Multiagent reasoning for social impact: Results from deployments for public health and conservation

MILIND TAMBE
Director, Ctr for Research on Computation & Society
Harvard University

Director "AI for Social Good"
Google Research India

@MilindTambe_AI
AI & Multiagent Systems Research for Social Impact

Public Health

Conservation

Public Safety and Security

Optimize Our Limited Intervention Resources
Optimizing Limited Intervention Resources

Public Health
- Social Networks & Bandits

Conservation
- Green security games

Public Safety & Security
- Stackelberg security games
Three Common Themes
Non-profits, Data-to-deployment pipeline, Multiagent systems

Empower non-profits to use AI tools; avoid being gatekeepers to AI4SI technology
Three Common Themes
Non-profits, Data-to-deployment pipeline, Multiagent systems

- Social networks
- Restless bandits
- Game theory

Immersion
- Data Collection

Predictive model
- Machine Learning

Prescriptive algorithm
- Multiagent Reasoning Intervention
Three Common Themes
Multiagent systems, Data-to-deployment pipeline, Interdisciplinary partnerships

Field test & deployment: Social impact is a key objective
Outline

Public Health

- Health information dissemination: Social networks
- Health program adherence: ML & Bandits
- COVID-19: Agent-based modeling

Conservation

- Cover papers from 2017-now [AAMAS, AAAI, IJCAI, NeurIPS...]
- PhD students & postdocs highlighted
Information dissemination & behavior change
Optimizing Limited Intervention (Social Worker) Resources

Prevent HIV in youth experiencing homelessness: HIV 10x housed population

- **Shelters**: Limited number of peer leaders to spread HIV information in social networks
- “Real” face-to-face interactions; not Facebook etc
Influence Maximization in Social Networks

- **Given:**
  - Social network Graph G
  - Choose K “peer leader” nodes
  - Assume: Independent cascade model of information spread

- **Objective:**
  - Maximize expected number of influenced nodes
Influence Maximization in Social Networks
Three Key Challenges Combined Together

- Uncertainty in propagation probability over edges
- Multi-step dynamic policies to handle peer leader “no shows”
- Unknown social network, limited query budget to SAMPLE network
Challenge 1: Uncertainty in Real-world Physical Social Networks

\[ P(C,D) = 0.4 \]

\[ \mu = 0.5 \]

\[ \mu \in [0.3, 0.7] \]
Robust Influence Maximization
(AAMAS 2017)

- Worst case parameters: a zero-sum game against nature

\[
\text{max}_{x \in \Delta^{|P|}} \text{min}_{\mu, \sigma} \sum x_p \frac{\text{Outcome}(p)}{\text{OPT}(\mu, \sigma)}
\]

**Algorithm**
Choose Peer Leaders \( p \in P \)
generating mixed strategy
“\( x \in \Delta^{|P|} \)”

vs

**Nature**
Chooses parameters \( \mu, \sigma \)
HEALER Algorithm
Robust Influence Maximization
(AAMAS 2017)

Theorem: Converge with approximation guarantees

- Equilibrium strategy despite exponential strategy spaces: Double oracle

<table>
<thead>
<tr>
<th>Nature’s oracle</th>
<th>Influencer’s oracle</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Influencer</strong></td>
<td></td>
</tr>
<tr>
<td>Policy #1: 0.8, -0.8</td>
<td>Policy #1: 0.8, -0.8</td>
</tr>
<tr>
<td>Policy #2: 0.7, -0.7</td>
<td>Policy #2: 0.7, -0.7</td>
</tr>
<tr>
<td>Policy #3: 0.6, -0.6</td>
<td>Policy #3: 0.6, -0.6</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th><strong>Params #1</strong></th>
<th><strong>Params #2</strong></th>
<th><strong>Params #3</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td>Policy #1</td>
<td>0.8, -0.8</td>
<td>0.3, -0.3</td>
</tr>
<tr>
<td>Policy #2</td>
<td>0.7, -0.7</td>
<td>0.5, -0.5</td>
</tr>
<tr>
<td>Policy #3</td>
<td>0.6, -0.6</td>
<td>0.4, -0.4</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th><strong>Params #1</strong></th>
<th><strong>Params #2</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td>Policy #1</td>
<td>0.8, -0.8</td>
</tr>
<tr>
<td>Policy #2</td>
<td>0.7, -0.7</td>
</tr>
<tr>
<td>Policy #3</td>
<td>0.6, -0.6</td>
</tr>
</tbody>
</table>

