Artificial Intelligence Methods for Social Good

M2-3 [Game Theory]:

Human Behavior Modeling

08-537 (9-unit) and 08-737 (12-unit)

Instructor: Fei Fang

feifang@cmu.edu

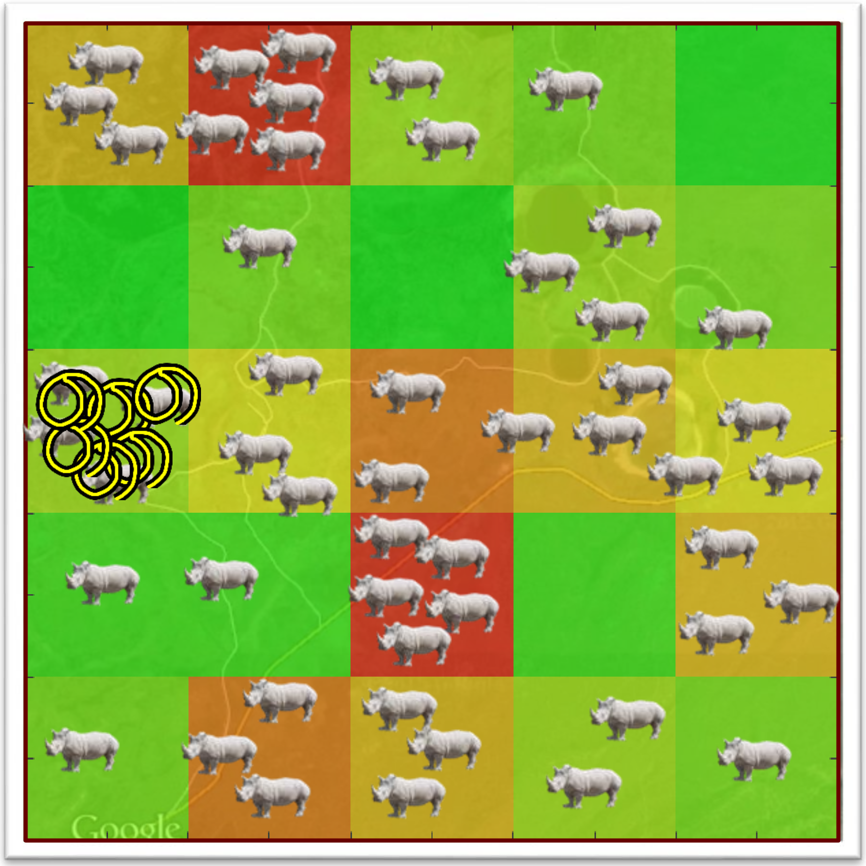
Wean Hall 4126

Challenges in Wildlife Conservation

- 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

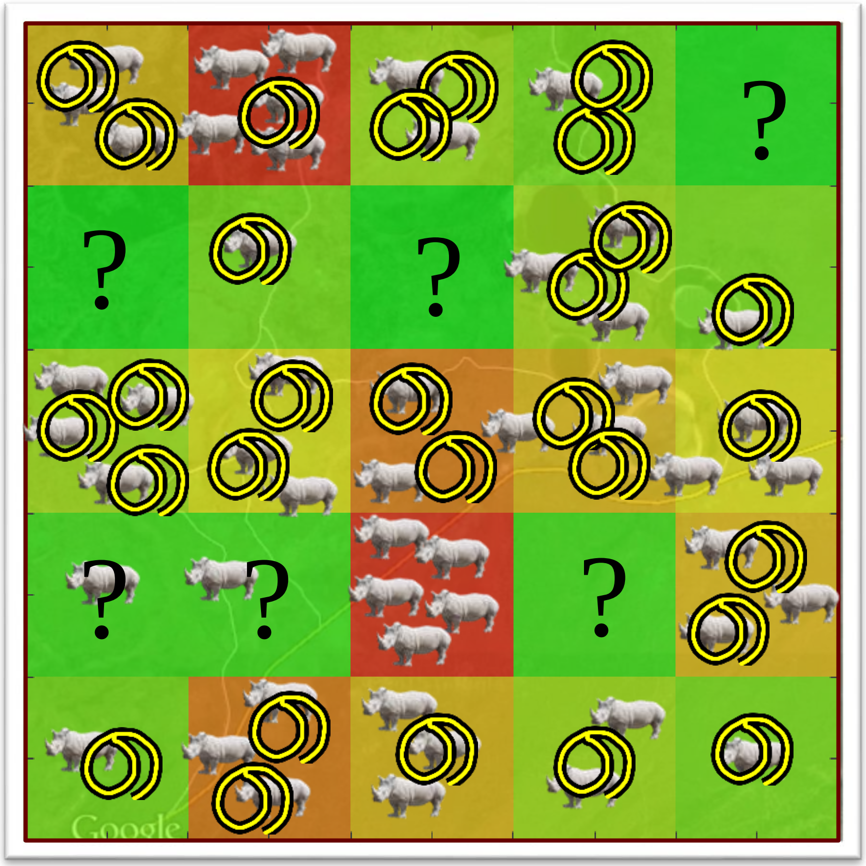
Challenges in Wildlife Conservation

- Perfectly rational (Maximize expected utility)? No!



Challenges in Wildlife Conservation

- Real-world data



Human Behavior Modeling & Learning

- 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]
- Latest Models
  - Incorporating delayed observation [Fang et al, 2015]
  - Bounded rationality in repeated games [Kar et al, 2015]
  - Two-layered model [Nguyen et al, 2016]
  - Decision tree-based model [Kar & Ford et al, 2017]
- PAWS

Human Behavior Modeling & Learning

- 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]
- Latest Models
  - Incorporating delayed observation [Fang et al, 2015]
  - Bounded rationality in repeated games [Kar et al, 2015]
  - Two-layered model [Nguyen et al, 2016]
  - Decision tree-based model [Kar & Ford et al, 2017]
- PAWS

PT: Prospect Theory

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

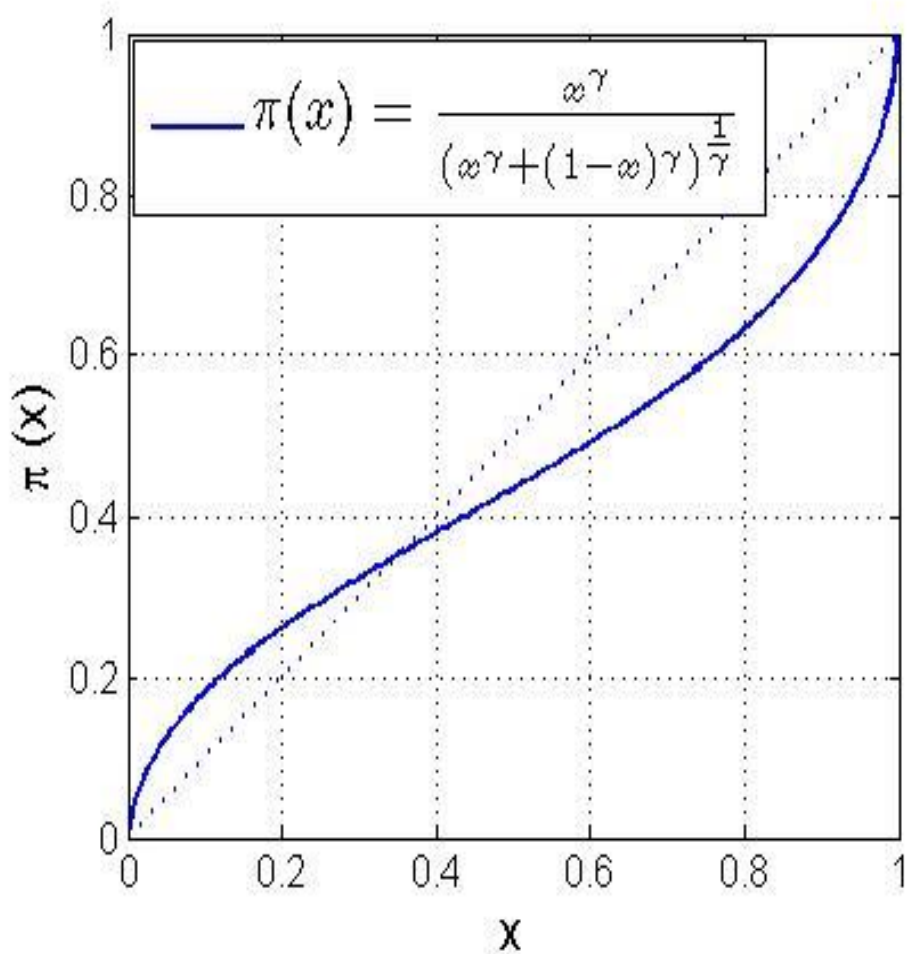
$$\text{prospect} = \sum_{i \in AllOutcomes} \pi(x_i) \cdot V(C_i)$$

