Managing Invasive Species in a River Network

Kim Hall, Majid Alkaee Taleghan, H.J. Albers, Mark Crowley, Tom Dietterich
Oregon State University
CompSust at OSU

- **eBird** – learning species distribution patterns from citizen science data
- **BirdCast** – predicting bird migration patterns
- **BudID** – automated categorization of bugs from image data
- wildlife corridor planning
- data cleaning for forest sensor networks
- forest fire control
- controlling invasive species
The Problem

- Managing plant ecosystem along a river network
- Competing native and invasive plant species
- Native and Invasive species spread dynamics
  - Local, spatial, stochastic
- Optimize for best outcome subject to budget constraints
Example River Network

- Slot States
  - Native plants, invasive plants, empty

- Actions in each Reach
  - Eradicate invasive plants
  - Eradicate and restore (replant) natives
  - Do nothing
States and Actions

• Slot States
  • Native plants, invasive plants, empty

• Actions in each Reach
  • Eradicate invasive plants
  • Eradicate and restore (replant) natives
  • Do nothing
Where We Fit In

- **Ecology**: focus on biological processes, postulate complete eradication (may be economically infeasible)

- **Economics**: focus on optimal control policy
  - Spatial spread often ignored/simplified greatly
  - Steady state analysis of spatial spread

- **Econ + CS**: Collaboration between Ecosystem Economics and Computer Science
  - Spread modelled as a conditional, spatial, stochastic process
  - Optimized as an MDP
    - Structure of the problem presents interesting computational challenges
Goal

Optimal policy describing placement of management actions over space and time

Economic Optimization

• Objective: reduce presence of invasive plants while minimizing costs
• Subject to annual management budget constraints and ecological processes
Optimization Problem

\[
\min_{a_{it}} \sum_{t} \sum_{i} \gamma^t c_{it}(n_{it}, a_{it})
\]

s.t. \[
\sum_{i} c_{it}(n_{it}, a_{it}) \leq b_t \forall t
\]

s. t. ecological model holds

Where:

\( \gamma \) Discount factor
\( c_{it} \) Cost function
\( n_{it} \) Invasive population size
\( a_{it} \) Management action
\( b_t \) Budget constraint
<table>
<thead>
<tr>
<th>Component</th>
<th>Expression</th>
</tr>
</thead>
<tbody>
<tr>
<td>States</td>
<td>$s \in S$</td>
</tr>
<tr>
<td>Actions</td>
<td>$a \in A$</td>
</tr>
<tr>
<td>Transition Dynamics</td>
<td>$\mathcal{T}(s'</td>
</tr>
<tr>
<td>Rewards</td>
<td>$r(s^t, a^t) : S \times A \rightarrow \mathbb{R}$</td>
</tr>
<tr>
<td>Discount Factor</td>
<td>$\gamma$</td>
</tr>
<tr>
<td>Policy</td>
<td>$\pi(a</td>
</tr>
</tbody>
</table>
Size of States and Actions

- Number of States and Actions are exponential in the network size
  - \( N \) is the number of states each slot can take on
  - \( M \) is the number of actions available in each reach

\[
|S| = \left( \frac{(N + H - 1)!}{H!(N-1)!} \right)^R \\
|A| = M^R
\]

| Number of Reaches (R) | Number of Slots (H) per Reach | Number of Actions: \( |A| \) | Number of States: \( |S| \) | Transition Model Size: \( |S| \times |S| \times |A| \) |
|------------------------|-------------------------------|------------------|-------------|-------------------|
| 3                      | 2                             | 27               | 216         | 1.0 x 10^7       |
| 5                      | 2                             | 243              | 7,776       | 1.4 x 10^{10}    |
| 3                      | 3                             | 27               | 1,000       | 2.7 x 10^7       |
| 5                      | 3                             | 243              | 10,000      | 2.4 x 10^{12}    |
Dynamics

• **Mortality**
  - Plants in each slot die with independently with probability $d$
  - Eradication of invasive plant, could fail stochastically

• **Propagule Generation**: Each surviving plant produces $g$ propagules deterministically

• **Propagule Dispersal**: upstream or downstream with
  - $P(\text{arrive at reach } j \mid \text{ started at reach } i)$

• **Site competition/colonization** at slot $h$:
  - If $h$ is occupied, no effect
  - Else propagules compete to colonize slot, bias $\beta$ in favour of invasive plants

