General rules for managing and surveying networks of pests, diseases, and endangered species

Iadine Chadès1,2, Tara G. Martina, Samuel Nicolb, Mark A. Burgman2, Hugh P. Possinghamb, and Yvonne M. Buckleya,b

*CSIRO Ecosystem Sciences, Ecocosciences Precinct, Dutton Park, QLD 4102, Australia; “University of Queensland, School of Biological Sciences, Queensland, 4072, Australia; and “School of Botany, University of Melbourne, Parkville, VIC 3052, Australia

Edited by Simon A. Levin, Princeton University, Princeton, NJ, and approved March 31, 2011 (received for review November 11, 2010)

The efficient management of diseases, pests, or endangered species is an important global issue faced by agencies constrained by limited resources. The management challenge is even greater when organisms are difficult to detect. We show how to prioritize management and survey effort across time and space for networks of susceptible–infected–susceptible populations. We present simple and robust rules of thumb for protecting desirable or eradicating undesirable, subpopulations connected in typical network patterns (motifs). We further demonstrate that these rules can be generalized to larger networks when motifs are combined in more complex formations. Results show that the best location to manage or survey a pest or a disease on a network is also the best location to protect or survey an endangered species. The optimal starting point in a network is the fastest motif to manage, where line, star, and island, and cluster motifs range from fast to slow. Managing the most connected node at the right time and maintaining the same management direction provide advantages over previously recommended outside–in strategies. When a species or disease is not detected and our belief in persistence decreases, our rules recommend shifting resources toward management or surveillance of the most connected nodes. Our analytic approximation provides guidance on how long we should manage or survey networks for hard-to-detect organisms. Our rules take into account management success, dispersal, economic cost, and imperfect detection and offer managers a practical basis for managing networks relevant to many significant environmental, biodiversity, and human health issues.


The authors declare no conflict of interest.

This article is a PNAS Direct Submission. Freely available online through the PNAS open access option.

1To whom correspondence should be addressed. E-mail: iadine.chades@csiro.au.

This article contains supporting information online at www.pnas.org/cgi/suppl/doi/10.1073/pnas.1016846108/DCSupplemental.

www.pnas.org/cgi/doi/10.1073/pnas.1016846108

1 of 6
efficient for network motifs. This is a key finding because eradication or effective control is generally possible only where a pest or a disease occupies a small number of nodes (38). Likewise, managing rare or threatened metapopulations is by definition a small network problem.

The first step in solving a resource allocation problem is defining the objective we seek to achieve. When managing pests and diseases, our objective is to minimize the number of infected nodes by taking actions to increase the probability of extinction of an infected node. When managing threatened species, our objective is to maximize the number of infected (occupied) nodes by taking actions that decrease the probability of extinction of an infected node. These equivalent problems have the same solution. The best local population to manage for an infestation is also the best place to manage for a threatened metapopulation.

At each time step, we assume the infection state of each node $i$ ($(s_i = \{\text{susceptible, infected}\})$) is observed. Once managed, a node can remain infected with probability $(1 - p_m)$ and infect susceptible nodes to which it is connected with a probability $p_d$. When management occurs, the node may become susceptible and can be reinfected with probability $1 - (1 - p_d)^\delta$, where $\delta$ is the number of infected nodes directly connected to the susceptible node. We assume that each infected node has a probability of becoming susceptible without management ($p_{e}$), to account for the natural, stochastic decline of the infection. We assign an economic cost, $c$, every year that results from a node being infected. Managing a node incurs a cost $c_m$. We consider that only one node can be managed at a time. This method is the simplest way to prioritize the order in which nodes are managed so that our objective is achieved in the most efficient way. We also demonstrate that our management rules are robust when several nodes can be managed at a time (Results and Methods). We assume uniform dispersal, management success, benefit, and cost across all nodes (Methods and Table S1). However, our method can accommodate changes and nonuniform assumptions in any of the following parameters: the probability of a node being infected if one neighbor is infected ($p_{d}$), the probability of a node becoming susceptible if infected and not managed ($p_{e}$), the probability of a node becoming susceptible if it has been managed ($p_{m}$), the cost of a node being infected ($c$), and the cost of managing one node ($c_m$).

