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

CRCS Center for Research on Computation and Society
at Harvard School of Engineering and Applied Sciences

Google Research

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

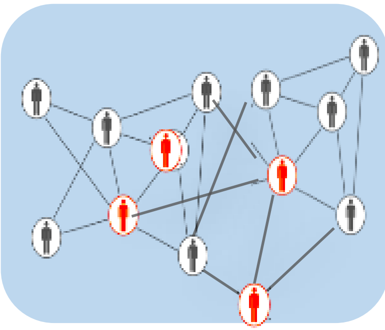
Public Health

Conservation

Public Safety
and Security

Optimize Our Limited Intervention Resources

Lesson #1:
Achieving Social Impact and AI Innovation Go hand-in-hand



Social Networks & Bandits

Public Health

Multiagent Systems Research



Green security games

Conservation

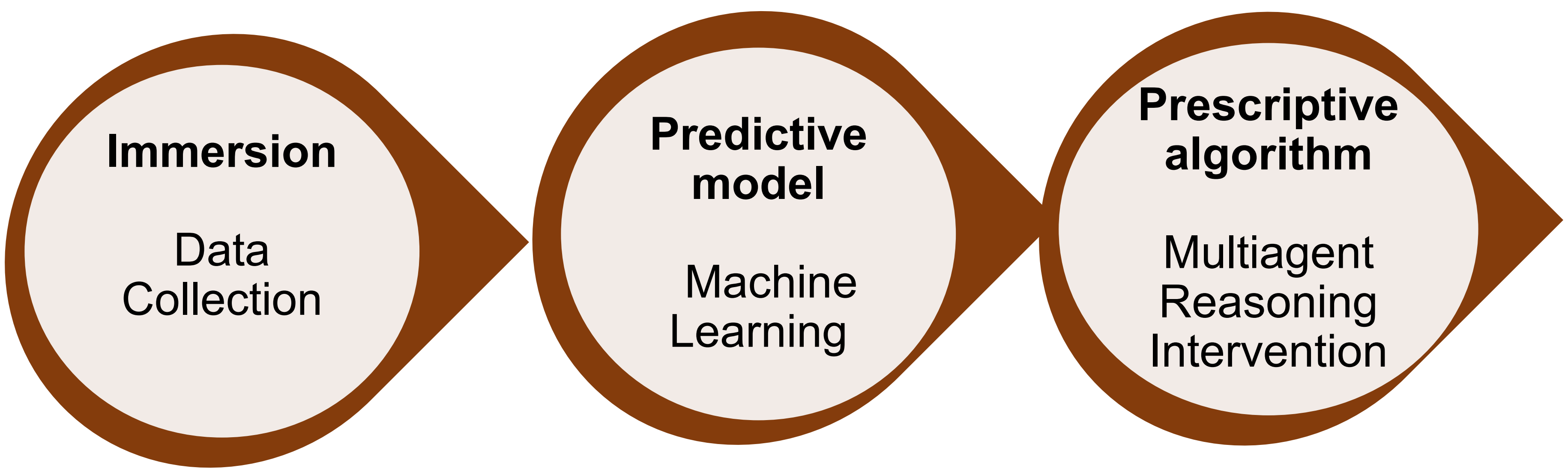
Stackelberg security games

Public Safety & Security

Lesson #2:
Partnerships with Communities, NGOs (non-profits) crucial

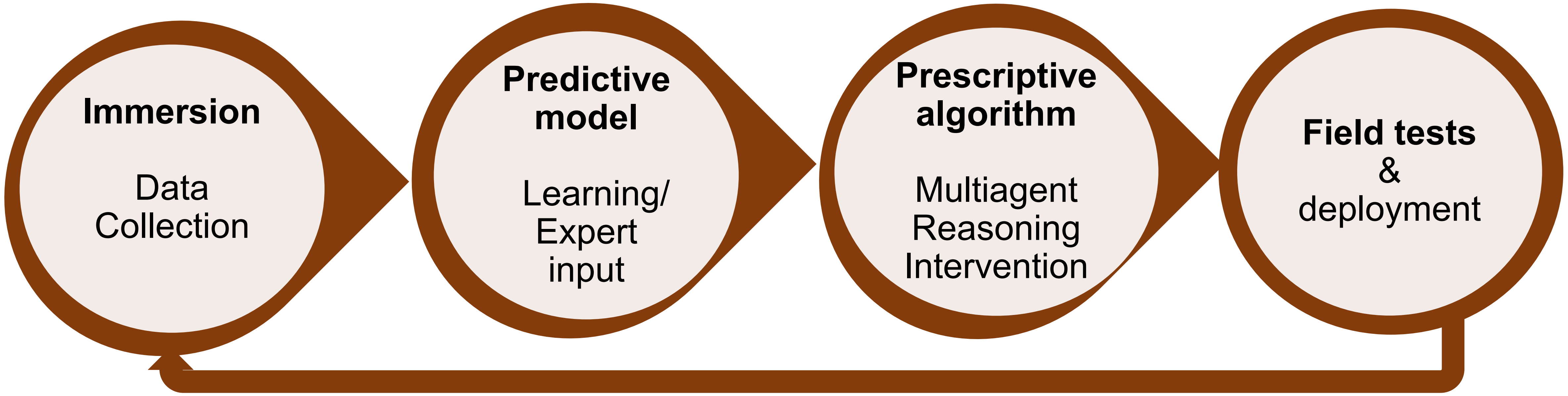
Empower non-profits to use AI tools; avoid being gatekeepers to AI4SI technology

Lesson #3:
Data-to-deployment pipeline; beyond improving algorithms



Lesson #3:
Data-to-deployment pipeline; beyond improving algorithms

Field test & deployment: Social impact is a key objective



Outline: Four Projects

Public Health

- ➢ *Social networks: HIV prevention*
- ➢ *Restless bandits: Maternal & child care*
- ➢ *Agent-based modeling: COVID-19 dynamics*

Conservation

- ➢ *Game theory, behavior modeling: Poaching prevention*

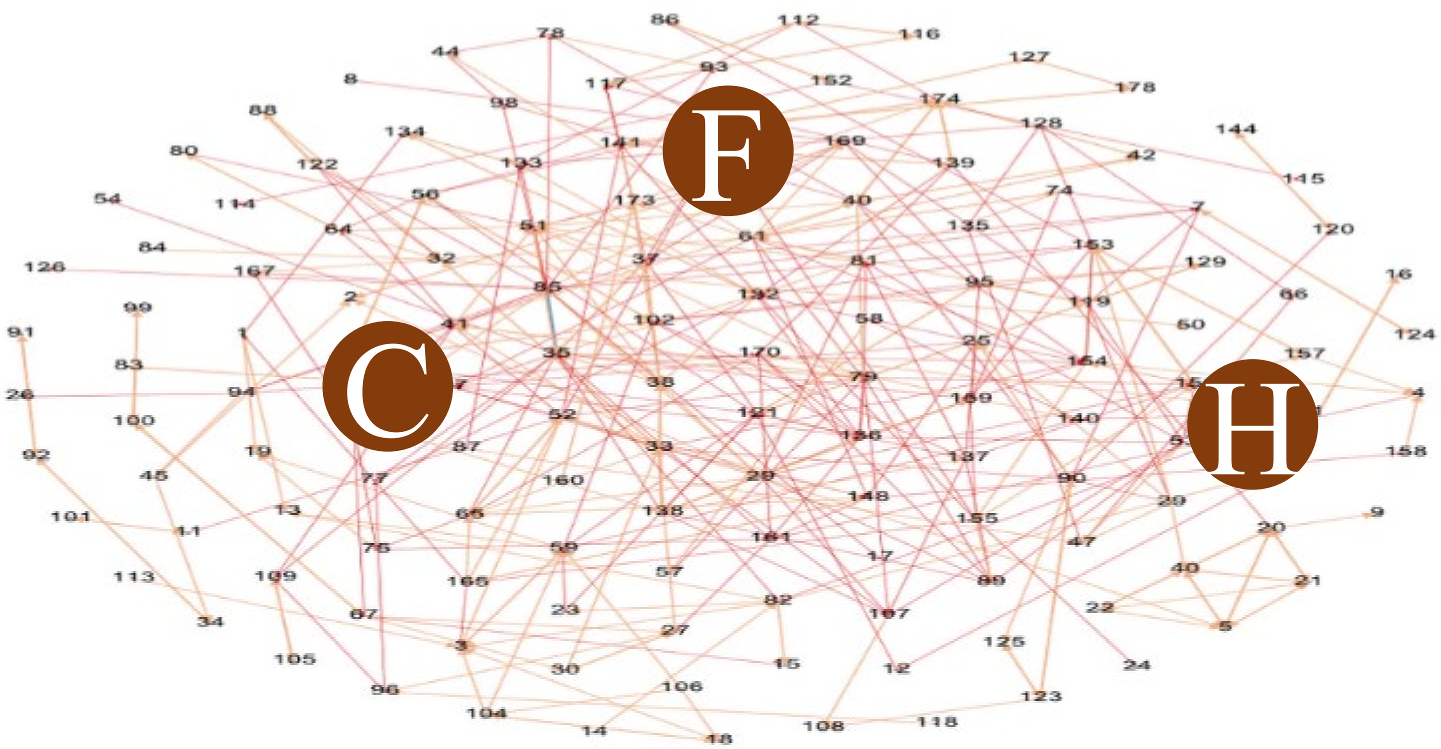
- Cover papers from 2017-now [AAMAS, AAAI, IJCAI, NeurIPS…]
- Focus on real world results; more simulations in papers
- PhD students & postdocs highlighted

Information dissemination & behavior change
Optimizing Limited Intervention (Social Worker) Resources

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

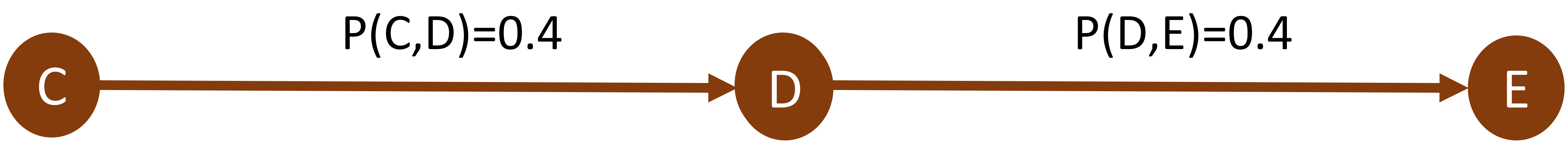
- ***Shelters**: Limited number of peer leaders to spread HIV information in social networks*
- *“Real” face-to-face interactions; not Facebook etc*

Influence Maximization in Social Networks



Select peer leader nodes to Maximize Expected Number of Influenced Nodes

- Independent cascade model: Propagation probability



Influence Maximization in Social Networks
Three Key Research Challenges

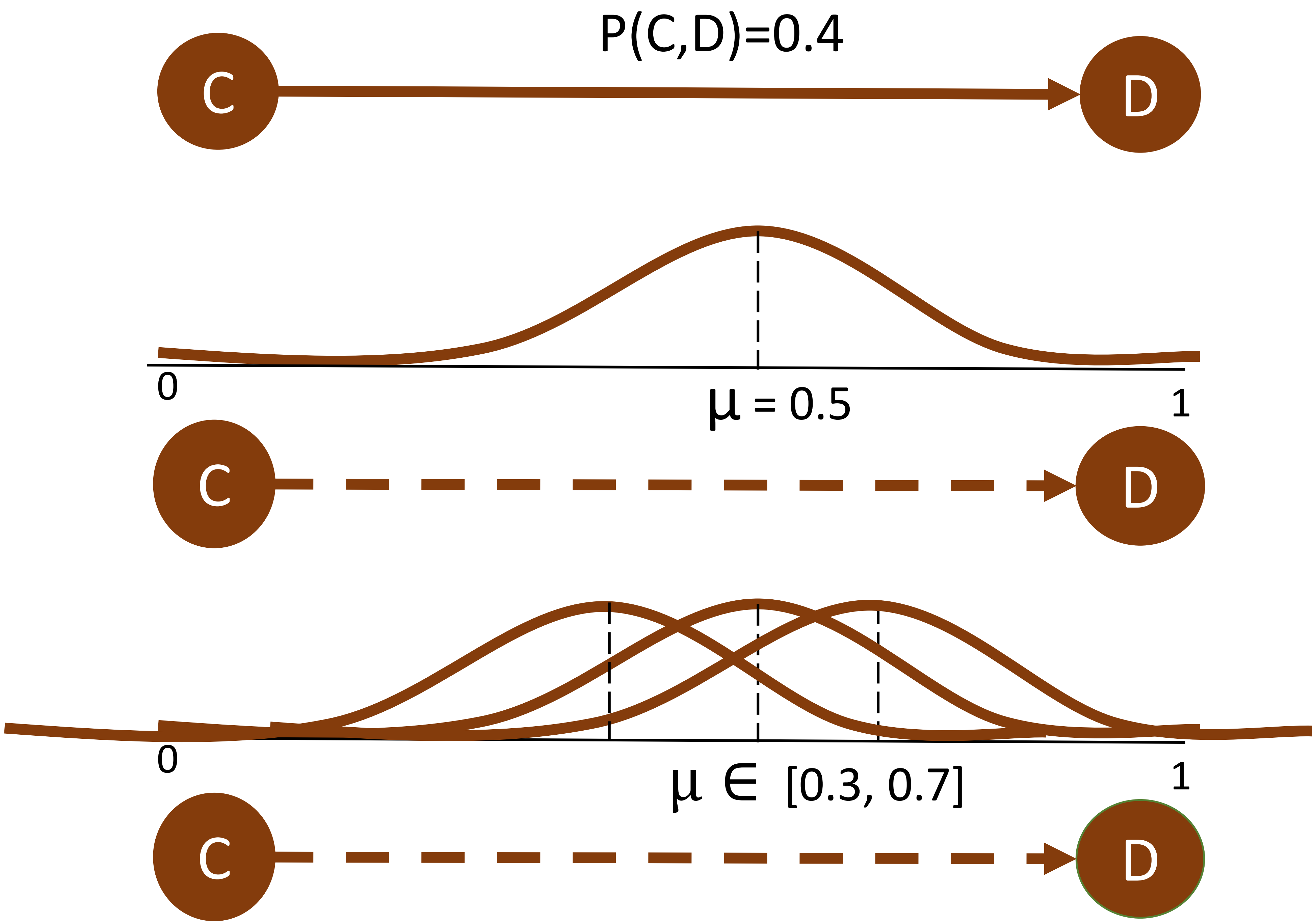
***Lesson #4**: Research challenges in AI for social impact?*

Lack of data & uncertainty is a key feature of AI for social impact

- Uncertainty in propagation probability over edges
- Multi-step dynamic policies to handle peer leader “no shows”
- Unknown social network, limited query budget to uncover network

Sketch some ways we solve these problems

Challenge 1: Uncertainty in Real-world Physical Social Networks



Robust Influence Maximization

(AAMAS 2017)

Wilder

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

$$\max_{x \in \Delta^{|P|}} \min_{\mu,\sigma} \sum x_p \frac{(Outcome(p))}{OPT(\mu,\sigma)}$$

Algorithm

Choose Peer Leaders p ∈ P generating mixed strategy "$x \in \Delta^{|P|}$"

vs

Nature

Chooses parameters

μ,σ

HEALER Algorithm
Robust Influence Maximization

(AAMAS 2017)

Wilder

Theorem: Converge with approximation guarantees

- Equilibrium strategy despite exponential strategy spaces: Double oracle

Nature

Influencer

	Params #1	Params #2	Params #3
Policy #1	0.8, -0.8	0.3, -0.3	0.4, -0.4
Policy #2	0.7, -0.7	0.5, -0.5	0.6, -0.6
Policy #3	0.6, -0.6	0.4, -0.4	0.7, -0.7

Influencer's oracle

	Params #1	Params #2
Policy #1	0.8, -0.8	0.3, -0.3
Policy #2	0.7, -0.7	0.5, -0.5
Policy #3	0.6, -0.6	0.4, -0.4

Nature's oracle

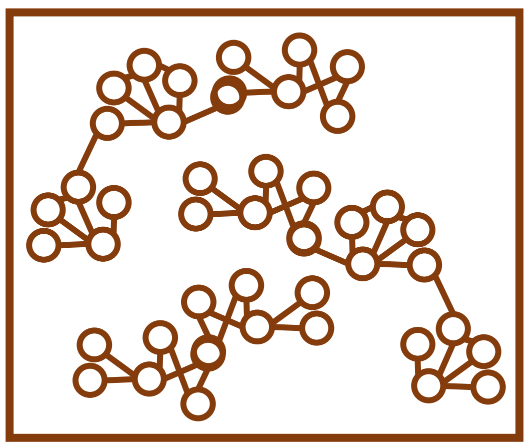
	Params #1	Params #2	Params #3
Policy #1	0.8, -0.8	0.3, -0.3	0.4, -0.4
Policy #2	0.7, -0.7	0.5, -0.5	0.6, -0.6
Policy #3	0.6, -0.6	0.4, -0.4	0.7, -0.7

Challenge 3:
Sampling Networks: Exploratory Influence Maximization
(AAAI 2018)

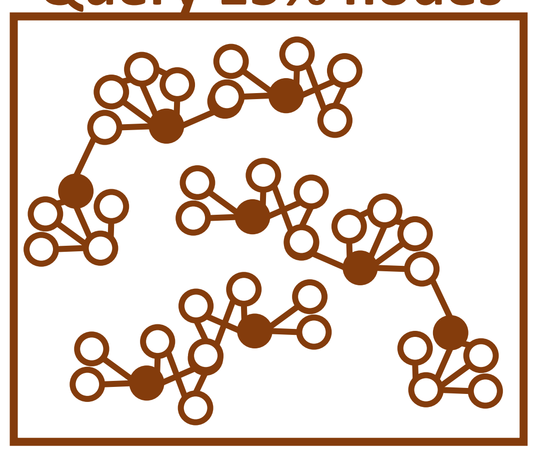
Wilder

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

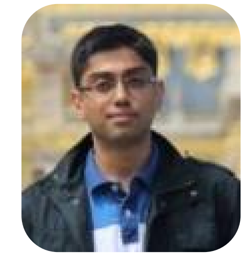
Sample node randomly & estimate size of its community;
Choose seeds from largest K communities

- 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

“CHANGE” with Homeless Youth

(IJCAI 2018)



