A Network-Driven Approach to Modeling the Spread of Ebola-type Epidemics

[1]Neil Slighton, [2]José M. Rico, [1]Emily Kallfelz, [3]Ji Qi, [3]Christopher G. Brinton
Departments of [1]Operations Research and Financial Engineering, [2]Computer Science, and [3]Electrical Engineering
Princeton University

***Abstract*—We propose new models for the spread of Ebola-type epidemics, considering networks both within and between countries. By modifying the traditional SIR model to capture the effects within each specific country, our Spatial SI(D/S) model overlays geographic information in the form of a graph topology to model the spread of diseases across boarders. In fitting the models to real-world data from the 2014-16 West Africa Ebola outbreak, we find that each is able to obtain low error in predicting infections over time, and that the use of spatial information can provide at least marginal improvements. We also show how our model parameters offer more insights into how these types of diseases are spread than does the SIR model, and propose an optimization problem for epidemic response strategies on a fixed budget that makes use of these parameters.**

I. Introduction

Epidemics have repeatedly shown an ability to explode in infected density if not properly contained, and have reaped massive consequences on populations in the process. The 2014-16 Ebola outbreak, for example, is estimated to have cost $4.55 Billion for Liberia, Sierra Leone and Guinea combined, slowing trade and plunging these countries into recessions in the process. It also closed schools in these countries for several months and led to more than 11,000 deaths [1].

The economic, political, and human costs of epidemics, coupled with limited funds of many countries to address them, motivates research in the development of models for the spread of diseases through populations. Such models can in turn be used to derive methods for preemptively combating the diseases. In particular, if the network of disease spread through communities can be properly estimated, then the people with the most influence (i.e., ability to infect others) could be the initial targets for cure or immunization [2], [3].

In this paper, we take a network-based approach to modeling Ebola-type epidemics, considering the spread of disease between geographic regions. We also formalize how one can respond to an epidemic using the parameters of our model.

A. Related Work

Research on epidemic modeling can be roughly divided into three categories: collection of data on outbreaks, modeling from data, and implementation of models. Work in collecting data generally relies on either (i) scraping information from health center and containment effort databases, or (ii) crowdsourcing methods, such as encouraging people to fill out surveys/questionnaires and/or to report cases [4]–[6].

As for modeling outbreaks, the famous three-compartment Susceptible-Infected-Recovered (SIR) model traces back to M'Kendrick's seminal 1925 paper [7]. In this model, susceptible individuals may become infected, and may eventually recover from the disease. Several extensions of this model have been proposed over the years, such as adding fourth susceptible (e.g., SIRS) or exposed (e.g., SEIR) compartments [8], using spatial data to account for the proximity of spread [3], and using population-dependent rates for the transitions between different species [9]. Many of these enhancements are motivated by properties of specific diseases like dengue fever [3] and influenza [4]. Our work can be viewed as an extension of the SIR model to account for factors specific to epidemics like Ebola, including (i) the possibility of becoming re-susceptible in addition to dying, and (ii) a spatial model for geographic spread on top of the baseline epidemic graph.

Research on the implementation of models has looked at the design of preemptive methods for minimizing epidemics. These include analyzing the effect of containment methods and immunization strategies, e.g., [10] [11]. Models have also been applied to analyze the efficacy of past policies, particularly for HIV prevention [12]. Motivated by this, we also propose an epidemic response strategy based on our model.

B. Organization and Contributions

In this work, we develop and evaluate new models for Ebola-type epidemics. We begin in Sec. II by formulating three models from a networking perspective. After deriving the traditional SIR model, we turn that into our SI(D/S) model which more closely reflects the dynamics of Ebola where individuals in a country can become infected multiple times. We then extend this into our Spatial (Graph-Based) SI(D/S) model, which overlays a digraph of infection rates between countries. Following this, in Sec. III we perform an evaluation of these models in terms of their ability to fit real-world data from an Ebola outbreak in three West African countries, in which we find that each obtains low error overall. We finally further motivate the parameters of our Spatial SI(D/S) model in Sec. IV by using them to develop an optimization problem for an epidemic response strategy.

II. Epidemic Model Formulation

In this section, we formalize our two network-based epidemic models. We first discuss preliminaries of the SIR

978-1-5386-0579-0/18/$31.00 ©2018 IEEE

model (Sec. II-A). We then build upon those in deriving our SI(D/S) (Sec. II-B) and Spatial SI(D/S) (Sec. II-C) models. The notation we use throughout is summarized in Table I.

