Deploying Novel Exploration Techniques (NETs) for Malaria Policy Interventions

An Investigation into Augmenting Human Decision-Making

Oliver Bent & Professor Steven Roberts
Machine Learning Research Group (MLRG)
oetbent@robots.ox.ac.uk

AIMS
Autonomous Intelligent Machines & Systems

Abstract

The task of decision-making under uncertainty is daunting, especially for problems with significant complexity. Healthcare policy makers globally are making decisions with challenging constraints and they have limited tools to help them make data-driven decisions. In this work we frame the process of finding an optimal malaria policy as a stochastic multi-armed bandit problem, and implement three agent based strategies to explore the policy space. We apply a Gaussian Process regression to the findings from each agent, naturally accounting for the stochasticity associated with simulating the vector-borne transmission of malaria. The Agent generated policy recommendations are compared with human made interventions taken from the literature.

Introduction

There has been significant progress in the prevention and control of malaria over the last 15 years. Reductions in mortality rate and number of new cases. But many countries in Sub Saharan Africa still rely heavily on external funding for malaria prevention and control, which in recent years is declining [4]. Separate decision makers (e.g., NGOs, Governments and Charities) must be able to explore the possible set of actions for appropriate malaria interventions for their environments. Such possible policies may include a mix of actions like the distribution of long-lasting insecticide-treated nets (ITNs), indoor residual spraying (IRS), vector larvicide in bodies of water, and malaria vaccinations. The space of possible policies for malaria interventions is daunting and inefficient for human decision makers to explore. This work uses the OpenMalaria codebase of stochastic transmission models for malaria simulation. Currently used by researchers to evaluate the impact of various malaria control interventions. OpenMalaria therefore provides a platform to create a simulation environment from which an AI agent may explore optimal policies for the control of malaria. This work will use a parameterisation of OpenMalaria models for Rachuonyo South district in Kenya [3].

Stochastic Multi-Armed Bandit

Exploring malaria policies from OpenMalaria simulations is posed as a multi-armed bandit problem; to efficiently determine high performing policies for a simulated human population.

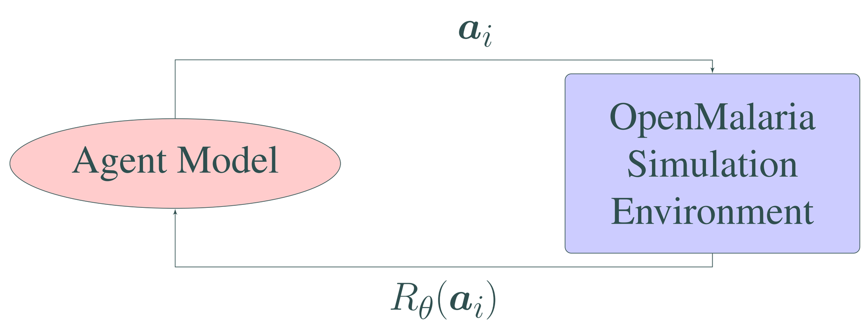


Figure 1: Policies $\boldsymbol{a}_i$ are chosen by the Agent Model which receives Rewards $R(\boldsymbol{a}_i)$

There is no access to state transitions at run-time for OpenMalaria simulations. Instead we solve for the problem of making a 'one-shot' policy recommendation for the intervention period (5 years).

Action

The main control methods used in Rachuonyo South district are: mass-distribution of long-lasting insecticide-treated nets (a_{ITN}); Indoor Residual Spraying (a_{IRS}) with pyrethroids; and the prompt and effective treatment of malaria. $\boldsymbol{a} = \{a_{ITN}, a_{IRS}\}$ where a_{ITN} and $a_{IRS} \in (0, 1]$, describe the coverage of the intervention for the simulated population.

Figure 2: Visual description actions: long-lasting insecticide-treated nets and indoor residual spraying

Reward

The reward for each policy $R_\theta(\boldsymbol{a}_i)$ is stochastic through the parameterisation of the simulation θ, which generates a randomised distribution of parameters for the OpenMalaria simulation. The magnitude of the reward is given by an economic cost-effectiveness analysis of the stochastic simulation output. Disability adjusted life years (DALYs) [1], are a metric defined by the total years of life lost due to fatality linked with contraction of the disease, and number of years of life with disability as a result of the disease. We use a discount factor $\gamma = 0.97$ to discount the value of future years of life lost, and a life expectancy of 46.6 years for Rachuonyo South District. We simulate two types of costs, the healthcare system costs (HSC), and intervention costs (IC). For each malaria episode that a patient seeks treatment, hospitals incur costs to treat the disease, to manage the patent's recovery process, and also to deal with the patient's death if that were to occur. Agent models receive rewards based on the cost effectiveness of a policy. This is a score often used by human researchers evaluating the impact of a policy, calculated as the cost per DALY averted (C_{DA}).

$$C_{DA} = \frac{HSC_{\text{int}} - HSC_{\text{no int}} + C_{\text{int}}}{DA} \tag{1}$$

Agent Models

Three different agents (GP-ULCB, Genetic Algorithm and Batch Policy Gradient) perform sequential batch exploration, towards optimisation of an unknown stochastic reward function R. Due to the computational expense of calculating $R(\boldsymbol{a}_i)$ and the size of A, we wish to find solutions of maximal reward in as few iterations i as possible. Approximating $\boldsymbol{a}^* = \text{argmax}_{\boldsymbol{a} \in A} R(\boldsymbol{a})$ without prohibitively expensive computation for all possible policies, using a subset $A_c \in A$ of the policy space.

