Optimal Wildlife Smuggling Interdiction

Rachel McLean

April 23, 2019

Thesis submitted for the degree of
Master of Philosophy
in
Applied Mathematics
at The University of Adelaide
Faculty of Engineering, Computer and Mathematical Sciences
School of Mathematical Sciences

THE UNIVERSITY of ADELAIDE

Contents

List of Tables

List of Figures

Signed Statement

I certify that this work contains no material which has been accepted for the award of any other degree or diploma in my name, in any university or other tertiary institution and, to the best of my knowledge and belief, contains no material previously published or written by another person, except where due reference has been made in the text. In addition, I certify that no part of this work will, in the future, be used in a submission in my name, for any other degree or diploma in any university or other tertiary institution without the prior approval of the University of Adelaide and where applicable, any partner institution responsible for the joint-award of this degree.

I give permission for the digital version of my thesis to be made available on the web, via the Universitys digital research repository, the Library Search and also through web search engines, unless permission has been granted by the University to restrict access for a period of time.

I acknowledge the support I have received for my research through the provision of an Australian Government Research Training Program Scholarship.

Signed: .... Date: ...23/04/2019...

Acknowledgements

Firstly, I would like to thank my supervisors, Professor Joshua Ross, Professor Matt Roughan, and Associate Professor Phill Cassey, for the time and effort they put into assisting me over the past two years. Secondly, I would like to thank Sarah Heinrich, for her expertise and providing data that greatly assisted the research. Finally, I wish to thank my friends and family, for their constant encouragement and support throughout my study.

Abstract

The illegal trade of wildlife is one of the greatest threats to the survival of many species. Pangolins are claimed to be the most heavily trafficked wild mammals in the world. There is a huge demand for their scales and meat in traditional medicines and food. Wildlife products are trafficked across the world to meet such demands. Smugglers are able to choose the route they take to move these commodities across international borders. Biosecurity resources are placed at ports of entry to intercept smugglers. In this thesis, we describe a model of the relationship between smugglers and biosecurity forces.

We begin by assuming that the smuggler picks a path that either minimises travel time, maximises the probability of going undetected, or some combination of these two objectives. Biosecurity agencies allocate resources to ports of entry to maximise their probability of catching the smuggler. However, we assume that the smuggler fully observes any resource changes and thus may alter their path to avoid detection. This results in a dynamic game between the two players. We aim to find the optimal biosecurity resource allocation.

However assuming all traffic is sent along a single path from each source is unrealistic. Thus we develop an extension to the model, where the illegal traffic is distributed across the possible paths and only some proportion of the traffic is sent along the optimal path.

We describe implementations of two algorithms that find the optimal resource allocation. These solutions lead to insights about the best approaches to interdict illegal trade across ports of entry.

We apply our model to a real world network to investigate it in a practical setting. We look at the trafficking of pangolins from Asian and African sources into Vietnam. Resources are allocated at five entry points into Vietnam. We determine the optimal allocations for different smuggler objectives. We consider robust allocations that are good for a majority of the assumed objectives. Spreading resources evenly between three specific ports of entry, out of the five, appears best when we just consider single paths from sources. When we consider multiple paths from the same source, the most robust allocation across smuggler objectives is to spread resources between two specific ports of entry. The ports of entry that are assigned resources in the robust allocations have many more paths into them compared to the other entry points. Thus this robust allocation makes logical sense, and perhaps may apply to situations other than the trafficking of pangolins from Asian and African sources into Vietnam.

Chapter 1

Introduction

The illegal trade of wildlife is a major threat to biodiversity. It is a highly specialised form of organised crime [8, 43, 48] that generates enormous profits, estimated to run into billions of dollars [27, 43]. Thousands of species are exploited, from elephants for ivory, to sea turtles for meat, skin, and shells. Many species have become endangered because of this trade.

The illegal wildlife trade also indirectly harms non-targeted wildlife. For example, invasive species may be introduced and harm native species, and animals are unintentionally killed in fishing gear (bycatch).

In addition to threatening the survival of populations, smuggling techniques frequently harm individual animals [42]. Live animals are put through horrific conditions and many die in transit [42]. They are also found sick, raising concerns that the illegal wildlife trade is a pathway for the introduction and spread of diseases [8, 18, 42].

Wildlife crimes are run through complex criminal networks that enforcement agencies often know little about [56]. This lack of information helps the criminals avoid enforcement attempts to interfere with their activity. Examples of players in these networks include poachers, distributers, sellers, and buyers [2, 56]. The illegal wildlife trade is often enabled by corrupt officials [8, 48, 49, 56, 57]. It is also not uncommon for wildlife trafficking to be linked to other crimes such as money laundering and passport fraud.

The illegal wildlife trade is a global problem. Commodities originate in one part of the world and are sold in international markets. Almost every country is involved in some way and smugglers of around 80 nationalities have been linked to incidents [48]. Countries act as sources, transits or destinations for these illegal wildlife commodities. The role that they play depends on the wildlife type. Particular wildlife groups have strong connections to certain regions in the world. For instance, mammals to Asia and Africa.

The illegal trade of wildlife is regulated through the Convention on International

Trade in Endangered Species of Wild Fauna and Flora (CITES) [12]. CITES is an agreement between governments that aims to protect the survival of species through regulation of international trade. Governments set legislation based on this agreement. CITES places species into three categories (Appendices I, II, and III) based on conservation status and the risk posed from trade. Despite protection from CITES, species are still illegally traded.

Pangolins are scaly mammals found in regions across Asia and Africa. They have been labelled as the most trafficked wild mammals in the world [10]. All eight species are listed as Vulnerable through to Critically Endangered on the IUCN Red List of Threatened Species (www.iucnredlist.org). Between 2010 and 2013, one million pangolins are estimated to have been trafficked [21]. There is a huge demand for their meat and scales in food and traditional medicines [10, 35]. Their skins are also turned into leather products [10]. Given the declining populations of Asian pangolins, African pangolins are increasingly being trafficked to Asian markets in order to satisfy the demand [10]. Pangolin numbers in the wild have reached such critically low levels that in 2016, CITES banned all international trade of wild pangolins [24]. Despite this, trade continues, pushing this animal closer to extinction. This thesis will use the illegal pangolin trade as a case study to motivate and explore the problem of wildlife trafficking.

The majority of enforcement work to combat international wildlife trafficking occurs at ports of entry, as opposed to markets [48]. Customs officers operate as the front line of enforcement at sea, land, and air borders. It can be challenging for officers with limited training to accurately identify protected species [56]. This is made particularly difficult when products have been heavily processed or deliberately mislabelled. Some studies report that as low as 5% - 10% of contraband is seized [8, 47, 49].

The resources assigned to combat international wildlife trafficking are always limited. There are fewer than 330 U.S. Fish and Wildlife Service agents appointed to inspect the millions of shipments arriving at U.S. ports and airports [27]. This is an enormous task. Thus these agent's efforts need to be allocated intelligently amongst ports of entry.

Certain ports of entry are more desirable to traffickers. Kurland and Pires [27] found that a high proportion of illegal wildlife goods attempting to enter the United States were seized at a small number of ports. Petrossian *et al.* [36] analysed illegal, unreported, and unregulated fishing seizures. Their results indicated that these fishing vessels were more likely to use ports with higher fishing vessel traffic and larger harbours. As well, they were more likely to stop at ports with greater corruption levels and less effective inspection systems. These factors are likely to reduce the risk of detection, and traffickers are well aware of those risks.

Traffickers are also selective of transit locations along routes. Most observed illegal routes via air follow common passenger routes that use major transport hubs [8]. Seizure analysis indicates that ports with less screening processes for contraband but high numbers of connecting flights are preferred by traffickers [8], once again to minimise the chance of detection.

As mentioned, a large amount of enforcement activity occurs at international borders. Limited enforcement resources need to be allocated strategically between ports of entry into a country. Traffickers prefer certain ports. But how might changes in security resource allocations alter the smuggler's port choice? Ideally, adding more resources increases detection. An informed smuggler may alter the path they use based on this. However, smugglers may also use paths that satisfy other criteria aside from low detection. For instance, majority of live animal traffickers appear to use paths with shorter travel times (*i.e.,* direct flights) and avoid stopovers [8]. So there is a dynamic between smugglers and biosecurity forces. We aim to model this dynamic and consequently determine the optimal resource allocation across ports of entry.

We consider a smuggler who attempts to travel from a source into a foreign destination country. This country can place biosecurity resources across ports of entry in their jurisdiction. We assume that the smuggler picks a path that either minimises travel time, maximises the probability of going undetected, or some combination of these two objectives. Biosecurity agencies allocate resources to maximise their probability of catching the smuggler, however smugglers may change their path to avoid detection. This results in a dynamic game between the two players. We aim to find the optimal biosecurity resource allocation.

Games between attackers and security agencies have been studied before [5, 17, 19, 20, 23, 29, 32, 46]. However we use a different method of allocating resources. Furthermore, most studies are motivated by the interception of drugs and weapons trafficking and few have looked at wildlife crime. Catching poachers was the main motivation of the papers that looked at wildlife crime [14, 30, 59]. As far as we know, none of these models have been applied to international wildlife trafficking. Wildlife trafficking has some features that separate it from other illicit trades. For example, live animals may need careful handling and fast delivery [39]. On the other hand, drugs can be placed in almost any type of packaging and left in transportation for large amounts of time.

We describe implementations and analyse two different evolutionary algorithms to find the optimal resource allocation in our model, in Chapter 4. The first algorithm is a Genetic Algorithm (GA) [54] and the second is an Induced Natural Selection Heuristic (INSH) [38]. We test the accuracy and run-time of the algorithms on random Waxman graphs. The accuracy is tested on small networks, and both algorithms find accurate solutions. The algorithms also both converge quickly to solutions, but the GA on average converges to better solutions.

To investigate our model in a practical setting, in Chapter 5 we develop a case study motivated by the trafficking of pangolins. There is a large demand for pangolins in Vietnam, so we consider commodities being smuggled from African and Asian sources into Vietnam. A network of potential transport pathways is created from flight data (www.oag.com).

An important element of the model is the detection probability at each node, and how this is influenced by additional resources. We create estimates of the proportion of illegal flow that went undetected through countries. These estimates come from analysis of international pangolin seizure records. To see how these estimates are influenced by resources, we consider the corruption perception index (www.transparency.org/research/cpi/overview) as a proxy for resources. We fit a logistic model to the detection estimates against the corruption perception index of their corresponding country. This function is assumed to describe how the addition and removal of resources changes the detection probability.

Our algorithms are used to find the optimal biosecurity allocation at Vietnamese ports of entry for different smuggler objectives. There are five airports that act as entry points into Vietnam. They are located in Ho Chi Minh City, Hanoi, Da Nang, Can Tho, and Nha Trang. We first assume that there is smuggler traffic from 189 sources across Asia and Africa. We find that as the smuggler becomes more detection sensitive, the optimal allocation is to spread resources evenly across Ho Chi Minh City, Hanoi, and Da Nang. This allocation appears robust across other smuggler objectives as well. Given we do not know the smugglers true objective, this would be the best allocation to implement. Ho Chi Minh City, Hanoi, and Da Nang have the most paths from sources into them. So spreading resources between these three nodes intuitively makes sense. However, to reach this conclusion, our model considered the possibility of smugglers choosing alternative ports of entry with lower detection. Despite this, the best allocation was still to use the other three ports of entry.

We find that most of the frequently used transit locations are major busy airports including Bangkok Suvarnabhumi International Airport and Paris Charles de Gaulle Airport. Once again, this result makes sense as these hub airports have more connections and thus are more likely to lie along paths from sources to destinations.

Realistically, there is likely to be different traffic volumes from sources. So we also consider the scenario of non-uniform traffic from sources into Vietnam. Traffic volumes are estimated from passenger air traffic data. The optimal resource allocations under non-uniform source traffic are different to the uniform source traffic allocations: less resources are now placed at Da Nang. The previous robust allocations do not perform as well in this case, however are still better than assigning equal resources to all five entry points.

Many assumptions are made in this model of the smuggler's movement. For instance, not all smugglers may take the same path from source to destination. This could be due to other objectives or different perceptions of parameters. So in Chapter 6, we introduce the *multi-path model*. Traffic is distributed among the source-destination paths, rather than going along a single path. The resulting optimal resource allocations approach the single path allocations as more traffic is directed along the optimal path. The most robust allocation is to spread resources evenly between Ho Chi Minh City and Hanoi. This is robust across both smuggler objectives (time and detection trade-off), as well as when different proportions of traffic take the optimal path.

So we have analysed the model through a real world study of trafficking into Vietnam and found robust allocations. These allocations generally assign more resources to ports of entry (either Ho Chi Minh City, Hanoi, or Da Nang) that have more connections to sources.

Though this case study looked at the smuggling of pangolins into Vietnam, the insights we found are likely to be the same in other transport networks. A robust allocation across different smuggler objectives, in both models, was to spread resources evenly between the ports of entry with the most paths into them. This is likely to be the same in other transport networks where some ports of entry have many more connections. It is a better allocation than spreading resources amongst all ports of entry, even if a few of the smaller ports have lower detection.

Our approach uses computational optimisation and focuses on intercepting trafficking. However, there are frameworks in the broader mathematical bioeconomics literature that also attempt to reduce illegal wildlife trade, trafficking, and poaching through other interventions [7, 13, 22]. For instance, comparing the return of investment of demand reduction campaigns to that of law enforcement [22], or the influence of economic incentives on hunting techniques, and the consequent ecological impacts.

Chapter 2

Related Work

We are motivated here by the goal to interdict illegal wildlife trade networks. To do so, we aim to model the behaviour and interactions of two opposing forces: smugglers and biosecurity forces. We assume that a smuggler moves through a transport network to reach one of several target nodes. Biosecurity forces must be strategically allocated to intercept the smuggler.

Problems where we act on networks to disrupt an adversary's activity are referred to as *network interdiction problems* [19, 23, 55]. Defenders place resources on a network to protect against the attacker who moves accordingly. This resource allocation problem is related to Stackelberg games [16, 33]. Stackelberg games involve two players, the leader (defender) and followers (attackers). The leader makes the first decision in the game and the followers fully observe this action before making their own decision.

A *network* is a structure $G(N, E)$ constructed by a set N of nodes and a set E of edges [45]. The edges connect the nodes together. In an *undirected network*, the edges have no directions. An edge between any pair of nodes means that either is accessible from the other along that edge. A *weighted network* is one with *weights* assigned to each edge. These weights may be physical distances between nodes or costs of travelling from one node to another.

A defining aspect of each model is how a smuggler chooses to move across a transport network. The smuggler is assumed to have an objective that governs their actions. These actions are usually made with full knowledge of the defender's decisions. Common smuggler objectives include minimising the shortest-path length [5, 17, 19, 23, 29, 32], maximising network flow [20, 55], or successfully reaching a target node with high reward [46].

There are various defender actions in the literature. An action generally involves choosing a subset of edges to interdict. Some typical examples of the effects of actions include removal of the smuggler's access to certain edges [5, 29], and increases in edge weights [17, 19, 32]. These edge weights are sometimes detection probabilities [32, 34]. Or in other models, they are simply distances between nodes [17, 19]. The resources assigned for interdiction are always limited. This reflects real world budgets. We will now discuss these various interdiction methods in more detail.

In many cases, interdiction is considered to be all or nothing. That is, if we place resources on the edge, it cannot be used by the smuggler, and so it is effectively removed from the network. Ball and Golden [5] studied the problem of finding a set of edges whose removal from the graph resulted in the greatest increase in shortest-path length between two nodes. Each edge was assigned a removal cost, and the budget for edge removal was fixed. They showed that the problem was NP-hard. Malik [29] considered a special case of this problem. Their goal was to find the set of k edges whose removal from the network created the greatest increase in shortest-path length. They proposed an algorithm to solve this problem. They also developed an algorithm to find the single edge that caused the greatest increase in shortest path. This algorithm was shown to have the same time complexity as Dijkstra's shortest path algorithm. However in practise, it may be difficult to remove whole edges from physical transport networks. For example, if an adversary is moving through a transport network where the edges are flights, it may be challenging to have 100% interdiction on that flight. It may be more realistic to have some increase in penalty along a chosen edge rather than total removal.

Rocco and Ramirez-Marquez [41] looked at the problem of minimisation of the maximum flow between a source and sink given limited resources. Their aim was also to determine which edges to interdict. As before, their interdiction variable was binary. Each edge was either interdicted or not. If an edge was not interdicted, then all flow was still allowed along it. They took an evolutionary optimisation approach to the problem, much as we will do. Their approach was based on three processes. The first was to produce possible solutions through a Monte Carlo simulation. Next, they calculated the maximum $s - t$ flow for each of the solutions. The final stage was to estimate how likely it was that each edge was in the final solution. But our approach does not use Monte Carlo simulations. Instead, we use two evolutionary algorithms, including a genetic algorithm, to sample new solutions and find the optimal allocation.

So far we have discussed the limitations of binary models where edges are removed or traffickers are always caught along an edge assigned a resource. Other models have different consequences of interdiction on edges. For example, some models increase the edge length in the presence of a resource.

Israeli and Wood [23] considered the problem of interdiction on edges to maximise the shortest-path length. An interdiction action on an edge either removed it or increased its length to some fixed amount. Each edge k was increased by d_k, $d_k > 0$, if interdicted. A certain amount of resource was required to interdict an edge. Once again, the total amount available was restricted. They formulated the problem as a mixed integer program that could be solved directly. This approach is more realistic than completely removing edges. However the interdiction variable on each edge was still binary. That is, either the required amount must be allocated to an edge for interdiction, or none at all. This is a limitation. The addition of resources may still increase the length. In our case, rather than a physical edge length, this is a weight (or cost) associated with each edge.

Wood [55] considered the problem where an attacker aims to maximise flow through the network which an interdictor attempts to minimise. The interdictor tries to achieve this through breaking edges. Each edge k is assigned an edge capacity of u_k units. The interdictor must spend r_k units of resource to remove the edge. These resources are limited. Edges were initially not allowed to be partially broken. As mentioned before, this is a limitation in such models. They developed an integer program and showed that it could be extended to generalisations of the problem such as multiple sources. However, they did extend their model to consider partial interdiction. They were able to assign $f_k r_k$ resources to edge k, where $0 \leq f_k \leq 1$. This reduced the edge capacity to $(1 - f_k)u_k$. However this linear decrease in capacity is not realistic in our case.

Another network security problem that is close to ours is considered by Morton *et al.* [32]. They were motivated by the detection of attempts to smuggle nuclear materials. The aim was to increase the probability of detecting a smuggler by allocating resources to a transport network. They considered the problem where a smuggler moved from a source node to a destination node via their maximum reliable path. Each edge (i, j), with no sensor present, was assigned a probability of traversing undetected, p_{ij}. The interdictor had to place a limited number of sensors on the edges. The interdictor's decision variable was x_{ij}. The variable x_{ij} was binary; 1 if a sensor was installed along edge (i, j), or 0 if no sensor was installed. Only a single sensor could be installed along each edge. The probability of travelling undetected along an edge with a sensor was $q_{ij} < p_{ij}$. The events of detection on distinct edges were assumed to be independent. Given these probabilities, the smuggler selected a path through the network to maximise their probability of going undetected. The interdictor had to select edges on which to place sensors to minimise this probability.

The authors addressed the idea that resources are also placed at nodes, such as airports and seaports. They incorporated this into the model by splitting such a node into two nodes separated by an edge. The edge between the nodes represented travel through that location.

In the stochastic version of this problem, the origin-destination pairs, (s, t), were known through a probability distribution. The defender's aim was then to minimise the expected value of the path evasion probabilities, taken over all origin-destination pairs.

They were motivated by the placement of radiation sensors at Russian customs to deter smuggling of nuclear material. In the model, the origin nodes were located inside Russia and the destination nodes outside of Russia. Morton *et al.* modelled this as a bipartite network. Suppose AD is the set of edges that sensors can be placed on. Then each origin-destination path $s - t$ contains one edge in AD. This means that along each $s - t$ path there was exactly one checkpoint. This property simplified their formulation.

Finally, they also considered the scenario where the smuggler and interdictor have different perceptions of the detection probabilities. However the interdictor remained aware of the smuggler's belief of detection probabilities and information used to select a path. But the smuggler was only aware of a subset of sensor locations. This seems like a more realistic scenario. Perfect knowledge of all resource allocations is unlikely.

Pan and Morton [34] then considered a generalisation of this problem. They removed the assumption of a bipartite transport network. The problem was once again framed as a stochastic mixed integer program and solved with an L-shaped decomposition.

This work was further extended in [31]. They assumed the budget for sensor installation to be unknown. Instead, the budget had a known probability distribution. The solution was then a priority list of locations where sensors should be installed. When the budget was revealed, the sensors were allocated according to their rank until the budget was exhausted.

However in all of these models, we weren't able to choose the amount of resources assigned to each edge. This is a limitation.

As we have mentioned, the binary interdiction variable is a limitation. We do not know what will happen if more resources are added to an edge. We briefly mentioned that Wood [55] considered partial interdiction on edges. There are some other works as well that look at continuously increasing edge lengths as resources are added.

Fulkerson and Harding [17] considered the problem of increasing edge lengths continuously to maximise the shortest-path length. Each edge was assigned a cost of increasing the length by a single unit. The total fund for increasing edge lengths was restricted. The interdiction variable was non-binary. In some networks it may be difficult to break or remove edges completely. In this case, we could specify the amount to pay to increase the distance of an edge. The change in distance was linear in the cost, however in our case this is unrealistic. Suppose that each edge has an associated probability of being caught if traversed. Then perhaps there becomes a point when the addition of more resources no longer increases the probability of detection (this must be below one anyway). A function with limiting returns on the addition of resources may be more suitable.

Golden [19] looked at a closely related problem. They had to increase the length of the shortest path in the network by τ units. The aim was to spend the minimum amount increasing edge lengths by increments of one unit. This problem was shown to reduce to a minimum cost flow problem. They also looked at interdiction across multiple source-destination shortest path pairs. Considering multiple sources is more realistic than a single source. However, they still had a linear cost function.

As well as these models of the interactions between smugglers and defenders, there are formal mathematical games that look at such relationships. However, many of the previously discussed limitations also apply to these games. Tsai *et al.* [46] considered a binary interdiction variable. They investigated the allocation of resources on a network to protect targets, for example, buildings connected by a road network. They framed the following scenario. An attacker assigns values to reaching one of multiple targets. The defender must allocate checkpoints to capture the attacker before he reaches the target. The attacker starts at one source node and moves through the network towards a target. The defender must place a limited number of checkpoints on the network edges. The attacker is captured if their path intersects with an edge where a checkpoint has been placed. An unsuccessful attack is associated with a payoff, and a successful attack with a penalty depending on the target. They assumed that the game was zero-sum. They developed a linear program to find the optimal strategy. The main limitation of this game is the binary placement of checkpoints on edges. Another limitation is the assumption that if a smuggler traverses an edge with a checkpoint, then they are caught. It is more likely that they are caught with some higher probability in the presence of a checkpoint.

Guo *et al.* [20] considered an interdiction action on an edge to intercept some proportion of flow. The game theoretic model aims to reduce the total flow through the network. There is a single 'inspection station' located on certain edge and nodes. If the inspection station i is operated, a constant $\tau_i \in [0, 1]$ is assigned to that station representing the proportion of flow interdicted. Equivalently, this is the probability that the trafficker will be caught along the edge/node. The defenders strategy is to choose which stations to operate. This is a binary variable. The attackers strategy is to choose the amount of flow assigned to each path, and this must satisfy capacity constraints along each edge. The attacker tries to maximise their utility which is the sum of successful flows on all paths. This problem was formulated as a linear program. However, it is difficult to solve as there is an exponentially large number of variables (paths and allocations). Thus a new algorithm was proposed to solve it. In this model, no flow is interdicted along edges where no inspection station is operated. So, the proportion interdicted along each edge is either 0 or $\tau_i \geq 0$. Only a single inspection station may be operated along any edge. The presence of an operating inspection station increases the probability of detection to some fixed amount. This model does not consider how the detection may change if we add multiple inspection stations to an edge. However in real life we may be able to do this.

Network security problems are motivated by a range of applications. These are usually drug and weapons smuggling. For instance, Jain *et al.* [25] tested their model on capturing an attacker moving through a road network in Mumbai. Guo *et al.* [20] proposed using their model to interdict drug flow across the U.S-Mexico border. Interception of nuclear weapons smuggling is another studied application [32, 34]. None of the studies discussed so far have been motivated by mitigating wildlife trafficking. However, there has been some work that applies game theoretic methods to wildlife crimes [14, 30, 59].

Illegal logging threatens forests and important wildlife habitats. Budgets for forest protection are limited. One method of forest protection is patrolling. These patrols are formed by various groups including volunteers and police, and there are often different costs associated with each group. McCarthy *et al.* [30] considered the problem of allocating resources to intercept illegal loggers on a network formed by roads and rivers in a forest region. Illegal loggers attempted to travel from sources to targets with different pay-offs. They aimed to select the best team for a patrol and the best allocation of this team across the network. Each type of team member (e.g. police or volunteer) had an associated cost and probability of successful interception. They could also only patrol a certain number of connected edges. In contrast to the other security games, multiple resources could be placed along edges. This represented coordination of different groups. Additional resources on an edge increased the detection probability. Similar to previous models, a single allocation of a resource increased the detection probability to some fixed amount. When multiple resources were added to an edge, the probability of non-detection was the product of the probabilities of non-detection of each resource type on that edge. The addition of more resources to the same edge or node is a decision that can be made by biosecurity officers. We consider this in our wildlife trafficking interdiction model. However instead of discrete changes in detection probability, we consider continuously changing this as a function of resources.

Poaching is another threat to the survival of many species. Yang *et al.* [59] developed a program with the aim of improving wildlife ranger patrols to catch poachers. They modelled the dynamic between poachers and rangers as a Stackelberg network security game. The ranger was the leader and a population of poachers were the followers. The ranger aimed to protect the wildlife by patrolling particular locations, and the poacher aimed to attack areas where a ranger was not present. An example of an attack in real life could be the placement of a snare. If a poacher attacked a patrolled area than they received a pay-off that was less then that with no ranger present. Given a ranger strategy, the probability that the poacher selected a certain target depended on the poacher's behavioural model. They incorporated heterogeneity into the poacher's behaviour model (*i.e.,* poachers make independent actions). A population of poachers are likely to make different target choices. They then created a framework that improves the behavioural model with the addition of new data, such as snare locations.

These games incorporate some interesting ideas such as allocating multiple resources to an edge and considering a population of smugglers that make independent assumptions. In our interdiction model we aim to address some of these concepts as well.

The main limitation of the existing models is the binary and discrete interdiction variables. Table 2.1 summarises the most relevant models discussed in this chapter. The interdiction action on an edge often changes the detection probability to some fixed amount. We aim to relax this in our model. Instead, we will consider how continuously adding resources to a node increases the detection probability. We will also consider non-linear functions for this. In many other models, the smuggler is assumed to take the path with the minimal probability of detection. This seems to be a reasonable assumption. However, we will also consider the objective to minimise travel time, as well as combinations of both minimising travel time and maximising the probability of non-detection. For wildlife trafficking, this may be more essential if the goods are live or frozen. In the following chapter, we formally introduce our wildlife smuggling interdiction model.

Paper	Interdiction Variable	Defender Objective
Ball and Golden [5]	binary	maximise shortest-path length
Malik [29]	binary	maximise shortest-path length
Rocco and Ramirez-Marquez [41]	binary	minimise network flow
Israeli and Wood [23]	discrete	maximise shortest-path length
Wood [55]	discrete continuous (linear decrease)	minimise network flow minimise network flow
Morton *et al.* [32]	discrete	minimise probability smuggler goes undetected
Pan and Morton [34]	discrete	minimise probability smuggler goes undetected
Fulkerson and Harding [17]	continuous (linear increase)	maximise shortest-path length
Tsai *et al.* [46]	binary	maximise defender reward
Guo *et al.* [20]	discrete	minimise network flow
McCarthy *et al.* [30]	discrete	maximise expected defender utility
Yang *et al.* [59]	discrete	maximise expected defender utility

Table 2.1: Properties of the most relevant interdiction models. A *binary* interdiction variable indicates that interdiction corresponds to effective removal of edges in the network. A *discrete* interdiction variable means that interdiction corresponds to an increase in the edge weights (*e.g.,* distance, or detection probability) by a fixed amount. A *continuous* interdiction variable means that interdiction corresponds to continuous choice of edge weights, subject to a constraint.

Chapter 3

Smuggling Interdiction Model

Smugglers transport wildlife from sources to destinations across the globe. They are able to choose the route they take to move these commodities across international borders. We want to maximise the probability of interception. Biosecurity resources are limited and so must be allocated intelligently.

This chapter introduces a novel two player game between smugglers and biosecurity agencies. The smuggler attempts to transport goods and their route is determined by some objective. The biosecurity agencies allocate resources to maximise their probability of catching the smuggler. However, the smuggler is assumed to have perfect knowledge so may alter their path to avoid detection. This results in a dynamic game between the two players. We aim to find the optimal biosecurity resource allocation.

This resource allocation problem is broken down into two parts. The first is to predict the path a smuggler takes given a resource allocation, and the second is to optimally allocate resources using this knowledge.

3.1 The Smugglers' Objectives

A smuggler attempts to move commodities from a source into some intended destination country. Certain routes or points of entry into this country may be more desirable for the smuggler to take.

We model the underlying transport choices as a network $G(N, E)$ where the set of nodes N consists of ports (*e.g.,* airports or seaports). Initially, we specify a single source node $s \in N$ and a set of destination nodes $D \subseteq N$ in the destination country. All the nodes, aside from the destination nodes, also act as transit nodes. Transit nodes facilitate paths from sources to destinations.

We later consider multiple sources, as it is more realistic that illegal products come from numerous origins. For example, wild pangolins are found in countries across both Asia and Africa.

The smuggler's goal is to enter a country by whichever destination node is optimal. Once inside the country, transportation is lower risk, so their goal is to get past border protection.

Each node $i \in N$ has a probability p_i that the smuggler is caught at that location. We assume that transit using edge (i, j) means the smuggler can be detected at either node i or node j. We assume that detection at any two nodes is independent such that the probability of detection at least one end of an edge is

$$p_{ij} = p_i + p_j - p_i p_j, \quad \forall i, j \in N, i \neq j.$$

This corresponds to the smuggler having a chance of being caught both entering and exiting a node; these probabilities are assumed to be equal. This model may not be completely realistic. Consider transport via air. It is likely security will be high when entering a foreign country, however security levels may be different if you are just stopping over at an airport and not exiting.

