

Analysis of Fleet Modularity in an Artificial Intelligence Based Defender-Attacker Game

Xingyu Li¹, Bogdan I. Epureanu¹

^a *Department of Mechanical Engineering
University of Michigan, Ann Arbor*

Abstract

Because combat environments change over time and technology upgrades are widespread for ground vehicles, a large number of vehicles and equipment become quickly obsolete. A possible solution for the U.S. Army is to develop fleets of modular military vehicles, which are built by interchangeable substantial components also known as modules. One of the typical characteristics of module is their ease of assembly and disassembly through simple means such as plug-in/pull-out actions, which allows for real-time fleet reconfiguration to meet dynamic demands. Moreover, military demands are time-varying and highly stochastic because commanders keep reacting to enemy's actions. To capture these characteristics, we formulated an intelligent agent-based model to imitate decision making process during fleet operation, which combines real-time optimization with artificial intelligence. The agents are capable of inferring enemy's future move based on historical data and optimize dispatch/operation decisions accordingly. We implement our model to simulate an attacker-defender game between two adversarial and intelligent players, representing the commanders from modularized fleet and conventional fleet respectively. Given the same level of combat resources and intelligence, we highlight the tactical advantages of fleet modularity in terms of win rate, unpredictability and suffered damage.

Keywords:

Multi-agent systems, Decision processes, Artificial intelligence, Strategic planning

1. Introduction

Military vehicles operate in a large variety of environments and scenarios resulting in a diverse set of requirements from fleet mix. The special functionalities of military vehicles and incessantly updated technologies make them hardly reusable after military operation (Shinkman, 2014). In order to reduce the wastage, the US Army requires that fleets of vehicles can be reutilized across a large array of military mission scenarios. Modular vehicles are then introduced by constructing from special components, which are named as modules (Dasch and Gorsich, 2016). Modules are assumed to be a special type of components which

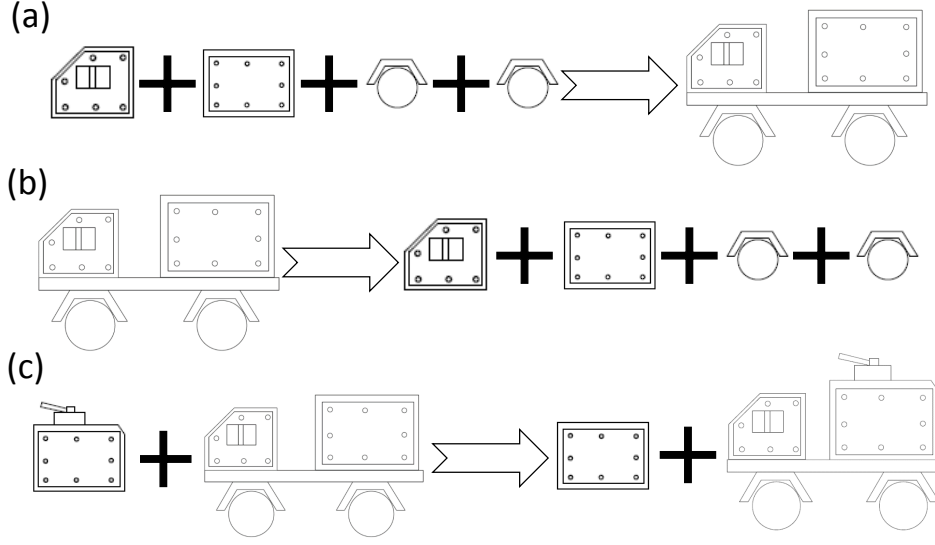


Figure 1: Operation actions for modularized vehicles: a) assembly, b) disassembly, c) reconfiguration.

can be easily coupled/decoupled through simple plug-in/pull-out actions on battle fields. This property enables vehicles to be quickly assembled, disassembled and reconfigured (ADR) (as shown in Fig.1) on field in reacting to demands.

Because of the close connection between operation strategy and fleet performance, researchers start investigating the potentials of modularity in the sights of performance boost during fleet operation. Bayrak et al. proposed a mathematical model to simulate modular fleet operation in a logistic mission scenario. They noticed a significant operational cost reduction after fleet modularization (Bayrak et al., 2016). In 2017, Li and Epureanu (Li and Epureanu, 2018) proposed an intelligent agent-based model in managing modular fleet operation. Agents are classified into three categories: camp, distributor, and supply. Different types of agents collaboratively and real-timely yield operational decisions in reacting to stochastic field demands (Li and Epureanu, 2017a). Later, Li and Epureanu also modeled the fleet operation as a dynamic system and implemented model predictive control to manage the system dynamics. Their results show that modular fleet exhibits a better robustness than conventional fleet in reacting to the disturbance and noise from battlefields (Li and Epureanu, 2017b).

However, most of previous research focused on a single decision maker who operates (play a game) against the environment. However, unpredictability from the enemy's reaction is essential and non-negligible in every form of warfare, which leads to an inability to forecast the outcome of actions or weakly perceived causal links between the events on the battlefield (Lynch, 2015). Furthermore, smart

system and artificial intelligence are playing an ever-increasing role in our daily lives. This trend does not spare military operations. Autonomous vehicles, especially Unmanned Aerial Vehicles, have been widely used to assist military operations (Landa, 1991; Jose and Zhuang, 2013; Evers et al., 2014). With no surprise, artificial intelligence will play a significant role in management of a large scale fleet of autonomous vehicles. Given the equivalent autonomous decision making system, our goal is to emphasize the synergy between modularity and autonomy by performing an attacker-defender game between conventional fleet and modular fleet.

The use of games in modeling the relationship between an attacker and a defender has a long history starting with the work of Dresher (Dresher, 1961). The variety of applications and research relate to issues in military OR and defense studies are rich (Kardes and Hall, 2005; Hausken and Levitin, 2009; Zhuang et al., 2010; Paulson et al., 2016). There are also several studies of attacker-defender games that consider the resource dependent strategies. Power use game theoretical approach for finding a defender’s resource allocation strategy for protecting sets from being destroyed by a strategic adversary, i.e., terrorist group. The defender is uncertain about the targets that enemy is likely to strike. Given the distribution of enemy’s behavior, they derived a pure strategy for defender that leads to a Bayesian Nash equilibria (Powell, 2007). Hausken and Zhuang consider a multi-period game where a defender can allocate resources to both defend its own resources and attack the resources of the enemy. Similarly, attacker can also determine the use of their resources for attacking the defender or protecting itself. They adopt a strategy pair which is proven to be a subgame perfect Nash equilibrium for a two-stage game and illustrates how strategy depends on changes in the enemy’s resources (Hausken and Zhuang, 2011).

Game theory related techniques are popular in the existing literature related to resource dependent attacker-defender game. However, these applications mainly focus on single period games or repeated games where the significant information from previous periods are ignored. Furthermore, heavy assumptions on previous approach, i.e., single resource type, sequential move, perfect information on enemy also make the previous research inapplicable to real-world military mission, where the demands and environment are stochastic and unpredictable (Shinkman, 2014; Lynch, 2015; Xu et al., 2016). Furthermore, the recent study also proves that the performance of modular fleet is heavily influenced by the optimality of operation decisions (Li and Epureanu, 2018, 2017a), which make the analytical solution intractable. A new method is required to investigate the tactical advantages of fleet modularity.

