Reminder

- Course project progress report 2: come to OH for discussions!
- HW5 due 4/4
- PRA6 due 4/16

Artificial Intelligence Methods for Social Good

Lecture 22

Human Behavior Modeling and Resource Allocation in Security Applications

17-537 (9-unit) and 17-737 (12-unit)

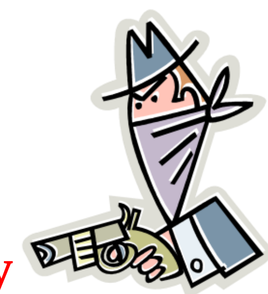
Instructor: Fei Fang

feifang@cmu.edu

Recap: Stackelberg Security Games

- Stackelberg Security game
  - Defender: Commits to mixed strategy
  - Adversary: Conduct surveillance and best responds
- Expected Utility

$$AttEU(i) = c_i P_i^a + (1 - c_i) R_i^a$$
$$DefEU(i) = c_i R_i^d + (1 - c_i) P_i^d$$



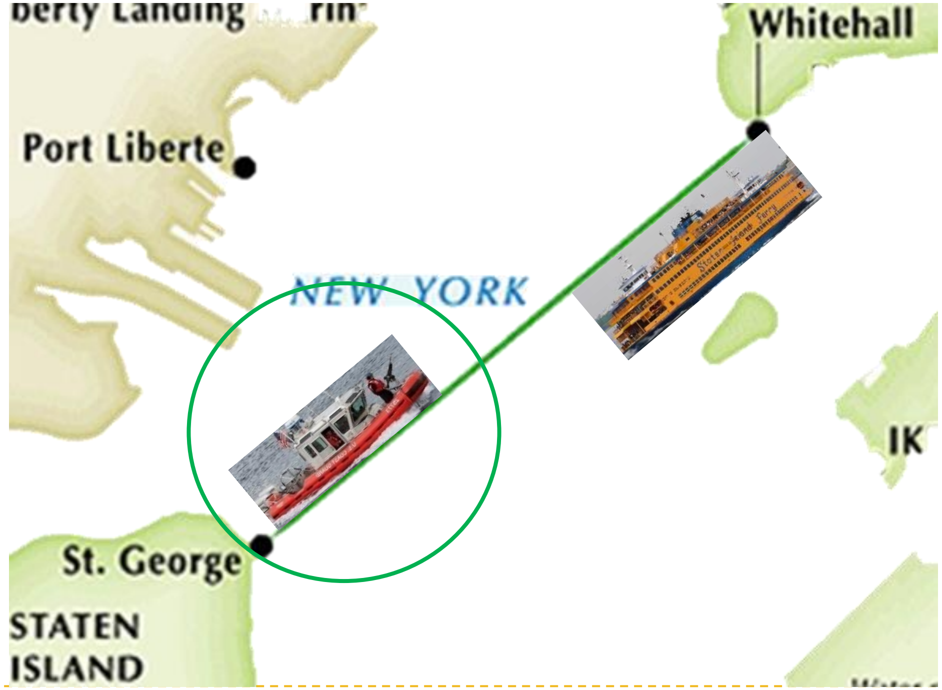
Adversary

Defender

		Target #1	Target #2
55.6%	Target #1	5, -3	-1, 1
44.4%	Target #2	-5, 4	2, -1

Recap: Game Theory for Ferry Protection

- Optimize the use of patrol resources



Green Security Domains

- How are these domains similar to / different from airport / port security?
  - Similarity:
  - Difference:

Environmental Resources

Endangered Wildlife

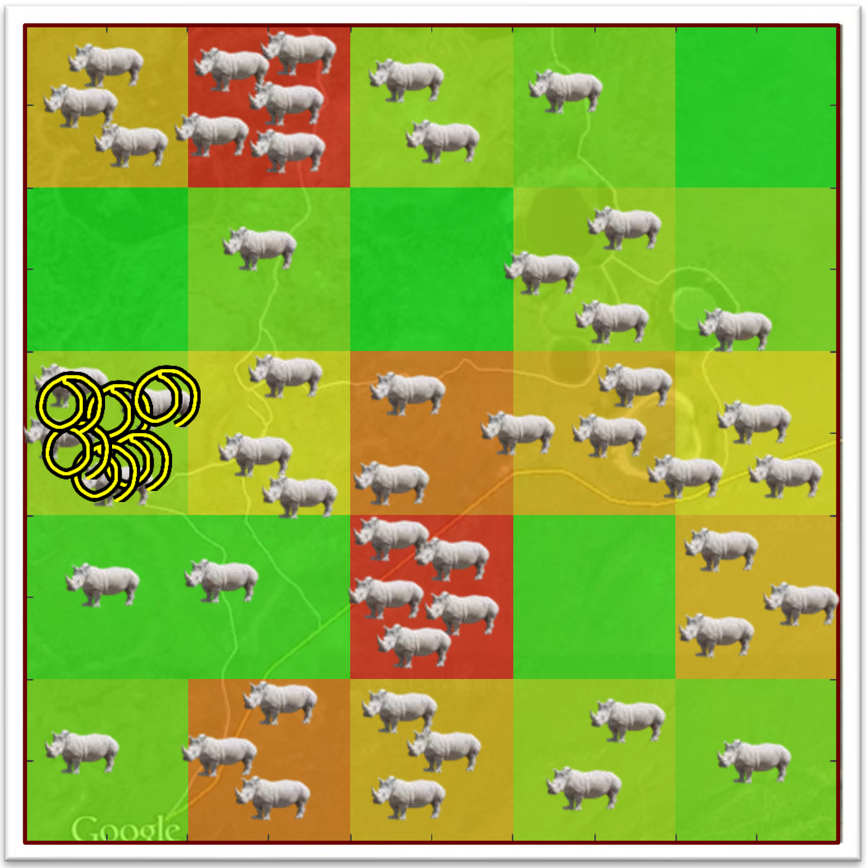
Fisheries

Challenges in Green Security Domains

- Frequent and repeated attacks
  - Not one-shot
- Attacker decision making
  - Limited surveillance / Less effort / Boundedly rational
- Real-world data
  - Sparse / Incomplete / Uncertainty / Noise
- Real-world deployment
  - Practical constraints
  - Field test

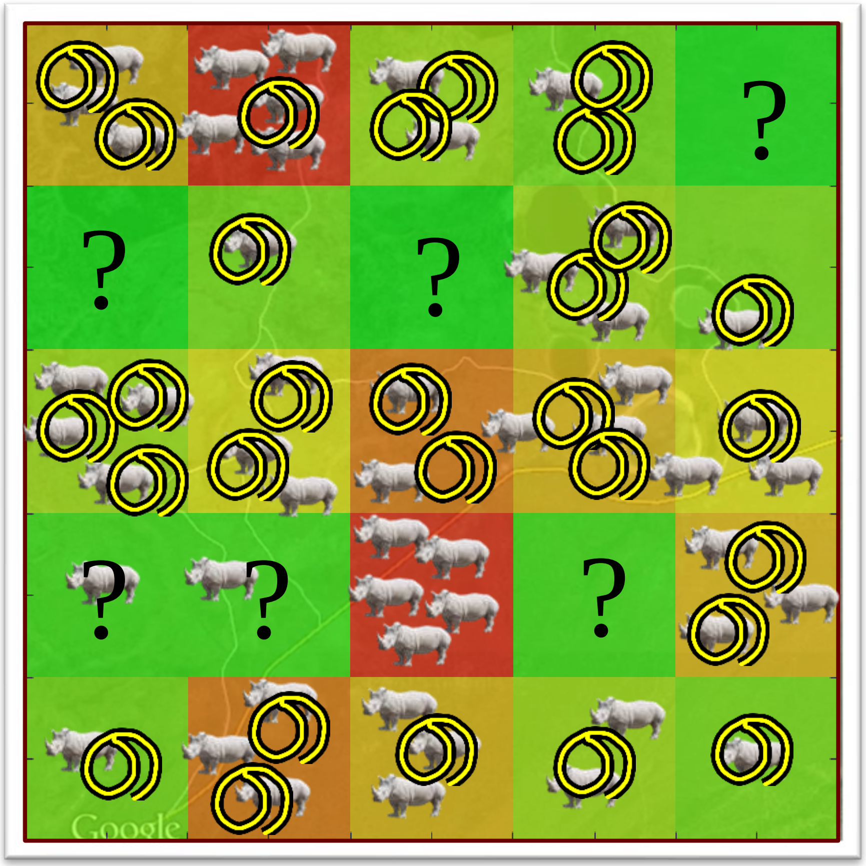
Challenges in Wildlife Conservation Domain

- Perfectly rational (Maximize expected utility)? No!



Challenges in Wildlife Conservation Domain

- Real-world data



Outline

- Modeling and Learning Human Behavior in Games
  - Uncertainty and Bias Based Models
  - Quantal Response Based Models
- PAWS Application
- Other Models (Optional)
- Discussion (Optional)

Learning Objectives

- Write down the mathematical formulation of
  - Prospect Theory
  - Quantal Response
  - Subjective Utility Quantal Response
- Understand and describe the high-level idea of
  - Anchoring bias
  - Epsilon-bounded rationality
- For PAWS application, describe the target problem, method used, evaluation criteria

Modeling and Learning Human Behavior in Games

- Uncertainty and Bias Based Models
  - Prospect Theory [Kahneman and Tvesky, 1979]
  - Anchoring bias and epsilon-bounded rationality [Pita et al, 2010]
  - Attacker aims to reduce the defender's utility [Pita et al, 2012]
- Quantal Response Based Models
  - Quantal Response [McKelvey and Palfrey, 1995]
  - Subjective Utility Quantal Response [Nguyen et al, 2013]
- Other Models (optional)
  - Incorporating delayed observation [Fang et al, 2015]
  - Bounded rationality in repeated games [Kar et al, 2015]

PT: Prospect Theory

- Option 1: 20% chance to get \$500
- Option 2: 100% chance to get \$100

- Which one will you choose?

- Option 1: 20% chance to lose \$500
- Option 2: 100% chance to lose \$100

- Which one will you choose?

PT: Prospect Theory

- Model human decision making under uncertainty
- Maximize the 'prospect' [Kahneman and Tvesky, 1979]

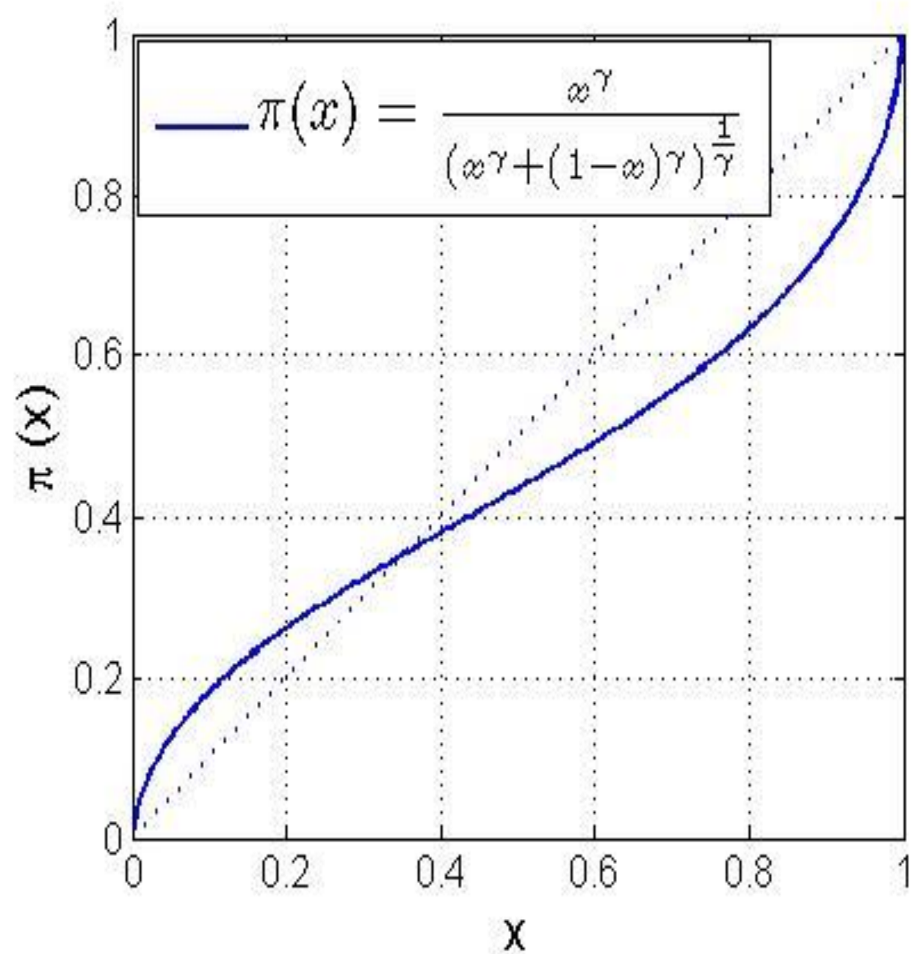
$$\text{prospect} = \sum_{i \in AllOutcomes} \pi(x_i) \cdot V(C_i)$$

  - π(·): weighting function
  - V(·): value function
- Defender: choose a strategy that maximizes DefEU when attacker best responds to the expected prospect (instead of AttEU)

Kahneman, D., & Tversky, A. (1979). Prospect theory: An analysis of decision under risk. Econometrica: Journal of the econometric society, 263-291.

PT: Prospect Theory

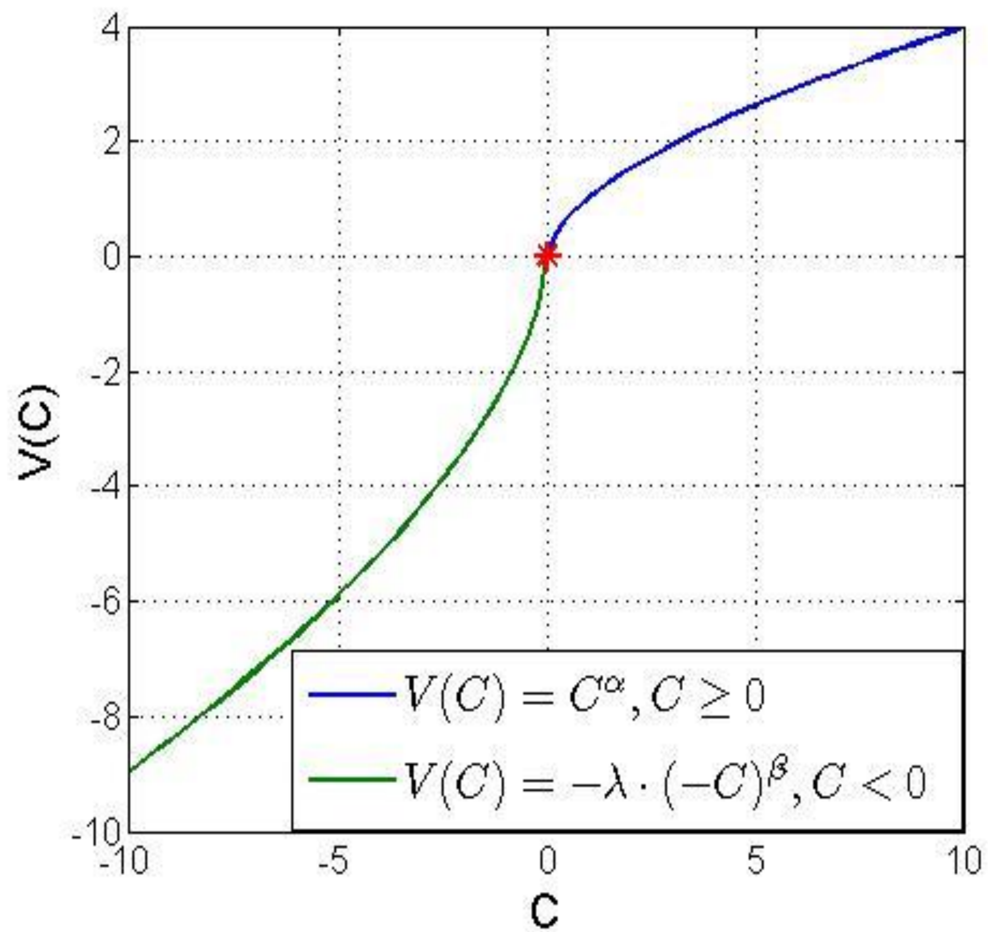
- Empirical Weighting Function
- Slope gets steeper as x gets closer to 0 and 1
- Not consistent with probability definition
  - $\pi(x)+\pi(1-x) < 1$
- Empirical value: $\gamma=0.64$ $(0<\gamma<1)$



Kahneman, D., & Tversky, A. (1979). Prospect theory: An analysis of decision under risk. Econometrica: Journal of the econometric society, 263-291.