Yadav

Wilder

Network Sampling

↓

Robust multi-step policy

↓

Peer leader selection

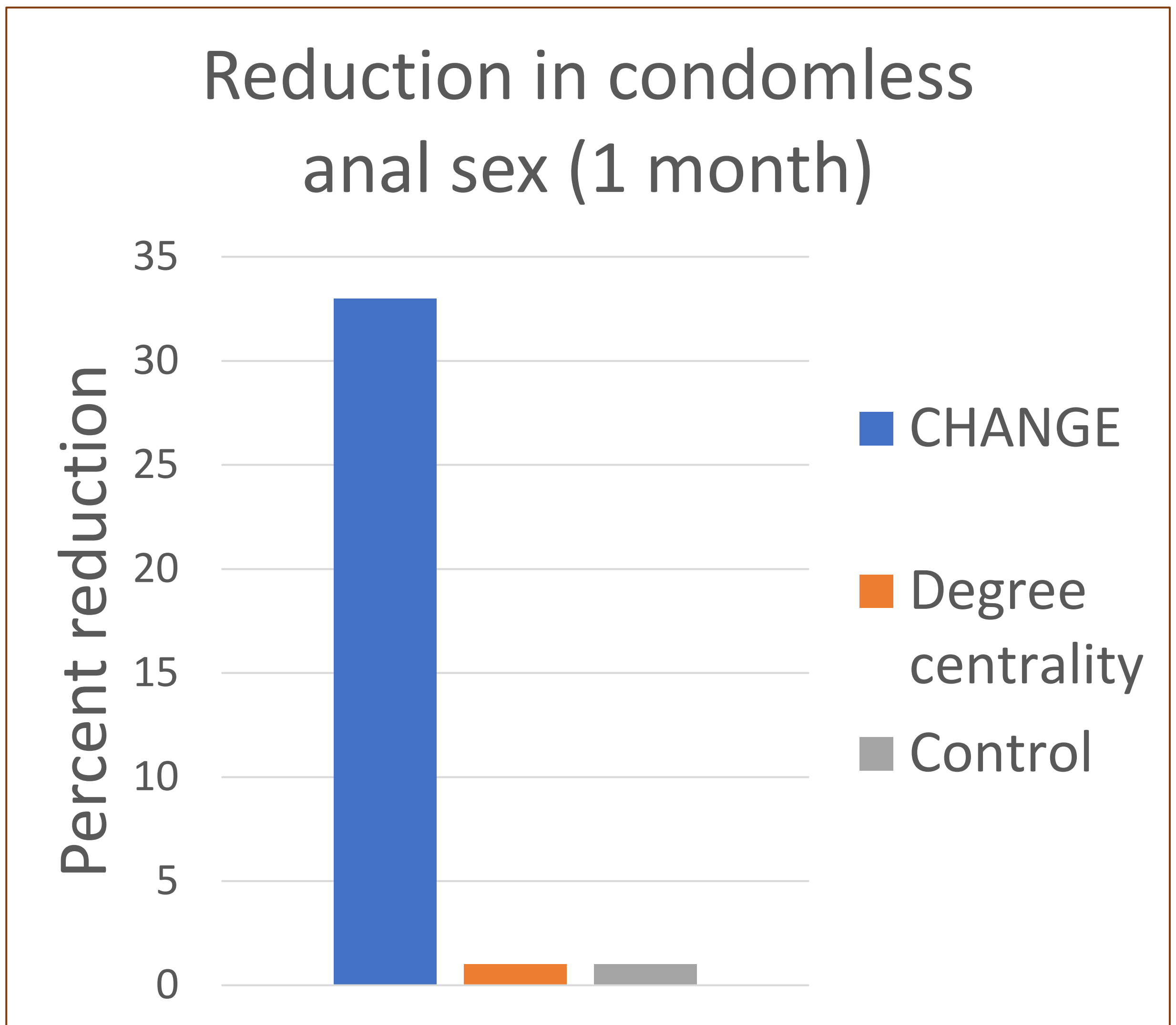
CHANGE

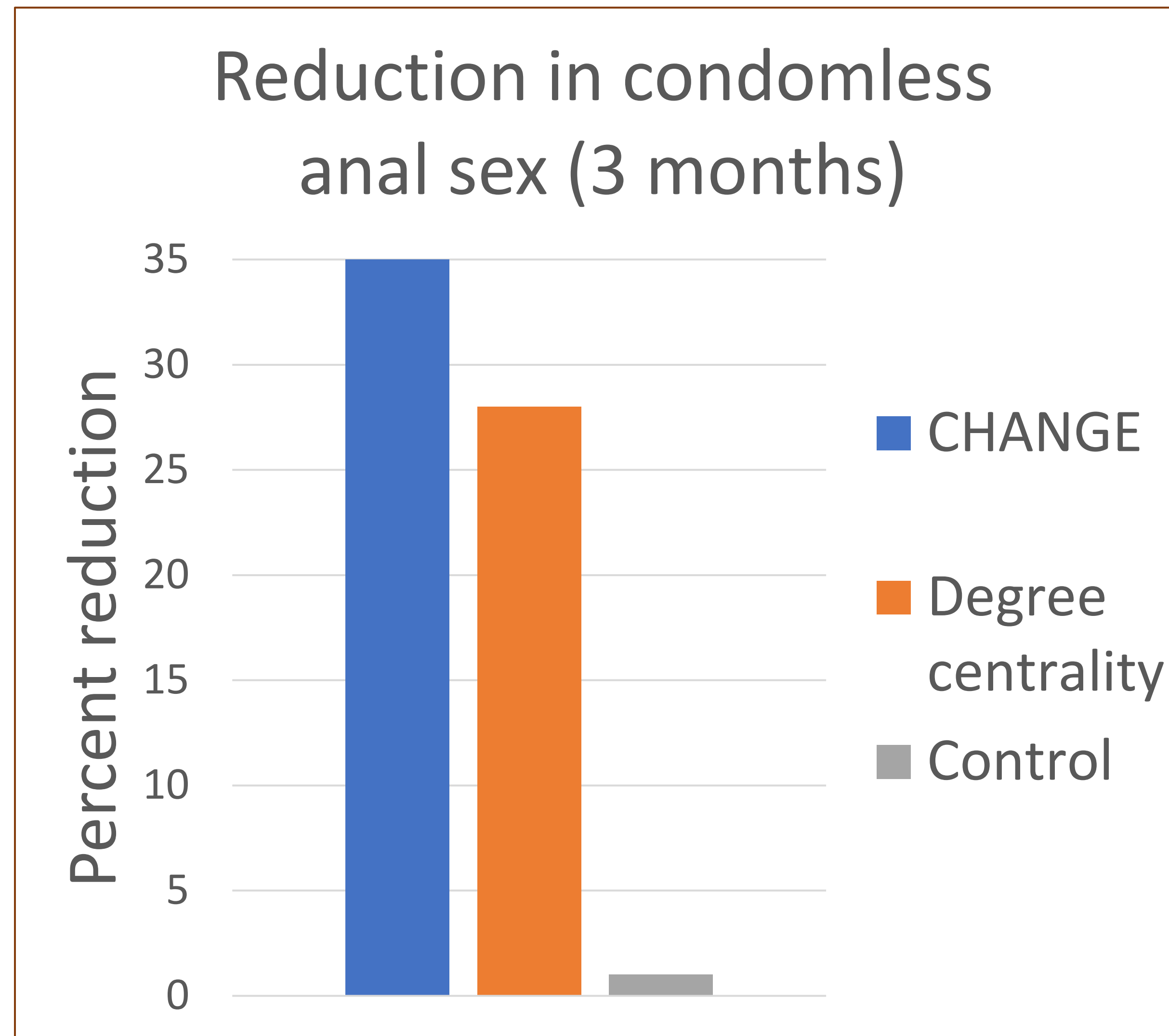
- 750 youth study with Prof. Eric Rice
- CHANGE vs Degree centrality vs Control
- Actual reduction in HIV risk behaviors?

Results of 750 Youth Study [with Prof. Eric Rice]
Actual reduction in HIV RISK Behavior?

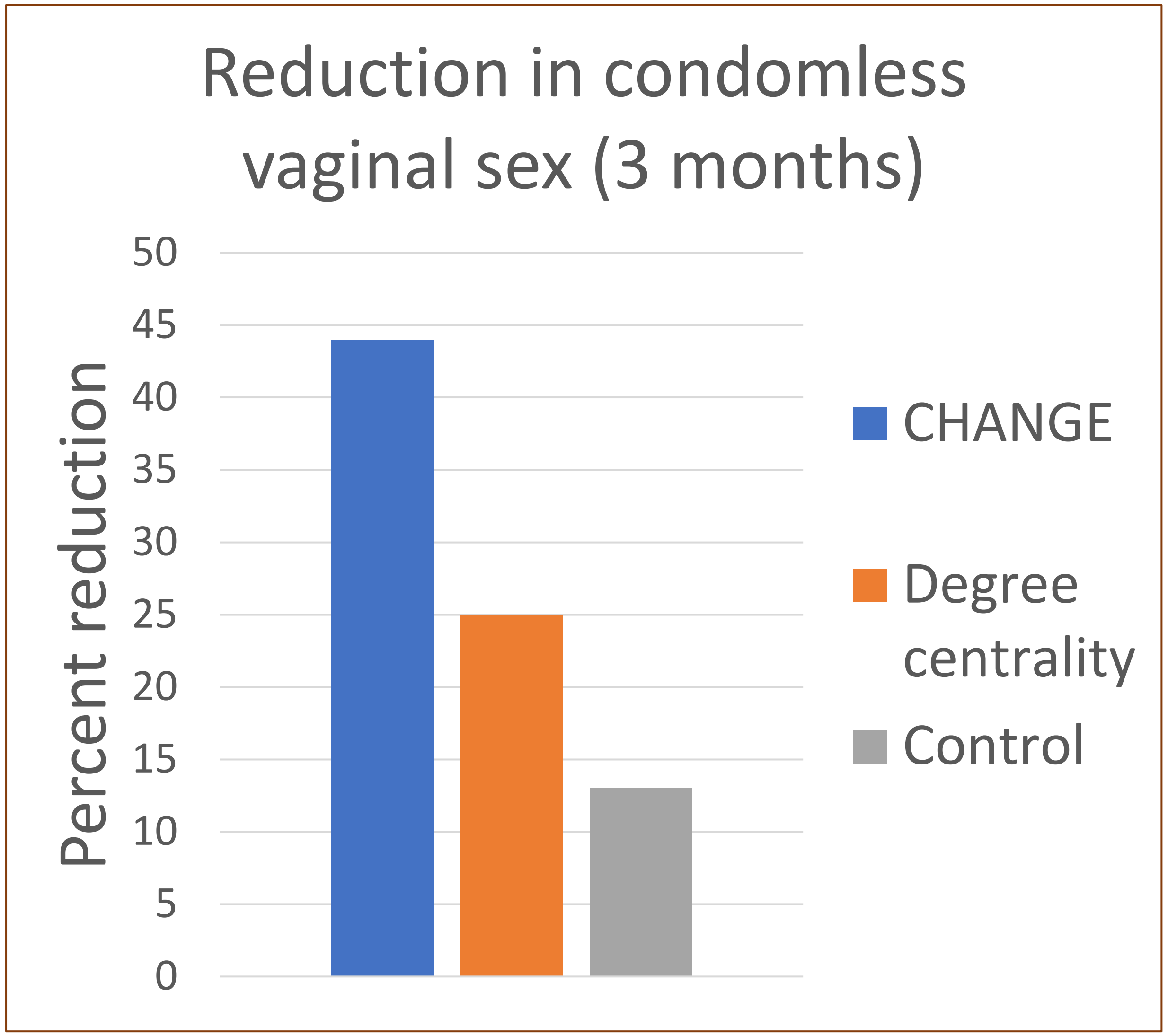
(AAAI 2021, Journal of AIDS/JAIDS 2021)

First large-scale application of influence maximization for public health



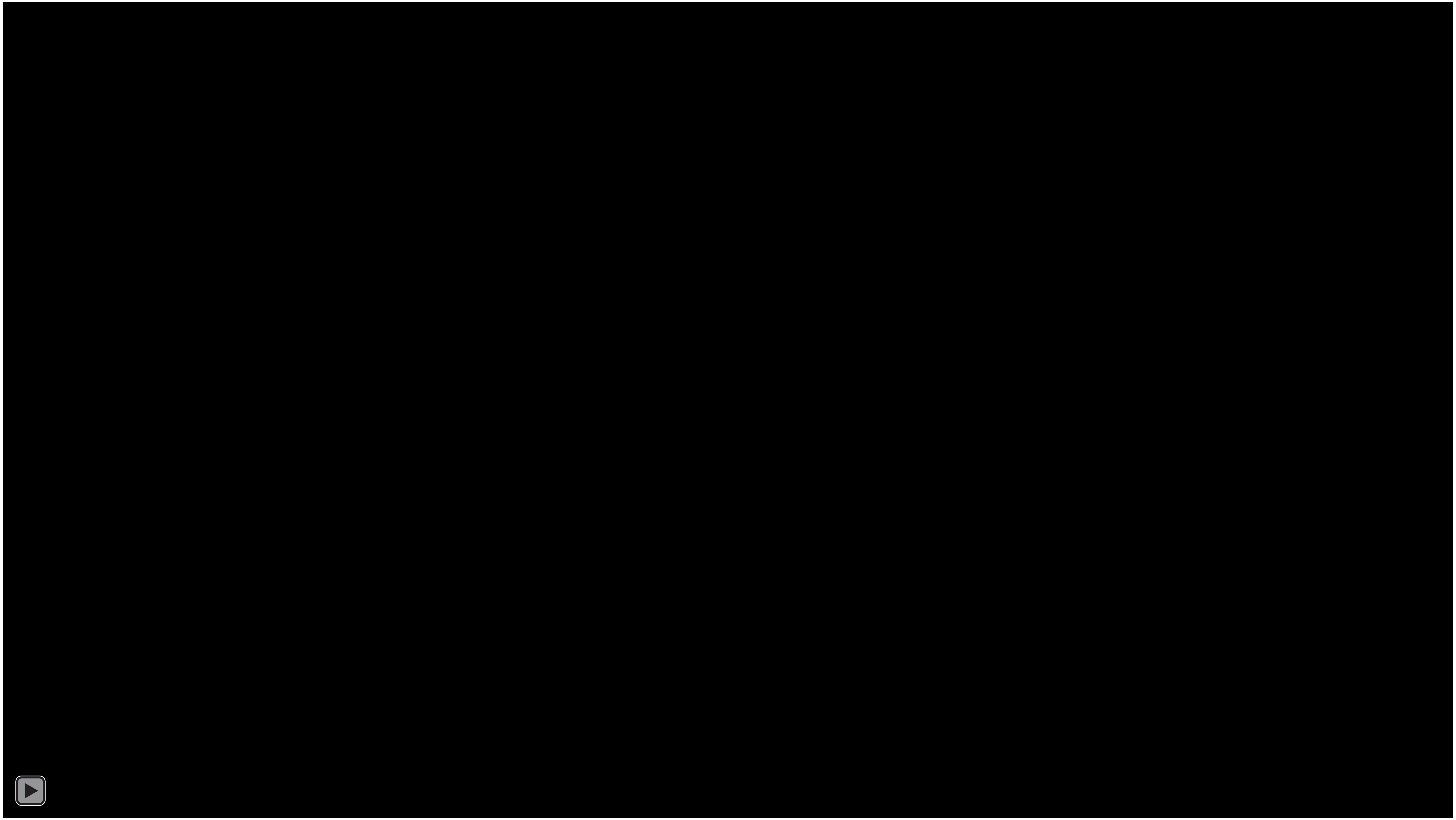


Results of 750 Youth Study [with Prof. Eric Rice]



**Statistical significance results in AAAI'21, JAIDS'21*

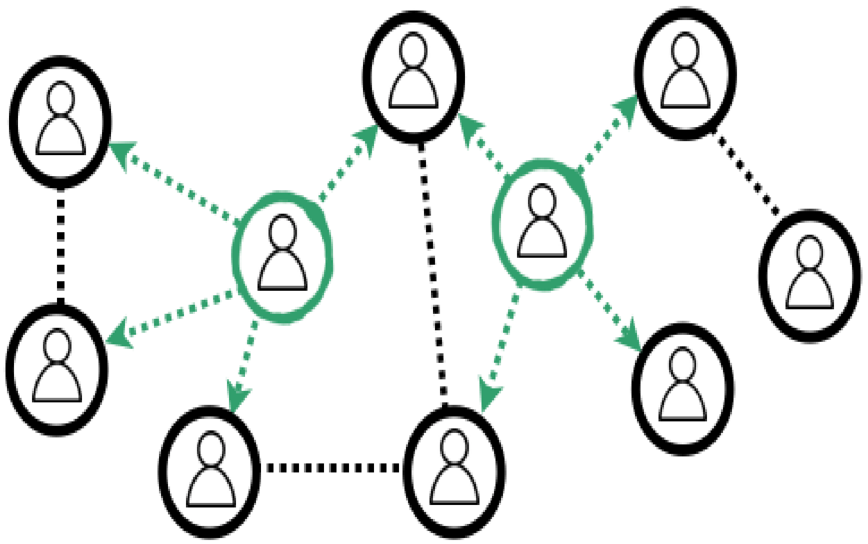
What our collaborators are saying:

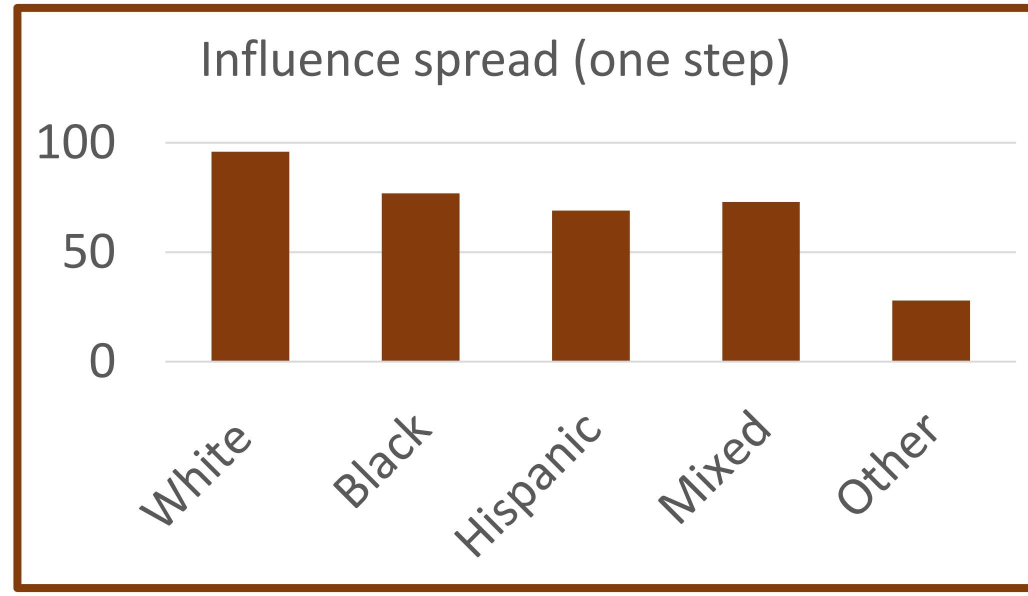


Next Steps: Fairness in Influence Maximization

(NeurIPS 2019, IJCAI 2019, AAAI 2021)