In this study, we go beyond existing literature and formulate an attacker-defender game and formulate an intelligent agent-based modeling techniques (Adhitya et al., 2007; Yu et al., 2009; Onggo and Karatas, 2016; Li and Epureanu, 2017a) to imitate the human-like decision making process. We combine opti-

mization techniques and artificial intelligence to enable each player (fleet) to make real-timely decisions based on experience and optimize their decisions accordingly. By selecting one player as modular fleet and another as conventional fleet, we emphasize the benefits of fleet modularity, including adaptability and unpredictability, through competing in a generalized mission scenario.

2. Game Formation

During the armed conflicts with Iraq and Afghanistan, the U.S. faced supply shortages due to exogenous supply chain disruptions (Xu et al., 2016). We borrowed this example and modeled it as a competition between two military forces. In this game, we assume each force is composed of a fleet of vehicles, i.e., fleet red and fleet blue. The goal of fleet operation is to satisfy the supply demands which randomly appear at battle fields. Each demand contains the due time, required materials, personnel, and target fleet to accomplish the supply. In order to satisfy the demand, we need to ensure the dispatched convoy, which are vehicles selected from fleet, owns enough capacities and firepower to guarantee the safety of transportation. For convenience, all demands are automatically converted into the attribute requirements for the convoy, i.e., fire power, water capacity, etc. We denote the demands received at time t as $\mathbf{r}(t)$. According to the due time of demand, attributes required to be satisfied at a future time $t + \tau$ can be obtained as $\mathbf{d}^x(t + \tau) = [d_1^x(t + \tau), d_2^x(t + \tau), \dots, d_{N_a}^x(t + \tau)]^T$, where $d_e^x(t + \tau)$ represents the attributes of type e to be satisfied before time $t + \tau$. Matrix $\mathbf{D}^x(t)$ can also be created and updated to record the demands to be satisfied in the planning horizon T_p , i.e., $\mathbf{D}^x(t) = [\mathbf{d}^x(t + 1), \mathbf{d}^x(t + 2), \dots, \mathbf{d}^x(t + T_p)]$. Correspondingly, attributes carried by on-field convoy at time t are $\mathbf{v}^x(t) = [v_1^x(t), v_2^x(t), \dots, v_{N_a}^x(t)]^T$.

The fleet targeted by demand becomes defender, with the goal of delivering a qualified convoy to field on time. The other becomes attacker automatically, which aims to disrupt the delivery of defender by dispatching an attacker convoy. In other words, each demand initializes a supply task for one fleet and attack task for another. Based on the demand, the role of the player dynamically changes correspondingly. Common rules of the game are assumed to be known to the players: delivery of a qualified convoy to the assigned battlefield makes the defender win and attacker lose. Because the damage from the attacker can reduce the attributes of the convoy, defender may lose the game even if a qualified convoy is dispatched.

We denote the resulted conflict between convoys as an event and assume the dispatch decision made from attacker and defender are simultaneous, which specifies the game as simultaneous-move. To simplify the problem, we assume that both fleets can simultaneously sense the demands regardless of its target. Thus, given the same probability to be selected as target fleet, each fleet will play equivalent times of defender or attacker to guarantee the fairness of the game. Fig. 2 illustrates the convoy competition in a

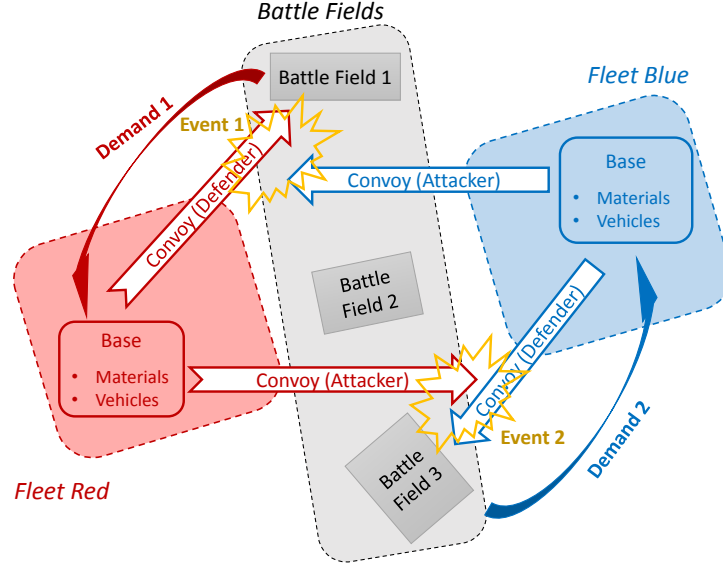


Figure 2: Layout of Battle Field

multiple battle fields scenario.

Following the assumptions from previous research (Azaiez and Bier, 2007; Wang and Bier, 2011), both attacker and defender are modeled as rational and strategic. Based on the simulation results, we summarized 10 types of dispatch strategies for attacker and defender respectively, as shown in Tab.1. For an attacker, firepower is the only attribute needed to win. Thus, we measure the strategy by k_a times of required firepower for enemy. For defender, they need to guarantee the delivered convoy to satisfy the demands with consideration of attribute losses from the enemy's attack. The strategy is a mixture of decisions in selecting of safety coefficients of fire power k_a and capacity k_c . We cluster all the dispatch orders that are less than the requirements in strategy 1, which means the defender gives up the game to save force once a strong enemy convoy is predicted.

In this study, the amount of damage is based on the comparison of the fire power carried by different convoys. The probability of damage of type i component for team red p_{di}^r and team blue p_{di}^b are,

$$p_{di}^r = \tanh(k_{di} \frac{v_f^b}{v_f^r}) \quad (1)$$

$$p_{di}^b = \tanh(k_{di} \frac{v_f^r}{v_f^b}), \quad (2)$$

where v_f^x is the amount of firepower carried by convoy x . k_{di} represents the damage factor for a component in type i . Each component in the convoy gets damage stochastically based on the calculated probability.

In order to create a fair game, we constrain the amount of supplies, and assume all damaged resources

Table 1: Dispatch strategies for attacker and defender

Attack Strategy	1	2	3	4	5
Range of k_a	$[0, 0.5)$	$[0.5, 1)$	$[1, 1.5)$	$[1.5, 2)$	$[2, 2.5)$
Attack Strategy	6	7	8	9	10
Range of k_a	$[2.5, 3)$	$[3, 3.5)$	$[3.5, 4)$	$[4, 4.5)$	$[4.5, \infty)$
Defense Strategy	1	2	3	4	5
Range of k_a	$[0, 1)$	$[1, 1.5)$	$[1, 1.5)$	$[1, 1.5)$	$[1.5, 2)$
Range of k_c	$[0, 1)$	$[1, 1.5)$	$[1.5, 2)$	$[2, \infty)$	$[1, 1.5)$
Defense Strategy	6	7	8	9	10
Range of k_a	$[1.5, 2)$	$[1.5, 2)$	$[2, \infty)$	$[2, \infty)$	$[2, \infty)$
Range of k_c	$[1.5, 2)$	$[2, \infty)$	$[1, 1.5)$	$[1.5, 2)$	$[2, \infty)$

are recoverable. Thus, the amount of resources for both fleets are constant, but the conditions of resources are dynamic. We penalize damage by a long time to recover. Recovery strategy for damaged vehicles is to replace all the damaged components by healthy ones. Once fleet modularity is considered, disassembly becomes another option in dealing with a badly damaged vehicle.

