Scalable, Stochastic and Spatiotemporal Game Theory for Real-World Human Adversarial Behavior

Milind Tambe

MURI Review

July 24, 2013

Program managers: Purush Iyer, Harry Chang

MURI Team

Investigator	Qualifs.	Discipline	Institution	#P/S
PI:Milind Tambe	Helen N. and Emmett H. Jones Professor in Engineering	Computer Science	Univ of Southern California	2
Andrea Bertozzi	Betsy Wood Knapp Chair for Innovation and Creativity	Math	Univ of California Los Angeles	2
P. Jeffrey Brantingham	Professor	Anthropology	Univ of California Los Angeles	1
Vincent Conitzer	Sally Dalton Robinson Professor of Computer Science	Computer Science	Duke University	1
Maria D'Orsogna	Associate Professor	Math	Cal State Univ Northridge	1
Richard John	Associate Professor	Psychology	Univ of Southern California	1
Rajiv Maheswaran	Research Asst Professor	Computer Science	Univ of Southern California	1
Michael McBride	Associate Professor	Economics	Univ of California Irvine	1
Martin Short	Asst Professor	Math	Georgia Tech (previous UCLA)	1

Scientific Objectives

Fundamentally new game theoretic concepts & approaches

- Based on models of real-world adversaries and their strategies:
  - *Social-cultural/cognitively-biased, act in space & time, in coalitions, with uncertainty*

Scientific barriers and opportunities:

- Quantitative behavior modeling in massive scale-games
  - *Previous work: single "rational" (strict payoff-maximizing) agents*
  - *Opportunity: Leverage new behavioral models, OR approaches for games*
- Games with geographically-dispersed actions, actions in networks
  - *Previous work: Limited work on spatiotemporal actions, massive networks*
  - *Opportunity: Leverage new spatiotemporal adversarial models (gangs, criminals)*
- Coalitions and uncertainty
  - *Previous work: coalitions, uncertainty, adversarial settings studied separately*
  - *Opportunity: Unified approach leveraging algorithmic advances on each frontier*

Technical Approaches: Overview

- Scalable behavioral game theory
  - *New quantitative models of boundedly rational players*
  - *Algorithms for massive scale games with models of bounded rationality*

- Spatiotemporal game theory
  - *Model acting in geographical space & time, on networks*
  - *Model games over populations*
  - *Solution approaches for equilibrium computation*

- Stochastic coalitional game theory *(not covered today)*
  - *Algorithms for player (mis)coordination, coalitions*
  - *Algorithms to address imperfect observation, execution, payoff noise*

Scientific Accomplishments to Date: What have we learned

- Background: Stackelberg (leader follower) games and networks
- Scalable algorithms for Stackelberg games on networks
  - *Rational follower*
  - *Boundedly rational follower*
- Spatio-temporal game theory
  - *Populations*
  - *Geographical space and time*

All of these research topics closely interact with each other

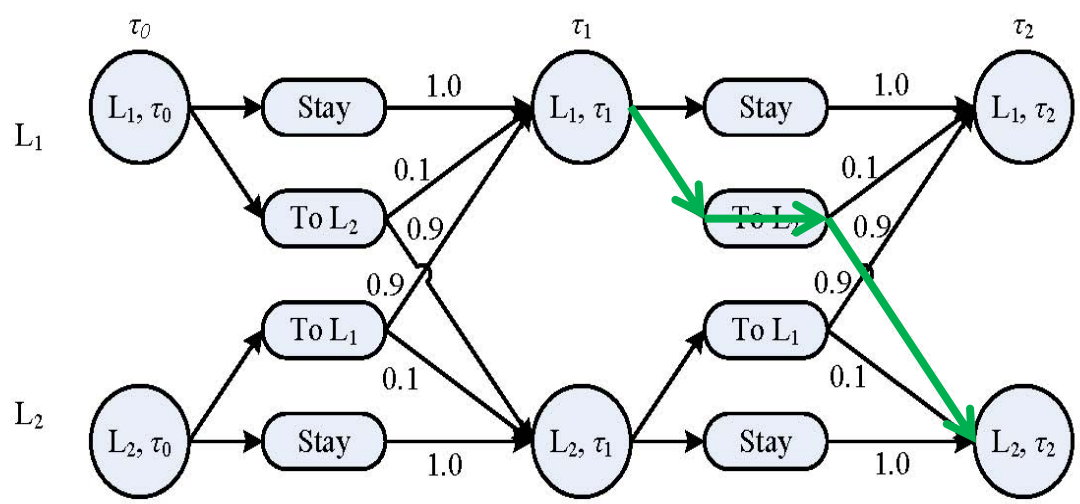
Stackelberg Games on Networks

- ✓ Stackelberg games
  - ✓ Leader: Commits to a mixed strategy
  - ✓ Follower: Chooses a strategy after observing leader's mixed strategy
  - ✓ Strong Stackelberg Equilibrium
    - *Leader commits to the optimal strategy assuming follower will:*
      - *Choose best response*
      - *Break ties in the leader's favor*

Stackelberg Games on Networks

- ✓ Stackelberg games
  - ✓ Leader: Commits to a mixed strategy
  - ✓ Follower: Chooses a strategy after observing leader's mixed strategy
  - ✓ Strong Stackelberg Equilibrium
    - *Leader commits to the optimal strategy assuming follower will:*
      - *Choose best response*
      - *Break ties in the leader's favor*

- ✓ Major subclass:
- ✓ Defender-Attacker games on networks
  - Strategies: Subgraphs
  - Network: Dynamic & stochastic
  - E.g., transportation, social networks



Games on networks: illustrative example

Scale-up in Stackelberg Games on Networks: Rational Followers

- ✓ Challenges: Real-world sized networks
  - Number of pure strategies exponential in size of network

Scale-up in Stackelberg Games on Networks: Rational Followers

- ✓ Challenges: Real-world sized networks
  - Number of pure strategies exponential in size of network

- ✓ Approach 1: Column Generation
  - **Theorem**: There exist solutions whose support (# pure strategies with positive probability) is at most T+1, given T # of targets (e.g., nodes in network).
    - *Rational & boundedly rational followers*