  - 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

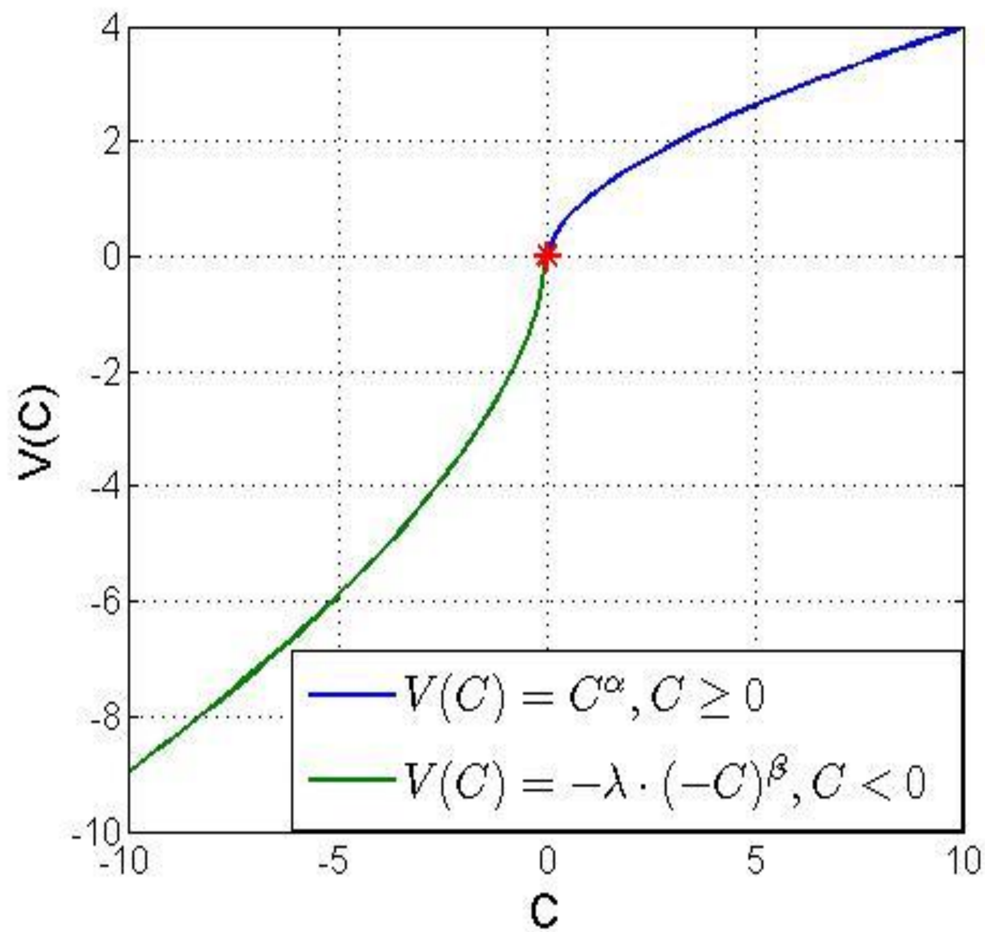
- Empirical Weighting Function
- Slope gets steeper as x gets closer to 0 and 1
- Not consistent with probability definition
  - $\pi(x)+\pi(1-x) < 1$
- Empirical value: $\gamma=0.64$ $(0<\gamma<1)$



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.

COBRA: Anchoring Bias and Epsilon-Bounded Rationality

- "epsilon optimality"
- Anchoring bias: Full observation ($\alpha = 0$) vs no observation ($\alpha = 1$)

$$\max_{x,q,\gamma,a} \gamma$$
$$s.t.\, x' = (1-\alpha)x + \frac{\alpha}{N}$$
a is attacker's highest expected utility given x'
$$q_j = 1 \text{ if } \mathrm{AttEU}_j(x') \geq a - \epsilon$$
$$\gamma \leq \mathrm{DefEU}_j(\mathrm{x}) \text{ if } q_j = 1$$

- 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.

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

Human Behavior Modeling & Learning

- 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]
- Latest Models
  - Incorporating delayed observation [Fang et al, 2015]
  - Bounded rationality in repeated games [Kar et al, 2015]
  - Two-layered model [Nguyen et al, 2016]
  - Decision tree-based model [Kar & Ford et al, 2017]
- PAWS

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.

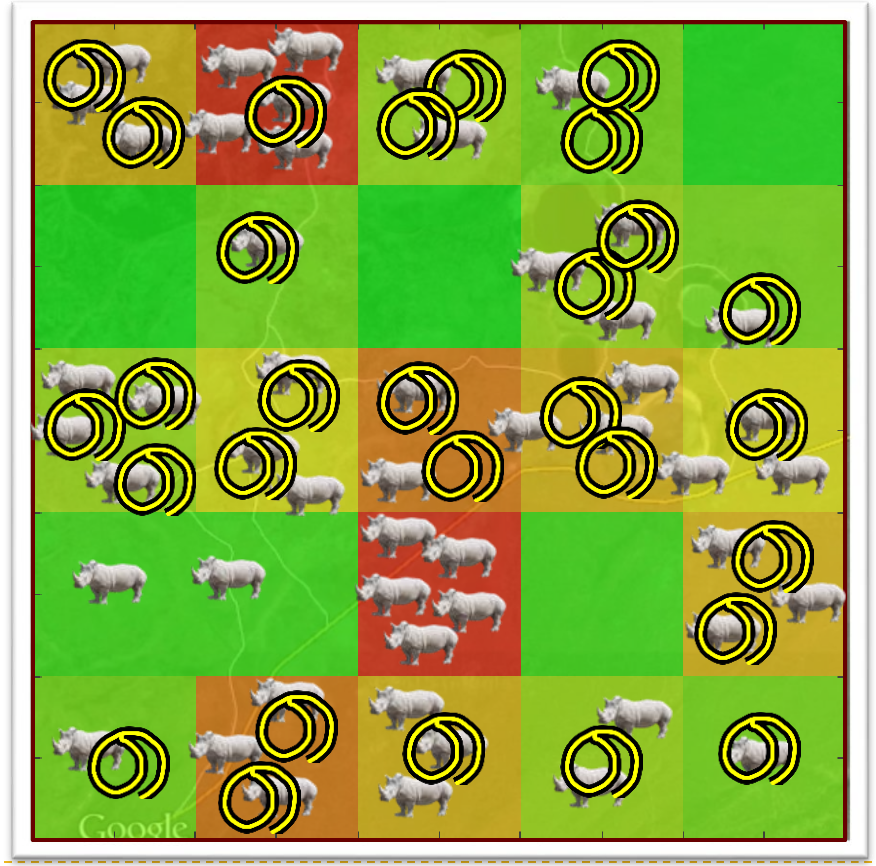
Quiz 1: Quantal Response Model

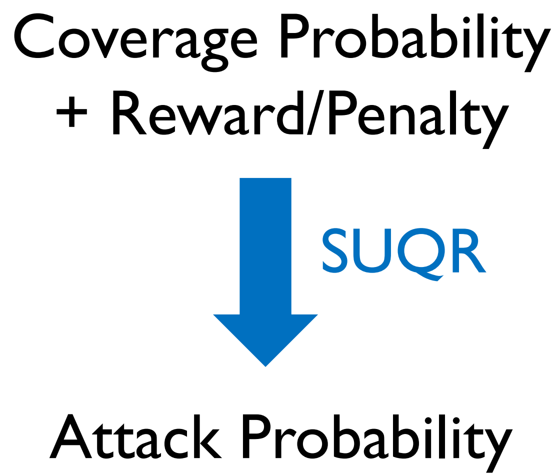
- If there are two choices (actions), what is the probability of choosing the first action if the player follows quantal response model with $\lambda = 0$?
  - 1
  - 0
  - $\frac{1}{2}$
  - $\frac{1}{e} \approx 0.368$