**Results**

When two infected nodes have the same probability of infecting each other, the optimal strategy is to manage either one. When several nodes are linked by a directed probability of dispersal, it is optimal to manage the upstream node before managing the downstream nodes (Fig. 1A). In the case of a line with bidirectional infection risk, the nodes at the extremity of the infection process have a lower probability of being reinfected than nodes with two neighbors. The optimal strategy is to start from either extremity of the network and then manage the nearest infected nodes sequentially (Fig. 1B). In the case of an island network, all nodes have an equal probability of being reinfected once managed; therefore, any node can be the starting point (Fig. 1C). Star networks are composed of a central node connected to satellite nodes. Infected satellite nodes are managed until the number of susceptible satellites is equal to or larger than the remaining number of infected satellites; it is then optimal to manage the central node before managing the remaining infected satellites (Fig. 1D). When managing star networks with lines of nodes of different lengths connected to a central node, the optimal strategy has three phases: (i) manage the longest lines from an outside–in direction until just over half of the nodes connected to the central node remain infected, (ii) manage the shortest infected line and the central node following an outside–in direction, and (iii) manage the remaining infected lines starting with the shortest lines following an inside–out direction.

Fig. 2 shows simple rules of thumb for more complex networks made up of multiple motifs. When an island is connected to a line, the risk of reinfection is minimized by starting management at the extreme node of the line and continuing to the connecting node, before managing the island. Once the connecting node is
susceptible, it is optimal to manage the infected nodes closest to the connecting node. A more complex configuration is an island connected to two lines (Fig. 2A). Maintaining the same management direction reduces the risk of reinfection and is therefore more efficient than managing the “outside” nodes first.

We define a cluster as an island with interconnected nodes. When managing connected clusters (Fig. 2B), it is optimal to start from the smallest cluster or, for the same number of nodes, the least connected cluster. For a connected island and star (Fig. 2C), the star is faster to manage when the number of satellites is equal to or smaller than the number of nodes of the island. In contrast, when the number of star satellites is greater than the number of island nodes, managing the island first reduces the chance of reinfection. In both cases, the rules of thumb derived from individual motifs can be applied. For a star connecting four substars (Fig. 3), the optimal rule of thumb for managing single star motifs holds.

We compared our rules of thumb to an outside–in control strategy by simulating the management of 1, 2, and 5 nodes at a time on large star networks of 101 nodes (Figs. S1–S3). Our optimal strategy outperforms the outside–in strategy, reducing the time to eradication and the management required by 31.2% (90.7 treatments, managing 1 node), 10.7% (20.4 treatments, managing 2 nodes), and 3.3% (5.6 treatments, managing 5 nodes). Our rules are also robust to nonuniform dispersal probabilities ($p_d$) and management success ($p_m$) and consistently outperform the traditional outside–in strategy (Table S2 and SI Methods).

We found that the relative benefits of our rules of thumb are greater when fewer patches are managed in each time step. Our rules of thumb outperform the outside–in strategy because reinfections are less likely to occur using our priority management order. The longer the time it takes to eradicate the infection, the more opportunity there is for reinfections to take place, and the greater the benefit of our rules of thumb over the outside–in strategy. By managing multiple patches in a time step, the time to eradication is reduced, making the performance of the two strategies more similar.

For all of the complex networks we examined, management starts with the fastest motif to manage, following the priority order line, star, island, and cluster. Management of the most connected node at the right time and maintaining the same management direction minimize the reinfection risk, providing potentially important practical advantages over outside–in strategies, particularly in the case of limited management resources. To determine the management starting point more easily we asked if a network centrality measure could substitute for our rules of thumb. Centrality measures are able to rank nodes in a network according to the topological features that they account for (see refs. 39–41 for more details). We calculated centrality measures for nine small networks (Table S3). We found that the closeness centrality measure correctly predicted the most appropriate node to start from for seven cases of nine but predicted incorrect starting nodes for the two networks characterized by connected clusters. Betweenness centrality, eigenvector centrality, and subgraph centrality (42) performed inadequately. Although closeness centrality often predicted the correct starting node, it did not predict an optimal priority management order, highlighting the continuing need for our optimization framework.