A. SIR as a Graphical Model

In contrast to initial work [7] that introduced the SIR model using state transitions, we instead view the problem from a network perspective. In particular, we seek to represent the population of possible infected individuals as a network.

Epidemic graph. Formally, define the complete edge weighted digraph $\mathbb{G}_P = (V, E)$, where $V_i \in V$ represents the vertex i and the weight of edge $E_{i,j} \in E$ represents the strength of the connection from vertex i to vertex j. In our case, each i is a person, and the weight of $E_{i,j}$, say $w_{i,j}$, measures how likely i is to infect j if i is already infected, i.e., the probability of i infecting j is proportional to $w_{i,j}$.

Note that defining exact formulas for $E_{i,j}$ and $w_{i,j}$ is elusive, since in reality it is a complex mix of factors such as proximity, frequency of physical contact, cleanliness, and so forth. We will instead optimize the weight values $w_{i,j}$ in Section III to fit the dataset of the modeled epidemic.

Temporal model. We seek a temporal, discrete-time model for epidemics. In doing so, we let $V_i(t)$ denote the compartment (i.e., state) of vertex i at time t, either susceptible (S) or infected (I) for now. $S(t)$ and $I(t)$ are accordingly the sets of vertices that have not and have been infected at time t. The third state recovered (R) will be added into the model shortly.

Viewing $w_{i,j}$ as the probability that i infects j in a given interval of time Δ, we assume that each infected will infect others independently and that on average $w_{i,j} < 1/|I(t)|\ \forall i, j$. With this, we have the following equation for the probability that a person will get infected in a time interval:

$$\begin{aligned} &P(V_j(t) \in I(t) | V_j(t) \in S(t-\Delta)) \\ &= \sum_{i=1}^{|V|} w_{i,j} \mathbb{1}\{V_i(t-\Delta) \in I(t-\Delta)\} \end{aligned}$$

where $\mathbb{1}\{\cdot\}$ is the indicator function and $t - \Delta$ is the index for the previous time period.

Now, a sick person in $I(t)$ will eventually either die or get better. This adds a third state for recovered vertices $R(t)$ at time t. A recovered vertex cannot be infected; technically speaking this state corresponds to the immune or deceased. To model the transition from I to R, we let $h(t)$ denote a time-varying function of the likelihood that an infected vertex recovers after time t spent infected. Formally, with t_i denoting the time at which vertex i is infected, we have

$$P(V_j(t) \in R(t) | V_j(t) \in I(t-\Delta)) = h(t - \Delta - t_i)$$

Deriving the traditional SIR model. We now derive the traditional, population-based SIR model from ours to aid in simulations. First, we find the average edge weight, say β_Δ and assign each edge $w_{i,j}$ to that weight. With uniform edge weights, we can rewrite our probability of infection as

$$P(V_j(t) \in I(t) | \cdots) = \sum_{i=0}^{|V|} \beta_\Delta \mathbb{1}\{V_i(t-\Delta) \in I(t-\Delta)\}$$

TABLE I
TABLE OF NOTATION USED IN THE PAPER.

Term	Description
V	Set of all vertices.
V^c	Set of all cliques in the spatial model.
E	Set of all edges.
E^c	Set of all edges between cliques.
$V_i(t)$	State of the ith vertex at time t.
$E_{i,j}$	The edge between vertices i and j.
$w_{i,j}$	The weight of edge $E_{i,j}$.
Δ	Size of time step.
$S(t)$	Susceptible set at time t. $S_j(t)$ for clique j.
$I(t)$	Infected set at time t. $I_j(t)$ for clique j.
$R(t)$	Recovered set at time t.
$D(t)$	Dead set at time t. $D_j(t)$ for clique j.
$h(t)$	Probability of transitioning from I to R at time t.
t_i	Time at which vertex i gets infected
t_a	Time at which the combating action occurred.
β_Δ	Discrete time model average edge weight.
α_Δ	Discrete time model rate of recovery/death.
β	Continuous model rate of infection.
α	Continuous model rate of recovery/death.
$\beta^c_{i,j}$	Rate of infection from clique i to j.
α^c_i	Rate of recovery/death in clique i.
ϵ	RMSE error metric.
$\bar{c}_T$	Total resources.
$\bar{c}$	Given set of resources.
$\Delta\beta_{i,j}$	Change in $\beta_{i,j}$.
$\Delta\alpha_i$	Change in α_i.
$f_\alpha(x)$	Function that maps $\delta_{\alpha,i}$ values to $\hat{c}$ vectors.
$f_\beta(x)$	Function that maps $\delta_{\beta,i,j}$ values to $\hat{c}$ vectors.