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

SUQR: Subjective Utility Quantal Response Model

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

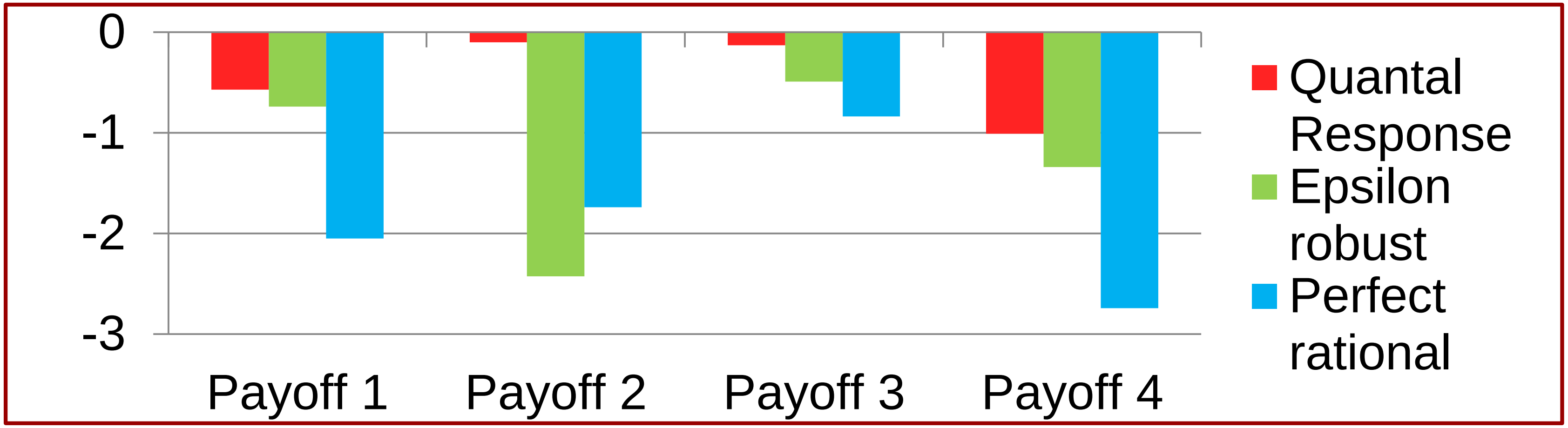




Nguyen, T. H., Yang, R., Azaria, A., Kraus, S., & Tambe, M. Analyzing the Effectiveness of Adversary Modeling in Security Games. In AAAI, 2013.

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.

Human Behavior Modeling & Learning

- 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]
- Latest Models
  - Incorporating delayed observation [Fang et al, 2015]
  - Bounded rationality in repeated games [Kar et al, 2015]
  - Two-layered model [Nguyen et al, 2016]
  - Decision tree-based model [Kar & Ford et al, 2017]
- PAWS Application

GSG: Incorporating Delayed Observation

Wildlife

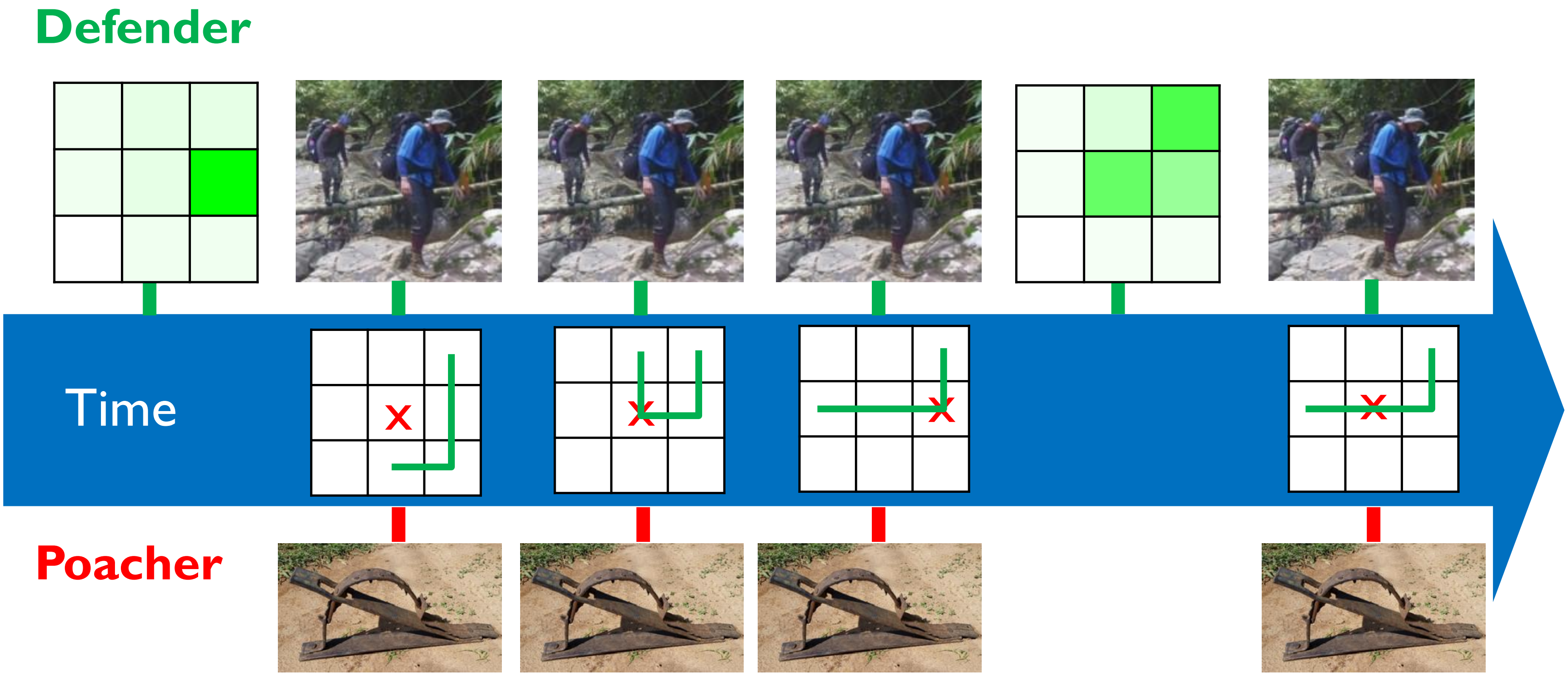
Forest

Fishery

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

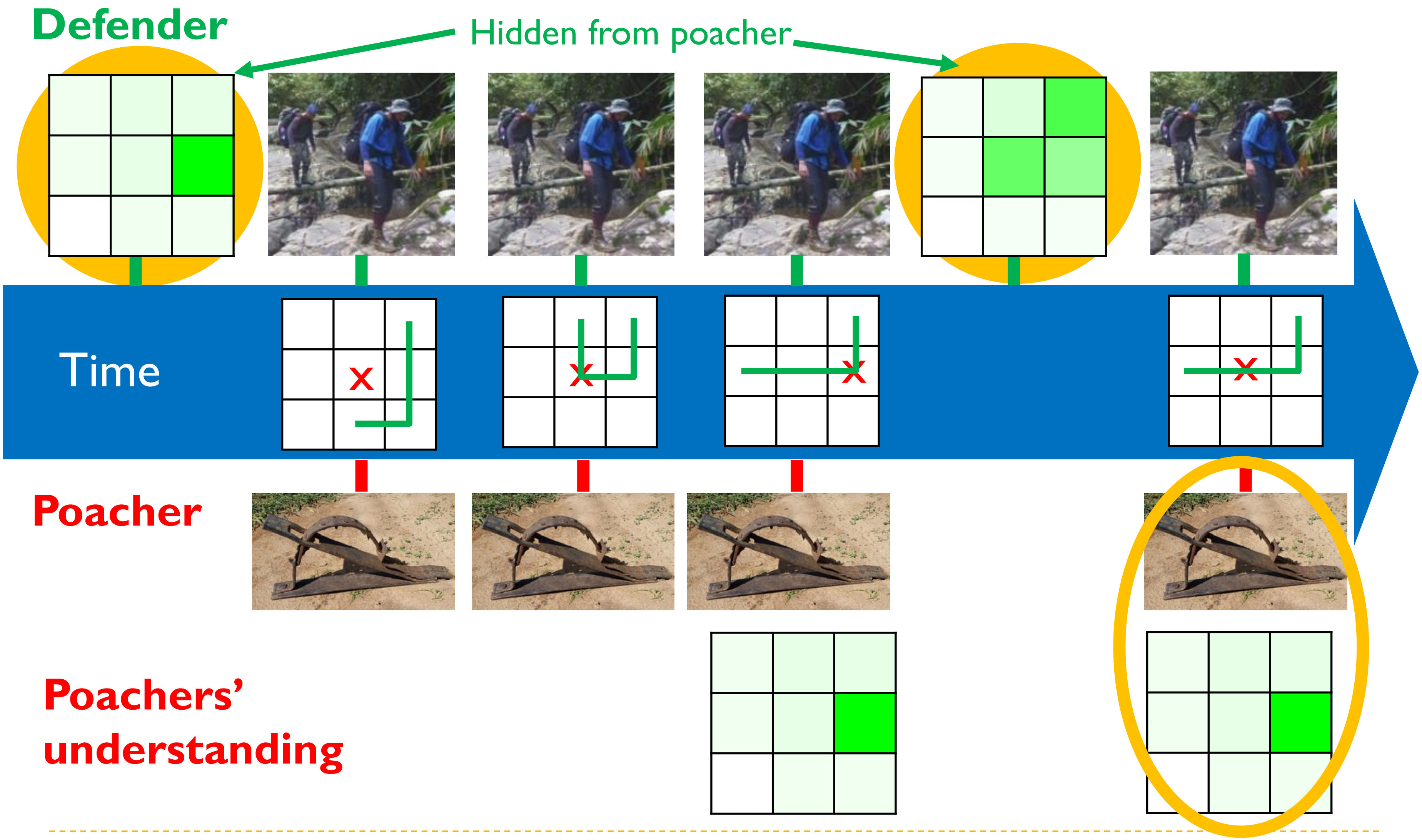
Fang, F., Stone, P., & Tambe, M. When Security Games Go Green: Designing Defender Strategies to Prevent Poaching and Illegal Fishing. In IJCAI, 2015.