Most pests, diseases, and threatened species cannot be detected perfectly (9). Apparently susceptible nodes may, in fact, be infected. We extend our framework to take into account our inability to detect subpopulations perfectly, using the partially observable Markov decision process (POMDP) (Methods). In this case, we consider three possible management actions: Manage a node, survey for the infection, or do nothing. We include in our optimization model the probability of imperfectly detecting infected nodes, $d_n$, when doing nothing or managing. Probability of detection increases when we survey a node, $c_n$, is set to half the cost of management.

When all of the nodes in a star are initially infected (Fig. 4), the order of management follows the rules of thumb we described under perfect detection. After this first round of management, if the infection is not detected in any node, a stochastic dynamic programming solution recommends doing nothing; however, when accounting for imperfect detection, our results recommend managing all of the nodes a second time for this star network. The number of times a node must be managed depends on the POMDP parameters used. As our belief in eradication increases, our solution shifts to surveying the most connected node. Finally, when our belief that the infection has been eradicated from the network has reached a threshold (0.82 in the case of Fig. 4), and in the absence of observing an infection, “do nothing” becomes the optimal action. When we observe an infection to be present in a node(s), the belief of eradication falls to zero and “manage” is recommended.

![Complex networks](image)

**Fig. 3.** Optimal management of a complex star network of a disease or an invasive or threatened species.

![Resource allocation recommended over time for the management of a hypothetical invasive species when infection may be undetected on a star network.](image)

**Fig. 4.** Resource allocation recommended over time for the management of a hypothetical invasive species when infection may be undetected on a star network (Inset). m_i, manage node i; s, survey node i; and dn, do nothing. The red line represents the probability (belief) of eradication of the infection from the network. The other lines represent the belief in eradication of the individual nodes. The probability of eradication starts at zero for each node because initially, the network is assumed to be infected.
When we do not have any prior information about the infection of the network, every infection state in every node is equally probable (Fig. S4). In this case, our solution recommends first managing all of the nodes at least once and more than once if the node is highly connected (e.g., node 2, Fig. S4). The action of surveillance then becomes optimal. In the absence of detection of the infection under surveillance, our belief that the infection is eradicated reaches a high value (e.g., probability of eradication in the network > 0.82). It is then recommended to do nothing.

Under imperfect detection, as our belief in eradication increases, our control strategy evolves through three phases: Manage, survey, and do nothing. As demonstrated in Chadès et al. (9), the problem of when to stop managing and surveying a single population of cryptic species is dependent on the relative cost of survey and management, the cost of the infection, the benefits of eradication, and the efficiency of management and surveillance determine the switching points (9). To avoid re-running the POMDP algorithms for every network studied, we build on this work and derive an analytic approximation by transforming our network POMDP problem into a single POMDP. To make the transformation, we assume the dispersal probability between nodes is a less important factor than the number of subpopulations (Methods). Although this assumption may not always hold, this approximation is the best indicator of how long we can afford to manage derived thus far and forms the basis for future theoretical work. We find that, in the absence of sighting, as the number of subpopulations increases, the time we should spend managing increases substantially irrespective of the economic costs (Fig. S5).

Discussion

We identify simple rules for the management of small networks under perfect and imperfect detection of susceptible–infected–susceptible network invasions, infections, and threatened metapopulations. Our findings provide guidance for managers in determining (i) where to prioritize management over time and space to achieve the most efficient control strategy and (ii) when the species or the disease is difficult to detect, how long they should invest in active management or in surveilling and when it is reasonable to stop investing limited resources. Our simulation results demonstrate the improvement in performance when following our rules of thumb over traditional outside–inside management approaches on large motif networks. Our rules are also robust to nonuniform distribution of dispersal and management efficiency. When deciding where we should prioritize management, a key finding is to start with the fastest motifs to manage and keep the same management direction through the network. Centrality measures are not good proxies for determining the priority order of management of nodes. Finally we derived an analytic approximation providing guidance on the time we should spend managing and surveying across a network in the absence of detection.

