An adaptive control strategy for the West Africa Ebola outbreak

NC STATE UNIVERSITY

Nicholas J. Meyer, Eric B. Laber, Krishna Pacifici, Brian J. Reich, John Drake
North Carolina State University, University of Georgia

Controlling Information Transfer on a Network

Networks pervade many areas of science [5] and provide the foundation for modeling the evolution of complex systems. We are often interested in how these systems evolve over time and seek a deep understanding of information exchange throughout the network. Examples include social networks, computer networks, ecological food webs, and infectious diseases. We are interested in understanding the dynamics of these complex systems and using this knowledge to control the spread of diseases.

The Ebola outbreak in West Africa began in December of 2013. As of January 16, 2015, there were over 21 thousand cases of Ebola with nearly 8.5 thousand deaths [3]. This more than deadly disease in West Africa poses a serious threat to regional and global public health. Consequently, the ability to contain the spread of Ebola is imperative. Managing the spread of a disease is a complex decision making problem with both spatial and temporal components. Developing optimal strategies for controlling complex systems, such as Ebola, requires solving a sequential decision making problem that is complicated by extremely high dimensions.

Observed Data

The Ebola outbreak in West Africa consists of two waves. Infections prior to May 2014 are considered part of the first wave. Two regions in Guinea, Guéckédou and Conakry, maintained the infection between waves. For Analysis, the infection is treated as beginning in these two regions on April 26, 2014.

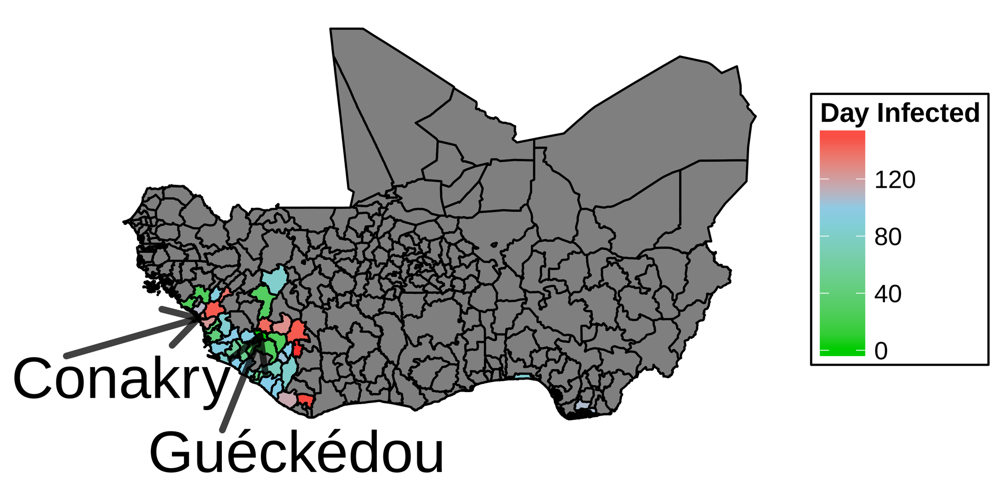

Figure: Day infected starting from April 26, 2014

Goals

1. Define a class of rich and interpretable treatment strategies.
2. Using scientific knowledge of the system, develop a low-dimensional parametric model of the system dynamics and estimate optimal strategy using simulation based optimization.
3. Construct a semi-parametric estimator of the optimal strategy using a Q-learning based approach minimizing the Bellman residual.
4. Stabilize the semi-parametric approach with a low-dimensional parametric anchor.

Notation

- Let $\mathcal{T} = \{1, 2, \ldots\}$ be the set of decision points. Let $\mathcal{L} = \{1, \ldots, L\}$ be the set of locations in the network.
- At each $t \in \mathcal{T}$, collect state information $\boldsymbol{S}^t = \{S_\ell^t, \ldots, S_L^t\}$ where $S_\ell^t \in \mathbb{R}^p$.
- At each $t \in \mathcal{T}$, choose actions $\boldsymbol{A}^t \in \{0,1\}^L$.
- A decision strategy π is a mapping from supp $\boldsymbol{S} \to \mathcal{B}_L$ where $\mathcal{B}_L$ is the set of all $\{0,1\}^L$ valued random variables.
- Let $Y_\ell^t \in \mathbb{R}$ be the outcome at time t for location ℓ.
- The optimal strategy $\pi^{opt} \in \Pi$ satisfies,

$$\mathbb{E}\left[\sum_{t\geq 1} \gamma^{t-1}\mathbf{1}^\intercal \boldsymbol{Y}^t(\pi^{opt})\right] \geq \mathbb{E}\left[\sum_{t\geq 1} \gamma^{t-1}\mathbf{1}^\intercal \boldsymbol{Y}^t(\pi)\right]$$

  for all $\pi \in \Pi$ where $Y(\pi)$ is the potential outcome under π.
- Goal: estimate π^{opt}