Rahmattalabi





Influence spread may cause disparity

Maxmin fairness:
NeurIPS2019

$$\min_{c \in C} u_c(A) \geq \gamma$$

Υ: 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

$$W_\alpha\left(u(A)\right)$$

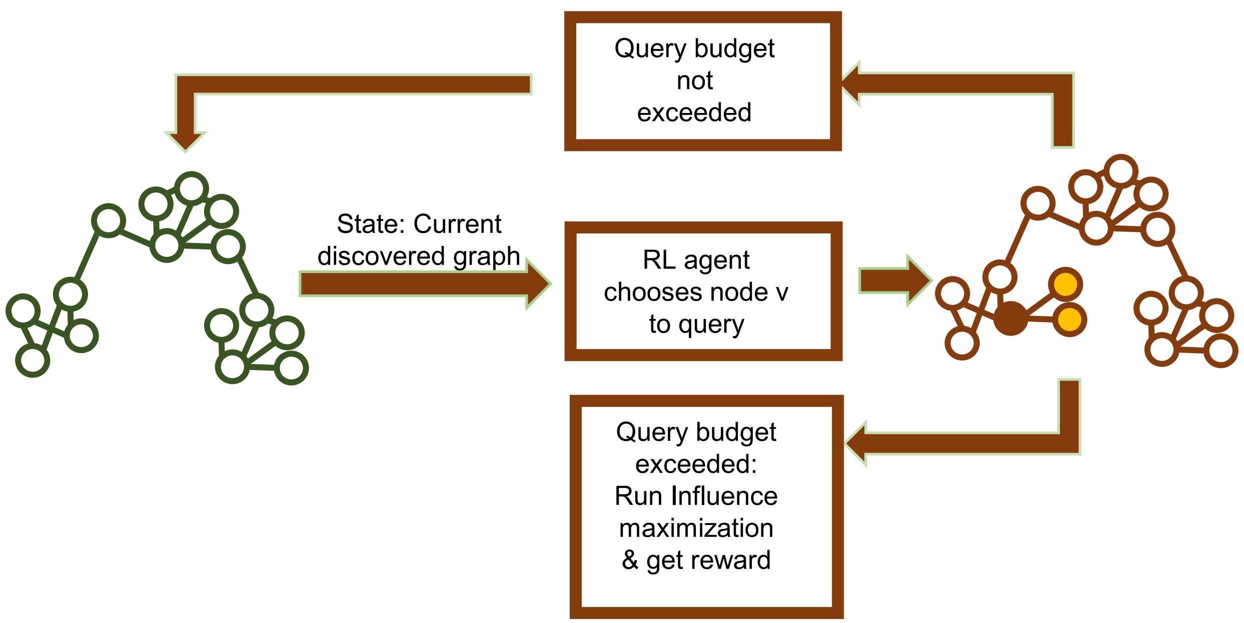
α controls fairness tradeoff; policymaker has choice

Next steps: Reinforcement Learning (RL)

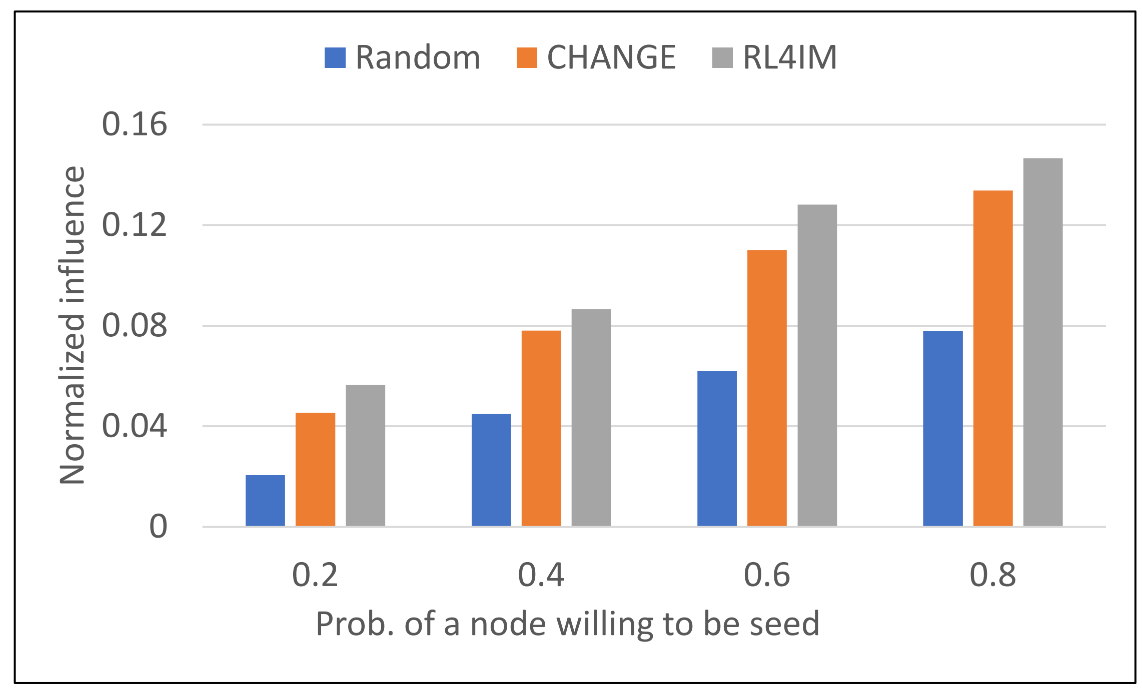
(AAMAS 2021 with IIT-Madras, UAI 2021)

Chen

RL for network sampling



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



RL speeds up Influence Maximization (RL4IM): RL4IM comparable performance to CHANGE, but negligible runtime

Outline

Public Health

- *Social networks: HIV prevention*
- *Restless bandits: Maternal & child care*
- *Agent-based modeling: COVID-19 dynamics*

Conservation

- *Game theory, behavior modeling: Poaching prevention*

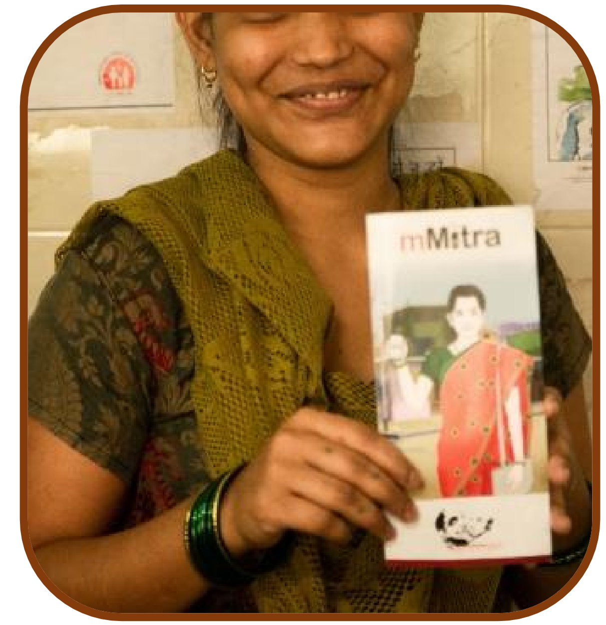
Motivating Restless Bandits
Health Program Adherence: Maternal & Child Care in India

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

armman
(Advancing Reduction in Mortality and Morbidity of Mothers, Children and Neonates)
Delivering India's Future

25 Million women

Weekly 2 minute AUTOMATED MESAGE to new/expecting moms



mMitra: Significant benefits
2.2 million women enrolled

- *Unfortunately, significant fraction 30-40% may become low-listeners*
- *Limited intervention resources: Service call to small number of beneficiaries*

Intervention Scheduling with Limited Resources: Motivating Restless Bandits

Example:

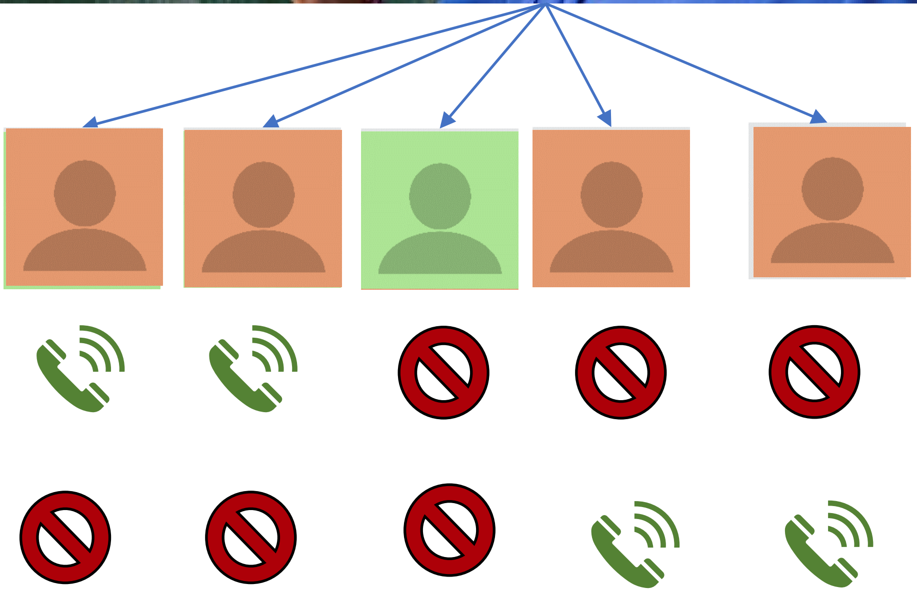
- *Large number N beneficiaries: 200000*
- *Choose K=4000 for service call per week?*
- *Maximize health messages listened to*

Photo Credit: IntraHealth International (CC BY-NC-SA 3.0 via https://www.intrahealth.org/)

Intervention Scheduling with Limited Resources: Motivating Restless Bandits

Example:

- *Large number N beneficiaries: 200000*
- *Which K=4000 for service call per week?*
- *Maximize number of messages listened to*



Challenges:

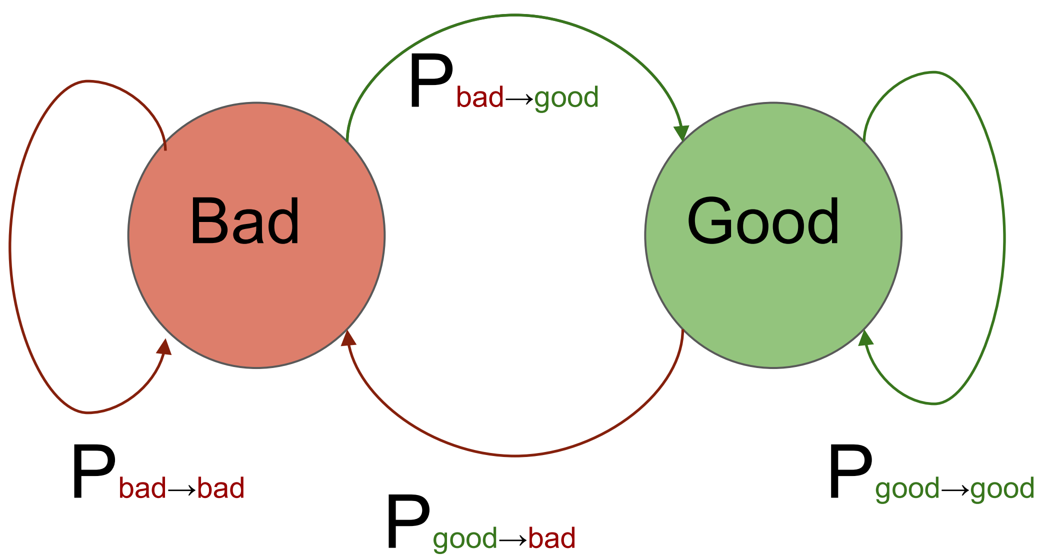
- *Call may not change beneficiary state*
- *Beneficiary may change state on their own*
- *Prioritize 4000 beneficiaries per week*

Restless bandit: *K of N arms per week*

Photo Credit: IntraHealth International (CC BY-NC-SA 3.0 via https://www.intrahealth.org/)

Restless Bandits Model: Each Arm is an MDP
Each Arm Models a Beneficiary

States of MDP



A"bad" state & a "good" state

Actions

Intervene or
Not intervene

Transition matrix

Not intervene:

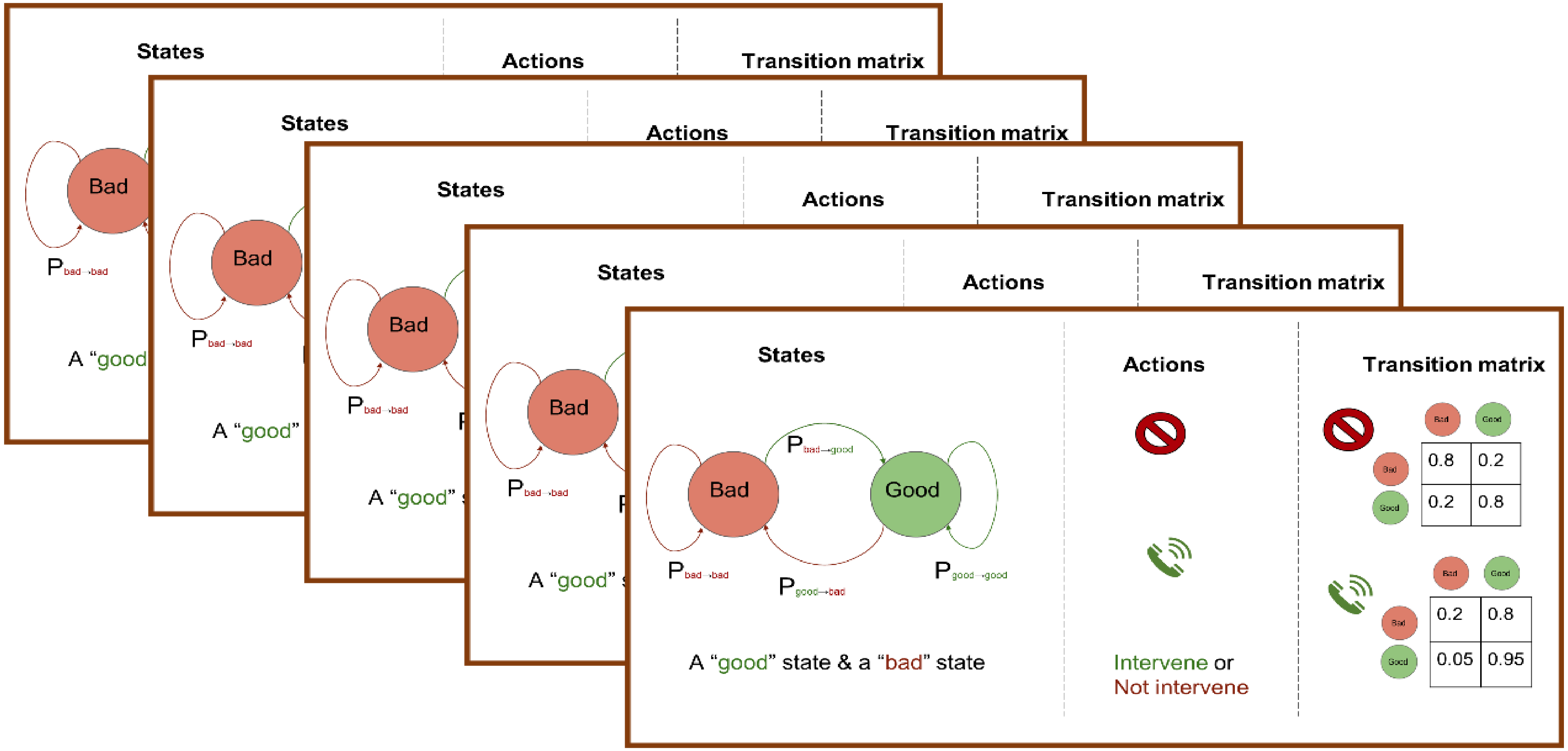
	Bad	Good
Bad	0.8	0.2
Good	0.2	0.8

Intervene:

	Bad	Good
Bad	0.2	0.8
Good	0.05	0.95

Restless Bandits Model
Whittle Index: Efficiently Select K out of N Beneficiaries

Mate



Compute Whittle index for current state of each arm: Computes benefit of intervention
Choose top K arms by benefit
Use (Qian et al 2016) algorithm

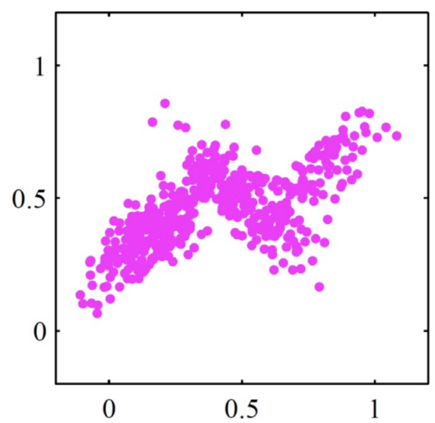
$W(s) = INF_\gamma \{\gamma : Q_\gamma(s,$  $) = Q_\gamma(s,$  $)\}$

Key Research Challenge
Unknown Transition Probabilities

Mate

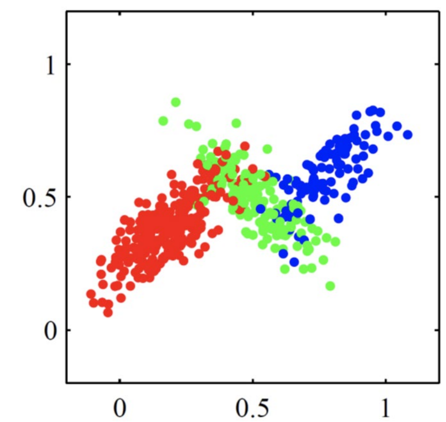
- **Limited previous beneficiary data:** features ***f*** + engagement sequence {(s, a, s'), … }
- Clustering compensates for lack of data, also speeds up Whittle index computation

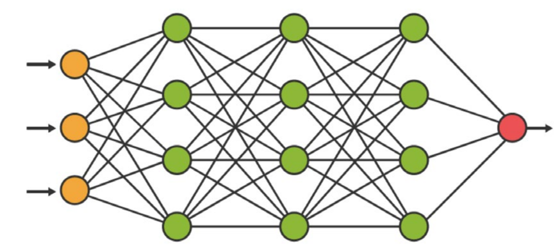
Training Step:
With historical batch data



Passive transition probability data

Fit a GMM or k-means

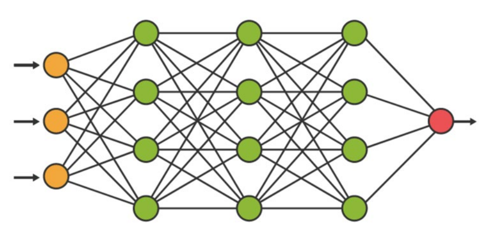




Learn a map from features → clusters

Testing Step:
New, unseen beneficiaries

features



Predict clusters
[0.3, 0.1, 0.6]

