of $B = 4$ cycles to patrol (choosing among around $\binom{10000}{4} \sim 10^{14}$ alternatives).

1. *How should the set of patrol cycles be selected?*
2. *How should rangers use the available data to dynamically improve their routes?*

*A good solution needs to combine the constant* **exploration** *of the reserve with visiting those places where rangers suspect environmental crimes are most likely to happen (***exploitation*** of previously acquired knowledge).*

Multi-armed bandits
(or how to solve an *exploration* vs
*exploitation* dilemma).

# Multi-armed bandits:

*A player has a set of $m$ slot machines (**arms**) numbered $1, 2, 3, \ldots, m$. At each turn the player chooses a machine, pulls the corresponding arm and gets a certain amount of money.*
**What strategy should the player use to maximize his return in $T$ turns?**

# Multi-armed bandits:

If we knew the mean $\mu_i$ of each arm, this problem would be very easy.

# Multi-armed bandits:

If we knew the mean $\mu_i$ of each arm, this problem would be very easy.

*We look for the best machine (arm)*

$$j^* := \mathrm{argmax}_{\mathrm{j \in [m]}}(\mu_j)$$

*and use it on every turn.*

# Multi-armed bandits:

If we knew the mean $\mu_i$ of each arm, this problem would be very easy.

*We look for the best machine (arm)*

$$j^* := \mathrm{argmax}_{\mathrm{j} \in [\mathrm{m}]}(\mu_j)$$

*and use it on every turn.*

The problem is that **the player does not know the means**. So he should use some of his time to try to learn which machines have a good return and some of his time to play in these machines.

# Multi-armed bandits:

If we knew the mean $\mu_i$ of each arm, this problem would be very easy.

*We look for the best machine (arm)*

$$j^* := \mathrm{argmax}_{\mathrm{j} \in [\mathrm{m}]}(\mu_j)$$

*and use it on every turn.*

The problem is that **the player does not know the means**. So he should use some of his time to try to learn which machines have a good return and some of his time to play in these machines.

*The* **multi-armed bandit** *is a fundamental problem because it abstracts the dilemma between* **exploration** *and* **exploitation***.*

# Solution strategies

A **strategy** $A$ is an algorithm that tells us which arm to choose in each turn based on the returns of previous moves.

# Solution strategies

A **strategy** $A$ is an algorithm that tells us which arm to choose in each turn based on the returns of previous moves.

**Definition.**

*The* **regret** *of a strategy $A$ after $n$ turns is defined as the profit loss resulting from using strategy $A$ and not the optimal strategy during the first $n$ turns.*

More precisely, if $T_i(n)$ denotes the number of times that arm $i$ has been selected during the first $n$ turns then

$$R(n) := \mu^* n - \sum_{j=1}^{m} \mu_j \mathbb{E}[T_j(n)]$$

where $\mu^* := \max_{i \in [m]} \mu_i$.

# Solution strategies

A **strategy** $A$ is an algorithm that tells us which arm to choose in each turn based on the returns of previous moves.

**Definition.**

*The* **regret** *of a strategy $A$ after $n$ turns is defined as the profit loss resulting from using strategy $A$ and not the optimal strategy during the first $n$ turns.*

More precisely, if $T_i(n)$ denotes the number of times that arm $i$ has been selected during the first $n$ turns then

$$R(n) := \mu^* n - \sum_{j=1}^{m} \mu_j \mathbb{E}[T_j(n)]$$

where $\mu^* := \max_{i \in [m]} \mu_i$.

The regret function is always nonnegative and we wish to make it **as small as possible**.