PT: Prospect Theory

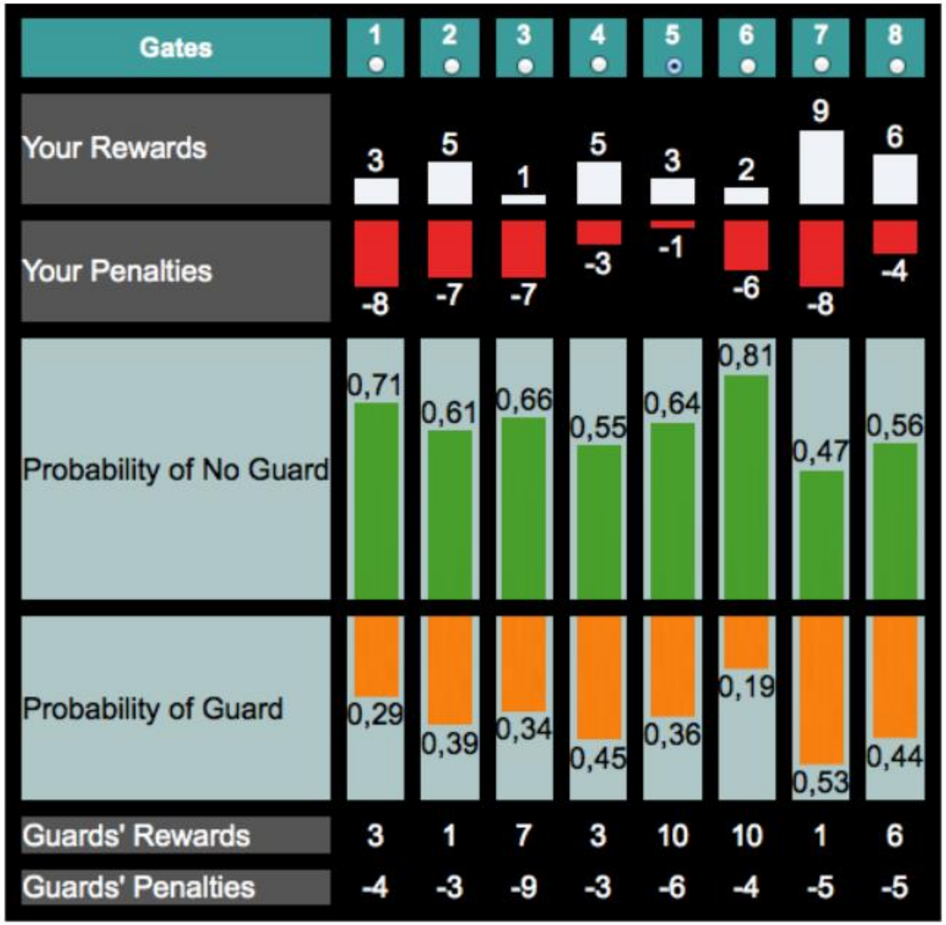
- Empirical Value Function
- Risk averse regarding gain
- Risk seeking regarding loss
- Empirical value: $\alpha=\beta=0.88, \lambda=2.25$



Kahneman, D., & Tversky, A. (1979). Prospect theory: An analysis of decision under risk. Econometrica: Journal of the econometric society, 263-291.

Human Subject Experiments

- Learn parameters from human subject experiments



Modeling and Learning Human Behavior in Games

- Uncertainty and Bias Based Models
  - Prospect Theory [Kahneman and Tvesky, 1979]
  - Anchoring bias and epsilon-bounded rationality [Pita et al, 2010]
  - Attacker aims to reduce the defender's utility [Pita et al, 2012]
- Quantal Response Based Models
  - Quantal Response [McKelvey and Palfrey, 1995]
  - Subjective Utility Quantal Response [Nguyen et al, 2013]
- Other Models (optional)
  - Incorporating delayed observation [Fang et al, 2015]
  - Bounded rationality in repeated games [Kar et al, 2015]

COBRA: Anchoring Bias and Epsilon-Bounded Rationality

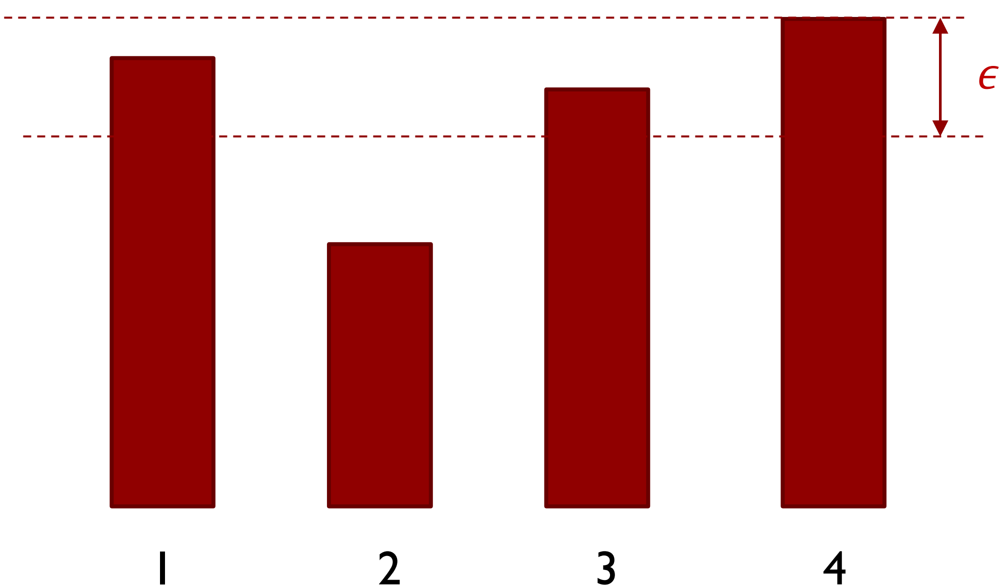
- Suppose you observe the defender's airport patrol strategy for 2 days, and find that the defender goes to terminal 1 in both days
- Which one of the following do you believe is closer to the actual strategy used by the defender?
  - (1,0)
  - (0.5,0.5)
  - (0.8,0.2)
- Anchoring bias: Full observation ($\alpha = 0$) vs no observation ($\alpha = 1$)

$$x' = (1 - \alpha)x + \frac{\alpha}{N}$$

Pita et al. Effective solutions for real-world stackelberg games: When agents must deal with human uncertainties. In AAMAS, 2009.

COBRA: Anchoring Bias and Epsilon-Bounded Rationality

- "epsilon optimality"
  - Any target whose expected utility is at least $AttEU^* - \epsilon$ may be attacked
  - Do not assume a specific target to be attacked



Pita et al. Effective solutions for real-world stackelberg games: When agents must deal with human uncertainties. In AAMAS, 2009.

COBRA: Anchoring Bias and Epsilon-Bounded Rationality

- Compute defender's strategy assuming anchoring bias and epsilon-bounded rationality

$$\max_{x,q,\gamma,a} \gamma$$

$$s.t.\, x' = (1-\alpha)x + \frac{\alpha}{N}$$

a is attacker's highest expected utility given x'

$$q_j = 1 \text{ if } \mathrm{AttEU}_j(x') \geq a - \epsilon$$

$$\gamma \leq \mathrm{DefEU}_j(\mathrm{x}) \text{ if } q_j = 1$$

Q: What values of α and ϵ will make it same as the basic Stackelberg Security Game setting?

- Human subject experiments: $\alpha = 0.37$ works best

Pita et al. Effective solutions for real-world stackelberg games: When agents must deal with human uncertainties. In AAMAS, 2009.

Modeling and Learning Human Behavior in Games

- Uncertainty and Bias Based Models
  - Prospect Theory [Kahneman and Tvesky, 1979]
  - Anchoring bias and epsilon-bounded rationality [Pita et al, 2010]
  - Attacker aims to reduce the defender's utility [Pita et al, 2012]
- Quantal Response Based Models
  - Quantal Response [McKelvey and Palfrey, 1995]
  - Subjective Utility Quantal Response [Nguyen et al, 2013]
- Other Models (optional)
  - Incorporating delayed observation [Fang et al, 2015]
  - Bounded rationality in repeated games [Kar et al, 2015]

MATCH: Attacker aims to reduce the defender's utility

- Attacker may deviate from the best response to reduce the defender's expected utility
- Choose a target to maximize

$$\frac{\text{Defender's utility loss due to deviation}}{\text{Adversary's utility loss due to deviation}}$$

- Defender: choose a strategy that maximize DefEU while bound the above value by β
- Experiments: $\beta = 1$

 Pita et al. A robust approach to addressing human adversaries in security games. In ECAI, 2012

Modeling and Learning Human Behavior in Games

- Uncertainty and Bias Based Models
  - Prospect Theory [Kahneman and Tvesky, 1979]
  - Anchoring bias and epsilon-bounded rationality [Pita et al, 2010]
  - Attacker aims to reduce the defender's utility [Pita et al, 2012]
- Quantal Response Based Models
  - Quantal Response [McKelvey and Palfrey, 1995]
  - Subjective Utility Quantal Response [Nguyen et al, 2013]
- Other Models (optional)
  - Incorporating delayed observation [Fang et al, 2015]
  - Bounded rationality in repeated games [Kar et al, 2015]

QR: Quantal Response Model

- Error in individual's response
  - Still: more likely to select better choices than worse choices
- Probability distribution of different responses
- Quantal best response:

$$q_j = \frac{e^{\lambda * \mathrm{AttEU}_j(x)}}{\sum_i e^{\lambda * \mathrm{AttEU}_i(x)}}$$

- λ: represents error level (=0 means uniform random)
  - Maximal likelihood estimation (λ=0.76)

McKelvey, R. D., & Palfrey, T. R. (1995). Quantal response equilibria for normal form games. Games and economic behavior, 10(1), 6-38.

Poll 1: Quantal Response Model

- If there are two choices (actions), what is the probability of choosing the first action if the player follows quantal response model with $\lambda = 0$?
  - A: 1
  - B: 0
  - C: $\frac{1}{2}$
  - D: $\frac{1}{e} \approx 0.368$
  - E: None of the above
  - F: I don't know

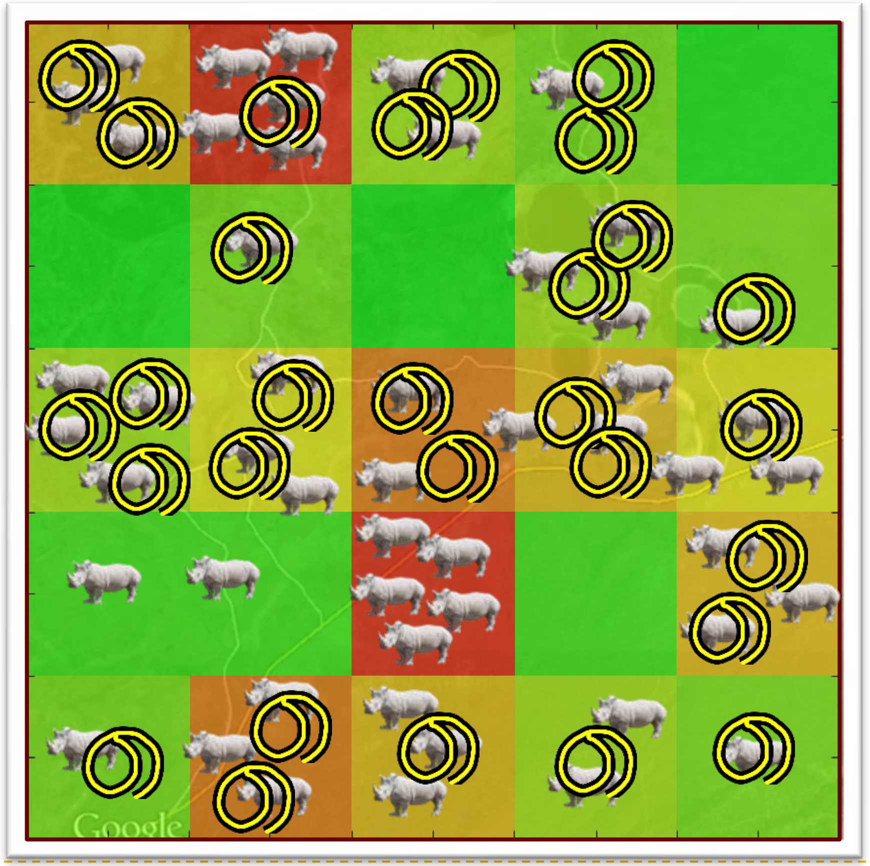
$$q_j = \frac{e^{\lambda * \mathrm{AttEU}_j(x)}}{\sum_i e^{\lambda * \mathrm{AttEU}_i(x)}}$$

Modeling and Learning Human Behavior in Games

- Uncertainty and Bias Based Models
  - Prospect Theory [Kahneman and Tvesky, 1979]
  - Anchoring bias and epsilon-bounded rationality [Pita et al, 2010]
  - Attacker aims to reduce the defender's utility [Pita et al, 2012]
- Quantal Response Based Models
  - Quantal Response [McKelvey and Palfrey, 1995]
  - Subjective Utility Quantal Response [Nguyen et al, 2013]
- Other Models (optional)
  - Incorporating delayed observation [Fang et al, 2015]
  - Bounded rationality in repeated games [Kar et al, 2015]

SUQR: Subjective Utility Quantal Response Model

- $\mathrm{SEU}_j = \sum_k w_k \times f_j^k,\ \ q_j = \frac{e^{\lambda * \mathrm{SEU}_j(x)}}{\sum_i e^{\lambda * \mathrm{SEU}_i(x)}}$



Coverage Probability
+ Reward/Penalty

Attack Probability

Nguyen, T. H., Yang, R., Azaria, A., Kraus, S., & Tambe, M. Analyzing the Effectiveness of Adversary Modeling in Security Games. In AAAI, 2013.