Date: 5/2/2021
Challenge 2: Contingency-Aware Influence Maximization
(AAMAS 2018a)

K = 4
1st time step

K = 4
2nd time step

HIDDEN STATE
Action
Choose nodes

POMDP Policy

Observation: Node presence

Partition POMDPS:
Exploit community structure

Date: 5/2/2021
Theorem: For community-structured graphs(*), sampling algorithm obtains a constant-factor approximation to the optimal influence spread using polylog(n) queries.

Data collection costly

Query 15% nodes

Sampling Algorithm
Sample node randomly & estimate size of its community; Choose seeds from largest K communities

- Query 15% of nodes in the population
- Output $K$ peer leader nodes to spread influence
- Perform similar to $OPT$, best influence spread with full network

(*)Community structured: drawn from a stochastic block model
“Sampling-HEALER”
Pilot tests with Homeless Youth
(IJCAI 2018)

Sampling-HEALER

Network Sampling

Robust multi-step policy

Peer leader selection

Observe peer leaders present/absent

12 peer leaders

<table>
<thead>
<tr>
<th>Sampling HEALER (Sampled Network)</th>
<th>HEALER (Full Network)</th>
<th>DEGREE CENTRALITY (Full Network)</th>
</tr>
</thead>
<tbody>
<tr>
<td>60 youth</td>
<td>62 youth</td>
<td>55 youth</td>
</tr>
</tbody>
</table>

Date: 5/2/2021
Results: Pilot Studies

Percent of non-Peer Leaders

- Informed
- Not Informed

Date: 5/2/2021
Results of 750 Youth Study [with Prof. Eric Rice]
Actual Change in Behavior?
(AAAI 2021)

First large-scale application of influence maximization for public health

---

**Reduction in condomless anal sex (1 month)**

- **SAMPLING HEALER**: 35%
- **Degree centrality**: 10%
- **Control**: 20%

**Reduction in condomless anal sex (3 months)**

- **SAMPLING HEALER**: 30%
- **Degree centrality**: 25%
- **Control**: 15%
Results of 750 Youth Study [with Prof. Eric Rice]

Reduction in condomless vaginal sex (3 months)

- SAMPLING HEALER
- Degree centrality
- Control
What our collaborators are saying:
Next Steps: Fairness in Influence Maximization
(NeurIPS 2019, IJCAI 2019, AAAI 2021)

Robust graph covering, maximize worst case coverage $\rightarrow$ Disparity

Maxmin fairness: NeurIPS2019
$$\min_{c \in C} u_c(A) \geq \gamma$$
$\gamma$: Max of minimum utility for any community

Diversity constraints: IJCAI2019
$$u_c(A) \geq U_c$$
$U_c$: Constraint from cooperative game theory

Inequity aversion: AAAI 2021
$$W_\alpha(u(A))$$
$\alpha$: controls fairness tradeoff; policymaker has choice
Next steps: RL for Influence Maximization in Social Networks
(with B. Ravindran & team, AAMAS 2020)

Network Family | Improve %
---|---
Rural | 23.76
Animal | 26.6
Retweet | 19.7
Homeless | 7.91
Outline

Public Health

- Health information dissemination: Social networks
- Health program adherence: ML & Bandits
- COVID-19: Agent-based modeling

Conservation
Intervention Reasoning: Active Adherence Monitoring

- Immersion
  - Data Collection
- Predictive model
  - Learning
- Prescriptive algorithm
  - Restless bandit
- Field tests & deployment
Intervention Reasoning: Active Adherence Monitoring