# Upper confidence bound policy (UCB):

In 2002 Auer, Cesa-Bianchi and Fischer propose the following strategy:

- *Initialization: Play each arm once.*
- *In each turn* $n \geq m + 1$ *do:*
  1. *Compute the quantities*

     $$\hat{x}_j := \overline{x_j} + \sqrt{\frac{2\ln(n)}{n_j}}$$

     *where* $\overline{x_j}$ *is the average return obtained by the* $j$*-th arm so far and* $n_j$ *is the number of times that the* $j$*-th arm has been used so far.*
  2. *Pull the* $t$*-th arm where* $t$ *is the index that maximizes* $\hat{x}_j$*. Write down the returns and update all your estimates.*

**Theorem. (Auer, Cesa-Bianchi, Fischer, 2002)**

*For every $m > 1$ and return distributions supported in $[0, 1]$, the UCB strategy satisfies the inequality*

$$R(n) \leq \left[ 8 \sum_{i:\mu_i < \mu^*} \frac{\log(n)}{\Delta_i} \right] + \left(1 + \frac{\pi^2}{3}\right) \left( \sum_{i=1}^{m} \Delta_i \right)$$

*with* $\Delta_i := \mu^* - \mu_i$

**Theorem. (Auer, Cesa-Bianchi, Fischer, 2002)**

*For every $m > 1$ and return distributions supported in $[0, 1]$, the UCB strategy satisfies the inequality*

$$R(n) \leq \left[ 8 \sum_{i:\mu_i<\mu^*} \frac{\log(n)}{\Delta_i} \right] + \left(1 + \frac{\pi^2}{3}\right) \left( \sum_{i=1}^{m} \Delta_i \right)$$

*with* $\Delta_i := \mu^* - \mu_i$

In particular, we have

$$\lim_{n\to\infty} \frac{R(n)}{n} = 0$$

**Theorem. (Auer, Cesa-Bianchi, Fischer, 2002)**

*For every $m > 1$ and return distributions supported in $[0, 1]$, the UCB strategy satisfies the inequality*

$$R(n) \leq \left[ 8 \sum_{i:\mu_i < \mu^*} \frac{\log(n)}{\Delta_i} \right] + \left(1 + \frac{\pi^2}{3}\right) \left( \sum_{i=1}^{m} \Delta_i \right)$$

*with* $\Delta_i := \mu^* - \mu_i$

The proof of the Theorem consists in verifying that for every sub-optimal arm $j$ we have $\mathbb{E}[T_j(n)] \leq \frac{8 \log(n)}{\Delta_j^2}$ via Hoeffding's inequality

**Theorem. (Auer, Cesa-Bianchi, Fischer, 2002)**

*For every $m > 1$ and return distributions supported in $[0,1]$, the UCB strategy satisfies the inequality*

$$R(n) \leq \left[8 \sum_{i:\mu_i<\mu^*} \frac{\log(n)}{\Delta_i}\right] + \left(1 + \frac{\pi^2}{3}\right)\left(\sum_{i=1}^{m} \Delta_i\right)$$

*with* $\Delta_i := \mu^* - \mu_i$

The proof of the Theorem consists in verifying that for every sub-optimal arm $j$ we have $\mathbb{E}[T_j(n)] \leq \frac{8\log(n)}{\Delta_j^2}$ via Hoeffding's inequality

It is known (Lai y Robbins, 1985) that **for every strategy** the inequality

$$\mathbb{E}[T_j(n)] \geq \frac{\log(n)}{D(p_j \| p^*)}$$

holds. so the UCB is asymptotically optimal

# Back to Jama Coaque:

Can we think of the rangers' problem as a MAB?

# Back to Jama Coaque:

Can we think of the rangers' problem as a MAB?
Yes to first approximation. Every day the ranger must choose a cycle hoping to find which ones do have illegal activity (return).

# Back to Jama Coaque:

Can we think of the rangers' problem as a MAB?
Yes to first approximation. Every day the ranger must choose a cycle hoping to find which ones do have illegal activity (return).
However, this approach has several problems:

- There are too many means to be estimated (if we simply assign one per cycle).
- By thinking of cycles as independent black boxes there is a lot of information that we are loosing (for example that different cycles share edges).
- To think that the cycles are independent of each other is a very weird assumption for very similar cycles.
- We would like to be more flexible and assign rangers to each shift.