Master Problem: LP with *few* pure strategies

⇩ ⇧

Slave Problem: Generates *new* pure strategy

Scale-up in Stackelberg Games on Networks: Rational Followers

✓ Challenges: Real-world sized networks
- Number of pure strategies exponential in size of network

✓ Approach 1: Column Generation
- **Theorem**: There exist solutions whose support (# pure strategies with positive probability) is at most T+1, given T # of targets (e.g., nodes in network).
  - *Rational & boundedly rational followers*

✓ Approach 2: Marginals for mixed strategies
- Size polynomial in # nodes & edges in the network
- When can we use marginals?
  - **Express expected utilities**

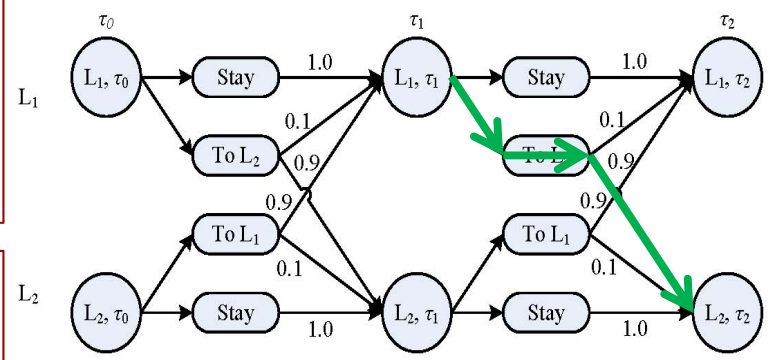
    Theorem: If utilities of game are **separable**, marginals are sufficient to express expected utilities

$$\sum_{i} \sum_{(s_i, a_i, s_i') \in t_i} U^d(s_i, a_i, s_i', \alpha, \lambda)$$

  - **Sample from marginals** (next slide)

Master Problem: LP with *few* pure strategies

Slave Problem: Generates *new* pure strategy



✓ LP using marginals

$$\max_{w,x,u} \sum_{\lambda \in \Lambda} p_\lambda u_\lambda + \sum_{s_i, a_i, s_i'} x_i(s_i, a_i, s_i') R_i(s_i, a_i, s_i')$$

$$x_i(s_i, a_i, s_i') = w_i(s_i, a_i) T_i(s_i, a_i, s_i'), \ \forall s_i, a_i, s_i'$$

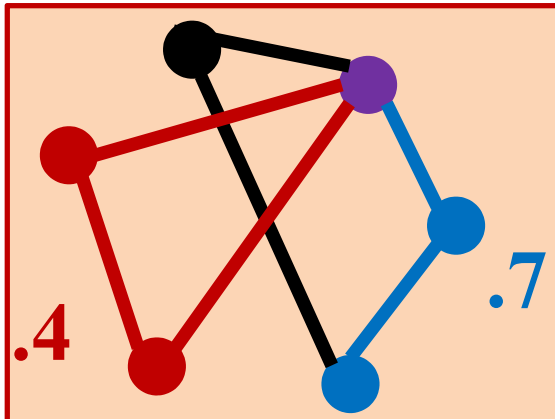
$$\sum_{s_i', a_i'} x_i(s_i', a_i', s_i) = \sum_{a_i} w_i(s_i, a_i), \ \forall s_i$$

$$\sum_{a_i} w_i(s_i^+, a_i) = \sum_{s_i', a_i'} x_i(s_i', a_i', s_i^-) = 1,$$

$$w_i(s_i, a_i) \geq 0, \ \forall s_i, a_i$$

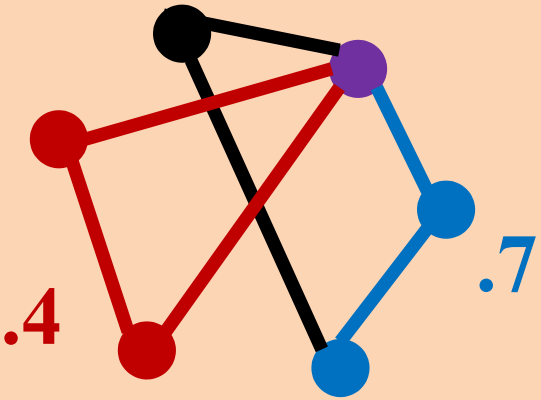
$$u_\lambda \leq x^T U_\lambda^d e_\alpha, \ \forall \lambda \in \Lambda, \alpha \in A$$