- **Immersion**
  - Data Collection

- **Predictive model**
  - Learning

- **Prescriptive algorithm**
  - Restless bandit

- **Field tests & deployment**
Health Program Adherence
Maternal & Child Care in India
(IJCAI 2021)

Woman dies in childbirth every 15 min; 4 of 10 children too thin/short

- 18 Million women
- Weekly 3 minute call to new/expecting moms
- mMitra: Significant benefits
  - 2.2 million women enrolled

➤ Unfortunately, significant fraction low-listeners or drop-outs
Passive Adherence Monitoring Maternal & Child Care in India

(With B Ravindran IIT Madras)

Classifier to predict beneficiaries drop out? So ARMMAN can focus interventions

- Results of pilot with 18000 beneficiaries: High precision, recall, accuracy
- Field trial with 8000 beneficiaries: Call intervention helps

- Prediction software in use to help 300,000 beneficiaries in mMitra
Passive Adherence Monitoring
Preventing Tuberculosis in India
(KDD 2019)

Tuberculosis (TB): ~500,000 deaths/year, ~3M infected in India

- Predict adherence risk from phone call patterns for early intervention?

TB Treatment
6 months of pills

Track adherence
via daily phone calls

Results Mumbai, India:
15,000 patients, 1.5 Million calls
Intervention Reasoning: Active Adherence Monitoring

- Immersion
  - Data Collection

- Predictive model
  - Learning

- Prescriptive algorithm
  - Multiagent Reasoning Intervention

- Field tests & deployment
Intervention Scheduling with Scarce Data: Active Adherence Monitoring
(NeurIPS 2020)

Challenge:
- Large number of patients \( N \)
- Can only call \( K \) patients per day.
- Which \( K \)?

Approach: Restless bandit
- Each arm (patient): POMDP
- Each POMDP: binary latent state \{0, 1\}
  - 0 = not-adhering; 1 = adhering

Goal: Policy for \( K \) patients to call per day

Photo Credit: IntraHealth International (CC BY-NC-SA 3.0 via https://www.intrahealth.org/)
Intervention Scheduling with Scarce Data: Collapsing Bandits (NeurIPS 2020)

Theorem (Whittle Index): Collapsing bandits are Indexable if threshold policies are optimal.

When arm not played (patient not called)
- No observation
- Instead, compute belief of adherence

When arm played: Uncertainty collapse
- Observe current state
- Belief of adhering next round

Exploit “collapsing” for fast algorithm: Fixed number of belief states
New Fast Algorithm: Collapsing Bandits

- Orders of magnitude speedup with no solution quality loss
- ORANGE = Best baseline
- Blue = Our model

![Graph showing runtime and solution quality](image)

Date: 5/2/2021
New Directions in Restless Bandits

**Fast algorithms for extending to:**

- **Multiple action types**  
  (AAMAS 2021a)

- **Risk aware restless bandits**  
  (AAMAS 2021b)

**Online learning**

- **Learning policies via Index Q-Learning**  
  (AAMAS 2021c)
New Directions in Restless Bandits

Restless bandits for intervention:

- 20000 subject trial
Outline

Public Health

- Health information dissemination: Social networks
- Health program adherence: ML & Bandits
- COVID-19: Agent-based modeling

Conservation
COVID-19: Agent-based Simulation Model

RESEARCH ARTICLE

Modeling between-population variation in COVID-19 dynamics in Hubei, Lombardy, and New York City

Bryan Wilder, Marie Charpignon, Jackson A. Killian, Han-Ching Ou, Aditya Mate, Shahin Jabbari, Andrew Perrault, Angel N. Desai, Milind Tambe, and Maimuna S. Majumder

PNAS October 13, 2020 117 (41) 25904-25910, first published September 24, 2020, https://doi.org/10.1073/pnas.2010551117