Low-Dimensional System Dynamics Model (M1)

- Assume the system is Markov and let $f(\boldsymbol{y}^t|\boldsymbol{s}^t, \boldsymbol{a}^t; \psi)$ and $g(\boldsymbol{s}^t|\boldsymbol{s}^{t-1}, \boldsymbol{a}^{t-1}; \theta)$ be postulated parametric distributions for $\boldsymbol{Y}^t$ and $\boldsymbol{S}^t$ respectively.
- Under the assumed model and strategy π, define

$$V^T(\pi; \psi, \theta) = \int \left(\textstyle\sum_{t=1}^T \gamma^{t-1}\mathbf{1}^\intercal \boldsymbol{y}^t\right) \textstyle\prod_{t=1}^T \{f(\boldsymbol{y}^t|\boldsymbol{s}^t, \boldsymbol{a}^t; \psi) \; g(\boldsymbol{s}^t|\boldsymbol{s}^{t-1}, \boldsymbol{a}^{t-1}) P[\pi(\boldsymbol{s}^t) = \boldsymbol{a}^t]\} d\lambda(\overline{\boldsymbol{y}}^T, \overline{\boldsymbol{s}}^T, \overline{\boldsymbol{a}}^T)$$

  to be the expected discounted reward until time T.

Low-Dimensional System Dynamics Model (M1) cont.

- If the model is correct, $\pi_{M1}^{opt} = \arg\max_{\pi\in\Pi} \lim_{T\to\infty} V^T(\pi; \psi^*; \theta^*)$ where ψ^* and θ^* are the true parameters.
- Approximate optimal strategy using estimates $\hat{\psi}$ and $\hat{\theta}$ and large T,

$$\hat{\pi}_{M1}^{opt} = \arg\max_{\pi\in\Pi} V_{M1}^T(\pi; \hat{\psi}; \hat{\theta}).$$

- Estimate $V_{M1}^T(\pi; \hat{\psi}; \hat{\theta})$ using Monte Carlo integration.
- (M1) provides a low-variance estimator of π^{opt}, but quality depends on the postulated system dynamics model.

Semi-Parametric Q-Learning Based Model (M2)

- Rather than parameterizing the entire system dynamics model, as in (M1), it is sufficient to model the Q-function, $Q^\pi(\boldsymbol{s}, \boldsymbol{a})$, defined as

$$Q^\pi(\boldsymbol{s}, \boldsymbol{a}) = \mathbb{E}^\pi\left[\sum_{k\geq 1} \gamma^{k-1}\mathbf{1}^\intercal \boldsymbol{Y}^{t+k-1}|\boldsymbol{S}^t = \boldsymbol{s}, \boldsymbol{A}^t = \boldsymbol{a}\right].$$

- Use the Bellman equation [6]

$$\mathbb{E}\left[\mathbf{1}^\intercal \boldsymbol{Y}^t + \gamma Q^\pi(\boldsymbol{S}^{t+1}, \pi\{\boldsymbol{S}^{t+1}\}) - Q^\pi(\boldsymbol{S}^t, \boldsymbol{A}^t)\right] = 0$$

  to construct an estimating equation for $Q^\pi(\boldsymbol{s}, \boldsymbol{a})$.
- Assume the Q-function is additive across locations, $Q^\pi(\boldsymbol{s}, \boldsymbol{a}) = \sum_{\ell\in\mathcal{L}} Q_\ell(\boldsymbol{s}, \boldsymbol{a}; \beta_\ell^\pi)$.
- Define an estimating equation $\Lambda_T(\beta, \pi)$ as

$$\Lambda_T(\beta, \pi) = \mathbb{E}^\pi \textstyle\sum_{t=1}^{T-1} \Big[\textstyle\sum_{\ell=1}^L (Y_\ell^t + \gamma Q_\ell\{\boldsymbol{S}^{t+1}, \pi[\boldsymbol{S}^{t+1}]; \beta_\ell\} - Q_\ell\{\boldsymbol{S}^t, \boldsymbol{A}^t; \beta_\ell\})\Delta_\beta Q_t \Big]$$

  where $\Delta_\beta Q_t = \{\Delta_{\beta_1} Q_1^\intercal(\boldsymbol{S}^t, \boldsymbol{A}^t; \beta_1), \ldots, \Delta\beta_L Q_L^\intercal(\boldsymbol{S}^t, \boldsymbol{A}^t; \beta_L)\}^\intercal$ and $\mathbb{E}^\pi$ denotes expectation with respect to the distribution of $\pi(\cdot)$ only.
- Define $\hat{\beta}^{\pi,\tau} = \arg\min_\beta \|\Lambda_T(\beta, \pi)\|_2^2 + \tau J(\beta)$ where $J(\beta)$ is a spatially explicit penalty smoothing over locations.
- The estimated optimal strategy as

$$\hat{\pi}_{M2}^{opt} = \arg\max_{\pi\in\Pi} Q(\boldsymbol{s}, \pi\{\boldsymbol{s}\}; \hat{\beta}^{\pi,\tau}).$$

- (M2) provides an estimator of π^{opt} that is consistent under weaker conditions than (M1), however, it requires more data to provide a high quality estimator.