Since $|I(t)|$ people are going to be sick at a given time t, we can simplify this further:

$$P(V_j(t) \in I(t) | V_j(t) \in S(t-\Delta)) = |I(t-\Delta)| \cdot \beta_\Delta$$

Now, note that the number of people who can be infected at any given time is the number of susceptible people. As a result, the expected number of people who are infected in one time step is

$$E\left[|S(t-\Delta)| - |S(t)|\right] = |S(t-\Delta)| \cdot |I(t-\Delta)| \cdot \beta_\Delta$$

To arrive at the traditional SIR model, we must further simplify the process of moving from state I to state R by assuming that $h(t)$ is uniform over all times, i.e., $h(t) = \alpha_\Delta$. Thus,

$$\alpha_\Delta = P(V_j(t) \in R(t) | V_j(t) \in I(t-\Delta))$$

We can see then quantify the expected number of people that recover in one timestep as

$$E\left[|R(t)| - |R(t-\Delta)|\right] = |I(t-\Delta)|\alpha_\Delta$$

Finally, we must make the functions continuous. As we take $\lim_{\Delta\to 0}$, the number of "chances" to infect approaches to infinity, but the strength of each chance will technically approach 0. Adjusting α and β to reflect this by specifying $\lim_{\Delta\to 0} \alpha_\Delta = \alpha$ and $\lim_{\Delta\to 0} \beta_\Delta = \beta$, we have that

$$\lim_{\Delta\to 0} E\left[|S(t-\Delta)| - |S(t)|\right] = \frac{dS}{dt} = -|S(t)||I(t)|\beta$$

$$\lim_{\Delta\to 0} E\left[|R(t)| - |R(t-\Delta)|\right] = \frac{dR}{dt} = |I(t)|\alpha$$

Fig. 1. High-level block diagram of the traditional SIR model.

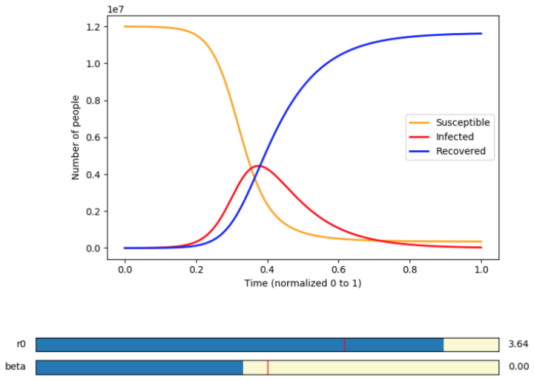

Fig. 2. Simulation of the SIR model. The sizes of $S(t)$, $I(t)$, and $R(t)$ are shown over time t with sliders for $r_0 = |S|/\alpha$ and β that show the value r_0 and β are set at to return generate these results. α and β are the rates of recovery and infection respectively.

These equations correspond exactly to those of the population-based SIR model, with the additional state $I(t) = -S(t) - R(t)$:

$$\frac{dI}{dt} = |S(t)||I(t)|\beta - |I(t)|\alpha$$

Model intuition. Overall, the SIR model splits the population into three groups: susceptible (S), infected (I), and recovered (R), with the flow between states summarized in Figure 1. A susceptible person can be infected by someone currently infected, and an infected person will eventually recover, either becoming immune or deceased. A person in the recovered category remains recovered and does not interact with the other categories. We give an example simulation of the SIR model for fixed values of α and β in Figure 2.

B. The SI(D/S) Ebola Model

We now modify the population-based SIR model to fit the dynamics of Ebola and Ebola-type diseases. We take two factors into consideration:

(1) Non-immunity. First, it is possible for a person to catch Ebola twice, i.e., surviving infection does not make a person immune [4]. This implies that an infected person in $I(t)$ can become susceptible again $S(t)$ or recover, but in this case, recovery implies death. We presume that the dead are buried or quarantined from the rest of the population and thus cannot infect or be infected. As a result, we replace $R(t)$ with a compartment $D(t)$, and add a transition between $I(t)$ and $S(t)$.

(2) Mortality rate. Second, studies have revealed that the probability of someone dying once contracting Ebola is 71% [13]. Therefore, for those exiting the I category, we assign a transition rate of 0.71 to the D category and a rate of 0.29 to re-entering the S category.