SUQR: Subjective Utility Quantal Response Model

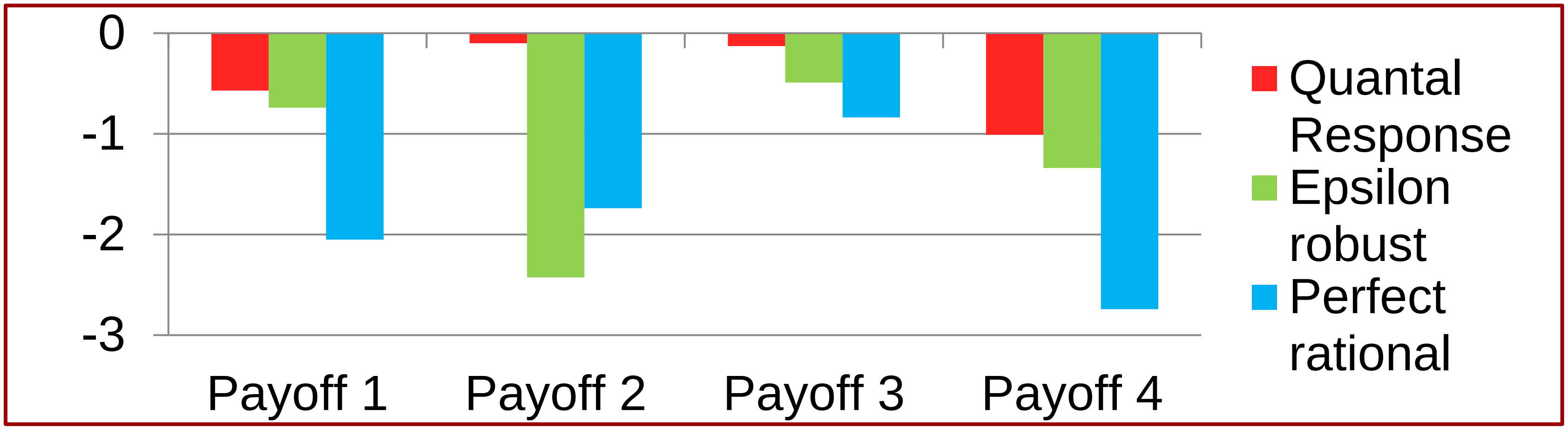
- Compute the optimal defender strategy

$$\max_x \sum_{t=1}^{T} \frac{e^{\lambda(w_1 x_t + w_2 R_t^a + w_3 P_t^a)}}{\sum_{t'} e^{\lambda(w_1 x_{t'} + w_2 R_{t'}^a + w_3 P_{t'}^a)}} (x_t R_t^d + (1 - x_t) P_t^d)$$

$$\text{s.t.} \sum_{t=1}^{T} x_t \leq K, 0 \leq x_t \leq 1 \qquad (3)$$

Comparison of Model Performance

- Prospect Theory < DOBSS < COBRA < Quantal Response < MATCH < SUQR



MATCH wins	Draw	QR wins
42	52	6

MATCH wins	Draw	SUQR wins
1	8	13

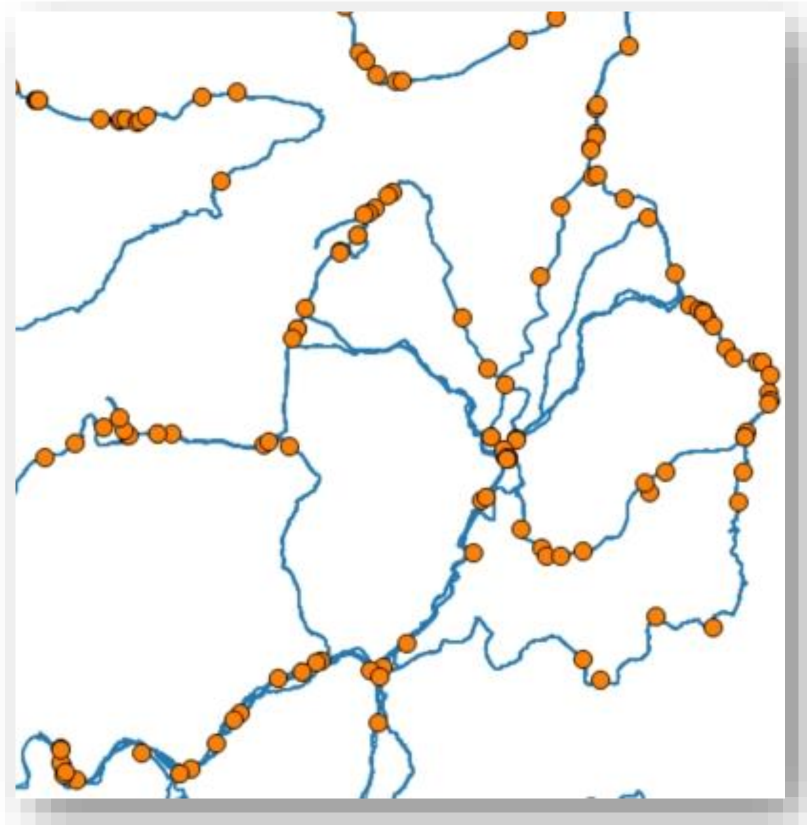
Nguyen, T. H., Yang, R., Azaria, A., Kraus, S., & Tambe, M. Analyzing the Effectiveness of Adversary Modeling in Security Games. In AAAI, 2013.

Outline

- Modeling and Learning Human Behavior in Games
  - Uncertainty and Bias Based Models
  - Quantal Response Based Models
- PAWS Application
- Other Models (Optional)
- Discussion (Optional)

LEARN POACHERS' BEHAVIOR MODEL

- Use SUQR with parameters learned from human subject experiments
- Q: Can we use data from previous patrols?



GAME-THEORETIC REASONING

GAME-THEORETIC PATROL STRATEGY DESIGN

- Challenge for PAWS: Payoff uncertainty
- ARROW algorithm (Nguyen et al. 15)
  - Behavioral minimax regret

Payoff uncertainty
Poacher behavior model

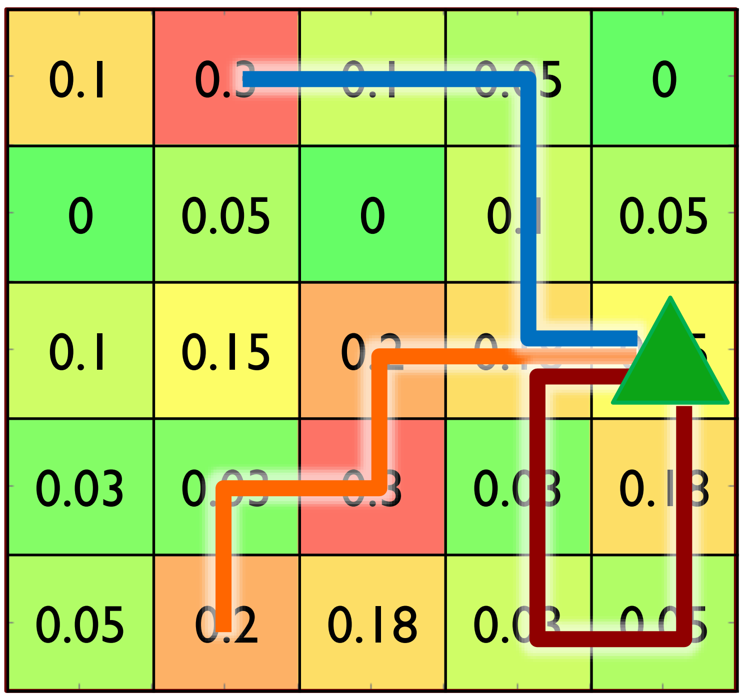
ARROW

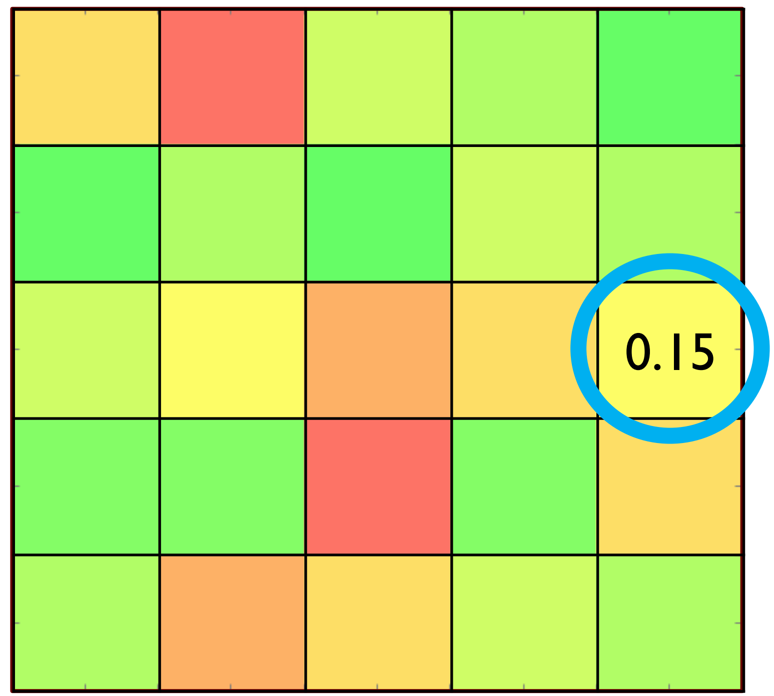
Coverage probability

0.1	0.3	0.1	0.05	0
0	0.05	0	0.1	0.05
0.1	0.15	0.2	0.18	0.15
0.03	0.03	0.3	0.03	0.18
0.05	0.2	0.18	0.03	0.05

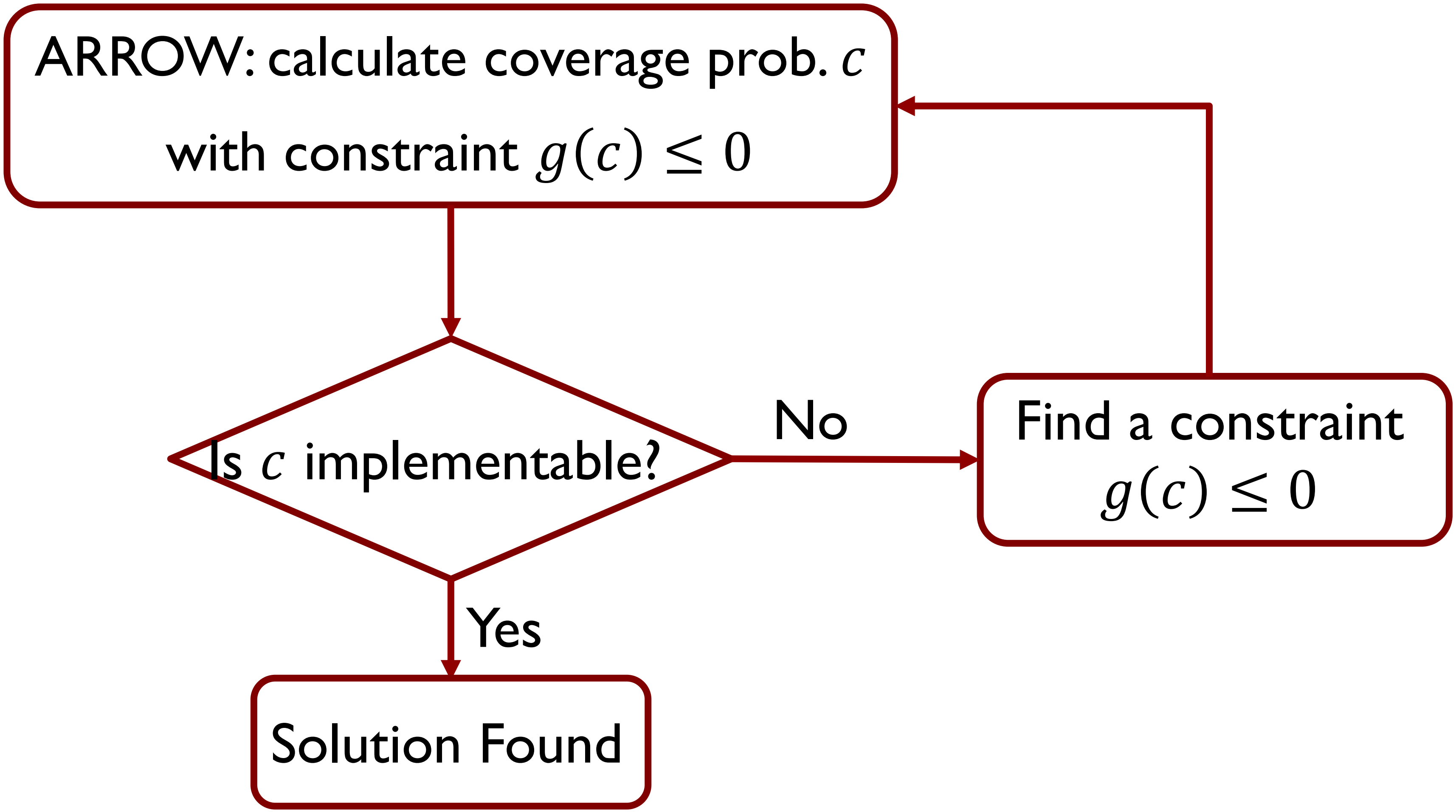
ROUTE PLANNING

- Coverage probability → route to take
- First challenge: Impossible to implement coverage





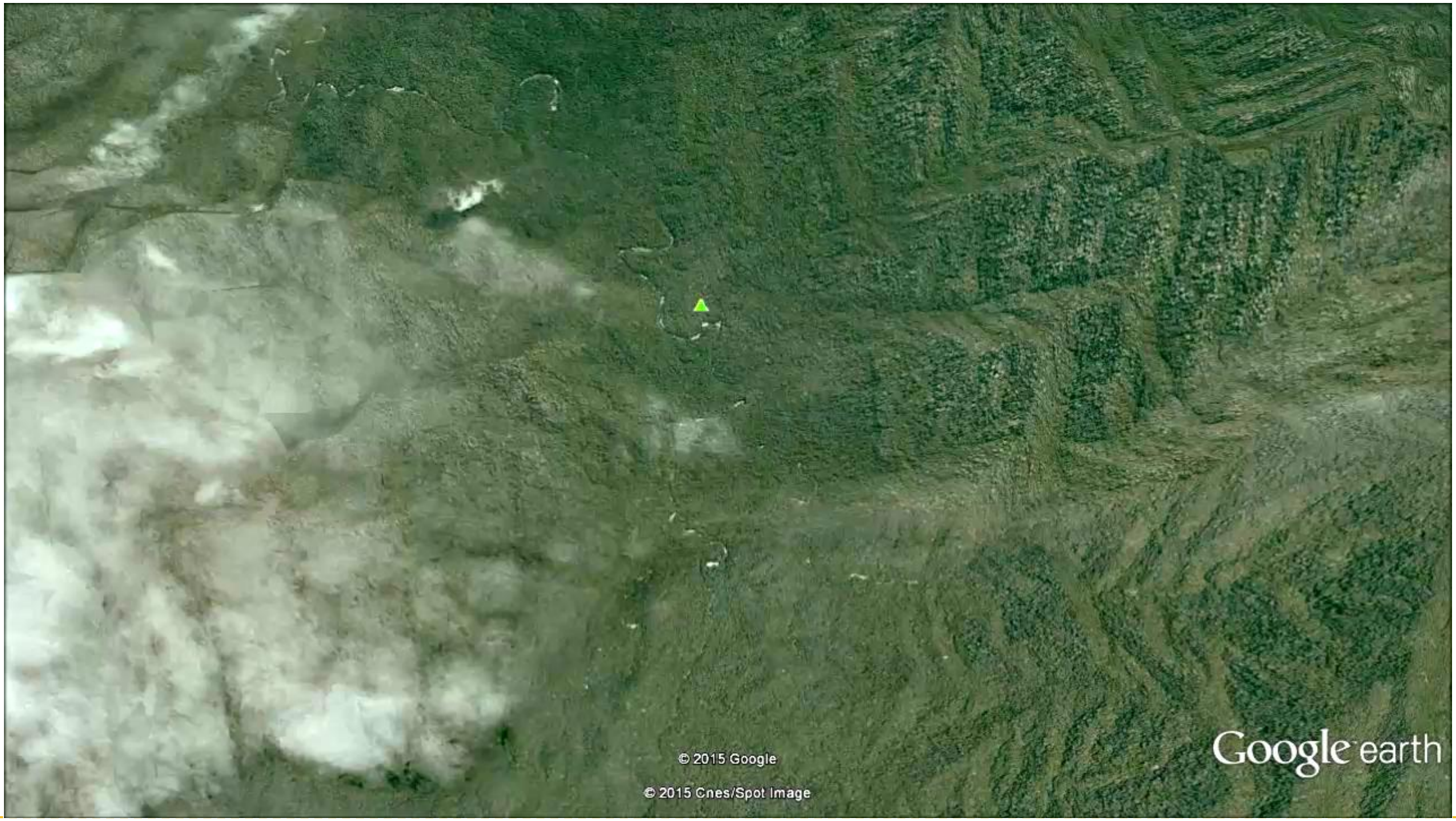
MODIFIED ARROW + BLADE



 Rong Yang, Albert Xin Jiang, Milind Tambe, Fernando Ordonez. *Scaling-up Security Games with Boundedly Rational Adversaries: A Cutting-plane Approach.* IJCAI'13

ROUTE PLANNING

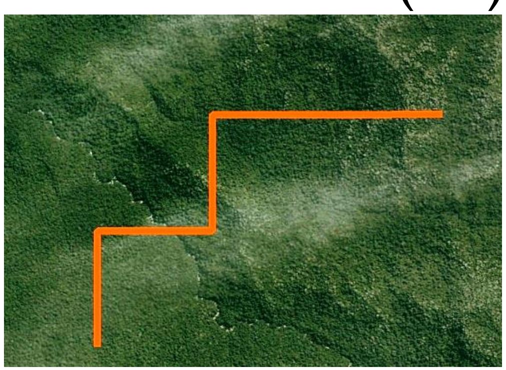
- Coverage probability $c \rightarrow$ route to take
- Second challenge: Route not compatible with terrain