Compute Whittle indices

Top k

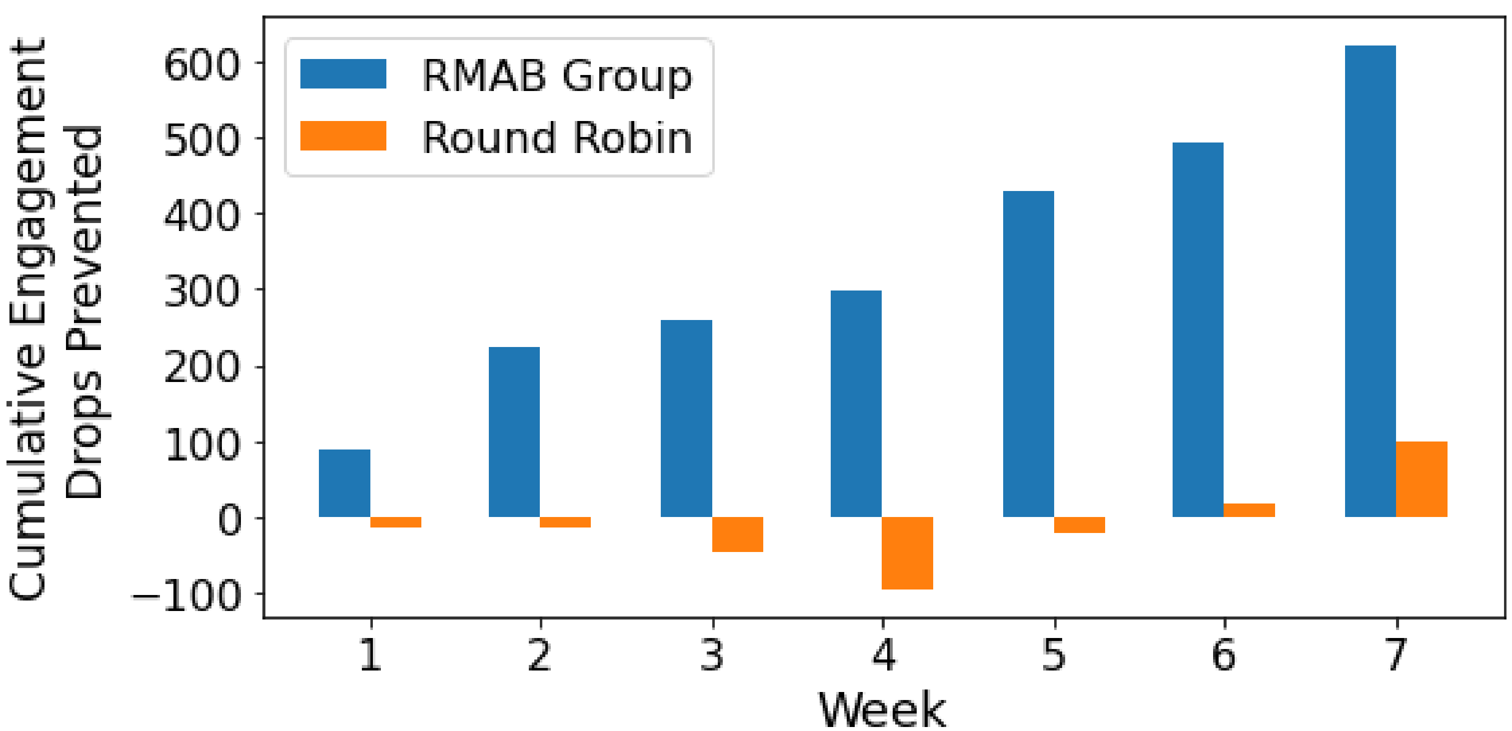
Results of 23000 Beneficiary Field Study

(Under submission)

Mate

First large-scale application: restless multiarmed bandits (RMAB) for public health

- 7667 beneficiaries per group:
  RMAB, Round-robin,
  Current-Standard-of-Care (CSOC)
- Pulled 225 arms/week for seven weeks
- How many more health messages listened to over Current-Std-of-Care (CSOC) group
- Statistical significance: linear regression model



	RMAB vs CSOC	RR vs CSOC	RMAB vs RR
% reduction in cumulative engagement drops	32.0%	5.2%	28.3%
p-value	0.044*	0.740	0.098†

New 100,000 Beneficiary Study

Transitioning software to ARMMAN

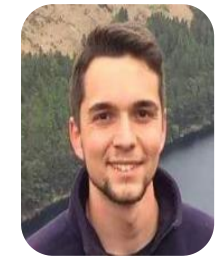
ARMMAN Feedback
Youtube: "AI for Social Good in partnership with ARMMAN"

"We are able to reach out to more and more women each week, and get them back into the fold and save lives, because of AI" – Dr Aparna Hegde

"I follow all the advice and take good care of my baby"

Next steps: Adherence Monitoring for Preventing Tuberculosis in India

(KDD 2019)



Killian

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

TB Treatment
6 months of pills

everwell

➢ *Which patients to call? Challenge of partial observability*

Collapsing Bandits:
Restless Bandits with Partial Observability
(NeurIPS 2020)


Mate

Killian

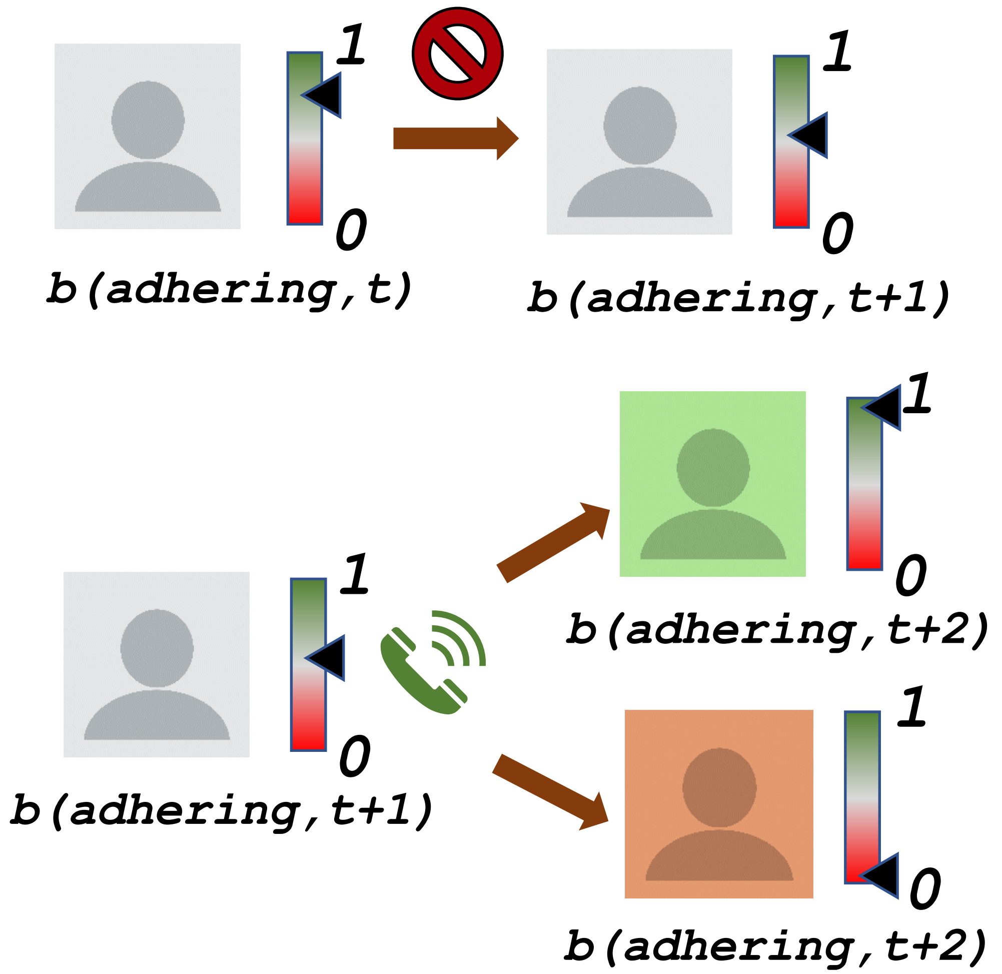
Theorem (Whittle Index): Collapsing bandits are Indexable if threshold policies are optimal.

When arm not played (patient not called)

- *No observation*
- *Instead, compute belief of adherence*

When arm played: Uncertainty collapse

- *Observe current state*



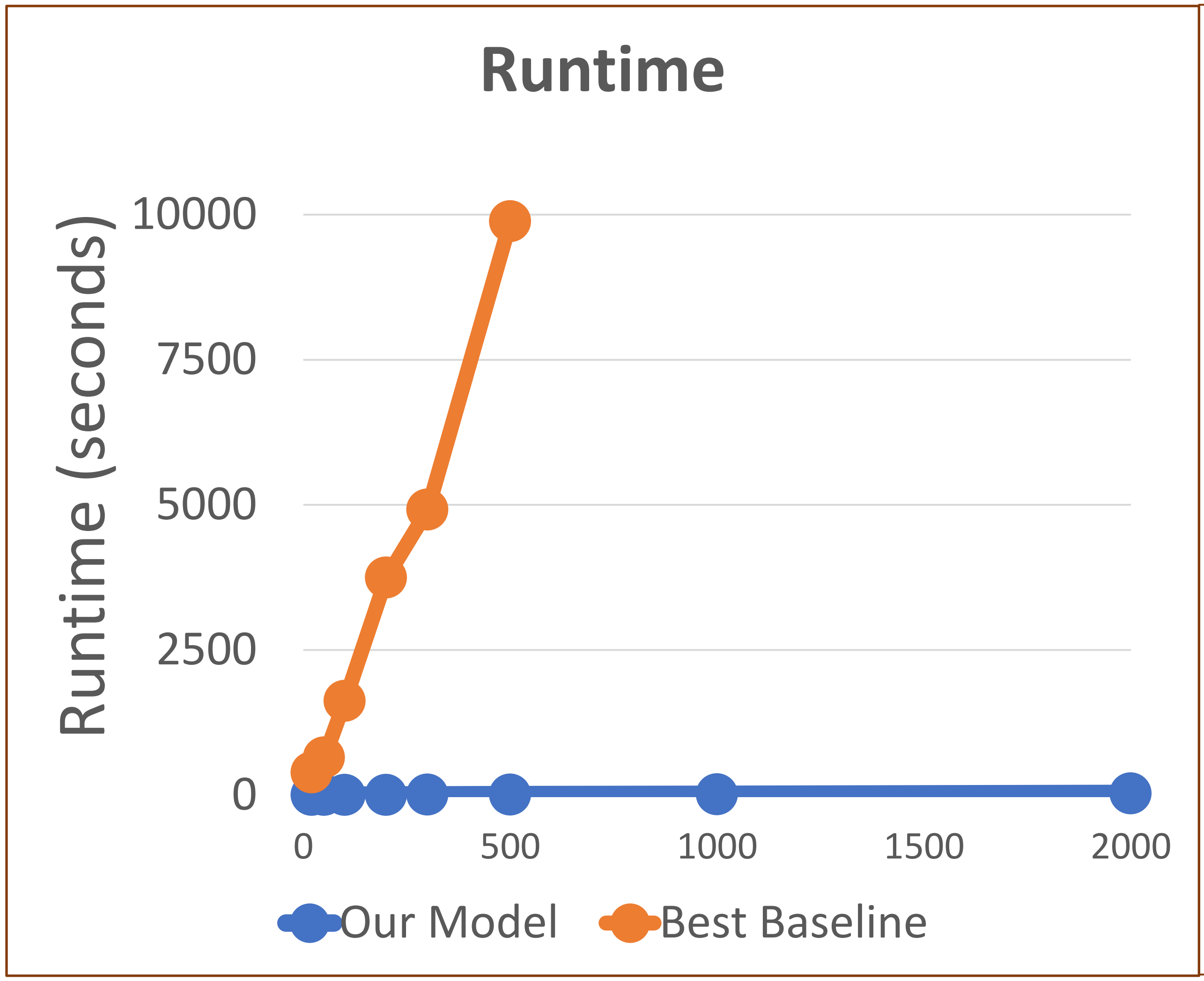
- ***Exploit "collapsing" for fast algorithm: Fixed number of belief states***

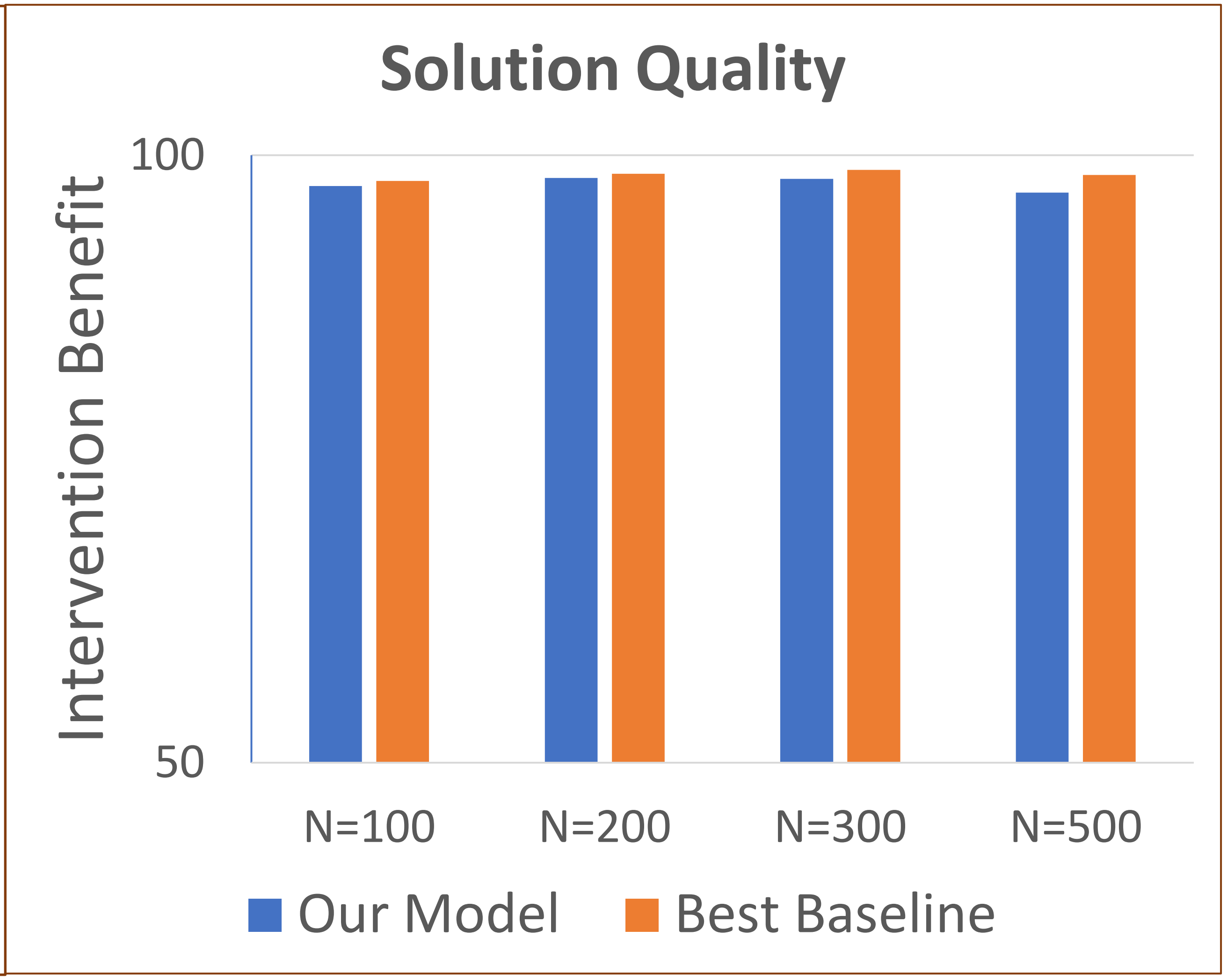
New Fast Algorithm:
Collapsing Bandits for Partial Observability


Mate


Killian

- *Orders of magnitude speedup with little solution quality loss*
- *ORANGE = Best baseline*
- *Blue = Our model*





Next Steps:
Decision-focused Learning in Restless Bandits

(AAMAS2020, NeurIPS 2020, NeurIPS 2021)

Wang

Data-to-deployment pipeline:

- **TWO STAGES**: Maximize learning accuracy, then maximize decision quality
- Maximizing learning accuracy ≠ Maximizing decision quality

Immersion

Problems/ Data Collection

Learning

Features to transition probabilities

Optimize

Restless bandits: Top K beneficiaries

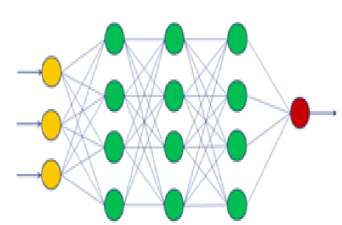
Evaluation & deployment

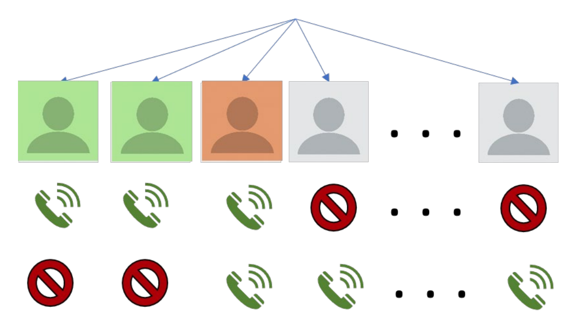
armman

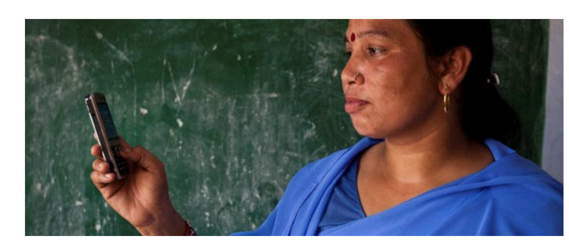
(Advancing Reduction in Mortality and Morbidity of Mothers, Children and Neonates)

Delivering India's Future

25 Million women







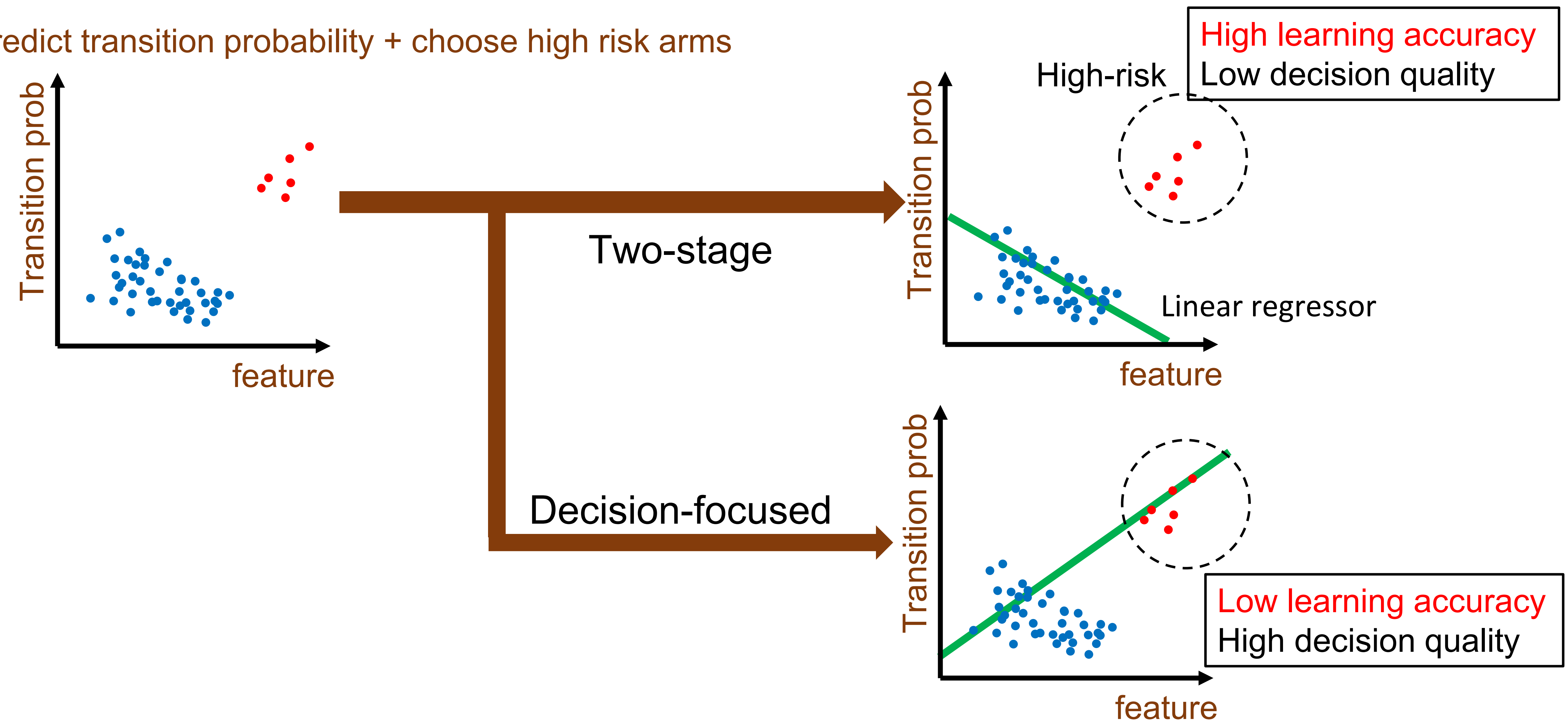
Next Steps:
Decision-focused Learning in Restless Bandits