# A model:

Back to **JamaCoaque**, we propose the following basic model:

*Illegal activity is a function of the terrain conditions (presence of certain tree species, proximity to bodies of water, etc.). We assume that such activities occur in each edge according to a Bernoulli r.v. with parameter* $p(e)$ *(sampled independently each day and among distinct edges).*

# A model:

Back to **JamaCoaque**, we propose the following basic model:

*Illegal activity is a function of the terrain conditions (presence of certain tree species, proximity to bodies of water, etc.). We assume that such activities occur in each edge according to a Bernoulli r.v. with parameter $p(e)$ (sampled independently each day and among distinct edges).*

The objective of the team of rangers is to select a subset $S$ consisting of $B$ cycles so that

$$r_p(S) := \mathbb{E}[r(S)] = \sum_{e \in \bigcup S} p(e)$$

is maximized.

# Knowing $p(e)$: Coverage problems.

**If the** $p(e)$ **were known numbers** then, selecting an optimal set $S$ becomes a variant of the **weighted coverage problem**.

# Knowing $p(e)$: Coverage problems.

**If the** $p(e)$ **were known numbers** then, selecting an optimal set $S$ becomes a variant of the **weighted coverage problem**.

*This problem is* **NP-hard** *so* **even knowing the probabilities** *the problem that rangers have to solve is computationally difficult.*

# Knowing $p(e)$: Coverage problems.

**If the** $p(e)$ **were known numbers** then, selecting an optimal set $S$ becomes a variant of the **weighted coverage problem**.

*This problem is* **NP-hard** *so* **even knowing the probabilities** *the problem that rangers have to solve is computationally difficult.*

However, there is a good **certified approximation algorithm**,

# Knowing $p(e)$: Coverage problems.

**If the** $p(e)$ **were known numbers** then, selecting an optimal set $S$ becomes a variant of the **weighted coverage problem**.

*This problem is* **NP-hard** *so* **even knowing the probabilities** *the problem that rangers have to solve is computationally difficult.*

However, there is a good **certified approximation algorithm**,

**Theorem. (Betancourt, -)**

*There exists a linear programming + sampling algorithm so that the obtained cycles are guaranteed to have weight at least* $(1-1/e)OPT$ *where* $OPT$ *is the true optimum of the problem.*

# Learning edge probabilities

In practice **rangers do not know the edge probabilities** $p(e)$. To learn them we will use the framework of **Combinatorial multi-armed bandits** [Chen, Wang, Yuan, 2013].

# Learning edge probabilities

In practice **rangers do not know the edge probabilities** $p(e)$. To learn them we will use the framework of **Combinatorial multi-armed bandits** [Chen, Wang, Yuan, 2013].

*CUCB Algorithm: Initialize. In turn* $n > 0$ *do:*

1. *Recompute estimates for the probabilities* $p_e$ *of illegal activities on each edge*

$$\hat{p}_e(n) := \min\left(\overline{p_e} + \sqrt{\frac{3\ln(n)}{2n_e}}, 1\right)$$

   *where* $\overline{p_e}$ *es the average illegal activity seen on edge* $e$ *so far and* $n_e$ *is the number of times that edge* $e$ *has been visited.*

2. *Rangers visit the cycles determined by* **our approximation algorithm**, *assuming the weights are given by the estimators* $(\hat{p}_e(n))_e$ *obtained in* (1).

# Approximate regret

**Definition.**

*We define the* **regret** *of any strategy $A$ as*

$$R(n) = n(1 - 1/e)OPT(p) - \sum_{t=1}^{n} \mathbb{E}_p \left[ r(S_t^A) \right]$$

# Approximate regret

**Definition.**

*We define the* **regret** *of any strategy A as*

$$R(n) = n(1 - 1/e)OPT(p) - \sum_{t=1}^{n} \mathbb{E}_p \left[ r(S_t^A) \right]$$

Observations:

- The regret depends on the (unknown) true probabilities $p$.
- Compares the optimum that one could guarantee practically (with a polynomial time approximation algorithm) with the output of our CUCB algorithm (on average).