COMPLEX TOPOGRAPHICAL INFORMATION

0.1	0.3	0.1	0.05	0
0	0.05	0	0.1	0.05
0.1	0.15	0.2	0.18	0.15
0.03	0.03	0.3	0.03	0.18
0.05	0.2	0.18	0.03	0.05

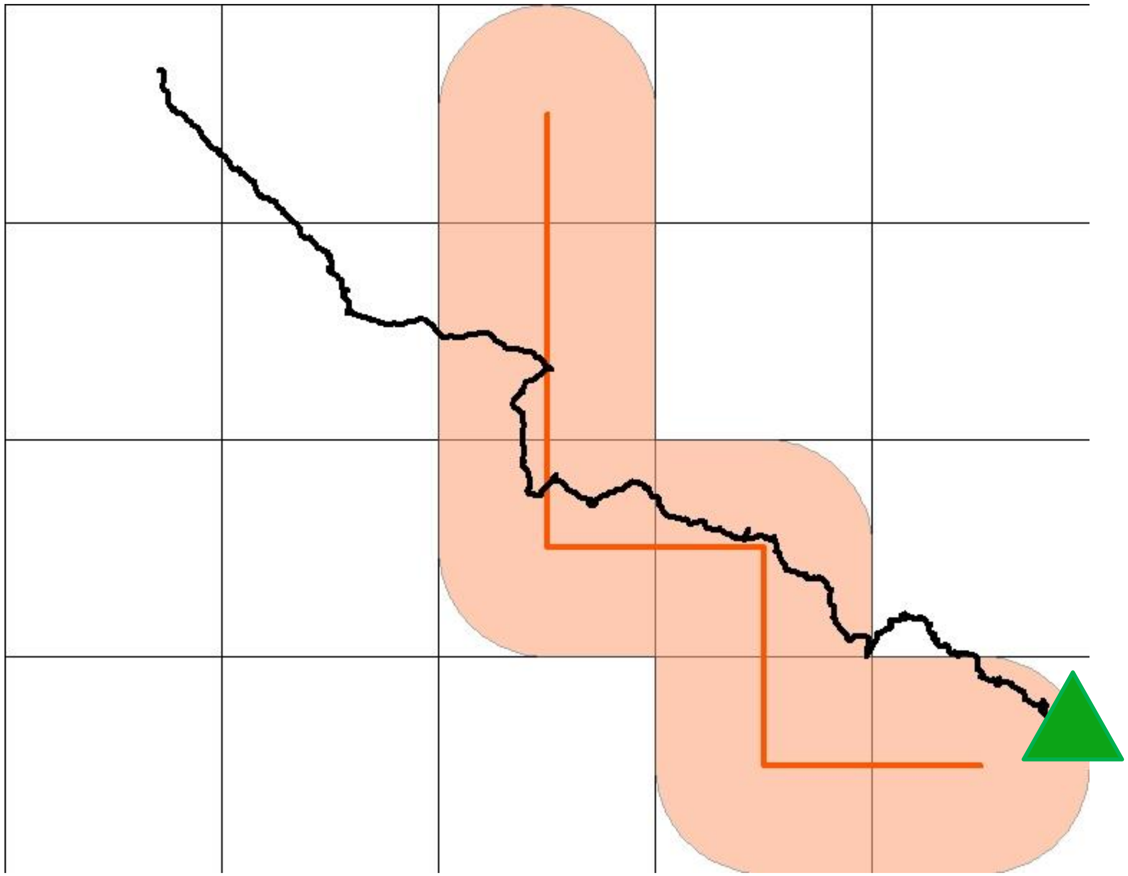
Patrol Route (2D)



Patrol Route (3D)

FIRST TESTS

- Test in Malaysia

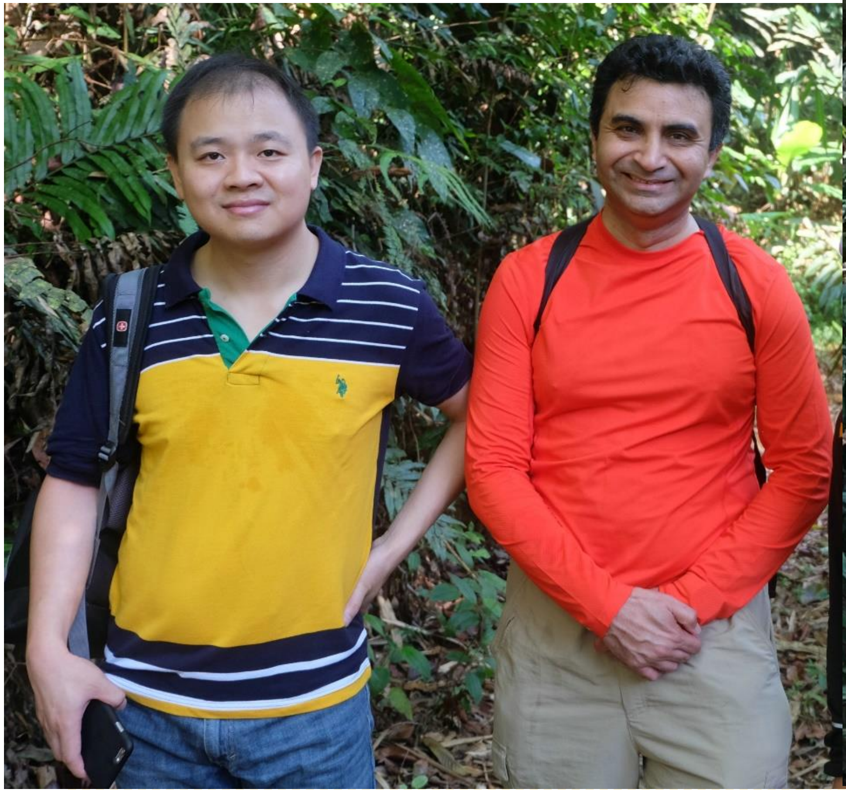


FIRST TESTS

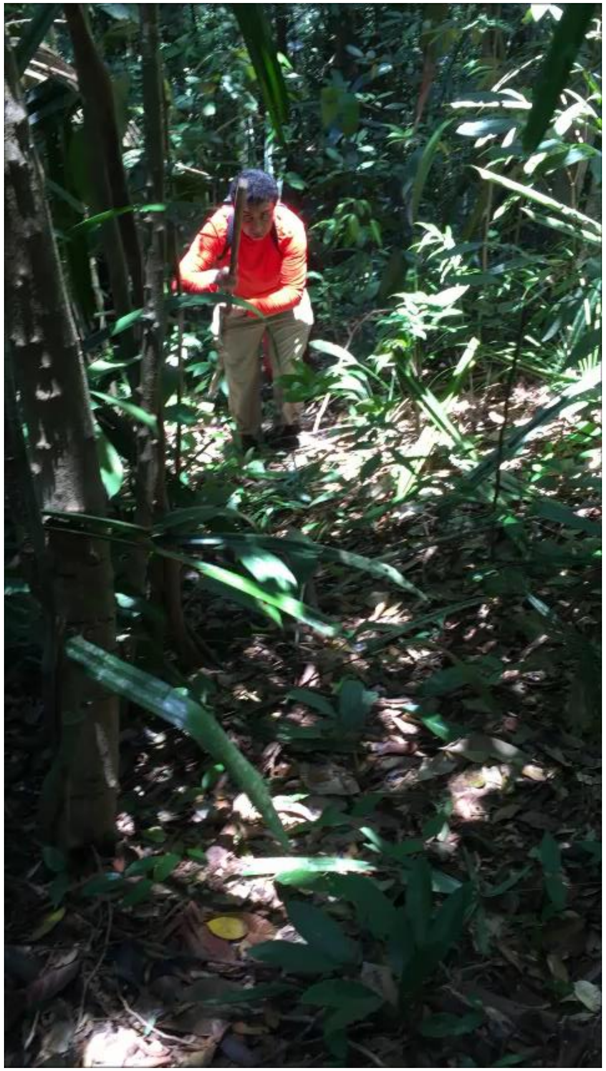
- Test in Uganda

TRIAL PATROL IN THE FIELD

- 8-hour patrol in April 2015: patrolling is not easy!

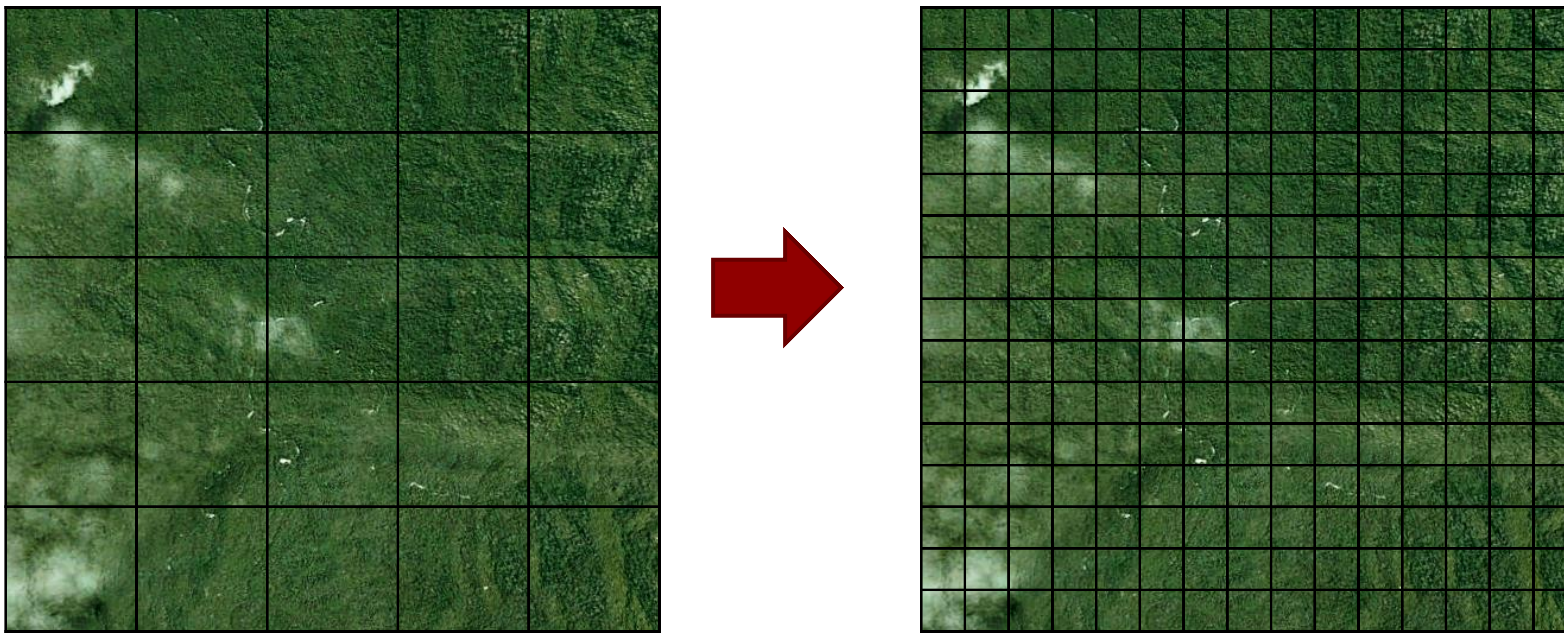



TRIAL PATROL IN THE FIELD



COMPLEX TOPOGRAPHICAL INFORMATION

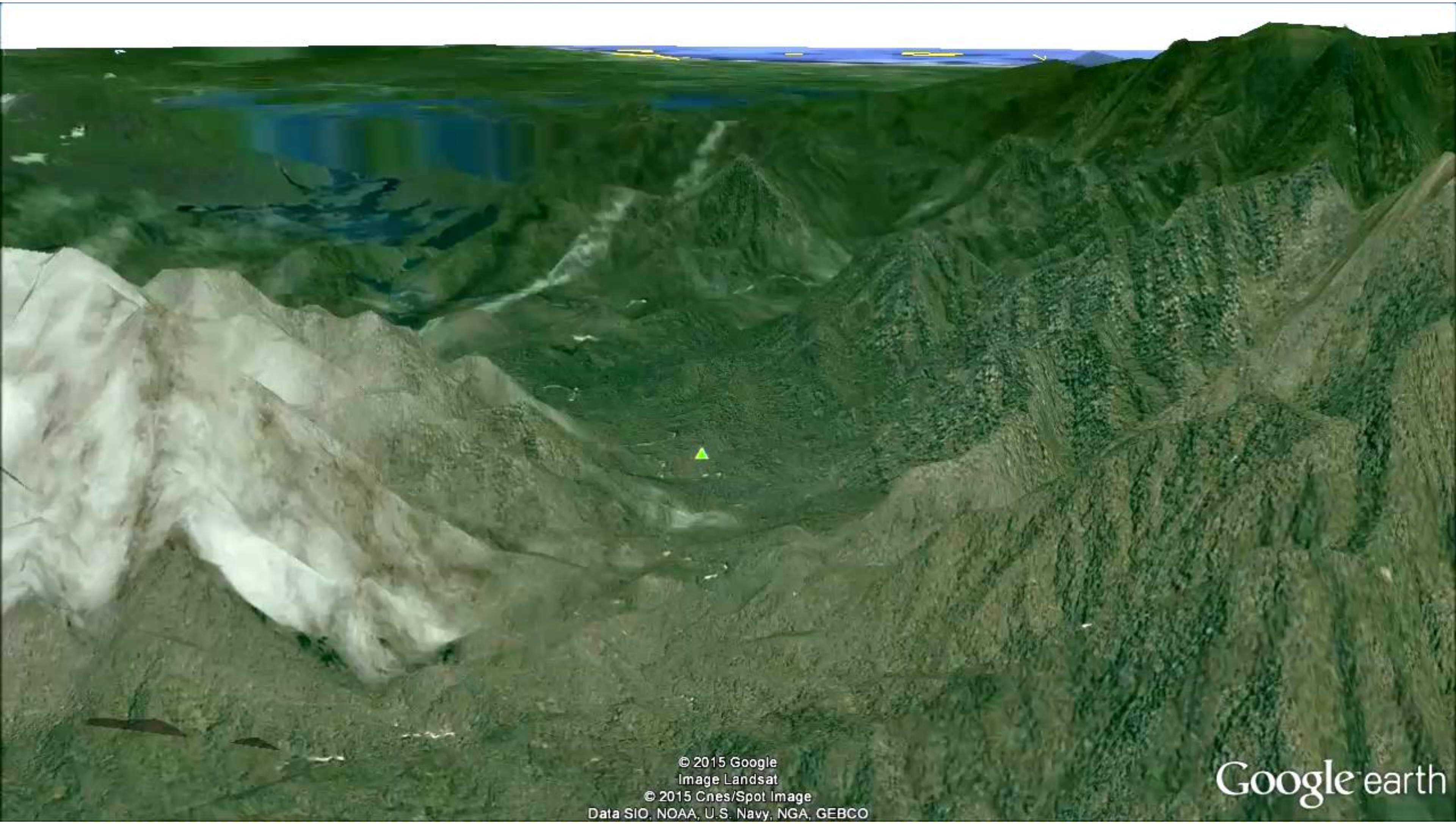
- Fine discretization → huge number of patrol routes
- Novel solution:
  - Focus on terrain features
  - Hierarchical modeling → virtual street map