As well as a probability of detection, each edge in the network has an associated travel time, t_{ij}. Live and frozen commodities may be time sensitive. Consequently, one potential objective of a smuggler is to minimise their total travel time. This objective is written mathematically as

$$\min_{r \in L_{s,D}} \sum_{(i,j) \in r} t_{ij},$$

where $L_{s,D}$ is the set of paths from the source node s to any destination node $d \in D$. This problem is a *shortest path problem* [45].

Alternatively, traffickers may determine paths based on the overall probability of detection. That is, they might want to maximise their probability of non-detection,

$$\max_{r \in L_{s,D}} \prod_{(i,j) \in r} (1 - p_{ij}).$$

This optimisation problem is equivalent to

$$\min_{r \in L_{s,D}} \sum_{(i,j) \in r} -\log(1 - p_{ij}).$$

This is also a shortest path problem with edge weights $-\log(1 - p_{ij})$.

A more likely scenario is that the smuggler chooses a route based on a combination of travel time and probability of detection. This combined objective can be expressed as minimising a weighted sum of time and log of the probability of non-detection. That is,

$$\min_{r \in L_{s,D}} \sum_{(i,j) \in r} -\alpha \frac{\log(1 - p_{ij})}{\pi} + (1 - \alpha) \frac{t_{ij}}{\tau}, \tag{3.1}$$

where $\pi = \max\{-\log(1 - p_{ij})\}$, $\tau = \max\{t_{ij}\}$ and $0 \leq \alpha \leq 1$. The values π and τ scale the two objective functions so that they have an even effect on the sum. The weight α controls the trade-off between travel time and probability of non-detection: if $\alpha = 1$, then the path maximises the probability of non-detection; alternatively if $\alpha = 0$, the path minimises travel time. For instance, travel time may be more important for live or frozen commodities. In this case α would be closer to zero. We rarely know the value of α. So in our analyses we will look at how the optimal paths change for different values of α.

Another significant assumption is that smugglers have full knowledge of the network parameters. It is likely smugglers have some sense of detection levels in different countries. However our perception of these levels may be different. We also assume that smugglers instantly observe any changes made to detection. That is, they observe where we place resources. Realistically, there is some lag between allocating resources and the smuggler's observations.

The optimisation problem (3.1) can be solved using graph shortest path algorithms such as *Dijkstra* and *Floyd-Warshall* [45]. Dijkstra is one algorithm that solves the *single source shortest path problem*. It finds the shortest path from a specified node to all other nodes in a connected graph. The edges must have non-negative weights, which is true in our model. However, Dijkstra does not identify multiple equal shortest paths between nodes. The algorithm's simplest implementation has complexity $O(|N|^2)$. Smarter implementations can have complexity $O(|E| + |N| \log |N|)$. The Floyd-Warshall algorithm solves the *all pairs shortest path problem*. It finds the shortest path between all pairs of nodes in a weighted graph. The algorithm's complexity is $O(|N|^3)$.

We have assumed that smugglers choose a path based on a combination of two objectives: minimise time and maximise the probability of non-detection. However, there may be other objectives we have not accounted for. This may mean the smuggler takes an alternative path. Moreover, we have assumed that all smugglers take the same path from a source. There is likely some variation in the path taken. To account for some of this uncertainty, we extend our work to the *multi-path* model in Chapter 6, where we look at distributing some traffic along non-optimal paths as well.

If we can predict a smuggler's path, then we can allocate resources to maximise the probability of interception. The next section will define a resource allocation problem to achieve this.

3.2 The Resource Allocation Problem

Biosecurity resources need to be allocated wisely, as the amount of available resources is limited. Countries are able to place these resources at locations in their jurisdiction.

A destination country places its resources at the *destination nodes*, D, *i.e.,* the nodes in the destination country that the smuggler may enter through. Let $\delta_i \geq 0$ be the resource allocation at node $i \in D$. Resources are constrained such that

$$\sum_{i \in D} \delta_i = B,$$

where B is the total amount of resources available. More resources will increase the overall probability of interception. Without loss of generality, we set B to one. So each resource allocation, δ_i, is between zero and one inclusive. This way, the allocation δ_i can be thought of as the proportion of resources assigned to node i.

Each node has an associated probability of detection. For destination nodes, this probability changes as resources are added and removed. We need to know how this probability of detection changes as a function of assigned resources, δ_i. Whilst we do not know what this function is, there are some reasonable assumptions surrounding it. First, we assume that the function is non-decreasing in δ_i; we do not expect additional resources to decrease the detection probability. Second, the function must be within the range $[0, 1]$, because it is a probability. However the function does not necessarily have to go from zero to one. In fact, a detection probability of one at any node is highly unlikely as this would mean that all illegal trade is detected. Also, we assume that the addition of resources has a limiting effect on the detection probability. So eventually there becomes a point where placing more resources on a node does not increase the probability that the smuggler goes undetected. One function that captures all of these properties is a logistic function,

$$p(\delta_i) = \frac{a}{1 + e^{(b - \delta_i)/k}},$$

where δ_i is the resource allocation at node i, a is the theoretical maximum probability of detection ($0 < a \leq 1$), and b and k are parameters that control the shape. This is not the only function that has these assumed properties. For example, exponential functions could fit the assumptions. However the logistic function is also motivated by data discussed in Chapter 5.

Let $\boldsymbol{\delta}$ be a vector of resource allocation where δ_i is the allocation at destination node i. We want to choose the resource allocation $\boldsymbol{\delta}$, that minimises the probability that a trade from source s goes undetected, $f_s(\boldsymbol{\delta})$. That is,

$$f_s(\boldsymbol{\delta}) \quad = \quad \min_{\boldsymbol{\delta}} \prod_{(i,j) \in r^*(\boldsymbol{\delta})} (1 - p_{ij}), \tag{3.2}$$

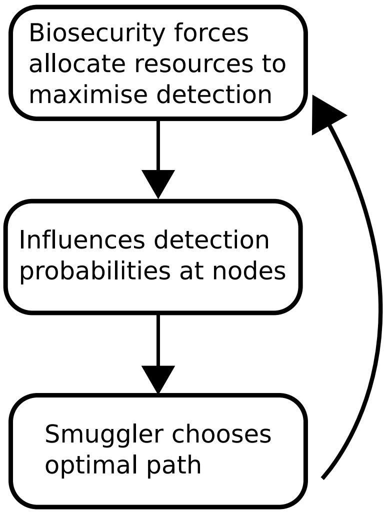


Figure 3.1: The dynamic game between smugglers and biosecurity forces. Biosecurity forces allocate resources among the destination nodes. They aim to maximise the probability of detecting a smuggler. This change in allocation influences the detection probabilities at nodes.

where $r^*(\boldsymbol{\delta})$ is the smuggler's optimal path given the resource allocation $\boldsymbol{\delta}$. This is the solution to the optimisation problem (3.1) where the p_{ij} values depend on $\boldsymbol{\delta}$.

An extension to this model is the *multiple-source problem.* Wildlife commodities do not originate from a single location. We need to consider illegal trades coming from a group of nodes. The objective function of the multiple-source problem is

$$\min_{\boldsymbol{\delta}} \quad \sum_{s \in S} w_s f_s(\boldsymbol{\delta}),$$

where $w_s \geq 0$ is a weight assigned to source s and $f_s(\boldsymbol{\delta})$ (the solution of problem (3.2)) is the probability of non-detection along the smuggler's optimal path from source s into D. The weight w_s can either represent demand, traffic, or how much we value intercepting a trade from source s. For example, if a species from a particular region is at high risk of extinction, we may want to put more effort into their protection. So we would assign these sources larger weights.

The smugglers act the same way in the multiple-source problem as in the single-source problem. A smuggler will take their optimal path given a source node, destination nodes, and resource allocation.

The actions of the smugglers and biosecurity forces directly affect one another when the smuggler aims to avoid detection (*i.e.,* when $\alpha > 0$). This dynamic relationship is captured in Figure 3.1. Biosecurity agencies place limited resources across destination nodes. When resources are placed at a destination node, this increases the detection probability. A detection-sensitive smuggler may adjust their path to optimise their objective. Recall this objective is some weighted combination of minimising travel time and maximising the probability of going undetected. Given this change in path, the biosecurity resources may be theoretically reallocated to maximise the probability of interception. There would exist a back and forth dynamic between the two forces. We aim to find the optimal resource allocation that minimises the probability that a smuggler goes undetected at the end of this game. That is, we aim to find a stable point which maximises the detection probability despite the smuggler's attempts to minimise it.

We have defined a new dynamic game between smugglers and biosecurity forces. We aim to explore some aspects of this dynamic. We will do this through a real world case study. However, we do not know the smuggler's trade-off between the two objectives, α. So we will look at how the optimal resource allocations change with α. These resource allocations may be very different across smuggler objectives. This would make choosing a resource allocation to implement in real life challenging. But, perhaps in our case study we can determine a robust allocation that is good for multiple smuggler objectives.

Recall we want to find the optimal biosecurity resource allocation. In the following chapter, we will describe implementations of two algorithms to find the optimal allocation. But first, we will illustrate the model on the small transport network in Example 3.2.1.

Example 3.2.1
Here we describe the resource allocation problem on a small network. Consider the 4-node network shown in Figure 3.2. Smugglers are attempting to move commodities from the source, node 1, into the destination country formed by destination nodes 3 and 4. The smuggler can choose which node they wish to enter the country through. The edge weights are travel times between the nodes. The numbers beside the nodes are the probabilities of detection at those nodes. This probability will change for nodes 3 and 4 depending on the resource allocations, δ_3 and δ_4 respectively.

The detection probability at a destination node is a function of resources. We use the detection function

$$p(\delta_i) = \frac{0.6}{1 + e^{(0.5-\delta_i)/0.2}},$$

such that the maximum probability of detection at any destination node is 0.6.

When the smuggler's objective is to minimise travel time ($\alpha = 0$), the optimal path is from node 1 directly to node 3. The optimal resource allocation is to place all resources on node 3, $\delta_3 = 1$ and $\delta_4 = 0$, as the smuggler does not care about their probability of being caught.

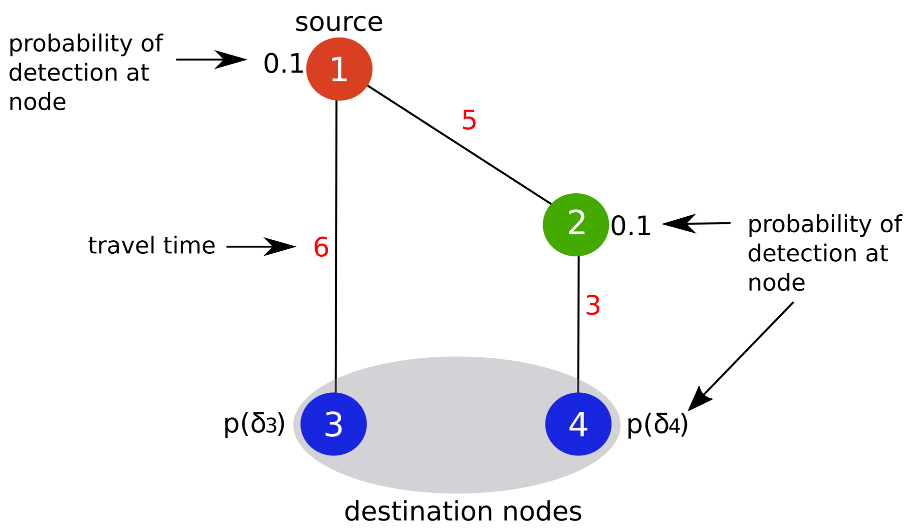


Figure 3.2: A transport network. The smugglers are trying to move from the red source node, 1, into the destination country (blue nodes 3 and 4). The baseline probabilities of detection at the source and transit node are listed next to nodes 1 and 2. The travel time is also displayed along each edge. The smuggler will take a different path depending on their objective, captured by the weight α.

Suppose that the smuggler cares equally about minimising travel time and maximising the probability of non-detection. This corresponds to setting $\alpha = 0.5$. To find the optimal solution, we generate $100,000$ allocations across the search space and evaluate the probability of detection for each allocation. These probabilities are shown in Figure 3.3. For this problem the minimum is approximately 0.4913 and the optimal resource allocation is $\delta_3 = 0.7386$ and $\delta_4 = 0.2614$.

Notice the jump in the function around $\delta_3 = 0.7$ shown in Figure 3.3. Often the detection function for other transport networks also contains jumps. These correspond to changes in the smuggler's optimal path. In this example network, the smuggler travels from node 1 to node 3 for $0 \leq \delta_3 < 0.7386$ until it becomes optimal for them to switch path.

Finally, we consider the scenario where the smuggler only tries to minimise detection, $\alpha = 1$. Once again, we evaluate the probability of detection at each of the $100,000$ allocations when $\alpha = 1$. The minimum probability of non-detection found is 0.5639 with the corresponding allocations of $\delta_3 = 0.6$ and $\delta_4 = 0.4$.

So we have observed three different allocations for $\alpha = 0$, $\alpha = 0.5$, and $\alpha = 1$. However we do not know the true value of α. All of the allocations assign more

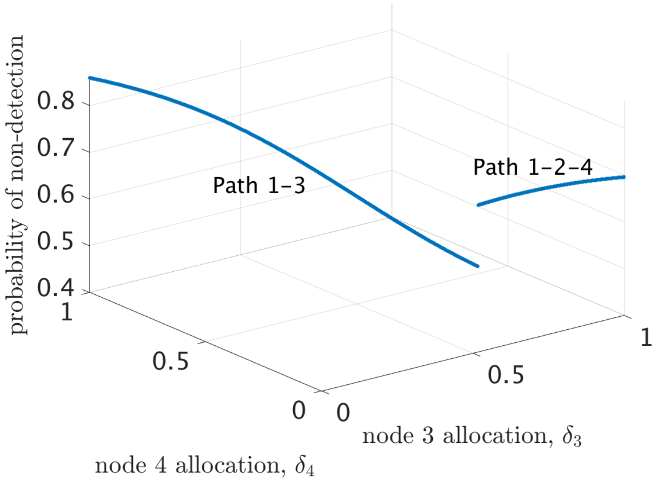


Figure 3.3: The total probability of non-detection, $f_1(\delta_3, \delta_4)$, when the smuggler is travelling in the transport network shown in Figure 3.2. The trade-off between detection and time is set to $\alpha = 0.5$. This indicates the smuggler cares equally about minimising time and maximising probability of non-detection. The function is not connected. Each piece corresponds to the smuggler travelling along a single path. When all resources are placed at node 4, the smuggler's optimal path is $1-3$. As we decrease the resources at node 4, there is a point after which it is optimal for the smuggler to take path $1-2-4$. We aim to find the resource allocation that gives the minimum of this function.

resources to node 3. There may be an allocation with this property that works well for multiple smuggler objectives. We investigate this idea in other transport networks in Chapter 5.

To find the best allocations so far, we have evaluated the detection probability given allocations across a fine grid space. However, as the network grows and there are more destination nodes, this method will become harder to use. Thus, in the next chapter we will describe algorithms that find the minimum of objective functions such as the function in Figure 3.3.

Chapter 4

Methods

In the previous chapter we defined an allocation problem: biosecurity forces place resources at nodes to minimise the probability that a smuggler goes undetected. We want to solve this optimisation problem.

There are a few different aspects of the problem that we need to consider. First, we have observed in the previous chapter, that the objective functions (proportion of trade that goes undetected) may have discontinuities in them that correspond to changes in the smuggler's optimal path. Thus we need an algorithm that finds the optimal allocation given these discontinuities.

Another crucial element will be finding the shortest path through the network, which is the smuggler's optimisation problem. There are existing algorithms such as Dijkstra and Floyd-Warshall [45] that we can utilise to do this. These shortest path evaluations will need to be incorporated into the algorithm we develop.

Our algorithm will need to be adaptable to other scenarios. We want to capture the relationship between resources and detection probabilities. We assume this to be a logistic function. However, in the future we may also want to test other detection functions. We also aim to generalise this model in Chapter 6. So an algorithm that adapts easily to this is ideal.

Also, we have a continuous search space. We want to be able to specify the amount of resources allocated to each node. Without loss of generality, this is a number between 0 and 1. This is a more flexible approach to previous work where detection is either binary or takes a discrete set of values [5, 20, 29, 32, 34, 55].

Evolutionary algorithms are highly suited to solving a wide range of optimisation problems. They begin with an initial population of solutions that gradually evolve to better regions of the search space [3]. They are versatile approaches to many problems that adapt to variations easily. In this chapter, we will describe implementations of two algorithms to solve the optimisation allocation problem: a Genetic Algorithm (GA) and an Induced Natural Selection Heurstic (INSH). The Genetic

Algorithm [54] is a popular class of evolutionary algorithms. The Induced Natural Selection Heuristic [38] was initially developed for optimal Bayesian experimental design. We consider two algorithms rather than a single one to increase confidence that they are finding near optimal solutions. Beyond this, we can compare aspects of them including accuracy and run-time, to find the best approach.

4.1 Genetic Algorithm

A Genetic Algorithm (GA) reflects the evolutionary biological process where DNA is shared between parent chromosomes to create offspring [54]. They are based on the evolutionary theory of natural selection where the fittest individuals are more likely to survive and pass on their genes.

Genetic algorithms operate by examining a set of possible solutions and choosing the best ones. There are five steps in this process: population initialisation, fitness evaluation, selection, crossover, and mutation. We describe them below.

Population Initialisation

The first step in a GA is to generate an initial set of individuals called a *population*. An individual solution is described by a vector $\mathbf{x}$ of length n, analagous to a chromosome in biology. We often randomly generate the initial population.

The population at each generation contains N individuals. This parameter must be chosen by the user. If N is too small, then not enough variation will be present in the population. However if N is too large, then the GA may take a long time to run.

Fitness Evaluation

The *fitness* of an individual is defined by a *fitness function* that determines how well the chromosome performs. This is in terms of the objective function. The probability that the individual is selected for the next phase is based on its fitness.

Selection

The selection process aims to pick the individuals that have the best fitness levels. These individuals then pass on their *good* genes to the next generation. Suppose that individual i has fitness g_i, where a larger g_i indicates better fitness. The probability that i is chosen for the sampling/reproduction phase is

$$p_i = \frac{g_i}{\sum_{j=1}^{N} g_j},$$

where N is the population size. These probabilities form a distribution from which the new population is randomly chosen. We sample with replacement N individuals from the existing population. This sampling method is called *roulette wheel selection*.

There are other methods of selection. For example, *tournament selection*. In tournament selection, subsets of individuals are chosen from the population at random. The fittest individual from each group is selected for reproduction. We choose to use roulette wheel selection as it is straightforward and will allow us to likely sample around the best solutions.

Crossover
Crossover is based on the exchange of DNA sections between parental chromosomes passed onto offfspring. The idea behind crossover is that we mix parts of our best solutions to create a mixture of individuals that combine good features of the parents. One example of a crossover method is *single point crossover*. Each pair of selected individuals crossover with some fixed probability k. A crossover point is picked randomly and parts of the individuals are interchanged. Figure 4.1 shows a crossover process. The crossover point was chosen to be between elemenst 3 and 4. Parents 1 and 2 are split at this point and parts from both of them are combined to form new individuals. The set of these individuals make up a new population.

There are many other crossover methods such as two point crossover. We consider single point crossover because it is the simplest method. It is also a very different method of generating allocations than the other algorithm we will test. We will formally introduce this algorithm in the following section. However, future work may involve testing the performance of other crossover methods.

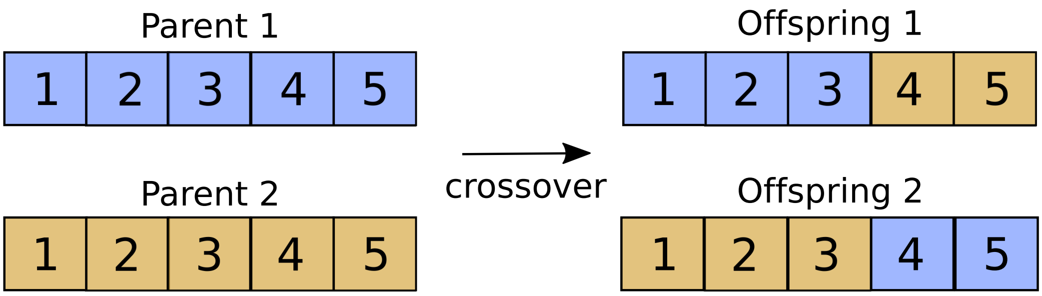


Figure 4.1: Crossover process. Two parent individuals are interchanged between crossover points 3 and 4 to create two new individuals.

Mutation
Mutations occur in new individuals with some probability. This mutation process is shown in Figure 4.2. When mutation occurs we randomly select one element of the individual and assign it a new random number. So each individual has a maximum of one mutation. In the canonical form, mutation occurs with low probability. It exists to introduce variation in the population of solutions to avoid getting stuck in a local minima. We test different probabilities of mutation later.

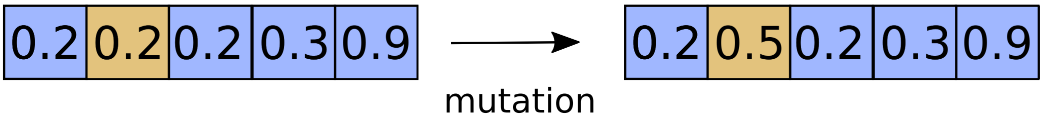


Figure 4.2: Mutation process. One element is randomly selected and assigned a uniformly chosen random number.

We will now define a genetic algorithm that is used to solve our network interdiction optimisation problem defined in Chapter 2. This genetic algorithm largely follows the steps outlined above. Our problem has the additional constraint that each resource allocation must sum to one as we have finite resources. This adds an extra phase to the algorithm where allocations are normalised.

This genetic algorithm is summarised in Algorithm 1. We begin by randomly generating a set of N allocations that sum to one. Each allocation is a vector δ where each element in the vector is the allocation, between zero and one, at a destination node. For example, an allocation for a network with three destination nodes could be $(0.3, 0.3, 0.4)$. Boundary cases are also added to this set. The boundary cases are allocations where you assign all resources to a single node. For example, $(1, 0, 0)$ is the solution of assigning all resources to the first destination node (and none to the second and third nodes).

For each allocation, $\boldsymbol{\delta}$, there is a corresponding proportion of illegal traffic that goes undetected through the network, $f(\boldsymbol{\delta})$. The function $f(\boldsymbol{\delta})$ is evaluated in four main steps. The first is to determine the probability of detection $p(\delta_i)$ at each of the destination nodes i. The second step is to evaluate the edge weights in the network which depend on these probabilities and the smuggler's trade off between minimising travel time and maximising the probability of non-detection, α. The third step is to compute the smuggler's optimal path. This is the solution to the smuggler's optimisation problem. We use Floyd Warshall's shortest path algorithm. This computes shortest paths between all pairs of nodes. The fourth step is to evaluate the probability of non-detection along the smuggler's optimal path from each source. From this, the total proportion undetected is calculated. This algorithm is detailed in Algorithm 2.

The selection of the next generation of individuals uses a fitness function. The fitness function that we use is

$$g(\boldsymbol{\delta}) = 1 - f(\boldsymbol{\delta}).$$

So we are more likely to select resource allocations that give smaller proportions of undetected trade. We then select N allocations based on their selection probabilities.

Allocation $\boldsymbol{\delta}$ is selected with probability

$$p(\boldsymbol{\delta}) \leftarrow \frac{g(\boldsymbol{\delta})}{\sum_{\bar{\boldsymbol{\delta}} \in A} g(\bar{\boldsymbol{\delta}})}, \quad \boldsymbol{\delta} \in A,$$

where A is the set of allocations (population).

For each pair of selected individuals, crossover then occurs with a fixed probability

Algorithm 1 Genetic Algorithm for Network Interdiction

Inputs: W, m, k, N
Generate initial population A of size N
for i = 1 to W **do**
    for each $\boldsymbol{\delta}$ in A **do**
        Set $g(\boldsymbol{\delta}) \leftarrow 1 - f(\boldsymbol{\delta})$ ▷ f is proportion undetected, see Algorithm 2
    end
    for each $\boldsymbol{\delta}$ in A **do**
        Set $p(\boldsymbol{\delta}) \leftarrow \frac{g(\boldsymbol{\delta})}{\sum_{\bar{\boldsymbol{\delta}} \in A} g(\bar{\boldsymbol{\delta}})}$
    end
    $A' \leftarrow$ N allocations from A chosen using roulette wheel selection
    $A'' \leftarrow \emptyset$
    for pairs $(\boldsymbol{\delta}_1, \boldsymbol{\delta}_2)$ in A' **do** ▷ crossover
        $u \leftarrow U(0, 1)$
        if $u \leq k$ **then**
            $(\boldsymbol{\delta}'_1, \boldsymbol{\delta}'_2) \leftarrow$ crossover$((\boldsymbol{\delta}_1, \boldsymbol{\delta}_2))$
            $A'' \leftarrow A'' \cup \boldsymbol{\delta}'_1 \cup \boldsymbol{\delta}'_2$
        else
            $A'' \leftarrow A'' \cup \boldsymbol{\delta}_1 \cup \boldsymbol{\delta}_2$
        end
    end
    for each $\delta \in A''$ **do**
        $u \leftarrow U(0, 1)$
        if $u \leq m$ **then** ▷ mutation
            $\boldsymbol{\delta} \leftarrow$ mutate$(\boldsymbol{\delta})$
        end
        $\boldsymbol{\delta} \leftarrow$ normalise$(\boldsymbol{\delta})$ ▷ normalisation
    end
    Set A to A''
end
Return: $A, f(A)$

Algorithm 2 Proportion of Illegal Trade that goes Undetected, $f(\boldsymbol{\delta})$

1: Inputs: $G(N, E)$, p_i, α, $\boldsymbol{\delta}$, S
2: Compute $p_j \leftarrow p(\delta_j), \quad \forall\ j \in D$ ▷ destination detection prob
3: Compute $p_{ij} \leftarrow p_i + p_j - p_i p_j, \quad \forall\ (i, j) \in E$ ▷ edge detection prob
4: Compute $q_{ij} \leftarrow -\alpha \frac{\log(1-p_{ij})}{\pi} + (1-\alpha)\frac{t_{ij}}{\tau}, \quad \forall\ (i, j) \in E$ ▷ edge weights
5: Compute shortest path r_s from each source s into D ▷ smuggler's optimisation
6: Compute $f_s \leftarrow \prod_{(i,j,)\in r_s}(1 - p_{ij})$ for each $s \in S$ ▷ path detection probability
7: $f(\delta) \leftarrow \frac{1}{\sum_{s\in S} w_s} \sum_{s\in S} w_s f_s$ ▷ proportion undetected
8: Return: $f(\boldsymbol{\delta})$

k. New individuals are more likely to be created for larger crossover probabilities. Parameters in a GA are often specific to the problem. However common values of crossover probabilities range from 0.5 to 1 [44]. We test the convergence of the GA for two high crossover probabilities, $k = 0.9$ and $k = 0.7$, in Section 4.3.

We crossover pairs of consecutive allocations rather than all combinations of allocations. For instance, we crossover the pairs $(\boldsymbol{\delta}_1, \boldsymbol{\delta}_2)$, $(\boldsymbol{\delta}_3, \boldsymbol{\delta}_4)$, and $(\boldsymbol{\delta}_5, \boldsymbol{\delta}_6)$ instead of $(\boldsymbol{\delta}_1, \boldsymbol{\delta}_2)$, $(\boldsymbol{\delta}_1, \boldsymbol{\delta}_3)$, $(\boldsymbol{\delta}_2, \boldsymbol{\delta}_3)$ and so forth. Single point crossover is used. A point is chosen uniformly and the elements are swapped around that point between the pair of individuals to create two new resource allocations. Individuals $\boldsymbol{\delta}_1$ and $\boldsymbol{\delta}_2$ crossover to produce new individuals $\boldsymbol{\delta}'_1$ and $\boldsymbol{\delta}'_2$. If crossover does not occur, then the two resource allocations, $\boldsymbol{\delta}_1$ and $\boldsymbol{\delta}_2$, are carried over to the next generation without alteration (unless mutation occurs).

The next process is mutation. Each allocation is mutated with probability m. A node is chosen uniformly and uniformly assigned a number between zero and one. We test different mutation probabilities in Section 4.3.

The final process is normalisation. This is needed because the GA reproduction processes cause solutions to no longer sum to a constant. Thus each allocation is normalised so that the resources sum to one. This process is shown in Figure 4.3.

The process is repeated for a fixed number of iterations, W. The number of iterations to achieve convergence of a solution will depend on the problem. For instance, problems with larger search spaces (more destination nodes) may require more iterations. This may also depend on the population size N. If we are sampling around a higher number of individuals at each iteration, then less iterations may be needed. We will investigate the convergence of solutions in Section 4.3.

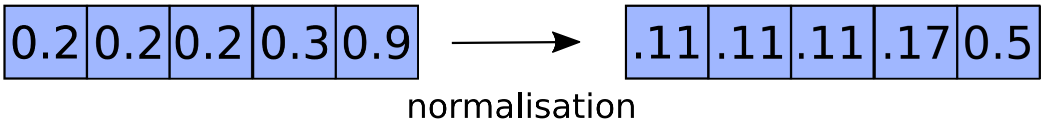


Figure 4.3: The normalisation process in our genetic algorithm. The allocations must sum to 1.

One extension to the basic genetic algorithm is elitism. Some of the fittest individuals are carried over to the next generation. The purpose of this is to not lose the current minimal resource allocation once it has been found. Elitism was added to the GA to solve this problem.

In the next section we describe an implementation of an Induced Natural Selection Heuristic to solve the wildlife smuggling network interdiction problem.

4.2 Induced Natural Selection Heuristic

The Induced Natural Selection Heuristic (INSH) is a search heurisic developed for optimal Bayesian experimental design [38]. It can be applied to a wide variety of optimisation problems. This algorithm has a different sampling process to the genetic algorithm. We choose the *best* solutions at each iteration and sample a set of solutions around this. It allows us to search in different regions of the space and hopefully avoid getting stuck in local minima.

The INSH specific to our network interdiction problem is detailed in Algorithm 3. The algorithm begins by generating a set of N resource allocations. These initial allocations are generated in the same way as in the GA. The proportion of undetected illegal traffic, f, given each resource allocation is evaluated (Algorithm 2). Resource allocations are accepted into the next generation if their corresponding proportion of undetected trade satisfies some acceptance criteria. Then s new resource allocations are sampled around these accepted ones as well as the current optimal solution. We set s so that we have N new solutions. The process is repeated for W iterations.

There are many ways to define the acceptance criteria. In our implementation we accept a fixed number, r, of resource allocations, δ. These allocations have the smallest proportions undetected, $f(\delta)$. Alternatively, allocations within some percentage of the current optimal objective can be accepted.

New allocations are sampled around the accepted ones using a perturbation function. A simple perturbation could be to give a random allocation to a uniformly chosen node in an allocation set. This is similar to the mutation process used in the GA. However we use a different sampling method. Recall that each resource allocation at a node must be between zero and one. The resource allocations across the

Algorithm 3 Induced Natural Selection Heuristic for Network Interdiction

Inputs: n, r, m, W, s
Generate initial set of resource allocations A of size n
for $i = 1$ to W **do**
  for each $\boldsymbol{\delta} \in A$ **do**
    Evaluate $f(\boldsymbol{\delta})$ ▷ f is proportion undetected, see Algorithm 2
  end
  Accept r allocations with the smallest f values.
  Sample s allocations around each accepted allocation and current optimal.
  Set A to be these new resource allocations.
end
Return: A, $f(A)$.

network must also sum to one. Suppose $\boldsymbol{\delta}$ is a resource allocation we want to sample around. Recall δ_i is the allocation at destination node i. We have the conditions $0 \le \delta_i \le 1$ and $\sum_i \delta_i = 1$. Now let $x_i = \sqrt{\delta_i}$. We perturb around this allocation such that $x_i^p = x_i + N(0, \sigma^2)$. Then the new resource allocation is

$$\delta_i' = \frac{(x_i^p)^2}{\sum_{i \in D}(x_i^p)^2},$$

where i is a node in the destination country. Squaring x_i^p ensures the allocation is non-negative. The resource allocation was perturbed using a Gaussian distribution with mean zero and variance σ^2. We test different variances in the next section.

4.3 Algorithm Analysis

We have proposed both a Genetic Algorithm and an Induced Natural Selection Heuristic to find a solution to the interdiction game. The GA and INSH are heuristics that are not guaranteed to find the optimal allocation, so we need to make sure they get close to the optimal. Thus in this section we test how the algorithms perform on simulated networks. We investigate the accuracy, convergence, and run-time of the algorithms. Both algorithms are implemented in MATLAB, and simulations are performed in MATLAB as well.

4.3.1 Waxman Graphs

We need to generate a large number of test networks to evaluate the performance of the algorithms. Waxman graphs are randomly generated graphs that incorporate distance between nodes [53]. The networks we use in Chapter 5 are real world networks where the nodes are cities and the edges are weighted by travel time (distance). Thus Waxman graphs incorporate one of the features (*i.e.,* distance) that we need in the interdiction game.

Waxman graphs are generated by randomly assigning n nodes to positions in a unit square. The distance between each pair of nodes is measured by their Euclidean distance, d_{ij}. The probability that there is an edge of length d_{ij} between nodes i and j is

$$p_{ij} = qe^{-sd_{ij}},$$

where $s \in [0, \infty)$ and $q \in (0, 1]$. We then place an edge between each pair of nodes i and j with probability p_{ij}. The edge between i and j is weighted by the distance, d_{ij}. The parameter s controls the influence of distance on the edge probability. The parameter q controls the edge density. We use these Waxman graphs to test our algorithms.

Let n be the number of nodes in the network. We need to set the parameters $n, q,$ and s. We set the number of nodes, n, to $n = 50$. The real networks we will eventually use are small, thus this gives us networks close in size to the ones we need in the real scenario. Also, one test we aim to do is create a fine grid of the search space, and evaluate the objective function at each point. Relatively small networks will enable this to be done faster. To set the parameters q and s, we consider a real world network. This transport network was constructed from flight data (www.oag.com). It had African and Asian sources and the destination country was Australia. There were 32 nodes and 142 edges. The edge density of the undirected network was

$$D = \frac{2e}{n(n-1)} = 0.2863,$$

where e is the number of edges. We generated 400 networks under the parameters $s = 5$ and $q = 50/n$. The value of s was chosen as if s is too small, we don't see much effect from distance, but if s is too large, the network becomes disconnected. The mean density of these networks was 0.2869. As this is close to the real world network density, we use these parameters in our following simulations.

4.3.2 Accuracy

In this section we define various interdiction problems and evaluate the accuracy and convergence of GA and INSH, described in Algorithms 1, 2, and 3.

We do not know the true optimal resource allocations of our problems. To evaluate the accuracy of the GA and INSH, we compare their solutions to a fine grid solution.

This grid covers the search space and contains $2,000,000$ allocations. We then evaluate the proportion of illegal traffic undetected under each of these allocations. The minimum of these gives a close approximation to the optimal solution.

There are a few other parameters that we must set in the interdiction problem. The first is the number of source nodes, which is set to two for our tests. We choose two because it is a more interesting problem than just a single source. The number of destination nodes is also set to two. This gives us a smaller sized search space which we aim to search by generating a fine grid that covers the space. We later consider problems with a larger number of destination nodes, but do not evaluate them across a fine grid. The probability of detection at each node is assigned a uniformly generated number between 0.2 and 0.5. We assume that high probabilities of detection are more unlikely. The detection probability at destination nodes is given by

$$p(\delta_i) = \frac{0.6}{1 + e^{(0.5-\delta_i)/0.2}},$$

where δ_i is the resource allocation at node $i \in D$. The final parameter we set is the assumed smuggler's trade-off between minimising travel time and maximising the probability of going undetected, α. This is set to $\alpha = 0.8$. This corresponds to the smuggler caring more about detection than time.

On the same 400 networks, we run the INSH and GA until they reach the grid search minimum or for $40,000$ iterations. In each iteration, 50 allocations are sampled. So, a maximum of $2,000,000$ allocations are evaluated. This is no more than the number of evaluations in the grid search. In the genetic algorithm we set the crossover probability, k, to 0.9 and the mutation probability, m, to 0.1. In the INSH we set the number of solutions we sample around, r, to 5. These parameters are chosen initially because they gave us solutions close to the optimal when tested on smaller cases. We evaluate the performance of the algorithms under different parameter choices shortly.

We calculate the differences between the grid search optimum and the minima found by the GA and INSH. The GA and INSH find better solutions than the grid search in 68.5% and 72% of cases respectively. For these cases, the difference is set to zero. The median and maximum differences are shown in Table 4.1. The median for both is zero. The maximum differences for both are quite small as well. The GA and INSH results are very close to the grid search minimum. There doesn't appear to be important differences between the INSH and GA in terms of accuracy in these problems.

These problems all had only two destination nodes and were tested on one set of parameters in the algorithms. We now test the convergence of the algorithms for problems with different sized search spaces and algorithm parameters. We consider problems with two, five, ten and twenty destination nodes. For the same parameters as in the exhaustive search, we run the algorithms for 5000 iterations and record the

Error	INSH	GA
Median	0	0
Maximum	1.996e-07	7.140e-05

Table 4.1: Error in algorithms compared to a fine grid search. 400 networks are tested. The error is the difference between the algorithms and grid search solutions. The difference was set to zero in cases where the algorithm solutions were better than the grid search. The differences for both algorithms were very small.

optimal at regular intervals.

For each network, the proportion decrease in solution is calculated. The decrease at iteration i is

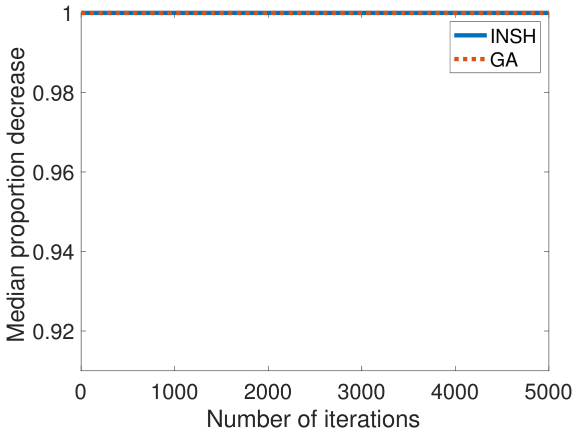
$$\frac{f^i}{f^0},$$

where f^i is the current minimum at the i^{th} iteration and f^0 is the minimum from the initial set of allocations.

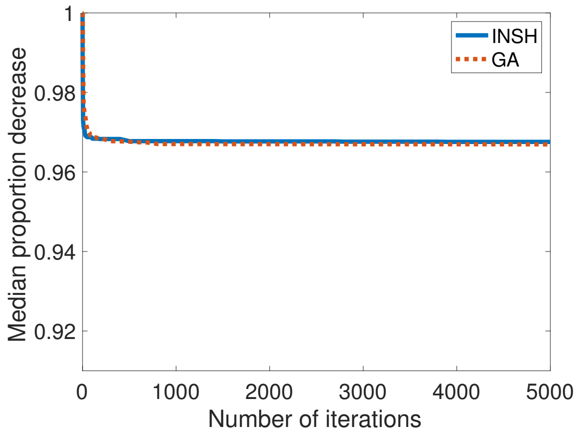
The initial set contains 50 allocations. Boundary cases are included in this set. These are the cases where all resources are placed at a single node. Randomly generated allocations made up the remainder of the initial set. To generate an allocation we create a vector of $|D|$ uniform random numbers between zero and one and normalised such that the elements sum to one (recall D is the set of destination nodes).

The median decreases for the 400 networks are shown in Figure 4.4. We first note that the solutions converge fast, especially in problems with ten or less destination nodes. The GA reaches a better median minimum than the INSH in the five, ten and twenty destination nodes problems. This suggests its sampling method is more effective for these problems than the method used in the INSH.

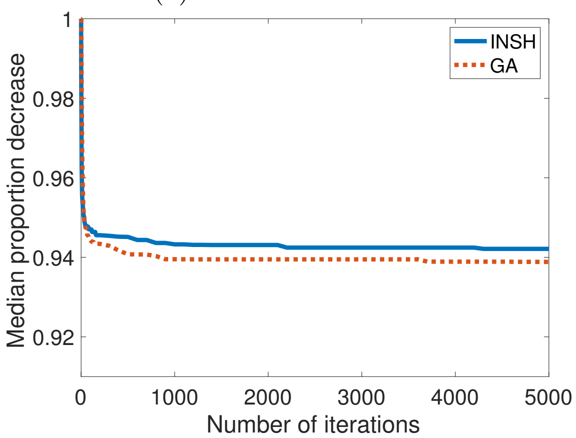
The median decreases in solution for both GA and INSH become greater as the number of destination nodes increases. The median decrease with GA for five destination nodes, at 5000 iterations, is around 0.97, and 0.92 for twenty destination nodes. This is because the search space is larger for problems with a higher number of destination nodes. So we see greater improvement beyond the initial set of random allocations. For the two destination node problems, the median does not decrease. This indicates that no improvement beyond the initial random set of allocations was found. One explanation for this is that there are less paths into a destination country with only two nodes. So it is more likely that an optimal solution is to place all resources at a single node. This allocation is part of the initial set. Some of these simulations did see improvement though. Figure 4.5 shows boxplots of the proportion decreases for the two destination node problems. We see that the median is one but the greatest improvement at 1000 iterations is 0.9. Thus there are solutions other



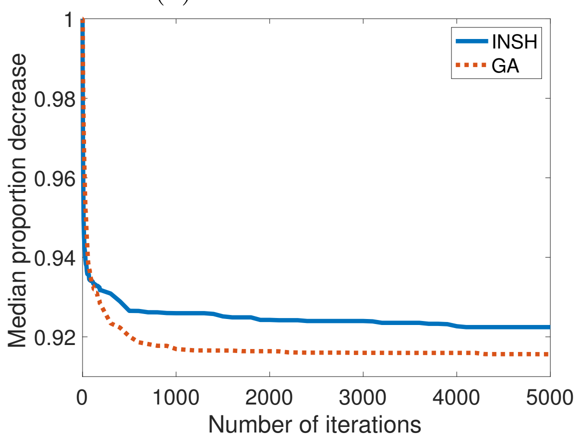
(a) Two destinations.



(b) Five destinations.

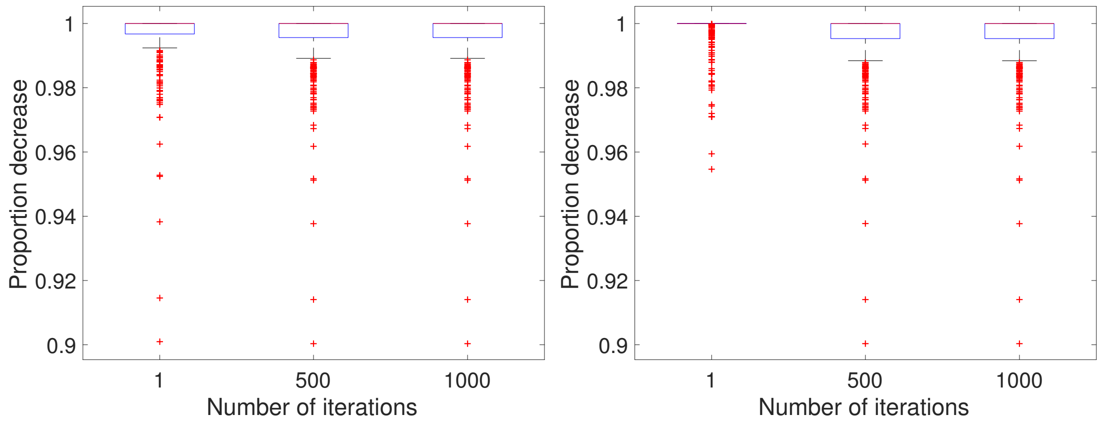


(c) Ten destinations.



(d) Twenty destinations.

Figure 4.4: Median proportion decreases in minimum for problems with with (a) two, (b) five, (c) ten and (d) twenty destination nodes. The median is taken from results from 400 networks. In each of the cases, GA converges to a better (lower) solution. We also see more improvement for problems with a higher number of destination nodes. Overall, there is not much improvement for the two destination case. This is because the search space is smaller. So we are able to find a good solution in the initial set of random allocations.



(a) INSH (b) GA

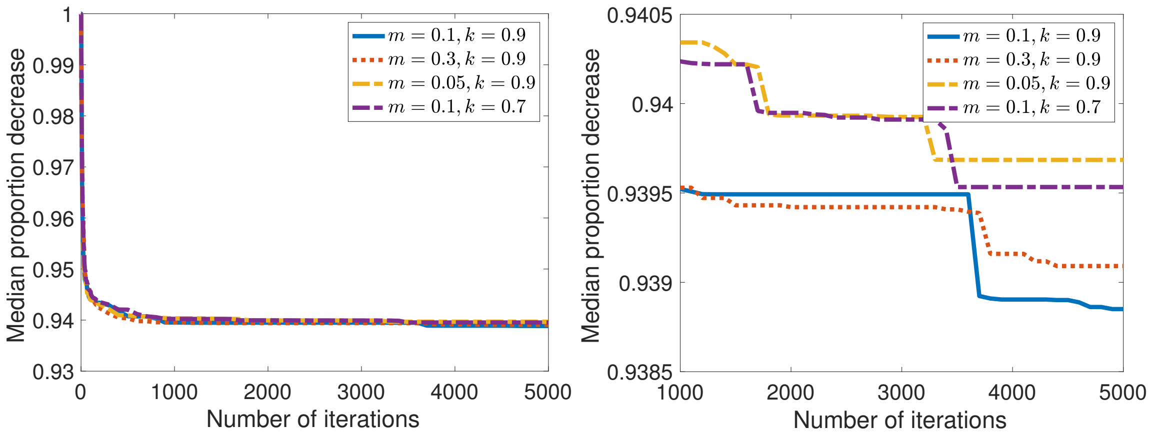
Figure 4.5: Proportion decreases in minimum for problems with two destination nodes for the (a) INSH and (b) GA algorithms. 400 networks were tested. The median proportion decrease is one.

than placing all resources on a single node in the two destination node problem.

These convergence results are for a single set of algorithm parameters. Perhaps the algorithms perform better under different parameters? The convergence of the INSH in networks with ten destination nodes for various parameters is shown in Figure 4.6. We test the convergence for different values of r and σ. Recall r is the number of solutions we choose to sample around at each iteration and σ is the variance of the normal distribution used to sample new solutions. Out of the four sets of parameters shown, $r = 5$ and $\sigma = 0.05$ provides convergence to a smaller minimum in 5000 iterations. However there isn't much difference between $\sigma = 0.05$ and $\sigma = 0.1$. Increasing σ to 0.2 increases the median proportion decrease.

The convergence of the GA in the same ten node destination networks is shown in Figure 4.7. We altered the probability of mutation m and crossover probability k. Overall, there isn't too much difference in solution between these parameter values. They all appear to have converged within 1000 iterations. If we look at the proportions between 1000 and 5000 iterations, the solution under $m = 0.1$ and $k = 0.9$ is the best. At 1000 iterations, it is approximately the same as that under $m = 0.3$ and $k = 0.9$, and better than the other two. It then decreases the most by 5000 iterations. Nonetheless, this difference is not large.

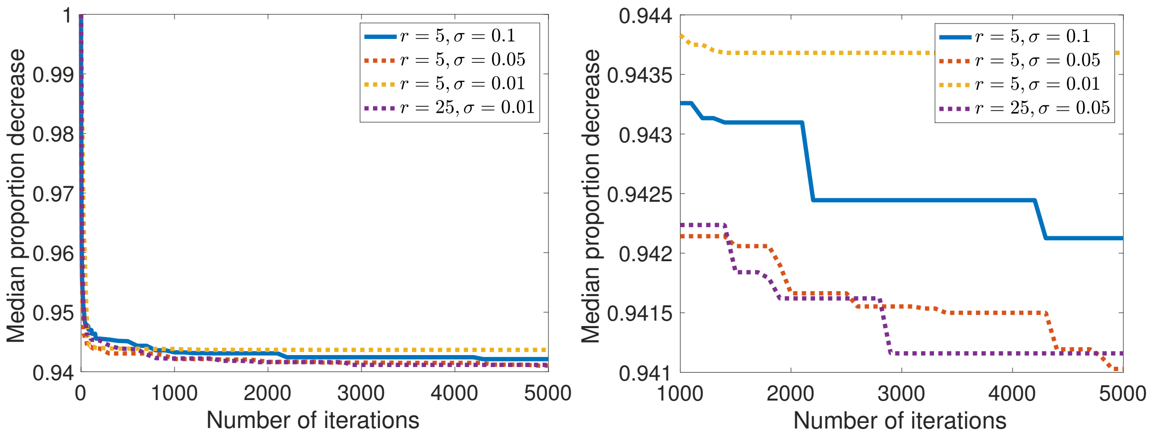
The analysis we have performed suggests that the GA converges to a better solution than the INSH. However we have not considered computational run-time. It may be the case that the GA reaches a better minimum but takes much longer to run. In the following section we will test this.



(a) Iterations between 0 and 5000.

(b) Iterations between 1000 and 5000.

Figure 4.7: GA convergence for difference mutation probabilities, m, and crossover probabilities, k. There are ten destination nodes. (a) The convergence when GA is run for a maximum of 5000 iterations. (b) A closer view of the convergence between 1000 and 5000 iterations. We see that the GA reaches the smallest median when $m = 0.1$ and $k = 0.9$.



(a) Iterations between 0 and 5000.

(b) Iterations between 1000 and 5000.

Figure 4.6: INSH convergence for different parameters. There are 50 nodes in the networks and ten of them are destination nodes. r is the number of solutions we chose to sample around at each iteration and σ is the variance of the normal distribution used to sample new solutions. (a) The convergence when INSH is run for a maximum of 5000 iterations. (b) A closer view of the convergence between 1000 and 5000 iterations. We see that the set $r = 5$ and $\sigma = 0.05$ reaches smallest median.

4.3.3 Run-Time

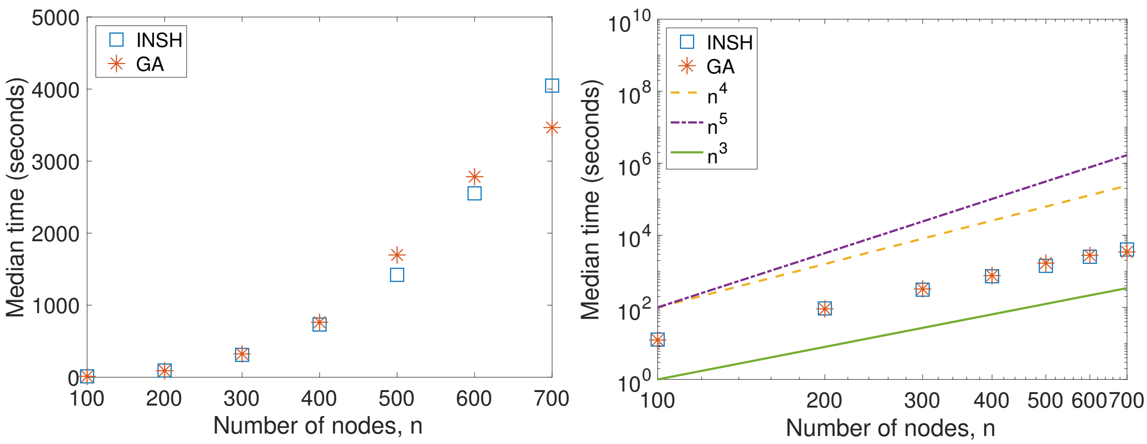
We perform some analysis on the computational time of the algorithms in this section.

We generate transport networks and detection probabilities in the same way as in the accuracy analysis. The GA and INSH with the same parameters are applied to solve the problem for each network. To see how the run-time scales as the network grows larger, we run each problem for 50 iterations for networks with different numbers of nodes. The GA and INSH times are shown in Figure 4.8 and boxplots in Figure 4.9. 100 networks of each size are tested.

Figure 4.8a shows the median INSH and GA times. The median times of networks of sizes 100, 200, and 300, are not much different between INSH and GA. However, there are some differences in median times for the larger sized networks. They appear to be similar on the log-log scale shown in Figure 4.8b. If we increase the number of nodes in the network past 700, we may see more a difference in the log-log times.

Using MATLAB profiling tools we see that the bottleneck of the code in a single iteration is the shortest path computation. We use Floyd-Warshall's algorithm [45] to compute shortest paths. Alternatively, Dijkstra's algorithm could be used. However we select Floyd-Warshall as we have access to efficient code of this algorithm. The complexity of Floyd-Warshall is $O(n^3)$.

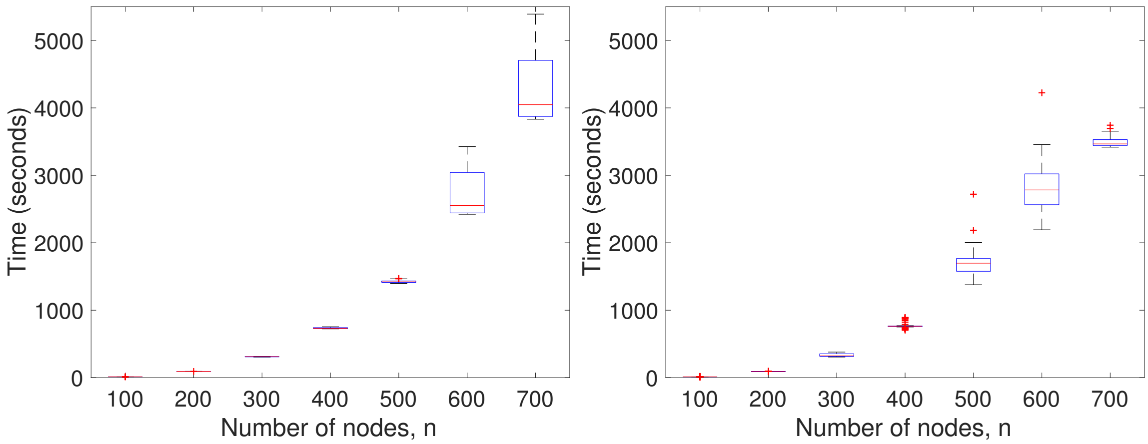
These results tell us about the run-time for a fixed number of iterations when we increase the number of nodes in the network. The number of destination nodes in the network remained fixed. As more destination nodes are added, the search space increases. So it might take longer for the algorithms to reach a good solution.



(a) Median times.

(b) Loglog plot of times.

Figure 4.8: (a) Median time and (b) log of the median time it takes the algorithms to run 50 iterations. The median was taken from 100 instances.



(a) INSH. (b) GA.

Figure 4.9: Boxplots of algorithm running times for 50 iterations of (a) INSH and (b) GA.

4.3.4 Conclusions

We tested the convergence of the algorithms for problems with different numbers of destination nodes. The GA converged to better solutions for problems with more destination nodes (ten and twenty). For the two and five destination node problems, there wasn't a large difference in convergence between the two algorithms. There also was not a major difference in run-time between the algorithms. Based on this, the GA performs better and we will use it to generate our results in the following chapter.

We need to choose the number of iterations of the GA to run. We can use the previous analysis on convergence to help us do this. The GA appeared to converge by roughly 1000 iterations for all of the sized problems we tested. The maximum number of destination nodes was twenty. The real world networks we will consider have five destination nodes. So, we will run the GA for 1000 iterations, which should be quite a conservative upper bound.

The population size will be set to $N = 50$. This is what it was for the convergence results. The best pair of mutation and crossover probabilities we tested was $m = 0.1$ and $k = 0.9.$, we will use these values.

One issue with setting the number of iterations for each new type of problem is that it may involve testing convergence every time. Ideally, some stopping rule should be implemented. A stopping rule would evaluate the current and past solutions at each iteration and determine if the current solution is "good enough". However we leave this approach for future work.

Chapter 5

Interdiction on a Real World Network: Vietnam Case Study

We have developed a dynamic game between smugglers and biosecurity forces in a theoretical setting. The problems in the previous chapter were constructed from randomly generated graphs in order to test our algorithms. In this chapter we construct a case study that uses real world networks based on flight routes and parameters based on fitting to data on pangolin seizures. This will investigate the interdiction model in a practical setting.

Wild pangolins are found in parts of Asia and Africa. There is a huge demand for them in Asian countries, in particular in China and Vietnam [10]. Vietnam was involved in a relatively high number of pangolin trafficking incidents between 2010 and 2015 [21]. Here we consider the scenario where smugglers are attempting to move pangolin commodities from African and Asian sources into Vietnam.

To construct a realistic case study we must incorporate real world data in the model. We need to identify potential pathways from African and Asian sources into Vietnam. Smugglers generally depend on air and sea transport to move illegal wildlife commodities between countries that are far away from one another [8, 49]. Here we will use flight data to create a network of potential pathways.

Another important aspect is the probability of detection at each node. We can analyse records of past pangolin seizure incidents to create estimates of the proportions of illegal flow that have gone through countries. We will also use these detection estimates to calibrate the detection probability function in terms of resource allocation.

In the next few sections, we discuss the data and methods used to construct the network and estimate the probability of detection. We then apply the model to the transport network and find the optimal resource allocation across Vietnamese cities.

5.1 Transport Network

Transport via air is a major reported method of smuggling illegal wildlife commodities [49]. Hence we construct a network of transport routes into Vietnam using flight data from OAG Aviation (www.oag.com). The data provides information on flights between 1997 and 2013. It is broken up into monthly spreadsheets. Each entry contains a record of either a direct flight or a route. These records contain the following information:

- carrier code, carrier name and corresponding flight number;
- departing and arriving airports;
- local departure and arrival times, and total elapsed travel time;
- number of seats on plane;
- monthly frequency of flight;
- route code;
- number of stops along route; and,
- whether the flight is international or domestic.

We consider international flights in the most recent year, 2013. Flight data before this year is not used, because we are using this data to produce a network of potential pathways, and are interested in interdiction in the present time. Thus we use flights only from the latest available year.

We will use this data to find paths from African and Asian sources into Vietnam. There are 17 Asian pangolin range states [4, 9, 11, 28]. These states are shown in Table 5.1. There are 31 African pangolin range states [37, 50, 51, 52], also shown in Table 5.1. We initially consider the airports (264 in total) that lie within these countries to be the source nodes. However, we remove the five airports in Vietnam as sources in the transport network, because we are interested in the smuggling of pangolins into Vietnam across ports of entry.

Paths with more hops will be less likely to be taken by smugglers. This is because there is an additional detection probability associated with the entry and exit of every node. One assumed objective of the smuggler is to pick the path that minimises the detection probability. Paths with more hops are also likely to have longer travel times and costs. The other smuggler objective is to minimise travel time. So by the construction of our model, the smugglers are more likely to choose paths with less hops. Thus we begin by considering all of the direct and two hop paths into

Vietnam from sources. As a result of this, we only consider sources within two hops of a Vietnamese airport. This will reduce the size of the transport network and allow us to compute optimal allocations faster. Consequently, 70 airports are removed as sources from the network as they are not connected to Vietnam within two hops. These are smaller airports that have fewer links and are not connected to transport hubs.

The other nodes in the network are transit and destination nodes. The transit nodes are the nodes that facilitate these 2-hop paths. The destination nodes are the airports in Vietnam that are within two hops of the sources: Ho Chi Minh City, Hanoi, Da Nang, Nha Trang and Can Tho. The edges connecting the network correspond to the flights along the 2-hop paths and direct edges between sources and destinations. The edges are weighted by the average travel time (minutes) from flights in 2013. There are 215 nodes in the network and 1283 edges.

There are 189 source nodes. So a high proportion of the nodes are also source nodes. This is because many of the nodes in Asia act as both a source and transit. Figure 5.1 shows all of the nodes in the transport network.

We aim to use this transport network in our model to find the optimal allocation across the Vietnamese destination nodes. Before we can do this we need to estimate the baseline probability of detection at each node in the network, as well as the parameters in the logistic detection probability function (3.2). In the following section we define estimates to achieve this.

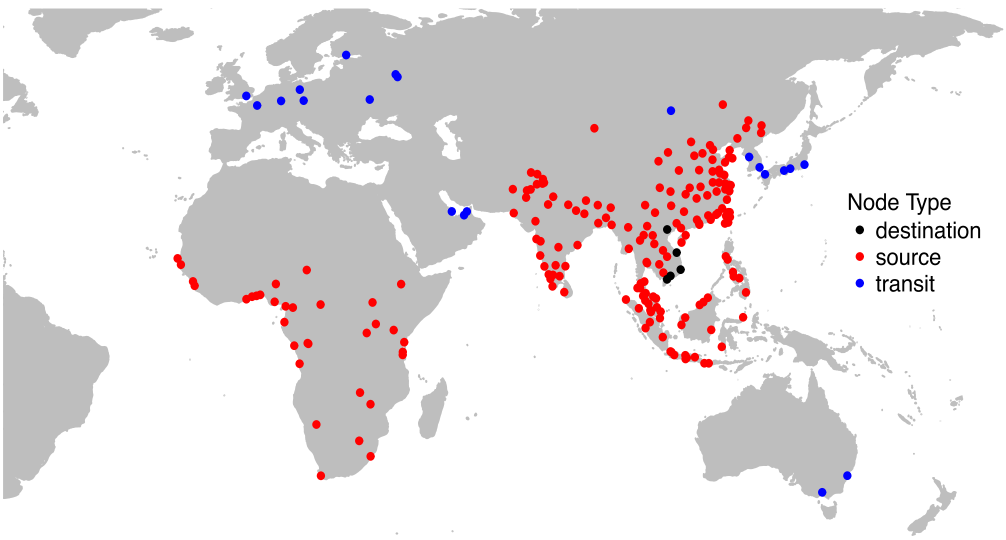


Figure 5.1: Nodes in the transport network. Sources are airports in specific Asian and African regions, and destinations are in Vietnam. Sources are connected to destinations by direct edges or via transit nodes (edges aren't displayed on map). Note that sources may also act as transit nodes.

	Range Country	Continent
1	Bangladesh	Asia
2	Nepal	Asia
3	India	Asia
4	Pakistan	Asia
5	Sri Lanka	Asia
6	Philippines (the)	Asia
7	Bhutan	Asia
8	Myanmar	Asia
9	Thailand	Asia
10	Lao People's Democratic Republic (the)	Asia
11	Vietnam	Asia
12	China	Asia
13	Taiwan (Province of China)	Asia
14	Hong Kong	Asia
15	Cambodia	Asia
16	Malaysia	Asia
17	Indonesia	Asia
18	Brunei Darussalam	Asia
19	Singapore	Asia
20	Cameroon	Africa
21	Senegal	Africa
22	Ghana	Africa
23	Uganda	Africa
24	Kenya	Africa
25	Tanzania, United Republic of	Africa
26	Angola	Africa
27	Cte d'Ivoire	Africa
28	Liberia	Africa
29	Sierra Leone	Africa
30	Guinea	Africa
31	Guinea-Bissau	Africa
32	Gambia (The)	Africa
33	Equatorial Guinea	Africa
34	Gabon	Africa
35	Congo	Africa
36	Congo (the Democratic Republic of the)	Africa
37	Rwanda	Africa
38	Burundi	Africa
39	South Sudan	Africa
40	Chad	Africa
41	Central African Republic (the)	Africa
42	Ethiopia	Africa
43	Namibia	Africa
44	Zambia	Africa
45	Zimbabwe	Africa
46	Botswana	Africa
47	Mozambique	Africa
48	Malawi	Africa
49	South Africa	Africa
50	Nigeria	Africa
51	Benin	Africa
52	Togo	Africa

Table 5.1: Pangolin range countries in Asia and Africa. Airports in these countries are the sources in the network.

5.2 Estimating the Probability of Detection

An important variable in the model is the probability of detection at each node and how this increases as resources are allocated to the node. The probabilities of detection at source and transit nodes are baseline fixed values. Whereas, the probability of detection at a destination node is a variable that depends on the resource allocation. In Chapter 3, we proposed a logistic function of the detection probability at a destination node in terms of resource allocation. We aim to fit the parameters of this function. We must first construct estimates of the probability of detection at different nodes.

Pangolin seizure data provides information on past trafficking routes and seizure locations. From this we can estimate the proportion of illegal trade that goes undetected through a country.

In this section, we first describe the pangolin seizure data and then define detection estimates and fit parameters of the logistic function.

5.2.1 Pangolin Seizure Data

Pangolin seizure data from [21] are used here to estimate node detection probabilities. The data contains records of international pangolin seizures between 2010 and 2015. The seizure incidents were collated from a range of sources. These sources included media reports, open data from CITES and publications from non-governmental organisations.

We only consider seizure records for incidents involving international source-destination routes. We remove domestic cases and incidents that do not contain information on the route, *i.e.,* data where only the seizure location is known. A total of 1200 cases are removed.

Information for all trafficking incidents analysed include:

- country seizure location;
- trade route information (origin, transit and destination countries);
- species (full scientific name or just genus *Manis* if only genus is known); and,
- incident month and year.

Information that was available for some incidents include:

- commodity type (13 categories, plus *other* and *unknown* groups);
- commodity count;

- commodity weight;
- estimated value;
- transport mode (air, land, sea, or unknown); and,
- details on suspects, *e.g.,* nationality, age, and name.

In total, we have 580 international illegal incidents. We will use the data on seizure location and trade route information.

5.2.2 Detection Probability Estimates

An estimate of the proportion of illegal trade that goes undetected through a country is found from this wildlife seizure data. Utermohlen and Baine [49] investigated wildlife trafficking in the air transport sector. They focussed on the trafficking of ivory, rhino horn, live reptiles and live birds. They defined an approximate enforcement success rate for various countries. The enforcement success rate, P^A, at country A was defined as

$$P^A = \frac{n^A}{n^A + d^A},$$

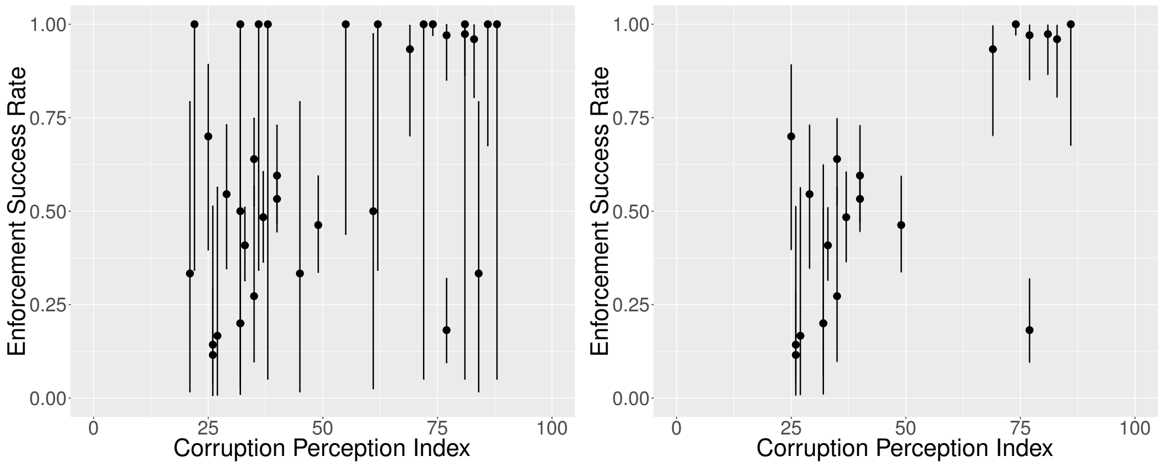
where n^A was the number of commodities that were seized at country A and d^A was the number of commodities that were seized at a later node on a route that passed through A. Note that there are biases in this measurement. For example, as $d^A \to 0$, $P^A \to 1$. If $P^A = 1$, this could either mean that node A has good enforcement, or countries further along the route have very low enforcement and so flow through A is not observed.

The variable d^A is an underestimate of the flow through A. So P^A is an overestimate of the true enforcement success rate. However it is still a useful means to understand the relationship between enforcement and detection.

We calculate the enforcement success rate for countries involved in pangolin trafficking incidents between 2010 and 2015. These success rate estimates are shown in Table 5.2. The USA and Switzerland both have enforcement success rates of one, because these countries only appeared as destination locations. So no commodities were observed to flow through them. However it is important to note that countries with high enforcement success rates are not necessarily only destination countries. Germany, Belgium and France all have high success rates but are mostly transit countries. For example, Germany seized commodities 33 times as a transit location and only five times as a destination location.

	Country	Success Rate	Incident Frequency	Seizure Frequency	CPI
1	Japan	1	1	1	72
2	Mali	1	1	1	32
3	Malta	1	3	3	55
4	Poland	1	2	2	62
5	Sri Lanka	1	2	2	36
6	Sweden	1	1	1	88
7	Switzerland	1	8	8	86
8	United Kingdom of Great Britain	1	1	1	81
9	United States of America	1	127	127	74
10	Zambia	1	1	1	38
11	Zimbabwe	1	2	2	22
12	Germany	0.97	38	37	81
13	Belgium	0.97	34	33	77
14	Netherlands	0.96	25	24	83
15	France	0.93	15	14	69
16	India	0.80	31	25	40
17	Uganda	0.78	9	7	25
18	Thailand	0.75	52	39	35
19	Nepal	0.67	18	12	29
20	Indonesia	0.6	50	30	37
21	China	0.56	221	122	40
22	Taiwan Province of China	0.5	2	1	61
23	Tanzania, United Republic of	0.5	2	1	32
24	Malaysia	0.48	52	25	49
25	Vietnam	0.44	87	38	33
26	Cambodia	0.33	3	1	21
27	Singapore	0.33	3	1	84
28	South Africa	0.33	3	1	45
29	Philippines	0.3	10	3	35
30	Pakistan	0.2	10	2	32
31	Togo	0.2	5	1	32
32	**Hong Kong**	**0.18**	**44**	**8**	**77**
33	Mozambique	0.17	6	1	27
34	Kenya	0.14	7	1	26
35	Cameroon	0.13	24	3	26

Table 5.2: Estimates of the pangolin trafficking enforcement success rate for countries. The estimates come from pangolin seizure data. CPI stands for corruption perception index. A higher CPI indicates lower levels of corruption. Hong Kong is highlighted because it has a relatively high CPI (77) but low enforcement success rate estimate (0.18).



(a) All countries. (b) Countries involved in at least 5 incidents.

Figure 5.2: Estimated enforcement success rate against corruption perception index (CPI) for countries involved in pangolin trafficking. Higher CPI values indicate lower corruption levels. (a) Estimates for all countries involved in trafficking incidents. (b) Estimates for countries involved in at least five international pangolin trafficking incidents across $2010 - 2015$.

We aim to find a model for the probability of detection as a function of resource allocation. Wildlife trafficking has been linked to corruption [57]. The *Corruption Perception Index* (CPI) is a value between 0 and 100 assigned to each country, yearly, that estimates the amount of perceived corruption present in the country (www.transparency.org/research/cpi/overview). A higher CPI indicates lower corruption levels.

Figure 5.2a shows a plot of the enforcement success rates against the CPI values for countries involved in pangolin trafficking incidents. We use the 2016 CPI values (www.transparency.org/research/cpi/overview). These were the latest values at the time of research and are shown in Table 5.2. However the CPI values don't appear to change very much from year to year. The median difference between 2016 and 2017 CPI values was one. The figure also shows 95% confidence intervals for these binomial estimates. Some of these intervals are wide due to countries being involved in only a few incidents, and so Figure 5.2b shows confidence intervals for countries involved in at least five pangolin trafficking incidents. The correlation coefficient of these two variables, when only countries involved in at least five incidents are included, is 0.726, indicating a relationship between the two quantities.

Given that we have observed a relationship between CPI and detection, we will use CPI as a proxy for resource allocation. A physical interpretation for this choice

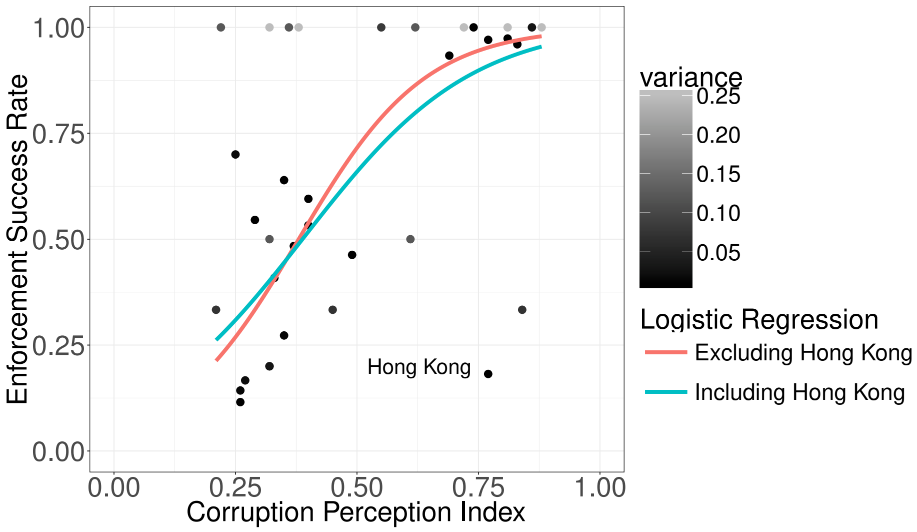


Figure 5.3: Logistic curves fitted to enforcement estimates including and excluding Hong Kong. The weights were the reciprocal of the variance. Nodes with lower variance are given a higher weighting.

could be countries with less perceived corruption (higher CPI), have more trained biosecurity agents placed at ports (more resources).

There are some enforcement success rate estimates that are close to or equal to one. These indicate that the countries intercept almost all illegal trade which is unrealistic. This is due the estimates being overestimates, particularly for countries that are primarily destinations. However, the current estimates give us information on relative differences. This is what we need to determine path choices relative to resource allocations. Still, scaling of the probabilities is required for more accurate estimates.

The points in Figure 5.2b appear to have a logistic shape. Recall we mentioned in Chapter 3 that the logistic function had the minimal required properties of a detection probability function. We now aim to calibrate the logistic function on these enforcement success rate estimates.

The model is fit to the data using weighted regression [15, 40] and the fit is shown in Figure 5.3. Weighted regression is used as the countries were involved in different numbers of incidents. The weight is the reciprocal of the variance of each binomial sample proportion, *i.e.*, the variance for country i is $\hat{\sigma}_i^{\,2} = p_i(1-p_i)/n_i$ where p_i is the success rate estimate and n_i is the number of incidents observed to flow through the country (including seizures). Each CPI is divided by 100 so that it is between 0 and 1. This is required as the resources assigned to each node are in this range. The logistic function is also fit under the constraint that it must be between 0 and 1.

Hong Kong shows some odd behaviour. It appears to be an outlier as it sits far below the line of best fit. Hong Kong has a relatively heavy weighting in the model. It was involved in 44 trafficking incidents. It has a high CPI (77) but low enforcement success rate estimate (0.18). Hong Kong has also recently introduced stricter laws around wildlife crime [6]. Earlier in 2018, they increased the maximum penalty of trafficking CITES Appendix I listed animals. The maximum imprisonment was increased from two years to ten years.

We consider Hong Kong to be an outlier and remove it from the data that was used in the regression. The resulting fit is shown in Figure 5.3. The overall fit is improved when we remove Hong Kong. The residual standard error when Hong Kong is included is 2.721 on 32 *df*, and 1.51 on 31 *df* when removed. This final curve is

$$p_i = \frac{1}{1 + \exp\left(\frac{0.38008 - \mathrm{CPI}_i}{0.13}\right)}, \tag{5.1}$$

and is used to estimate the baseline detection probabilities at source and transit nodes. The function is also used to predict the probability of detection at a node given a resource allocation δ_i between 0 and 1. This function is

$$p(\delta_i) = \frac{1}{1 + \exp\left(\frac{0.38008 - \delta_i}{0.13}\right)}. \tag{5.2}$$

In the next section we apply the interdiction model, with these detection functions, to real world networks.

5.3 Case Study Results

Recall we have defined a scenario where smugglers are moving pangolin commodities from African and Asian sources into Vietnam. We assume that smugglers pick the route that comes from some combination of minimising travel time and maximising the probability of non-detection. They move commodities on the underlying transport network defined in Section 5.1. We aim to find the optimal resource allocation across Vietnamese destination nodes. We will use the Genetic Algorithm defined in Chapter 4 to find this optimal allocation.

We first consider the *African* and *Asian source networks* separately to observe differences. The *African source network* has sources only in Africa and the *Asian source network* has sources only in Asia.

We first assume uniform traffic from all sources. This corresponds to setting every source weight to one, $w_i = 1$ for all sources i. Realistically, the same amount of traffic will not come from every source. We consider non-uniform traffic in Section 5.4.1.

The probability of detection function is the fitted logistic curve from Section 5.2. The detection probability at a node where all resources have been placed ($\delta_i = 1$) is 0.9916. This is an unrealistically high detection probability for a single node. We investigate reducing these probabilities in Section 5.3.2. We are also allowing the possibility of no biosecurity forces to be placed at destination nodes ($\delta_i = 0$). However, there is still a positive detection probability, 0.0510, at a destination node with no resources. It is unlikely that no security resources would be placed at an airport. We can think of this problem as allocating extra resources on top of baseline security resources.

We set the total budget of resources to one. That is,

$$\sum_{i \in D} \delta_i = 1$$

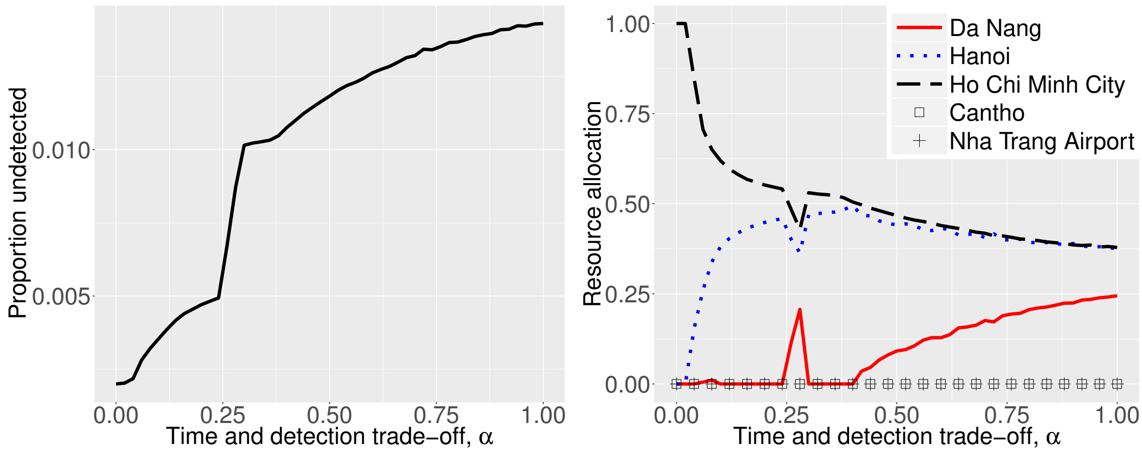
The allocation δ_i can be thought of as the proportion of resources assigned to node i.

Another parameter of the model is α. This is the smuggler's trade-off factor between minimising travel time and maximising the probability of non-detection. We do not know the value of this parameter. So we will find the optimal resource allocation for α increasing from 0 to 1 in 0.02 increments.

To find the optimal solutions we use the GA as described in Chapter 4. This is because the GA appeared to converge to better solutions than INSH. We set the GA crossover probability to 0.9, mutation probability to 0.1, and the population size to 50. We fix the number of iterations of the algorithm to 1000. Our case study has five destination nodes. In our simulations, the GA appeared to converge to a solution within 500 iterations under five destination node problems, so 1000 iterations is a conservative choice to ensure convergence.

Figure 5.4 shows the proportion of illegal traffic that goes undetected when the optimal resource allocation is applied to the African source network. As α increases, the proportion of illegal traffic that goes undetected also increases. This makes sense as α indicates the weight the smuggler applies to avoiding detection. So as α increases the proportion undetected will be greater due to the smuggler's efforts. The proportion undetected across the network is very small. This is the case because the logistic function is close to one for moderate values of δ. So it is possible to intercept a high proportion of flow through a node.

Figure 5.4 shows the optimal resource allocation in the African source network. Resources are placed across three nodes. These are airports in Ho Chi Minh City, Hanoi and Da Nang. No other Vietnamese airports could be reached within two hops from African sources. So these receive no allocation of resources. When $\alpha = 0$, the smuggler's objective is to minimise travel time and the optimal strategy is to place all resources at Ho Chi Minh City. As α increases from 0 to 0.3, the general trend is to place increasingly more resources at Hanoi and less at Ho Chi Minh City. Hanoi has



(a) Optimal proportion undetected.

(b) Optimal resource allocation.

Figure 5.4: The (a) proportion of illegal traffic that goes undetected for the African source network under the (b) optimal allocation. The greater α, the more the smuggler weights avoiding detection as opposed to time. In these cases, a higher proportion goes undetected.

the second highest number of connections, making it easily accessible to smugglers as they care more about detection. There is an irregularity around $\alpha = 0.25$. It appears that suddenly more resources are required at Da Nang. However, we will see shortly that the proportion of undetected trade remains unaffected when no resources are placed at Da Nang. Finally, as α increases above 0.4, the optimal allocation is to place continuously more resources at Da Nang and less at the other two.

In order to understand the results we also consider how the source-destination paths from Africa vary. Figure 5.5 shows the transit country frequencies for the African source network. Across α there isn't very much change in transit frequency. This is because there is a limited number of pathways going from Africa to Vietnam. However there are some slight changes. For smuggler objective weights around $\alpha = 0.25$, there is a very small decrease in the number of trades passing through Dubai and an increase in the number passing through Beijing. There is also a jump in the proportion undetected around this mark.

As well as observing how the use of transits change with α, the locations themselves are interesting. Six out of the eight transit locations were listed in the top 30 busiest airports (by passenger volume in 2015) [1]. These airports were Suvarnabhumi Airport, Dubai International Airport, Frankfurt Airport, Paris-Charles de Gaulle Airport, Singapore Changi Airport, and Beijing Capital International Airport. Given these airports have more connections, they are more likely to lie along

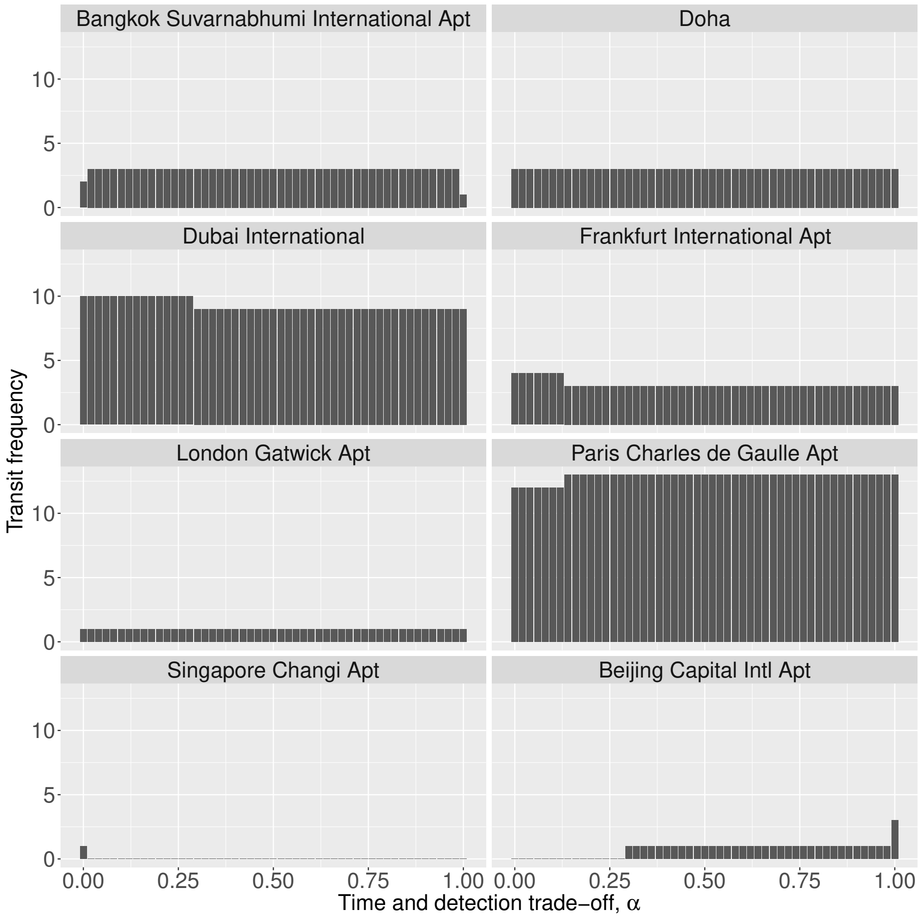
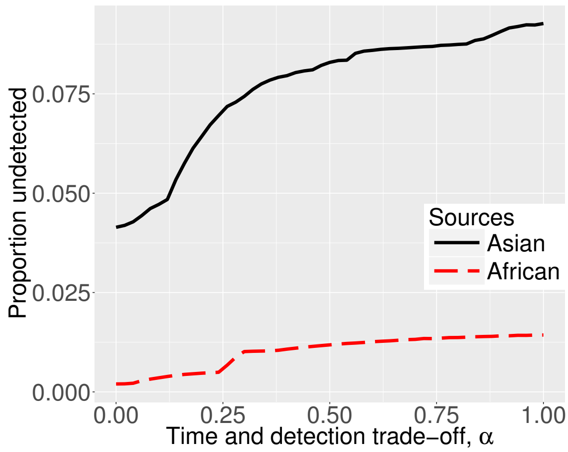


Figure 5.5: The transit frequency of countries for increasing α with African sources. An α closer to one indicates a higher weighting on maximising probability of non-detection in the smuggler's objective.

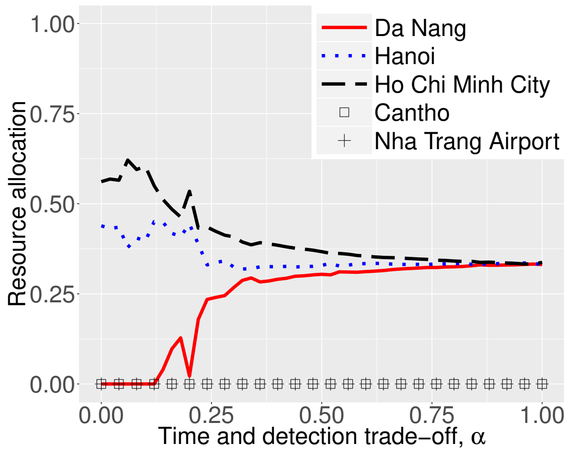
optimal paths.

We have seen various allocations across α. What happens if we use non-optimal resource allocations? In particular, it will be interesting to look at the region containing the irregularity around $\alpha = 0.25$. In Section 5.3.1 we apply non-optimal allocations and look at the corresponding proportions of undetected illegal traffic. Also, we do not know the true value of the weight, α. So we will aim to find a robust allocation across α.

We now consider the Asian source network. The optimal resource allocations are shown in Figure 5.6. We see that the optimal strategy is different to that of the African source network. For $\alpha > 0.5$, the optimal allocation is roughly to spread all resources evenly across three airports: Ho Chi Minh City, Hanoi, and Da Nang.



(a) Optimal proportion undetected.



(b) Optimal resource allocation.

Figure 5.6: The (a) proportion of illegal traffic that goes undetected for the Asian source network under the (b) optimal resource allocation. The optimal allocations at Can Tho and Nha Trang are always zero. This is because they have fewer connections. The optimal allocation appears to be to spread resources more evenly across Da Nang, Hanoi, and Ho Chi Minh City as α increases and the smugglers avoid detection more.

When $\alpha = 0$, the optimal solution is to place resources at two airports (Ho Chi Minh City and Hanoi) rather than only at Ho Chi Minh City which was the strategy for the African source network. This is because both Hanoi and Ho Chi Minh City are closest to the Asian sources in the transport network.

The proportion that goes undetected in the Asian source network is also shown in Figure 5.6. A slightly higher proportion goes undetected than in the African source network. The median path detection probability for the African source network is 0.0021. The median for the Asian source network is 0.0715. So, the probability of detection along an optimal path from Africa into Vietnam is more likely to be larger as all of the paths have intermediate stopovers. There are more chances that the smuggler is caught at these stopovers.

The transit frequencies along paths from Asian sources were also found and are shown in Figure 5.7. These frequencies vary a lot more across α than the ones in the African source network. For α close to zero, a high proportion of paths (40/161) pass through Hong Kong. As α increases, the number of trades that pass through Hong Kong decreases. This is because Hong Kong has a relatively high fitted probability of detection (0.9525). The smugglers pick alternative paths that do not pass through Hong Kong as they become more sensitive to maximising the probability of non-detection. Recall that Hong Kong was an outlier in the model fitted in Section 5.2. It

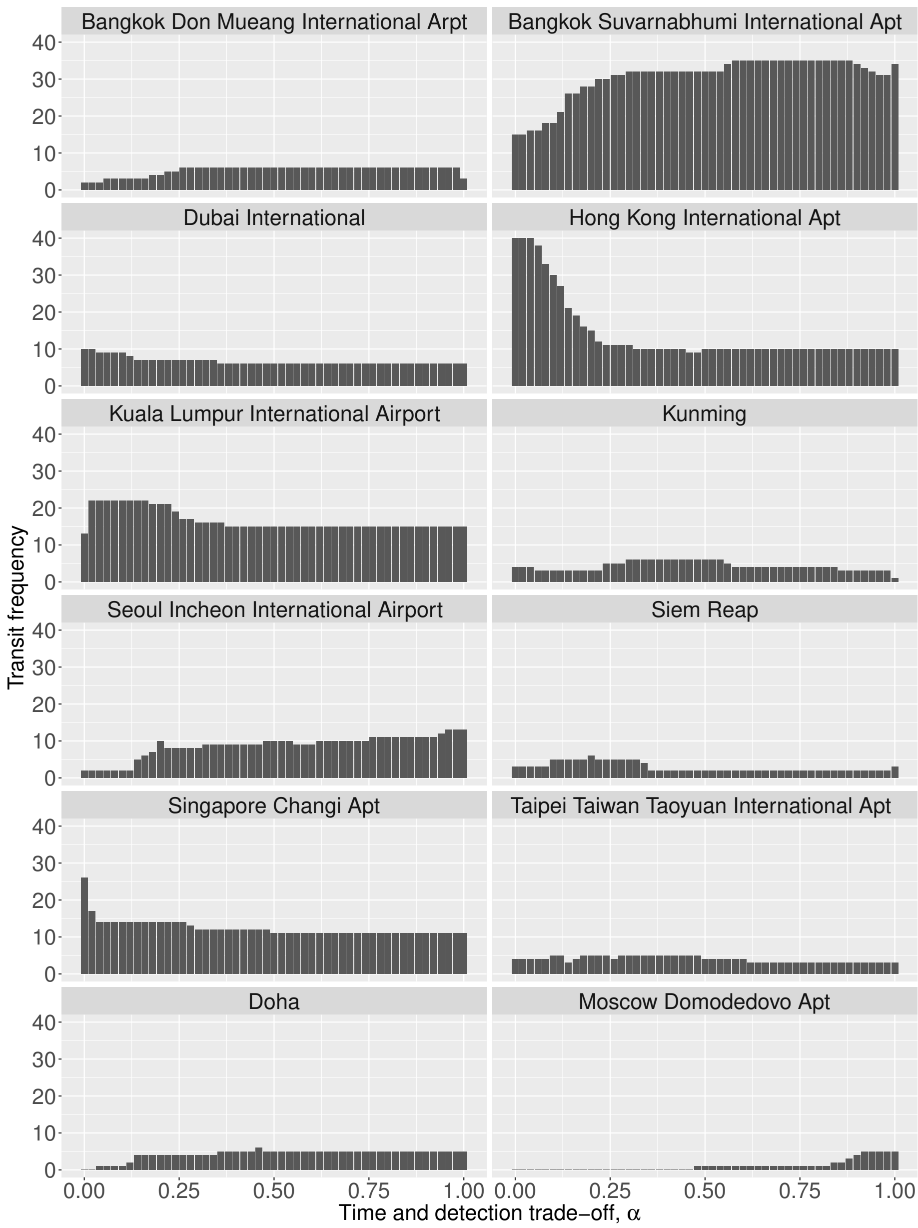
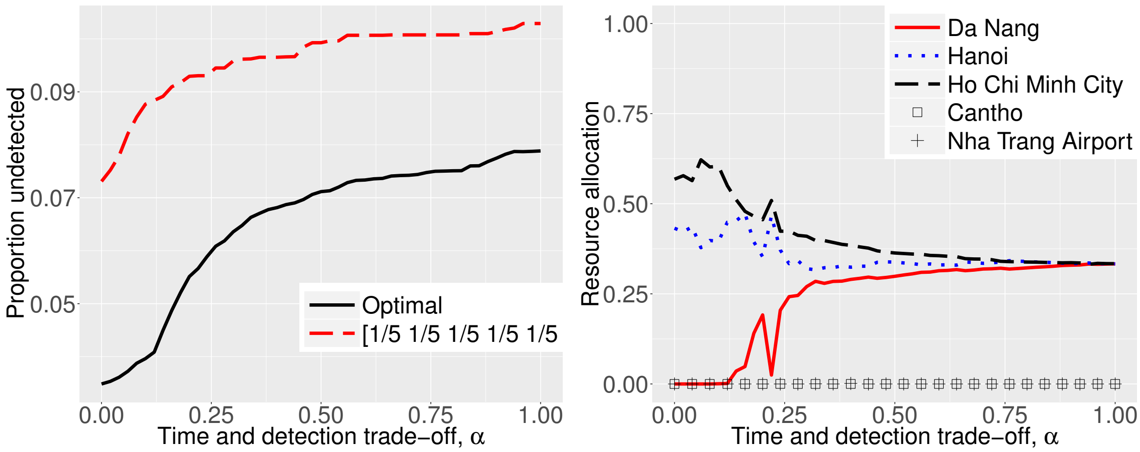


Figure 5.7: The transit frequency of countries for increasing α with Asian sources. Only airports that had frequencies greater than four in at least one α are shown. In total 30 airports were used as transit nodes.



(a) Optimal proportion undetected (b) Optimal resource allocation

Figure 5.8: The (a) proportion of undetected illegal trade for the combined Asian and African source network and the (b) optimal allocation. It is very similar to the Asian source resource allocation. This is because the detection probabilities along paths from Asia to Vietnam are more likely to be lower than that along African source paths.

had a high CPI but low enforcement success rate estimate. However its fitted value affects these results quite a lot. In Section 5.3.1 we investigate the effect of changing Hong Kong's detection probability on the optimal allocation.

Half of these transits were listed in the top 30 list of busiest airports (by passenger volume in 2015) [1]. These airports were Suvarnabhumi Airport, Dubai International Airport, Hong Kong International Airport, Kuala Lumpur International Airport, Seoul Incheon International Airport, and Singapore Changi Airport. We also see that these airports are used more frequently as stopovers than the other locations. This is likely as they have more connections between sources and destinations.

Finally, we consider the combined network of Asian and African sources. This optimal resource allocation is very similar to the one for just the Asian sources in Figure 5.6. When we implement the optimal Asian source allocation on the combined source network, the proportion of undetected trade is shown in Figure 5.9. The two proportions appear almost the same for every value of α, but with a slight difference. One reason why the combined network resource allocation looks like the Asian source allocation is that there are more Asian sources. There are approximately five times as many Asian sources, with 33 African source nodes compared to 161 Asian ones.

The true weights of the sources are unknown. To explore the impact of different weightings, we put different relative weights on the African and Asian sources.

There would be approximately equal traffic into Vietnam from Africa and Asia if the weighting of the African sources was $w_i = 5$. The optimal resource allocation for the problem with a weighting of five on the African sources is shown in Figure 5.10. The optimal resource allocation still appears to match the Asian allocation. This is because the probability of detection along paths from Africa into Vietnam is likely to be higher. So a greater increase in weighting is required for a resource allocation to look more similar to the optimal solution of the African source network. Figure 5.10 shows the optimal allocations for increasing weight assigned to African sources. We see that Da Nang is assigned less resources as the weightings are increased. This is more similar to the African source allocation.

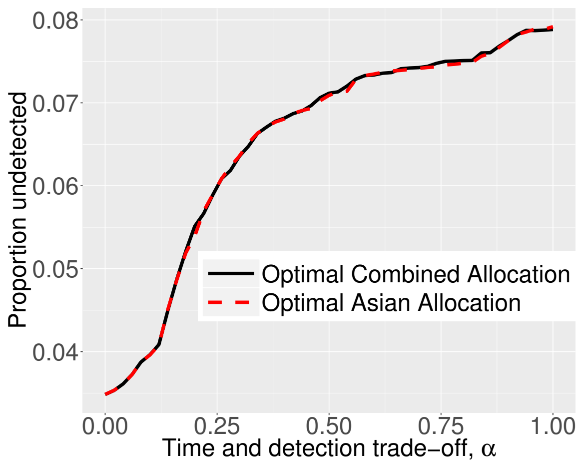
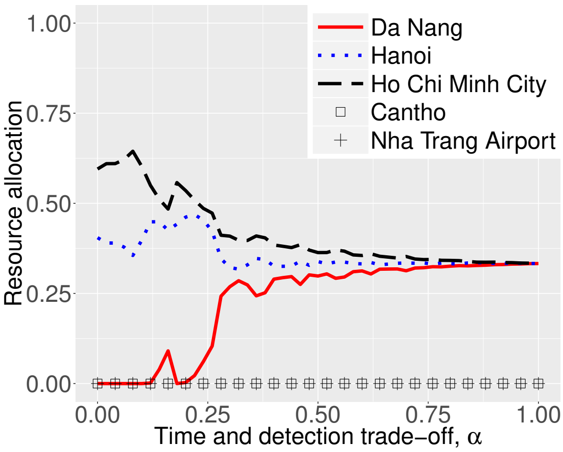
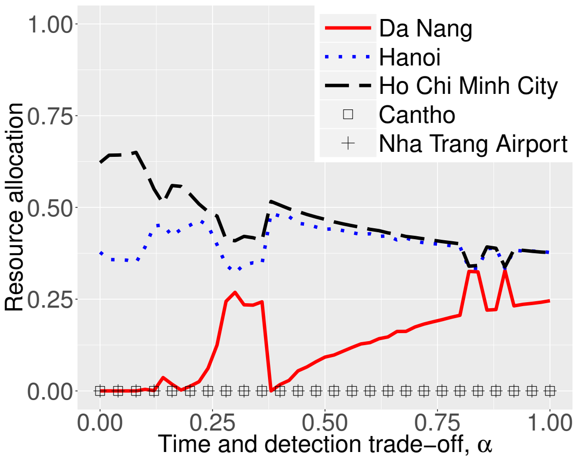


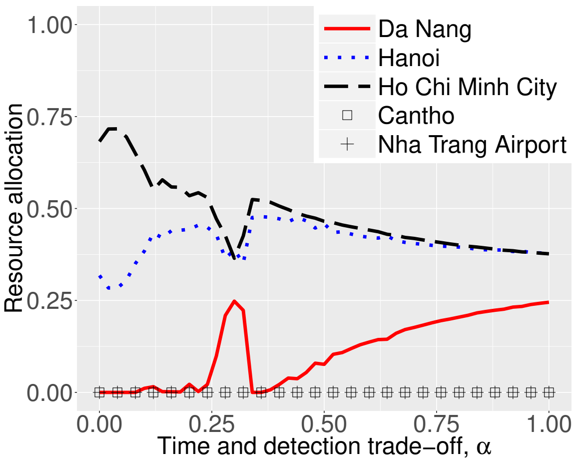
Figure 5.9: The proportion of illegal traffic that goes undetected when the optimal resource allocation for the Asian source network is applied to the whole network with both Asian and African sources. The proportions are very close to one another but not exactly the same.



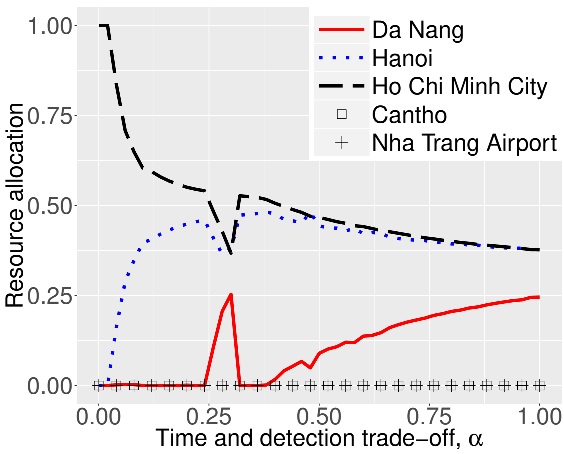
(a) Weighting of $w_i = 5$.



(b) Weighting of $w_i = 10$.



(c) Weighting of $w_i = 25$.



(d) Weighting of $w_i = 200$.

Figure 5.10: The optimal allocation for different weightings of African sources in the combined network. The optimal allocation is to place less resources at Da Nang for heavier weightings of the African sources. This approaches the optimal African source network allocation.

5.3.1 Sensitivity of Resource Allocations

In the previous section we looked at optimal allocations for various values of α, the weight controlling the trade-off between the smugglers two objectives. However we rarely know whether the smuggler cares more about minimising travel time or maximising the probability of non-detection Moreover, not all smugglers will have the same objective. Given this lack of knowledge, we would not know what allocation to implement. But perhaps there is a resource allocation that performs well across a range of α. In this section we investigate the sensitivity of the resource allocations to α. That is, we consider the loss of optimality resulting from a difference in the smugglers true α than the value we assume.

We consider the proportion of illegal traffic that goes undetected for four different allocations. We will represent each allocation as a vector,

$$(a_1, a_2, a_3, a_4, a_5),$$

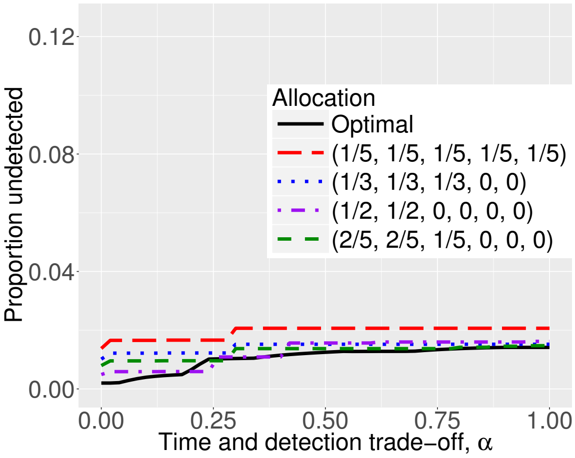
where a_1, a_2, a_3, a_4, and a_5 are the allocations at Ho Chi Minh City, Hanoi, Da Nang, Nha Trang, and Can Tho respectively.

The first allocation is to spread resources evenly across Ho Chi Minh City, Hanoi, and Da Nang, $(1/3, 1/3, 1/3, 0, 0)$. This was the optimal allocation on the Asian and combined source networks when the smuggler avoided detection ($\alpha = 1$). The second is to allocate resources evenly between Ho Chi Minh City and Hanoi, $(1/2, 1/2, 0, 0, 0)$. These two airports were always allocated more resources in the optimal solutions. The third allocation is to assign most of the resources to Ho Chi Minh City and Hanoi, and some at Da Nang, $(2/5, 2/5, 1/5, 0, 0)$. The final allocation is to assign 1/5 of the resources to every destination node, $(1/5, 1/5, 1/5, 1/5, 1/5)$. We look at how this allocation performs against the optimal as it is a naive approach.

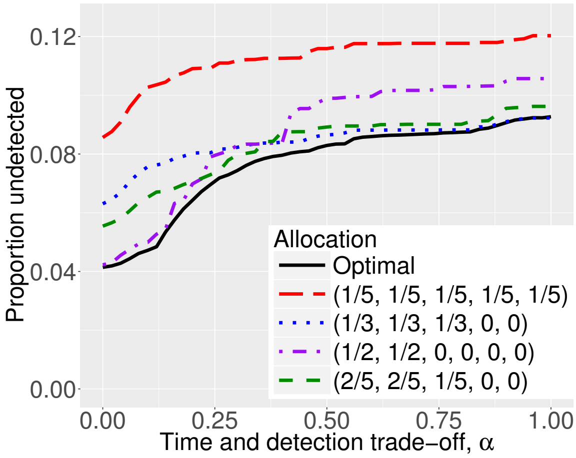
We set the resource allocations to the four just described and calculate the proportion of illegal trade that goes undetected in both Asian and African source networks for α increasing from 0 to 1 by 0.02 increments. The results are shown in Figure 5.11. The proportion undetected is not very sensitive to changes in the resource allocation. The allocation $(2/5, 2/5, 1/5, 0, 0)$ appears closest to the optimal in the African source network. Overall there is not much change in the proportion undetected in the African network.

Now consider the Asian source network sensitivity shown in Figure 5.11. The resource allocations affect the proportion undetected more in this network than the African source allocation. The allocation $(1/3, 1/3, 1/3, 0, 0)$ appears to match the optimal closely for $\alpha > 0.4$. The strategy $(1/2, 1/2, 0, 0, 0)$ appears to match the optimal better than the previous allocation for $\alpha < 0.4$. Overall, the allocation $(2/5, 2/5, 1/5, 0, 0)$ appears best across the full range of α.

The sensitivity of the combined network is shown in Figure 5.12. Once again, the



(a) African Sources.



(b) Asian Sources.

Figure 5.11: Proportion of undetected illegal trade for fixed and optimal allocations for the (a) African source network and the (b) Asian source network. The allocation $(2/5, 2/5, 1/5, 0, 0)$ is closest to the optimal in the African source network. However, there is little difference between the allocations. This allocation is also best on the Asian source network across α.

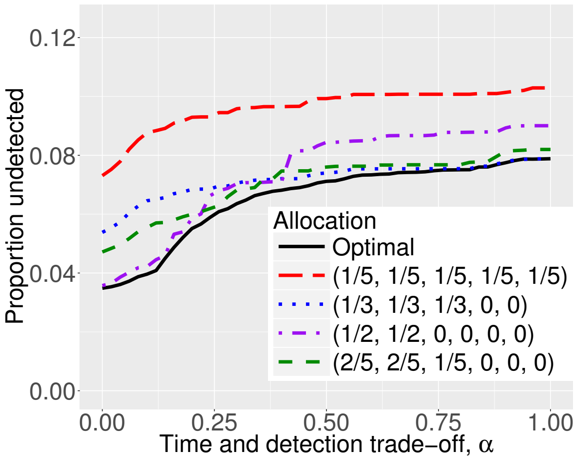


Figure 5.12: Proportion of undetected illegal trade for fixed and optimal allocations for the combined network. These proportions are very similar to the Asian source ones in Figure 5.11. The allocation $(1/2, 1/2, 0, 0, 0)$ appears close to the optimal in the Asian source network for $\alpha < 0.25$. For $\alpha > 0.4$, the allocation $(1/3, 1/3, 1/3, 0, 0)$ appears better. The best allocation across all α is $(2/5, 2/5, 1/5, 0, 0)$.

Resource Allocation	α	Average Difference in Proportion Undetected		
		African	Asian	Combined
(1/3, 1/3, 1/3, 0, 0)	$0 \leq \alpha \leq 1$	0.00367	0.00756	0.00655
(1/2, 1/2, 0, 0, 0)	$0 \leq \alpha \leq 1$	0.00224	0.0106	0.00873
(1/5, 1/5, 1/5, 1/5, 1/5)	$0 \leq \alpha \leq 1$	0.00879	0.0359	0.0309
(2/5, 2/5, 1/5, 0, 0)	$0 \leq \alpha \leq 1$	0.00209	0.00643	0.00534
(1/2, 1/2, 0, 0, 0) $(1/3, 1/3, 1/3, 0, 0)$	$0 \leq \alpha \leq 0.4$ $0.4 < \alpha \leq 1$	0.00185	0.00294	0.00235

Table 5.3: The average difference in proportion of illegal traffic that goes undetected between the optimal and fixed allocation. We found this average on three different networks: the African source network, the Asian source network, and the combined African and Asian source network. The best allocation across all α for the Asian and Combined networks is $(2/5, 2/5, 1/5, 0, 0)$. The best for the African network is (1/2, 1/2, 0, 0, 0).

sensitivity looks quite similar to the Asian source network. The African sources have little effect on the results.

We now calculate the average difference between the proportion undetected under both the optimal and specified resource allocations. This average difference is the sum of absolute differences between the optimal and specified resource allocation proportions,

$$d(\boldsymbol{\delta}) = \frac{1}{51} \sum_{i=0}^{50} \left| f_{\alpha_i}(\boldsymbol{\delta}^*) - f_{\alpha_i}(\boldsymbol{\delta}) \right| ,$$

where $\alpha_i = i/50$, $\boldsymbol{\delta}^*$ is the optimal strategy given α_i, $\boldsymbol{\delta}$ is the specified strategy and f is the proportion undetected given α_i and $\boldsymbol{\delta}$. These differences are shown in Table 5.3.

We could implement a rule based on these measurements if we had a range that α was likely to be contained in. For all networks, the single best allocation was to place a majority of the resources across Ho Chi Minh City and Hanoi, and a small fraction at Da Nang, $(2/5, 2/5, 1/5, 0, 0)$. However we also considered a mix of allocations across α. The best mix we looked at was to place resources across Ho Chi Minh City and Hanoi for $\alpha \leq 0.4$, and to place resources across Ho Chi Minh City, Hanoi and Da Nang for $\alpha > 0.4$. These allocations are useful if α lies close to 0 or 1. However as α moves closer to 0.4 and we become more uncertain about where it lies, it is harder to determine which resource allocation is better.

However, overall we have found some allocations that are fairly robust. The most robust allocation across all α, in the combined network, was to place 2/5 of resources at Ho Chi Minh City and Hanoi, and 1/5 at Da Nang. Though, the allocation of spreading resources evenly amongst these three locations was also reasonably robust. Ho Chi Minh City, Hanoi, and Da Nang have more 2-hop connections to the sources than Can Tho and Nha Trang. Hence, spreading resources across these airports makes sense. This type of allocation, spreading resources evenly amongst ports of entry with more connections, may generalise to other transport networks as well.

Probability of Detection at Hong Kong

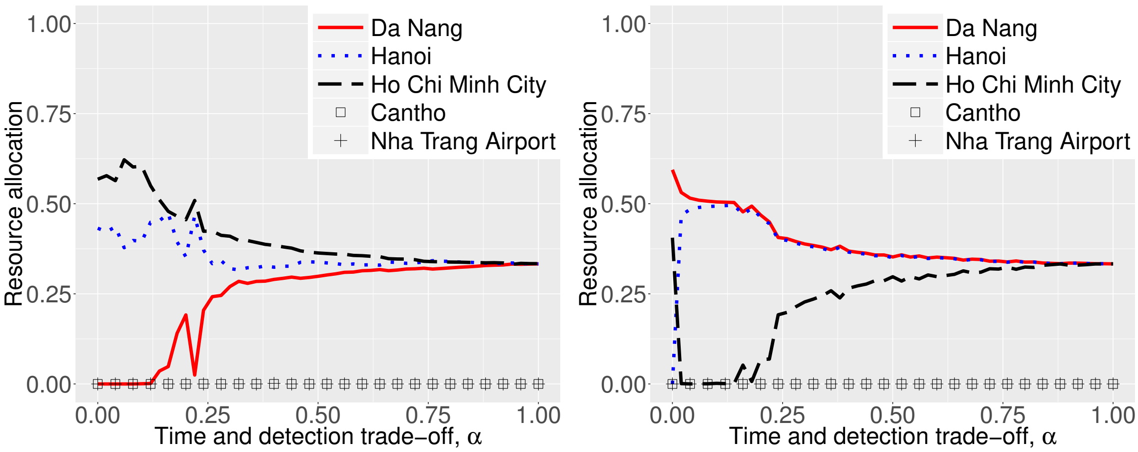
There is some uncertainty in the node detection probability estimates. One source of uncertainty comes from illegal traffic that we did not observe. Also, some countries have only been involved in a few incidents. We accounted for some of this by weighting each estimate by the reciprocal of its variance in our regression. We also have bias in the estimates that arises from the roles each country play in the trafficking incidents. For example, countries that mainly import pangolin commodities will have positively-biased enforcement success rates. Given the uncertainty, we are interested to know whether changes in transit detection probabilities from our predictions result in different optimal allocations.

Detection at Hong Kong is of particular interest. In the previous section we saw that Hong Kong was a frequently used stopover from Asian sources into Vietnam. For instance when the smuggler travelled to minimise time, 1/4 of the source to destination paths went though Hong Kong. As the smuggler weighted avoiding detection more, the frequency of Hong Kong as a stopover decreased. This is because Hong Kong has a high fitted probability of detection (0.9525). However recall that Hong Kong was treated as an outlier when we fitted the probability of detection curve in Section 5.2. Its raw value was quite low (0.18).

In this section we find the optimal allocation again using Hong Kong's raw value instead. We are interested to see if the optimal allocations change greatly given Hong Kong played a large role but its detection probability was far from the observed value. This will allow us to test the sensitivity of the proportion of undetected trade and optimal allocation under a change in transit detection probability. This is important as there is uncertainty in the node detection probabilities.

The optimal resource allocations for each version of Hong Kong's probability of detection is shown in Figure 5.13. We see that the resource allocations are quite different. Ho Chi Minh City has less resources placed on it for all α when the Hong Kong detection probability is its raw value. For $\alpha = 0$, you would expect the two resource allocations to be the same as the smuggler only cares about minimising travel time. But due to the stochasticity of the algorithms used to find the solutions, these allocations are different.

Figure 5.14 shows the proportions of illegal flow that go undetected for a few



(a) Hong Kong Detection Probability is Fitted Value.

(b) Hong Kong Detection Probability is Raw Value.

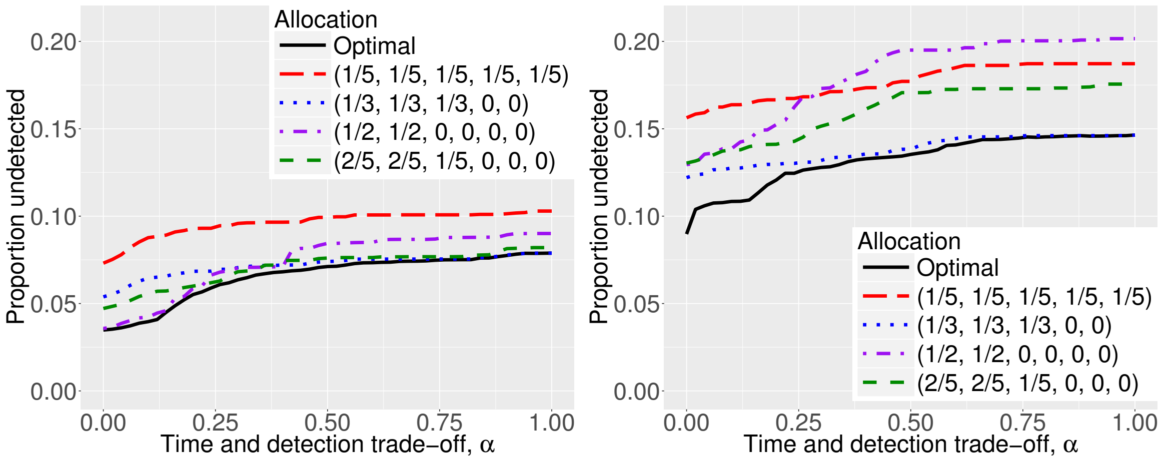
Figure 5.13: Resource allocations at Vietnamese airports when the detection probability at Hong Kong is (a) its fitted value and (b) its raw value. The fitted value is 0.9525 . The raw detection probability is 0.18. The sources are Asian cities. The two sets of allocations are quite different. Da Nang is assigned more resources when the detection is less at Hong Kong.

resource allocations for both probabilities of detection at Hong Kong. We see that more flow goes undetected when the optimal strategy is in play for the Hong Kong raw detection probability then the fitted Hong Kong detection probability. This is because the raw Hong Kong detection probability is much less than the fitted value.

In Figure 5.15 we see that the frequency of Hong Kong as a transit country increases as α increases. This contrasts the behaviour of the system under the fitted Hong Kong detection probability where Hong Kong was used less as a transit when α increased.

The sensitivity of the resource allocations are also shown in Figure 5.14. The resource allocation (1/3, 1/3, 1/3, 0, 0) fits close to the optimal for $\alpha > 0.25$ similar to how it fits the optimal proportion undetected curve for the Hong Kong fitted value. The allocation $(2/5, 2/5, 1/5, 0, 0)$ does not perform as well. As we do not know the true detection probability at Hong Kong, the allocation (1/3, 1/3, 1/3, 0, 0) is far more robust than the others.

So we have seen that greatly changing the detection probability of a frequently used stopover does affect the optimal resource allocation. This is not surprising. However, the allocation of spreading resources evenly across three nodes, Ho Chi Minh City, Hanoi, and Da Nang, remains robust in this scenario. Previously, the



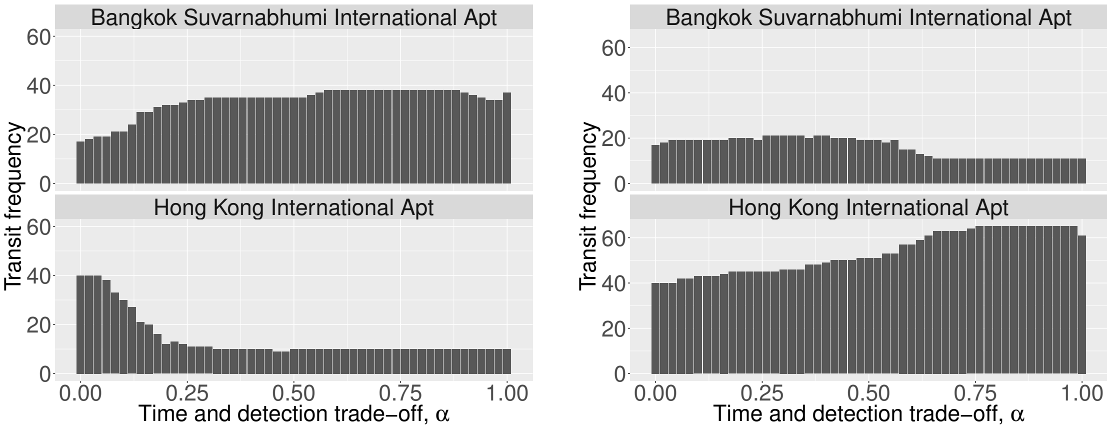
(a) HK fitted detection probability

(b) HK raw detection probability

Figure 5.14: Sensitivity when the detection probability at Hong Kong is (a) its fitted value and (b) its raw value in the Asian source network. The fitted probability is 0.9525. The raw value is 0.18. The allocation $(1/3, 1/3, 1/3, 0, 0)$ is best when the detection probability at Hong Kong is its raw value.

allocation of spreading resources across only two nodes was not too bad. But in this case, it performs the worst out of the tested allocations for detection-sensitive smugglers (greater α). The allocation of spreading resources evenly across Ho Chi Minh City, Hanoi, and Da Nang may be best as we have seen here.

Given that Hong Kong has recently increased penalties for wildlife trafficking [6], it may be the case that it currently has low detection and this scenario was discussed in this section. However, as Hong Kong enforces the new laws, the results from the previous section may represent the changes we begin to see. For example, Hong Kong was used less as a transit location for detection-sensitive smugglers, despite being a transport hub.



(a) Hong Kong Fitted Detection Probability (b) Hong Kong Raw Detection Probability

Figure 5.15: Transit frequency at Bangkok and Hong Kong when the detection probability at Hong Kong is (a) its fitted value and (b) its raw value. Increasingly less routes stopover at Hong Kong when it has the higher fitted detection probability. These routes instead go through Bangkok.

5.3.2 Reducing Maximum Detection Probabilities

The total proportion that went undetected through the network in the previous section was very small. This is because we used a detection probability curve with a maximum of one. It is very unlikely that the detection probability is ever this high. Realistically, it is lower. In this section we investigate the effects of scaling the node detection probabilities to reduce this maximum. We want to see if there is a change in the optimal allocation. If we were to implement this allocation in real life, the detection probabilities would likely be smaller and so it is important to test for any differences.

The scaled detection probability at node i, $\hat{p}_i$, is

$$\hat{p}_i = cp_i, \tag{5.3}$$

where p_i is the original unscaled detection probability at node i and c, $0 < c \leq 1$, is the scaling factor.

Suppose that the smuggler only cares about maximising the probability of going undetected, $\alpha = 1$. Figure 5.16 shows the proportion that goes undetected through the network when we scale all of the node detection probabilities. The proportion undetected when no scaling is present is 0.079. The total proportion undetected through the network increases as we decrease the detection probabilities. Reducing the detection probabilities by a factor of ten, *i.e.,* $c = 0.1$, means 0.81 of illegal trade goes undetected. This figure is much more realistic than 0.079.

This curve is clearly non-linear. This is because the proportion undetected function (see Equation 3.2) is not linear in the probability of detection. It contains products of probabilities of non-detection along edges. However, this curve sits below the linear one shown in Figure 5.16. As the node detection probabilities are decreased, the proportion undetected is less than the linear value. So if, for example, we half all node detection probabilities, *i.e.,* $c = 0.5$, the total proportion undetected through the network is now around 0.3. This is better than what if would have been if there was a linear relationship.

Figure 5.17 shows the optimal resource allocations and proportions undetected when the detection probabilities are decreased by a factor of ten, *i.e.,* $c = 0.1$. The proportions undetected are now higher than the original non-scaled ones as expected.

The optimal resource allocations are different to the ones when no scaling is present for small values of α. More resources are now placed at Da Nang. However, this does not influence the total proportion of undetected trade. Figure 5.17 shows the proportion that goes undetected when we implement the optimal allocations from the non-scaled network to the scaled one. There is very little difference between the two proportions.

Once again, we would like a robust allocation across smuggler objectives. Given the proportion of undetected trade when we apply the original allocation to the net-

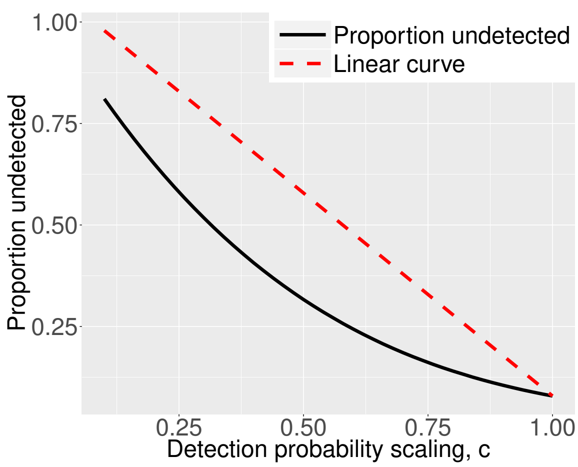
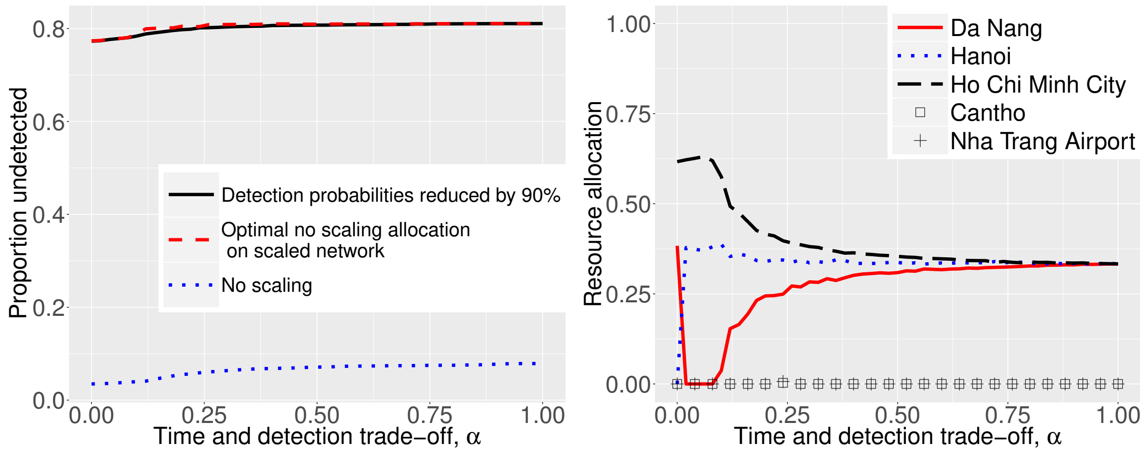


Figure 5.16: Proportion of illegal trade that goes undetected through the network for different scalings of the node detection probabilities. The smuggler solely aims to minimise detection. The proportion-undetected curve is not linear.



(a) Proportion undetected. (b) Optimal resource allocation.

Figure 5.17: (a) Proportion of undetected illegal trade when node detection probabilities are decreased by a factor of ten, *i.e.*, $c = 0.1$, and the (b) optimal resource allocation. The sources are in both Asia and Africa. The optimal resource allocations are different than the non-scaled allocations for smaller values of α. However, the proportion undetected is not sensitive to these differences. Also shown in (a) is the proportion undetected when the optimal allocation under no scaling is applied to the network with scaling. There is little difference between the two proportions.

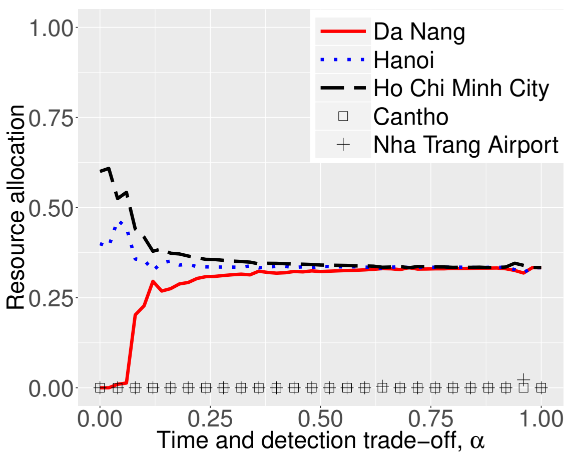
work with scaled detection probabilities remains unaffected, we assume the specified allocations would perform the same on these networks. Recall the two robust allocations are $(1/3, 1/3, 1/3, 0, 0)$ and $(2/5, 2/5, 1/5, 0, 0)$.

So, we have looked at the effect of reducing the node detection probabilities. We did this as the previous probabilities were unrealistically high. The resulting optimal resource allocations changed slightly. This could be due to the stochasticity in the algorithms as the objective functions may have become flatter with lower detection values. However, we still detected the same proportion of illegal traffic when we used the original allocations on the network with reduced probabilities. This means that we may not need to consider the scaling of the detection probabilities when determining the optimal resource allocation. If we require an accurate estimate of the proportion of undetected illegal trade then this may become more important. But for assigning resources in this model it is not.

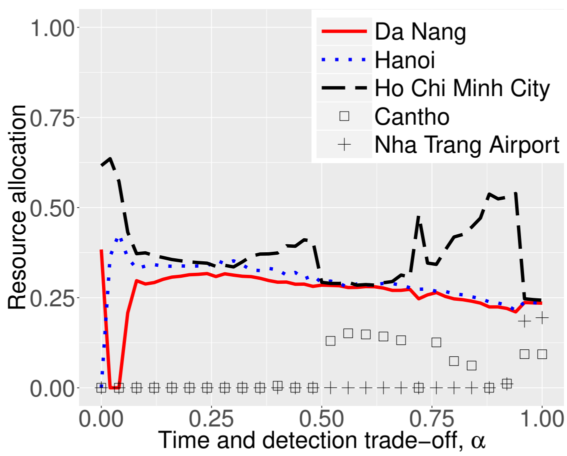
Reducing Detection at Non-destination Nodes

Previously, we reduced the detection probabilities across all nodes. However we have mentioned that perhaps the detection probabilities at transit locations are lower than at destination nodes because security procedures on transiting passengers are not centered on uncovering trafficking [8]. Thus, here we find the optimal allocations when source and transit detection probabilities are reduced. This may be a more realistic scenario and so we want to compare the optimal allocations.

We reduce the original detection probabilities as before (see Equation 5.3). Figure 5.18 shows the optimal allocation when source and transit detection probabilities are reduced by factors of two and ten, *i.e.*, $c = 0.5$ and $c = 0.1$ respectively. When $c = 0.5$, the optimal allocation is similar to the one under no scaling. However, the optimal allocation when we reduce the detection probabilities to 0.1 of their original values, is very different. Resources are now allocated at Can Tho for larger smuggler detection weights (greater values of α). Previously, no resources were allocated to this city. This is explained by the underlying connections in the transport network and the node detection probabilities. Can Tho is connected directly to Taiwan. There is an increasing number of paths into Vietnam that go through Taiwan as α increases. There is a decrease in the number of paths via Thailand. Thailand has a lower corruption perception index than Taiwan. So previously Thailand was more likely to

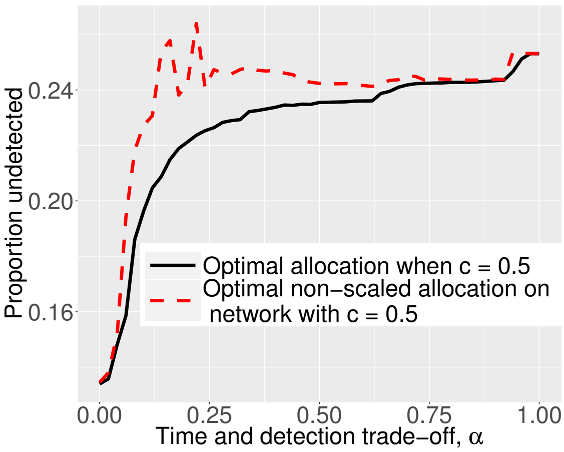


(a) $c = 0.5$.

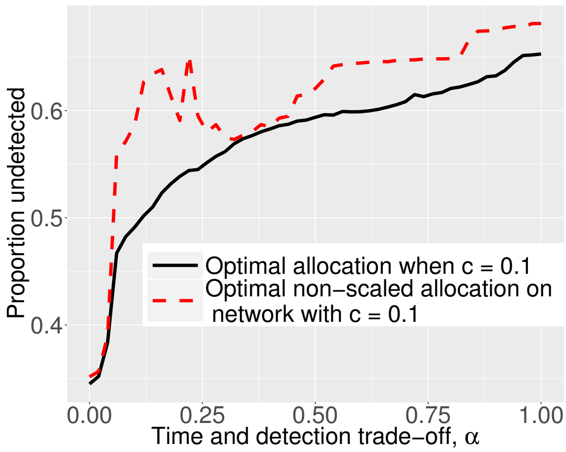


(b) $c = 0.1$.

Figure 5.18: Optimal allocation when non-destination node detection probabilities have been scaled with (a) $c = 0.5$ and (b) $c = 0.1$. The allocation when we half the detection probabilities ($c = 0.5$) does not vary too much from the original allocation. However the optimal allocation is very different when we decrease the detection probabilities to 0.1 of their original values.



(a) $c = 0.5$.



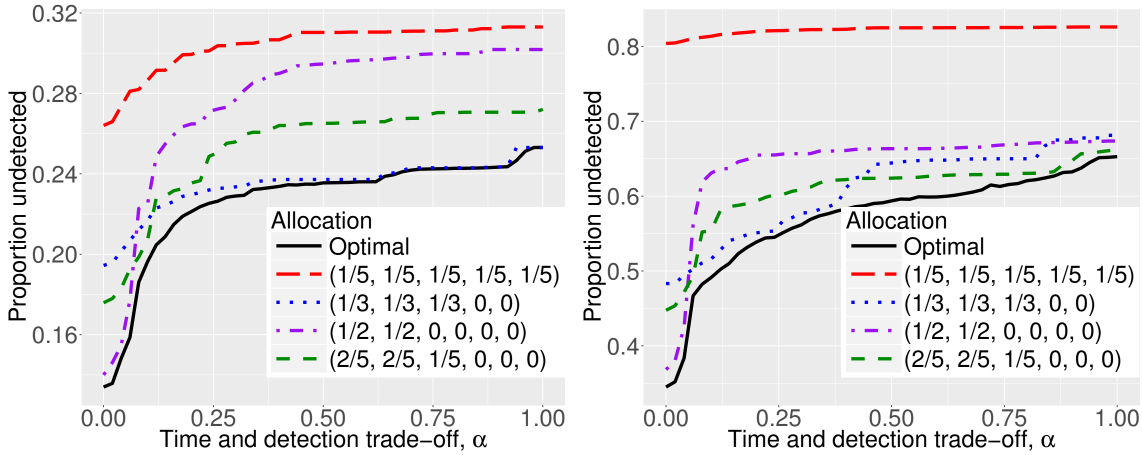
(b) $c = 0.1$.

Figure 5.19: The proportion of undetected illegal trade through the network with scaled non-destination node detection probabilities for (a) $c = 0.5$ and (b) $c = 0.1$. The proportion under the optimal solution as well as the original allocation from the non-scaled network is shown. The original allocation does not intercept as much illegal trade. The proportion undetected under these scaled detection probabilities appears sensitive to these changes in allocation.

lie along the smuggler's optimal route. Now entry via Can Tho is more desirable as there are less resources placed there and Taiwan has a lower detection probability.

Though the optimal allocations look different, our original allocations may still be sufficient at interdiction under these new detection probabilities. Figure 5.19 shows the proportion of undetected illegal trade when the non-destination node probabilities are reduced. Also shown is the proportion undetected when the original optimal allocation is applied to the network with scaled non-destination node detection probabilities. We see that the proportion undetected increases more when we apply the non-scaled optimal allocation. This is because the smugglers are more sensitive to changes in detection at the destination node given the detection at source and transits is lower. So, if we assumed that the underlying probabilities were our original ones, and assigned the corresponding optimal allocations, then if the underlying non-destination nodes were actually lower we would intercept less. This is not ideal.

However recall we don't know the true value of α. What we would prefer is an allocation across α that performs well. Figure 5.20 shows the proportion of illegal trade that goes undetected for various allocations. The allocation that still performs best overall is to spread resources evenly across three nodes: Ho Chi Minh City, Hanoi, and Da Nang.



(a) $c = 0.5$. (b) $c = 0.1$.

Figure 5.20: Proportion of undetected illegal trade under various allocations when the non-destination nodes have been scaled down by (a) $c = 0.5$ and (b) $c = 0.1$. Under both cases, the best allocation across α is to spread resources evenly across three nodes: Ho Chi Minh City, Hanoi and Da Nang.

So we have seen that when we reduce non-destination node detection probabilities, the proportion of undetected illegal trade through the network is more sensitive to deviations from the optimal allocation. We are interested in this as security may be lower at non-destination locations. We don't know the true ratio of non-destination node detection probabilities to destination ones. However we can say that when we reduce the non-destination node detection probabilities to 0.5 and 0.1 of their initial values, the most robust allocation across α that we tested is to spread resources evenly across Ho Chi Minh City, Hanoi, and Da Nang.

5.4 Source Traffic Model

The previous section and results assumed uniform traffic from all sources in a given category into Vietnam. It is more likely that the levels of traffic vary between sources. In this section we construct a case study with non-uniform source traffic.

We require a model to assign traffic to sources but data on this for every source in our network is unavailable. Instead we propose to use a model with as much realism as possible from the data that is available. We do this using a model for the amount of traffic between two nodes in the network.

Gravity models have been used before to estimate volumes of trade between two cities [26]. The trade flows between origins and destinations are assumed to be functions of the economic masses of each city, commonly gross domestic product (GDP), and the geographical distance between them. The general form of a gravity model is

$$T_{ij} = \alpha_1 Y_i^{\alpha_2} Y_j^{\alpha_3} f(d_{ij}),$$

where T_{ij} is the trade between exporter i and importer j, Y_i is the GDP of trader i and $f(d_{ij})$ is a function of the geographical distance between i and j. We expect less traffic between nodes that are further apart. This is captured by the distance function. Typically we assume an exponential penalty on distance.

Though the GDP was used in the general gravity model, we do not use this to model traffic. Instead we use the total flow of air traffic in and out of each node. We do this because we have this information for every source and destination node as the network itself was created from the same dataset.

Let T_{ij} be the passenger volume/traffic from airport i to airport j. Our gravity model for predicting air traffic is

$$T_{ij} = K \frac{O_i^a D_j^b}{e^{\beta d_{ij}}}, \tag{5.4}$$

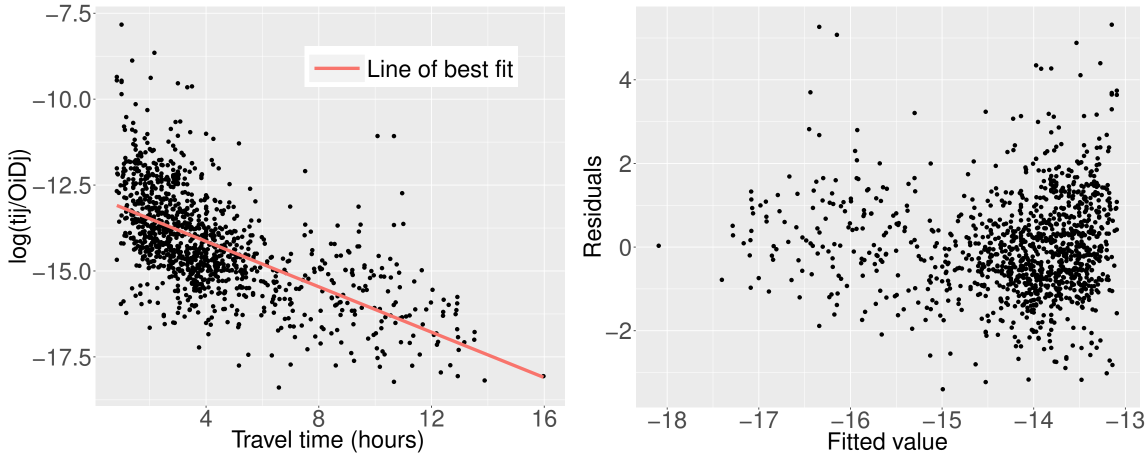
where O_i is the total traffic leaving airport i, D_j is the total traffic entering airport j, d_{ij} is the distance/time between the airports, and a, b and β are parameters. For simplicity we set $a = b = 1$, but we still need to choose β.

We will use the OAG flight data discussed in Section 5.1 to fit the β parameter of the gravity model. We are assuming this data will provide sufficient information on the traffic into and out of airports, allowing us to estimate β.

We can take the log of the model above to get

$$\ln T_{ij} = \ln O_i + \ln D_j + \beta d_{ij} + C.$$

We then take all of the edges and nodes in the Vietnam case study network described in Section 5.1. We know the air traffic between two nodes that share a



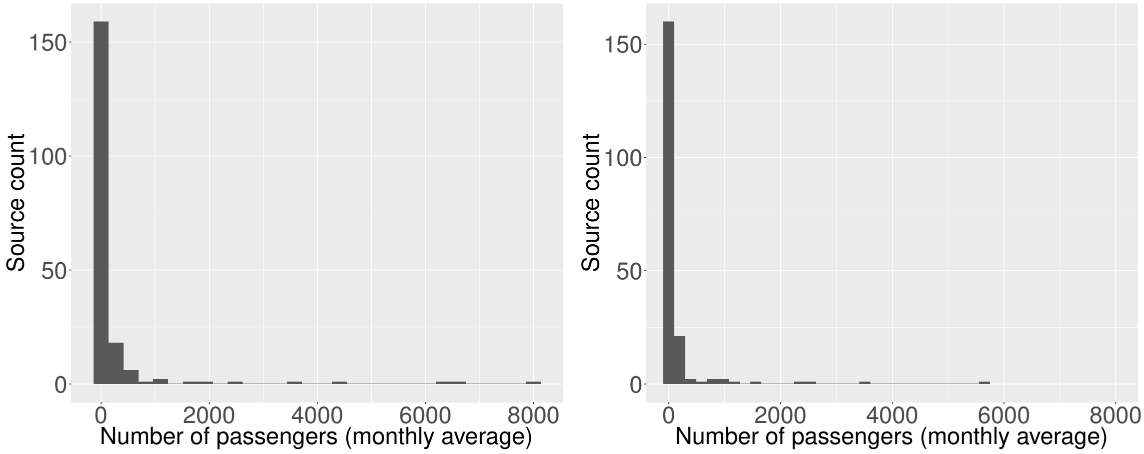
(a) Linear fit to air traffic data. (b) Residuals vs fitted plot.

Figure 5.21: The (a) linear regression fit and (b) residuals. The regression is used to fit the parameter β in the gravity model.

β	0.33023
Constant, C	-12.81818
R^2	0.3561
Residual Std. Error	1.187 (df = 1176)

Table 5.4: Linear regression output of gravity model.

direct edge from the data. The model is then fitted to these traffic and distance values with linear regression. The resulting fit is shown in Figure 5.21. A plot of the residuals is shown in Figure 5.21 and the regression output in Table 5.4. The fit appears sufficient for our purposes. There isn't any strong trend in the residuals vs fitted plot. The model has the parameter $\beta = 0.33023$. We use this model to predict the traffic between sources and destinations in the network. The traffic into Ho Chi Minh City and Hanoi from each source is shown in Figure 5.22. There are a few large traffic volumes but most are quite small. To reduce the effect of such large volumes, normalisation may be required.



(a) Ho Chi Minh City (b) Hanoi

Figure 5.22: Estimated traffic from sources into (a) Ho Chi Minh City and (b) Hanoi. Many sources have small predicted traffic. There are a few that have large values. These sources with more traffic will have a greater influence on the optimal resource allocations.

5.4.1 Interdiction Model with Source Traffic

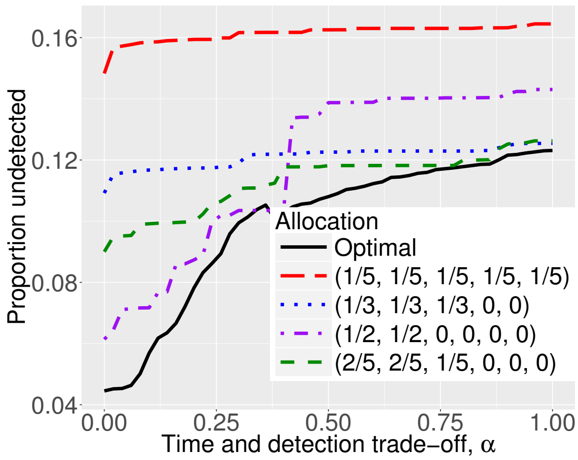
We can now incorporate the traffic into our case study. Each source $i \in S$ is weighted by its total estimated traffic into Vietnam,

$$w_i = \sum_{j \in D} T_{ij},$$

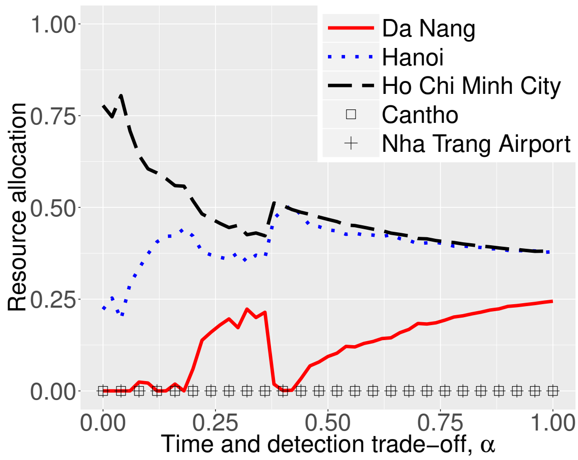
where T_{ij} is the traffic from node i to node j.

As we saw in the previous section, some sources have heavy traffic volumes. Also, the Asian sources have more traffic than the African ones as they are closer geographically to Vietnam. To address these issues we normalise the traffic within each source group (African and Asian) by dividing by the maximum traffic.

We then use these source traffic weights in our model. The results are shown in Figure 5.23. The optimal allocation across α is to place more resources at Ho Chi Minh City and Hanoi, less at Da Nang, and none at the other two nodes. These allocations are different than the ones in the uniform traffic problem. More resources are placed at Da Nang in the uniform problem. This is explained by the network of heavily weighted Asian sources. The weights of sources with weightings of at least 0.1 are shown in Table 5.5. Some of these sources do not have direct connections to Da Nang. For example, Bangkok, which has a relatively large traffic weight, is only connected to Ho Chi Minh City and Hanoi.



(a) Proportion undetected.



(b) Optimal resource allocation.

Figure 5.23: The (a) proportion of undetected illegal trade with normalised source weights from the traffic model under the (b) optimal allocation. The optimal allocations are different to the ones under uniform source traffic. Here we place less resources at Da Nang for higher values of α. This is because some of the heavier weighted Asian sources do not have connections to Da Nang. The proportion undetected also appears slightly more sensitive to changes in allocation.

Figure 5.23 shows proportions of undetected trade given different allocations when the sources are weighted using the traffic model. The proportion undetected is more sensitive to these changes than when uniform traffic was used. When the smuggler is more time sensitive, the allocation of spreading resources between two cities is much better than spreading between three. However, when the smuggler is more detection sensitive, it is better to allocate some resources to Da Nang. The most robust allocation tested overall is to assign $2/5$ of resources to Ho Chi Minh City and Hanoi, and $1/5$ to Da Nang.

As many sources have small weights, we consider the model with just sources that have weightings greater than 0.1. The optimal resource allocation is shown in Figure 5.24. It appears to be very similar to the optimal allocations when all sources are included.

So we have seen that incorporating source traffic influences the optimal resource allocation. The proportion of illegal trade that goes undetected becomes more sensitive to changes away from the optimal. The allocation of spreading resources across three specific nodes has appeared the most robust so far. However, in the non-uniform traffic case, it does not perform as well when the smuggler is more time sensitive (*i.e.*, for smaller values of α). But it still performs better than the allocation of spreading

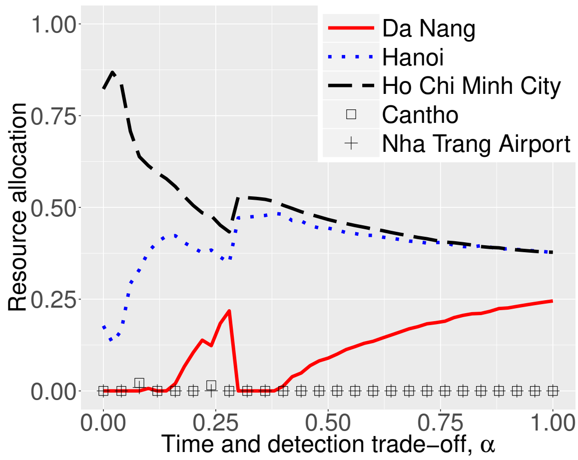


Figure 5.24: Optimal resource allocation on network with relatively heavily weighted source nodes. There are 16 source nodes (rather than 189).

resources across all nodes.

Given the differences in allocation between the uniform and non-uniform source traffic models, traffic from sources should definitely be considered when allocating resources.

Source Airport	Traffic Weight
Hong Kong International Apt	1
Taipei Taiwan Taoyuan International Apt	0.4473
Singapore Changi Apt	0.6564
Shanghai Pudong International Apt	0.2013
Bangkok Suvarnabhumi International Apt	0.7248
Kuala Lumpur International Airport	0.4385
Manila Ninoy Aquino International Apt	0.2026
Beijing Capital Intl Apt	0.1520
Guangzhou	0.1265
Addis Ababa	1
Nairobi Jomo Kenyatta International Apt	0.8913
Lagos	0.1292
Mombasa	0.1039
Dar es Salaam	0.1218
Entebbe	0.1633
Kigali	0.1157

Table 5.5: Sources with weights that are greater than 0.1. These weightings affect the optimal resource allocation. Bangkok has a relatively heavy weighting and has no connections to Da Nang. Da Nang is assigned less resources in this non-uniform traffic case.

5.5 Discussion

In this chapter we have developed a case study that tested our model in a more realistic setting. We explored the influence of changing parts of the model on the optimal resource allocation. In particular, we considered the effects of having different groups of sources, altering estimated node detection probabilities, and assigning non-uniform traffic from sources.

The case study considered the illegal trade of pangolins into Vietnam from Asian and African sources. We first found optimal resource allocations on networks with Asian and African sources separately. There were different optimal resource allocations for both networks. More resources were assigned to Da Nang in the Asian source network. This was because the African sources had less connections to Da Nang. So an alternative strategy may be better if African pangolins require more protection.

However, we saw that sources located in Africa were more likely to have higher probabilities of detection along their paths into Vietnam. This is because they all had intermediate stopovers. There was an additional detection probability associated with each stopover. So the detection probability at a destination node at the end of these paths influenced the overall detection less. Note that the transit nodes still had high detection probabilities. Perhaps more realistically these detection probabilities should be lower if they just act as a transit.

We also saw the locations that were used as transits along routes. These locations changed for different smuggler objectives. However, they largely remained to be major busy airports. This agrees with the literature that traffickers tend to exploit the larger, busier airports [8].

The total proportions of illegal trade that went undetected were initially quite small. This was because our node detection probability function allowed high detection probabilities, which is unrealistic. We reduced all the node detection probabilities. The optimal resource allocation appeared to be different for cases when the smuggler was very time sensitive. However when we applied the original optimal allocation to this scaled network, the total proportion undetected did not deviate greatly from the optimal. This was good as it means we don't have to consider this scaling too much when we aim to allocate resources, just the relative differences between nodes. However if we need to report the total proportion undetected through the network, then the node detection probability estimates need to be more accurate.

We also found the optimal allocation when non-destination node detection probabilities were decreased. Detection at non-destination nodes may be lower in real life as the goods are not leaving an airport and so go through less security checks. The proportion of undetected illegal trade was more sensitive to changes away from the optimal allocation.

Given we don't know the smuggler's weight between detection and time, α, we found the proportion that went undetected across α for specific allocations. For the scenarios just described, the most robust allocations, allocated resources to three specific destination nodes: Ho Chi Minh City, Hanoi, and Da Nang. The first robust allocation was to assign resources evenly among Ho Chi Minh City, Hanoi, and Da Nang. The second robust allocation was to assign 2/5 of the resources to both Ho Chi Minh City and Hanoi, and 1/5 to Da Nang. We expect that small perturbations from these allocations, would still give robust solutions. However, thorough analysis of the sensitivity of them will be left for future work. These three cities had the most connections in the transport network. So these robust allocations make intuitive sense. However, this may not necessarily have been the case. The other ports of entry, Can Tho and Nha Trang, were assigned no resources in this allocation. But, they were still connected (indirectly) to major hub airports like Hong Kong.

We then looked at non-uniform traffic from sources. The optimal allocations with non-uniform traffic were different to those with uniform traffic. Less resources were now allocated to Da Nang. The proportion undetected for this model was also more effected by resource allocation changes. The allocation of spreading resources across three nodes, took the proportion undetected from the network further away from the optimal than in previous scenarios. However it was still much better than allocating evenly across all five nodes.

Many sources in the non-uniform traffic case were assigned almost zero resources. When we removed sources with less than 0.1 traffic from the network, the resulting optimal allocation remained the same as with all of the sources. So perhaps in the future we could remove sources from the network that have low weightings as they have little effect on the results. This would increase the efficiency of the algorithms used to find the optimal allocations.

There are limitations to these results. First, we considered only five entry points into Vietnam. This came from constructing a transport network assuming sources were within two hops of a Vietnamese airport. Realistically, there are more entry points. We also only considered airports. Other entry points would include harbours and border crossings. However, our case study on this smaller network still revealed insights. The most robust allocation was to allocate resources to three out of the five destination nodes. These were the nodes with more connections. If you were to include other entry points, the ones with a greater number of connections would likely get more resources. However, you would also need to consider the traffic into the entry points.

There is some uncertainty in our detection probability estimates. However we did test the sensitivity of the detection probability at Hong Kong on the optimal resource allocation. Hong Kong was explored in particular as its fitted detection probability was substantially higher than its raw one. Despite this large change in detection, the allocation to spread resources evenly across Ho Chi Minh City, Hanoi, and Da Nang remained robust across multiple smuggler objectives.

We did not consider the scenario where smugglers and biosecurity forces have different perceptions of network parameters. For example, smugglers may not have full knowledge of resource allocations, and thus a different view of detection probabilities. Perhaps this would change the optimal allocation. It would likely influence the total proportion of undetected illegal trade. Although, the allocation of assigning resources across the three nodes may still be robust.

The smuggler's path choice is very sensitive to small differences in time and detection between potential paths. For example, if the flight between a source and Ho Chi Minh City is 2 hours, and the one to Hanoi is 2.1 hours, the smuggler will always take the one to Ho Chi Minh City. It is more likely that this small time difference does not influence the path choice so much.

We also did not directly take into account the frequency of flights into Vietnam. If a higher number of flights connected a source and a particular Vietnamese airport, then you would expect it to be more likely that the smuggler takes this path.