Several assumptions are also used to simplify this problem by reserving a reasonable level of fidelity:

1. Each fleet can accurately observe and record the damage occurred in its convoy.
2. Each fleet can accurately observe the composition of enemy's convoy in every event, i.e., number and types of vehicles.
3. Convoy will return to base immediately after task.
4. Mission success will be reported to both fleets once it is completed.
5. No other type of vehicle damage is considered besides attack.
6. All components are recoverable.
7. Damage independently occurs based on the probability of damage.
8. Inventory status is updated every hour and accessible to all the agents.
9. The sequence of vehicle assembly, disassembly, and recovery is ignored.

3. Problem Statement

In order to explore the advanced tactics from fleet modularity, we need to first consider the way a human commander performs yield operation decisions. With simplification, the main procedure involved in human-like decision making (Chen and Barnes, 2013) are 1) perceive field information, 2) analyze enemy's behavior based on the received information, 3) optimally schedule the on-field actions to beat enemy. Through these procedures, decision makers could adaptively adjust their dispatch strategy and operation plan based on the learned enemy's behaviors, i.e., what kind of strategy they might adopt in a specific situation. Combining with fleet modularity, how to imitate the human-like decision making process with guaranteed decision optimality and efficiency remain to be solved.

Another existing challenge is the management of inventory with high diversity. We denote a vehicle with one or more than one damaged components as damaged vehicles, which generates numerous types of vehicle damage. For example, a single vehicle with 5 different components has 2^5 types of vehicle damage, which leads to distinct recovery strategy. Furthermore, given the limited working capacity and inventory status, it is still challenging to effectively and efficiently schedule the operations and select the recovery strategy in reaction to the stochastically arrived demands while maintaining reasonably healthy stock levels.

4. Agent-Based Model

With consideration of computational load and battlefield decision making process, we present an agent-based model to automatically yield adaptive tactics and real-timely plan for operational actions accordingly. To simplify the notation, we describe the approach from the standpoint of fleet blue, and proposed an approach to beat the enemy, i.e., fleet red. Three types of agents models are created to perform different functionalities. The decisions making process is then achieved by the cooperation of three types of agents. The interconnections are shown in Fig. 3.

1. **Inference Agent:** analyze enemy's historical behaviors, forecast enemy's future actions.
2. **Dispatch Agent:** optimize dispatch order based on inference.
3. **Base Agent:** optimally plan for the operation actions to satisfy the dispatch order.

4.1. Inference Agent

As simultaneous-move is considered, it is critical to forecast the enemy's actions to counter. As combat resources and workshop capacity are limited, it is possible to get cues from the enemy's historical

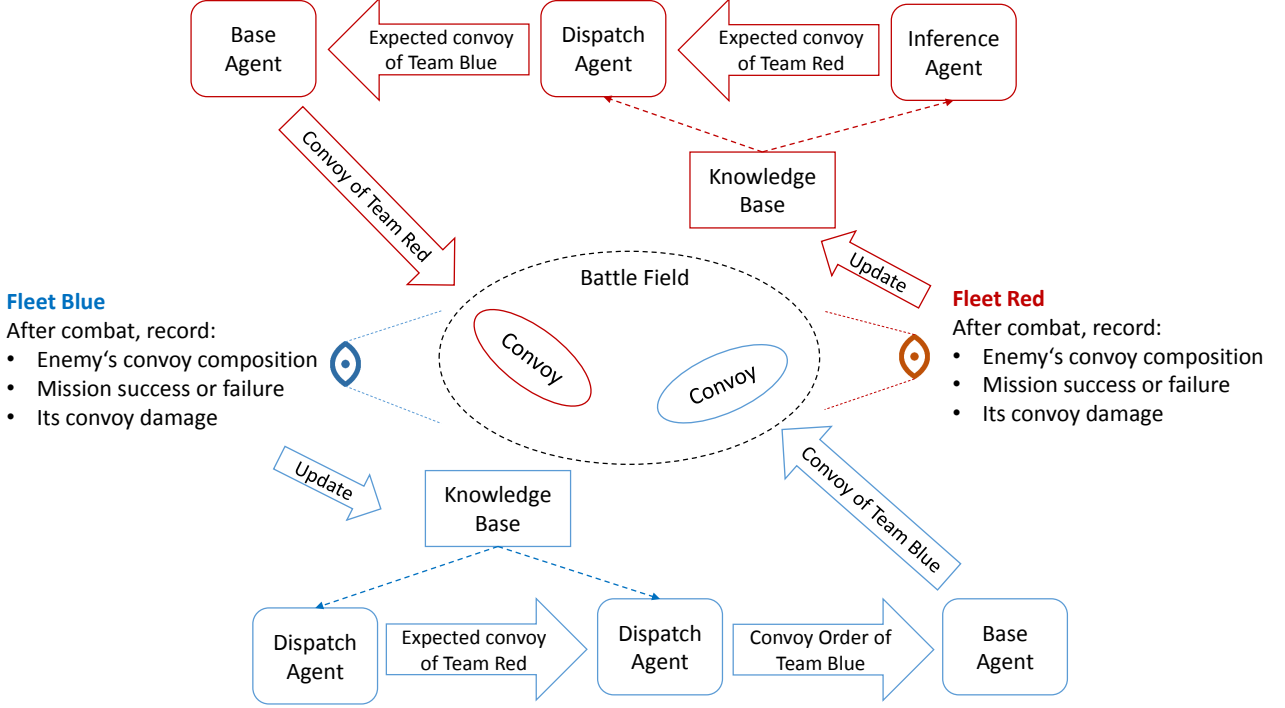


Figure 3: Agent-based model in attacker-defender game

dispatch actions in inference. For example, if enemy dispatched a significant size of convoy in the short past, it is possible to conclude that the enemy is not capable of grouping up a strong force in the short future. Meanwhile, existing damage in enemy's resources can also be analyzed by comparing the fire power of dispatched convoys in historical events. The amount of damage is also useful for decision maker to infer enemy's available forces.

The information that can be used for inference is very limited, including demand records $D(t)$, our previous dispatch, $v^b(\tau)$, and enemy's previous dispatches $v^r(\tau)$. Dispatch decisions depend on optimization algorithm, inference and personality of commander, which leads to a remarkable nonlinearity in modeling the decision making process. Meanwhile, as commander needs to adjust its strategy after learning from the enemy. It requires the prediction model capable of exchanging information from outsides, and make corrections once needed. We adopt techniques from artificial intelligence to solve this problem.

Recurrent neural network is a well known machine learning model that could capture the dependencies of time-series data (Mikolov et al., 2010). Compared to the neural network, it can memorize a certain period of historical data and analyze its influence on the future. Long short-term memory model networks is one of the popular RNN, which is capable of learning long-term dependencies which is widely implemented in the neural language without gradient vanishing problem in RNN. In this study, we implement a variant

model of RNN, long short-term memory (LSTM) as our predictive model to capture the correlations in enemy's sequential decisions. The model is widely used in forecasting based on sequential data, including, stock market price (Chen et al., 2015; Di Persio and Honchar, 2016; Fischer and Krauss, 2017), traffic (Ma et al., 2015; Zhao et al., 2017), etc. In this study, we model the inference of the enemy's strategy as a classification problem, where each class corresponds to a strategy. The inputs $\mathbf{x}^{(m)}$ of training data is the record for each event, including enemy's dispatched convoy, our dispatched convoy and received demand, which is a time-series data recording all the information during review horizon T_b . The outputs $y^{(m)}$ of the training example is the actual dispatch strategy adopted by the enemy. The architecture of LSTM used in this study is shown in Fig. 4.