GSG: Incorporating Delayed Observation



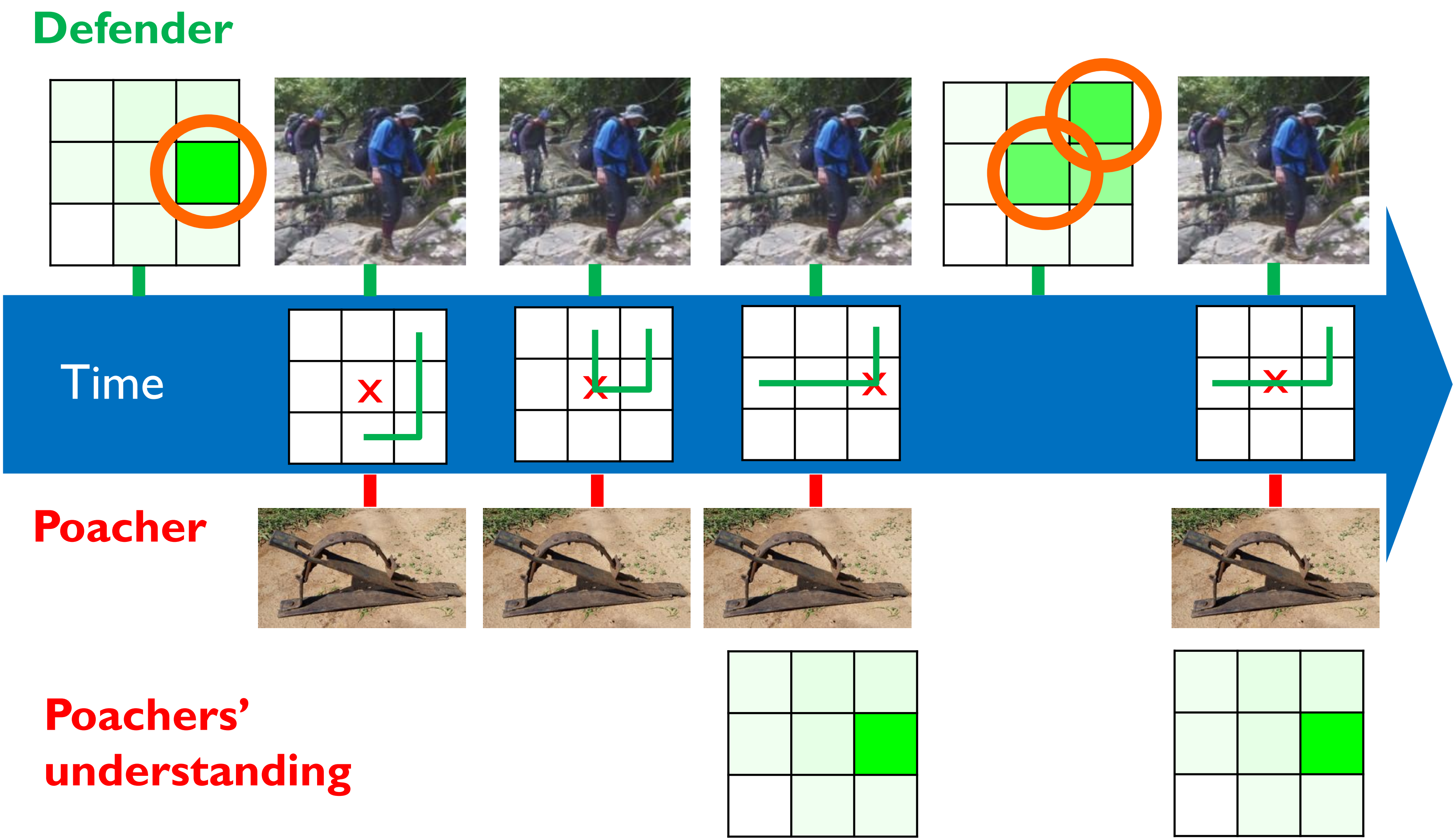
Fang, F., Stone, P., & Tambe, M. When Security Games Go Green: Designing Defender Strategies to Prevent Poaching and Illegal Fishing. In IJCAI, 2015.

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

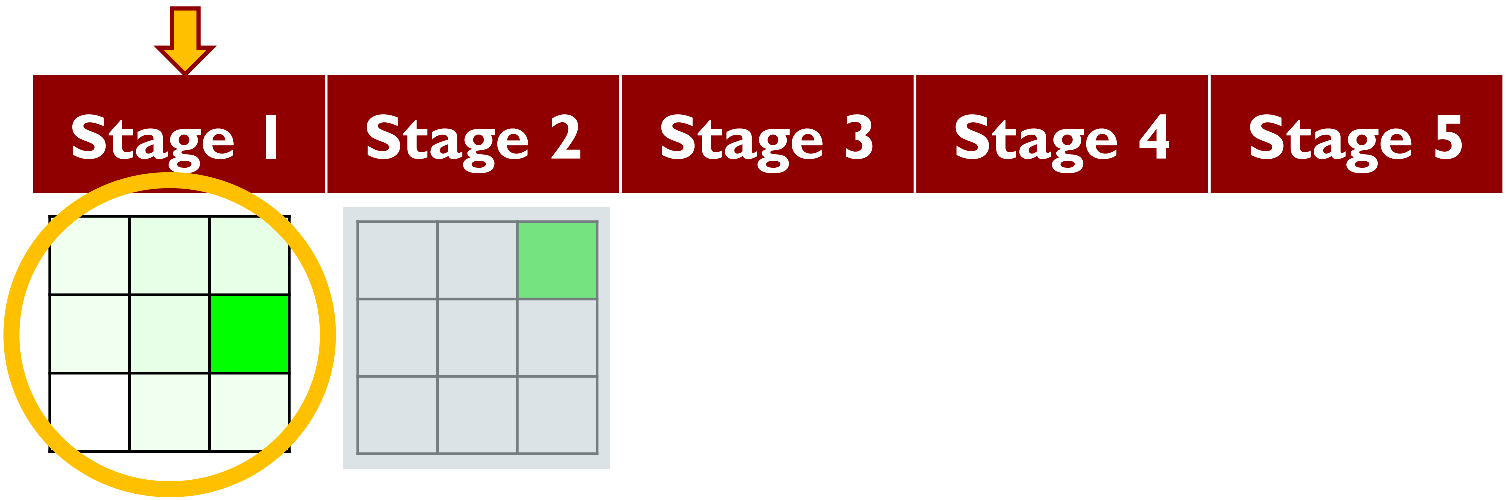
- A Green Security Game (GSG) is a T stage game where the defender protects N targets against L attackers. Defender chooses a mixed strategy c^t in stage t.