Network dispersal model from Muneepeerakul et al. 2007
Dynamics Models

Dispersal Model
Probability of propagule leaving reach $i$ and arriving in reach $j$ is proportional to the rate of propagule survival upstream/downstream and the distance travelled

$$K_{i,j} = Cu_{i,j}^{NU} d_{i,j}^{ND}$$

Competition Model
Probability that species $k$ wins in slot $s$ is equal to $k$’s proportion of the total number of propagules arriving in slot $s$ modified by a weighted factor $\beta$.

$$p_{\text{invasive}} = \frac{\beta g_{\text{invasive}}}{\beta g_{\text{invasive}} + g_{\text{native}}}$$

$$p_{\text{native}} = 1 - p_{\text{invasive}}$$
Dispersal Model

Probability of propagule leaving reach $i$ and arriving in reach $j$

$$K_{ij} = Cu^{NU_{ij}} d^{ND_{ij}}$$

Where:

- $C$ – is a normalization constant
- $u$ – is the upstream propagule survival rate
- $d$ – is the downstream propagule survival rate
- $NU_{ij}$ – is the number of upstream reaches between reach $i$ and $j$
- $ND_{ij}$ – is the number of downstream reaches between reach $i$ and $j$
Competition Model

\[ p_{\text{invasive}} = \frac{\beta g_{\text{invasive}}}{\beta g_{\text{invasive}} + g_{\text{native}}} \]

\[ p_{\text{native}} = 1 - p_{\text{invasive}} \]

- \( p_{\text{species}} \) - Probability that species wins
- \( g_{\text{species}} \) - Number of propagules of species
- \( \beta \) - “competitive advantage” of an invasive seed versus a native seed
  - 1.0, 1.5, 2.0...
Estimating the Transition Model

• It’s easy to write a simulator for drawing samples of the dispersal and competition processes.

• But computationally intractable to compute the exact transition probabilities $\mathcal{T}(S'|S, A)$.
  • Estimate transition probabilities by drawing a large number of samples from the simulator.
Error Bounds on Transition Model

\[ Pr \left( \max_{s'} \left| \mathcal{T}(s' \mid s, a) - \mathcal{T}(s' \mid s, a) \right| < \epsilon \right) > 1 - \delta \]

- Confidence interval with width of \( \varepsilon \)
- 1-\( \delta \) probability of being within interval
- This is a very loose bound : \( |S| \) is large
- Future Work:
  - Tighter bounds that account for missing states from simulations
  - Approximate algorithms with PAC guarantees on bounds
Optimization

• Once estimate of $\hat{T}(S' | S, A)$ is obtained perform Value Iteration on action-value function $Q^*(s, a)$
Interpreting the Policy

• Direct Examination of Optimal Policy

• Run optimal policy forward – collect stats from many simulated trajectories
  • Time to reach steady state
  • Frequency with which completely invaded
  • Frequency with which uninested states are reached

• Future Work
Comparing to Rule of Thumb Policies

• Managers and Ecology/Economics Literature suggest:
  • **Triage**: treat most invaded reaches first
  • **Chades, et al.**: upstream first; extreme nodes first (one reach treated per period)
  • Treat **leading edge** of spread
Comparing policy pathways

Chades

Leading edge

Optimal

Triage up to down

- empty
- native
- invader

☐ eradicate

○ restore
Results: Costs

Total Costs

- Large pop, up to down
- Shades
- Leading Edge
- Optimal
Results

• Optimized spatial policy can outperform aspatial rules of thumb policies

• Spatial characteristics of the system under invasion are relevant to optimal management
  • strength of downstream vs upstream dispersal
  • presence of long distance dispersal changes policy
Future Ecology/Economics work

• Better ecological models of competition needed
• More investigation of space-time interactions
• Stochastic arrivals from outside network
• Richer objective functions
  • Separate competitiveness and colonization probabilities
• Model human dispersal of invasive plants via boating
Future Computational Work

- **Memory**: value iteration with partial transition model loaded into memory

- **Data re-use**: minimizing calls to expensive simulators when learning model

- **Compact Representations**: more compact spatial representations of states and policies
  - Larger problem sizes
  - Improved policy interpretation
  - Relational learning to distill general rules from policy

- **Bounded Approximations**: PAC-style algorithms with bounds on results
  - Estimate values of states directly through simulation
References


This work funded as part of the US National Science Foundation project in Computational Sustainability.
Thank You

Questions?