Tracking disease outbreaks from sparse data with Bayesian inference

Bryan Wilder, Michael Mina, Milind Tambe

1 John A. Paulson School of Engineering and Applied Sciences, Harvard University
2 T.H. Chan School of Public Health, Harvard University

bwilder@g.harvard.edu, mmrna@hsph.harvard.edu, milind_tambe@harvard.edu
COVID Testing Policy: Accuracy vs Ease
(Science Advances, 2020) with Prof. Michael Mina

- Range of tests entering market, varying sensitivity/cost: Quantity vs Quality?
  - qRT-PCR ("gold standard"): Detect viral concentration of $10^3$/mL, $50-100
  - Antigen strip ("Less sensitive"): $10^6$/mL, $3-5

Rapid turnaround time & frequency more critical than sensitivity for COVID-19 surveillance

<table>
<thead>
<tr>
<th>Less sensitive; Cheap &amp; fast turnaround</th>
<th>More sensitive; Costly &amp; slow turnaround</th>
</tr>
</thead>
</table>

<table>
<thead>
<tr>
<th>Total infections</th>
<th>Total infections</th>
<th>Total infections</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

- Rapid turnaround time & frequency more critical than sensitivity for COVID-19 surveillance.
- qRT-PCR ("gold standard"): Detect viral concentration of $10^3$/mL, $50-100
- Antigen strip ("Less sensitive"): $10^6$/mL, $3-5
COVID Testing Policy: Impact

- Covered in NYT, WaPo, Time, The Atlantic, The Hill, etc
- Allowed epi collaborators to advocate to FDA/CDC
Outline

Public Health

Conservation

- Protect wildlife, forests, fisheries: Game-focused learning
- Integrating real time data for protection: Signaling games
Protecting Conservation Areas: Green Security Games
(IJCAI 2015)

Snare or Trap

Wire snares

Date: 5/2/2021
From Stackelberg Security Games to Green Security Games
(IJCAI 2015)

- Stackelberg security games (SSG)
- With boundedly rational poachers
- Learn adversary response model at targets “i”

<table>
<thead>
<tr>
<th></th>
<th>Area1</th>
<th>Area2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Area1</td>
<td>4, -3</td>
<td>-1, 1</td>
</tr>
<tr>
<td>Area2</td>
<td>-5, 5</td>
<td>2, -1</td>
</tr>
</tbody>
</table>
Learning Adversary Response Model: Uncertainty in Observations

- Ranger patrol
- Animal density
- Distance to rivers / roads / villages

Probability of snare Per 1 KM Grid Square

Area habitat
Area slope

Training: Filtered Datasets

Patrol Effort
0
1
2

Predict: Ensemble of Classifiers

C0
C1
C2
PAWS: First Pilot in the Field
(AAMAS 2017)

- Two 9-sq.km areas, infrequent patrols
- Poached elephant
- 1 elephant snare roll
- 10 Antelope snares
PAWS Predicted High vs Low Risk Areas: 3 National Parks, 24 areas each, 6 months
(ECML PKDD 2017, ICDE 2020)

Queen Elizabeth National Park

Murchison Falls National Park

Srepok Wildlife Sanctuary

Date: 5/2/2021
PAWS Real-world Deployment
Cambodia: Srepok Wildlife Sanctuary
(ICDE 2020)

Date: 5/2/2021

2019 PAWS: 521 snares/month

vs

2018: 101 snares/month

2021 PAWS

1,000 snares found in March
PAWS GOES GLOBAL with SMART platform!!

Protect Wildlife
800 National Parks
Around the Globe

Cross River, Nigeria
Sapo, Liberia
Kafue, Zambia
Gonarezhou, Zimbabwe
Limpopo, Mozambique
Srepok, Cambodia
Royal Belum, Malaysia
Direction #1: Integrating Real-Time “SPOT” Information (IAAI 2018)

Goal: automatically find poachers

Date: 5/2/2021
Drone Used to Inform Rangers

- \( \text{Prob(ranger arrives)} = 0.3 \)  [poacher may not be stopped]
- Deceptive signaling to indicate ranger is arriving
Drone Used to Inform Rangers