(AAMAS2020, NeurIPS 2020, NeurIPS 2021)

Wang

- Maximizing learning accuracy ≠ Maximizing decision quality
- Decision-focused learning: Modify loss function to directly maximize decision quality

Predict transition probability + choose high risk arms



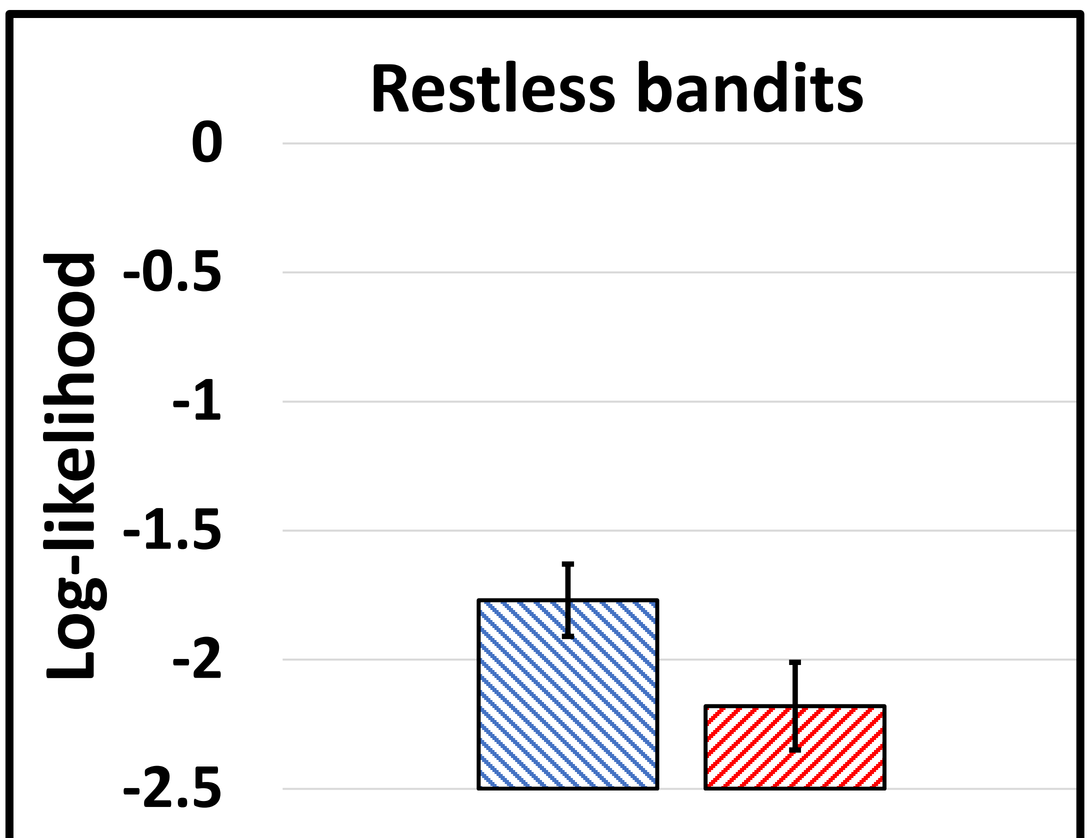
Next Steps:
Decision-focused Learning in Restless Bandits

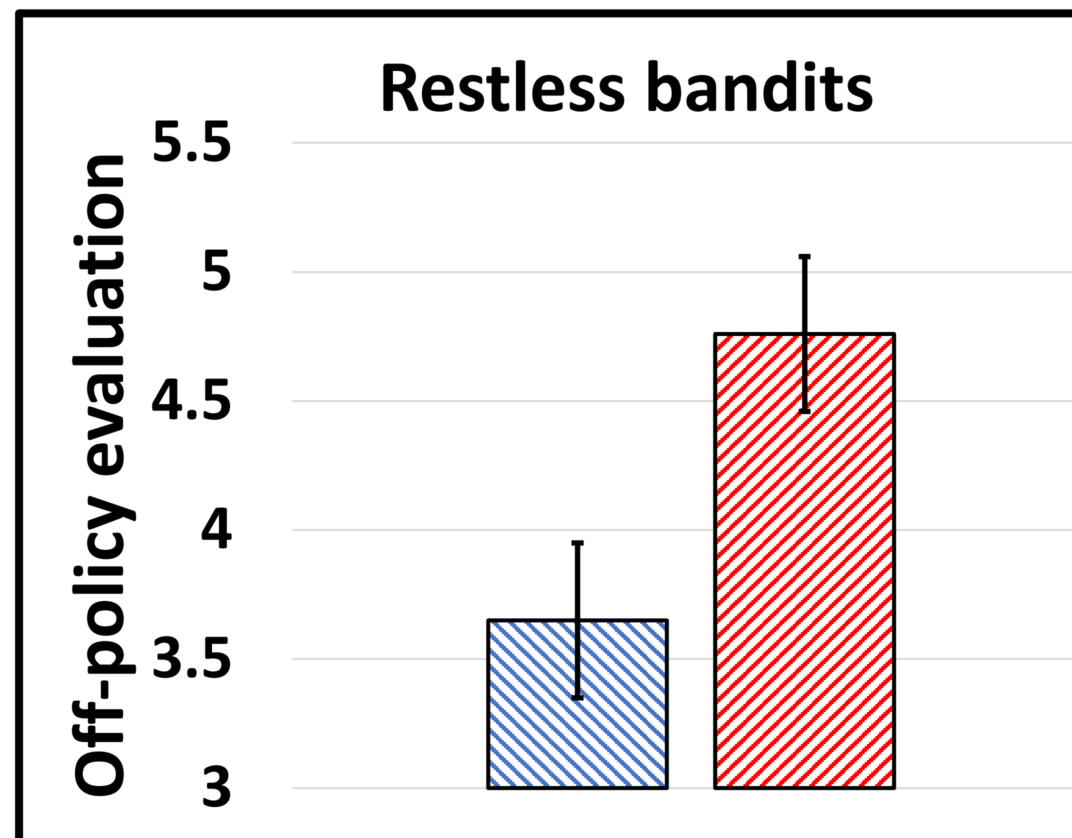
(AAMAS2020, NeurIPS 2020, NeurIPS 2021)

Wang

- Maximizing learning accuracy ≠ Maximizing decision quality
- Decision-focused learning: Modify loss function to directly maximize decision quality
- Working on ARMMAN

two-stage : $\frac{\partial \text{ MDP accuracy}}{\partial \text{ model}}$ **decision-focused** : $\frac{\partial \text{ quality}}{\partial \text{ MDP}} \frac{\partial \text{ MDP}}{\partial \text{ model}}$



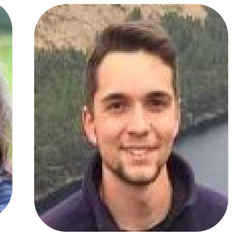


Next Steps in Restless Bandits

(AAMAS 2021a, KDD 2021, IJCAI 2021, AAMAS 2021b)

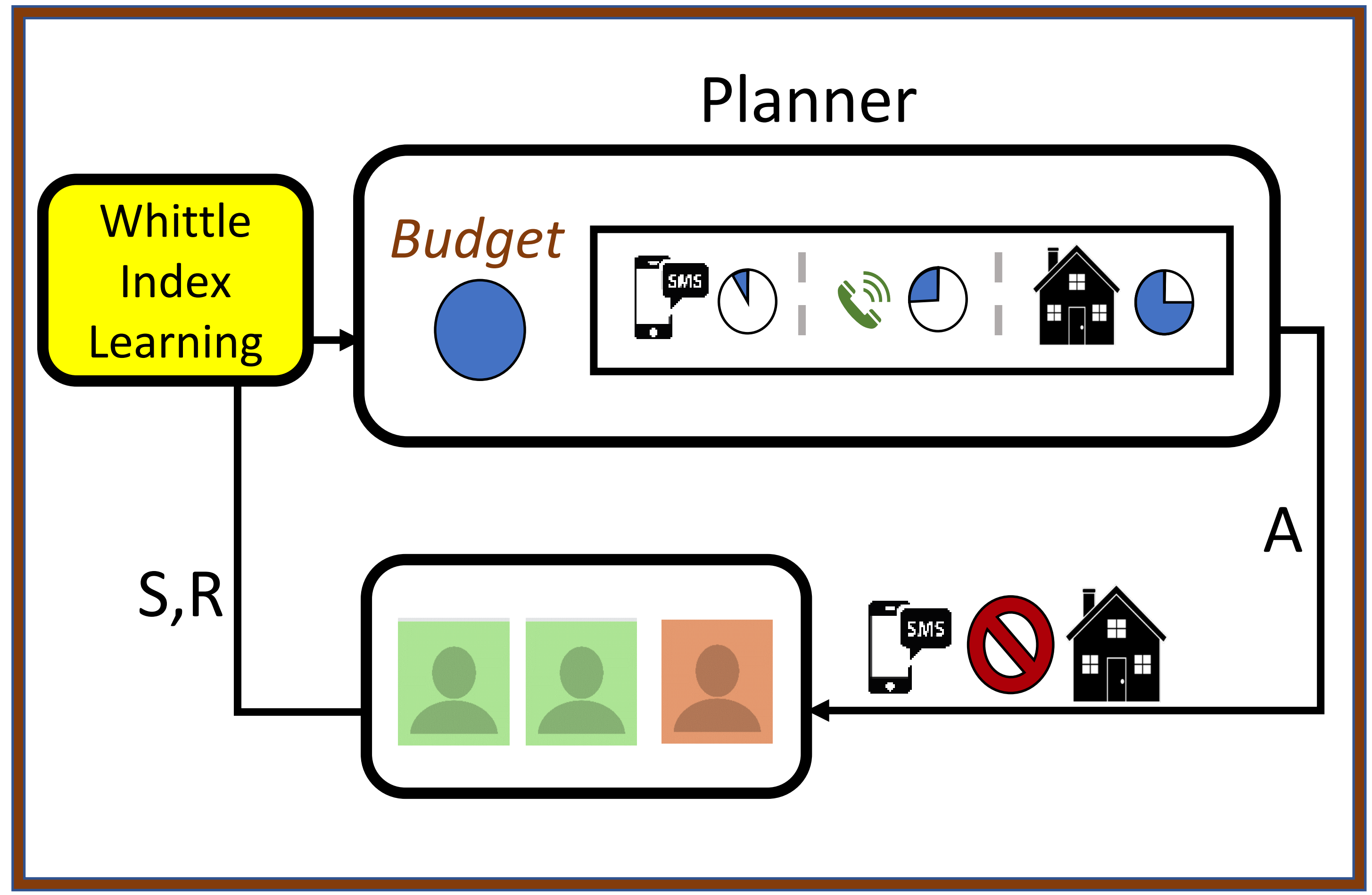

Mate


Biswas


Killian

Online learning with multiple actions (no past data):

- *Policies: index Q-Learning*



Fast Planning

- *Risk aware restless bandits*
- *Robust restless bandits*

COVID-19: Agent-based Simulation Model

Wilder

AAAS Become a Member

ScienceAdvances Contents

SHARE

RESEARCH ARTICLE | CORONAVIRUS

Test sensitivity is secondary to turnaround time for COVID-19 s

Daniel B. Larremore[1,2,*], Bryan Wilder[3], Evan Lester[4,5], Soraya Shehata[5,6], James M. Burke[4], James A. Hay[7,8], ...

+ See all authors and affiliations

Science Advances 01 Jan 2021:
Vol. 7, no. 1, eabd5393
DOI: 10.1126/sciadv.abd5393

The New York Times

THE MORNING NEWSLETTER

Where Are the Tests?

Other countries are awash in Covid tests. The U.S. is not.

PNAS Proceedings of the National Academy of Sciences of the United States of America

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.2010651117

35th AAAI Conference on Artificial Intelligence

A Virtual Conference

February 2–9, 2021

Tracking disease outbreaks from sparse data with Bayesian inference

Bryan Wilder,[1] Michael Mina[2], Milind Tambe[1]

[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, mmina@hsph.harvard.edu, milind_tambe@harvard.edu

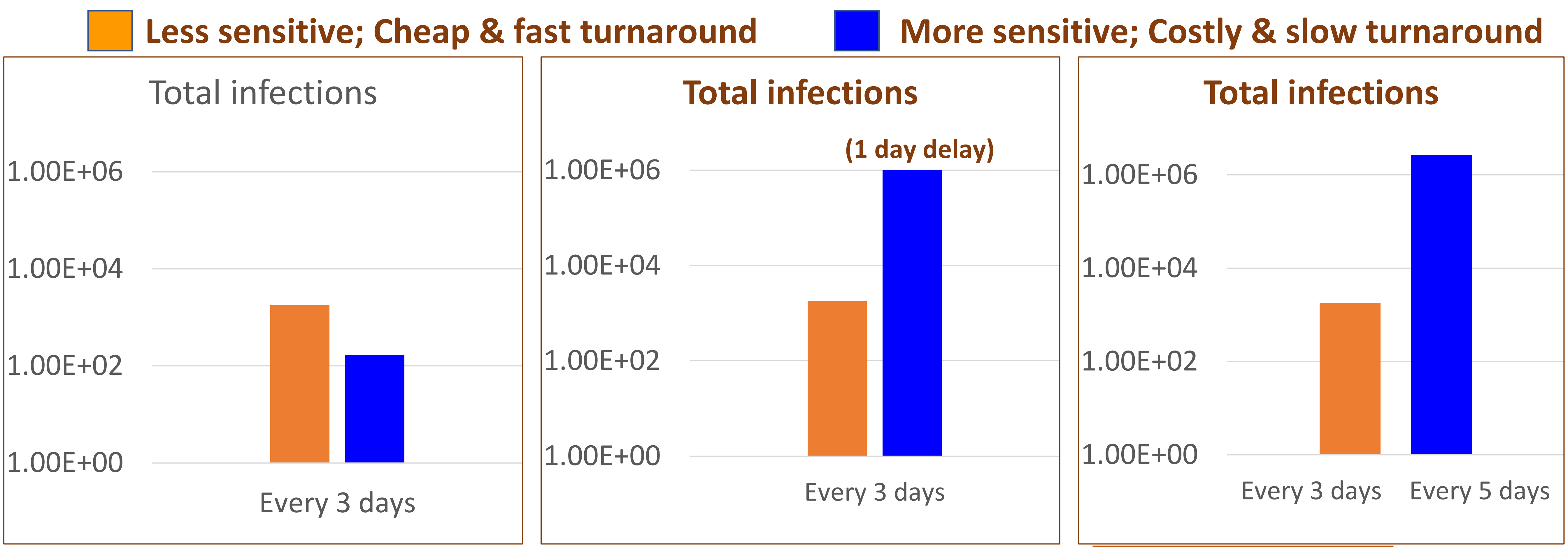
Wilder

COVID Testing Policy: Accuracy vs Ease

(Science Advances, 2020) with Prof. Michael Mina

- Tests varying sensitivity/cost: which one to use?
  - 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



COVID Testing Policy: Impact

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


TIME
PERSON OF THE YEAR 2020 FRONTLINE HEALTH WORKERS RACIAL-JUSTICE ORGANIZERS NEWSLETTER

IDEAS • COVID-19

America Needs to Radically Rethink Our COVID-19 Testing Approach


The New York Times

The Coronavirus Outbreak › LIVE Latest Updates Maps and Cases The Latest Vaccine Information The Thanksgiving Effect

'It's Kitchen Sink Time': Fast, Less-Accurate Coronavirus Tests May Be Good Enough


Popular Latest *The Atlantic* My Account Give a Gift

The Plan That Could Give Us Our Lives Back

The U.S. has never had enough coronavirus tests. Now a group of epidemiologists, economists, and dreamers is plotting a new strategy to defeat the virus, even before a vaccine is found.

Outline

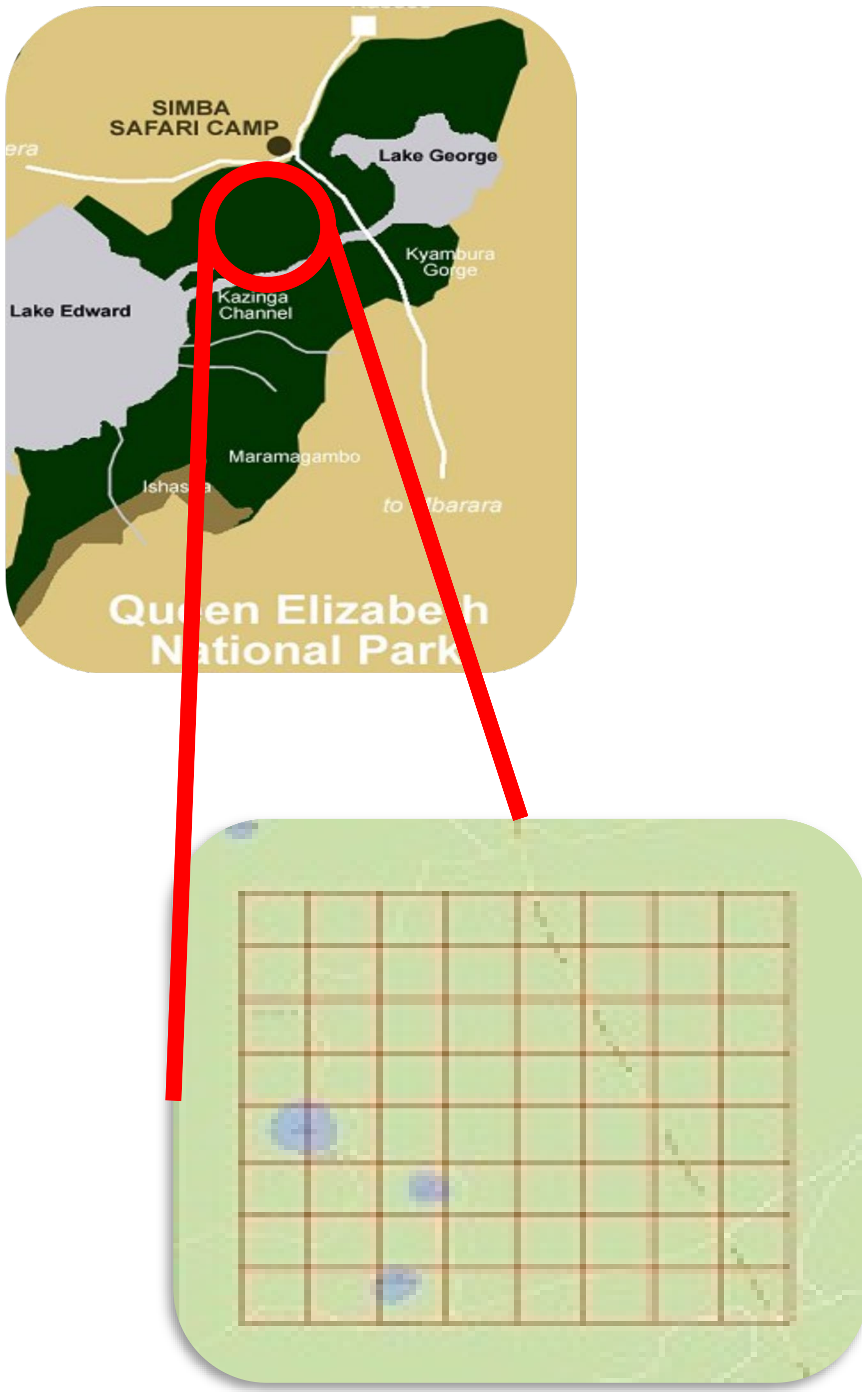
Public Health

- *Social networks: HIV prevention*
- *Restless bandits: Maternal & child care*
- *Agent-based modeling: COVID-19 dynamics*