Stackelberg Games on Networks: Complexity Results

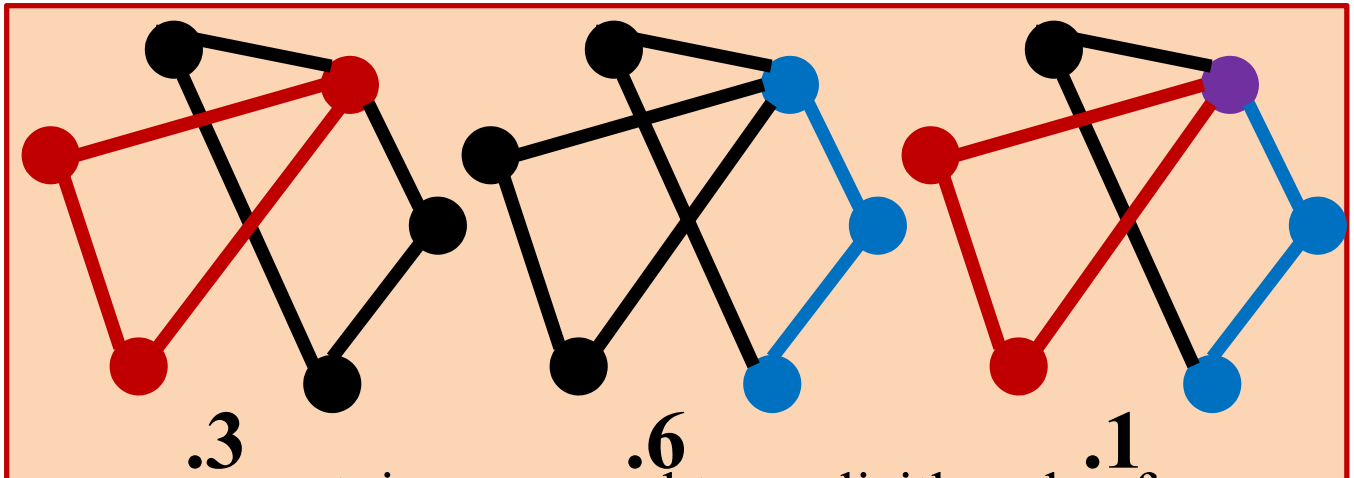


Does it suffice to solve for marginal probabilities only…

Stackelberg Games on Networks: Complexity Results

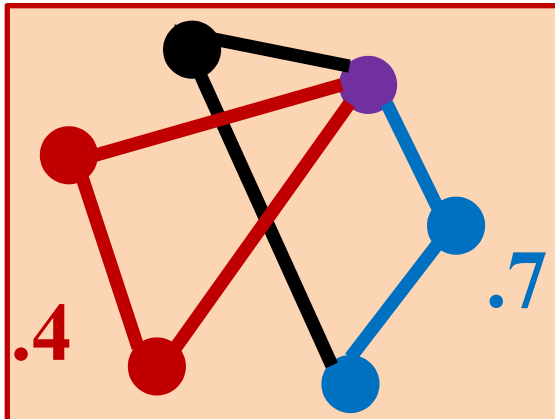


Does it suffice to solve for marginal probabilities only…

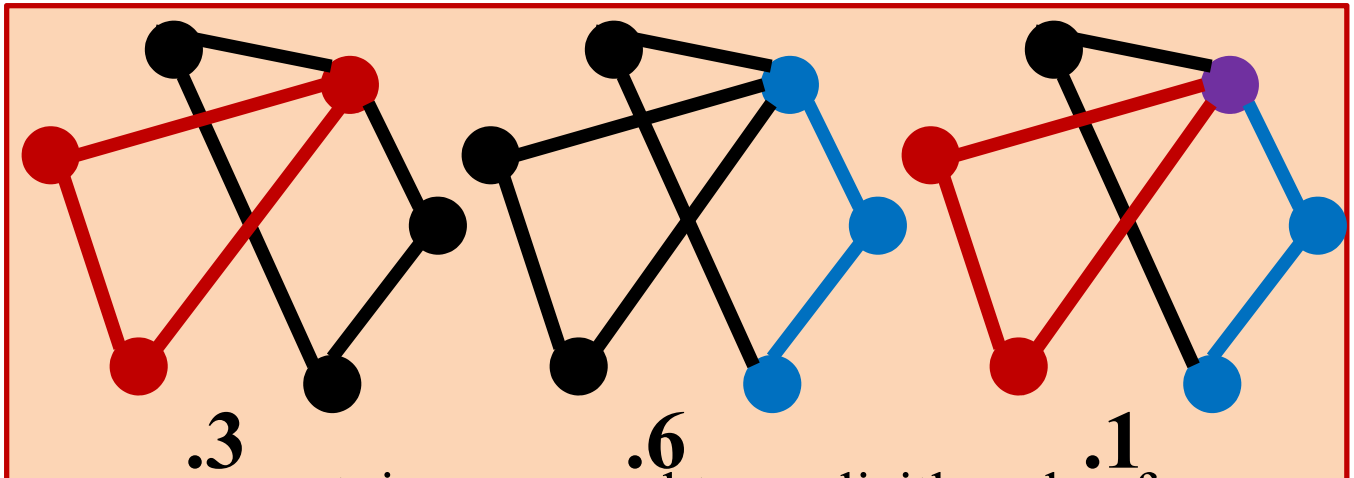


… or not, i.e., we need to explicitly solve for probabilities of full assignments? **(combinatorial explosion in formulation!)**

Stackelberg Games on Networks: Complexity Results



Does it suffice to solve for marginal probabilities only…



… or not, i.e., we need to explicitly solve for probabilities of full assignments? **(combinatorial explosion in formulation!)**

HOMOGENEOUS RESOURCES

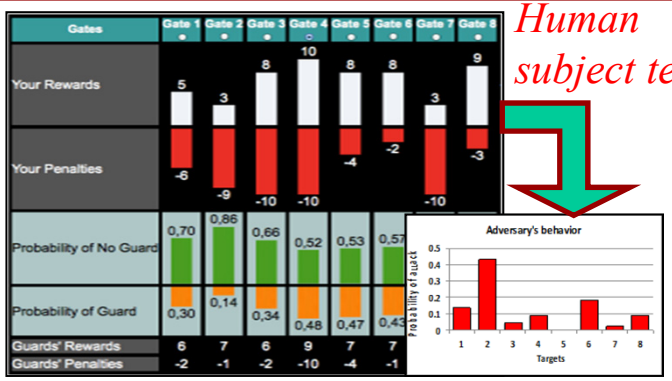
Graph \ Schedule	Edge	Path (start root)	Path (pass through root)	Path (general)
Path	Yes/P (KCP 2010)	Yes/P (Th 1)	Yes*/P (Th 2)	Yes*/P (Th 2)
Tree	Yes/P (KCP 2010)	Yes/P (Th 1)	No/? (Ex 3)	No/NP-h (Ex 3/(LCL 2006))
General	No/P (KCP 2010)	No/NP-h (Ex 2/Th 4)	No/NP-h (Ex 3/Th 4)	No/NP-h (Ex 3/(LCL 2006))

HETEROGENEOUS RESOURCES

Graph \ Schedule	Edge	Path (start root)	Path (pass through root)	Path (general)
Path	No/NP-h (Ex 1/Th 3)	Yes/P (Th 1)	?	No/NP-h (Ex 1/Th 3)
Tree	No/NP-h (Ex 1/Th 3)	Yes/P (Th 1)	No/NP-h (Ex 3/Th 5)	No/NP-h (Ex 3/(LCL 2006))
General	No/NP-h (Ex 1/Th 3)	No/NP-h (Ex 2/Th 4)	No/NP-h (Ex 3/Th 5)	No/NP-h (Ex 3/(LCL 2006))