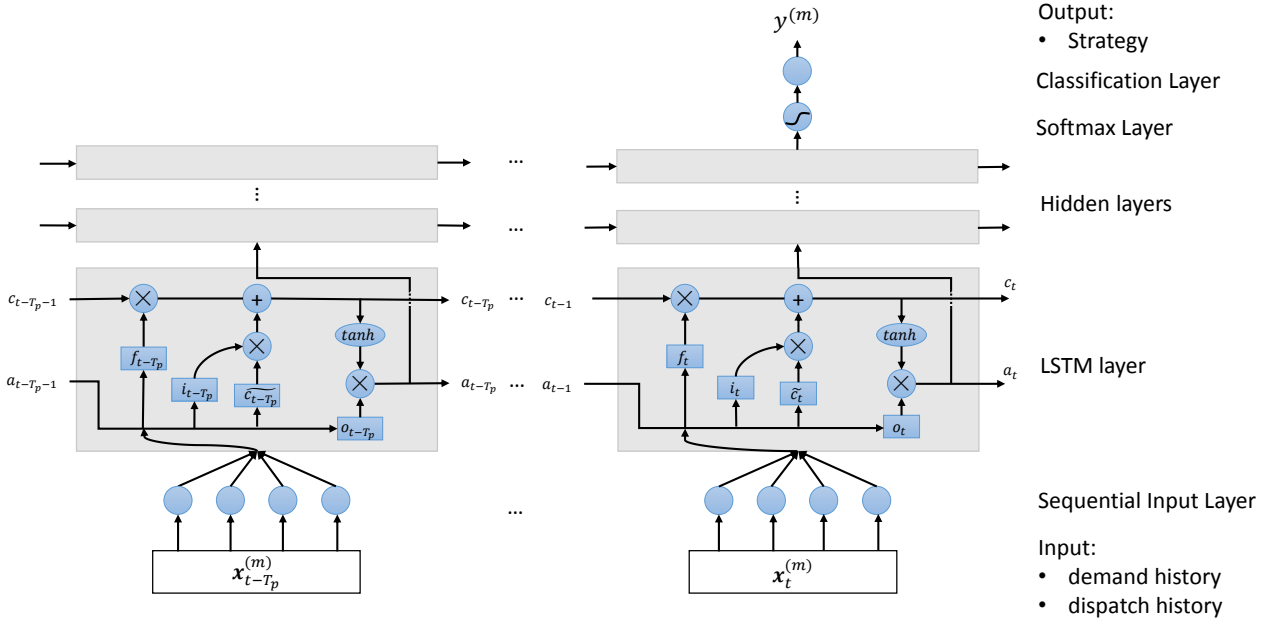


Figure 4: LSTM architecture

The status of LSTM at time t are described by input gate \mathbf{i}_t , forget gate \mathbf{f}_t , output gate \mathbf{o}_t , cell state \mathbf{c}_t and activation at i^{th} hidden layer \mathbf{a}_t^i . The forward propagation can be described by

$$\mathbf{i}_t = \sigma(\mathbf{W}_i[\mathbf{a}_{t-1}, \mathbf{x}_t^{(m)}] + \mathbf{b}_i), \quad (3)$$

$$\mathbf{f}_t = \sigma(\mathbf{W}_f[\mathbf{a}_{t-1}, \mathbf{x}_t^{(m)}] + \mathbf{b}_f), \quad (4)$$

$$\mathbf{o}_t = \sigma(\mathbf{W}_o[\mathbf{a}_{t-1}, \mathbf{x}_t^{(m)}] + \mathbf{b}_o), \quad (5)$$

$$\mathbf{c}_t = \mathbf{i}_t * \tanh(\mathbf{W}_c[\mathbf{a}_{t-1}, \mathbf{x}_t^{(m)}] + \mathbf{b}_c) + \mathbf{f}_t * \tanh(\mathbf{W}_c[\mathbf{a}_{t-2}, \mathbf{x}_{t-1}^{(m)}] + \mathbf{b}_c), \quad (6)$$

$$\mathbf{a}_t^0 = \mathbf{o}_t * \tanh(\mathbf{c}_t), \quad (7)$$

$$\mathbf{a}_t^1 = \mathbf{W}_1 \mathbf{a}_t^0 + \mathbf{b}_1, \quad (8)$$

$$\mathbf{a}_t^2 = \mathbf{W}_2 \mathbf{a}_t^1 + \mathbf{b}_2, \quad (9)$$

$$\vdots$$

$$\mathbf{a}_t^{N_h} = \mathbf{W}_{N_h} \mathbf{a}_t^{N_h-1} + \mathbf{b}_{N_h}, \quad (10)$$

$$\mathbf{p}^{(m)} = \sigma_m(\mathbf{a}_t^{N_h}), \quad (11)$$

where, \mathbf{W}, \mathbf{b} are weights and biases to be obtained through training. σ is the sigmoid function. N_h is the number of hidden layer. σ_m is the softmax function. $\mathbf{p}^{(m)}$ records the estimated probability of each class based on inputs from training sample m and weights of the model, i.e., $\mathbf{p}^{(m)} = [p_1^{(m)}, p_2^{(m)}, \dots, p_{N_c}^{(m)}]^T$. The loss function is represented by a cross entropy equation:

$$E^m = - \sum_{i=1}^{N_c} \gamma_i \log(p_i^{(m)}), \quad (12)$$

where, γ_i is a binary indicator (0 or 1), with value 1 if class label i is the correct classification and 0 otherwise. Thus, the training of the model is to minimize the sum of entropy of the training set to find the best model parameters through backward propagation (Hecht-Nielsen, 1992), i.e.,

$$\min_{\mathbf{W}, \mathbf{b}} \sum_m E^m, \quad (13)$$

thus, enemy's behavior can be forecasted through Eq.14.

$$\bar{y}^r = \{i \mid 1 \leq i \leq N_c, \forall 1 \leq j \leq N_c : p_i^{(m)} \geq p_j^{(m)}\}. \quad (14)$$

Based on the predicted enemy's strategy \bar{y}^r , we can calculate the possible enemy's dispatch order $\bar{\mathbf{v}}^r$ by using the upper bound of strategy of Tab. 1.

4.2. Dispatch Agent

The goal of convoy dispatch is to determine the desired attributes that need to be carried by our convoy to maximize the win rate. Based on game formation, a convoy with higher attributes, especially in fire power, indicates a higher chance to win. However, as resources are limited, the less vehicles are ordered, the higher chance that the order can be achieved by the inventory planner. The convoy dispatch should be carefully planned to guarantee the win rate of current mission without overdrawing resources.

To avoid the overuse or underutilization of available attributes, it is important to have an accurate evaluation of our current forces before making dispatch decisions. Available attributes in the future depend on inventory status and on-base action scheduling. As time delays exist in operations and future demands

are only partially observable. It is hard to infer the actual planning that will be made in the short future, which may change the available attributes at dispatching time totally. This difficulty becomes one of the challenges in placing the dispatch order. Even we ascertain that our convoy order can be achieved by base, it is still uncertain for them to determine the probability that this convoy can win. Because all the players do not know how the vehicle gets damaged, it requires players to speculate the damage mechanism based on the experience.