Details of the number of nodes, the costs and benefits of managing nodes, and the dispersal probability drive the optimal management strategies. In ecology and economics, evaluating the priority order of which nodes to manage is useful only for networks that exhibit a clear structure. Highly connected networks will not benefit from our approach as the amount of resources invested to manage the threat will be more important than the order of nodes managed. In other words networks need to be sparsely connected to benefit from our approach. If a structured network is found but the number of nodes is too large to be solved following our approach, we recommend using an abstraction process or decomposition methods to extract small subnetworks (45–47). An alternative approach is to use a model that does not include network structure at all, such as a diffusion process (48).

Managers of diseases might find our rules counterintuitive as we do not recommend managing the most connected nodes first. There are several plausible explanations for this. First, our rules provide a priority order of management from the initial condition of all nodes being infected. This order might not correspond to the starting management point of some diseases, e.g., when an infected node is in a wider network of nodes infected. Although our optimal strategy defines an optimal action for any infection stage, it is impossible to derive a unique priority order that applies to every infection stage. Second, although common, the assumption that a node can represent the state of a subpopulation might not reflect epidemic dynamics; however, we can consider a node as a representation of an individual and our rules continue to apply to a network of infected individuals. Third, we assume the structure of the network is known. Therefore, the infection cannot spread to unknown connected nodes and managing the most connected node does not bring any additional advantage. Finally, the susceptible–infected–susceptible network model might not capture the complexity of some diseases where vaccinated or recovered and immune states are possible; other models such as susceptible–infected–recovered (SIR) might be more appropriate.

We present general solutions for the management of connected subpopulations of pests, diseases, and threatened species arranged in a SIS network. The efficient management of diseases, pests, or threatened species is a globally important issue, especially in agencies constrained by limited resources. The challenge of managing new incursions is even greater when the species or disease is cryptic and impossible to detect perfectly. To date, most of the literature in network theory has focused on making predictions about large complex systems such as food webs, epidemics, and gene regulation (49, 50). We provide general rules of thumb for setting management priorities in small networks, where our chances of eradication or recovery of a species are maximized (51). Our results provide a practical basis to manage SIS networks relevant to many significant biosecurity, human health, and environmental issues.

Methods

When there is perfect detection, finding the optimal management strategy for an infected network can be solved using Markov decision process (MDP) and stochastic dynamic programming (SDP) (33). In our case MDP is inconvenient for two reasons: The computational complexity grows exponentially with the number of nodes and SDP methods mostly provide a single optimal solution. Determining alternative optimal solutions is essential when deriving rules of thumb. We overcome these disadvantages by modeling our SIS network management problem as a factored Markov decision process (FMMDP) (26), taking advantage of the structure of our problem and fast optimization methods from artificial intelligence.

A factored MDP is a MDP that explicitly represents the independent relationships between variables of the system (conditional independence).
A factored MDP is defined by four elements: a factored set of states (S), a finite set of actions (A), a transition probability function (P), and a reward function (R). Let S be the state of the network at any given time t and s represent the nodes infected or the susceptible state: \( S = s_1, s_2, \ldots, s_n \), where the state variable \( s_i \in S_i \) represents the state of node \( i \). The state has \( 2^n \) possible states. Let \( a \in A \) be the decision manager can make at any given time \( t \) and a represent the nodes managed by the decision makers: \( a = a_1, a_2, \ldots, a_n \) where the components \( a_i \in A_i = \{ \text{manage, do nothing} \} \), the set of decisions available in node \( i \). Because we are working in the context of management, we assume only one node can be managed at each time step. Our simulation results have shown that when we increase the number of nodes we can manage at a time, our rules continuously outperform the outside-in strategy (Results and SI Methods). As currently defined, our model and approach will still hold when more resources are available, allowing the management of several nodes at a time, the size of the action set grows with \( n \), limiting the use of our methods to very small networks.