Proof techniques: linear programming, Birkhoff-von Neumann theorem [1946] & its generalizations (e.g., [Budish et al. AER 2013]), NP-hardness reductions, …

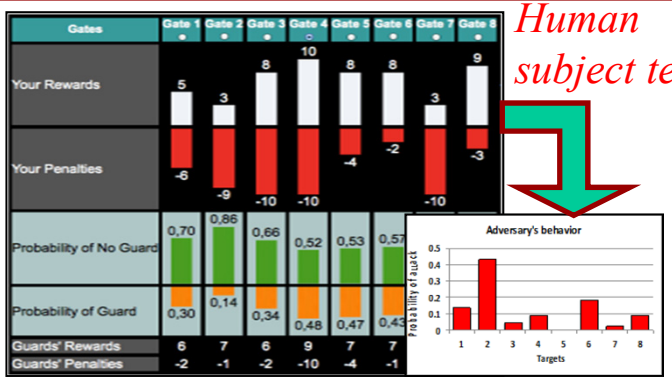
Scale-up of Stackelberg games: Boundedly Rational Followers



Human subject tests

Scale-up of Stackelberg games: Boundedly Rational Followers

- ✓ Challenges: Algorithms given behavioral models
  - Non-convex optimization given QR model of adversary



Human subject tests

- ✓ Quantal Response (QR)

$$q_t = \frac{e^{\lambda U_t^a}}{\sum_{t'=1}^{T} e^{\lambda U_{t'}^a}} = \frac{e^{\lambda (x_t P_t^a + (1-x_t) R_t^a)}}{\sum_{t'=1}^{T} e^{\lambda (x_{t'} P_{t'}^a + (1-x_{t'}) R_{t'}^a)}}$$

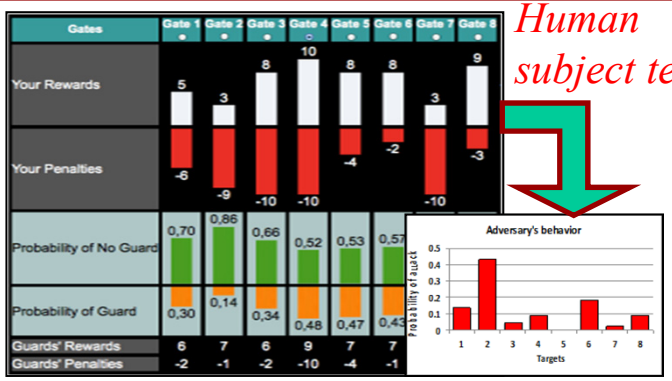
$$\max_x \sum_{t=1}^{T} \frac{e^{\lambda (x_t P_t^a + (1-x_t) R_t^a)}}{\sum_{t'=1}^{T} e^{\lambda (x_{t'} P_{t'}^a + (1-x_{t'}) R_{t'}^a)}} (x_t R_t^d + (1-x_t) P_t^d)$$

Scale-up of Stackelberg games: Boundedly Rational Followers

- ✓ Challenges: Algorithms given behavioral models
  - Non-convex optimization given QR model of adversary

- ✓ Approach 1: Binary search + linear approximation
  - **Theorem**: Let x^* be defender strategy computed, p^* global optimal defender expected utility

$$0 \le p^* - Obj_{P1}\left(x^*\right) \le 2C_1 \frac{1}{K} + \left(C_2 + 1\right)\varepsilon$$



Human subject tests

- ✓ Quantal Response (QR)

$$q_t = \frac{e^{\lambda U_t^a}}{\sum_{t'=1}^{T} e^{\lambda U_{t'}^a}} = \frac{e^{\lambda(x_t P_t^a + (1-x_t) R_t^a)}}{\sum_{t'=1}^{T} e^{\lambda(x_{t'} P_{t'}^a + (1-x_{t'}) R_{t'}^a)}}$$

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

Scale-up of Stackelberg games: Boundedly Rational Followers

✓ Challenges: Algorithms given behavioral models

- Non-convex optimization given QR model of adversary

✓ Approach 1: Binary search + linear approximation

- **Theorem**: Let x^* be defender strategy computed, p^* global optimal defender expected utility

$$0 \le p^* - Obj_{P1}\left(x^*\right) \le 2C_1 \frac{1}{K} + \left(C_2 + 1\right)\varepsilon$$

✓ Approach 2: Cutting-plane

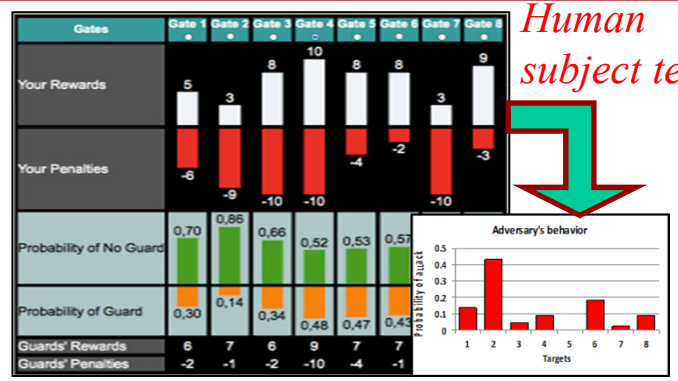
- **Separation Oracle**:

  Lemma 1: The hyper-plane generated by the separation oracle is a ***deep cut*** touching the feasible convex hull X_f

- **Heuristic**: Use gradient of the objective function

$$\sum_{i \in T} z_i \quad \rightarrow \quad \sum_{i \in T} \left(\nabla_i F\left(\tilde{x}\right) + \xi\right) z_i$$

  Lemma 2: A marginal $\tilde{x}$ is feasible iff the minimum of the corresponding Weighted-Separation Oracle LP is zero