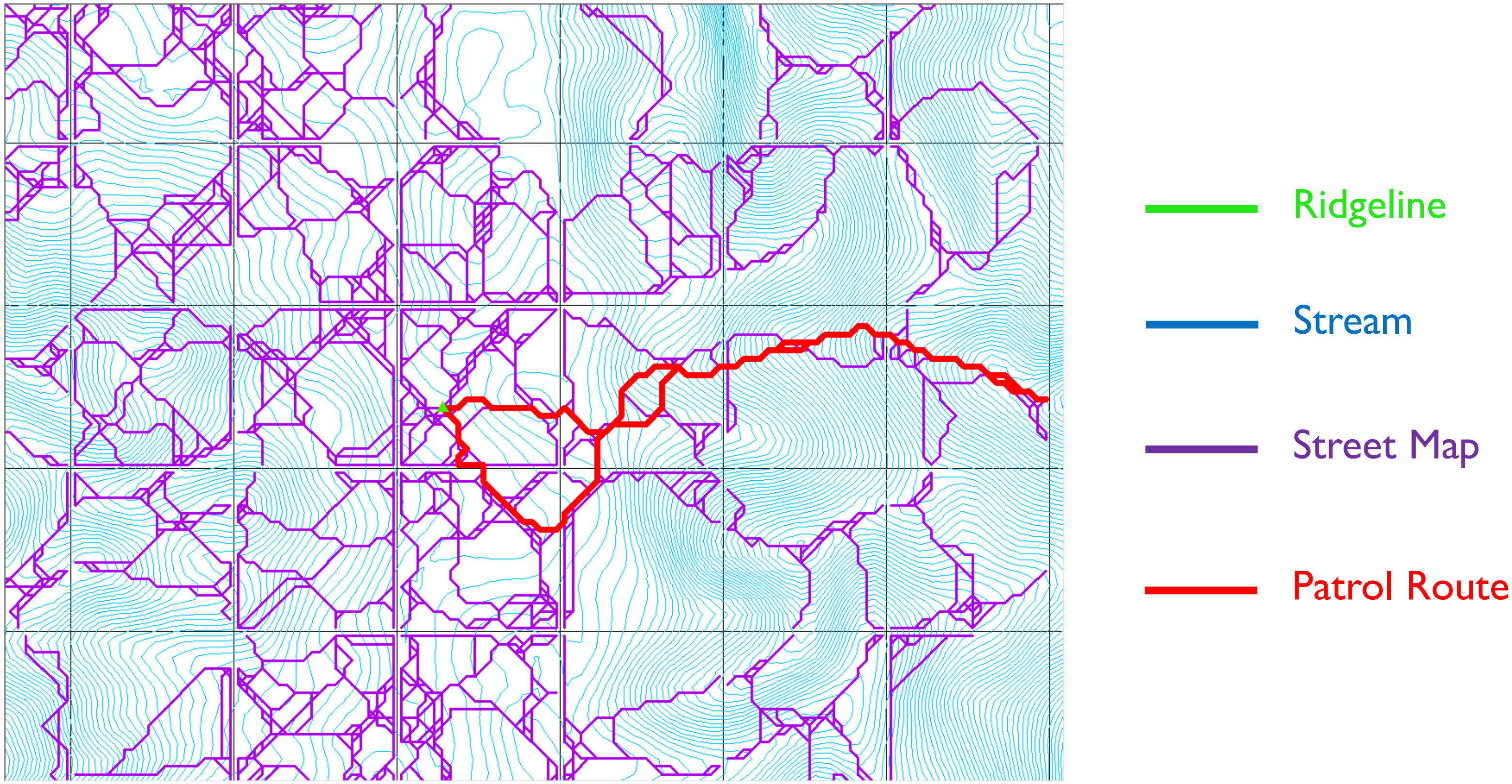
COMPLEX TOPOGRAPHICAL INFORMATION

- Terrain feature, e.g., ridgeline

ROUTE PLANNING



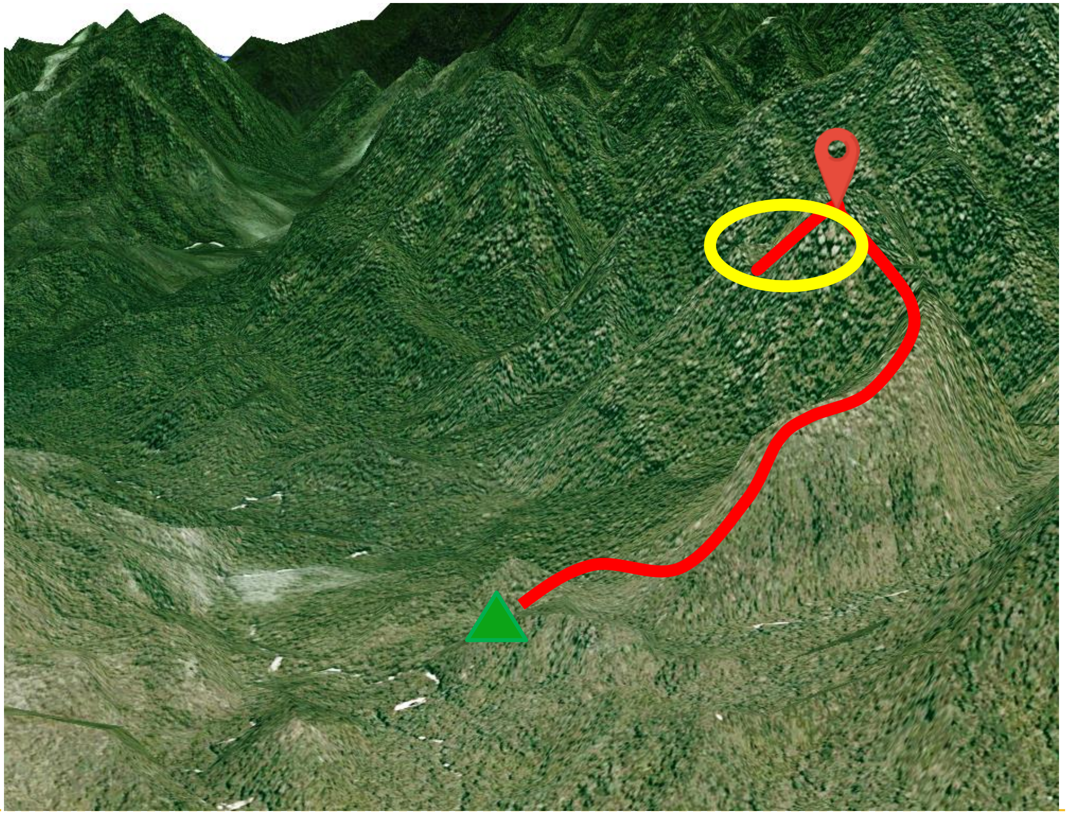
HIERARCHICAL MODEL



- Attacker action: choose a grid cell to place snares
- Defender action: choose a path on the street map

BEFORE REAL-WORLD DEPLOYMENT

- Practical constraints (1)
  - Short downhill followed by returning uphill is annoying

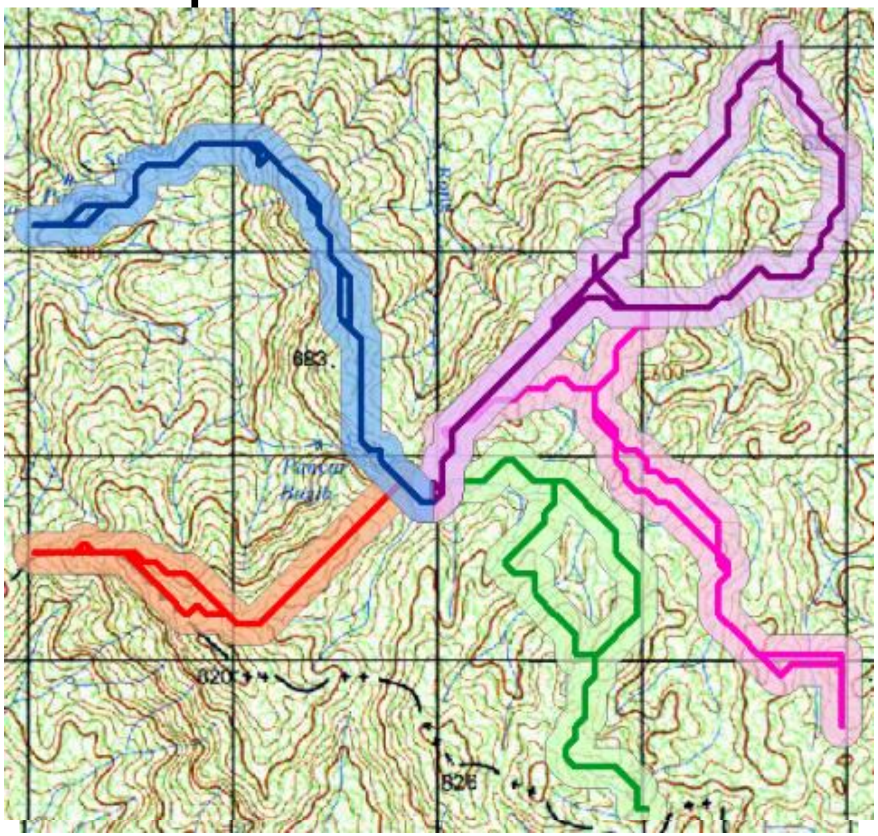


BEFORE REAL-WORLD DEPLOYMENT

- Practical constraints (II)
  - Patrol time = 5 hours = walking time + recording time

EXAMPLE OUTPUT OF PAWS

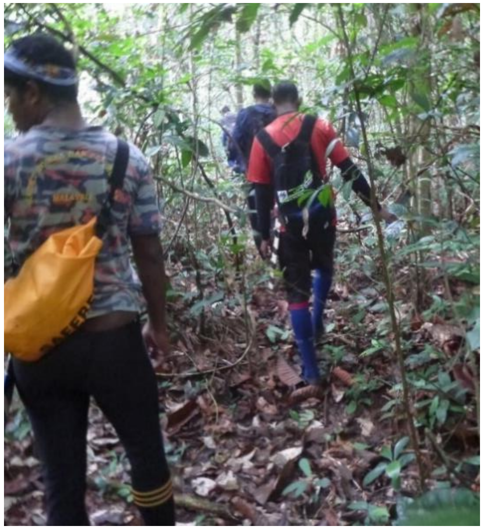
- 1 day patrol starting from a base camp
- Sample one route according to the probability every



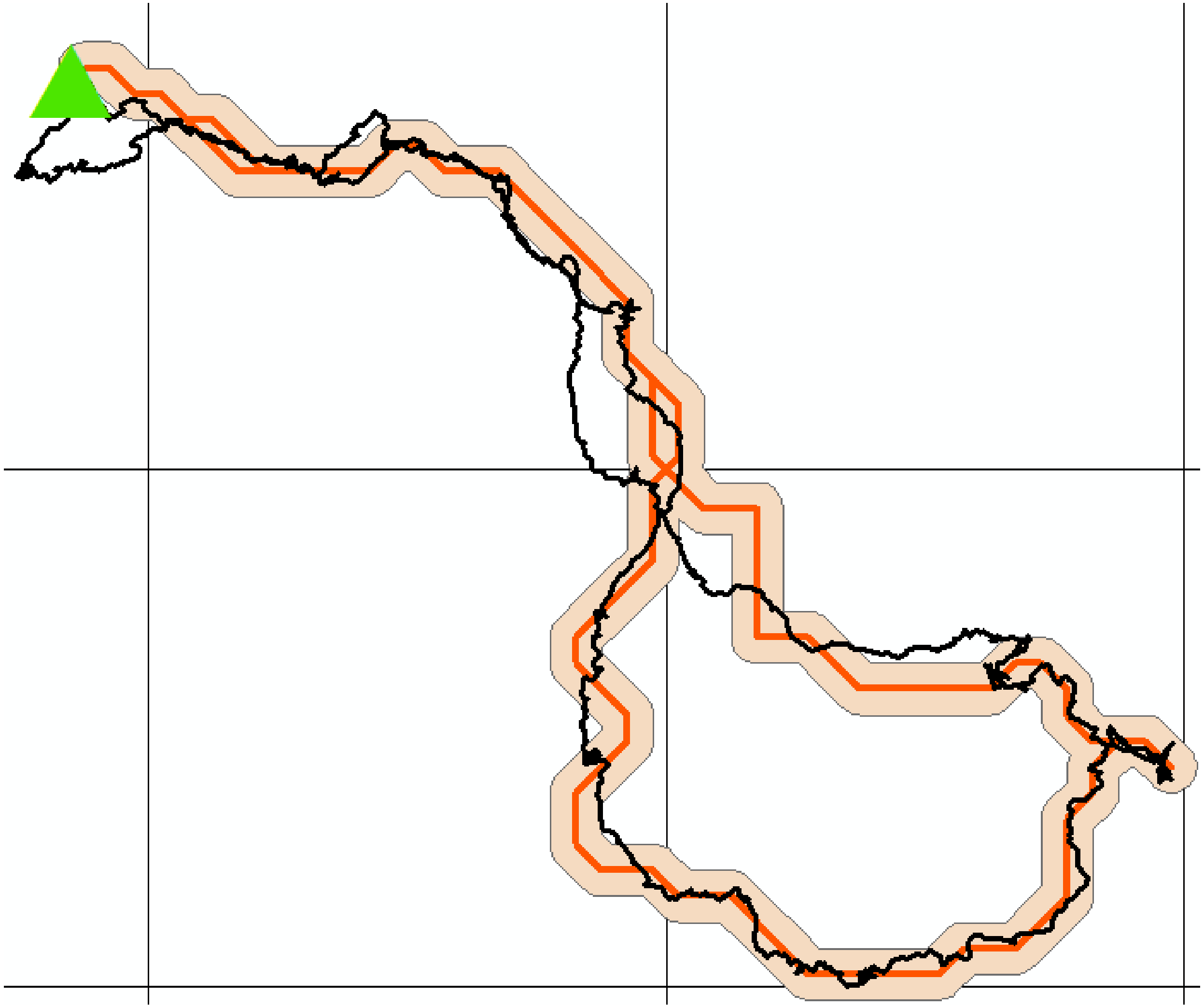
suggestedRoute	Prob=0.58
suggestedRoute	Prob=0.16
suggestedRoute	Prob=0.12
suggestedRoute	Prob=0.08
suggestedRoute	Prob=0.06

PAWS PATROLS IN THE FIELD

Basic Information of PAWS Patrols	
Average Trip Length	4.67 Days
Average Number of Patrollers	5
Average Patrol Time Per Day	4.48 hours
Average Patrol Distance Per Day	9.29 km