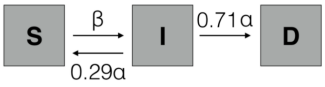

Fig. 3. High-level block diagram of the SI(D/S) Ebola model.

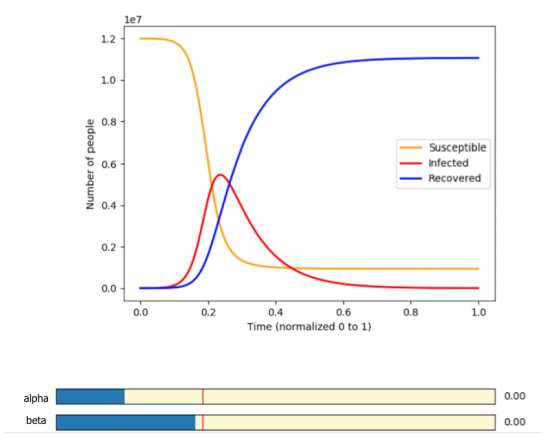

Fig. 4. Simulation of the SI(D/S) Ebola Model. The sizes of $S(t)$, $I(t)$, and $D(t)$ are shown over time t, with sliders for α and β the rates of recovery/death and infection respectively.

The modified population growth rate equations are thus:

$$\frac{dS}{dt} = -|S(t)||I(t)|\beta + 0.29|I(t)|\alpha$$

$$\frac{dI}{dt} = |S(t)||I(t)|\beta - |I(t)|\alpha$$

$$\frac{dD}{dt} = 0.71|I(t)|\alpha$$

We denote this model as the SI(D/S) Ebola model, which we summarize in Figure 3. Figure 4 gives an example simulation of this model for fixed values of α and β.

C. The Spatial SI(D/S) Ebola Model

Like the SIR model, the SI(D/S) Ebola model assumes that each person in the population is connected with equal strength and all recover identically fast. In reality, the connection strengths $w_{i,j}$ and the speed of recovery/death will vary person to person. Estimating these parameters individually is clearly intractable, however, without detailed information on each person's contact network within the population of interest.

On the other hand, we can expect strong variations in parameters to exist between different regions. Since datasets on Ebola split cases by country, such variations can be incorporated by allowing for country-specific α and β parameters.

Model overview. This model defines parallel and interdependent SI(D/S) models for each country. In doing so, it splits the larger population into distinct but dependent populations, and models the spread of epidemic both within a country and between countries. For example, Liberia will have an SI(D/S) Ebola model for itself, but people in this country will have additional avenues of contracting Ebola from other countries.

Model intuition. Each country has its own unique health care policy, level of development, infrastructure, and so forth that

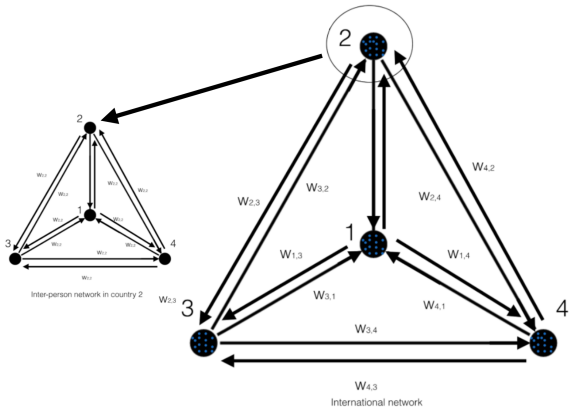

Fig. 5. Visualization of the two-layered network hierarchy of the Spatial SI(D/S) Ebola Model. This has four countries and four people in the second.

intuitively will lead to similar α and β values for those living within a country, but different values across countries. There are exceptions to this, such as international travelers who have more frequent contact with those in other countries, but as we are interested in population dynamics we may assume these internal variations will be absorbed by everyone in the country equally. Returning to our graph-based intuitions for epidemic modeling, we can say that people are organized into similarly-connected cliques based on their country, and different cliques (countries) are in turn connected in a higher-level graph. We can allow all edges between a vertex in clique x and a vertex in clique y to take the weight of the edge connecting clique x and y, while within a clique, we can assume that all vertices are connected with uniform weights. The networks within a country can function like the SI(D/S) Ebola model, with α and β unique to the given country. This two-layered network structure is visualized in Figure 5 for the case of four countries.