- $\text{Prob}(\text{ranger arrives}) = 0.3$  [poacher may not be stopped]
- Deceptive signaling to indicate ranger is arriving
Drone Used to Inform Rangers

- $\text{Prob}(\text{ranger arrives}) = 0.3$  [poacher may not be stopped]
- Deceptive signaling to indicate ranger is arriving
- Must be strategic in deceptive signaling

![Image of drone and ranger]

$\text{Prob}(\text{ranger}) = 0.3$
Exploiting Informational Advantage
Defender Knows Pure & Mixed Strategy
(AAAI 2018, AAAI 2020, AAMAS 2021)

Si-G Model: Stackelberg Security Games with Optimal Deceptive Signaling

- Poacher best interest to “believe signal” even if know 50% defender deception
- Recent work used RL for deception policy generation (AAMAS 2021)
Three new directions

Direction #2: Tailor ML Predictions to Ultimate Objective

Minimize \( \sum_{i \in T} q_{\text{empirical}} \log \hat{q} \)

Maximize accuracy in Adversary target values *estimates*

Prescription: Plan patrol Coverage Game Theory

Optimize
Game-Focused Learning: Modifies Loss Function via Downstream Objective
(AAAI 2019, AAAI 2020, NeurIPS19, NeurIPS20, AAMAS20…)

Data Collection

Prediction

Prescription

Field

Prediction: Help maximize
*Defender’s expected utility*

Prescription: Plan patrol
Coverage Game Theory

\[
\sum (1 - p_i(\hat{q})) q_{\text{empirical}}
\]
Another View:
Game-Focused Learning: End-to-End Method

Game-focused gradient descent:

$$\frac{\partial \text{obj}(\text{decision})}{\partial \text{weights}} = \frac{\partial \text{prediction}}{\partial \text{weights}} \frac{\partial \text{prediction}}{\partial \text{decision}} \frac{\partial \text{obj}(\text{decision})}{\partial \text{decision}}$$
Game-Focused Learning: Simulations Murchison Falls National Park

Expected utility vs. Number of training games

- Game-focused
- Two-stage
Direction #3: Data Scarce Parks

**Data-rich parks**: build predictive models to plan patrols

**Data-scarce parks**: conduct patrols to detect illegal activity and collect data to improve the predictive model

Srepor, Cambodia
43,269 patrol observations
2013 – 2018

Royal Belum, Malaysia
824 patrol observations
June – August 2018
**LIZARD: Multiarmed Bandit**
Lipschitz Arms with Reward Decomposability (AAAI 2021)

**Theorem:** Given $N$ targets, Lipschitz constant $L$, and time horizon $T$, the regret bound of the LIZARD algorithm is $\text{Reg}(T) \leq O\left(\frac{L^3 N T^2}{3} (\log T)^{\frac{1}{3}}\right)$:

- **Input:** N Targets with features, T Time, stochastic poacher places snares at targets
- **Output:** Specify patrol effort per target $\leq$ budget B
- Reduce regret wrt $OPT$, optimal patrol effort, for capturing snares

Lizard exploits decomposability, smoothness, monotonicity
Achieving social impact & AI innovation go hand in hand

Data to deployment: Not just improving algorithms, new AI4SI evaluation

Important to step out of the lab and into the field

Embrace interdisciplinary research -- social work, conservation

Lack of data is the norm, a feature; part of the project strategy

Empower non-profits to use AI tools; avoid being gatekeepers to AI4SI tech
Key Collaborators on Papers Referenced
(In the order papers referenced)

- Eric Rice (USC)
- Nicole Immorlica (MSR)
- Yair Zick (UMASS, Amherst)
- Balaraman Ravindran (IIT-Madras)
- Amit Sharma (MSR)
- Maia Majumder (Harvard)

- Michael Mina (Harvard)
- Daniel Larremore (Colorado)
- Andy Plumptre (Cambridge)
- Rohit Singh (WWF)
- Phebe Vayanos (USC)
- Bistra Dilkina (USC)

Collaborate to realize AI’s tremendous potential to
Improving society & fighting social injustice

@MilindTambe_AI