PAWS PATROLS IN THE FIELD



PAWS PATROLS IN THE FIELD

Animal Footprint

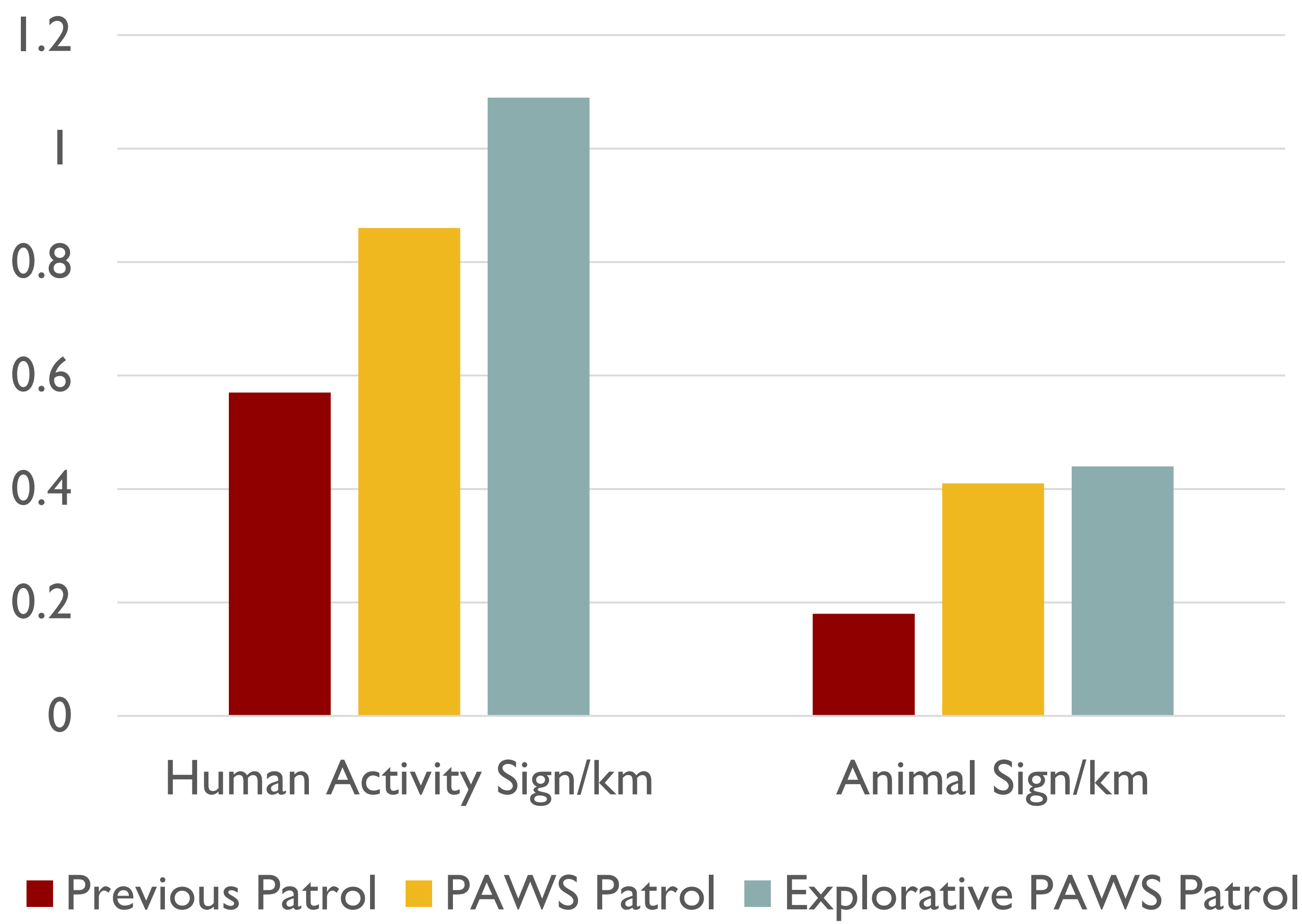
Tree Mark

Camping Sign

Tiger Sign

Lighter

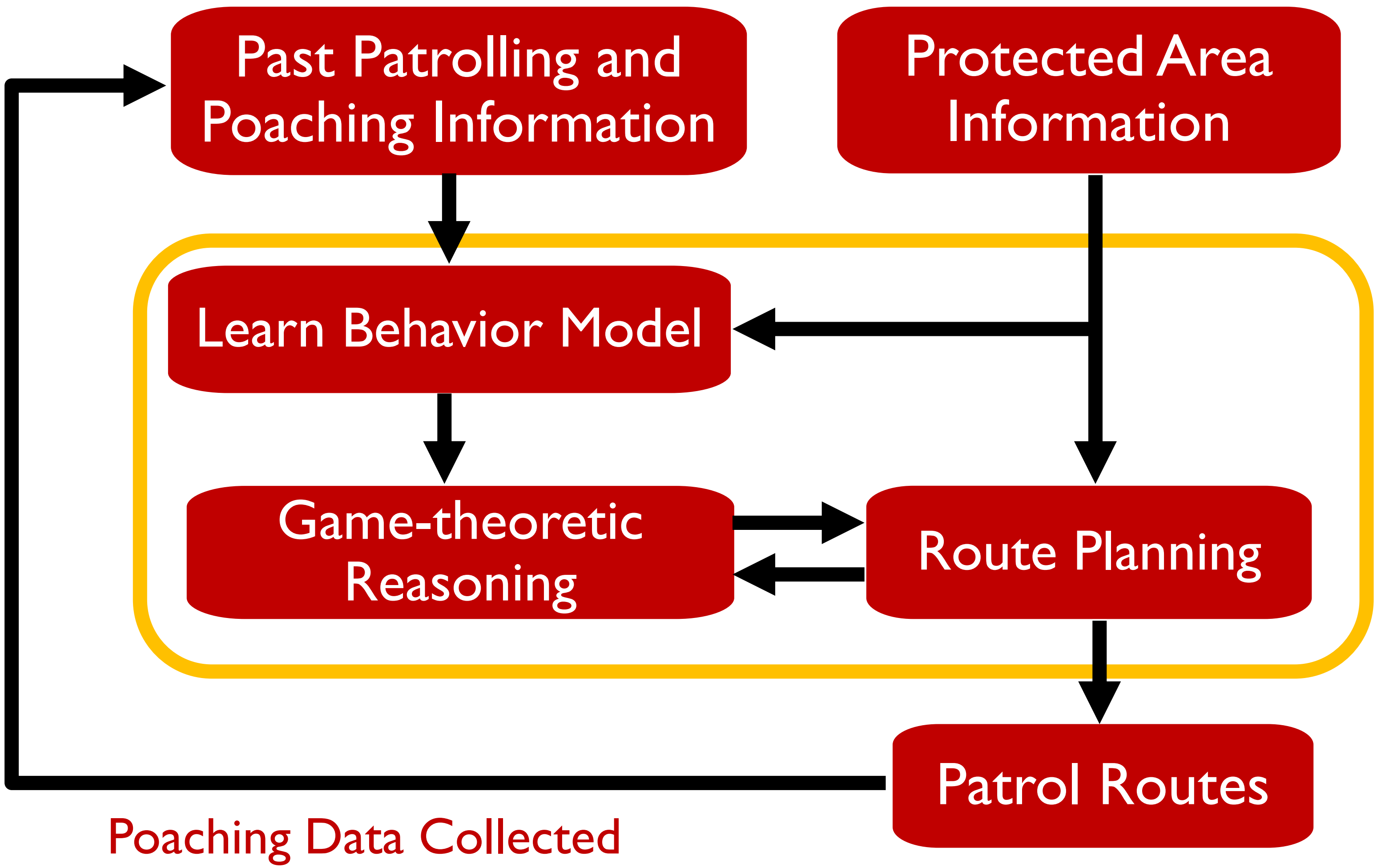
PAWS PATROLS IN THE FIELD



FUTURE DEPLOYMENT

- Queen Elizabeth National Park in Uganda
- Tested in Spring 2014
- PAWS with CAPTURE tool: Deploy later this year

PAWS SUMMARY



Outline

- Modeling and Learning Human Behavior in Games
  - Uncertainty and Bias Based Models
  - Quantal Response Based Models
- PAWS Application
- Other Models (Optional)
- Discussion (Optional)

Modeling and Learning Human Behavior in Games

- Uncertainty and Bias Based Models
  - Prospect Theory [Kahneman and Tvesky, 1979]
  - Anchoring bias and epsilon-bounded rationality [Pita et al, 2010]
  - Attacker aims to reduce the defender's utility [Pita et al, 2012]
- Quantal Response Based Models
  - Quantal Response [McKelvey and Palfrey, 1995]
  - Subjective Utility Quantal Response [Nguyen et al, 2013]
- Other Models
  - Incorporating delayed observation [Fang et al, 2015]
  - Bounded rationality in repeated games [Kar et al, 2015]

GSG: Incorporating Delayed Observation

Wildlife

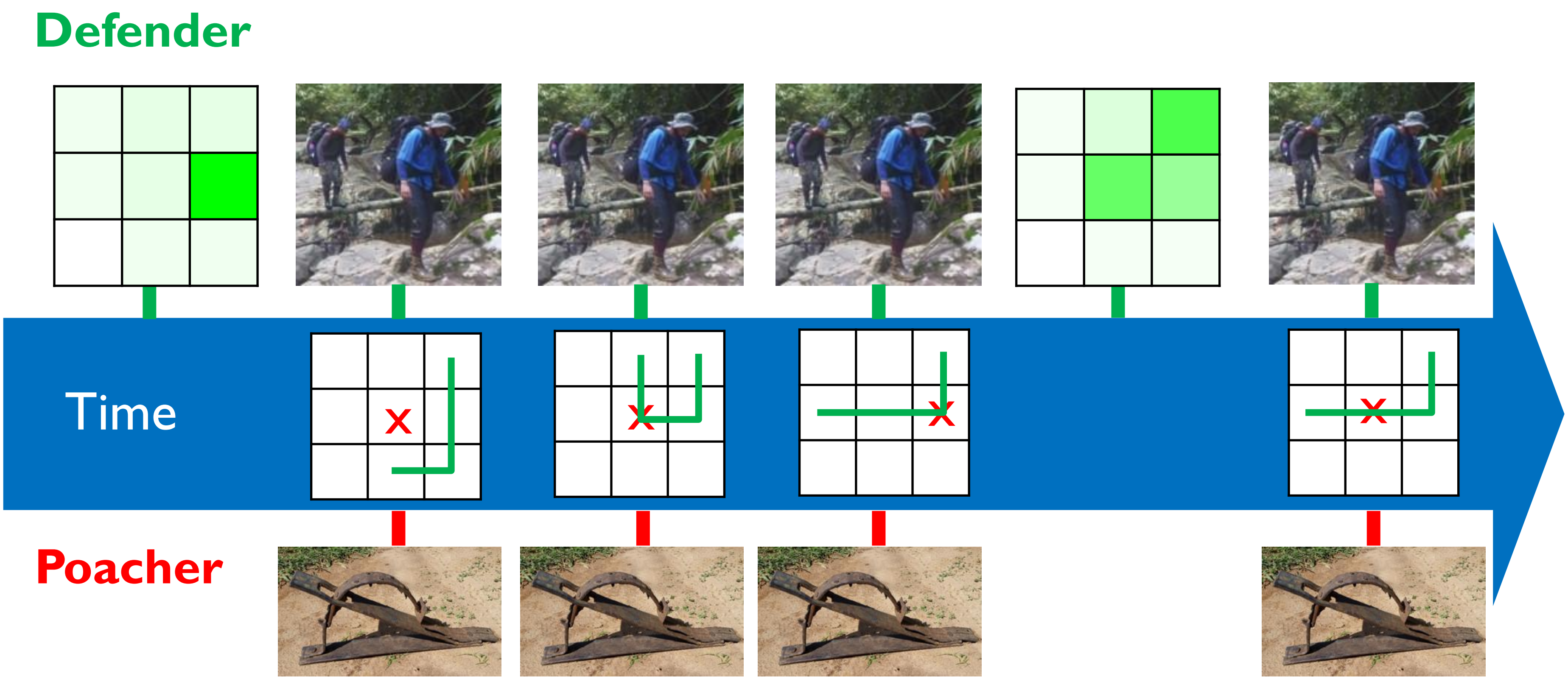
Forest

Fishery

- Frequent and repeated attacks
  - Not one-shot / More data
- Attacker decision making
  - Limited surveillance / Less effort / Boundedly rational
- New model: Green Security Games

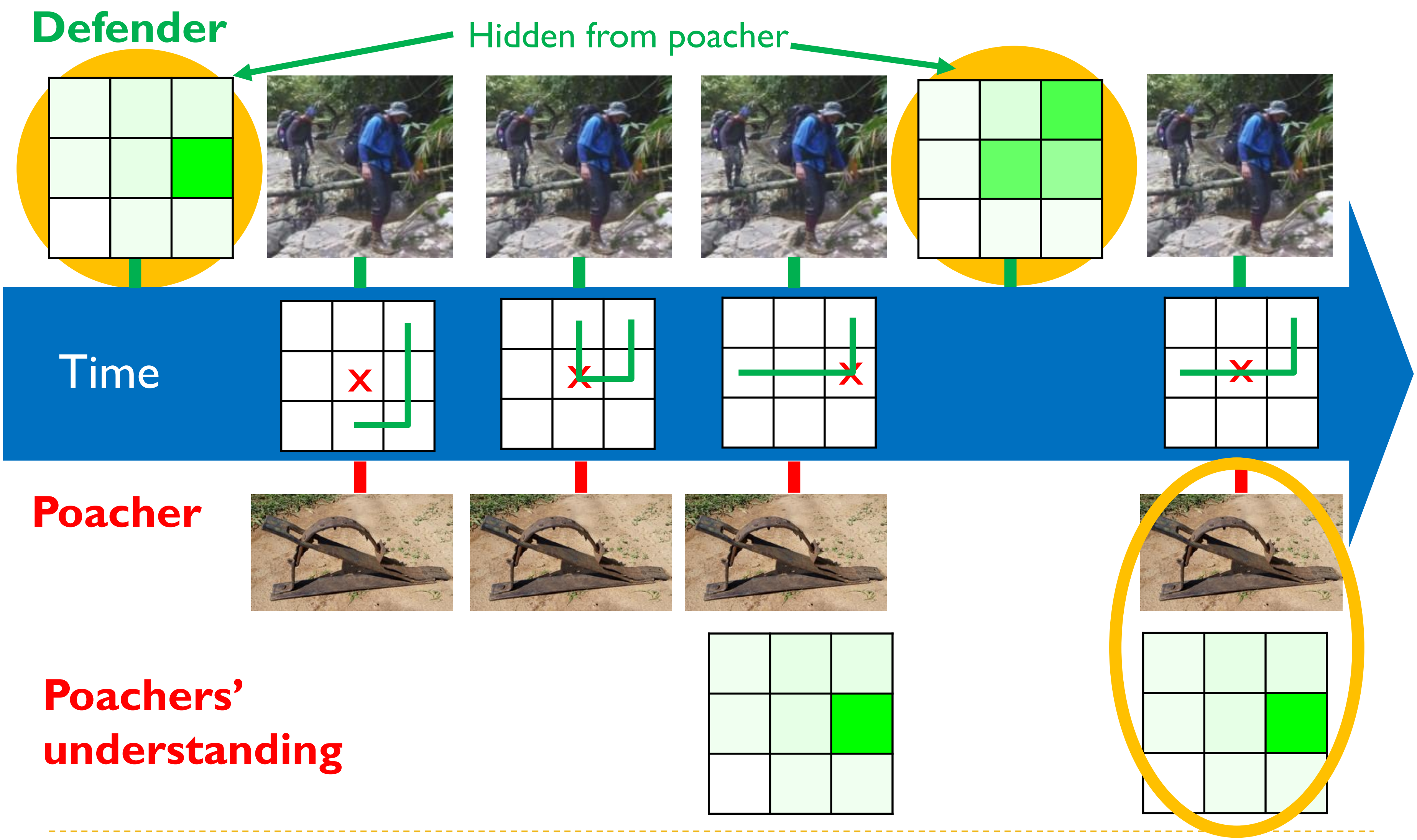
Fang, F., Stone, P., & Tambe, M. When Security Games Go Green: Designing Defender Strategies to Prevent Poaching and Illegal Fishing. In IJCAI, 2015.

GSG: Incorporating Delayed Observation



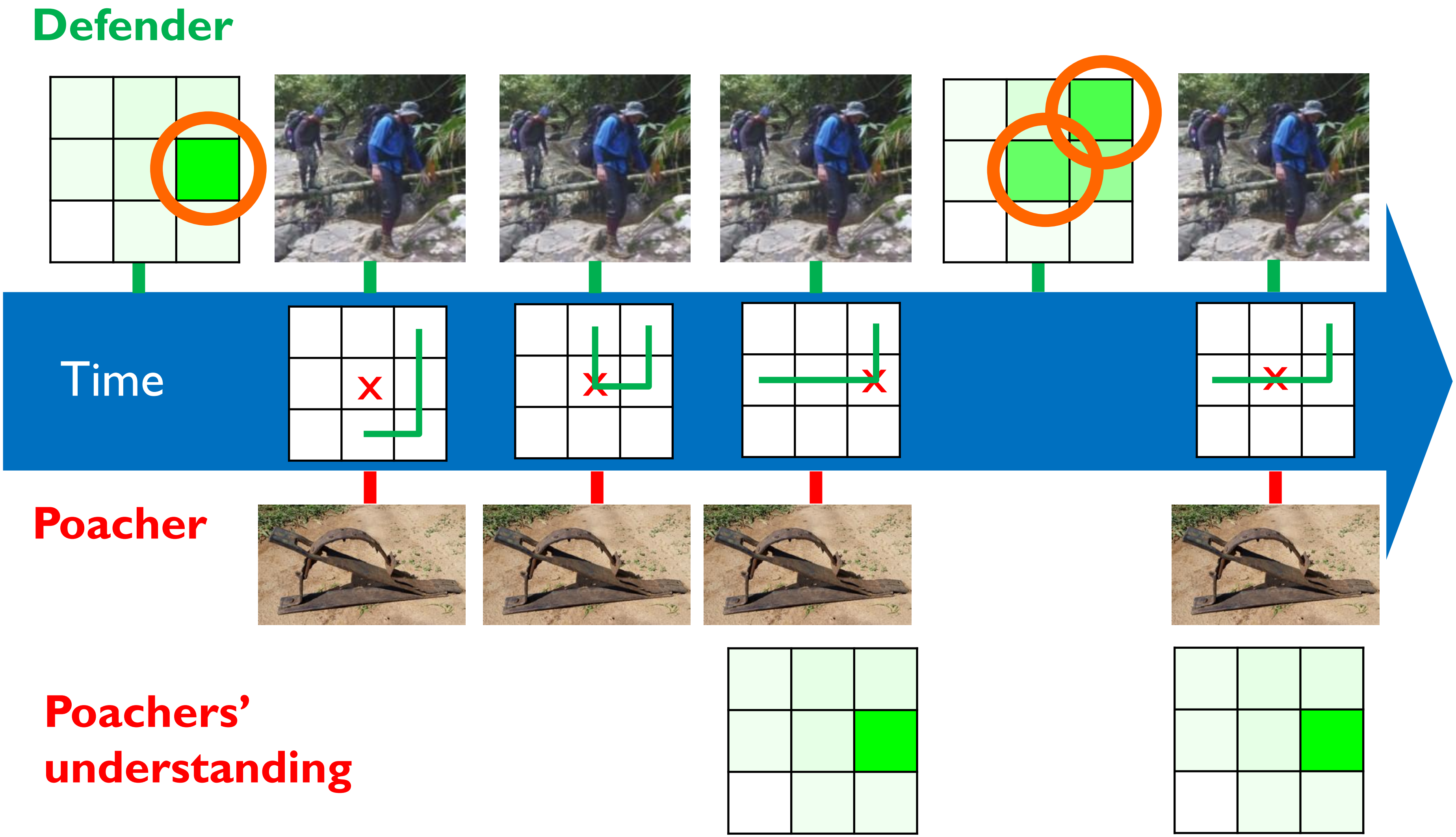
Fang, F., Stone, P., & Tambe, M. When Security Games Go Green: Designing Defender Strategies to Prevent Poaching and Illegal Fishing. In IJCAI, 2015.