Conservation

- *Game theory, behavior modeling: Poaching prevention*

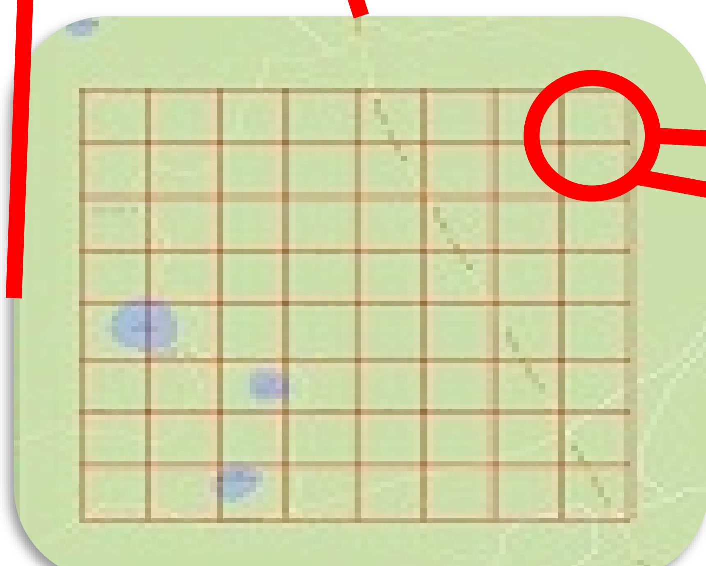
Patrols to Reduce Snaring in Wildlife Parks

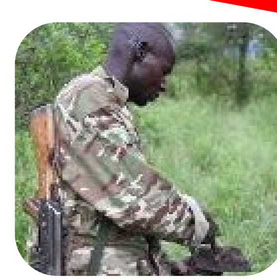


Snare or Trap

Wire snares

Stackelberg Security Games to Prescribe Patrols

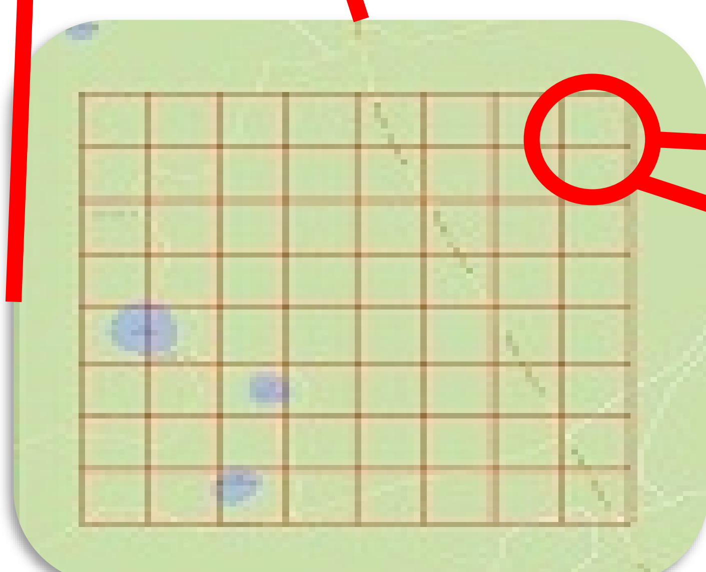


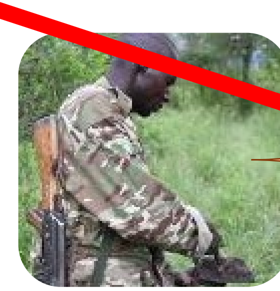


	Area1	Area2
Area1	4, -3	-1, 1
Area2	-5, 5	2, -1

Stackelberg Security Games to Prescribe Patrols

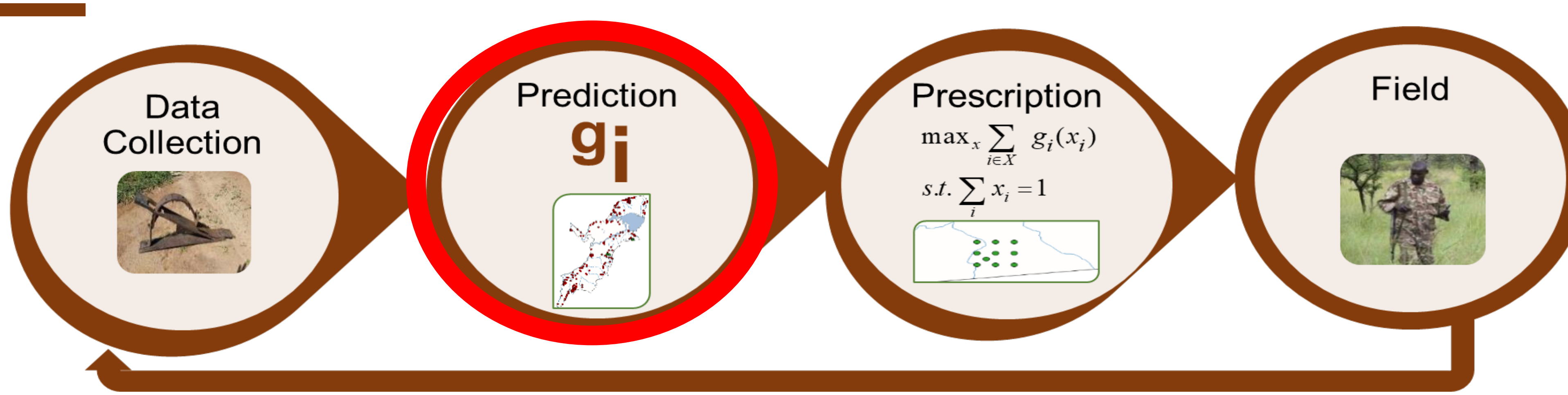
- *Randomized (mixed) strategy for rangers*
- *Bounded rational poacher model: learn via past poaching data*



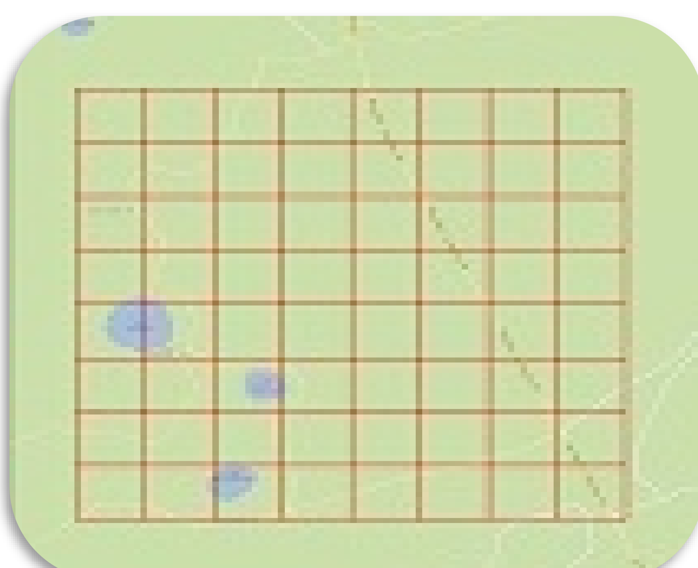


	Area1	Area2
Area1	4, -3	-1, 1
Area2	-5, 5	2, -1

Stackelberg Security Games to Prescribe Patrols



- *Bounded rational poacher model: learn via past poaching data*

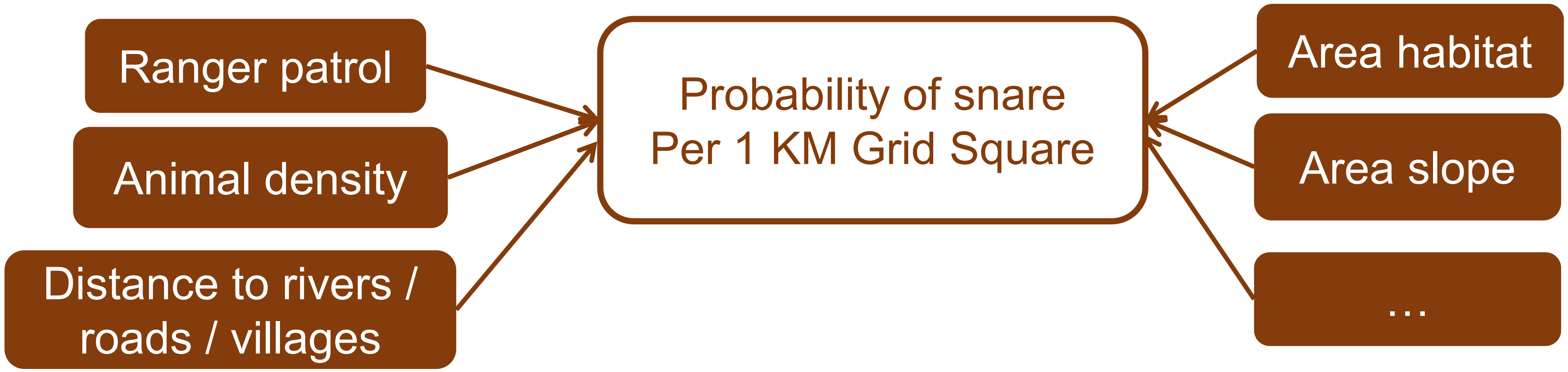


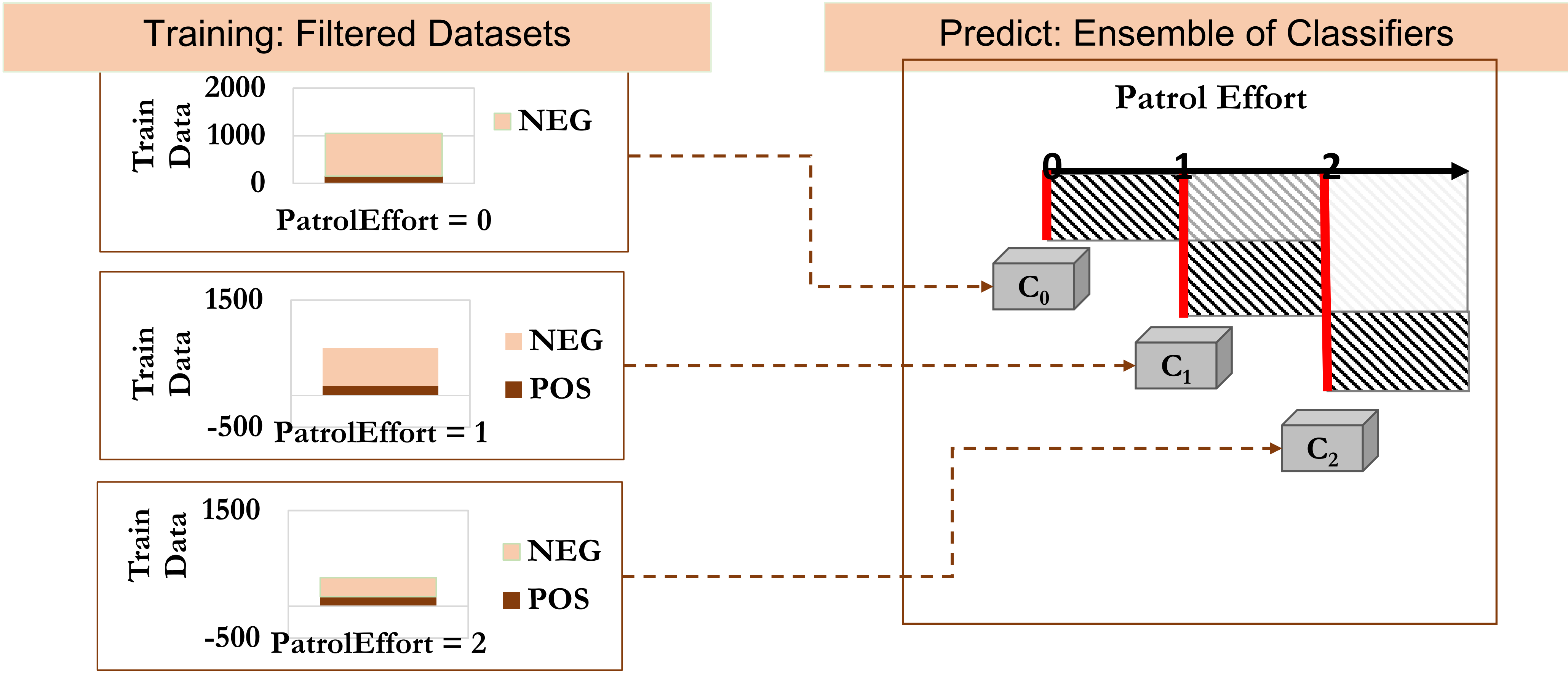
	Area1	Area2
Area1	4, -3	-1, 1
Area2	-5, 5	2, -1

Learning Adversary Response Model: Uncertainty in Observations


Nguyen


Gholami



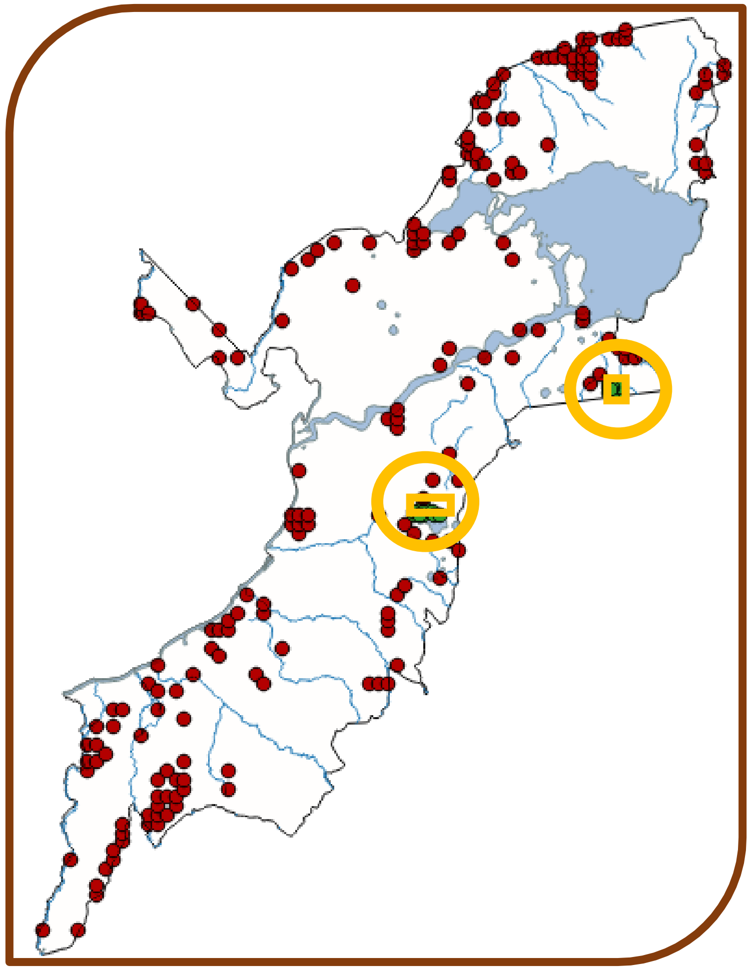


PAWS: First Pilot in the Field
(AAMAS 2017)

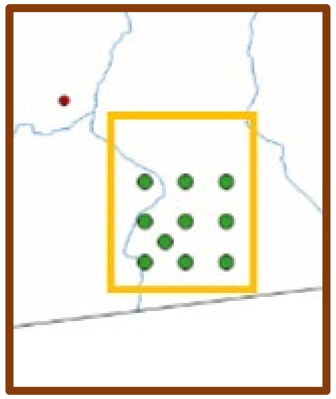

Ford

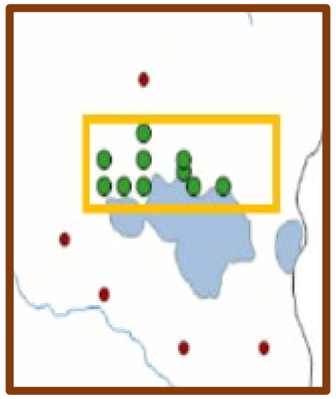
Gholami

- Two 9-sq.km areas, infrequent patrols

- Poached elephant
- 1 elephant snare roll
- 10 Antelope snares



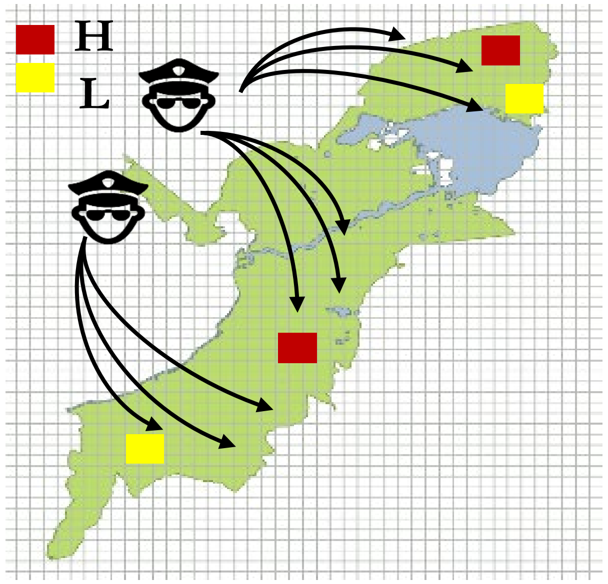


PAWS Predicted High vs Low Risk Areas: 3 National Parks, 24 areas each, 6 months

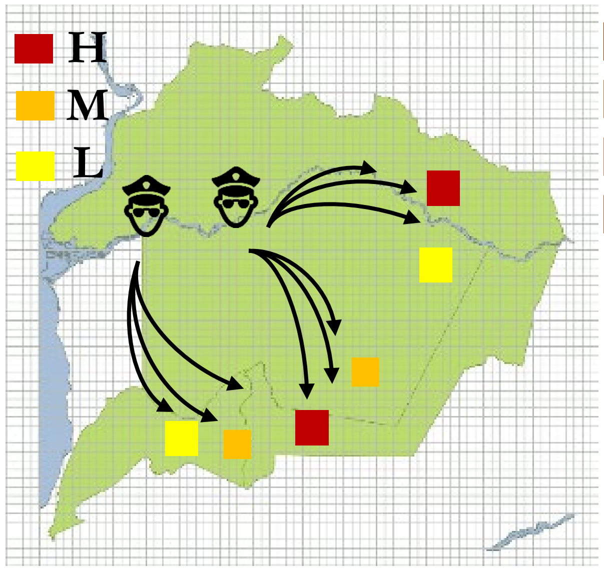
(ECML PKDD 2017, ICDE 2020)

