Integrating Policy Optimization in Epidemic Modeling

PI: Azza Abouzied, NYU Abu Dhabi (azza@nyu.edu)

Abstract

To support a policy optimization framework that suggests a plan of interventions to counteract the spread of an epidemic, the underlying epidemiological modeller must be (i) *scalable*, to allow the selection of an optimal policy from thousands of simulations of different interventions (ii) *transparent*, to allow policy-makers to reason about the model's policy choices and predictions and (iii) customizable, to allow policy-makers to ensure that disease spread dynamics and intervention implementations reflect the realities of their administrative locales. We propose to build an epidemic simulator that supports these goals. In particular, we address the design of a domain specific language that allows the declarative and easy specification of disease models and interventions, and tracks the provenance of a wide range of disease and intervention parameters to ensure reproducibility.

1   Project Description

Developing policies to suppress the spread of and mitigate the severity of the novel coronavirus pandemic is a top global public health priority. While a multitude of non-pharmaceutical interventions currently exist (case isolation, household quarantine, school or workplace closures, travel restrictions, aggressive contact tracing and testing, increasing hospital bed capacity), as well as potential pharmaceutical interventions such as vaccines and curative or prophylactic antivirals, the societal, economic and health costs, efficacy and implementation specifics of each intervention is unique to a given locale. Our goal is to help policy-makers decide when, how and for how long one should impose travel restrictions or closures, how much testing or bed capacity should one build up to, as well as which pharmaceutical treatments to purchase to control epidemics in a cost-effective manner with respect to the unique characteristics of an administrative locale.

Our strategy to provide such support is to model not only the disease through standard, deterministic, epidemiological compartment models but also to model the cost and impact of several, different interventions on disease spread. Unlike prior work, our framework explores combinations of interventions over time, using Monte Carlo simulations, to determine an optimal, cost-effective, plan of timed and parameterized interventions. For example, our framework might propose how much testing capacity up to a limit should a policy-maker plan for (and when to increase and decrease this capacity) as the disease progresses, taking into consideration available and potential future resources, as well as the impact of other concurrent interventions. Current models assume a fixed intervention set and then demonstrate its effectiveness, leaving the complex, combinatorial task of putting forward a cost-effective policy that combines and plans different interventions up to policy-makers [6].

1.1   Research Challenges

Our project tackles three main research challenges, which we briefly describe below. Of the three, this proposal focuses only on the last challenge but we describe all challenges to demonstrate our holistic approach to integrating policy optimizations within epidemic modeling.

1. ***Dealing with uncertainty.***   With a novel disease many key parameters are uncertain. For example, the exact case fatality rate of COVID-19 is still unclear with some countries reporting rates lower than 1.2% (UAE) and others reporting rates as high as 13.9% (Italy) [2]. The issue is confounded by many factors including how well positive cases are identified in the first place and whether cause of death is attributed to COVID-19 or other [4]. Moreover, as interventions are introduced, their efficacy can be quite uncertain. For example, the sensitivity and specificity of serology tests in identifying the presence of antibodies, varies widely across different testing kits [4]. In spite of this uncertainty, policy-makers still need to design adequate policies that minimize loss of life at reasonable cost. Uncertainty is inherent in our proposed modeling framework: for each locale, epidemiologists can specify a plausible range of disease and intervention parameters using currently available data.

2. ***Customization to a specific locale.*** How a disease spreads varies widely across geographic regions. Most epidemic models account for this by using underlying information such as a city's population density, demographics, transportation networks, border- and air-travel models, etc. The impact of an intervention also varies widely across regions depending on societal and economic factors. For example, the efficacy of government directives such as stay-at-home orders is heavily influenced by a society's trust in authorities and the effectiveness of public health risk communication [1]. Not only is the efficacy of an intervention unique to a locale but also its cost. Countries with shared open borders and similar demographics, such as Belgium and Netherlands, have opted to implement different interventions unique to their locales. It is thus crucial to allow the data-driven fine-tuning of disease and geographic parameters as well as of interventions. *We are specifically interested in the unique geographic, societal and economic characteristics of the UAE.* We already use publicly available data on cases, fatalities, hospital and testing capacity, population density per geographic area, transportation models, etc. We hope to work closely with the UAE's key decision-makers and epidemiologists to better customize our disease models and interventions to the UAE.