Anchoring (M2) with (M1)

- (M1) provides a stable low-variance treatment strategy, however the assumed system dynamics model from (M1) is almost surely misspecified.
- We propose (M2), a more flexible strategy consistent under weaker assumptions.
- Combining both strategies, we use (M1) when data are scarce to make reliable decisions, then at subsequent decision points test the performance of (M1) and (M2) to determine when (M2) has sufficient information for training.
- Under the hypothesis that the system dynamics model from (M1) is correctly specified, generate the null distribution of the Q-function and compare the observed value against the 95th percentile.

Priority Score

- Define $\mathcal{R} = \{R(\boldsymbol{s}, \boldsymbol{a}; \eta); \eta \in E\}$ where each element is a mapping supp $\boldsymbol{S}^t \times \{0,1\}^L \to \mathbb{R}^L$. If location j is treated, then $R_j(\boldsymbol{s}, \boldsymbol{a}; \eta) = -\infty$.
- For non-negative integers m, M define

$$\mathcal{U}_\ell^t(\boldsymbol{s}, \boldsymbol{a}; \eta, m, M) = \mathbb{1}\left\{R_\ell(\boldsymbol{s}, \boldsymbol{a}; \eta + \xi_M^t) \geq R_{(m)}(\boldsymbol{s}, \boldsymbol{a}; \eta + \xi_M^t)\right\}$$

  where $\{\xi_j^t\}_{j\geq 1}$ are independent Gaussian random variables.
- At time t assume budget C^t, for some $k \leq C^t$, $\pi^{(j)}(\boldsymbol{s}; \eta)$ is a random binary vector uniformly distributed on

$$\{\, \boldsymbol{a} \in \{0,1\}^L : \boldsymbol{a}^\intercal\mathbf{1} = \lfloor jC^t/k \rfloor, (1 - \boldsymbol{a})^\intercal \pi^{(j-1)}(\boldsymbol{s}; \eta) = 0, \; \mathrm{diag}(\boldsymbol{a}) \cdot \mathcal{U}^t(\boldsymbol{s}, \pi^{(j-1)}[\boldsymbol{s}; \eta], C^t - \lfloor (j-1)C^t/k \rfloor, j\mathbf{1}) = \pi^{(j-1)}(\boldsymbol{s}; \eta)\}$$

  for $j = 1, \ldots, k$ and $\pi^{(0)} = \mathbf{0}$.
- One possible priority score form is $\mathcal{R}_\ell(\boldsymbol{s}, \boldsymbol{a}; \eta) = \phi_\ell(\boldsymbol{s}, \boldsymbol{a})^\intercal \eta$ where ϕ_ℓ is a vector of features for location ℓ containing relevant information about the system.

Features

- The features provide flexibility to the class of strategies. Construction of features should use existing scientific knowledge and theory to maximize the potential.
- Our features are functions of:
  - Model based predictions of spread for the disease
    - Estimate under multiple models to increase robustness
  - Measures of centrality [2] and data depth [4]
  - Quantitative summaries of density to account for spill-over effect

Simulation Experiment: Ebola

- Goal: compare the performance of (M1) using a priority score based strategy against heuristic and myopic strategies.
- Setup:
  - Simulate 257 days of disease spread.
  - Constant generative model.
  - Total 16 regions treated, 8 preventative and 8 active.
  - Minimize expected proportion of infected regions after 257 days.
  - 300 Monte Carlo replications.