The dynamics of the system are captured by a matrix \( P \) that contains the transition probabilities of moving from any state to any other state of the system under different actions following the SI process (Introduction and Fig. S6). In classical MDP, the solution methods for all transition probabilities are represented and stored using \( I \) matrices of \( S \times S \) dimensions. Each element of these matrices represents a transition probability: \( P(s_{t+1} | s_t, a) = \sum_{s_{t+1}} P(s_{t+1} | s_t, a) \). A factored MDP takes into account the independence between state variables explicitly, using a dynamic Bayesian network implemented as an augmented MDP with a factored set of observations (Algorithm). The MDP model is an augmented FMDP with a factored set of observations and their corresponding Markov decision process (POMDP). We automatically generate the corresponding FMDP model and solve the optimization process using the fast symbolic perseus method (28). In our case, solving an FMDP is finding a strategy \( \pi : X \times T \rightarrow A \), mapping an allocation of recourse \( a \in A \) given a current belief state \( b \in B \) and a time step \( t \in T \). An optimal strategy minimizes the expected sum of costs or rewards \( R \) over a finite time horizon, \( T \). This expected summation is also referred to as the value function \( V(S) \). A value function essentially ranks strategies by assigning a real value to each choice.

We have also derived an analytic solution for a population of invasive species following previous work by Chadas et al. (9) (SI Methods). We also derived an analytic approximation for a population of invasive species following previous work by Chad, Possingham, and McAllister (2006). Our approach is innovative and constitutes a key point of our work. Factored MDP and POMDP allow us to (i) compactly represent the optimization problems using trees and dynamic Bayesian networks, (ii) implement efficient computation using algebraic decision diagrams, and (iii) identify management rules by using decision trees to represent all of the optimal solutions in a structured way.

We derived an analytic solution for a population of invasive species following previous work by Chadas et al. (9) (SI Methods). We also derived an analytic approximation for a population of invasive species following previous work by Chad, Possingham, and McAllister (2006). Our approach is innovative and constitutes a key point of our work. Factored MDP and POMDP allow us to (i) compactly represent the optimization problems using trees and dynamic Bayesian networks, (ii) implement efficient computation using algebraic decision diagrams, and (iii) identify management rules by using decision trees to represent all of the optimal solutions in a structured way.

In summary, being able to model our optimization problems as a factored MDP (under perfect detection) or factored POMDP (under imperfect detection) is innovative and constitutes a key point of our work. Factored MDP and POMDP allow us to (i) compactly represent the optimization problems using trees and dynamic Bayesian networks, (ii) implement efficient computation using algebraic decision diagrams, and (iii) identify management rules by using decision trees to represent all of the optimal solutions in a structured way.

We derived an analytic solution for a population of invasive species following previous work by Chad, Possingham, and McAllister (2006). Our approach is innovative and constitutes a key point of our work. Factored MDP and POMDP allow us to (i) compactly represent the optimization problems using trees and dynamic Bayesian networks, (ii) implement efficient computation using algebraic decision diagrams, and (iii) identify management rules by using decision trees to represent all of the optimal solutions in a structured way.

In reality many pests, diseases, and threatened species cannot be detected perfectly. A factored partially observable Markov decision process (POMDP) model is an augmented FMDP with a factored set of observations (28) and an observation function (O). Similar to a factored MDP, a factored POMDP is a POMDP that explicitly represents the independent relationships between variables of the system. In our case the POMDP tackles the problem of deciding the most appropriate management action given the difficulty of detecting an infected node. We define the finite set of local observations for each node at time \( t \), \( z_i \) \{Absent, Present\} and their corresponding observation function \( o_1 \) that maps to each state–action pair a probability distribution over \( z_i \). In other words, the probability of detection given that the node is infected and that the previous decision is to do nothing is defined by \( o_1(Present | Infected, Do nothing) \).