- A GSG attacker is characterized by his memory length Γ, coefficients $\alpha_0, \ldots, \alpha_\Gamma$ 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

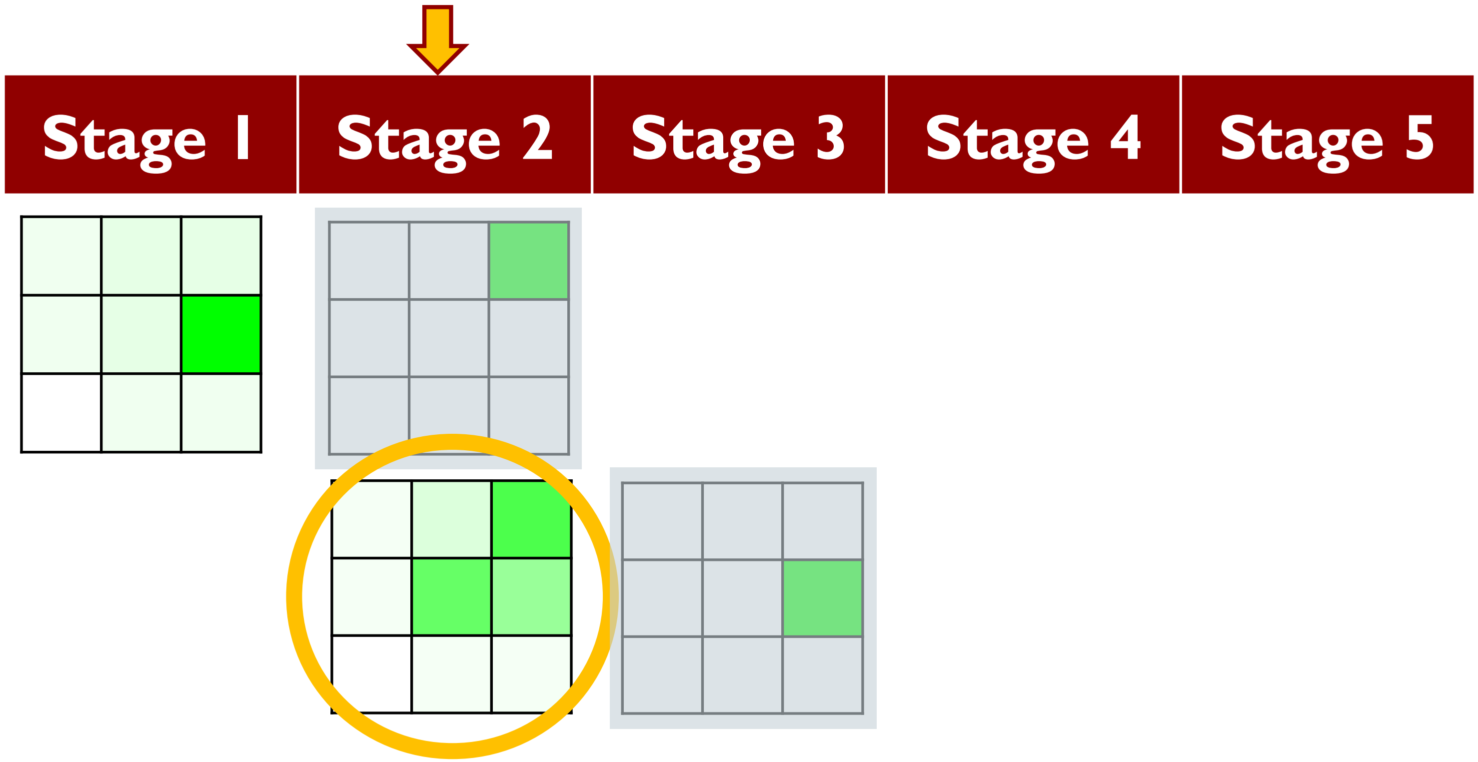
- Plan Ahead – M (PA-M)
- Plan ahead M stages



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

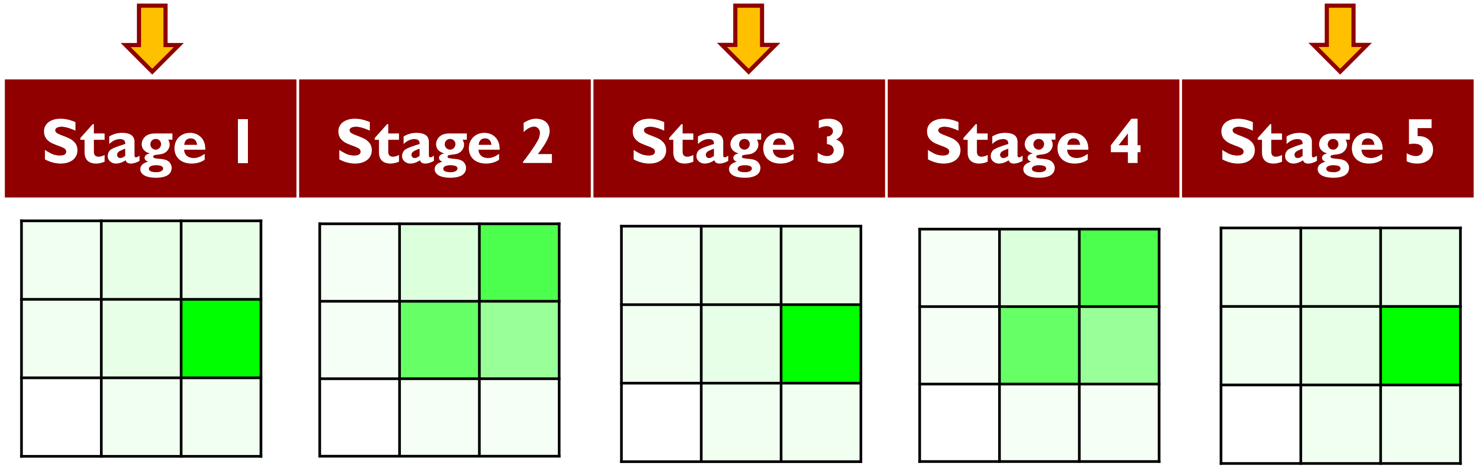
- Plan Ahead – M (PA-M)
- Plan ahead M stages



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

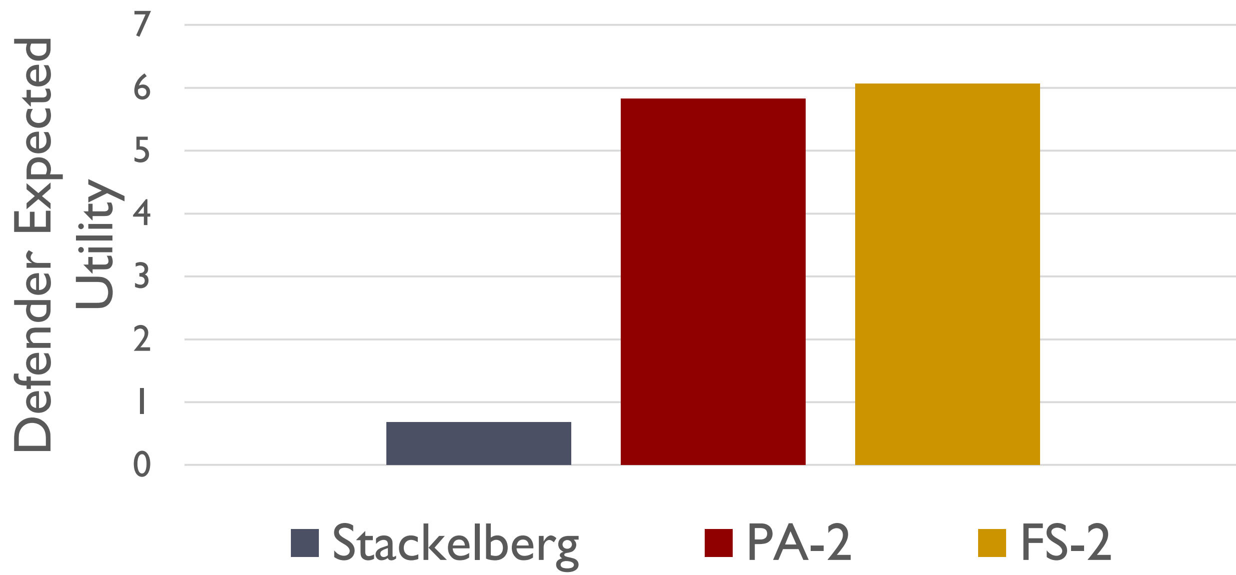
GSG: Incorporating Delayed Observation

- 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.