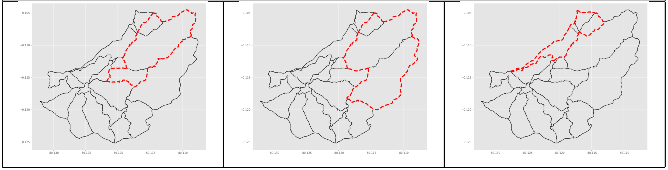
**Theorem. (Betancourt, - )**

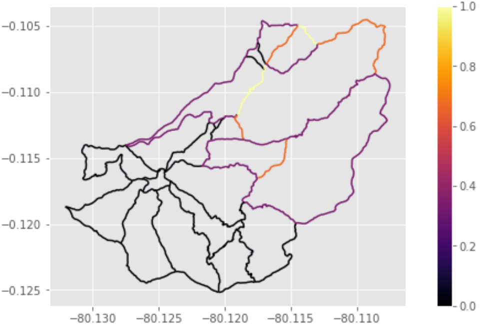
*If m is the number of trails then the CUCB satisfies*

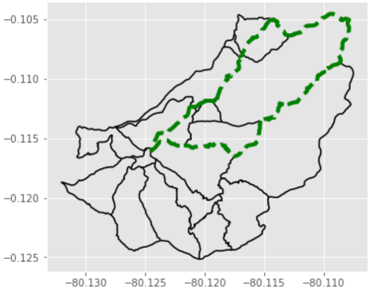
$$R(n) \leq \left[ \frac{6 \log(n)}{\Delta_{\min}^2} + \frac{\pi^2}{3} + 1 \right] m\Delta_{\max} = O(m \log(n)).$$

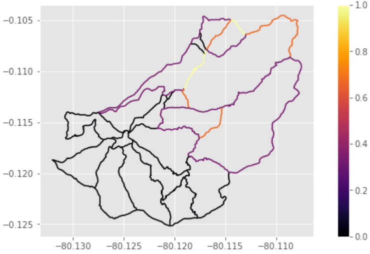
# Example:

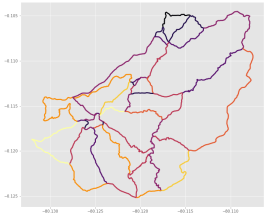
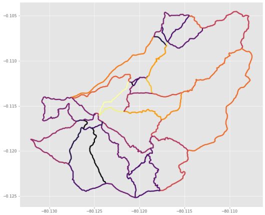
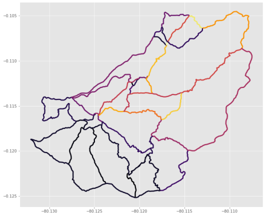
Poacher routes



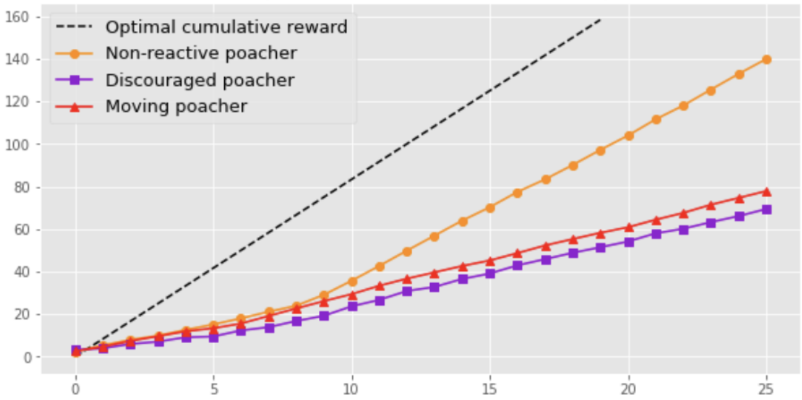





| CUCB patrol frequencies | | |
|---|---|---|
| $t = 10$ | $t = 50$ | $t = 180$ |
|  |  |  |

# Returns during 25 days:



# Combined with accoustic monitoring systems: