Decentralised Decision Making in Defence Logistics

Slava Shekh
Defence Science and Technology Group
slava.shekh@dsto.defence.gov.au

Michelle Blom
The University of Melbourne
michelle.blom@unimelb.edu.au

Abstract

We anticipate that future Defence logistics will involve increasingly sophisticated supply chains with a greater reliance on autonomous systems. At the same time, Defence environments are shrouded in uncertainty and pose a considerable challenge to the successful operation of these systems. We present and discuss a variety of techniques for decentralised decision making in Defence logistics environments. Cognitive partnerships are proposed as the next step towards improving decision making and harnessing the strengths of human-machine collaboration for logistics applications.

1 Introduction

Defence supply chains are often structured in point-to-point arrangements, where goods are distributed along linear pathways from origin to destination. These supply chains are designed with simplicity in mind to minimise the cognitive and physical burden on human operators. Even so, it is often challenging for humans to maintain complete situational awareness and control of the logistics flow, and the occurrence of stockout events (failure to meet demand for inventory) is already a reality [Yoho *et al.*, 2013].

With the increasing adoption and pervasiveness of autonomous systems [Murphy and Shields, 2012], we anticipate that future Defence logistics will have a strong reliance on autonomy. This will enable the use of increasingly sophisticated supply chains and distribution networks. A logistics network offers diversity and redundancy in the supply routes, as well as the ability to stockpile supplies at multiple locations throughout the network in preparation for possible future contingencies.

Defence logistics environments are fundamentally dynamic and uncertain. Any autonomous system operating in such an environment must be able to make sound decisions in the face of uncertainty [Scholz and Reid, 2015]. More explicitly, autonomous logistics systems should be able to deal with the temporary or permanent loss of nodes and arcs in the network, the loss of transportation systems, and an underlying uncertainty in the future demand of consuming nodes. Failure to operate under these conditions will inhibit the deployment and use of autonomous systems in Defence logistics.

If autonomous systems are to succeed in the presence of uncertainty, the use of decentralised decision making is critical. Failure of specific system entities (e.g. nodes and transports) should have limited impact on other system entities and on their ability to make robust decisions. By decentralising the decision making to the entities themselves, each part of the system operates in a relatively independent way and is less likely to be affected by system failures. Even under circumstances where communication between transports and nodes is severely impaired, if the entities can function independently, then the overall system is more resilient and mission requirements can continue to be satisfied with a diminishing pool of resources.

In a setting where decision making is fundamentally decentralised, the autonomous entities themselves could be machines, humans, or some combination of both. Perhaps the most promising area of autonomy is this final case, where human and machine capabilities are combined to form cognitive partnerships for problem solving [Lange *et al.*, 2013]. As we will discuss, many of our decentralised logistics algorithms have been developed based on our human experiences and intuitions. Cognitive partnerships would allow future decision making to arise from a collective of human and machine intelligence based on intuitions that may not be realisable by either entity operating alone.

2 The Network Distribution Problem

We investigate network-based logistics through a real-time network distribution problem (NDP). This is an extension of the more well-known inventory routing problems (IRPs) [Bell *et al.*, 1983] that form a core component of supply chain optimisation. An IRP is defined on a graph, in which one or more suppliers are connected, via arcs and intermediate nodes, to a number of consumers, each with inventory holding capacities and costs. Each consumer consumes inventory at a rate that potentially varies over the course of some time horizon. The problem is to determine how much inventory to produce at each supplier, the quantity of inventory to deliver to consumers across the horizon, and the manner in which this inventory is transported. The objective is to minimise routing and storage costs while meeting demand at consumers.

The NDP is a dynamic inventory routing problem with a single supplier (with an infinite supply of a single resource), and one or more consumers (sink nodes). A basic instance of the NDP is shown in Figure 1. Each sink consumes resource at a rate (units/day) that is known a priori. At unknown times, however, this rate increases significantly for a period of time. The timing, size, and duration of these surges become known only a short time before they arise. A finite set of transports (with varying capacity and speed) are required to continually transport inventory from the supplier to each sink, throughout the horizon, in a way that prevents stockouts (where there is insufficient inventory to meet demand) from occurring at sink nodes. Inventory may be delivered to any node that has capacity to store it (for stockpiling), and collected from any (non-sink) node at which it has been stockpiled. Our sole objective is to minimise stockouts arising at sinks (costs are not modelled).

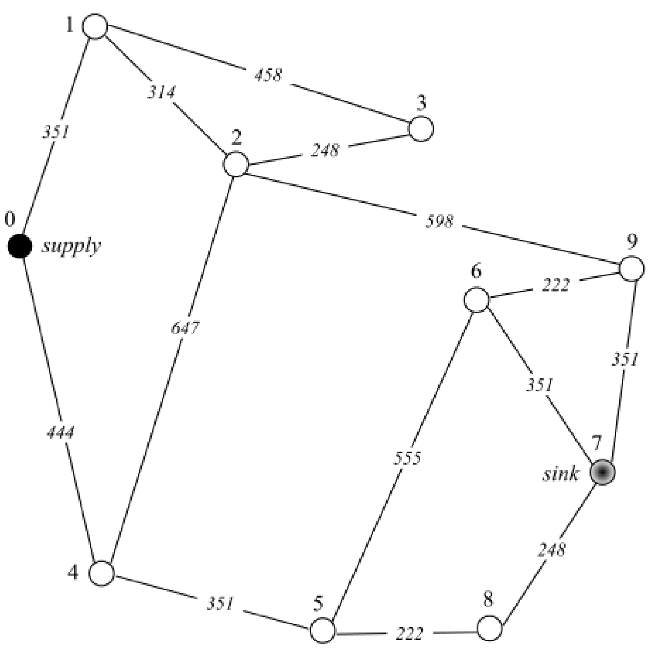

Figure 1: An instance of the network distribution problem with 10 nodes, including 1 supply and 1 sink. Each link is weighted based on the travel distance required to traverse that link.

The NDP described above has the key elements of a logistics problem (now referred to as the base NDP), but it does not represent a realistic military scenario. Therefore, we now describe two NDP variants that involve the same underlying problem but contain some additional features which provide a stronger link to Defence logistics.

2.1 Fuel Scenario

One NDP variant involves a fuel distribution network supplying several military bases with refined petroleum products. The crude oil is extracted at one or more offshore oil platforms and transported to onshore oil terminals using marine oil tankers (and assumed to be refined somewhere in the process). Road tankers are then used to distribute the fuel through a network of fuel farms for delivery to one of several military bases that have an ongoing demand for fuel. All transports and routes can experience short periods of downtime ranging from a few hours to a few days. An opposing force is also present nearby, which threatens the distribution network and causes surges in fuel demand for friendly forces with limited warning time. Compared to the base NDP, this scenario is more challenging due to the presence of an active adversary, multiple supply nodes and multiple transport types that need to cooperate, which increase the complexity of the decision making.

2.2 HADR Scenario

Another scenario involves Defence providing humanitarian assistance and disaster relief (HADR) in response to an international humanitarian crisis. Relief supplies need to be pre-positioned at mounting bases using road transports, and then delivered to international relief centres using a combination of air and sea transports. There are three types of relief supplies (food, medical supplies and fuel) and the relief centres have a combination of known and unknown (surge) demand over the course of the operation. Akin to previous scenarios, all transports and routes can experience short periods of downtime. Decision making in the HADR scenario requires more intricate logistics planning due to the diversity of supply items (each of which can be sourced from a different location) and an increased number of transport types with vastly different capabilities (such as mode, speed, and capacity).

3 Autonomous Decision Making

The scenarios described above demonstrate the kind of logistics environments in which decentralised decision making is being considered, yet these examples are not exhaustive. Future scenarios may involve other combinations of features (e.g. multiple supply items in an adversarial setting) or other features altogether (e.g. limits on arc throughput). In general, we cannot envisage all possible features that future scenarios might have. Thus, our primary goal is to explore techniques which can provide high generalisation performance across a range of logistics scenarios. Finding optimal solutions within those scenarios is currently a secondary consideration.

A key requirement of any technique we consider is its ability to function in a decentralised setting. If the decision making entities are fundamentally decentralised and operating independently, there is a higher chance that the overall system can cope with unforeseen future events and system failures. In addition, some logistics environments could be partially observable, meaning that each entity will maintain its own belief state (a representation of the state of the network and the expected deliveries of other transports), and information will always be decentralised.

In the following subsections, we discuss a variety of techniques for decision making in these challenging logistics environments. Initially, we consider mathematically rigorous approaches, by modelling the problem as a Markov decision process (MDP) and a mixed-integer program (MIP), but show that these methods are not practical for the NDP. This is followed by a discussion of general heuristic techniques and other approximate methods.

3.1 Markov Decision Process

An MDP describes a system in terms of: a set of states; actions that can be performed in those states; the probabilistic transitions that occur as actions are performed in each state; and a reward function that prescribes the utility or reward received when actions are performed in each state. The goal is to find a policy that maps each state to the best action to perform in that state. An optimal policy maximises the total expected sum of rewards received by following the policy.

To model the base NDP as an MDP, the state space must consider: the inventory levels at each node, the state of each arc (active or unavailable), and the location and inventory of each transport. A typical instance of the NDP could contain 15 nodes (each with 30 units capacity), 10 arcs, and 4 transports (10 units capacity). Even in a simplified setting in which each transport action takes a single time period, the number of required states equals $30^{15} \times 2^{10} \times 15^{4} \times 10^{4}$. This is a prohibitively large number and therefore modelling the NDP as an MDP is unlikely to be practical.

3.2 Mixed-Integer Programming

A mathematical optimisation method such as mixed-integer programming could be used to model and solve a static instance of the NDP, where the overall network state and demands are known. Using the model of Savelsbergh and Song [2008] as a starting point, we create a planner that solves a MIP each time a transport reaches a decision point.

To assess whether such a planner is feasible, we apply the MIP to a simple instance of the base NDP containing 15 nodes, 1 sink and 1 transport. Unfortunately, since our sole objective is to minimise stockouts, the MIP produces solutions without stockouts but with other undesirable behaviour, such as transports moving to arbitrary nodes instead of delivering inventory. We could continually refine the constraints and objectives to guide the MIP towards more desirable solutions, but this will only increase its complexity. Even for this static scenario, the MIP already contains hundreds of thousands of constraints and takes several minutes to compute. For a dynamic scenario, the computation time could be much higher, because the MIP would need to be recomputed across many realisations of uncertainty. Larger and more realistic scenarios will require even more computation time, so using a MIP to solve the NDP quickly becomes infeasible.

Although the MIP was unsuccessful in solving the entire NDP, it has been successfully used to select stockpile locations. As we will discuss in the next section, one of the key decisions for an algorithm is selecting suitable nodes in the network to use as stockpiles in preparation for unknown future demand (surge). Using a MIP, we can solve a maximum set covering problem, where stockpiles are selected to provide maximum coverage to the sink nodes (in terms of storage and proximity).

3.3 Heuristic Techniques

Due to the drawbacks of using rigorous methods (MDPs and MIPs) for NDP decision making, we now consider heuristic techniques. Since NDP scenarios can vary greatly, different heuristics will be ideal for different scenarios. Our end goal is to combine these heuristics into a heuristic framework, which automatically selects the best method to use in a given situation, and therefore provides a way of generalising the heuristics to multiple classes of logistics problems.

All heuristic methods have been implemented in a decentralised way, where each transport uses a specified heuristic to makes its own decisions over time. Each transport and node also maintains a belief of the state of all other nodes in the network and of the intentions (expected future deliveries) of all other transports. In partially observable environments, these belief states are updated through the interaction of nodes and transports that are situated at the same location. In fully observable environments, all transports and nodes are assumed to have complete knowledge.

Heuristics are evaluated by using them across many instances of the same scenario. Each instance is generated with different realisations of uncertainty via Monte Carlo simulation. Heuristics that lead to fewer stockouts across these instances are considered to be superior. The detailed experimental results are not included in this paper due to space constraints, but we provide a summary of the types of scenarios where each heuristic performed best.

Shuttling-Based Planners

The first heuristic is a shuttling method that instructs each transport to collect resource from a supply node, and deliver it to the most-in-need sink. The most-in-need sink is the sink that will have the least inventory upon arrival of the transport (on the basis of known demand and scheduled deliveries, while not considering unknown future events). Upon delivery of the inventory, each transport returns to a supply node to collect another load. If a sink does not have the storage capacity to accept a transport's inventory, the transport is sent to the most-in-need of the alternate sinks. Our experiments suggest that shuttling heuristics are effective in small point-to-point networks with limited storage locations between sources and sinks.

Surge-Based Planners

Our next heuristic adapts the shuttling method by instructing transports that have arrived at a sink with no room for additional inventory to return to a supply node, delivering spare inventory to intermediate nodes on its return journey. In periods of surge activity, when one or more sinks is experiencing a surge demand, transports are allowed to collect inventory from these stockpiles for delivery to a sink. Variants of this surge-based heuristic include allowing transports to wait at at-capacity sink nodes until they are able to deliver their inventory, or travelling to an alternate sink to attempt a delivery (instead of immediately returning to a supply node). Each surge-based planner also has a smart stockpiling variant, which uses a MIP to pre-compute stockpile locations that provide maximum coverage to sinks. Our experiments suggest that surge-based planners are effective in large networks with many interconnected nodes, where they consistently achieve fewer stockouts than shuttling-based planners.

Communications Rover

In partially observable environments, each transport maintains its own belief state and these belief states are only updated through the interaction of nodes and transports that are situated at the same location. In this setting, the previous planners can be extended by allocating one transport as a 'communications rover'. This rover repeatedly travels between key nodes (sources and sinks) and other transports, spreading information about node states and transport intentions throughout the network. Ensuring all nodes and transports are regularly visited by the communications rover is one way of mitigating against the constraints imposed by partial observability.

Push and Pull Variants

Another strategy is to allocate different heuristics to different transports. One approach that could be effective is to allocate some subset of the transports as 'stockpile pushers' and the remainder as 'stockpile pullers'. Stockpile pushers would collect resources from source nodes and deliver them to pre-computed stockpile locations (using the MIP approach described above), while stockpile pullers would select suitable stockpile locations to collect inventory from and deliver that inventory to the most-in-need sink based on current priorities. Further experiments are required to evaluate the effectiveness of this approach.

3.4 Monte Carlo Tree Search

Given an MDP, Monte Carlo Tree Search (MCTS) methods provide an online mechanism for selecting actions to perform in any given state. This is achieved by building a tree of possible future states that can be reached from the current state and assigning values to states based on expected future reward. The NDP decision tree could be constructed based on the actions available to transports at each decision point. One of the challenges is determining how to evaluate the value of certain states. Using the standard MCTS approach (Monte Carlo simulation) has shown to be problematic for the NDP in terms of runtime. We are now exploring other methods, such as using the degree of stockout or total inventory at sinks, as a way of evaluating states without simulation, which could make MCTS viable for NDP decision making.

4 Towards Human-Machine Collaboration

Our future work will focus on further refining existing algorithms and exploring other techniques for decision making. In particular, MCTS has shown promising initial results but further work is required to determine whether it is a viable technique for the NDP.

The current results show that heuristics are an effective tool for decision making in current NDP scenarios, so a logical step would be to combine them into a heuristic framework with more general decision making capabilities. Given a new (unseen) situation, knowledge of similar situations could be used to inform the decision. Although heuristics are based on human intuitions, combining them into a decision making framework for new scenarios could lead to novel approaches that human decision makers may not have considered. If the framework is an extension of human decision making, the human operators are also more likely to understand and trust the autonomous system.

A natural extension of this heuristic framework is to explore cognitive partnerships in support of human-machine decision making. Consider the example where the NDP (as presented in this paper) does not exist in isolation, but is part of a complex logistics system made up of human and machine entities. In a scenario with constrained communication, the humans may rely on machines to provide critical information during the logistics planning process, which may be difficult or time consuming for the humans to obtain on their own. Establishing cognitive partnerships between humans and machines would provide an effective mechanism for sharing this information and establish a critical role for machines in the decision making process. In particular, logistics plans could be developed in a shared capacity, where the core plan is created by human logisticians and parts of the plan are then augmented with input and recommendations from machines based on a shared understanding of the situation. These kinds of human-machine partnerships could provide a way of harnessing the strengths of all participants to provide enhanced decision making and increasingly robust logistics plans.

Acknowledgments

This project has been funded through a Research Agreement between DST Group (Australia) and the University of Melbourne on "Decentralised Algorithms for the Network Distribution Problem" (MYIP # 6810). The authors would like to thank Don Gossink, Miguel Ramirez, Tim Miller, Adrian Pearce, Nir Lipovetzky and Liz Sonenberg for their contributions to the project.

References

[Bell *et al.*, 1983] Bell, W. J., Dalberto, L. M., Fisher, M. L., Greenfield, A. J., Jaikumar, R., Kedia, P., Mack, R. G., & Prutzman, P. J. Improving the distribution of industrial gases with an on-line computerized routing and scheduling optimizer. *Interfaces*, 13(6), 4-23, 1983.

[Lange *et al.*, 2014] Lange, D. S., Gutzwiller, R. S., Verbancsics, P., & Sin, T. Task models for human-computer collaboration in supervisory control of teams of autonomous systems. *Cognitive Methods in Situation Awareness and Decision Support*, 97-102, 2014.

[Murphy and Shields, 2012] Murphy, R., & Shields, J. The role of autonomy in DoD systems. *Defense Science Board Task Force Report*, 2012.

[Savelsbergh and Song, 2008] Savelsbergh, M., & Song, J. H. An optimization algorithm for the inventory routing problem with continuous moves. *Computers & operations research*, 35(7), 2266-2282, 2008.

[Scholz and Reid, 2015] Scholz, J., & Reid, D. So, where are all the robots? A rationale for strategic research. *DST Group Research Report 0420*, 2015.

[Yoho et al., 2013] Yoho, K. D., Rietjens, S., & Tatham, P. Defence logistics: an important research field in need of researchers. *International Journal of Physical Distribution & Logistics Management*, 43(2), 2013.