Human Behavior Modeling & Learning

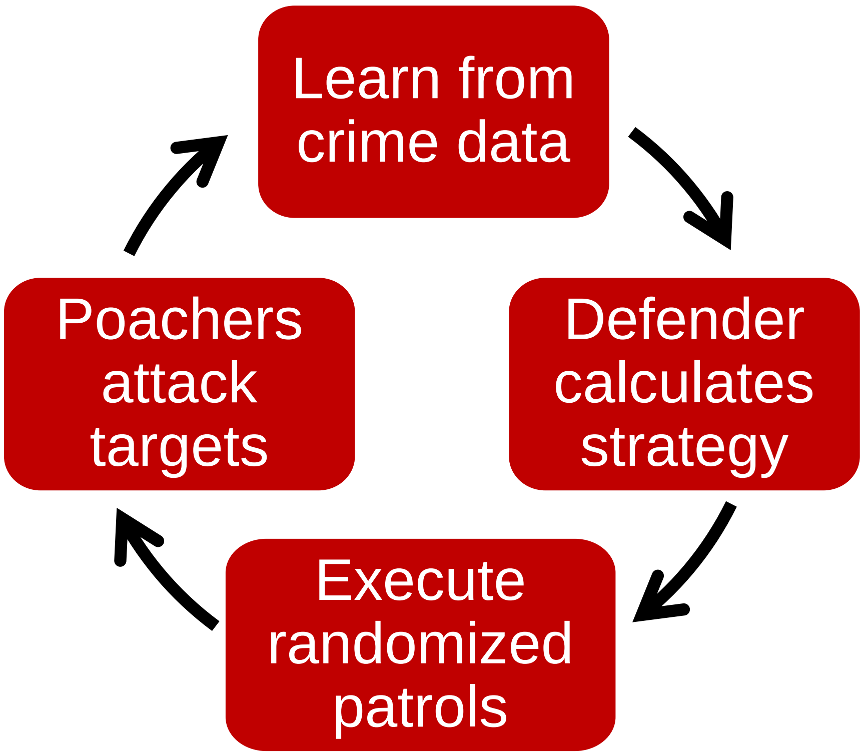
- Uncertainty and Bias Based Models
  - Prospect Theory [Kahneman and Tvesky, 1979]
  - Anchoring bias and epsilon-bounded rationality [Pita et al, 2010]
  - Attacker aims to reduce the defender's utility [Pita et al, 2012]
- Quantal Response Based Models
  - Quantal Response [McKelvey and Palfrey, 1995]
  - Subjective Utility Quantal Response [Nguyen et al, 2013]
- Latest Models
  - Incorporating delayed observation [Fang et al, 2015]
  - Bounded rationality in repeated games [Kar et al, 2015]
  - Two-layered model [Nguyen et al, 2016]
  - Decision tree-based model [Kar & Ford et al, 2017]
- PAWS Application

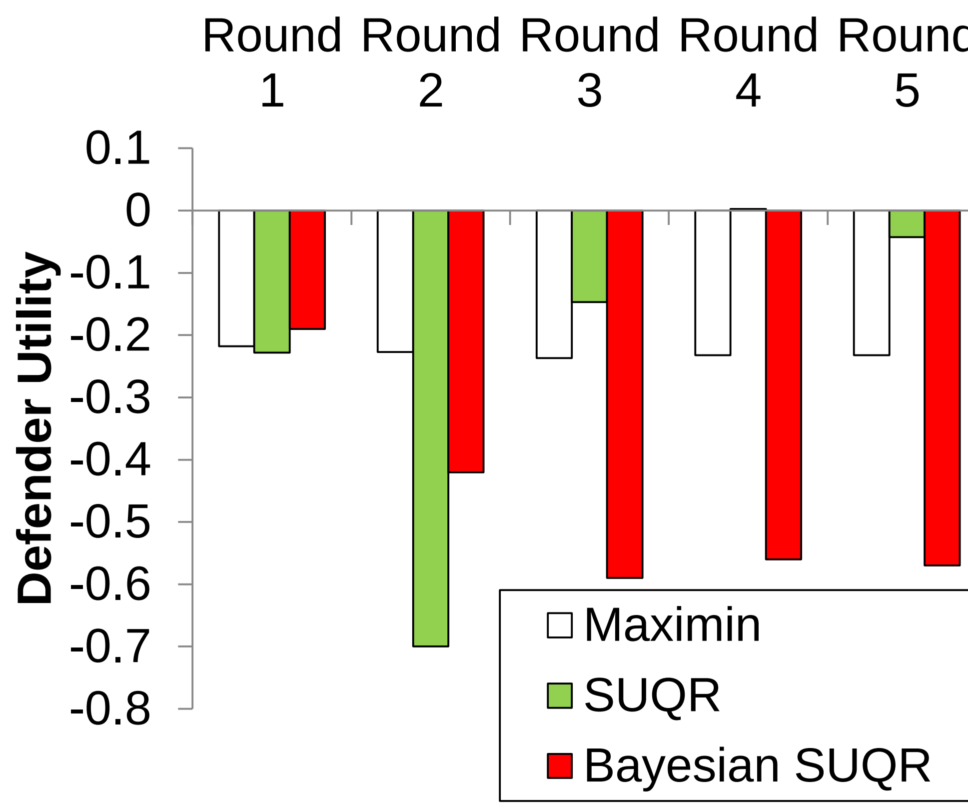
SHARP: Bounded Rationality in Repeated Games

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

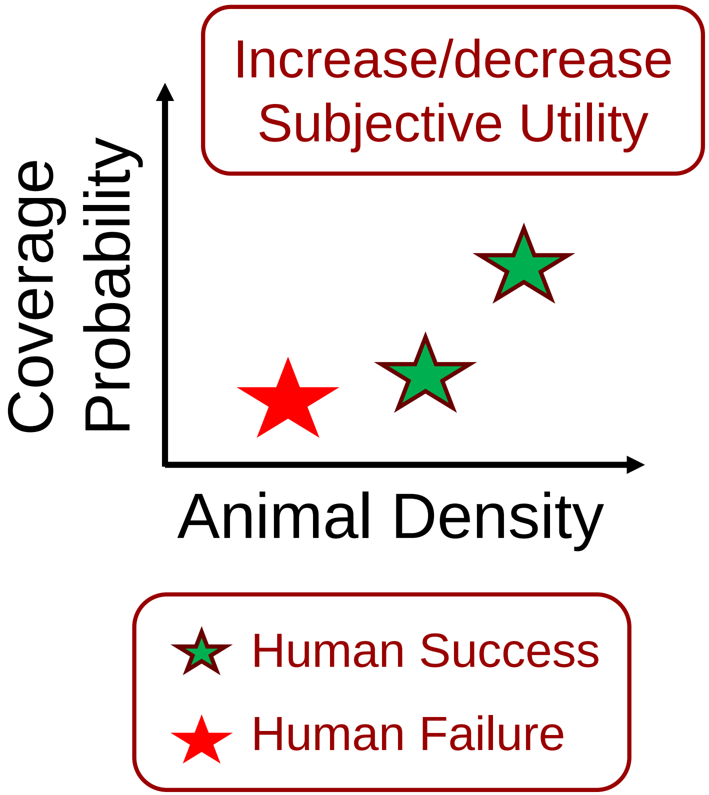
Repeated games on AMT: 35 weeks, 40 human subjects 10,000 emails!

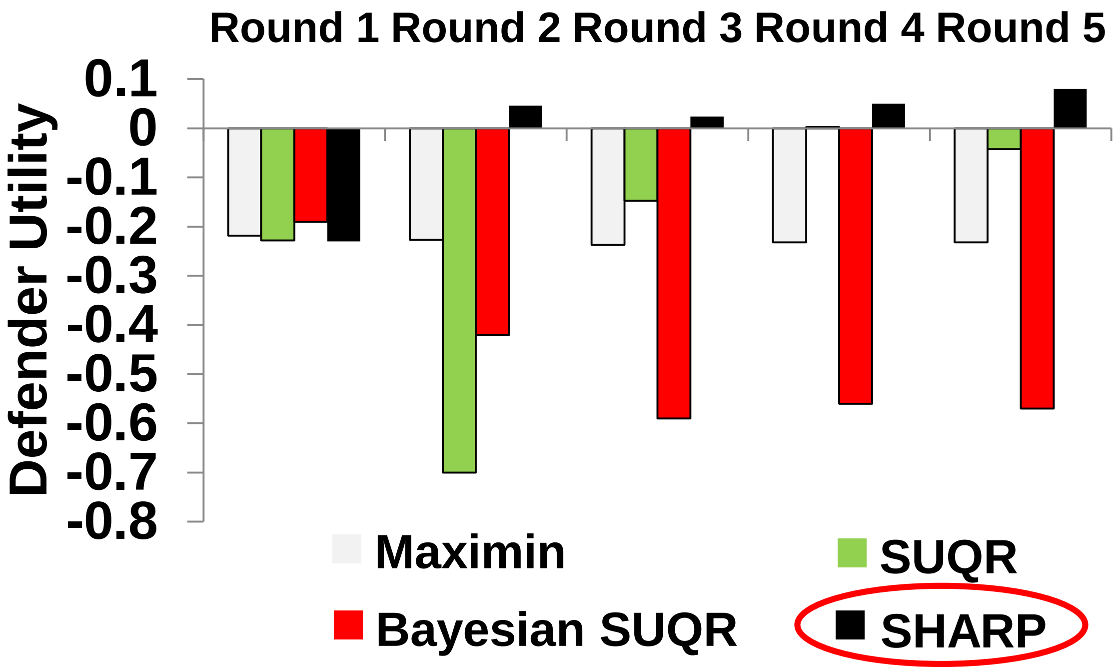




Kar, D., Fang, F., Delle Fave, F., Sintov, N., & Tambe, M.A game of thrones: when human behavior models compete in repeated Stackelberg security games. In AAMAS, 2015