This raises another limitation. We may not have accounted for other smuggler objectives. It is likely that smugglers take paths that are not the optimal from our model. The assumption that smugglers always take the optimal path given some trade-off between time and detection is a large part of the current model. In the next chapter we aim to develop a new model that relaxes this assumption.

Chapter 6

Multi-Path Model

We have assumed so far that all smugglers travel along a single optimal route depending on their source, destination, and time-detection weight, α. However, the assumption that all smugglers choose their route in this way is unrealistic. We do not have complete information around what influences their decision. It is possible that there are objectives other than detection and time that we have not accounted for. For example, pangolin commodities may be part of a larger global shipment with other illegal goods of different species and that impacts the route taken. Also, cost may be another issue, *i.e.,* different flights may cost different amounts. As well as that, biosecurity forces and smugglers may have different perceptions of variables such as detection probability. So we may attempt to predict the path of a smuggler using our knowledge, but in practise they may choose an alternative option. Our model is also sensitive to small differences between path distances. If there are two paths of very similar lengths, then the smuggler will always pick the smaller one and this may not always be the case in real life. Thus we expect some variation among their choices.

So it is more likely that multiple paths from the same source are used by a group of smugglers. We aim to incorporate this idea in the model. Consider the scenario where traffic is sent from a source to the destination country along different paths. How might these paths be chosen? The case where traffic is evenly distributed to all paths is unlikely and not particularly interesting. Perhaps more traffic moves along certain paths that appear more desirable to the smuggler?

We will refer to the previous model as the *single path model*. This chapter defines a new model of the smugglers movement. Multiple paths are used to transport commodities from the same source. This model is called the *multi-path model*.

Suppose that wildlife commodities are being smuggled from a source $s \in S$ into a destination country D. Let $P_{s,D}$ be the set of all paths between s and the destination nodes D. We assume that more traffic will be sent along more desirable paths, *i.e.,* paths with smaller weights. The weight of a path depends on the smuggler's assumed objective, which is some combination of minimising travel time and maximising the probability of non-detection. Recall the weight α controls the trade-off between the two objectives. For example, more traffic will be sent along paths with short travel times when the smuggler's objective is to minimise travel time (*i.e.,* $\alpha = 0$). Conversely, if the smuggler maximises their probability of non-detection (*i.e.,* $\alpha = 1$), then more traffic is sent along paths with lower detection probabilities.

Splitting traffic among paths is a research problem in another domain: internet traffic. Xu and Chiang [58] distributed traffic among paths by penalising longer paths with an exponential function. They defined the proportion of traffic along the i^{th} path from nodes u to t, $x^i_{u,t}$, to be inversely proportional to the exponential of its path length, $p^i_{u,t}$,

$$x^i_{u,t} = \frac{e^{-p^i_{u,t}}}{\sum_j e^{-p^j_{u,t}}}.$$

We use a similar function to split our source-destination traffic. Our traffic proportion from source s into destination D along the i^{th} path is

$$x^i_{s,D} = \frac{e^{-\beta q^i_{s,D}}}{\sum_{j\in P_{s,D}} e^{-\beta q^j_{s,D}}}, \tag{6.1}$$

where $q^i_{s,D}$ is the weight along the i^{th} path from s into D and $P_{s,D}$ is the set of paths from s into D. The parameter β controls the proportion of traffic that gets distributed along the optimal path. When $\beta = 0$, the traffic is split evenly between all paths. More traffic is placed on the optimal path as β increases. In the limit as $\beta \to \infty$, we solve the single path model.

The path weights are calculated the same way as in the single path model. This weight is

$$q^i_{s,D} = \sum_{(x,y)\in i} -\alpha\frac{\log(1-p_{xy})}{\pi} + (1-\alpha)\frac{t_{xy}}{\tau},$$

where α controls the trade-off between the two objectives. An α closer to one indicates that the smuggler is more detection sensitive. An α closer to zero means the smuggler cares more about minimising travel time. This may be important for transportation of live or frozen commodities.

The multi-path biosecurity objective is then to place resources across the destination nodes to minimise

$$f \quad = \quad \min_{\{\boldsymbol{\delta}\}} \quad \frac{1}{\sum_{s\in S} w_s}\sum_{s\in S} w_s \sum_{i\in P_{s,D}} x^i_{s,D} f^i_{s,D}(\boldsymbol{\delta}), \tag{6.2}$$

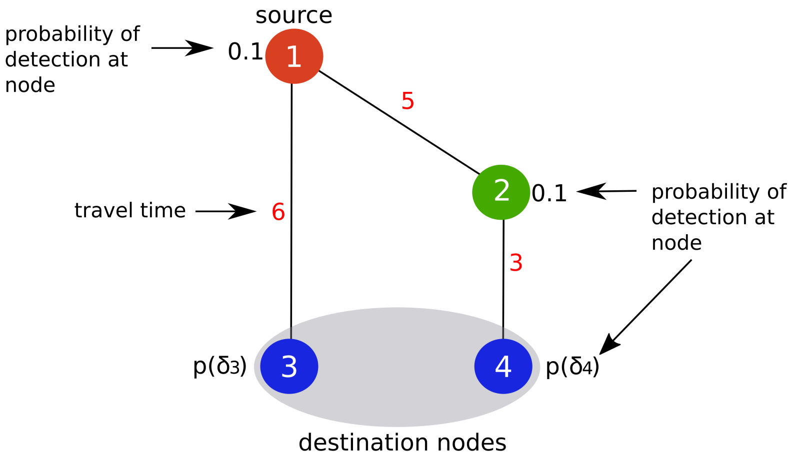


Figure 6.1: A transport network. Smugglers attempt to move illegal goods from the red source node into the destination country (blue nodes). There are baseline probabilities of detection at the source and potential transit nodes. The travel time is also displayed along each edge. Illegal traffic is split between the two paths into the destination country.

where w_s is the weight or total traffic from source s and $f^i_{s,D}(\boldsymbol{\delta})$ is the probability of non-detection along the i^{th} path from s into D given the resource allocation $\boldsymbol{\delta}$. The value f is the total proportion of illegal traffic that goes undetected through the network.

This new model has removed one of the fundamental assumptions of the single path model. We no longer need to compute the smuggler's shortest path. However, we require all path lengths from sources to destinations. This process is discussed in the following section.

We can still apply the GA and INSH methods to find the optimal resource allocation in the multi-path model. This is also detailed in the next section. We are interested in the optimal allocations for different values of β. We will investigate the effect that multiple paths has on the Vietnam case study in Section 6.2. But we first discuss the multi-path model on a small transport network.

Consider the small transport network in Chapter 3 (Example 1.2.1). It is shown again in Figure 6.1. We assume that the smuggler cares equally about time and non-detection. So α is set to 0.5. We apply the new model to the network for $\beta = 3$ and $\beta = 700$. The probability of non-detection given different allocations at node 3 is shown in Figure 6.2. The non-detection probability under the single path model is also shown in Figure 6.2. The most important thing we notice is there is no jump

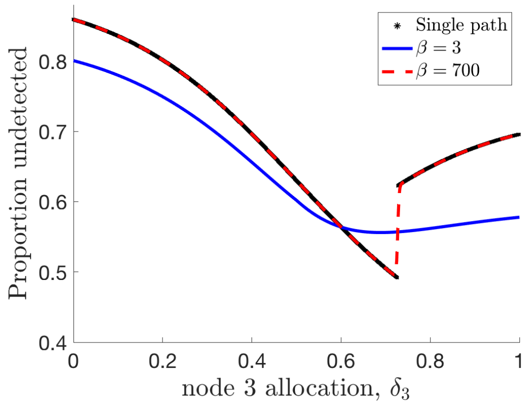


Figure 6.2: Proportion of illegal traffic undetected, given the resource allocation at node 3, on the underlying transport network shown in Figure 6.1. The smuggler's optimal objective weight is set to $\alpha = 0.5$. There is a discontinuity under the single path model that comes from a switch in smuggler's optimal path. The curve of $\beta = 700$ is almost the same as the single path model except there is no discontinuity as a tiny amount of traffic goes on the alternative path. The proportion undetected curve for the smaller β value, $\beta = 3$, is smoother.

in the function like there was in the single path model. This is because traffic is sent along multiple paths. The jump in the single path corresponds to a switch in path. We don't see this sudden change in the multi-path model. For $\beta = 3$, traffic is distributed between the two paths. For $\beta = 700$, most traffic is sent along the optimal path. However there is still a tiny fraction of traffic (0.006%) that takes the alternative path.

It is interesting that the proportion undetected for $\beta = 700$ is less than that of $\beta = 3$ between $0.6 < \delta_3 < 0.75$. Higher β values mean there is more traffic on the smuggler's optimal path. At $\delta_3 = 0.7$, the probability of detection along the path 1-3 is 0.5052 and along 1-2-4 is 0.6114. For $\beta = 700$, all traffic is still sent along the path 1-3 even though this path has a higher probability of detection. This is because the smuggler cares equally about minimising time and detection. The path has a lower travel time though it has a higher detection probability. For $\beta = 3$, some traffic is still sent along the non-optimal path, which has a lower detection probability.

For cases where the smuggler cares only about detection, $\alpha = 1$, the proportion of illegal trade that goes undetected is always higher for larger values of β. This is because more traffic is sent along the path with the smaller detection probability. When $\alpha = 1$, the proportions undetected for $\beta = 700$ and $\beta = 3$ are shown in Figure 6.3. We see that a higher proportion goes undetected for the larger value of β.

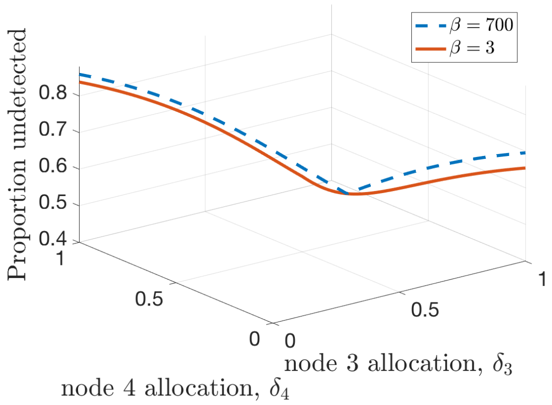


Figure 6.3: The proportion of illegal traffic undetected for the multi-path model with $\alpha = 1$. The underlying transport network is shown in Figure 6.1. Larger values of β will have higher proportions undetected as more traffic is directed along paths with smaller probabilities of detection. This behaviour is not always true for $\beta < 1$ as the smuggler also cares about travel time.

6.1 Computing the Optimal Resource Allocation

We aim to find the optimal biosecurity resource allocation of the multi-path model. The difference between the two models is the way traffic is distributed among paths. In the single path model, we assume that all traffic moves along the optimal path from a source into the destination country. This path is the solution of a graph shortest path problem. We find it using Floyd-Warshall's shortest path algorithm [45]. In the multi-path model, traffic is directed along multiple paths from sources to destinations. The proportion of traffic directed along a path depends on the weight of the path. These weights are still determined by the smuggler's assumed trade-off between time and detection, α. However we no longer find the shortest path. Instead, we evaluate all source-destination paths and their weights.

The methods used to calculate the proportion of undetected trade are different between the models. Nevertheless, we can still use our evolutionary algorithms, GA and INSH described in Chapter 4, to find the optimal allocation in the multi-path model. A high level description of the steps involved is shown in Algorithm 4. We begin by finding all paths from sources to destinations. We then generate an initial set of resource allocations and evaluate the proportion of undetected trade. New allocations are sampled around the old allocations using sampling methods in INSH or GA. These methods were discussed in detail in Chapter 4. We iterate for a fixed number of iterations.

We need to find all source-destination paths. Generally, finding all paths is a

Algorithm 4 INSH/GA for Multi-Path Optimal Allocation

Compute all source-destination paths using Algorithm 5
Generate initial set of resource allocations $\{\boldsymbol{\delta}\}$
for each iteration **do**
  Evaluate proportion of undetected trade, $f(\boldsymbol{\delta})$, for each resource allocation
  Sample new allocations using INSH or GA methods
end

Algorithm 5 Source-Destination Paths

Inputs: $G(V, E)$, S, T, D
$P_{O,D} \leftarrow \emptyset$
for each node $s \in S$ **do** ▷ 1-hop paths
  for each node $d \in D$ **do**
    if s-$d \in E$ **then**
      $P_{O,D} \leftarrow P_{O,D} \cup s\text{-}d$
    end
  end
end
for each node $s \in S$ **do** ▷ 2-hop paths
  for each node $t \in T$ **do**
    for each node $d \in D$ **do**
      if s-$t \in E$ & t-$d \in E$ **then**
        $P_{O,D} \leftarrow P_{O,D} \cup s\text{-}t\text{-}d$
      end
    end
  end
end
Return: $P_{O,D}$

challenging problem. There are an exponential number of simple paths. However we are able to do this easily in our application. Recall we constructed the transport network by assuming that all sources were within two hops of a destination node. It was possible for a path to be longer than this due to connections between transits, but, all of the optimal paths in the single path model had a maximum of two hops. So, in this multi-path model we will assume that every path has a maximum of two hops. Thus we know all the paths will either be direct links or have a single transit. We can generate every source-transit-destination combination and check whether the path exists. The method is outlined in Algorithm 5. Let S, T, and D be the sets

Algorithm 6 Multi-Path Proportion of Illegal Trade That Goes Undetected

Inputs: $G(V, E)$, $P_{O,D}$, α, β, $\{w_i\}$, $\{p_i\}$, $\boldsymbol{\delta}$
Compute edge weights given α, $\boldsymbol{\delta}$, and $\{p_i\}$
Compute weight along each path, $q^i_{s,D}$
$f \leftarrow 0$
for each s in S **do**
  for each i in $P_{s,D}$ **do**
    Compute $x^i_{s,D} \leftarrow \frac{e^{-\beta q^i_{s,D}}}{\sum_{j \in P_{s,D}} e^{-\beta q^j_{s,D}}}$
    Compute $f^i_{s,D}(\boldsymbol{\delta}) \leftarrow \prod_{(x,y,)\in i}(1 - p_{x,y})$
    $f \leftarrow f + w_i x^i_{s,D} f^i_{s,D}(\boldsymbol{\delta})$
  end
end
$f \leftarrow f / \sum_{i \in S} w_i$
Return: f

of source nodes, transit nodes, and destination nodes respectively. To determine all 2-hop paths, we evaluate $|S| \times |T| \times |D|$ possible paths. Our real world network is not large. It has less than 300 nodes. So, generating all possible paths and checking if they exist can be easily done. We only need to find the paths once in the multi-path model. In comparison, the shortest path evaluations have to be done for each resource allocation in the single path model. However, we still must re-compute the weights along paths given new resource allocations.

Algorithm 6 describes how we calculate the proportion of trade that goes undetected. We distribute traffic along alternative routes where the proportion of traffic directed along the i^{th} path from source s into D is $x^i_{s,D}$ (see Eq. (6.1)). This is easily calculated. We then compute the detection probabilities along the edges in each path. To calculate the total proportion undetected through the network we multiply by the source traffic divided by the total traffic.

We will use the GA in the following section to find optimal resource allocations. The GA was chosen over the INSH as it was used to find the single path optimal allocations. We assign the population size, N, to 50 as before. To choose the other parameter values to use in our real world investigation, we test three parameter sets on a case study scenario when $\beta = 15$, $\beta = 100$, and $\beta = 400$. The details of the transport network and detection probabilities used are as in Chapter 5. We assume uniform source traffic from African and Asian sources. The mean proportion decrease in minimum at each iteration for different parameter sets are shown in Figure 6.4. The

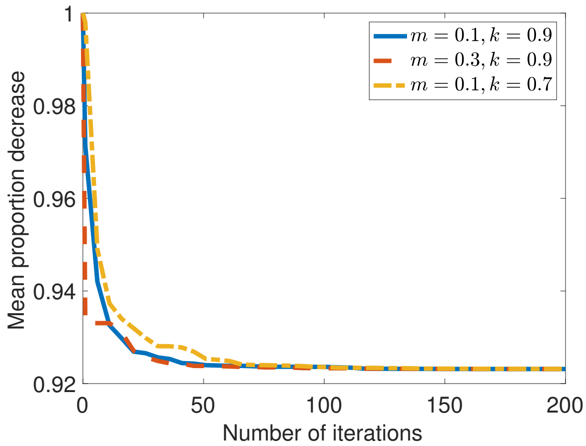


Figure 6.4: Optimal proportion of illegal trade that goes undetected as we run more iterations for different GA parameter sets. The mean is taken across minimums when $\beta = 15$, $\beta = 100$, and $\beta = 400$. The mutation probability is m and the crossover probability is k. All parameter sets converge to the same proportion.

sets with $k = 0.9$ converge faster. Therefore, we will assign the crossover probability to 0.9 and the mutation probability to 0.1 when we apply the GA in the following section.

Further work may involve more tests to see how these algorithms perform in a general setting. Properties tested could include accuracy, time and convergence. The algorithms should also be tested on more transport networks. Here we only looked at the convergence results on a single transport network. When doing this, we would also need to consider the number of hops in paths from sources to destinations. Recall we have assumed maximum path lengths of two hops and Algorithm 5 only finds paths with this maximum length. Thus we would need to consider the hop path lengths when constructing transport networks. For instance, this would need to be done to ensure there were two hops paths from sources to destinations. The effect that increasingly longer hop paths have on computational run time would also interesting to look at.

In the following section, we will use the GA to find the optimal resource allocations on a real world network. We will investigate how sending different volumes of traffic influences the optimal allocation, and hopefully find a robust allocation.

6.2 Case Study: Smuggling into Vietnam

In the last chapter we considered a smuggler who takes their optimal route from various Asian and African sources into Vietnam. As previously discussed, this is an unlikely scenario. For instance, smugglers may have slightly different objectives from what we have assumed, or different perceptions of detection probabilities. Thus we introduced the multi-path model to account for that fact that other paths are likely to be used as well. We aim to explore the effects of sending illegal traffic along multiple paths from a given source on the optimal resource allocation. Once again, we will consider this in a practical situation, extending on our case study from Chapter 5. We want to see if the robust allocations still perform well in this new model, or if better allocations arise.

Recall our case study considered the trafficking of pangolins from Asian and African sources into Vietnam. There were five entry points into Vietnam to which limited biosecurity resources were assigned: Ho Chi Minh City, Hanoi, Da Nang, Can Tho, and Nha Trang. In most scenarios, no resources were placed at Can Tho and Nha Trang. Instead, the resources were allocated between Ho Chi Minh City, Hanoi, and Da Nang.

We apply our multi-path model to the same transport network. We consider both Asian and African sources together and start with the assumption of uniform source traffic. We will consider non-uniform source traffic later in this section.

One parameter of the multi-path model is β. This parameter controls how much traffic is sent along the smuggler's proposed optimal route. In order to gain an understanding of how β influences this, we find the average proportion of traffic sent along the optimal route across all sources for different β values. Figure 6.5 shows these proportions. When $\beta = 0$, a very small proportion of traffic, on average, is sent along the optimal route. The proportion of traffic on the optimal route depends on the smuggler's objective weight, α. First consider the scenario where the optimal smuggler route minimises detection. This corresponds to $\alpha = 1$. The average proportion sent along the optimal route does not reach one. This is because the optimal allocation when $\alpha = 1$ is to spread resources evenly across five nodes. This gives those three destination nodes the same detection probabilities and results quite often in the same path weights into Vietnam. However, this is not the case for lower values of α. When $\alpha = 0.5$, we see that the average proportion on the optimal path is almost one at $\beta = 500$. When we weight time more, there is greater variation between path weights. So, more traffic is directed along the optimal path for smaller values of β.

Overall, this tells us that for values of β less than 50, traffic is split fairly evenly across paths as a smaller proportion is directed along the optimal. When β reaches 500, majority of traffic is sent along the optimal path. However if there are paths of similar weight, then the traffic is still split among these paths.

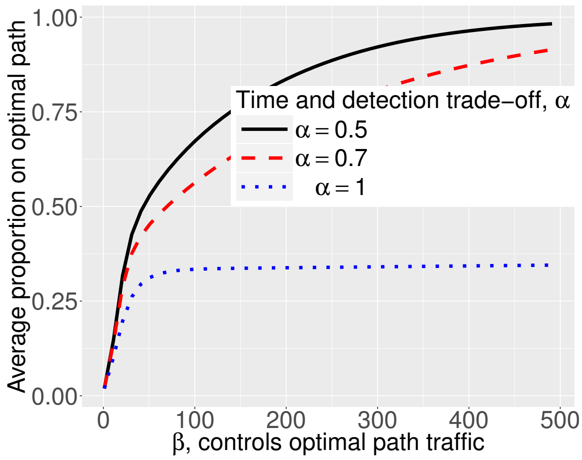
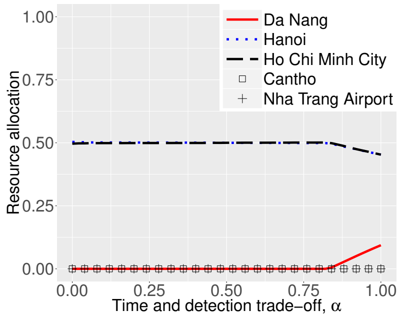


Figure 6.5: The average proportion of traffic sent along the optimal route from a source. The average proportions vary for different values of α. When $\alpha = 1$, there are multiple paths with the similar small weights, and thus traffic is distributed among these paths rather than all going on a single optimal path.

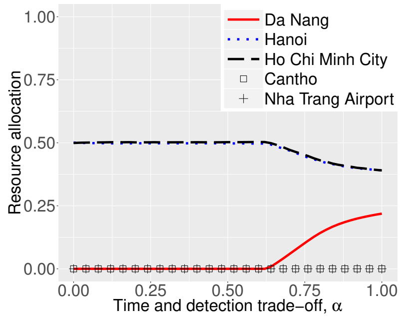
We are interested to know what the optimal resource allocations are for different values of β. We don't know the true value of β. More likely, there isn't a single value. But perhaps there is a robust allocation across different values of β as well as α. We have already found such allocations under different scenarios in the single path model. We will test these allocations in the multi-path model. But first, we will determine the optimal allocations across β to gain an understanding of how the allocations are influenced by this parameter.

Figure 6.6 shows the optimal biosecurity resource allocations at nodes in Vietnam. The single path model optimal resource allocation is also shown. As β increases, the multi-path model optimal allocation approaches the single path one. This occurs as more traffic is directed along the optimal route from a source. For $\beta = 10$ and $\beta = 15$, the optimal allocation is to place half of the resources at Ho Chi Minh City and Hanoi for most values of α. For these smaller β values, the traffic is roughly evenly distributed along the paths. There were 1021 paths into Ho Chi Minh City and 890 into Hanoi. But, the probability of detection along paths into Hanoi was less than that of Ho Chi Minh City. This is why slightly more resources need to be put at Hanoi even though less paths go into it. Once $\beta = 30$, we begin to see an allocation more similar to the single path optimal allocation.

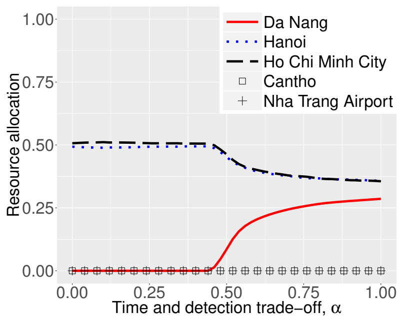
For $\beta = 500$, the optimal allocation does not look quite the same as the single path allocation. Around $\alpha \approx 0.2$, there is a large jump in the single path optimal allocation. However we do not see this in the multi-path allocations as a small fraction of traffic is still sent along non-optimal routes. This is, hopefully, a more realistic scenario.



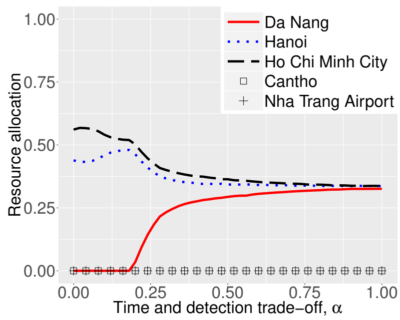
(a) $\beta = 10$.



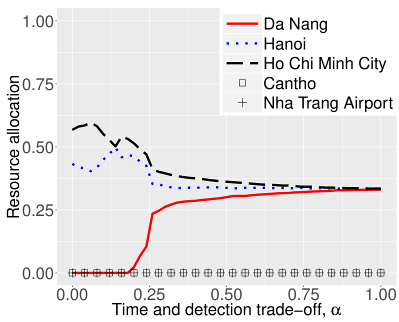
(b) $\beta = 15$.



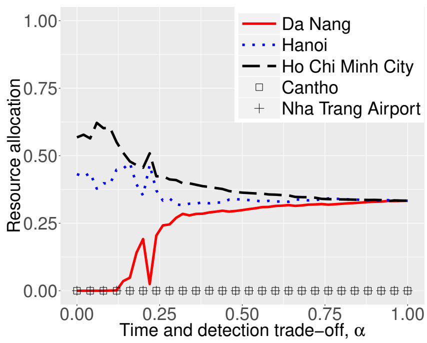
(c) $\beta = 30$.



(d) $\beta = 200$.



(e) $\beta = 500$.



(f) Single Path Model.

Figure 6.6: The (a)-(e) optimal resource allocations for different values of β in the multi-path model. Traffic is from African and Asian sources into Vietnam. The parameter β controls the proportion of traffic directed along the optimal path given the smuggler's objective weight, α. As β increases, the optimal resource allocation transitions to (f) the optimal of the single path model. This is because more traffic is sent along the optimal route for greater values of β. The jumps in the single path allocation around $\alpha = 0.2$ are also not present in the multi-path allocations. This is because a small amount of traffic is still sent along non-optimal routes.

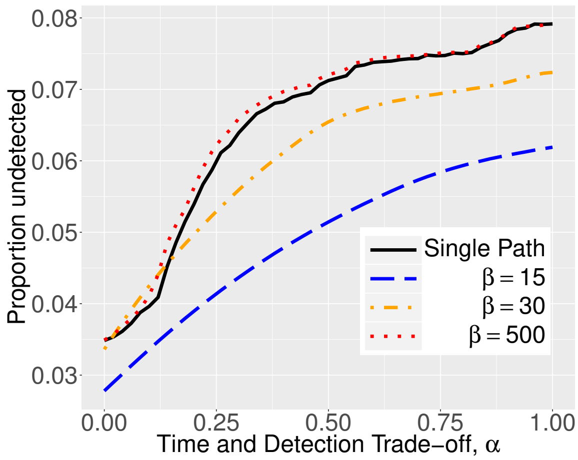


Figure 6.7: The proportion that goes undetected in the single path model and for different values of β in the multi-path model. We see that as β increases, the proportion undetected approaches that of the single path model.

We don't expect a small change in α to result in such a different resource allocation. Note that there is still some small noise in the allocations across α. This is due to the stochasticity in the GA used to find the optimal allocation.

The proportion of illegal traffic that goes undetected for various values of β is shown in Figure 6.7. Generally, as β increases, the proportion of illegal trade undetected also increases. This is because more traffic is directed along the smuggler's optimal path. For larger values of α, this is a path that minimises detection and so less trade is detection by biosecurity agencies.

The optimal allocations for $\beta = 15$ and 500 are quite different. However, there may be an allocation that results in proportions of undetected trade close to the optimal for both values of β. Given we don't know β, this would be an ideal allocation to implement in real life.

In the previous chapter we found proportions of trade that went undetected given fixed allocations across α. Recall each allocation was represented as the vector

$$(a_1, a_2, a_3, a_4, a_5),$$

where a_1, a_2, a_3, a_4, and a_5 are the allocations at Ho Chi Minh City, Hanoi, Da Nang, Nha Trang, and Can Tho respectively. The first robust allocation we found was to spread resources evenly between Ho Chi Minh City, Hanoi, and Da Nang. The second was to assign 2/5 of resources to Ho Chi Minh City and Hanoi, and 1/5 to Da Nang.

We find the proportions of trade that go undetected across α in the multi-path model for fixed resource allocations. Figure 6.8 shows these proportions for

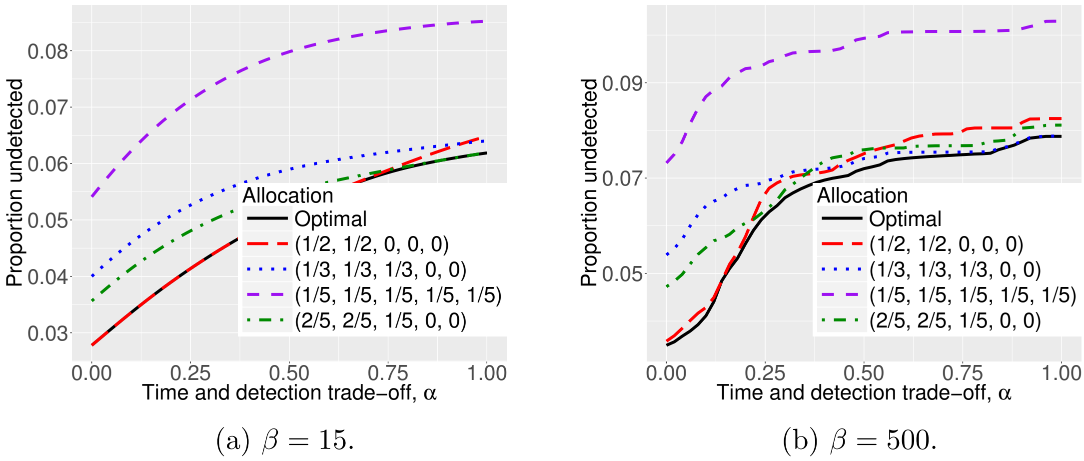


Figure 6.8: Proportions of illegal trade that go undetected for various allocations in the multi-path model when (a) $\beta = 15$ and (b) $\beta = 500$. The sensitivity of different allocations appear very similar for $\beta = 15$ and $\beta = 500$.

$\beta = 15$ and 500. Spreading resources evenly between Ho Chi Minh City and Hanoi, $(1/2, 1/2, 0, 0, 0)$, performs better than the other allocations when $\beta = 15$. It also performs well for more time-sensitive smuggler objectives (smaller α) when $\beta = 500$. For more detection-sensitive smugglers, the best allocation is to spread resources evenly between three nodes when $\beta = 500$. However, the allocation of spreading resources between Hanoi and Ho Chi Minh City still performs well here. This allocation may be robust for other values of β as well. It it effective for $\beta = 15$ and 500 even though they have quite different corresponding optimal allocations.