Model formulation. To formalize this model, define the complete weighted digraph $\mathbb{G}^c = (V^c, E^c)$ where vertex $V_i^c \in V^c$ represents a clique/country of vertices. Each edge $E_{i,j}^c \in E^c$ represents infection probability of individuals from country i to country j. The weight of edge $E_{i,j}^c$ corresponds the the magnitude of the β value between all people from country i to j. We define the $|V^c|$ by $|V^c|$ matrix β^c to contain these weights, i.e., $\beta_{i,j}^c$ is the value of β from i to j. By this specification, $\beta_{i,i}^c$ corresponds to the β value internal to i, i.e., the infection rate between people within the country.

Different from β, the recovery rate parameter α is not defined between cliques. We define the length $|V^c|$ vector α^c to contain these values for each clique, with α_i^c being that for country i. Now, denoting the set of susceptible, infected and dead vertices in clique i as S_i, I_i and D_i, respectively, our system of differential equations becomes

$$\frac{dS_j(t)}{dt} = 0.29\alpha_j^c I_i(t) + \sum_{i=1}^{|V^c|} \left(-\beta_{i,j}^c I_i(t) S_j(t)\right)$$

$$\frac{dI_j(t)}{dt} = -\alpha_j^c I_i(t) + \sum_{i=1}^{|V^c|} \left(\beta_{i,j}^c I_i(t) S_j(t)\right)$$

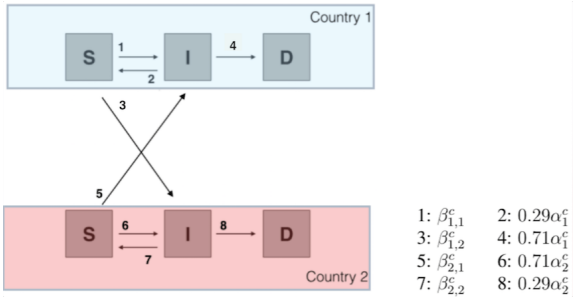

Fig. 6. High-level block diagram of the Spatial SI(D/S) Ebola model, with the interdependencies of model states shown for two countries.

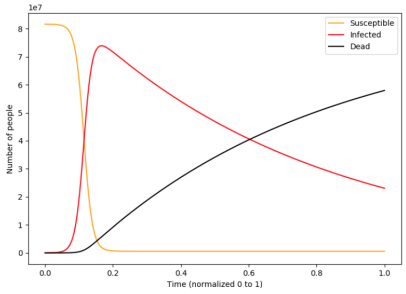

Fig. 7. Simulation of the Spatial SI(D/S) Ebola model over time for 10 countries. The S, I and D model states are shown as totals across all countries.

$$\frac{dD_j(t)}{dt} = 0.71\alpha_j^c I_i(t)$$

We call this model the Spatial (Graph-Based) SI(D/S) Ebola model. The interdependencies between model states are depicted in Figure 6 for the case of two countries. Figure 7 gives an example simulation of this model, plotting the aggregate S, I and D across countries.

III. Model Evaluation

We now compare the proposed SI(D/S) models with the SIR model in terms of their abilities to fit a real-world Ebola outbreak in different countries. We describe our dataset (Sec. III-A) and implementation (Sec. III-B) before presenting and discussing the results (Sec. III-C).

A. Dataset

We obtained our dataset from the open source Humanitarian Data Exchange [13]. It consisted of 17,585 entries, with each entry giving the cumulative number of Ebola cases (either confirmed, probable, or suspected) for a country on a particular date recorded. There are 259 distinct dates over the course of two years, from 2014 to 2016, in roughly uniform 3-day increments. A total of 12 different countries across Europe, Africa, and small portions of North America appear in the dataset. However, only three of them reported enough cases to be effectively modeled: Guinea, Liberia, and Sierra Leone.

The format of the data was already quite close to the output function $I(t)$ of our models, so we did not have to perform substantial transformations prior to evaluation. In particular, by making $I(t)$ cumulative and down-sampling it to the number

TABLE II
BEST FIT PARAMETERS AND ϵ ERROR METRICS OBTAINED FOR EACH COUNTRY AND MODEL. BOTH THE SI(D/S) AND SPATIAL SI(D/S) MODELS MAKE MARGINAL IMPROVEMENTS OVER THE SIR MODEL.