SHARP: Bounded Rationality in Repeated Games

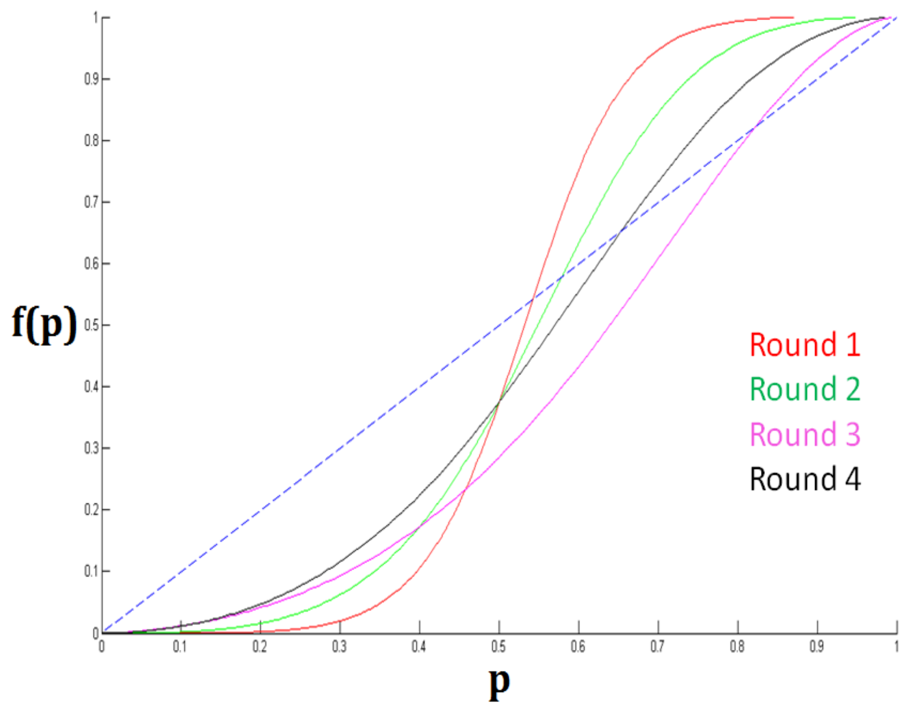




Kar, D., Fang, F., Delle Fave, F., Sintov, N., & Tambe, M.A game of thrones: when human behavior models compete in repeated Stackelberg security games. In AAMAS, 2015

SHARP: Bounded Rationality in Repeated Games

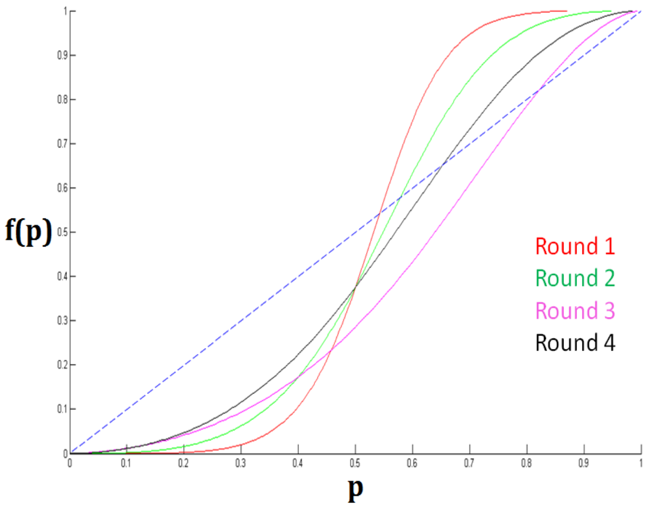
- Adversary's probability weighting function is S-shaped.
  - Contrary to Prospect Theory



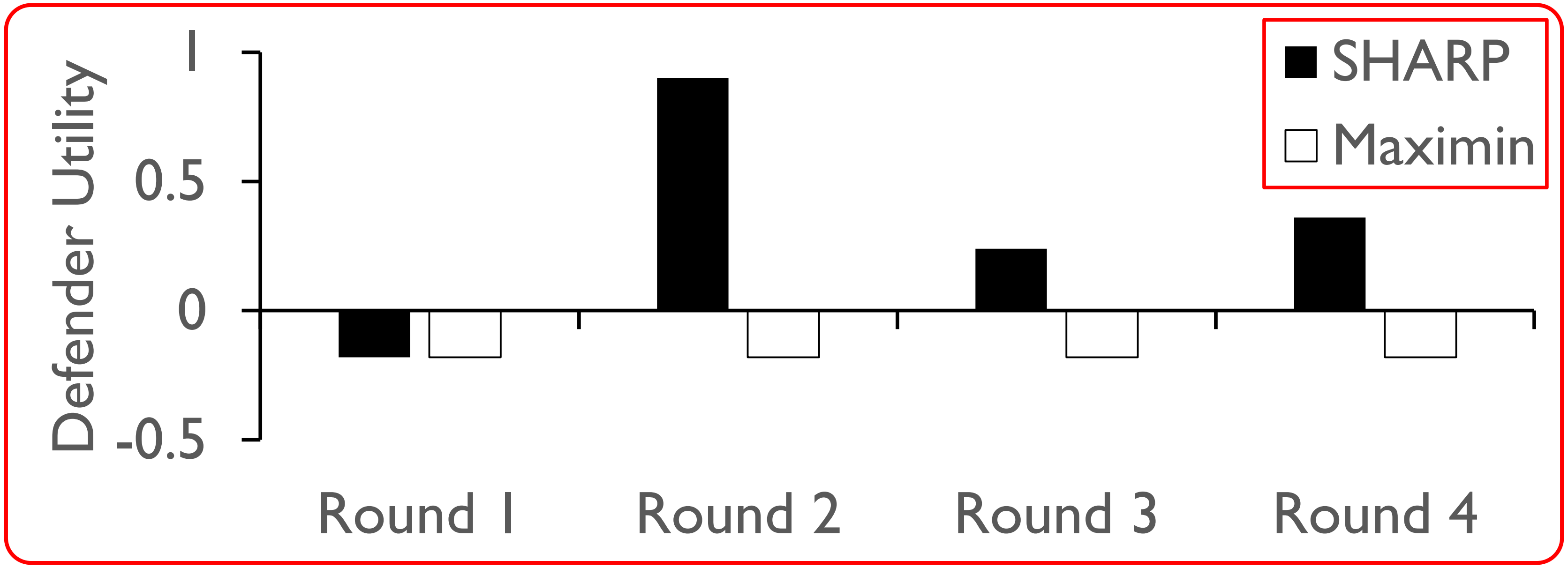
Kar, D., Fang, F., Delle Fave, F., Sintov, N., & Tambe, M.A game of thrones: when human behavior models compete in repeated Stackelberg security games. In AAMAS, 2015

Quiz 2: SHARP

- According to the learned weighting function, which is S-shaped, the human players are ________ the probability of getting caught when the probability is low
  - Over-estimating
  - Under-estimating



SHARP: Bounded Rationality in Repeated Games



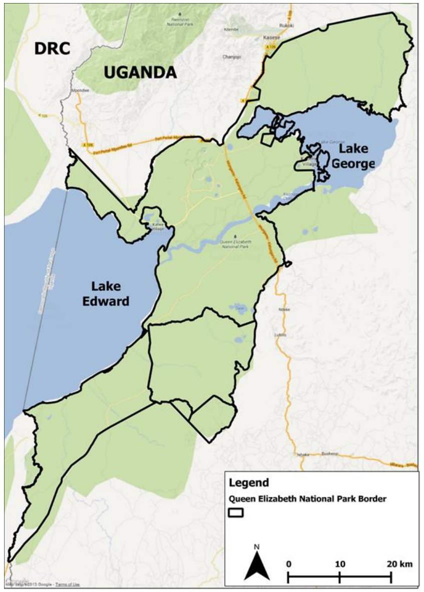
Kar, D., Fang, F., Delle Fave, F., Sintov, N., & Tambe, M.A game of thrones: when human behavior models compete in repeated Stackelberg security games. In AAMAS, 2015