To see how these allocations perform for more values of β, we calculate the average difference between proportions of undetected illegal trade under the optimal and fixed allocations across α. This average proportion is

$$v^{\beta}(\boldsymbol{\delta}) = \frac{1}{51} \sum_{i=0}^{50} \left(f^{\beta}_{\alpha_i}(\boldsymbol{\delta}^*) - f^{\beta}_{\alpha_i}(\boldsymbol{\delta}) \right),$$

where $\alpha_i = i/50$, $\boldsymbol{\delta}^*$ is the optimal allocation given α_i, $\boldsymbol{\delta}$ is the fixed allocation and f is the proportion undetected given α_i, $\boldsymbol{\delta}$ and β in the multi-path model. Figure 6.9 shows these proportions for various values of β. The allocation that performs best across β is to spread resources evenly across Ho Chi Minh City and Hanoi. This is different to the robust allocations we found in the single path model. This is the case as in the multi-path model, we are allowing all paths into Vietnam to carry commodities. Ho Chi Minh City and Hanoi have the most paths into them. So, this allocation is best overall. With that being said, allocating a little more of the

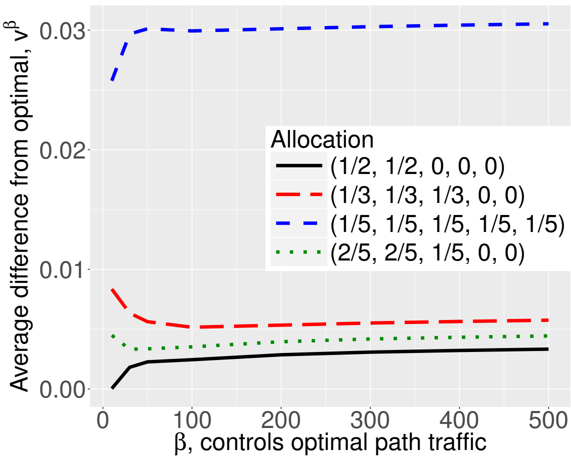


Figure 6.9: The average difference between proportion of undetected traffic under specific and optimal allocations. The most robust allocation is to assign resources evenly between Ho Chi Minh City and Hanoi. However, spreading resources evenly among Ho Chi Minh City, Hanoi, and Da Nang, is also fairly robust.

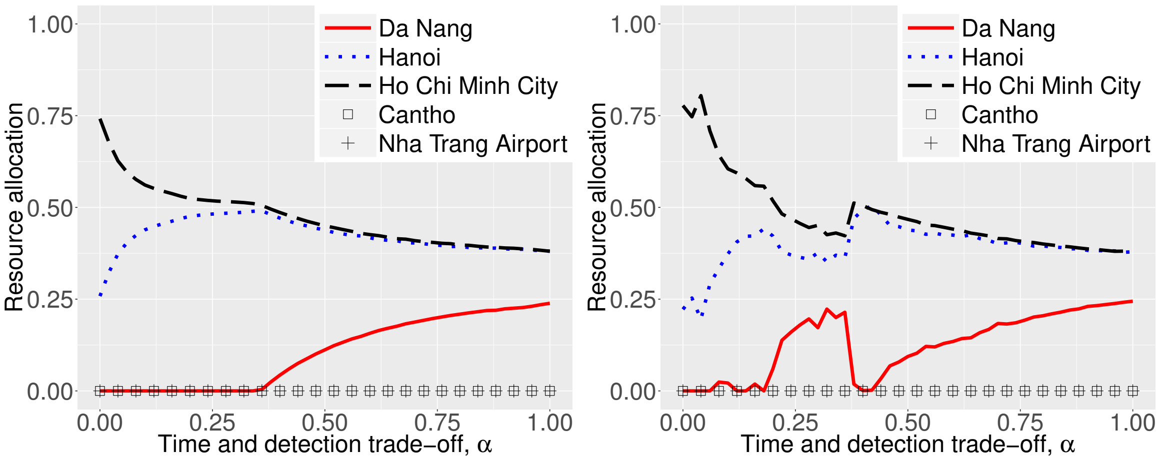
resources at Da Nang as well, and spreading resources evenly between Ho Chi Minh City, Hanoi, and Da Nang, do not appear too bad.

But how do these allocations perform on the other scenarios we tested in the previous section? Recall we looked at scaling node detection probabilities and non-uniform traffic from sources.

Figure 6.10 shows the optimal allocations given different traffic volumes from different sources in the multi-path model for $\beta = 200$. Traffic from sources were estimated in Section 5.4.1. As expected, the resource allocations are smoother across α. We no longer have the spike in allocation at Da Nang around $\alpha = 0.25$.

Figure 6.11 shows the difference between the optimal and the proportion of illegal trade that goes undetected for the fixed allocations given non-uniform source traffic in the multi-path model. Here we see that the most robust allocation is to assign half of your resources to Ho Chi Minh City and Hanoi. However, the allocation of 2/5 at these locations and 1/5 at Da Nang also performs well.

The detection probabilities in the previous simulations had a maximum of one. This is unrealistically high. It is unlikely that a node is able to intercept almost all illegal trade. Thus, we want to make sure these allocations are robust when we reduce the detection probabilities at nodes. Figure 6.12 shows the average difference across α between the optimal and fixed allocation proportions, when all node detection probabilities are reduced to 0.1 of their initial values. We see that the allocation of spreading resources between Ho Chi Minh City and Hanoi is still the most robust allocation.



(a) $\beta = 200$.   (b) Optimal.

Figure 6.10: Resource allocation with normalised weights of sources from traffic model for the case where (a) $\beta = 200$ and (b) when all smugglers take the same optimal path. There is no spike in allocation at Da Nang when $\beta = 200$.

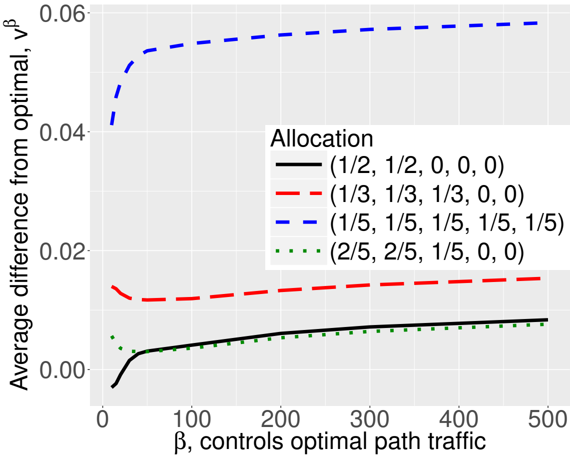


Figure 6.11: Difference between optimal and specific allocation proportion of undetected trade with non-uniform source traffic. The most robust allocations are to assign resources between Ho Chi Minh City and Hanoi, or 2/5 of resources at these locations and 1/5 at Da Nang.

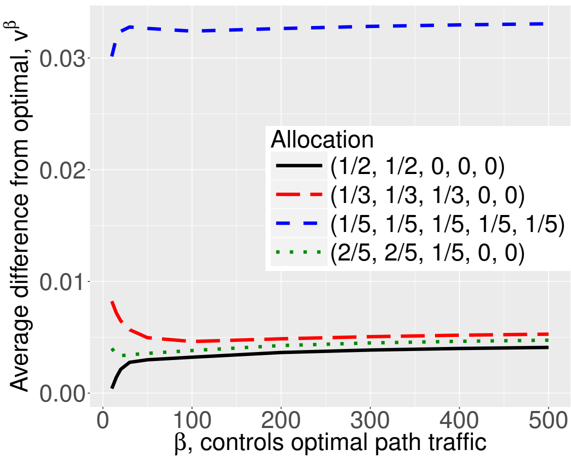


Figure 6.12: Average difference between optimal and fixed allocation proportions of undetected trade when all node detection probabilities are reduced to 0.1 of their initial value, taken across different smuggler objectives weights, α. The best allocation is to spread resources evenly between Ho Chi Minh City and Hanoi. This has proven to be the most robust allocation across the other tested scenarios as well.

6.3 Conclusions

We have extended our model to consider traffic from a source split along multiple paths, to account for some of the limitations of the single path model. We applied this model to the case study from Chapter 5 to investigate it in a practical setting. The parameter, β, in the model controls the proportion of traffic sent along the optimal path into the destination country from a source. The proportion of traffic along a path depended on the weight of that path.

We found that for smaller values of β, *i.e.*, when traffic was spread evenly across paths, the optimal solution was generally to assign resources evenly between two cities, Ho Chi Minh City and Hanoi. As β increased, and more traffic was sent along the optimal path, the optimal allocations for different smuggler objectives approached that of the single path model.

Previously, we saw that small changes in the smuggler's objective α, sometimes had very different resource allocations. Our new model does not exhibit this behaviour for smaller values of β. In the case study, we start to see jumps emerge in the optimal resource allocations once β reaches 500.

We found that the most robust allocation across β and α was to allocate resources evenly between Ho Chi Minh City and Hanoi. This was the case as all paths could potentially be used into Vietnam and these two locations had the most paths into them. This robust allocation was different to the ones found in the single path model.

In those allocations, more resources were assigned to Da Nang. When we assign 1/3 of the resources to Da Nang, the average proportion of undetected trade moves away from the optimal. However, the average difference in proportion of undetected illegal trade between the optimal allocation and spreading resources evenly between Ho Chi Minh City, Hanoi, and Da Nang was less than 0.01 for a wide range of β values. So, it still appears fairly robust in this model.

Chapter 7

Discussion and Future Work

Most enforcement activity to combat international wildlife trafficking takes place at ports of entry [56]. Resources are limited and thus must be allocated intelligently amongst these locations. Such resources influence the detection levels at ports. However, smugglers may respond to observed changes in detection by taking alternative routes into a country. We modelled the dynamic between smugglers and biosecurity agents in two models. In Chapter 3, we described the single path model: biosecurity forces place resources to minimise the probability that a smuggler goes undetected, and the smuggler takes their optimal route determined by some objective. We modelled their objective by a combination of minimising travel time and maximising the probability of non-detection. In Chapter 6, we described the multi-path model, where traffic from a source is distributed amongst paths into the destination country.

We described implementations of two algorithms, GA and INSH, that find the optimal resource allocations. These allocations provide insights into the best ways to intercept illegal wildlife trade across ports of entry.

The GA and INSH are heuristics, and so are not guaranteed to find the optimal allocation. Thus in Chapter 4, we tested the accuracy and convergence of the algorithms to find the optimal allocation in the single path model. We compared the algorithm solutions to the optimum found from a fine grid search in 400 networks. The algorithms both found solutions better or close to the optimum from the grid search. As well, they both converged quickly to solutions when we tested problems with different numbers of destination nodes. The GA, on average, converged to a better solution.

There is more testing that could be done to explore the properties of the algorithms. For instance, we tested our algorithms on random Waxman graphs. Perhaps more testing on different networks would be required for more confidence regarding convergence and accuracy. However, our goal was to perform actual tests on as realistic scenarios as possible, which we did in Chapters 5 and 6.

We also did not thoroughly investigate computational time. However, the real world networks for this problem were small so we found this not to be an issue. We also raised the idea of implementing a stopping rule so we wouldn't have to test for convergence for each new problem. One example could be to iterate until no new optimum has been found for some fixed number of iterations. It would also be interesting to test how the computational time increases with the number of sources. There were 189 sources in our case study. However, many sources were assigned almost no traffic when we considered non-uniform source traffic. When we removed these sources from the non-uniform source traffic problem, the optimal resource allocation was very similar to the allocation where all sources were included. So, we could potentially remove sources with little traffic if it greatly reduces computational run-time.

We investigated the interdiction model in a real world setting in Chapter 5. We constructed a case study motivated by the smuggling of pangolins into Vietnam from Asian and African sources. The pathways were constructed from flight data. We looked at airports in Vietnam that were within two hops from the sources. This gave rise to five entry points into Vietnam: Ho Chi Minh City, Hanoi, Da Nang, Can Tho, and Nha Trang. From this we determined the optimal biosecurity resource allocation across Vietnamese airports under different scenarios.

We observed different optimal allocations across various weightings of smuggler objectives, α. However, we rarely know where the smuggler sits on this spectrum. Luckily, we were able to determine a few robust allocations that performed well across α. One such robust allocation was to spread resources evenly between three out of five Vietnamese airports: Ho Chi Minh City, Hanoi, and Da Nang (with no additional resources at the other two nodes). These locations had far more paths into them than Can Tho and Nha Trang. Having more connections is likely to mean that the airports have a higher flow of traffic through them. Some studies [36, 49] reported that smugglers of particular wildlife species may rely on busier transport hubs. This may indicate that more resources should be placed at these locations. Our results largely support that. This is interesting as our method for reaching this conclusion was completely different. Rather than analysing seizure data, we considered a model that accounted for smugglers using alternative pathways. This has not been done before in this context. Though it makes logical sense, the allocation of spreading resources evenly across the three most connected nodes, was not always guaranteed to be robust. It may have been better to assign resources across all nodes. It would be interesting to develop more case studies to test whether this allocation is robust on other transport networks, as we expect it would be. We leave this for future work.

We could see how the routes, including transit locations, changed as the smuggler went from time sensitive to detection sensitive in the simulations. We found that the majority of locations that were used frequently as stopovers were hub airports. This was the case across all smuggler objectives. It is likely because the hubs had a high number of connections. This also agrees with studies [8, 49] that report transport hubs are exploited by traffickers.

The robust allocation was to spread 1/3 of the resources to Ho Chi Minh City, Hanoi, and Da Nang. We don't expect small perturbations from this allocation to result in great increases in undetected trade. The similar allocation of 2/5 of the resources at Ho Chi Minh City and Hanoi, and 1/5 at Da Nang, also appeared to be robust.

These robust allocations were determined assuming uniform traffic from 189 sources across Africa and Asia. We observed different optimal allocations when we considered non-uniform traffic from sources. The non-uniform source traffic affected the optimal resource allocations as a few sources had heavy traffic weightings. These sources were only connected to particular Vietnamese airports. Though, in this case study, the spreading of resources between the three nodes did not perform too poorly. This allocation intercepted around 5% less of the illegal trades than the optimal for time-sensitive smugglers. This was the greatest difference between the two allocations. In comparison, the allocation of spreading resources evenly across all five nodes intercepted around 10% less than the optimal allocation. If we were to use this model to inform policy, it would be best to consider traffic specific to the problem in order to get the best possible robust allocation across different smuggler objectives. Estimates of wildlife traffic from seizure data would be ideal rather than our model which used air traffic. Alternatively, interdiction of trades from sources may be weighted more heavily if the species survival is more at risk at certain locations.

The single path model provides some insights into optimal allocations but makes many assumptions and has limitations. Most importantly, it assumes that a smuggler takes their optimal route. We assume that the smuggler's objective is to either minimise travel time, maximise the probability of going undetected, or some combination of these two objectives. However, the smuggler may have alternative objectives that result in different paths. As well as that, if we consider multiple smugglers travelling from the same source, there is likely to be some variation in the paths they take. They may have different perceptions of model parameters such as detection. We introduced the multi-path model in Chapter 6 to account for some of these limitations. We then tried to find a robust allocation across different smuggler objective weights, α, as well as different multi-path scenarios. The most robust allocation was to spread resources evenly between two nodes: Ho Chi Minh City and Hanoi. These two airports had the most paths into them. The multi-path model assesses a smuggler potentially moving traffic along all of these paths. So it is not a surprise that the most robust allocation is to assign resources to both of these locations. Though, the allocation of spreading resources across the three main ports, Ho Chi Minh City, Hanoi, and Da Nang, is still fairly robust in this model.

There are definitely more paths into Vietnam than the ones we considered. We restricted the number of paths when we constructed the transport network from flight data. We constructed the underlying transport network by taking sources within two hops of the Vietnamese airports. However, we included all of the connections between sources, transits, and destinations. So this made it possible for paths to be more than two hops. Despite this, all the optimal paths were two hops or less. So even if we had included longer paths, they would likely not have been used. But there are other ports of entry that we could have considered. For example, sea ports and land border crossings. Wildlife contraband is also smuggled through these locations. Future work could investigate adding these ports of entry to the transport network. Perhaps detection probabilities depend on the type of port, or certain shipments are more likely to come through a particular form of entry point. We suspect that you would see similar robust allocations across various smuggler objectives: resources spread evenly among a subset of nodes that each have a higher number of connections.

On a similar note, we only considered paths of two hops in the multi-path model. However in the single path model it was possible to have longer paths. Perhaps if we evaluated paths with more hops, we would observe different optimal allocations. Although, the two airports in Ho Chi Minh City and Hanoi are still likely to have the most paths going into them. So our current allocation may also be robust when we consider paths of length greater than two hops. But this would be another extension to investigate.

Another interesting idea is to look at the trafficking of distinct wildlife commodities. It may be better for the smuggler to take a particular path for certain commodities. The different commodity types may have different detection probabilities. For instance, a large shipment of pangolin scales may have a different detection probability than a live pangolin. This extension would be fairly straightforward to incorporate into the existing model. However, if applied to the single path model, it would require evaluating the shortest path for each commodity as they have different detection probabilities. This would increase computational run-times. So perhaps more efficient methods are required to do this on a larger scale. The detection probabilities for various commodities would also need to be estimated.

Our case study was motivated by the trafficking of pangolins. The results and optimal allocations found may be different for other wildlife species travelling from Africa and Asia into Vietnam. The area of the case study that directly used pangolin seizure data was the calibration of the detection probability function to pangolin enforcement success rate estimates. These estimates were biased due to the role that countries played in pangolin trafficking incidents. For example, the U.S. had an enforcement success rate estimate of one, as it only appeared as a destination country in the seizure data. Thus, no illegal trade was observed to flow through it. In fact, the estimated enforcement success rates found from other seizure data sets [8, 49] were different. However, we still believe the calibrated detection probability function would be sufficient for capturing how detection increases with resources. But, we also used this function to determine the detection at source and transit nodes, using a country's CPI. These detection probabilities would likely be different for other species. We did determine the optimal allocation when Hong Kong's detection probability was decreased. The most robust allocation remained to be to allocate resources evenly among Ho Chi Minh City, Hanoi, and Da Nang. So perhaps this allocation will be robust to changes in transit node detection probabilities.

Future work may consider different perceptions between smugglers and enforcement agents of parameters such as detection probabilities. Smugglers and biosecurity forces are likely to have different views of these values. This may influence the optimal allocations. Similarly, we assumed that the smuggler changes path instantaneously once we set an allocation. This is not necessarily true. There may be some time between allocating resources and observing a change in the smuggler's behaviour. This may affect our values of the proportion of illegal trade that goes undetected.

Finally, an interesting extension could be to consider sharing resources between countries. If we consider two destination countries, is it better overall for each of them to assign resources to their own entry points, or does sharing resources decrease the total proportion of illegal trade that goes undetected?

Bibliography

[1] List of top 30 airports in the world. https://www.world-airport-codes.com/world-top-30-airports.html. Accessed: 23-11-2018.

[2] J. Ayling. What sustains wildlife crime? Rhino horn trading and the resilience of criminal networks. *Journal of International Wildlife Law & Policy*, 16(1):57–80, 2013.

[3] T. Back. *Evolutionary Algorithms in Theory and Practice: Evolution Strategies, Evolutionary Programming, Genetic Algorithms.* Oxford University Press, 1996.

[4] J. Baillie, D. Challender, P. Kaspal, A. Khatiwada, R. Mohapatra, and H. Nash. Manis crassicaudata. http://dx.doi.org/10.2305/IUCN.UK.2014-2.RLTS.T12761A45221874.en., 2014. The IUCN Red List of Threatened Species 2014: e.T12761A45221874.

[5] M.O. Ball and B.L. Golden. Finding the most vital arcs in a network. *Operations Research Letters*, 8(2):73–76, 1989.

[6] A. Butcher. Hong Kong increases wildlife smuggling jail terms. https://gbtimes.com/hong-kong-increases-wildlife-smuggling-jail-terms?media=42950, 2018. GBTIMES: Bringing China Closer.

[7] J.E. Byers and E.G. Noonburg. Poaching, enforcement, and the efficancy of marine reserves. *Ecological Applications*, 17(7):1851–1856, 2007.

[8] C4ADS. In plane sight: Wildlife trafficking in the air transport sector. Technical report, ROUTES Partnership report, authored by C4ADS, 2018.

[9] D. Challender, J. Baillie, G. Ades, P. Kaspal, B. Chan, A. Khatiwada, L. Xu, S. Chin, R. KC, H. Nash, and H. Hsieh. Manis pentadactyla. http://dx.doi.org/10.2305/IUCN.UK.2014-2.RLTS.T12764A45222544.en., 2014. The IUCN Red List of Threatened Species 2014: e.T12764A45222544.

[10] D. Challender and and JEM. Baillie C. Waterman. Scaling up pangolin conservation. Technical report, IUCN SSC Pangolin Specialist Group Conservation Action Plan. Zoological Society of London, London, UK., 2014.

[11] D. Challender, T. Nguyen Van, C. Shepherd, K. Krishnasamy, A. Wang, E. Panjang B. Lee, L. Fletcher, S. Heng, J. Seah Han Ming, A. Olsson, A. Nguyen The Truong, Q. Nguyen Van, and Y. Chung. Manis javanica. `http://dx.doi.org/10.2305/IUCN.UK.2014-2.RLTS.T12763A45222303.en.`, 2014. The IUCN Red List of Threatened Species 2014: e.T12763A45222303.

[12] CITES. What is CITES? `https://cites.org/eng/disc/what.php`.

[13] R. Damania, E.J. Milner-Gulland, and D.J. Crookes. A bioeconomic analysis of bushmeat hunting. *Proceedings of the Royal Society B: Biological Sciences*, 272(1560):259–266, 2005.

[14] F. Fang, P. Stone, and M. Tambe. When security games go green: Designing defender strategies to prevent poaching and illegal fishing. In *Proceedings of the 24th International Conference on Artificial Intelligence*, IJCAI'15, pages 2589–2595. AAAI Press, 2015.

[15] J.D. Flora, R.L. Copp, and D Tholen. Use of weighted least squares in relating component outage rates to age of vehicle. *Accident Analysis and Prevention*, 12(2):119–129, 1980.

[16] D. Fudenberg and J. Tirole. *Game Theory.* Massachusetts Institute of Technology, 1991.

[17] D.R. Fulkerson and G.C. Harding. Maximising the minimum source-sink path subject to a budget constraint. *Mathematical Programming*, 13:116–118, 1977.

[18] S.N. Godoy and E.R. Matushima. A survey of diseases in passeriform birds obtained from illegal wildlife trade in so paulo city, brazil. *Journal of Avian Medicine and Surgery*, 24(3):199–209, 2010.

[19] B. Golden. A problem in network interdiction. *Naval Research Logistics Quarterly*, 25:711–713, 1978.

[20] Q. Guo, B. An, Y. Zick, and C. Miao. Optimal interdiction of illegal network flow. In *Proceedings of the Twenty-Fifth International Joint Conference on Artifical Intelligence*, pages 2507–2513, 2016.

[21] S. Heinrich, T.A. Wittman, J.V. Ross, C.R. Shepherd, D.W.S. Challender, and P. Cassey. The global trafficking of pangolins: A comprehensive summary of seizures and trafficking routes from 2010-2015. Technical report, TRAFFIC, Southeast Asia Regional Office, Petaling Jaya, Selangor, Malaysia., 2017.

[22] M.H. Holden, D. Biggs, H. Brink, P. Bal, J. Rhodes, and E McDonald-Madden. Increase anti-poaching law-enforcement or reduce demand for wildlife products? A framework to guide strategic conservation investments. *Conservation Letters*, 0(0):e12618.

[23] E. Israeli and K.R. Wood. Shortest-path network interdiction. *Networks*, 40(2):97–111, 2002.

[24] IUCN. What does the new trade ban mean for pangolin conservation? `https://www.iucn.org/news/secretariat/201610/what-does-new-trade-ban-mean-pangolin-conservation`. Accessed: 23-11-2018.

[25] M. Jain, V. Conitzer, C. Laohaphan, and M. Tambe. Security scheduling for real-world networks. In *Proceedings of the Workshop on Multiagent Interaction Networks*, pages 215–222, 2013.

[26] M. Kabir, R. Salim, and N. Al-Mawali. The gravity model and trade flows: Recent developments in econometric modeling and empirical evidence. *Economic Analysis and Policy*, 56:60–71, 2017.

[27] J. Kurland and S.F. Pires. Assessing U.S. wildlife trafficking patterns: How criminology and conservation science can guide strategies to reduce the illegal wildlife trade. *Deviant Behavior*, 38(4):375–391, 2017.

[28] L. Lagrada, S. Schoppe, and D. Challender. Manis culionensis. `http://dx.doi.org/10.2305/IUCN.UK.2014-2.RLTS.T136497A45223365.en.`, 2014. The IUCN Red List of Threatened Species 2014: e.T136497A45223365.

[29] K. Malik. The k most vital arcs in the shortest path problem. *Operations Research Letters*, 8(4):223–227, 1989.

[30] S. McCarthy, M. Tambe, C. Kiekintveld, M. L. Gore, and A. Killion. Preventing illegal logging: Simultaneous optimization of resource teams and tactics for security. In *Proceedings of the Thirtieth AAAI Conference on Artifical Intelligence (AAAI)*, pages 3880–3886, 2016.

[31] D. P. Michalopoulos, D. P. Morton, and J. W. Barnes. Prioritizing network interdiction of nuclear smuggling. In H. I. Gassmann, S. W. Wallace, and W. T. Ziemba, editors, *Stochastic Programming: Applications in Finance, Energy and Logistics*. World Scientific, 2013.

[32] D.P. Morton, F. Pan, and K.J. Saeger. Models for nuclear smuggling interdiction. *IIE Transactions*, pages 3–14, 2005.

[33] M.J. Osborne and A. Rubinstein. *A Course in Game Theory*, volume 1. The MIT Press, January 1994.

[34] F. Pan and D.P. Morton. Minimising a stochastic maximum-reliability path. *Networks*, 52(3):111–119, 2008.

[35] S. Patel and S.Y. Chin, editors. *Proceedings of the Workshop on Trade and Conservation of Pangolins Native to South and Southeast Asian, Singapore Zoo*, 2008.

[36] G.A. Petrossian, N. Marteache, and J. Viollaz. Where do "undocumented" fish land? an empirical assessment of port characteristic for IUU fishing. *European Journal on Criminal Policy and Research*, 21:337–351, 2015.

[37] D. Pietersen, C. Waterman, L. Hywood, P. Rankin, and D. Soewu. Smutsia temminckii. `http://dx.doi.org/10.2305/IUCN.UK.2014-2.RLTS.T12765A45222717.en.`, 2014. The IUCN Red List of Threatened Species 2014: e.T12765A45222717.

[38] D.J. Price, N.G. Bean, J.V. Ross, and J. Tuke. An induced natural selection heuristic for evaluating optimal bayesian experimental designs. *Computational Statistics and Data Analysis*, 126:112–124, 2018.

[39] P. Reuter and D. ORegan. Smuggling wildlife in the americas: scale, methods, and links to other organised crimes. *Global Crime*, 18(2):77–99, 2017.

[40] J.A Rice. *Mathematical Statistics and Data Analysis*. Thomson Brooks/Cole, 2007.

[41] C. Rocco and J.E. Ramirez-Marquez. Deterministic network interdiction optimization via an evolutionary approach. *Reliability Engineering & System Safety*, 94:568 – 576, 2009.

[42] G.E. Rosen and K.F. Smith. Summarizing the evidence on the international trade in illegal wildlife. *EcoHealth*, 7(1):24–32, 2010.

[43] N. South and T. Wyatt. Comparing illicit trades in wildlife and drugs: An exploratory study. *Deviant Behavior*, 32(6):538–561, 2011.

[44] M. Srinivas and L. M. Patnaik. Adaptive probabilities of crossover and mutation in genetic algorithms. *IEEE Transactions on Systems, Man, and Cybernetics*, 24(4):656–667, 1994.

[45] K. Thulasiraman and M. N. S. Swamy. *Graphs: Theory and Algorithms*. John Wiley and Sons, INC, 2011.

[46] J. Tsai, Z. Yin, J. Kwak, D. Kempe, C. Kiekintveld, and M. Tambe. Urban security: Game-theoretic resource allocation in networked physical domains. In *Twenty-Fourth Conference on Artificial Intelligence (AAAI)*, pages 881–886, 2010.

[47] UNEP. Stolen apes. the illicit trade in chimpanzees, gorillas, bonobos and orangutans. a rapid response assessment. Technical report, United Nations Environment Programme. GRID-Arendal, 2013.

[48] UNODC. World wildlife crime report: Trafficking in protected species. Technical report, 2016.

[49] M. Utermohlen and P. Baine. Flying under the radar: Wildlife trafficking in the air transport sector. Technical report, C4ADS and Reducing Opportunities for Unlawful Transport of Endangered Species (ROUTES). USAID, 2017.

[50] C. Waterman, D. Pietersen, L. Hywood, P. Rankin, and D. Soewu. Smutsia gigantea. `http://dx.doi.org/10.2305/IUCN.UK.2014-2.RLTS.T12762A45222061.en`, 2014. The IUCN Red List of Threatened Species 2014: e.T12762A45222061.

[51] C. Waterman, D. Pietersen, D. Soewu, L. Hywood, and P. Rankin. Phataginus tetradactyla. `http://dx.doi.org/10.2305/IUCN.UK.2014-2.RLTS.T12766A45222929.en`, 2014. The IUCN Red List of Threatened Species 2014: e.T12766A45222929.

[52] C. Waterman, D. Pietersen, D. Soewu, L. Hywood, and P. Rankin. Phataginus tricuspis. `http://dx.doi.org/10.2305/IUCN.UK.2014-2.RLTS.T12767A45223135.en`, 2014. The IUCN Red List of Threatened Species 2014: e.T12767A45223135.

[53] B. Waxman. Routing of multipoint connections. *IEEE Journal on Selected Areas in Communications*, 6(9):1617–1622, 1988.

[54] D. Whitley. A genetic algorithm tutorial. *Statistics and Computing*, 4(2):65–85, 1994.

[55] K.R. Wood. Deterministic network interdiction. *Mathematical and Computer Modelling*, 17(2):1–18, 1993.

[56] World Bank. Tools and resources to combat illegal wildlife trade. Technical report, World Bank Group, Washington, D.C., 2018.

[57] T. Wyatt, K. Johnson, L. Hunter, R. George, and R. Gunter. Corruption and wildlife trafficking: Three case studies involving Asia. *Asian Journal of Criminology*, 13(1):35–55, 2018.

[58] D. Xu and M. Chiang. Link-state routing with hop-by-hop forwarding can achieve optimal traffic engineering. In *Proceedings IEEE/ACM Transactions on Networking*, 2011.

[59] R. Yang, B. Ford, M. Tambe, and A. Lemieux. Adaptive resource allocation for wildlife protection against illegal poachers. In *Proceedings of the 2014 International Conference on Autonomous Agents and Multi-agent Systems*, pages 453–460, 2014.