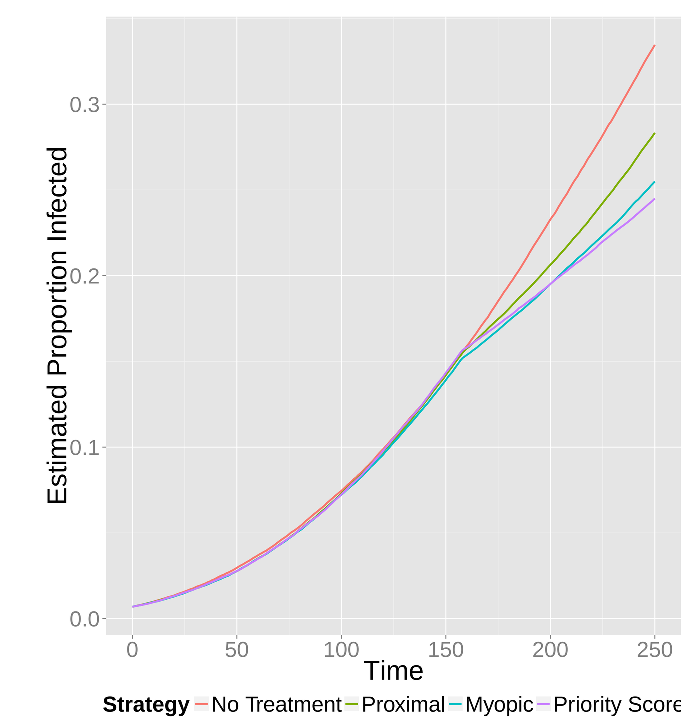

Figure: Simulation results for West Africa Ebola outbreak

Simulation Experiment: Toy Structures

- Running simulations on toy structures can provide important additional information. In this setting, the network structure is known and specific structural features can be incorporated.
- The chosen structure is known as scale-free [1]. The defining characteristic is the proportion of locations with k edges decays exponentially with k. Many networks, such as the web, exhibit this property and contain locations that are much more highly connected.
- Setup:
  - 100 locations.
  - Simulate 15 steps of disease spread.
  - Constant generative model.
  - Total 6 locations treated, 3 preventative and 3 active.
  - Minimize expected proportion of infected locations after 15 steps.
  - 300 Monte Carlo replications.

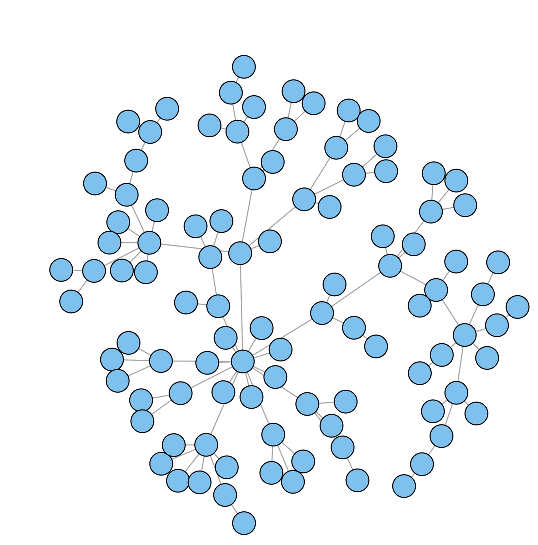

Figure: Network map of the toy structure

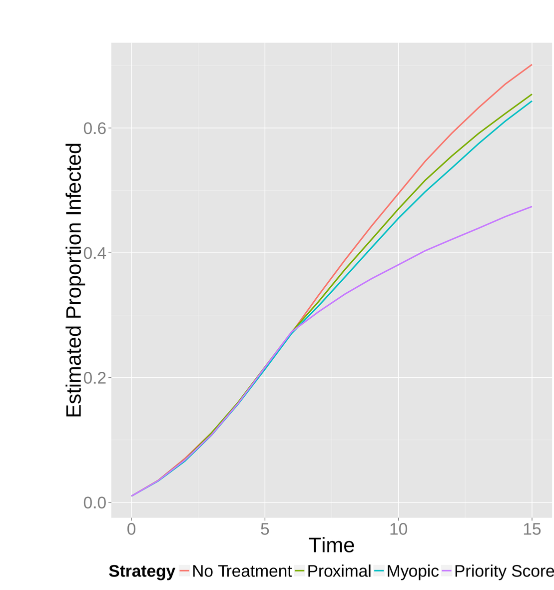

Figure: Simulation results for toy structure

Works Cited

[1] Albert-László Barabási and Réka Albert. Emergence of scaling in random networks. *Science*, 286(5439):509–512, 1999.

[2] Reinhard Diestel. *Graph Theory, 4th Edition*, volume 173 of *Graduate texts in mathematics*. Springer, 2012.

[3] Center for Disease Control. 2014 ebola outbreak in west africa - case counts, January 2015.

[4] John W Tukey. Mathematics and the picturing of data. *Proceedings of the international congress of mathematicians*, 1975.

[5] Steven H. Strogatz. Exploring complex networks. *Nature*, 410(6825):268–276, March 2001.

[6] Richard S Sutton and Andrew G Barto. *Reinforcement learning: An introduction*, volume 1. Cambridge Univ Press, 1998.

njmeyer@ncsu.edu