GSG: Incorporating Delayed Observation



Fang, F., Stone, P., & Tambe, M. When Security Games Go Green: Designing Defender Strategies to Prevent Poaching and Illegal Fishing. In IJCAI, 2015.

GSG: Incorporating Delayed Observation



Fang, F., Stone, P., & Tambe, M. When Security Games Go Green: Designing Defender Strategies to Prevent Poaching and Illegal Fishing. In IJCAI, 2015.

GSG: Incorporating Delayed Observation

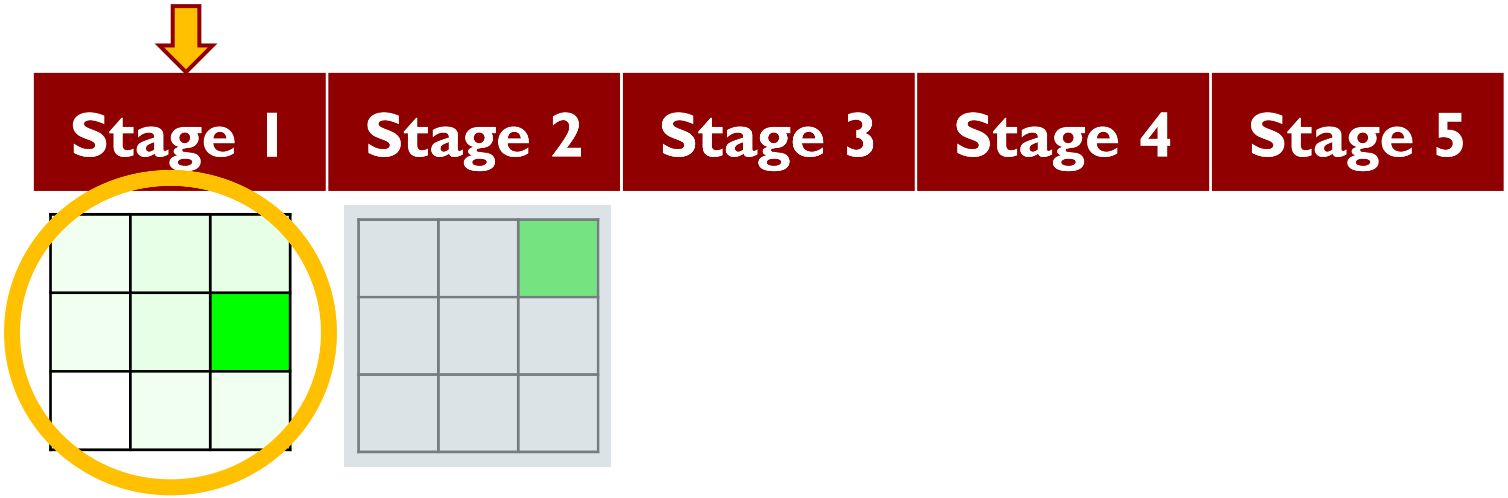
- A Green Security Game (GSG) is a T stage game where the defender protects N targets against L attackers. Defender chooses a mixed strategy c^t in stage t.

- A GSG attacker is characterized by his memory length Γ, coefficients $\alpha_0, \dots, \alpha_\Gamma$ and SUQR model parameter ω. In stage t, he responds to a convex combination of defender strategy in recent $\Gamma + 1$ rounds: $\eta_t = \sum_{\tau=0}^{\Gamma} \alpha_\tau c^{t-\tau}$

Fang, F., Stone, P., & Tambe, M. When Security Games Go Green: Designing Defender Strategies to Prevent Poaching and Illegal Fishing. In IJCAI, 2015.

GSG: Incorporating Delayed Observation

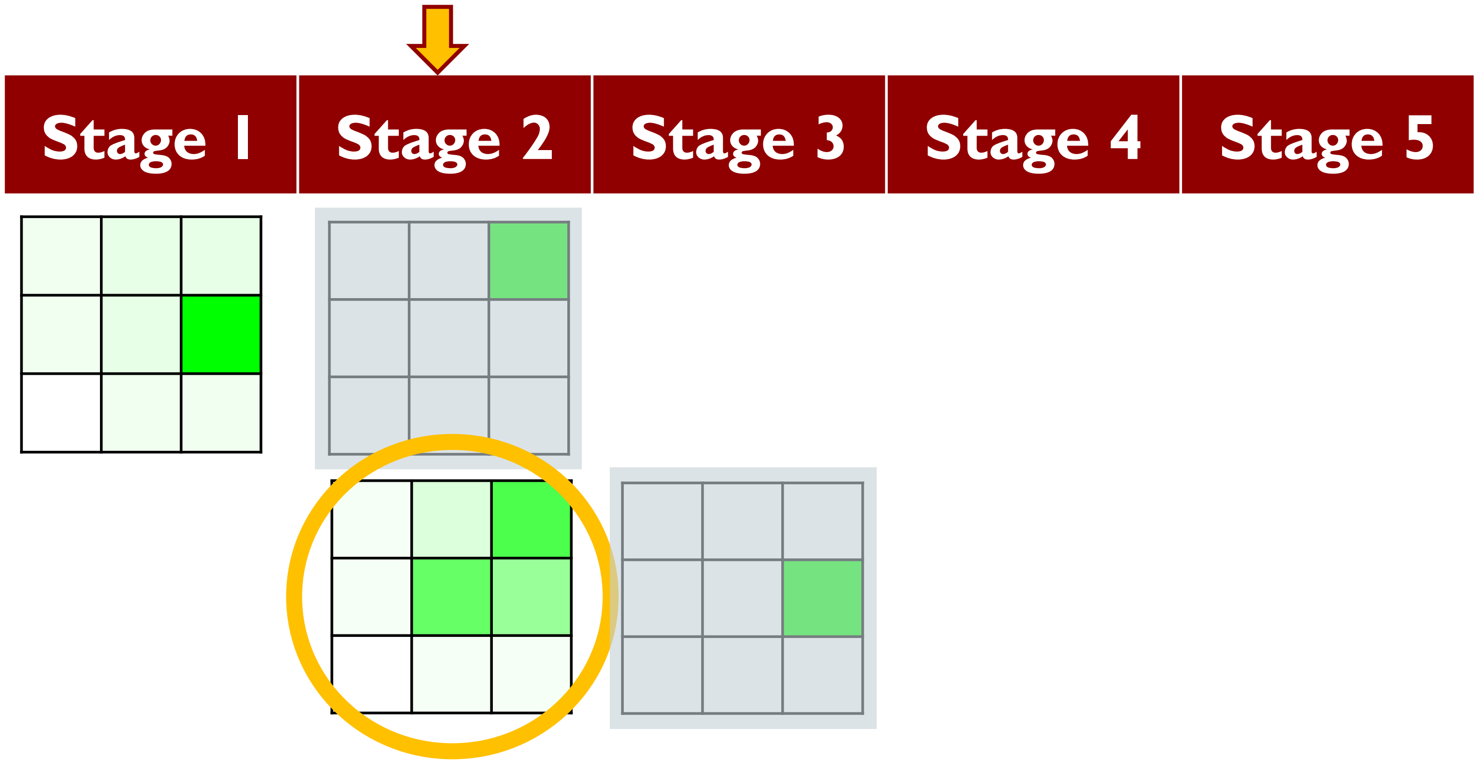
- Plan Ahead – M (PA-M)
- Plan ahead M stages



 Fang, F., Stone, P., & Tambe, M. When Security Games Go Green: Designing Defender Strategies to Prevent Poaching and Illegal Fishing. In IJCAI, 2015.

GSG: Incorporating Delayed Observation

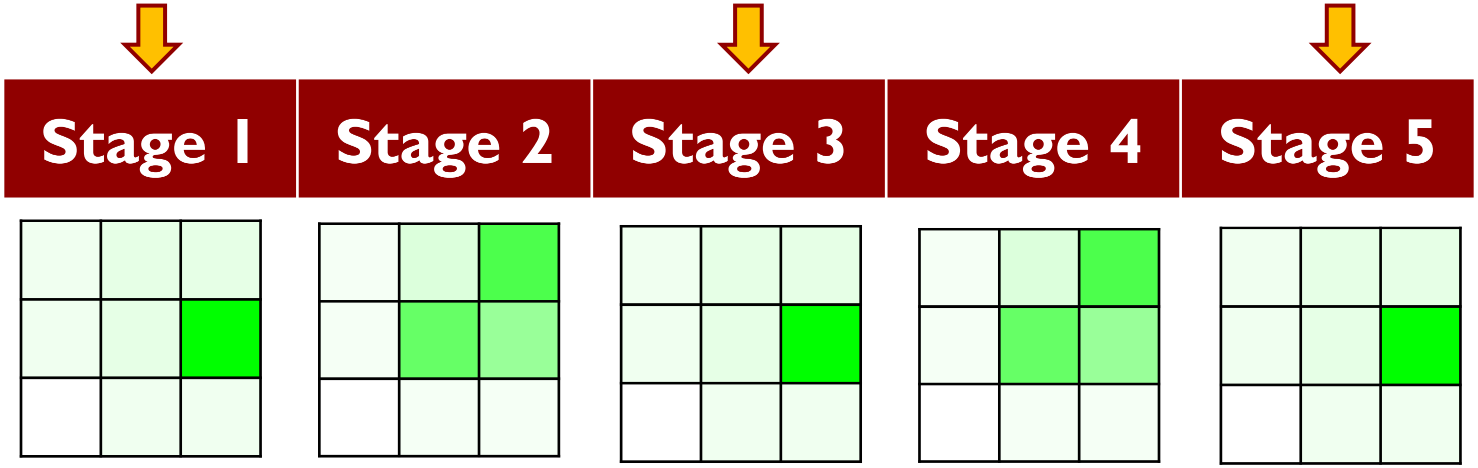
- Plan Ahead – M (PA-M)
- Plan ahead M stages



Fang, F., Stone, P., & Tambe, M. When Security Games Go Green: Designing Defender Strategies to Prevent Poaching and Illegal Fishing. In IJCAI, 2015.

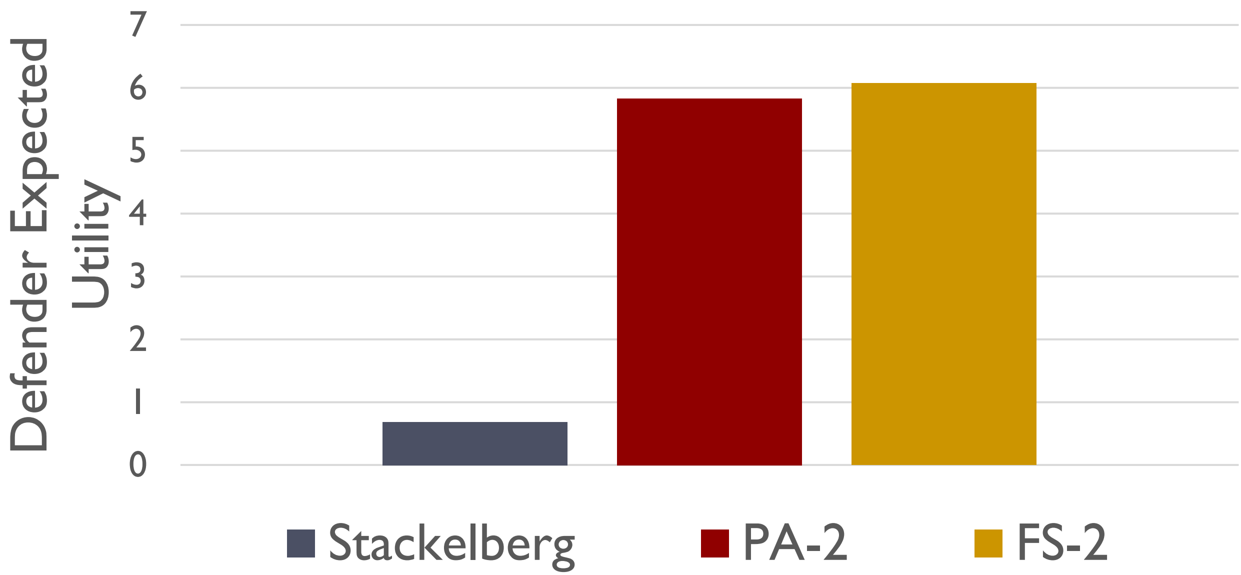
GSG: Incorporating Delayed Observation

- An alternative: Fixed Sequence – M (FS-M)
- Use M strategies repeatedly



Fang, F., Stone, P., & Tambe, M. When Security Games Go Green: Designing Defender Strategies to Prevent Poaching and Illegal Fishing. In IJCAI, 2015.

GSG: Incorporating Delayed Observation



- **Theorem 3**: In a GSG with T rounds, for $\Gamma < \mathrm{M} \leq T$, there exists a cyclic defender strategy profile $[s]$ with period M that is a $(1 - \frac{\Gamma}{T})\frac{Z-1}{Z+1}$ approximation of the optimal strategy profile in terms of the normalized utility, where $Z = \left\lfloor \frac{T-\Gamma+1}{M} \right\rfloor$

Fang, F., Stone, P., & Tambe, M. When Security Games Go Green: Designing Defender Strategies to Prevent Poaching and Illegal Fishing. In IJCAI, 2015.