To resolve this problem, we decoupled the estimation of event success in two parts, which are feasibility of order p_f^b and conditional success rate if order is feasible p_s^b . The probability of winning an event for convoy blue p_w^b can be calculated.

$$p(\text{win}) = p(\text{win}|\text{order is feasible})p(\text{order is feasible})$$

$$p_w = p_s^b p_f^b. \quad (15)$$

4.2.1. Feasibility

Based on the historical records, the feasibility of order can be flagged by 0(infeasible) and 1 feasible based on the comparison between dispatch order $\mathbf{do}^b(t)$ and actually dispatched convoy $\mathbf{o}_v^b(t)$. We denote the order is feasible if $\mathbf{do}^b(t) \leq \mathbf{M}_{va}\mathbf{o}_v^b(t)$, is infeasible otherwise. As optimizations are implemented in operation planning, the relationship between factors and feasibility is complex and nonlinear. We implement a neural network model (Hagan et al., 1996; Atsalakis et al., 2018; Rezaee et al., 2018) to capture these nonlinear inter-connections, as shown in Fig. 5. The outputs of training set is the feasibility of the order (1 for feasible, 0 for infeasible).

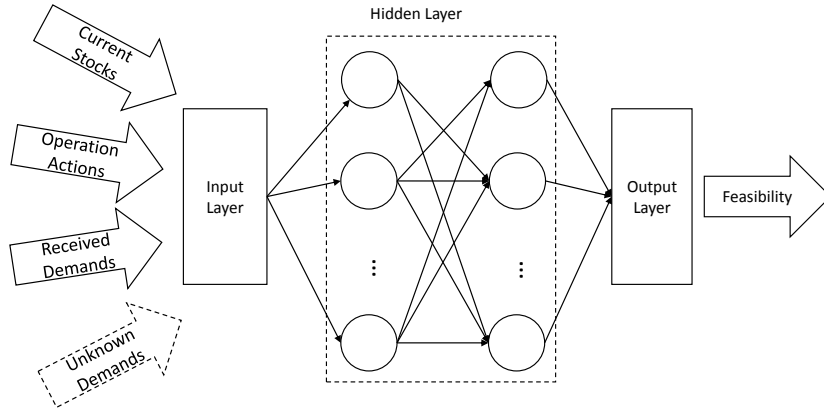


Figure 5: Neural network model for feasibility training

With enough training, the model is capable of evaluating the feasibility of dispatch order across diverse

operation situations. To capture the changing of inventory operation strategy, we periodically retrain the model based on the latest operation information. The relationship between model inputs and feasibility rate can be described by Eq. 16,

$$p_f = f_f(\mathbf{do}^b(t), \mathbf{s}^b(t), \delta \mathbf{s}^b(t+1), \dots, \delta \mathbf{s}^b(t+T_p), \mathbf{D}^b(t)), \quad (16)$$

where, $\mathbf{s}^b(t)$ records the number of vehicles and component stocks on base at time t . $\delta \mathbf{s}(t+\tau)$ is the changes in inventory stocks at time $t+\tau$ from unfinished actions.

4.2.2. Conditional Success Rate

Vehicle damage plays an important role in determining the success of a mission. However, it is driven by a stochastic process and will vary according to the changes in terrain, operational preparations and soldiers' reactions. It requires the model to be able to capture the complexity in the damage mechanism. Combining with the nonlinearity in predicating the success, we adopted another neural network model for success rate forecasting, as shown in Fig. 6.

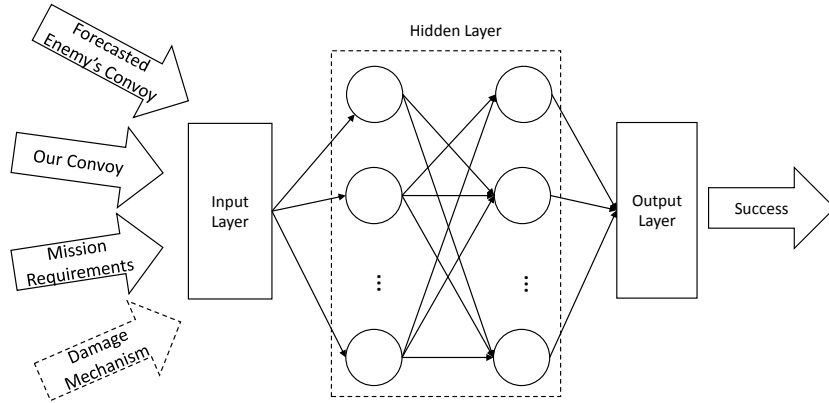


Figure 6: Neural network model for success rate training

Similarly, the outputs of training set are success reports of previous events (1 for success, 0 for fail). Given forecasted enemy's convoy attributes $\bar{\mathbf{v}}^r(t)$, the trained model will yield conditional win rate for a certain dispatch order and mission requirements. The model is capable of capturing the changes in damage mechanism by continuously feeding in the latest combat information and results. By denoting the trained neural network model for success as f_s^b , the probability of success can be also calculated.

$$p_s^b = f_s^b(\mathbf{do}^b(t), \bar{\mathbf{v}}^r(t), \mathbf{D}(t)). \quad (17)$$

4.2.3. Optimization

For each dispatch order $\mathbf{do}^b(t)$, the above approach provides the way to estimate the probability of success and feasibility based on predicted enemy's behavior, demand information and inventory status. An optimization model can be used to seek the optimal dispatch order to maximize win rate or minimize failure rate, i.e., $J = 1 - p_b^w$. Combining with Eqn. 16,17, a nonlinear programming model can be formulated to seek the optimal dispatch order.

$$\begin{aligned} \min_{\mathbf{do}^b(t)} \quad & 1 - [f_f(\mathbf{do}^b(t), \mathbf{s}^b(t), \delta \mathbf{s}^b(t+1), \dots, \delta \mathbf{s}^b(t+T_p), \mathbf{D}(t))][f_s^b(\mathbf{do}^b(t), \bar{\mathbf{v}}^r(t), \mathbf{D}(t))] \\ \text{s.t.} \quad & (a) \mathbf{do}^b(t) \geq 0, \end{aligned} \tag{18}$$

where, $\mathbf{do}^b(t)$ is the decision variable that specifies the desired attributes to be carried by the convoy. Thus, the number of decision variables is the number of attribute types (3 ~ 10). However, because of the non-convexity in objective function, it is intractable to get global optimum by the gradient-based approach. In this study, we implement a pattern search technique to yield optimal dispatch decisions.

As the minimized failure rate can be any value in the range of $[0, 1]$, dispatch agent should be capable of giving up the mission once a very high failure rate is calculated. There is also a stream of literature studying risk preferences in repeated and evolutionary games (Roos and Nau, 2010; Lam and Leung, 2006; Zhang et al., 2018). We define the ϵ_f as a customizable parameter to represent the minimal failure rate that can be tolerated, which filters the dispatch order $\mathbf{do}^b(t)$ as

$$\mathbf{do}^b(t) = \begin{cases} = \mathbf{0}, & \text{for } (1 - p_w) > \epsilon_f, \\ = \mathbf{do}^b(t) & \text{for } (1 - p_w) \leq \epsilon_f. \end{cases} \tag{19}$$

Thus, a convoy can be dispatched only when commander is confident enough. In this study, risk aversion behavior is purely related to ϵ which is constant during operation. As a future work, it is also interesting to vary ϵ to seek an advanced fleet operation strategy, i.e., combination of risk-prone and risk-averse (Roos and Nau, 2010).

4.3. Base Agent

Based on the behavior analysis from inference agent and dispatch order suggestion from dispatch agent, base agent is the one to plan operational actions to accomplish the orders. Li et al. proposed a model predictive control based approach to real-timely schedule the operation actions in reacting to the received demands. However, they did not consider the possible damage that occurs during the fleet operation (Li

and Epureanu, 2018). In this section, we further their research by considering the possible damage during fleet operation and manage the inventory based on the resulting diverse conditions. For convenience, we simplify the notation of operation actions for fleet x from \mathbf{o}^x to \mathbf{o} in this section as no enemy is considered.

As resources for each player are limited and repairable, it is important to schedule the operation actions properly to recover damaged resources and increase utility rate. It is also essential to allocate the capacity properly to balance between order satisfaction and damage recovery. In this section, we first model military fleet operation as a time-varying dynamical system. Then, a model predictive control is proposed to manage system dynamics thus achieving operation management.

4.3.1. Dynamical System

The dynamics in fleet operation is mainly located at the changes of inventory stocks and remaining demands, in terms of

1. Vehicle stocks, $\mathbf{I}_v = [I_{v_1}, I_{v_2}, \dots, I_{N_v}]$
2. Module/component stocks, $\mathbf{I}_m = [I_{m_1}, I_{m_2}, \dots, I_{N_m}]$
3. Damaged vehicles, $\mathbf{I}_{dv} = [I_{dv_1}, I_{dv_2}, \dots, I_{N_{dv}}]$
4. Damaged components, $\mathbf{I}_{dc} = [I_{dc_1}, I_{dc_2}, \dots, I_{N_c}]$
5. Unsatisfied demands, $\mathbf{I}_a = [I_{a_1}, I_{a_2}, \dots, I_{N_a}]$.

Although healthy vehicles and damaged vehicles are recorded in the similar structures, their meanings are totally different. For healthy stocks and damaged components, the subscript of variable is the type of vehicle/component; the value of variable indicates the number. For damaged stocks, the subscript is the index of damaged vehicles, which is created based on the vehicle receive date. Binary values are used to represent the status of the vehicle, where, 1 represents that the damaged stocks remain to be repaired; 0 indicates the damaged stock is recovered or not received yet. Vehicle type $\mathbf{v}_{lt}(t)$, number of damaged components $\mathbf{v}_{ldc}(t)$ and healthy components $\mathbf{v}_{lc}(t)$ in the damaged vehicle of type l are also recorded as a reference for the repair. These data are time variant as the number and type of damaged vehicles keep changing with newly occurred damage and vehicle recovery. These data structure can bypass the numerous states incurred from diverse vehicle damage patterns, as we distinguish the vehicles with different damage as different damaged vehicles. We create a state for each newly arrived damaged vehicle and remove the corresponding state once the damaged vehicle is recovered, i.e., state value changes from 1 to 0.

Vehicle conditions are reported to base agent as one of the inputs, we summarize all inputs to the system as

1. Returning healthy vehicles, $\mathbf{a}_r = [a_{r_1}, a_{r_2}, \dots, a_{r_{N_v}}]$,

2. Returning damaged vehicles, $\mathbf{a}_d = [a_{d_1}, a_{d_2}, \dots, a_{d_{N_{dv}}}]$,
3. Dispatch order from dispatch agent, $\mathbf{do} = [do_1, do_2, \dots, do_{N_a}]$.

Based on the characteristics of fleet operation, the operational actions to be determined are also distinct. For conventional fleet, the operation actions include

1. Convoy dispatch, $\mathbf{o}_v = [o_{v_1}, o_{v_2}, \dots, o_{v_{N_v}}]$
2. Recovery of damaged vehicle, $\mathbf{o}_{dr} = [o_{dr_1}, o_{dr_2}, \dots, o_{dr_{N_{dv}}}]$
3. Recovery of damaged component, $\mathbf{o}_c = [o_{c_1}, o_{c_2}, \dots, o_{N_c}]$.

The dynamics of vehicle stocks of type k , component stocks of type i , damaged vehicle of index l , damaged components of type i , and remaining attributes of type h are shown by Eq. 20, 21, 22, 23, 24 respectively

$$I_{v_k}(t+1) = I_{v_k}(t) - o_{v_k}(t) + a_{r_k}(t) + \sum_l v_{lkt}(t) o_{dr_l}(t - \tau_{v_l}), \quad (20)$$

$$I_{c_i}(t+1) = I_{c_i}(t) + o_{c_i}(t - \tau_{c_i}) - \sum_l v_{lc_i}(t) o_{dr_l}(t), \quad (21)$$

$$I_{dv_l}(t+1) = I_{dv_l}(t) - o_{dr_l}(t) + a_{d_l}(t), \quad (22)$$

$$I_{dc_i}(t+1) = I_{dc_i}(t) - o_{c_i}(t) + \sum_l v_{ldc_i}(t) o_{dr_l}(t - \tau_{dr_l}), \quad (23)$$

$$I_{a_h}(t) = do_h(t) - \sum_k M_{v_k a_h} o_{v_k}(t). \quad (24)$$

By introducing fleet modularity, several additional operation actions are available, in terms of

1. Vehicle assembly, $\mathbf{o}_a = [o_{a_1}, o_{a_2}, \dots, o_{a_{N_v}}]$
2. Vehicle disassembly, $\mathbf{o}_d = [o_{d_1}, o_{d_2}, \dots, o_{d_{N_v}}]$
3. Vehicle reconfiguration, $\mathbf{o}_r = [o_{12}, o_{13}, \dots, o_{N_v N_v - 1}]$
4. Damaged vehicle disassembly, $\mathbf{o}_{dd} = [o_{dd_1}, \dots, o_{dd_{N_{dv}}}]$.

With consideration of these actions, dynamic equations become

$$\begin{aligned} I_{v_k}(t+1) = & I_{v_k}(t) - o_{v_k}(t) + a_{r_k}(t) + \sum_{k' \neq k} o_{k'k}(t - \tau_{k'k}) - \sum_{k' \neq k} o_{kk'}(t) \\ & + \sum_l v_{lkt}(t) o_{dr_l}(t - \tau_{dr_l}) + o_{a_k}(t - \tau_{a_k}) - o_{d_k}(t), \end{aligned} \quad (25)$$

$$\begin{aligned} I_{c_i}(t+1) = & I_{c_i}(t) + o_{c_i}(t - \tau_{c_i}) - \sum_k M_{v_k c_i} [o_{a_k}(t) - o_{d_k}(t - \tau_{d_k}) \\ & + \sum_{k' \neq k} o_{k'k}(t) - \sum_{k' \neq k} o_{kk'}(t - \tau_{kk'})] + \sum_l v_{lc_i}(t) o_{dd_l}(t - \tau_{dd_l}), \end{aligned} \quad (26)$$

$$I_{dv_l}(t+1) = I_{dv_l}(t) - o_{dr_l}(t) - o_{dd_l}(t) + a_{d_l}(t), \quad (27)$$

$$I_{dc_i}(t+1) = I_{dc_i}(t) - o_{c_i}(t) + \sum_l v_{ldc_i}(t) o_{dr_l}(t - \tau_{dr_l}) + \sum_l v_{ldc_i}(t) o_{dd_l}(t - \tau_{dd_l}) \quad (28)$$

$$I_{a_h}(t) = do_h(t) - \sum_k M_{v_k a_h} o_{v_k}(t). \quad (29)$$

Because of the delays in operation actions, current inventory stocks might be influenced by previously-determined actions. In other words, the current actions may impact the stock level in the future. Thus, we define the state of the system by all inventory statuses that might be influenced by current actions, $\mathbf{s}(t)$, i.e.,

$$\mathbf{s}(t) = [\mathbf{I}_t(t), \mathbf{I}_{t+1}(t), \mathbf{I}_{t+2}(t), \mathbf{I}_{t+3}(t) \dots \mathbf{I}_{t+\tau_{max}}(t)]^T. \quad (30)$$