Human subject tests

✓ Quantal Response (QR)

$$q_t = \frac{e^{\lambda U_t^a}}{\sum_{t'=1}^{T} e^{\lambda U_{t'}^a}} = \frac{e^{\lambda\left(x_t P_t^a + (1-x_t) R_t^a\right)}}{\sum_{t'=1}^{T} e^{\lambda\left(x_{t'} P_{t'}^a + (1-x_{t'}) R_{t'}^a\right)}}$$

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

✓ Separation Oracle LP

$$\min_{a,z} \quad \sum_{i \in T} z_i$$

$$s.t. \quad z + Aa \ge \tilde{x}; \quad z - Aa \ge -\tilde{x}$$

$$-BAa \ge -b$$

$$\sum_{A_j \in A} a_j = 1, \quad a_j \ge 0, \quad \forall A_j \in A$$

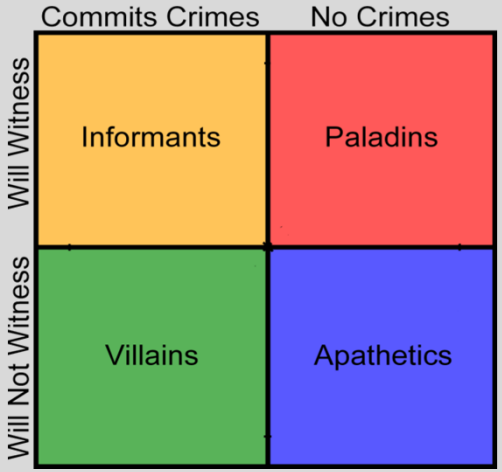
Outline

- Background: Stackelberg (leader follower) games and networks
- Scalable algorithms for Stackelberg games on networks
  - *Rational follower*
  - *Boundedly rational follower*
- Spatio-temporal game theory
  - *Populations*
  - *Geographical space and time*

Understanding Human Play in Adversarial Games

- ✓ Challenges: Understand how players coordinate on seemingly sub-optimal strategies, e.g., costly punishment

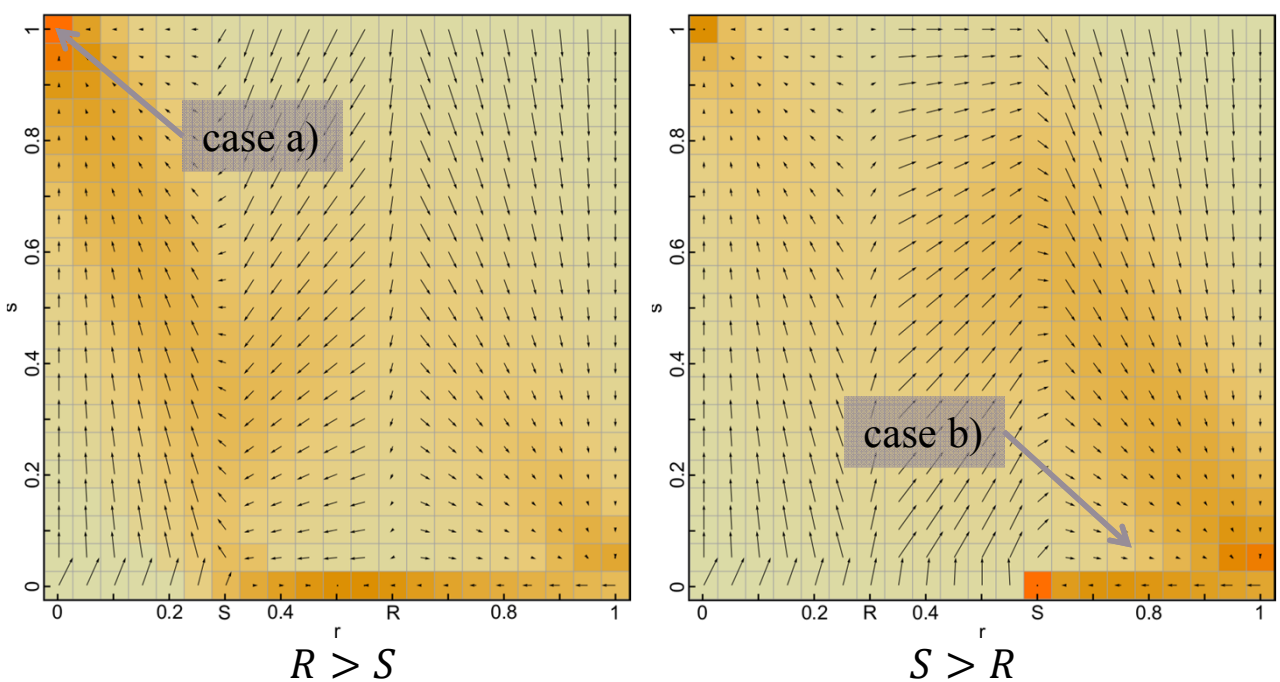
✓ Game Strategies



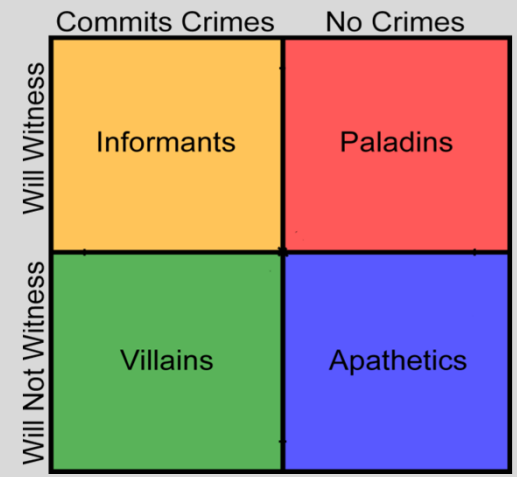
Understanding Human Play in Adversarial Games

- ✓ Challenges: Understand how players coordinate on seemingly sub-optimal strategies, e.g., costly punishment