Modeling and Learning Human Behavior in Games

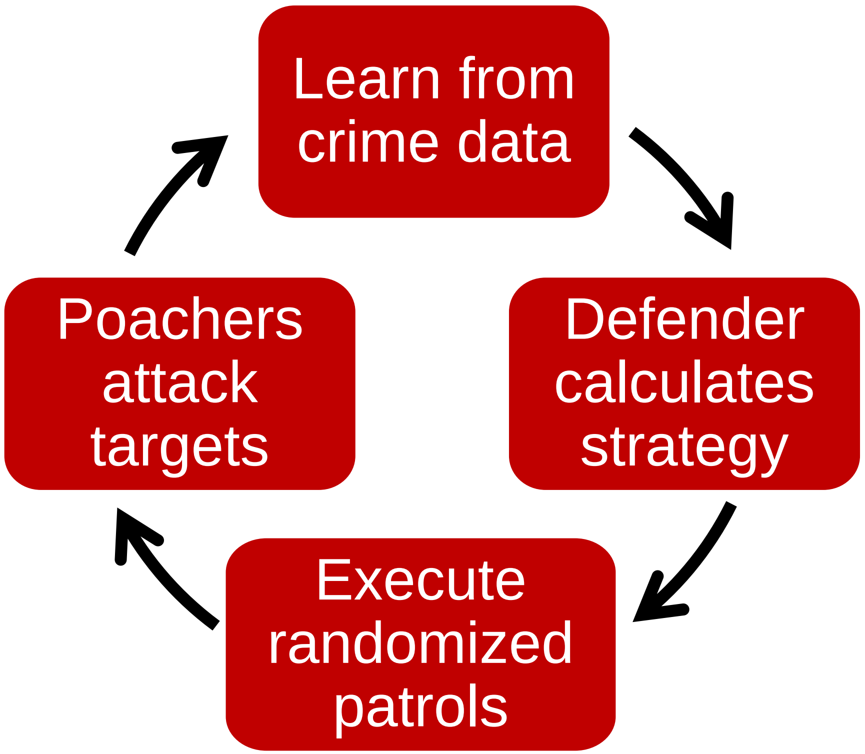
- Uncertainty and Bias Based Models
  - Prospect Theory [Kahneman and Tvesky, 1979]
  - Anchoring bias and epsilon-bounded rationality [Pita et al, 2010]
  - Attacker aims to reduce the defender's utility [Pita et al, 2012]
- Quantal Response Based Models
  - Quantal Response [McKelvey and Palfrey, 1995]
  - Subjective Utility Quantal Response [Nguyen et al, 2013]
- Other Models
  - Incorporating delayed observation [Fang et al, 2015]
  - Bounded rationality in repeated games [Kar et al, 2015]

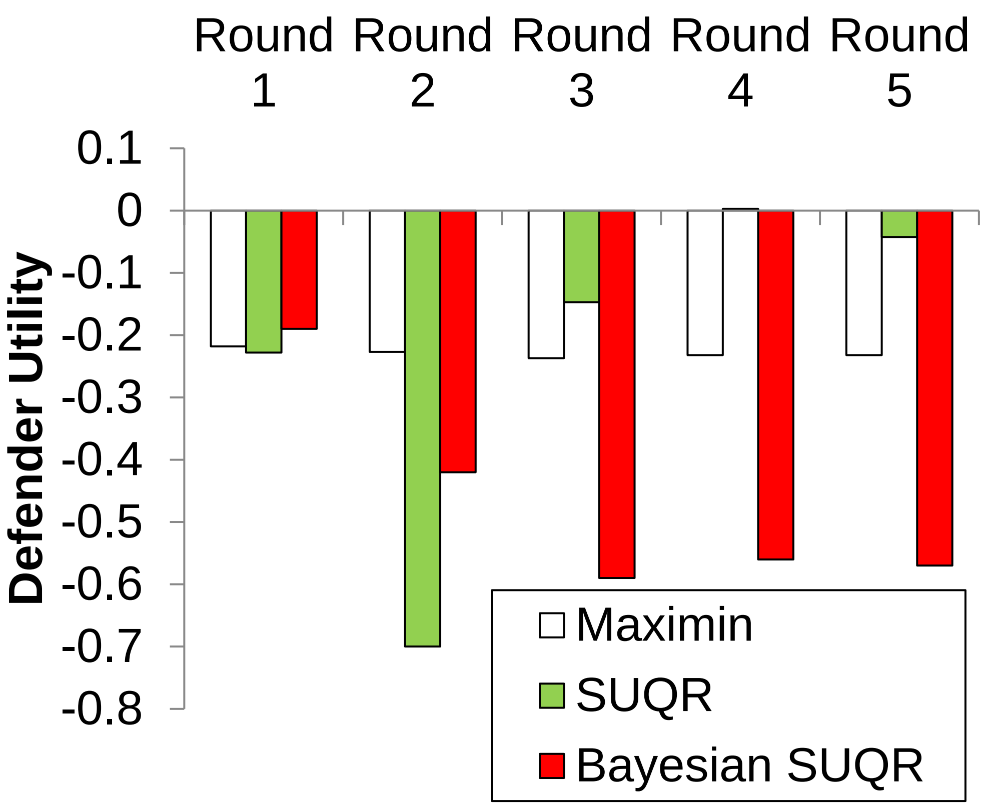
SHARP: Bounded Rationality in Repeated Games

Kar, D., Fang, F., Delle Fave, F., Sintov, N., & Tambe, M. A game of thrones: when human behavior models compete in repeated Stackelberg security games. In AAMAS, 2015

SHARP: Bounded Rationality in Repeated Games

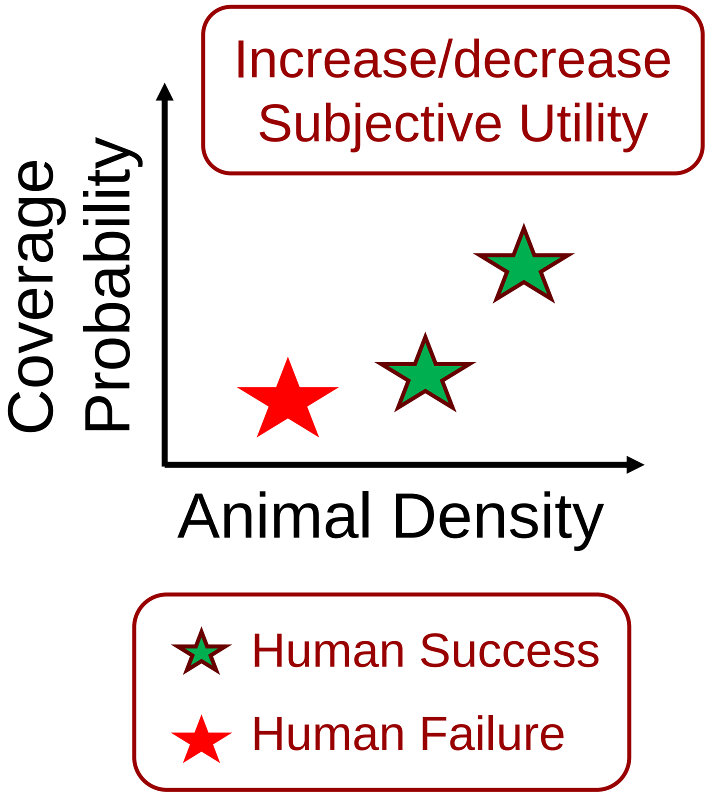
Repeated games on AMT: 35 weeks, 40 human subjects 10,000 emails!

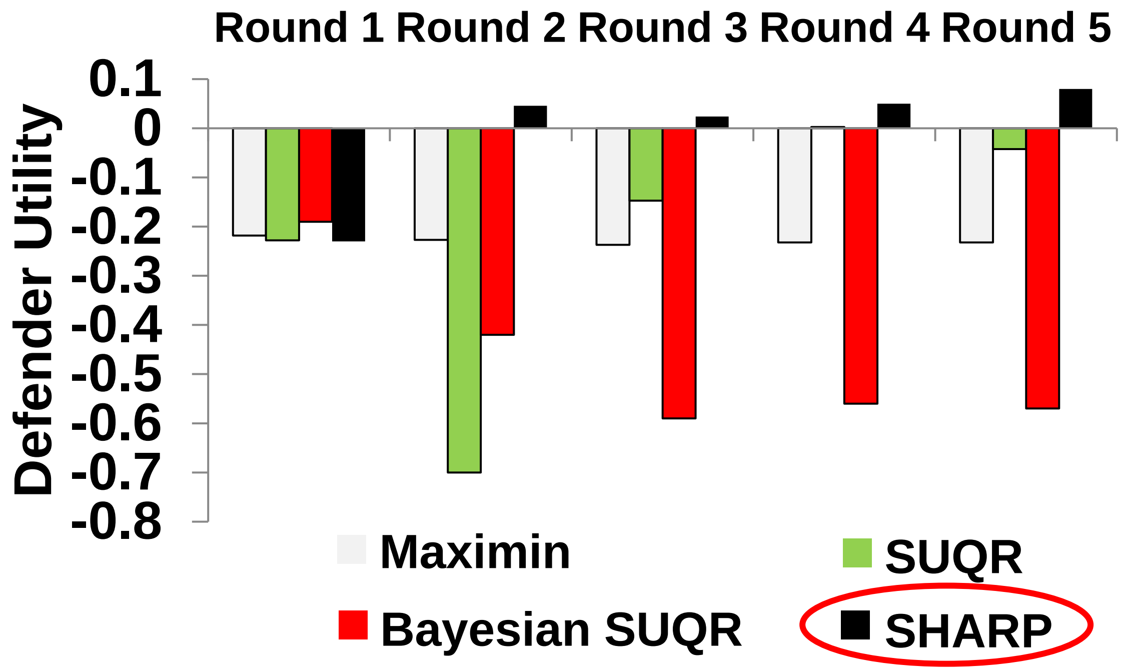




Kar, D., Fang, F., Delle Fave, F., Sintov, N., & Tambe, M.A game of thrones: when human behavior models compete in repeated Stackelberg security games. In AAMAS, 2015

SHARP: Bounded Rationality in Repeated Games

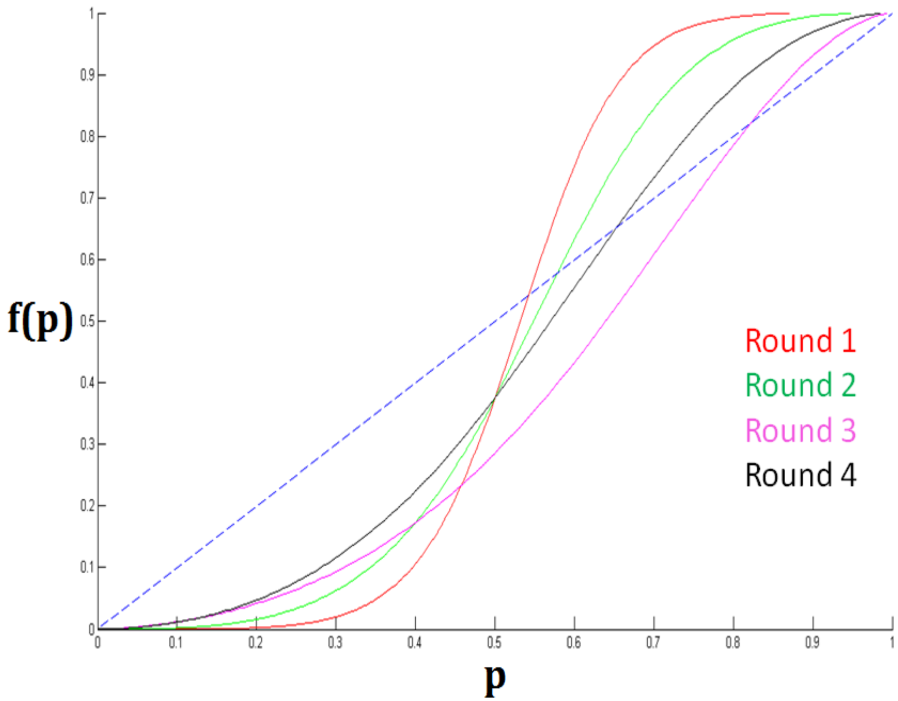




Kar, D., Fang, F., Delle Fave, F., Sintov, N., & Tambe, M.A game of thrones: when human behavior models compete in repeated Stackelberg security games. In AAMAS, 2015

SHARP: Bounded Rationality in Repeated Games

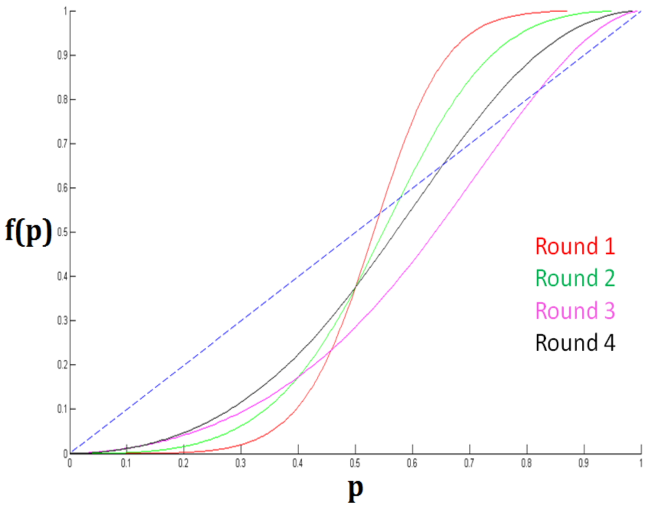
- Adversary's probability weighting function is S-shaped.
  - Contrary to Prospect Theory



Kar, D., Fang, F., Delle Fave, F., Sintov, N., & Tambe, M. A game of thrones: when human behavior models compete in repeated Stackelberg security games. In AAMAS, 2015

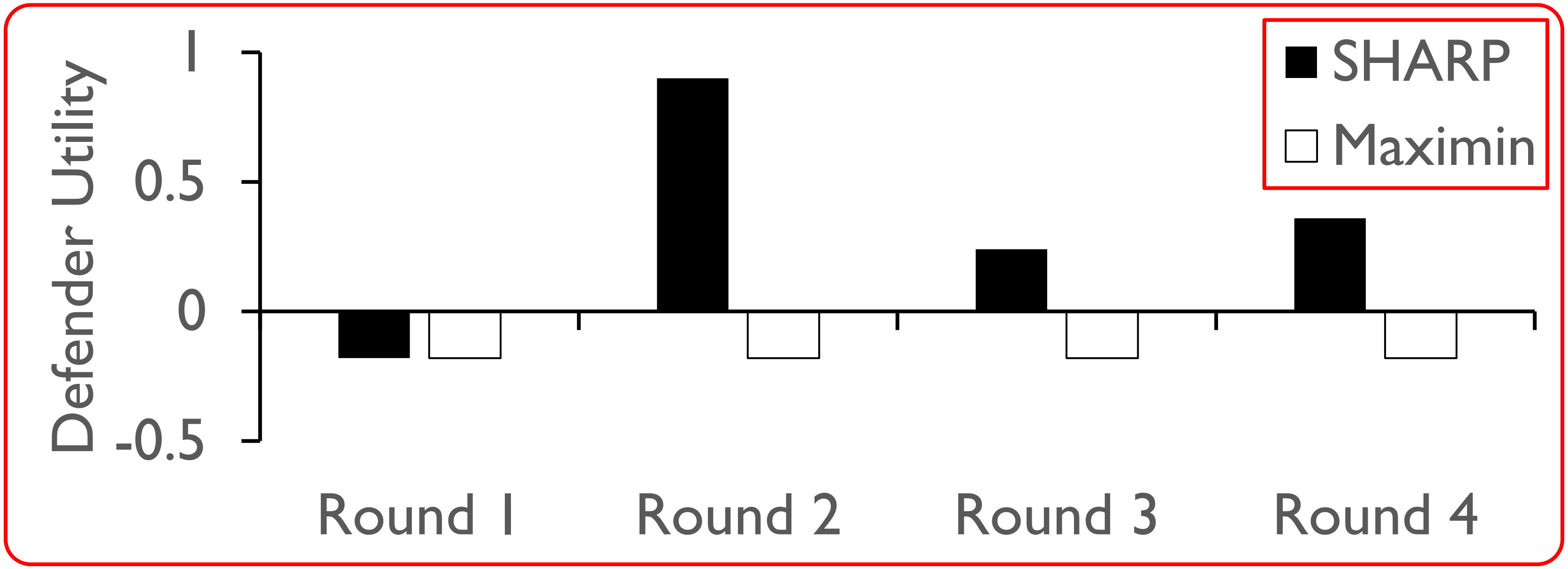
SHARP: Bounded Rationality in Repeated Games

- Q: According to the learned weighting function, which is S-shaped, the human players are <u>over/under?</u>-estimating the probability of getting caught when the probability is low



SHARP: Bounded Rationality in Repeated Games





Kar, D., Fang, F., Delle Fave, F., Sintov, N., & Tambe, M.A game of thrones: when human behavior models compete in repeated Stackelberg security games. In AAMAS, 2015

Other Models

- Cognitive Hierarchy
- Instance-based Learning Theory (IBLT)

Discussion

- Limitations of the models introduced today?