Country	Model	α	β	Error (ϵ)
Guinea	SIR	0.0057347	4.4870×10^{-10}	0.044131
	SI(D/S)	0.0057313	4.4845×10^{-10}	0.044143
	Spatial	0.0057384	*See Table III*	0.043784
Liberia	SIR	0.0262572	5.7840×10^{-9}	0.087516
	SI(D/S)	0.0263041	5.7923×10^{-9}	0.087320
	Spatial	0.0261867	*See Table III*	0.087340
Sierra Leone	SIR	0.024466	3.4113×10^{-9}	0.070857
	SI(D/S)	0.0244759	3.4128×10^{-9}	0.070731
	Spatial	0.0244659	*See Table III*	0.070615

TABLE III
TUNED β^c MATRIX FOR THE SPATIAL SI(D/S) MODEL.

	1: Guinea	2: Liberia	3: Sierra Leone
1	4.484×10^{-10}	4.825×10^{-15}	6.251×10^{-15}
2	6.933×10^{-12}	5.766×10^{-9}	3.037×10^{-12}
3	2.149×10^{-13}	2.745×10^{-14}	3.411×10^{-9}

of data points in the original data set, we were able to develop an error metric (described next) that compares our model prediction of $I(t)$ with the data of these three countries.[1]

B. Parameter Selection and Error Metric

Each model has parameters that we can adjust to fit the data. The SIR and SI(D/S) models both have two for each country – α and β – while the Graph-Based SI(D/S) model has $N(N+1)$ parameters – α_i^c and $\beta_{i,j}^c$ – where N is the number of countries.

Error metric. To select these parameters, we look to minimize the following root mean squared error metric:

$$\epsilon = \frac{\sqrt{\sum_{t=0}^{t_f} \left(\hat{I}(t) - I(t)\right)^2 / t_f}}{\sum_{t=0}^{t_f} I(t)/t_f}$$

where $\hat{I}(t)$ and $I(t)$ are the cumulative numbers of infected cases from the model and the data, respectively, and t_f is the number of time periods. Though the infected category itself does not fully encapsulate all model states, it is tied to both the S and D categories such that over time, we can expect that incorrect predictions in these other categories will be reflected.

For each model, we discretize time into $t_f = 10,000$ steps. Given that the dataset spans around 2 years, this gives a time step Δ in the range of 1-2 hours. We found that the models generally performed better with shorter values of Δ, and that $10,000$ was a reasonable tradeoff between model error and computational complexity.

Tuning SIR and SI(D/S). To tune the SIR and SI(D/S) models, we perform a grid-search procedure similar to [14]. In particular, for each country, we successively refine a point (α^k, β^k) over iterations k. In each iteration, we pixelate 100 values over a decreasing range and choose the pair with the lowest ϵ. We repeat until the decrease in error between two iterations is negligible.

[1]Our code for scraping and parsing the data can be found at https://github.com/jmrico01/ele-381 under file get_data.py.

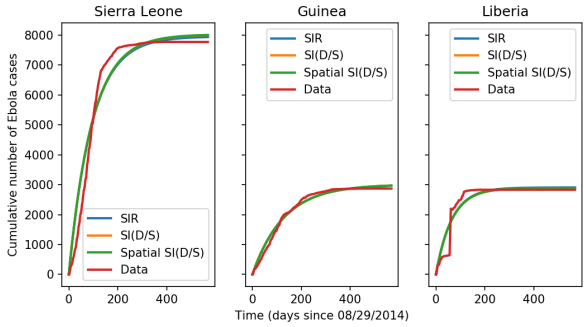

Fig. 8. Plot of the three models over time in predicting the cumulative number of infected, compared with the empirical data. Each model tends to overshoot initially and undershoot later, which could be used in early detection.

Tuning Spatial SI(D/S). With three countries, the Spatial SI(D/S) model has 12 parameters, making a grid search procedure computationally intractable. As a result, to tune this model, we start with the optimized parameters from the SI(D/S) model for each country, with cross terms set to zero:

$$\alpha_i^c = \alpha_i^{\text{SI(D/S)}} \qquad \beta_{i,i}^c = \beta_i^{\text{SI(D/S)}} \qquad \beta_{i,j}^c = 0 \;\; i \neq j$$