We use input matrices \mathbf{B}_τ to connect the current actions at time t to inventory level at a later time $t + \tau$. Furthermore, damage on stocks keep changing along with time. The matrices that connect to previous states \mathbf{A} , actions \mathbf{B}_τ and inputs \mathbf{C} are also time-varying matrices. Thus, the system dynamics for both fleets can be written as

$$\mathbf{I}(t+1) = \mathbf{I}(t) + \sum_{\tau} [\mathbf{B}_{\tau}^c(t) \mathbf{o}^c(t-\tau)] + \mathbf{C}^c(t) \mathbf{a}(t-\tau), \quad (31)$$

$$\mathbf{I}(t+1) = \mathbf{I}(t) + \sum_{\tau} [\mathbf{B}_{\tau}^m(t) \mathbf{o}^m(t-\tau)] + \mathbf{C}^m(t) \mathbf{a}(t-\tau), \quad (32)$$

where

$$\mathbf{o}^c(t) = [\mathbf{o}_v(t), \mathbf{o}_{dr}(t), \mathbf{o}_c(t)], \quad (33)$$

$$\mathbf{o}^m(t) = [\mathbf{o}_v(t), \mathbf{o}_{dr}(t), \mathbf{o}_c(t), \mathbf{o}_a(t), \mathbf{o}_d(t), \mathbf{o}_r(t), \mathbf{o}_{dd}(t)]. \quad (34)$$

Thus, a state space model can be created to record the influence from the actions at a single time point to the states in the short future, as shown in Eq. 35

$$s(t+1) = A(t)s(t) + B(t)u(t) + C(t)a(t), \quad (35)$$

$$\mathbf{A}(t) = \begin{bmatrix} \mathbf{0}_{n_s(t) \times n_s(t)} & \mathbf{I}_{n_s(t) \times n_s(t)} & \mathbf{0}_{n_s(t) \times n_s(t)} & \cdots & \mathbf{0}_{n_s(t) \times n_s(t)} \\ \mathbf{0}_{n_s(t) \times n_s(t)} & \mathbf{0}_{n_s(t) \times n_s(t)} & \mathbf{I}_{n_s(t) \times n_s(t)} & \cdots & \mathbf{0}_{n_s(t) \times n_s(t)} \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ \mathbf{0}_{n_s(t) \times n_s(t)} & \mathbf{0}_{n_s(t) \times n_s(t)} & \mathbf{0}_{n_s(t) \times n_s(t)} & \cdots & \mathbf{I}_{n_s(t) \times n_s(t)} \\ \mathbf{0}_{n_s(t) \times n_s(t)} & \mathbf{0}_{n_s(t) \times n_s(t)} & \mathbf{0}_{n_s(t) \times n_s(t)} & \cdots & \mathbf{I}_{n_s(t) \times n_s(t)} \end{bmatrix}, \quad (36)$$

$$B(t) = [B_0(t), B_1(t), B_2(t), \dots, B_{\tau_{max}}(t)]^T, \quad (37)$$

$$\mathbf{C}(t) = [\mathbf{C}_0(t), \mathbf{C}_1(t), \mathbf{C}_2(t), \dots, \mathbf{C}_{\tau_{max}}(t)]. \quad (38)$$

4.3.2. System Control

The goal of system control is to meet the received dispatch orders on time. In the decision making process, the predictions of future system are always involved. For example, given several dispatch orders, one may want to know what are the influences from satisfying one order on others. Compared to the classical control methodologies, e.g., PID control, MPC makes better use of future information and adapts to the system changes (Li and Epureanu, 2017b). We separate this section into two parts, future state prediction and optimization of operation decisions.

Future State Prediction

Because of time delays in the operation actions, the operation decisions made at the current time have to guarantee the match between the attributes of the dispatched convoy and ordered attributes. Given

1. current system states $\mathbf{s}(t) = [\mathbf{s}(t+1), \mathbf{s}(t+2), \mathbf{s}(t+3), \dots, \mathbf{s}(t+t_p)]^T$
2. operation actions in the future $\underline{\mathbf{q}}_t = [\mathbf{o}(t), \mathbf{o}(t+1), \mathbf{o}(t+2), \dots, \mathbf{o}(t+t_p-1)]^T$
3. system input, $\underline{\mathbf{a}}_t = [\mathbf{a}(t), \mathbf{a}(t+1), \mathbf{a}(t+2), \dots, \mathbf{a}(t+t_p-1)]^T$,

The future system states $\underline{\mathbf{s}}_{t+1}$ are predictable by iteratively substituting Eq. 35. Thus, we can express the $\underline{\mathbf{s}}_{t+1}$ as a function of $\underline{\mathbf{q}}_t$

$$\underline{\mathbf{s}}_{t+1} = \mathbf{P}(t)\mathbf{s}(t) + \mathbf{H}(t)\underline{\mathbf{q}}_t + \mathbf{G}(t)\underline{\mathbf{a}}_t \quad (39)$$

$$\mathbf{P}(t) = [\mathbf{A}(t), \mathbf{A}^2(t), \mathbf{A}^3(t), \dots, \mathbf{A}^n(t)], \quad (40)$$

$$\mathbf{H}(t) = \begin{bmatrix} \mathbf{B}(t) & \mathbf{0} & \dots & \mathbf{0} \\ \mathbf{A}(t)\mathbf{B}(t) & \mathbf{B}(t) & \dots & \mathbf{0} \\ \dots & \dots & \dots & \dots \\ \mathbf{A}(t)^{n-1}\mathbf{B} & \mathbf{A}^{n-2}(t)\mathbf{B}(t) & \dots & \mathbf{B}(t) \end{bmatrix}, \quad (41)$$

$$\mathbf{G}(t) = \begin{bmatrix} \mathbf{C}(t) & \mathbf{0} & \dots & \mathbf{0} \\ \mathbf{A}(t)\mathbf{C}(t) & \mathbf{C}(t) & \dots & \mathbf{0} \\ \dots & \dots & \dots & \dots \\ \mathbf{A}(t)^{n-1}\mathbf{B} & \mathbf{A}^{n-2}(t)\mathbf{C}(t) & \dots & \mathbf{C}(t) \end{bmatrix}, \quad (42)$$

with $\mathbf{P}(t)$ being the matrix that connects the future system outputs with current system states, and $\mathbf{H}(t), \mathbf{G}(t)$ being the matrix that connects the system outputs with the future operation actions and inputs respectively. Although the dynamical system is changing along with time, we assume it is constant at each decision making time. The system keeps updating to ensure we optimize the operation actions based on accurate system status.