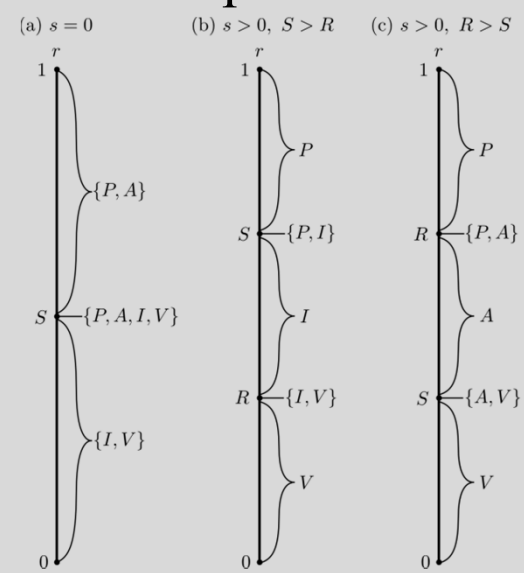
- ✓ Approach: Evolutionary strategy adjustment based on best-response function, parameters S and R
  - **Theorem**: Under best-response dynamics, there are two possible long-term behaviors
    a) coordination on the non-cooperative state $r = 0, s = 1$
    b) coordination on a cooperative state with $r \gtrsim S > R$



- ✓ Game Strategies



- ✓ Best Response Function

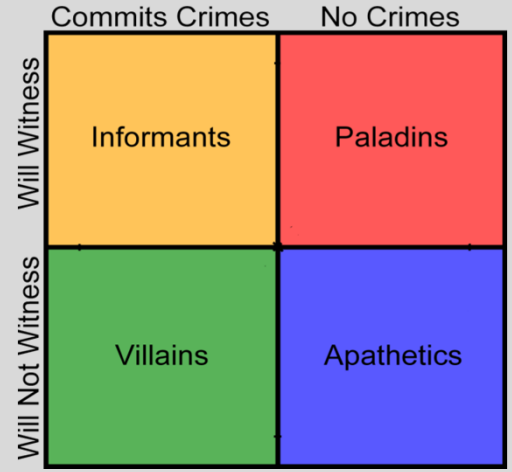


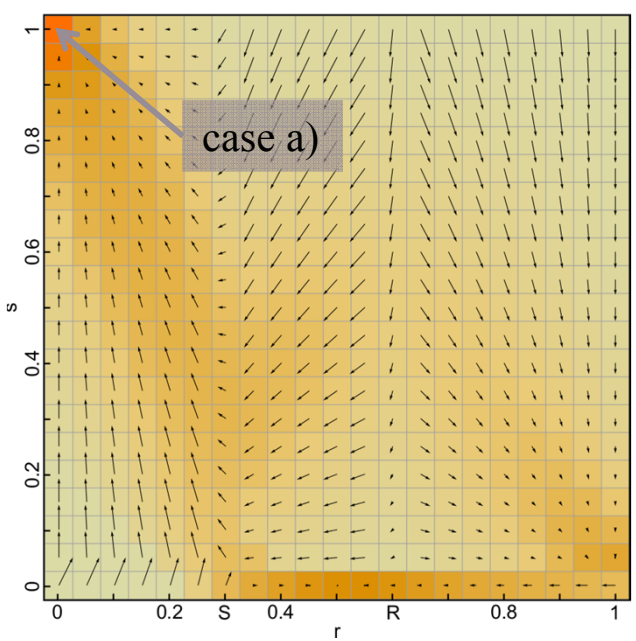
Understanding Human Play in Adversarial Games

- ✓ Challenges: Understand how players coordinate on seemingly sub-optimal strategies, e.g., costly punishment

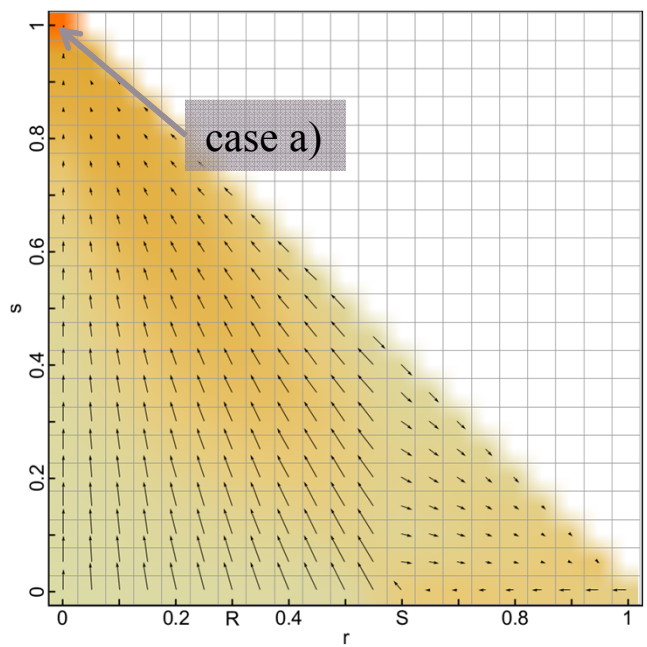
- ✓ Approach: Evolutionary strategy adjustment based on best-response function, parameters S and R
  - **Corollary**: If strategy I is removed, only case a) is possible

✓ Game Strategies



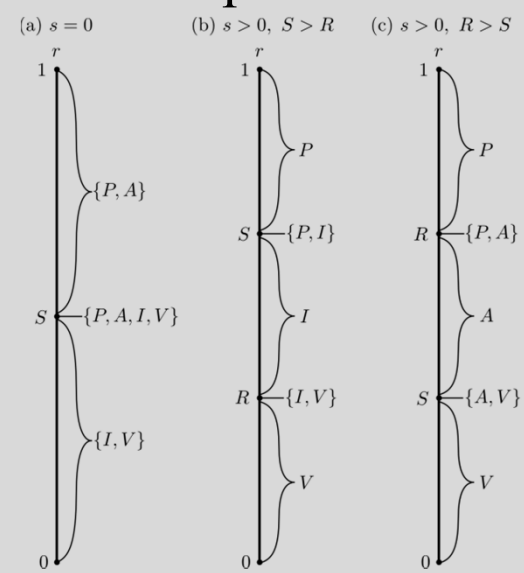


$R > S$



$S > R$, no I

✓ Best Response Function



Cops on the Dots: Instantaneous Optimal Defender Strategy

✓ Challenges: Understand optimal placement of police in residential burglary model.