Xu

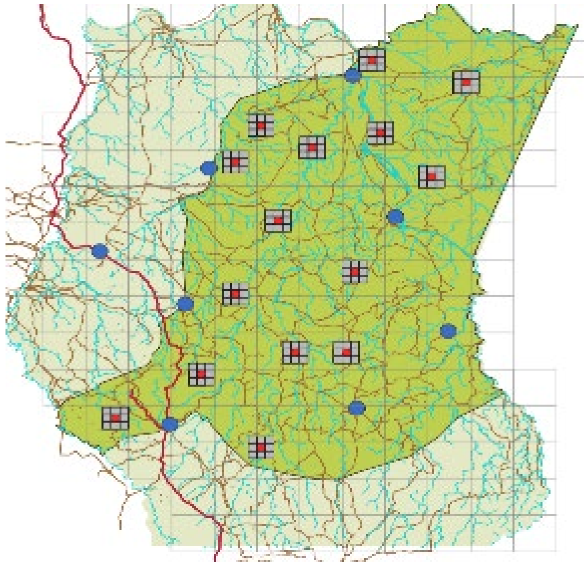
Gholami



Queen Elizabeth National Park

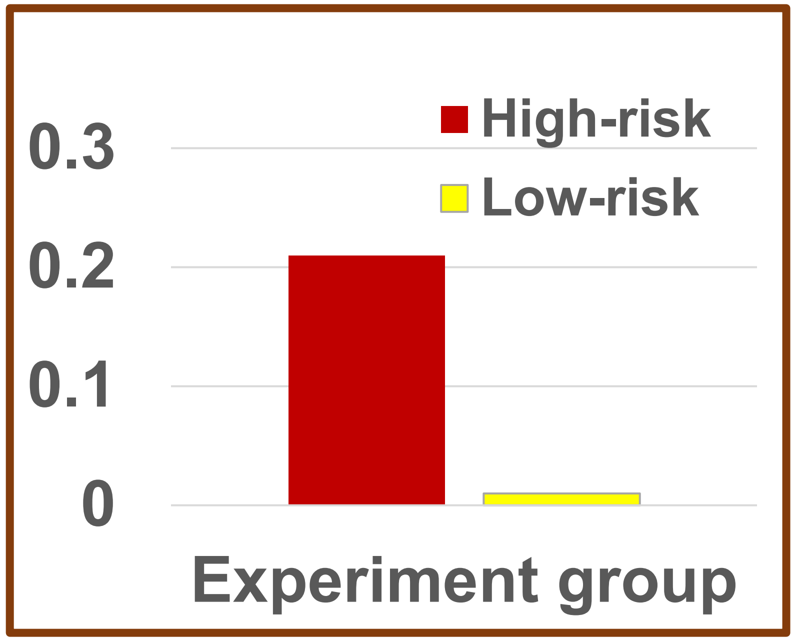


Murchison Falls National Park

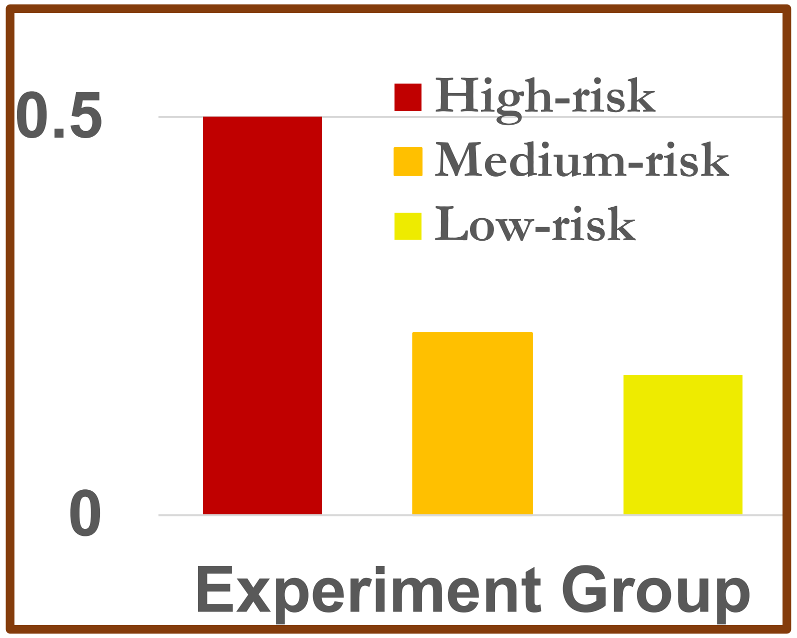


Srepok Wildlife Sanctuary

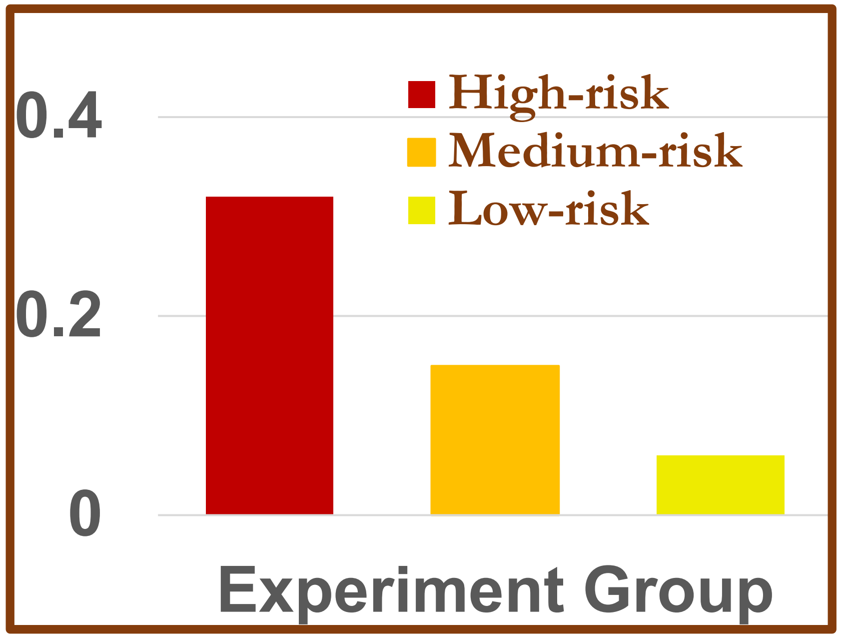
Snares per patrolled sq. KM



Snares per patrolled sq. KM



Snares per patrolled sq. KM



PAWS Real-world Deployment Cambodia: Srepok Wildlife Sanctuary

(ICDE 2020)

Xu

2019 PAWS: *521 snares/month*

vs

2018: *101 snares/month*

2021 PAWS

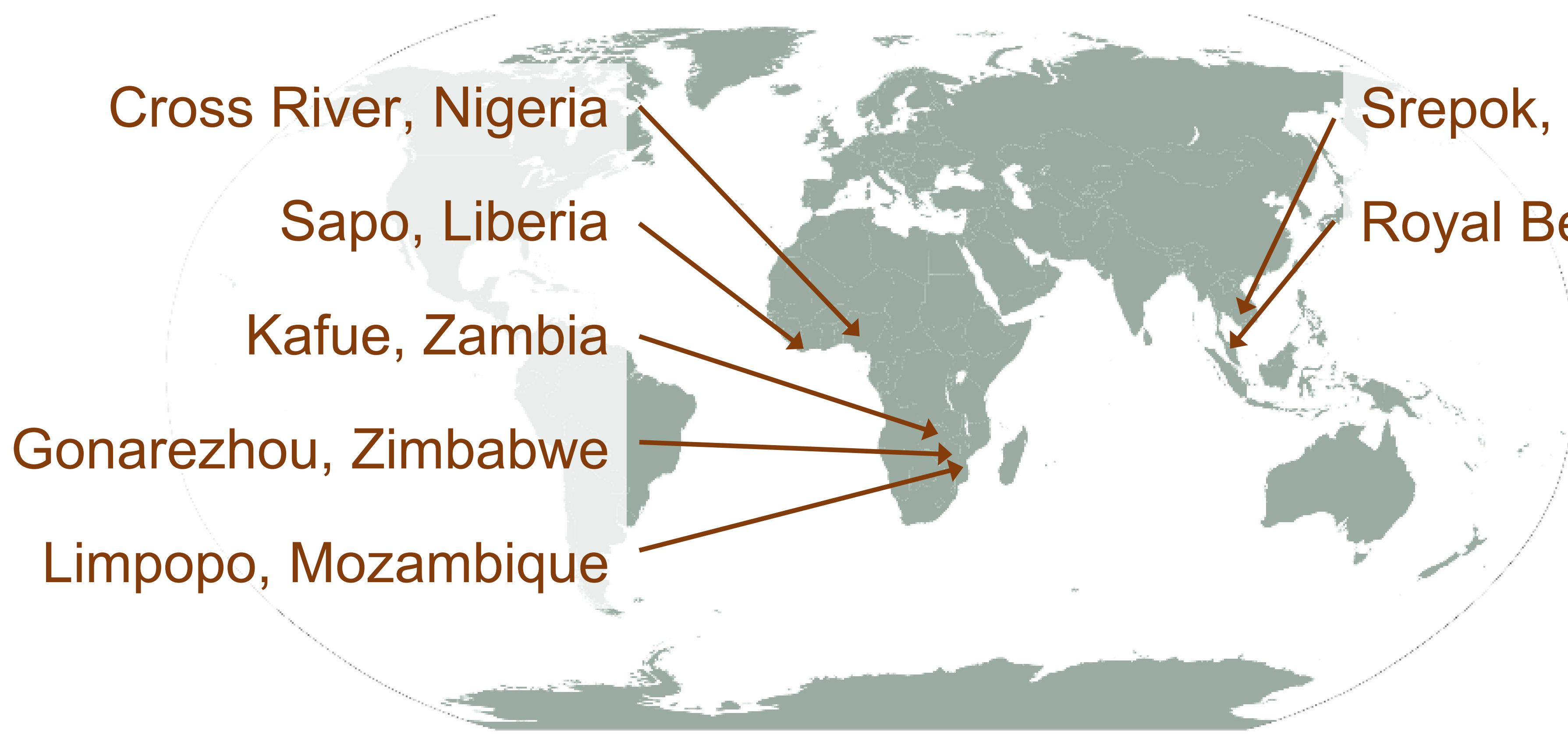
1,000 snares found in March

PAWS GOES GLOBAL with SMART platform!!

Xu

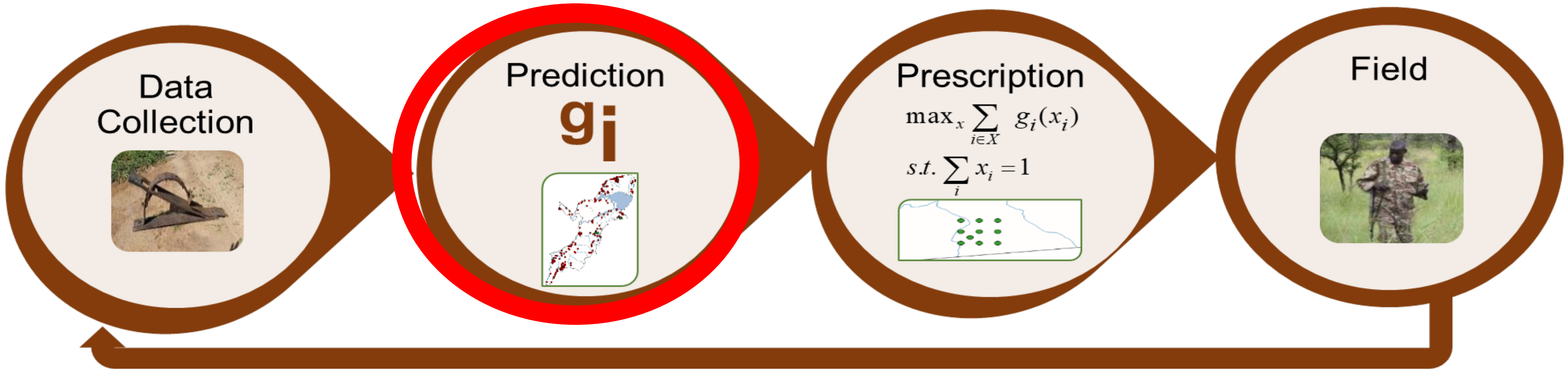
SMART WWF PEACE PARKS FOUNDATION · THE GLOBAL SOLUTION · WCS

Protect Wildlife
800 National Parks
Around the Globe

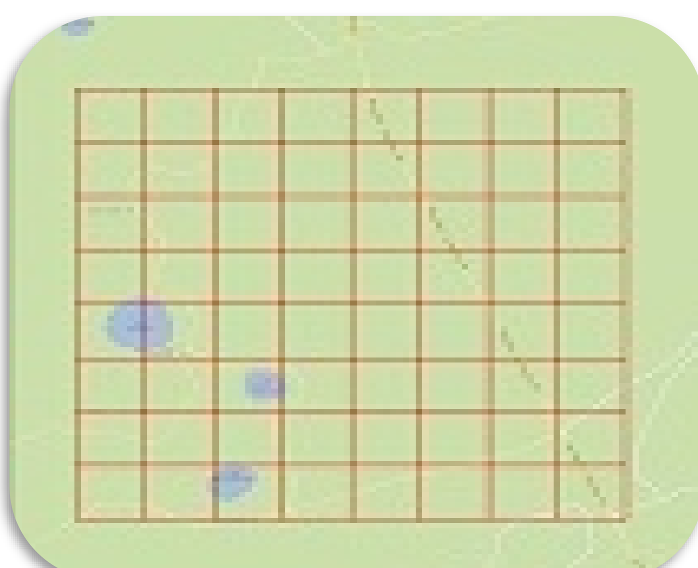


Stackelberg Security Games to Prescribe Patrols

Xu



- *Do poachers get deterred by patrols?*



	Area1	Area2
Area1	4, -3	-1, 1
Area2	-5, 5	2, -1

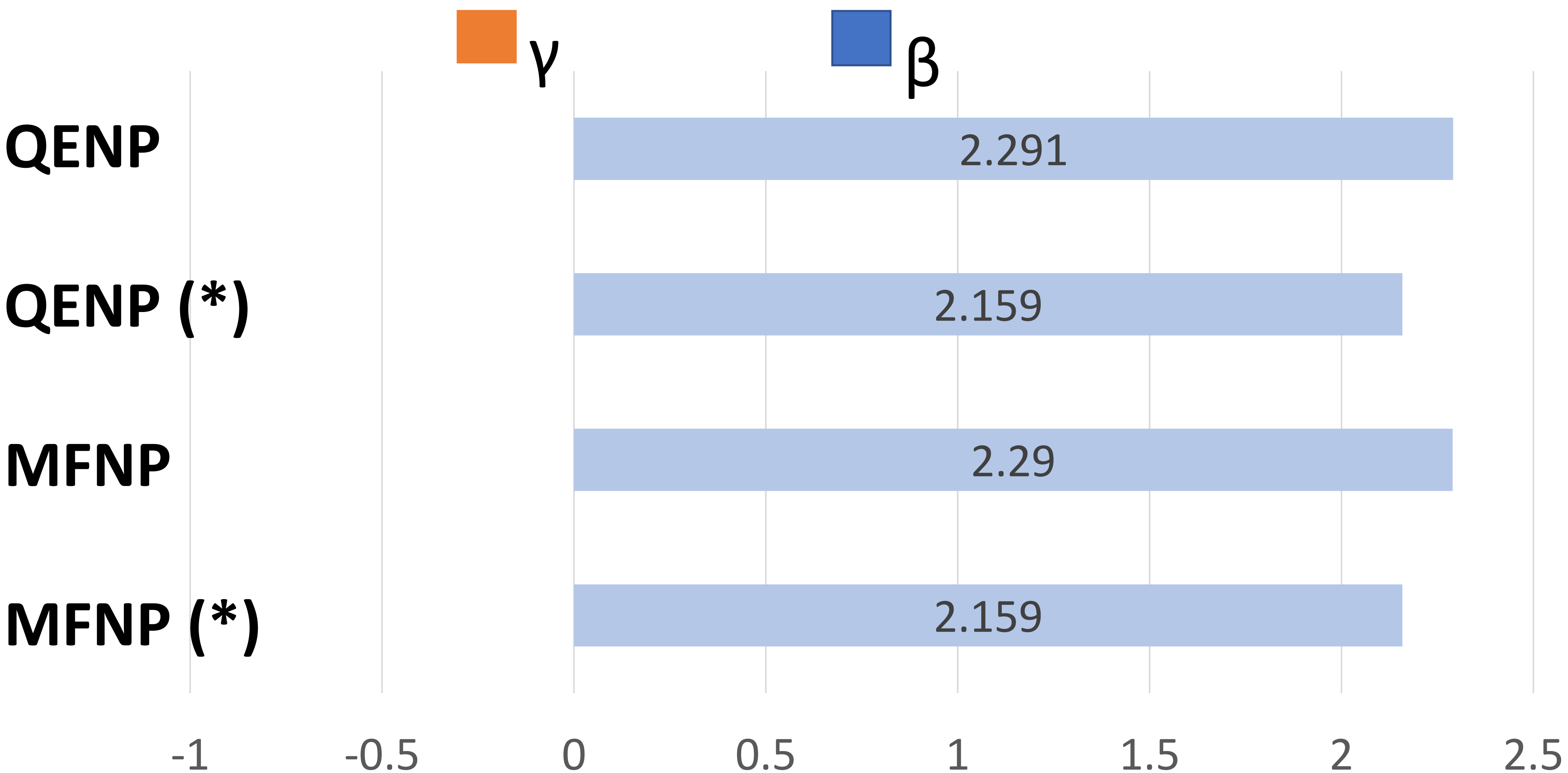
Is Adversary observing & Reacting to Patrols?
YES! Adversaries deterred by patrols


Xu


Perrault

- Logistic regression model

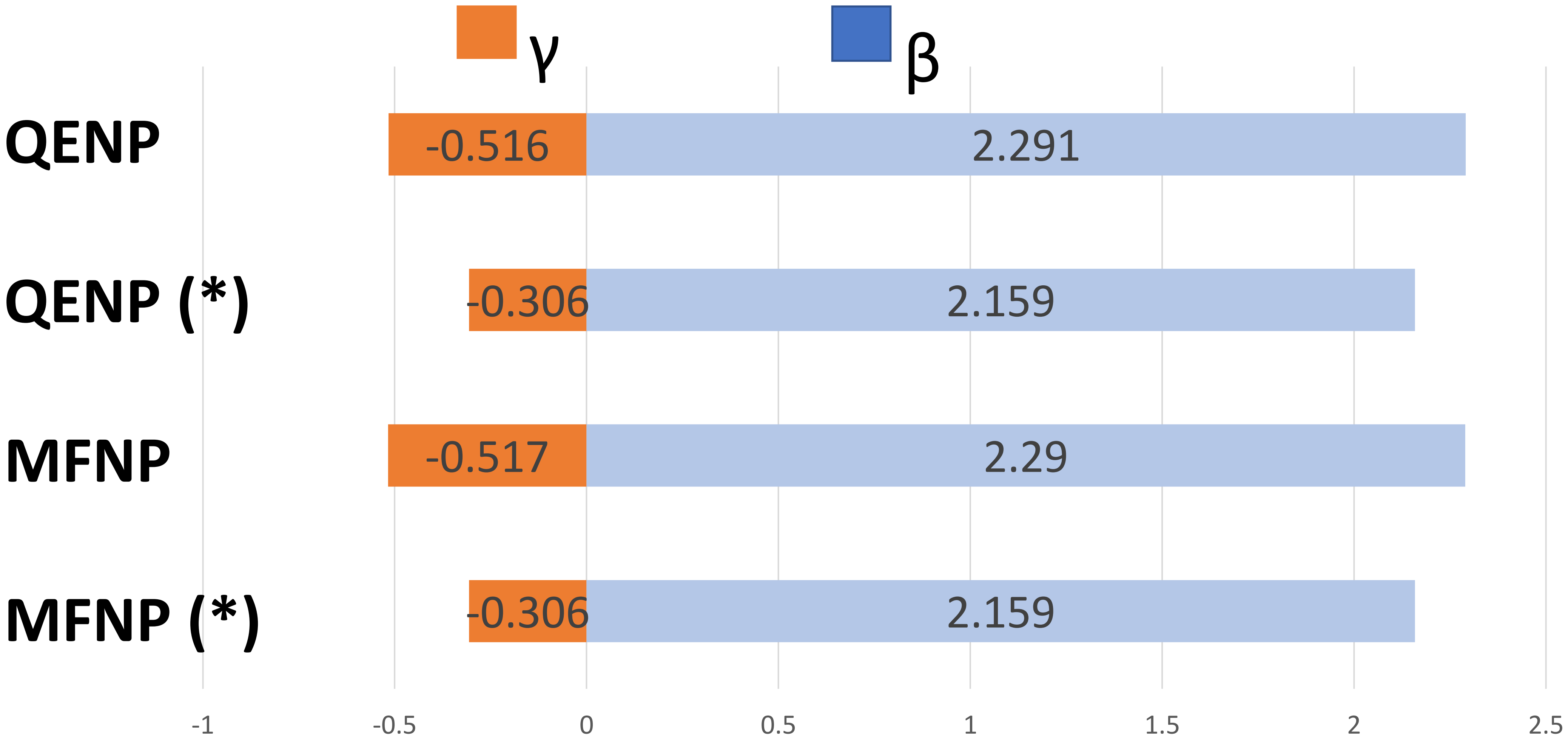
$$a_i + \gamma \cdot \texttt{past\_effort} + \beta \cdot \texttt{current\_effort}$$