Results

Human decisions are made using OpenMalaria as a research tool [2]. Researchers state the current policy of 56% a_{ITN}, 70% a_{IRS} for Rachuonyo South is the most cost-effective with regards to C_{DA}, while they recommended that increasing to 80% a_{ITN} and 90% a_{IRS} would have the greater health impact. In this work we use the same stochastic parameterisation θ of OpenMalaria, but instead answer the question for the decision maker: what policy decision can be made for the next 5 years to improve cost-effectiveness? Our results indicate that the best strategies *reduce* a_{IRS}, while maintaining the levels of a_{ITN}. These findings are extracted from the surface maxima of the posterior mean $\mu(\boldsymbol{a})$ through gaussian progress regression of rewards $R_\theta(\boldsymbol{a})$ collected by each respective agent. Figure 3 presents these surfaces. For reference, the cost of the human recommended policy (80% a_{ITN} and 90% a_{IRS}) is an order of magnitude higher than our recommended strategies.

GP-ULCB				Genetic Algorithm				Batch Policy Gradient			
Policy	C_{DA}	DA	C_{int}	Policy	C_{DA}	DA	C_{int}	Policy	C_{DA}	DA	C_{int}
$\{\mathbf{61,55}\}$	-9.94	1343	1400	$\{\mathbf{55,38}\}$	13.74	1744	845	$\{\mathbf{55,41}\}$	14.56	1507	1616
$\{56,53\}$	-9.83	1297	1203	$\{60,39\}$	14.25	1833	634	$\{55,5\}$	14.58	1527	1359
$\{65,55\}$	-9.68	1286	1726	$\{55,48\}$	14.56	1734	1739	$\{55,30\}$	14.58	1469	1574

Table 1: Top performing policies. Policy: $\{a_{ITN}\%, a_{IRS}\%\}$, C_{DA}: Cost per DALY Averted USD, DA: DALYs Averted, C_{int}: Intervention Costs USD. *notes*: -ve values indicate health-system savings greater than intervention cost

Evaluation

These methods give a comprehensive evaluation of exploring the cost-effectiveness space for a policy of two interventions. Such insight is often missing from empirical studies for malaria interventions, which may seek to determine how much of single intervention may be implemented to maximise a particular performance metric. A limitation is that we did not explore deploying interventions at different times of year, or if multiple policies could have been concurrently deployed in a population. Furthermore, the simulation environment did not permit interventions to be targeted to subsets of the population (e.g. young children). Finally, this work is specific to one studied location in Western Kenya, and the generalisation of the agents' insights across expansive environments e.g. Sub-Saharan Africa is yet to be explored.

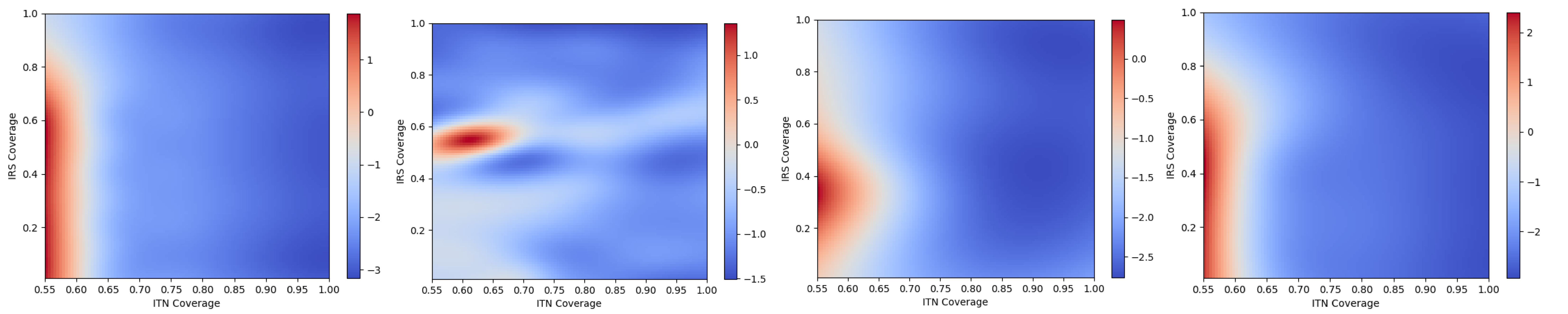
References

[1] Christopher J. L. Murray and Alan D. Lopez. The global burden of disease: a comprehensive assessment of mortality and disability from deceases, injuries and risk factors in 1990 and projected to 2010. *Harvard University Press*, 1:1–35, 1996.

[2] Erin M Stuckey, Jennifer Stevenson, Katya Galactionova, Amrish Y Baidjoe, Teun Bousema, Wycliffe Odongo, Simon Kariuki, Chris Drakeley, Thomas A Smith, Jonathan Cox, and Nakul Chitnis. Modeling the cost effectiveness of malaria control interventions in the highlands of western Kenya. *PLoS ONE*, 9(10), 2014.

[3] Erin M Stuckey, Jennifer C Stevenson, Mary K Cooke, Chrispin Owaga, Elizabeth Marube, George Oando, Diggory Hardy, Chris Drakeley, Thomas A Smith, Jonathan Cox, and Nakul Chitnis. Simulation of malaria epidemiology and control in the highlands of western Kenya. *Malaria Journal*, 11(1):357, oct 2012.

[4] Peter Winskill, Mark Rowland, George Mtove, Robert C Malima, and Matthew J Kirby. Malaria risk factors in north-east Tanzania. *Malaria Journal*, 10(1):98, 2011.



(a) Even Sample of Batch size B
(b) GP-ULCP
(c) Genetic Algorithm
(d) Batch Policy Gradient

Figure 3: Plots of $-\log(R_\theta(\boldsymbol{a}))$ under different Agent model selection of $\boldsymbol{a}_i$. Showing relative performance for a decision maker of different policies where there are existing interventions of $a_{ITN} = 0.55$ and $a_{IRS} = 0.7$