Human Behavior Modeling & Learning

- 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]
- Latest Models
  - Incorporating delayed observation [Fang et al, 2015]
  - Bounded rationality in repeated games [Kar et al, 2015]
  - Two-layered model [Nguyen et al, 2016]
  - Decision tree-based model [Kar & Ford et al, 2017]
- PAWS Application

Real-World Data

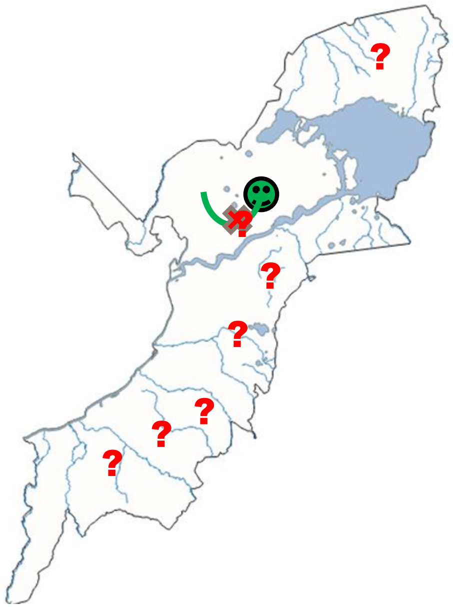
- Queen Elizabeth National Park
  - 1,978 sq. km
  - 2003-2015
- Geospatial Features
  - Terrain (e.g., forest, slope)
  - Distance to {Town, Water, Outpost}
- Ranger Coverage
- Crime Observations



Nguyen et al. Capture: A new predictive anti-poaching tool for wildlife protection. In AAMAS, 2016

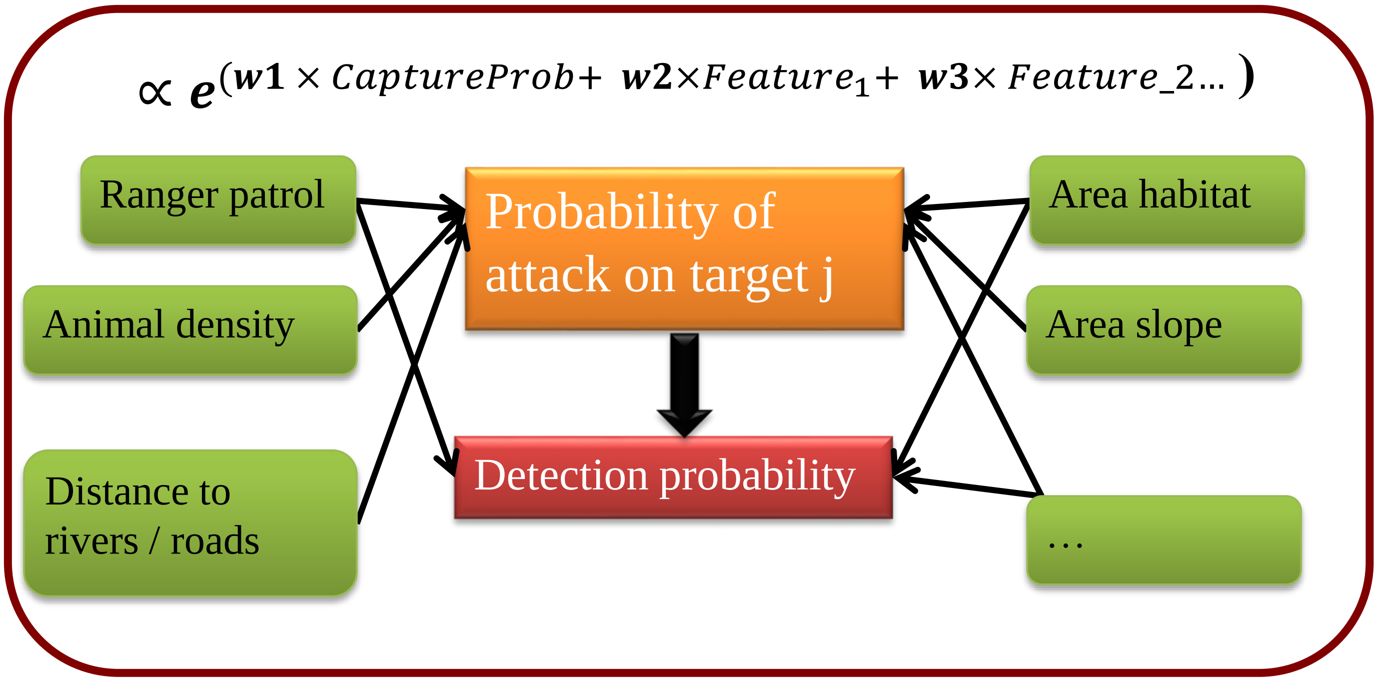
Real-World Data: Challenges

- “Missing” poaching data
  - Limited patrol resources (silent victims)
  - Imperfect observations (e.g., hidden snares)
- Consequences
  - Uncertainty in negative labels
  - Class imbalance



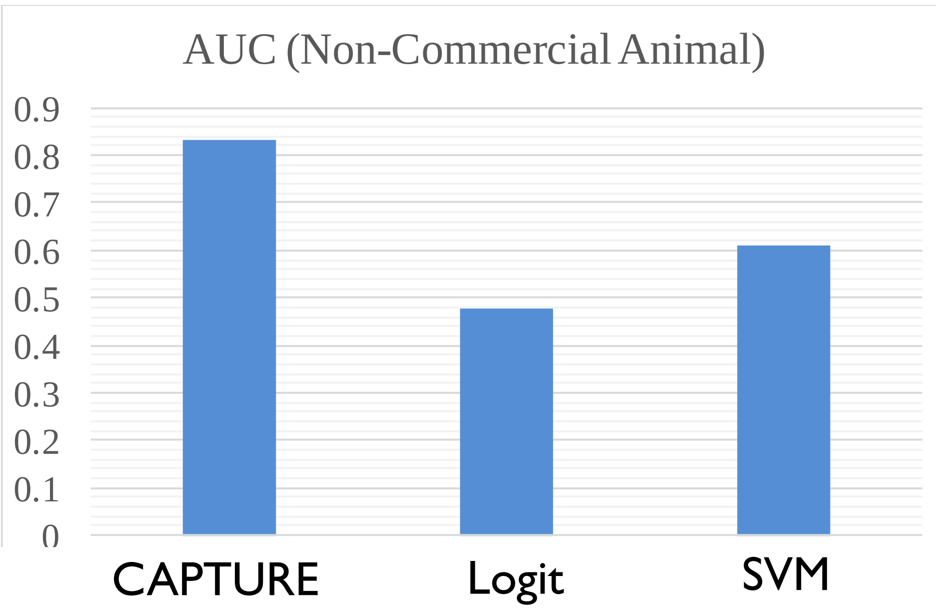
Nguyen et al. Capture: A new predictive anti-poaching tool for wildlife protection. In AAMAS, 2016

CAPTURE: Two-Layered Model

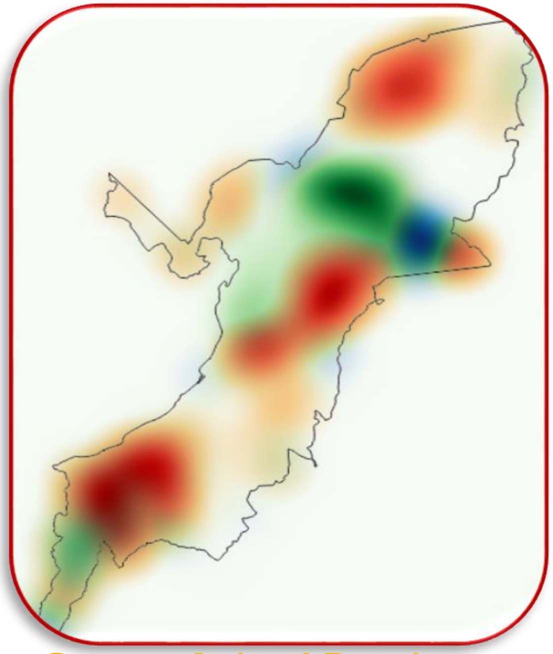


Nguyen et al. Capture: A new predictive anti-poaching tool for wildlife protection. In AAMAS, 2016

CAPTURE: Two-Layered Model



Dry Season (June-August 2008)



Green – Animal Density;

Blue – Defender Strategy;

Red – Observed Attack Probability

Quiz 3: Real-World Data Challenge

- Which of the following are challenges in the real-world data collected through anti-poaching patrols?
  - Limited amount of data
  - Uncertainty in negative labels
  - Class imbalance
  - Uncertainty in positive labels