Is Adversary observing & Reacting to Patrols?
YES! Adversaries deterred by patrols

- Is adversary observing & reacting to patrols? Logistic regression model

$$a_i + \gamma \cdot \texttt{past\_effort} + \beta \cdot \texttt{current\_effort}$$



MIRROR: Handling Uncertainty in Poacher Model
Simulation Results (UAI 2021)


Xu

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

$$\max_{x \in \Delta^{|P|}} \min_{\mu,\sigma} \sum x_p \frac{(Outcome(p))}{OPT(\mu,\sigma)}$$

Algorithm

Choose Patrol strategy p ∈ P generating mixed strategy "$x \in \Delta^{|P|}$"

vs

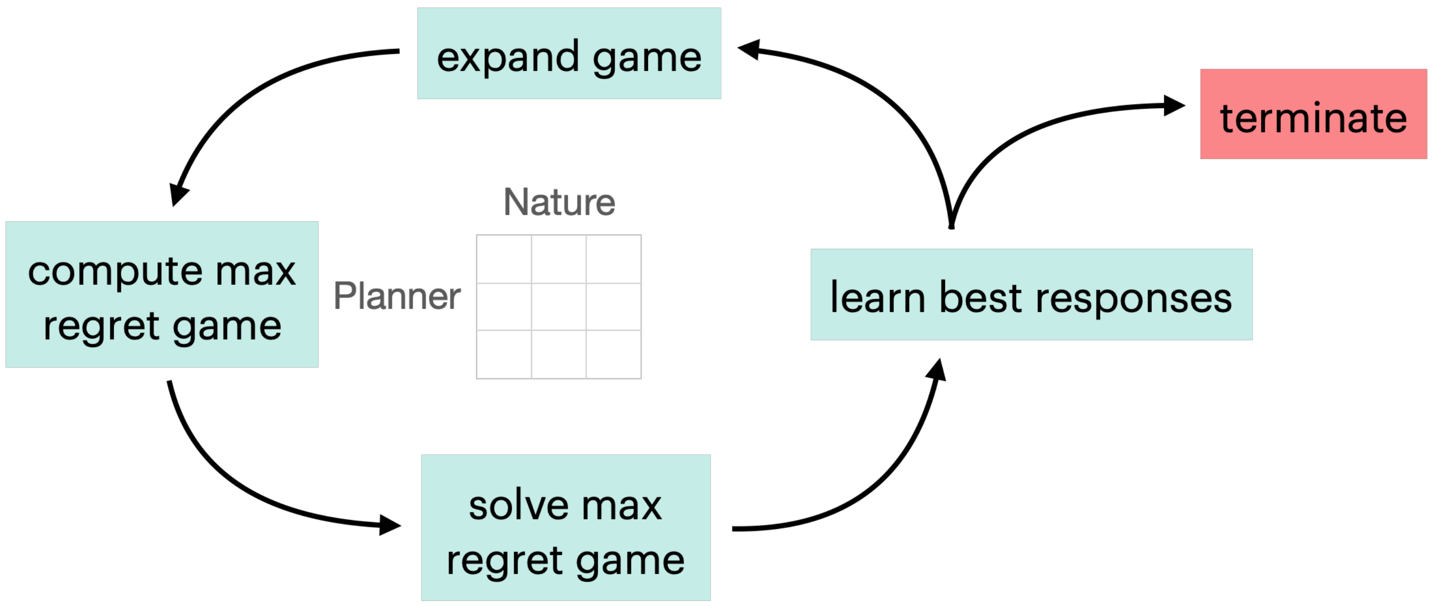
Nature

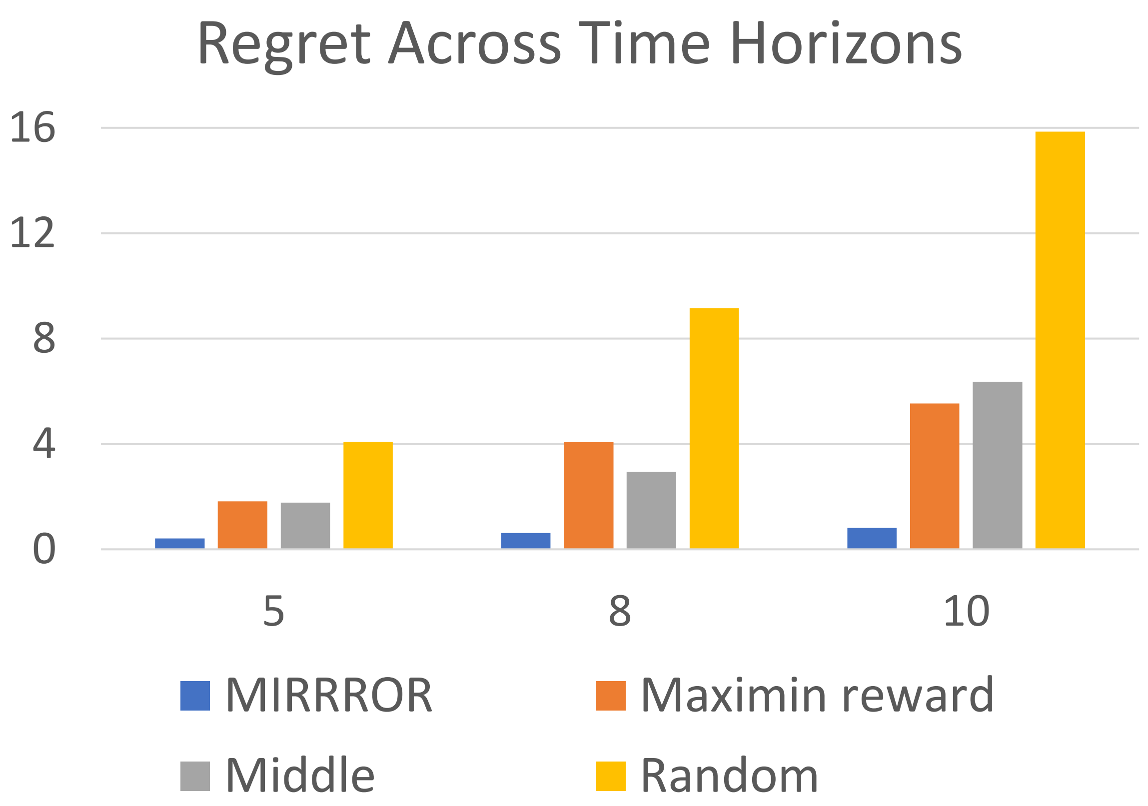
Chooses parameters of poacher model

μ,σ

MIRROR: Deterrence-Based Patrol Planning
Simulation Results (UAI 2021)

- Double oracle: Iteratively solve for equilibrium
- Final strategy is guaranteed to minimize max regret



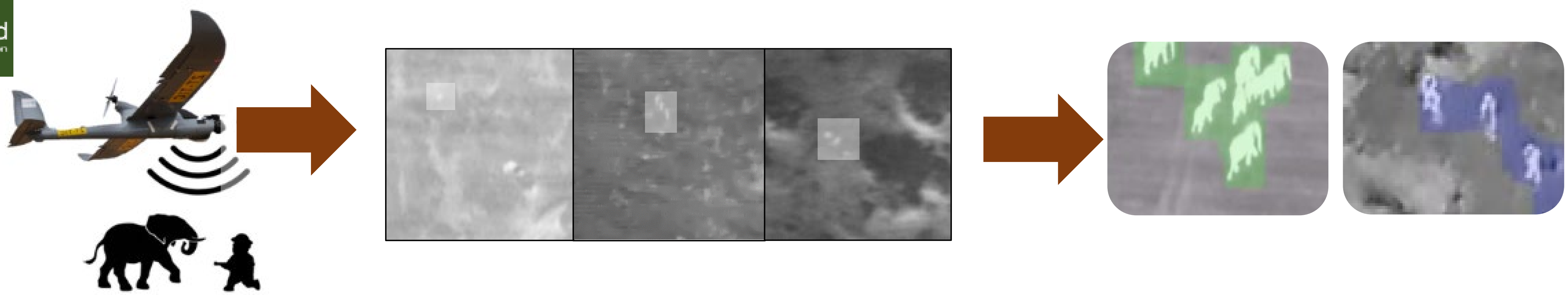


Next Steps:
Integrating Real-Time “SPOT” Information

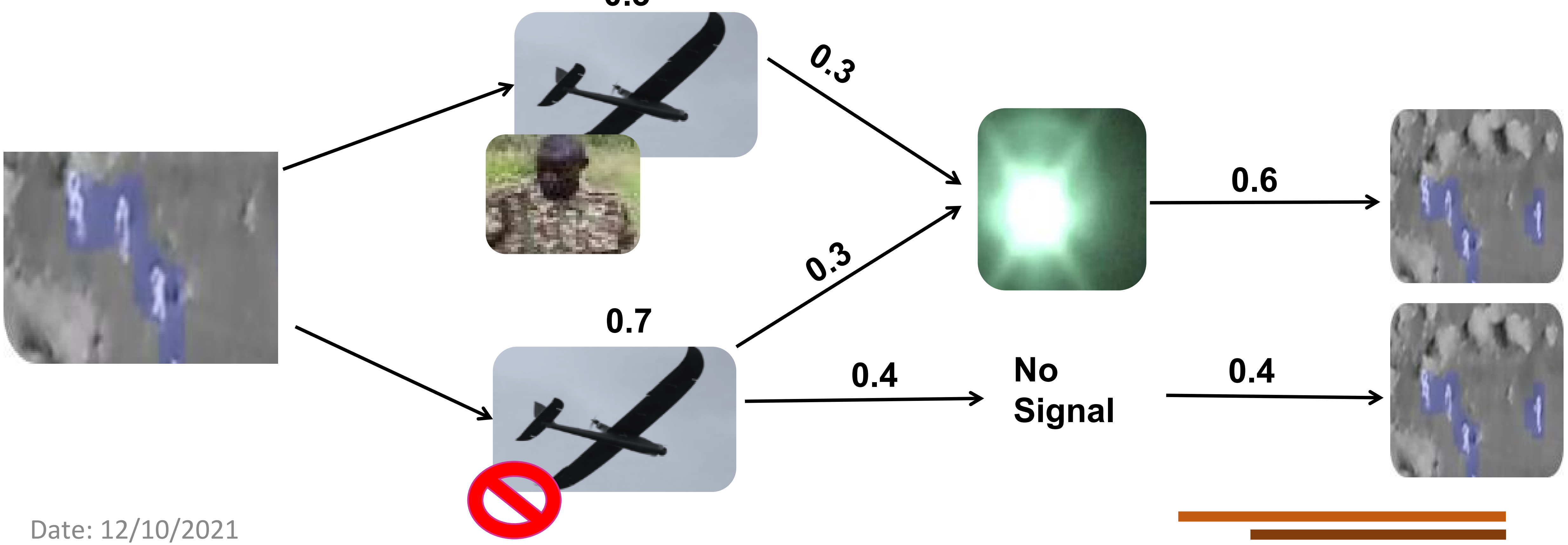
(IAAI 2018, AAAI 2018, AAAI 2020)


Bondi

Air Shepherd
The Lindbergh Foundation



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



Next Steps: Data Scarce Parks

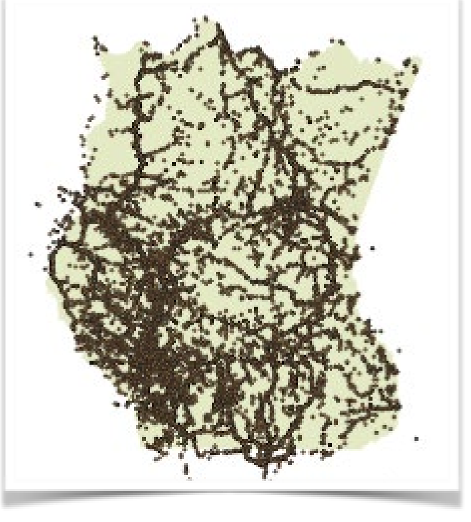
Xu

exploitation

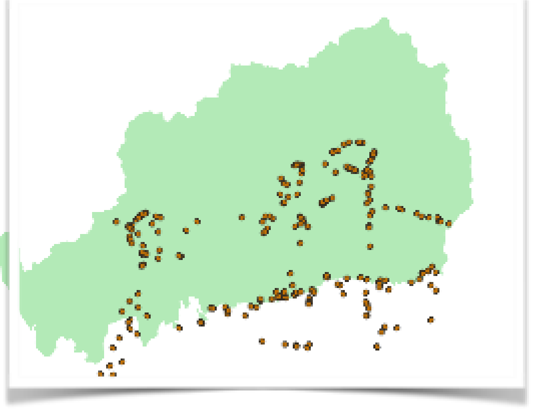
Data-rich parks: build predictive models to plan patrols

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

exploration



Srepok, 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)


Xu

Theorem: With time horizon T, regret bound of LIZARD is $Regret(T) \leq O\left(T^{\frac{2}{3}}\right)$

LIZARD algorithm exploits decomposability, smoothness, monotonicity

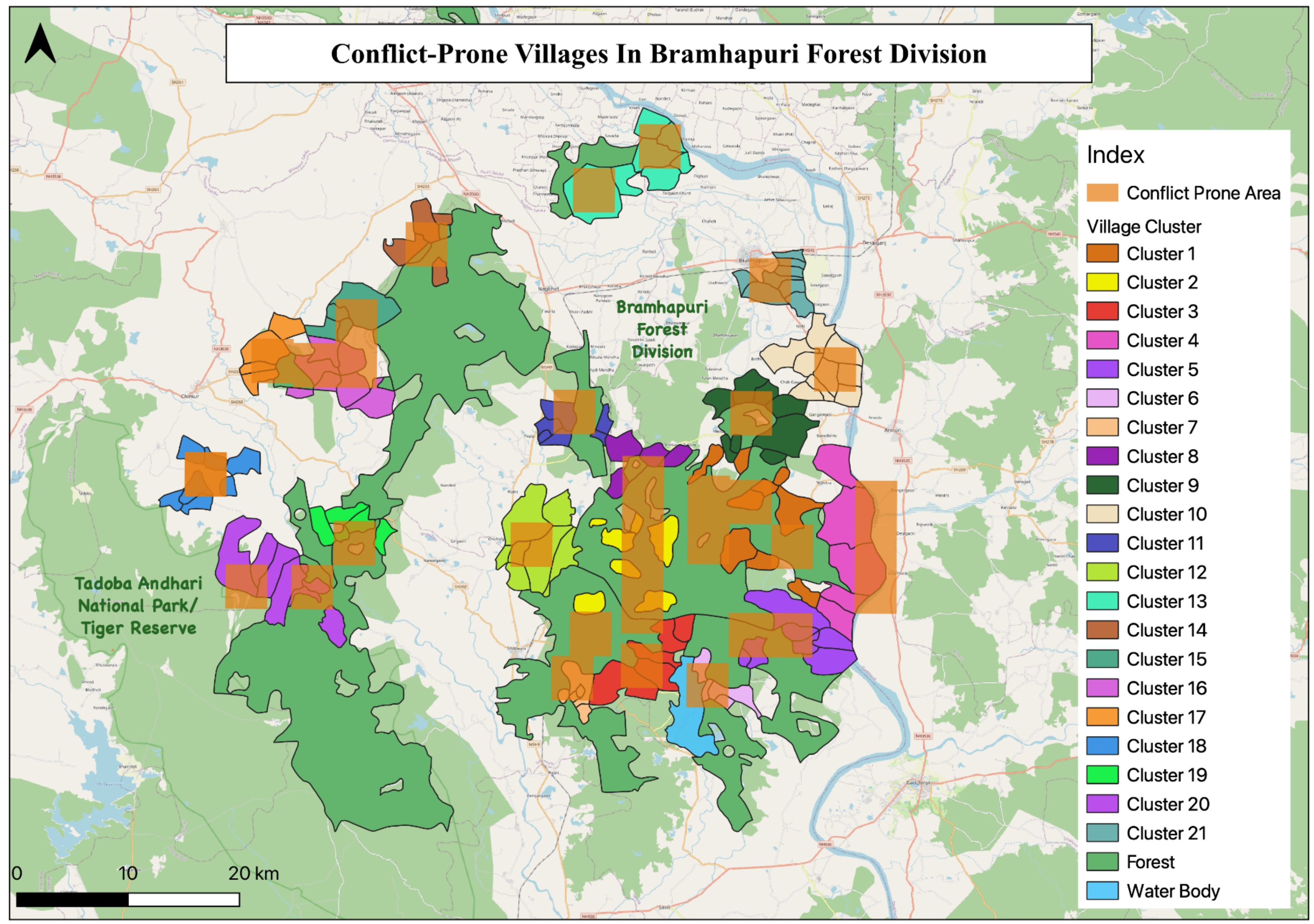
- *Input:* N targets with features, stochastic poacher places snares at targets
- *Output:* Patrol effort per target $\leq$ budget B
- Reduce regret wrt OPT, optimal patrol effort, for capturing snares

Preventing Human-Wildlife Conflict

(Joint work with P. Varakantham, WCT)

Ghosh

- Most forest areas in India are multi-use: wild animals & humans co-habit, conflict
- Our predictions used to distribute funds in Bramhapuri division, Maharashtra

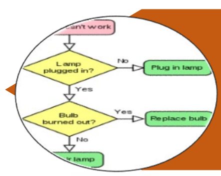


Future: AI for Social Impact (AI4SI)


Achieving social impact & AI innovation go hand in hand


Empower non-profits to use AI tools; avoid being gatekeepers to AI4SI tech

Data to deployment: Not just improving algorithms


Important to integrate AI innovations in NGO normal workflow


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

THANK YOU

#AIforSocialImact

@MilindTambe_AI