- APPROACH: well-known burglary model from UCLA group – add policing as a deterrence

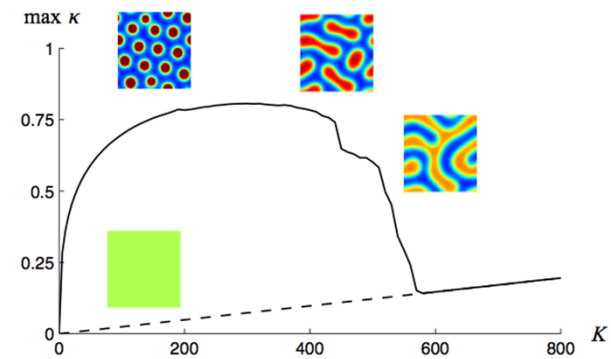
$$\frac{\partial A}{\partial t} = \eta \triangle A + d(\kappa)\rho A - A + A_0,$$
$$\frac{\partial \rho}{\partial t} = \nabla \cdot (\nabla \rho - 2\rho \nabla \log A) - d(\kappa)\rho A + d(\kappa)\overline{B},$$
$$\kappa = \arg\min \left\{ \int_\Omega d(k)\rho A \, dx : k \in L^1(\Omega), k \geq 0, \int_\Omega k \, dx = K \right\}.$$

Theorem 2.2. *Let ρA be non-negative everywhere, positive on a set of positive measure, and continuous. Then there exists $\lambda > 0$ such that*

$$\kappa = (d')^{-1}(-\lambda/\rho A)\chi_{U(\lambda,\rho A)}, \tag{11}$$

where $U(\lambda, f) = \{x \in \Omega : f(x) > -\lambda/d'(0)\}$. This λ is the Lagrange multiplier associated to the L^1 constraint (10) and as such is the unique solution to the nonlinear equation

$$G(\lambda) := \int_{U(\lambda,\rho A)} (d')^{-1}(-\lambda/\rho A) dx = K. \tag{12}$$



Thm: For all choices of parameters and deterrence functions there exists a K* such that the system with uniform crime rates is linearly stable for all K>K*.

✓ Despite the rigorous theory for linear stability, numerics show a complicated bifurcation structure of solutions for lower levels of K <K* - often the case in real world problems with limited police resources.

Contagion Theory for Evacuation

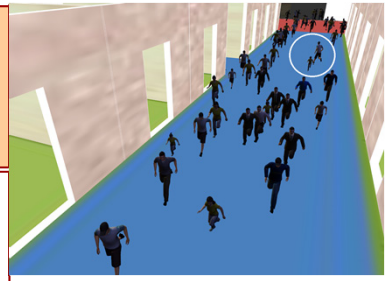
✓ Challenges: Predict compression wave in evacuation scenario.

- APPROACH: 1D particle model based on ASCRIBE model used by USC group Tsai et al. Develop rigorous theory in continuum limit.

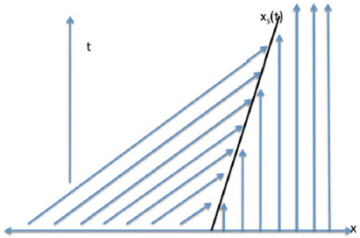
$$\dot{q}_i = \gamma(q_i^* - q_i), \quad q_i^* = \sum_{j,|x_j - x_i|<R} q_j, \quad \dot{x}_i = q_i.$$

THEOREM: (shock formation, speed prediction) Given an initial configuration moving to the right with particles of average density ρ_L for $x<0$ and particles of average density ρ_R for $x>0$ and respective fear levels $q_L>q_R$, the solution to the problem at later times is a singular shock with speed:

$$s^* = \frac{\sqrt{\rho_L} q_L + \sqrt{\rho_R} q_R}{\sqrt{\rho}_R + \sqrt{\rho}_L}$$



ESCAPES(early work)



Compression wave

✓ "Trampling" rate of agents

$$\frac{dM_\rho}{dt} = s^*[\rho] - [Q]$$

Bracket denotes jump across the shock. We measure both the jump in the density and the jump in the fear level specified at time zero.

✓ Theory related to pressureless gas dynamics and sticky particles. New theory for nonlocal effects from the ASCRIBE model.

Potential Breakthroughs

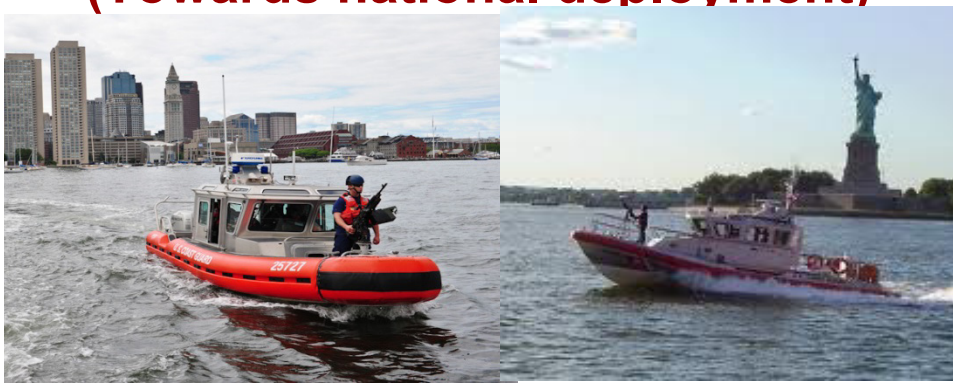
- First Stackelberg solution method to scale to million node networks
- First results on complexity for Stackelberg games on networks
- Algorithms for robustness to adversary without behavioral data
- First Algorithmic Experimental game theory results:
  - *Incorporate Quantitative Behavioral Models into Stackelberg Games*
- New Spatiotemporal Game-Theoretic Models:
  - *Combining predictive crime models with game theoretic approaches*
- First Macroscopic multidimensional models for Contagion Prediction
- Quantitative Metrics to evaluate Deterrence values of alternate strategies
- First game theoretic model to study recidivism of criminal offenders and optimize resources for punishment vs. rehabilitation.
- First experimental results on evolution of crime in adversarial game