Naturally, this results in the same error values for each country as the SI(D/S) model, which we can then use as a baseline to improve upon. With these initial values in place, we begin spreading each $\beta_{i,i}^c$ randomly along its row, ensuring at each step that each row always sums to the original $\beta_i^{\text{SI(D/S)}}$, i.e., using $\beta_i^{\text{SI(D/S)}}$ as a benchmark for the total magnitude of infection spread. More formally, in each step k we set each $(\beta_{i,j}^c)^k$ to a uniformly-selected random fraction of the $(\beta_{i,i}^c)^k$ in the same row, and subtract these from $(\beta_{i,i}^c)^k$ such that $\sum_{j=0}^{N} (\beta_{i,j}^c)^k = \beta_i^{\text{SI(D/S)}}$. We repeated this until we obtained an error value close to the original SI(D/S) model, but with non-zero cross terms $\beta_{i,j}^c$. Next, we repeatedly selected a parameter, any α_i^c or $\beta_{i,j}^c$, and adjusted it by a uniformly-selected random fraction of its value. We repeated this process until the difference in the sum of the errors across countries between two iterations was negligible.[2]

C. Results and Discussion

Tables II and III give the best fit parameters identified for each model by country, and the corresponding model errors ϵ. Figure 8 compares the cumulative values of $I(t)$ over time in each of the cases. Overall, we see that each model obtains a rather close fit to the actual data, and that both the SI(D/S) and Spatial SI(D/S) models make improvements over the SIR model, albeit marginally.

More specifically, the average errors over all countries for SIR, SI(D/S), and Spatial SI(D/S) are 6.750%, 6.740%, and 6.725%, respectively. The incremental improvements are

[2]The code to simulate and optimize each model can be found at https://github.com/jmrico01/ele-381, under files models.py, optimize_single.py, and optimize_spatial.py.

small, but the SI(D/S) and Spatial SI(D/S) models have the added improvements of interpretability. In particular, they both can be expected to give α and β parameters that are more well-tuned to the real world, since they employ an empirically measured death rate. The Spatial SI(D/S) provides the richest information: the $\beta^c_{i,j}$ values provide insights into how quickly the disease spreads within and between countries while α^c_i values give information on the death or recovery rates by country. For example, the fact that $\beta^c_{2,1}$ is three orders of magnitude larger than $\beta^c_{1,2}$ in Table III indicates that Ebola spread from Liberia to Guinea may be more prevalent than from Guinea to Liberia, which is also consistent with Liberia having substantially higher internal α^c_i values in Table II. Such information can be of utility to Ebola prevention measures.

Note also in Figure 8 that each model tends to overshoot initially and undershoot later on. A closer fit is always more ideal, but note that this characteristic can actually be useful for early detection of the spread of the disease.

IV. Epidemic Response Strategy

To further motivate the development of the Spatial SI(D/S) Ebola model, we now briefly propose a framework for using the results of this model in resource allocation strategies for epidemic response. We define the success metric for such a strategy to be the number of people that die after the strategy is implemented; denoting this value as ϕ, we have

$$\phi = \lim_{t\to\infty} \sum_{j=1}^{|V_c|} \left(|D_j(t)| - |D_j(t_a)|\right)$$

where t_a is the time at which the combating action is implemented and $D_j(t)$ is the cumulative number of deceased predicted by the spatial model in country j at time t, given conditions after t_a. The objective is to minimize ϕ.

Now, actions taken against the epidemic are subject to resource constraints, which could be monetary, policy, or other in nature. Assume that these costs types are mutually independent and that there are M types. The resources used are then represented with a vector $\bar{x}$ of length M, where each entry corresponds to the amount of one resource. We also let $\bar{x}_T$ denote the vector of total resources available, i.e., $\bar{x} \leq \bar{x}_T$.

Next, we assume that using a certain vector of resources $\bar{x}$ corresponds to a change in a given α^c_i or $\beta^c_{i,j}$ value by some $\Delta\alpha_i$ or $\Delta\beta_{i,j}$ respectively. Formally, we define $\mathbb{R}^{|V^c|} \to \mathbb{R}^M$ and $\mathbb{R}^{|V^c|^2} \to \mathbb{R}^M$ functions $f_\alpha(x)$ and $f_\beta(x)$ as the mapping of $\Delta\alpha_i$ and $\Delta\beta_{i,j}$ to $\bar{x}$ cost vectors. These functions can be defined and estimated for a particular set of countries and costs. This gives the following optimization problem at the current time t_c:

$$\begin{aligned}
\text{minimize} \quad & \phi = \lim_{n\to\infty} \sum_{j=1}^{|V_c|} |D_j(n)| - |D_j(t_c)| \\
\text{subject to} \quad & \sum_{i=1}^{|V_c|}\sum_{j=1}^{|V_c|} f_\beta(\Delta\beta_{i,j}) + \sum_{i=1}^{|V_c|} f_\alpha(\Delta\alpha_i) \leq \bar{x}_T \\
\text{variables} \quad & \Delta\beta_{i,j} \in \mathbb{R}, \Delta\alpha_i \in \mathbb{R}, \ \forall i,j \in V^c
\end{aligned}$$

The result of this optimization is a set of $\Delta\beta^\star_{i,j}$ and $\Delta^\star\alpha_i$, i.e., the changes that correspond to the optimal response strategy.

