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.

Action

The main control methods used in Rachuonyo South district are: mass distribution of long-lasting insecticide-treated nets (LLINs), Indoor Residual Spraying (IRS) with pyrethroids, and the prompt and effective treatment of malaria. \( \alpha = (\alpha_{LLIN}, \alpha_{IRS}) \) where \( \alpha_{LLIN} \) and \( \alpha_{IRS} \) are the coverage of the intervention for the simulated population.

Reward

The reward for each policy \( R(a_i) \) is stochastic through the parameterisation of the simulation \( \theta \), which generates a randomised distribution of parameter 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 disability 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 patient’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 (\( C_{DA} \)).

\[
C_{DA} = \frac{HS_{HSC} + HS_{IC} + IC_{int}}{DA}
\]

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 \( R(a_i) \), and the size of \( A \), we wish to find solutions of maximal reward in as few iterations as possible. Approximating \( a_i \equiv argmax_{a} R(a) \) without prohibitively expensive computation for all possible policies, using a subset \( A \), of the policy space.

Results

Human decisions are made using OpenMalaria as a research tool [2]. Researchers state the current policy of 56% ITN, 70% IRS for Rachuonyo South is the most cost-effective with regards to DALY, while they recommended that increasing to 80% ITN and 90% IRS would have the greater health improvement. In this work we use the same stochastic parameterisation \( \theta \) 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 IRS, while maintaining the levels of ITN. These findings are extracted from the surface maxima of the posterior mean \( \mu(a) \) through gaussian process regressions of rewards \( R(a_i) \) collected by each respective agent. Figure 3 presents these surfaces. For reference, the cost of the human recommended policy (60% ITN and 90% IRS) is an order of magnitude higher than our recommended strategies.

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