Why is this an important area

- Global, National, Hybrid and Local Security Threats in Defense, Economics, Crime and Cyber environments are growing:
  - *Larger in scale (more entities over wider areas)*
  - *Faster in action (time to organize malicious actions is shorter)*
  - *Computationally sophisticated (in observation and coordination)*
- The resources we have to address these issues are bounded and also growing harder to manage manually.
- We need to develop computational theories, models and methods that can help support and improve our security infrastructure in all domains.
- To do this effectively, we need to extend the science of game theory and in particular Stackelberg games to support to address the scale, stochasticity and sophistication of the adversaries we will be facing in the 21st century.

Transitions: Game Theory for Security (2013)

PROTECT: In use by the US Coast Guard (Towards national deployment)



PROTECT: patrolling the ports of Boston and New York

- ✓ Challenges: Spatial and Temporal Constraints; Coordination between boats and helicopters

PROTECT – Staten Island ferry: 60,000 passengers a day

- ✓ Challenges: Continuous dimensions; Scalability

TRUSTS: currently evaluated by TSA and Los Angeles Sheriff's Department

Fare Evasion Trials: Daily tests on the LA metro system

- ✓ Challenges: Robustness against uncertainty; Game Theory vs Uniform Random

Full Scale Exercise: 23 Teams patrolling the Los Angeles Metro System for 12 Hours

- ✓ Challenges: Coordination of the different teams; Robustness of the schedules

Budget

Period	Fiscal Yr.	Duration	Amount
Base	2011	4 months	$520,833
	2012	12 months	$1,250,000
	2013	12 months	$1,250,000
	2014	8 months	$729,167
	Total	36 months	$3,750,000
Option	2014	4 months	$520,833
	2015	12 months	$1,250,000
	2016	8 months	$729,167
	Total	24 months	$2,500,000

Dates, locations, overall results of major reviews or meetings

- **Kickoff Meeting** (Sept 2011, Los Angeles)
  - *Key guidance: Good start. Foster collaboration across universities*
  - *Response:*
    - Regular meetings in LA area, phone calls, joint papers

- **Year 2 Meeting** (Sept 2012, Los Angeles)
  - *Key guidance: Good progress. Foster collaboration and basic research; non-collaborating Co-PI should not be part of MURI.*
  - *Response:*
    - One specific Co-PI no longer part of MURI;
    - New collaborations with joint USC-UCLA papers

- **Year 3 Meeting** (Nov 2013, Los Angeles)

Thank you

- Milind Tambe
- tambe@usc.edu

Territorial Developments based on graffiti: a statistical mechanics approach

Challenges: To construct a lattice model for two adversarial gangs interacting only via graffiti fields

Our 2D lattice Hamiltonian – η gangs, g graffiti:

$$-\mathscr{H}(\mathbf{s}) = J\sum_{\langle i,j\rangle}\eta_i g_j + K\sum_i \eta_i g_i + \alpha\sum_i \eta_i^2 - \lambda\sum_i g_i^2,$$

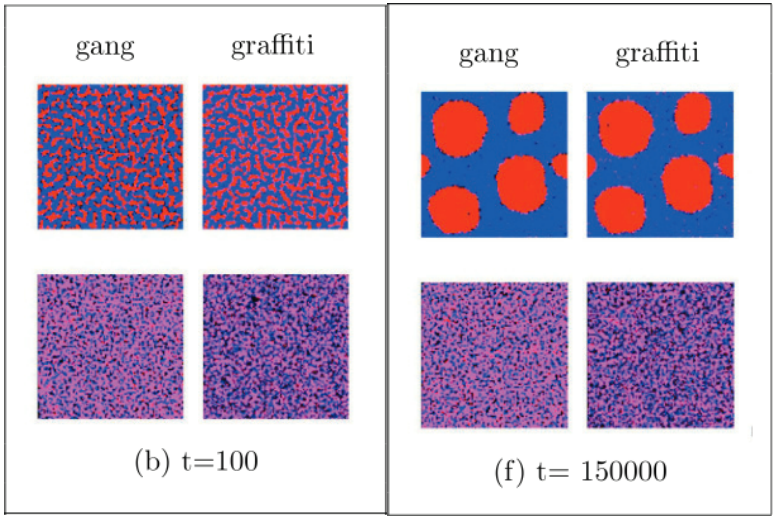
Couplings: red and blue gang/graffiti (offisite J; onsite K) , graffiti decay λ, gang proliferation α

Each site η=(−1 red, 0 none, 1 blue) gang
g > 0 blue, g < 0 red graffiti
b = fraction occupied, n = excess red/blue G = graffiti imbalance

Theorem: A phase transition exists between disordered "well mixed" gangs and ordered, "clustered" ones. The separatrix is at J_{cr}. The transition between phases is first order for $\alpha<\alpha_{cr}$ and continuous otherwise. J_{cr} and α_{cr} are

$$\alpha_{cr} = -2\ln 2$$

$$J^2 = \frac{\lambda}{e^{-\alpha}+2} = J_{cr}^2$$



(b) t=100 (f) t= 150000

Upper rows: low λ
long term graffiti clustering

Lower rows: high λ
short term graffiti no clustering

Mean Field Equations:

$$-\frac{\partial\Phi}{\partial G} = Jn - 2\lambda G = 0,$$

$$-\frac{\partial\Phi}{\partial b} = \alpha + \log(1-b) - \frac{1}{2}\log\left(\frac{b^2-n^2}{4}\right) = 0,$$

$$-\frac{\partial\Phi}{\partial n} = JG - \frac{1}{2}\log\left(\frac{b+n}{b-n}\right) = 0.$$