3. ***Easy specification and accurate modeling of the impact of different interventions.*** Current epidemic simulators require a fair amount of custom programming to introduce or combine interventions [5, 3]. Custom scripts hinder the overall transparency of the model and make it difficult to transfer the model from one locale to another or to easily generalize its findings. What is required is a domain specific language (DSL) to simplify the specification of interventions over models. By eliminating the time-consuming and error-prone effort of manually scripting the effect of different intervention sets and by ensuring the transparency and interpretability of intervention specifications, an easy-to-understand DSL ensures the reproducibility of a framework's proposed policies. A clear specification with well-defined intervention parameters such as costs, limits, etc. allows different stakeholders to reason about and easily modify interventions to better suit their administrative locales. In the following section 1.2, I articulate specific technical challenges that this proposal hopes to tackle.

1.2   Preliminary Work and Methods

The progression of infectious diseases with an incubation period such as COVID-19 is often modeled by the progression of individuals through compartments such as **S**usceptible, **E**xposed, **I**nfectious, and **R**ecovered in the SEIR compartment model. The *infectious rate*, β, describes the rate of spread ($S \xrightarrow{\beta} E$): it represents the probability of transmitting disease between a susceptible and an infectious individual. The *incubation rate*, σ, is the rate at which latent individuals becoming infectious ($E \xrightarrow{\sigma} I$). The average duration of incubation is $1/\sigma$. Finally, the *recovery rate*, γ represents the rate of infected individuals recovering ($I \xrightarrow{\gamma} R$) and is determined by the average duration of infection.

This basic model can be elaborated to include more compartments to better model disease characteristics such as vaccinated individuals, fatalities, re-infection, etc. Moreover, distinct sub-populations can be modeled by different SEIR models that interact with each other. For example the Spatio-Temporal epidemiological modeler [5] creates distinct compartments for each geographic region where the change in the population of exposed individuals in each region takes into account only the population of infectious individuals in neighboring regions. This simulates the uniform mixing of individuals across transportation borders.

Compartment models can be further categorized into two types: (i) deterministic models, or (ii) stochastic, agent-based models. In deterministic models changes across a population are expressed as deterministic differential equations. In stochastic models, each individual (or agent) i has a probability of transitioning from one compartment to another determined by $\beta_i, \sigma_i, \gamma_i$ parameters, their personal characteristics (such as workplace, age) and their contact network (such as the state of other agents they share a workplace with, live with, etc.). Both models make different trade-offs.

Deterministic models are simple to describe and easy to scale as the net outcome is determined by applying a series of differential equations across time. This inherent scalability makes them suitable for our overall policy-optimization framework as we simulate the effect of thousands of different policies to determine the most optimal plan of interventions.

By contrast, the agent-based model updates the state of each agent at every time step and hence does not scale well. Most agent-based models have an upper limit on the population size ($\approx 10^6$ individuals). Their main benefit is that they allow fine-grained control of agent and disease behavior, which in turn enables a more intuitive specification

of how disease spreads in schools, specific workplaces, shared living quarters, etc and how specific interventions can curtail such spread. Agent-based models, however, can be extremely complex to design and implement. An often cited UK agent-based model for COVID-19 [6] has roughly 450 parameters that describe individuals and how they interact as well as a complex code-base to implement a variety of disease spread dynamics (e.g. spread in hotels vs. other spaces) and interventions. Epidemiologists often use both models in tandem applying deterministic models at larger scale (e.g. across an entire country) and stochastic models at smaller scale (e.g. within a county).