V. Conclusion and Future Work

In this paper, we developed three models for the spread of Ebola-type epidemics from a network perspective: the traditional SIR, the SI(D/R), and the Spatial SI(D/R). In evaluating their abilities to fit real-world data of Ebola in three countries, we found average errors of 6.750%, 6.740%, and 6.725% respectively, which are small overall and show a trend of improvement. Importantly, the SI(D/R) and Spatial SI(D/R) models have learnt parameters that give insights into the spread of the disease. We motivated these parameters further in using those based on the Spatial SI(D/S) model in the proposal of an epidemic response resource allocation strategy.

In future work, we aim to (i) derive a more general parameter optimization procedure for the Spatial SI(D/S) model, (ii) fit the epidemic response strategy to specific cases, and (iii) test the generalizability of our models to other epidemics.

References

[1] Z. Mullan, "The cost of ebola," http://www.thelancet.com/pdfs/journals/langlo/PIIS2214-109X(15)00092-3.pdf, 2015.

[2] L. Meyers, "Contact network epidemiology: Bond percolation applied to infectious disease prediction and control," *Bulletin of the American Mathematical Society*, vol. 44, no. 1, pp. 63–86, 2007.

[3] G. Chowell, P. Diaz-Duenas, J. Miller, A. Alcazar-Velazco, J. Hyman, P. Fenimore, and C. Castillo-Chavez, "Estimation of the reproduction number of dengue fever from spatial epidemic data," *Mathematical biosciences*, vol. 208, no. 2, pp. 571–589, 2007.

[4] C. Guerrisi, C. Turbelin, T. Blanchon, T. Hanslik, I. Bonmarin, D. Levy-Bruhl, D. Perrotta, D. Paolotti, R. Smallenburg, C. Koppeschaar *et al.*, "Participatory syndromic surveillance of influenza in europe," *The Journal of infectious diseases*, vol. 214, no. suppl_4, pp. S386–S392, 2016.

[5] L. F. Lopes, J. Zamite, B. Tavares, F. Couto, F. Silva, and M. J. Silva, "Automated social network epidemic data collector," in *INForum informatics symposium. Lisboa*, 2009.

[6] W. H. Organization, "Hiv/aids global epidemic data and statistics," www.who.int/hiv/data/global_data/en/, 2008.

[7] A. M'Kendrick, "Applications of mathematics to medical problems," *Proceedings of the Edinburgh Mathematical Society*, vol. 44, pp. 98–130, 1925.

[8] H. W. Hethcote, "The mathematics of infectious diseases," *SIAM review*, vol. 42, no. 4, pp. 599–653, 2000.

[9] J. O. Lloyd-Smith, D. George, K. M. Pepin, V. E. Pitzer, J. R. Pulliam, A. P. Dobson, P. J. Hudson, and B. T. Grenfell, "Epidemic dynamics at the human-animal interface," *science*, vol. 326, no. 5958, pp. 1362–1367, 2009.

[10] A. Lima, M. De Domenico, V. Pejovic, and M. Musolesi, "Exploiting cellular data for disease containment and information campaigns strategies in country-wide epidemics," *arXiv preprint arXiv:1306.4534*, 2013.

[11] D. Zhao, L. Wang, S. Li, W. Zhen, W. Lin, and G. Bo, "Immunization of epidemics in multiplex networks," *Plos One*, vol. 9, no. 11, pp. 1–5, 2014.

[12] T. Brown and W. Peerapatanapokin, "The asian epidemic model: a process model for exploring hiv policy and programme alternatives in asia," *Sexually Transmitted Infections*.

[13] "Number of ebola cases and deaths in affected countries," https://data.humdata.org/dataset/ebola-cases-2014.

[14] C. G. Brinton and M. Chiang, "Mooc performance prediction via clickstream data and social learning networks," in *Computer Communications (INFOCOM), 2015 IEEE Conference on*. IEEE, 2015, pp. 2299–2307.