In this work, we hope to build scalable epidemiological models that lend themselves well to policy optimizations. To this end, we hope to create a deterministic epidemic model simulator with the following features. The simulator:

1. Enables the incorporation of additional sub-population compartments to model disease spread within certain demographics (e.g. school children) in conjunction with demographic-specific interventions (e.g. school closures). The system needs to provide a *declarative* language interface that not only describes sub-population compartments but also how they interact with each other. Unlike agent-based models, the interaction is implemented via modified differential equations rather than agent state updates. We believe this approach offers the best scalability-complexity trade-off.

2. Provides an intervention specification language that describes interventions by (i) how they effect transmission rates, (ii) how they interact with other interactions (additive or multiplicative combined effects), (iii) cost functions, (iv) and other parameters such as daily limits or resource capacity, efficacy, etc.

3. Uses an agent-based approach to estimate intervention parameters. Determining the global effect of an intervention on transmission rates within a deterministic model can be challenging especially if an intervention is best described by its local effect on an agent's network. We can model such interventions on a smaller scale agent-based model to infer its global effect. These agent-models can be auto-generated from a higher-level graph-based language that describes relevant individual interactions and intervention effects.

4. Uses a provenance framework that maps interventions and parameters to their data source or derivation process. This ensures the reliability and reproducibility of the overall model.

1.3   Potential Impact

Our proposed epidemic simulator and policy optimization framework will allow policy-makers in the UAE and other regions to construct optimal intervention plans tailored for the unique characteristics of the society, geography and economy of the area. Through clear and easy-to-interpret specifications and implementations of disease spread and intervention effects, policy-makers can map policy choices to transparent and reproducible models backed by available data and robust to uncertainties.

2   Resources

2.1   Key team members

The following members are currently involved in this research project:

- Co-investigator: Prof. Dennis Shasha (Courant Institute of Mathematical Sciences, New York University)
- Consultant: Dr. Whitney Bagge, Disease Ecologist (EcoHealth Alliance, NY)
- Research Assistant: Anh Mai (NYUAD Alum)

2.2   Budget Justification

We anticipate the following costs:

- On premise GPU server: 10,000 USD
- Cloud computational costs (Amazon GPU P2.xlarge instances): 15,000 USD

Modeling and policy optimization require significant computational resources. Currently, on a 48-core machine, finding an optimal policy with 2-4 interventions takes 14 days. An in-house GPU server will help speed up computations especially when building and testing our framework. The cloud servers will help us run multiple experiments simultaneously.

References

[1] Declan Terence Bradley, Mariam Abdulmonem Mansouri, Frank Kee, and Leandro Martin Totaro Garcia. A systems approach to preventing and responding to covid-19. *EClinicalMedicine*, 21:100325, 2020.

[2] John Hopkins: Coronavirus Resource Center. Mortality analyses: How does mortality differ across countries? https://coronavirus.jhu.edu/data/mortality, May 2020.

[3] Van den Broeck W, Gioannini C, Gonçalves B, Quaggiotto M, Colizza V, and Vespignani A. The gleamviz computational tool, a publicly available software to explore realistic epidemic spreading scenarios at the global scale. *BMC Infect Dis*, 2011.

[4] Martine Denis, Valerie Vandeweed, Rein Verbeeke, Anne Laudisoit, Laure Wynants, and Diane Van Der Vliet. COVIPENDIUM: information available to support the development of medical countermeasures and interventions against COVID-19. *Transdisciplinary Insights*, May 2020.

[5] Judith V. Douglas, Simone Bianco, Stefan Edlund, Tekla Engelhardt, Matthias Filter, Taras Günther, Kun (Maggie) Hu, Emily J. Nixon, Nereyda L. Sevilla, Ahmad Swaid, and James H. Kaufman. Stem: An open source tool for disease modeling. *Health Security*, 17(4):291–306, 2019. PMID: 31433284.

[6] Neil M Ferguson, Daniel Laydon, Gemma Nedjati-Gilani, and et al. Report 9: Impact of non-pharmaceutical interventions (npis) to reduce covid19 mortality and healthcare demand. *Imperial College London*, 2020.