Cost Function

The optimization of fleet operation originates from two facts: 1. convoy with insufficient attributes suffer a remarkable risk in losing the mission; 2. convoy with redundant attributes can also deteriorate the overall fleet performance from utility reduction. Furthermore, we also consider several operational costs that may be significant in the real-world fleet operation. As a summary, the costs of interest are,

1. Attribute redundancy cost, c_{oh}
2. Attribute insufficiency cost, c_{uh}
3. ADR action cost, $c_{ak}, c_{dk}, c_{kk'}$
4. Recovery cost, c_{ci}, c_{dri}, c_{ddi}
5. Inventory holding cost, $c_{hc_i}, c_{hv_k}, c_{hdc_i}, c_{hdv_l}$

Therefore, the cost function is shown as

$$\begin{aligned}
 J = & \sum_{\tau=t}^{\tau=t+T_p-1} \left[\sum_h c_{oh} I_{ah}^+(\tau) + \sum_h c_{uh} I_{ah}^-(\tau) + \sum_l (c_{dri} o_{dri}(\tau) + c_{ddi} o_{ddi}(\tau)) \right. \\
 & \left. + \sum_k [c_{ak} o_{ak}(\tau) + c_{dk} o_{dk}(\tau)] + \sum_{k' \neq k} c_{kk'} o_{kk'}(\tau) + \sum_i c_{ci} o_{ci}(\tau) \right] \\
 & + \sum_{\tau=t+1}^{\tau=t+T_p} \left[\sum_i c_{hc_i} I_{ci}(\tau) + \sum_k c_{hv_k} I_{vk}(\tau) + \sum_i c_{hdc_i} I_{dc_i}(\tau) + \sum_l c_{hdv_l} o_{dv_l}(\tau) \right],
 \end{aligned} \tag{43}$$

where $I_{ah}^+(\tau), I_{ah}^-(\tau)$ are non-negative variables created to remove the nonlinearity, which satisfies that

$$I_{ah}(\tau) = I_{ah}^+(\tau) - I_{ah}^-(\tau). \tag{44}$$

We record the insufficient and redundant attributes during the planning horizon as $\underline{I}_{\underline{a}}^+$ and $\underline{I}_{\underline{a}}^-$ respectively. The holding costs and actions related costs are also aggregated as $\underline{c}_{\underline{h}}$ and $\underline{c}_{\underline{q}}$. By substituting in Eq.39, we created a mixed-integer programming model to optimize operational decisions

$$\begin{aligned}
 \min_{\underline{I}_{\underline{a}}^+, \underline{I}_{\underline{a}}^-, \underline{q}_t} & \quad \underline{c}_o \underline{I}_{\underline{a}}^+ + \underline{c}_u \underline{I}_{\underline{a}}^- + \underline{c}_{\underline{h}} \underline{X}_s [\underline{P}(t) \underline{s}(t) + \underline{H}(t) \underline{q}_t] \\
 \text{s.t.} & \quad (a) \quad \underline{q}_t \geq 0 \text{ and integer} \\
 & \quad (b) \quad \underline{X}_s \underline{s}_{t+1} \geq 0 \\
 & \quad (c) \quad \sum \underline{o}(t) \leq \bar{P}, \quad \forall t \\
 & \quad (d) \quad \underline{I}_{\underline{a}}^+ - \underline{I}_{\underline{a}}^- = \underline{X}_a [\underline{P}(t) \underline{s}(t) + \underline{H}(t) \underline{q}_t + \underline{G}(t) \underline{a}_t] \\
 & \quad (e) \quad o_{ddi}(t) + o_{dri}(t) \leq 1, \quad \forall t,
 \end{aligned} \tag{45}$$

where, $\mathbf{X}_s, \mathbf{X}_a$ are index of inventory stocks and remaining dispatch orders in states respectively. Constraint (a) ensures that all operational decisions are non-negative and integer; (b) indicates that the amount of inventory stocks are non-negative; (c) ensures that the on-base ADR actions are always constrained by the maximum action capacity \bar{P} ; (d) preserves the balance between auxiliary variables and remaining orders to be satisfied; (e) specifies that each damaged vehicle can only be recovered by one recovery strategy. As cost function and constraints are linear and the number of decision variables is huge, we first implement a cutting-plan to reduce the decision space and then use the integer programming solver to get the solution. Time required for decision making of each time point is less than 1 second for operating 5 types of modular vehicles with planning horizon as 12 hours.

5. Numerical Illustrations

In this section, we provide numerical illustrations in a generalized mission scenario to study the different impacts of modularity on fleet performance. In general, it may be difficult to estimate the parameters accurately. However, we believe that it may be possible to get reasonable estimates for these parameters by using expert judgments and data from the existing literature. In this study, we assume the resources provided to fleet operation is constant and equal, which can be imagined as a competition of two fleets at an isolated island. One of them is conventional fleet; the other is modular fleet. Initially, ten of each type of vehicle and component are provided to both fleets. Demands randomly occur at battle field based on Poisson distribution with time interval as 10 hours. Demands include personnel capacity d_p , material capacity d_m and fire power d_f , which are generated based on Gaussian distribution as shown in Eq. 46, 47, 48. Because of the lack of diversity in the existing designs of modular vehicles, we borrowed five types of modular vehicles as well as six resulted modules from (Li and Epureanu, 2017b). The attributes carried by each vehicle are summarized in Tab. 2.

$$d_p \sim \mathcal{N}(40, 15) \quad (46)$$

$$d_m \sim \mathcal{N}(50, 20) \quad (47)$$

$$d_f \sim \mathcal{N}(30, 10) \quad (48)$$

The costs of insufficiency and redundancy are created based on the heuristic rules. For example, convoy usually suffers a high risk of failure once attributes of dispatched convoy are less than ordered. Thus, the cost for attribute insufficiency is assigned much higher than attribute redundancy, i.e., $c_{o_h} \gg c_{u_h}, \forall h$. The costs for operation actions are created based on the difficulties and time required.

Table 2: Mapping between vehicles and attributes

Vehicle Type	1	2	3	4	5
Fire power	1	3	8	0	6
Material capacity	2	6	2	2	8
Personnel capacity	4	1	0	10	5

We assigned the time required for module assembly and disassembly as constant vectors τ_{ma}, τ_{md} . Vehicle assembly/disassembly time τ_{va}, τ_{vd} is calculated by summing up all the time required for its components. Similarly, for repair and reconfiguration, we first sum all the actions required to process each individual components in the vehicle. We assume that the interface between components are well-designed to achieve quick vehicle reconfiguration, where assembly and disassembly time for all types of components are 1 hour and 0.5 hour respectively. We assume that on-base ADR actions are proceeded in a generalized work station, thus, the number of stations determines the amount of available capacity. In this study, the number of available work stations for both fleets is assigned as 10.

We propose a discrete event model to simulate fleet competition for three years. We separate the mission into two parts, which are stochastic stage (1st year) and learning stage (2nd and 3rd year). In stochastic stage, dispatch agent randomly picks up a dispatch strategy based on Tab.1 and passes this decisions to base agent. First-year operations generate time-series data, including combat history, feasibility records, etc., which are important inputs for the learning model. Training of learning model starts at the beginning of learning stage, where inference agent and dispatch agent make decisions based on the historical enemy’s behavior. Learning models are also updated monthly to ensure they reflect the enemy’s latest behavior.

6. Fleet comparison

In this section, we compare the fleet performance between modular fleet and conventional fleet in different stages. As one of the important metrics in measuring the fleet performance, we first compare the probability of win based on the results from multiple simulation, which is shown in Fig.7. According to the plot, conventional fleet outperforms modular fleet at the stochastic stage. However, once the intelligence of agent is introduced, i.e., both fleets entered learning stage, modular fleet gradually receives more wins of the game. A well separation can also be noticed along with the learning stage, which indicates a solid

leading position of modular fleet during the learning stage. To explain these results, we first compare the attributes carried by the actual dispatched convoy between fleets, and then compare the estimation accuracy in inference of enemy, order success and feasibility.

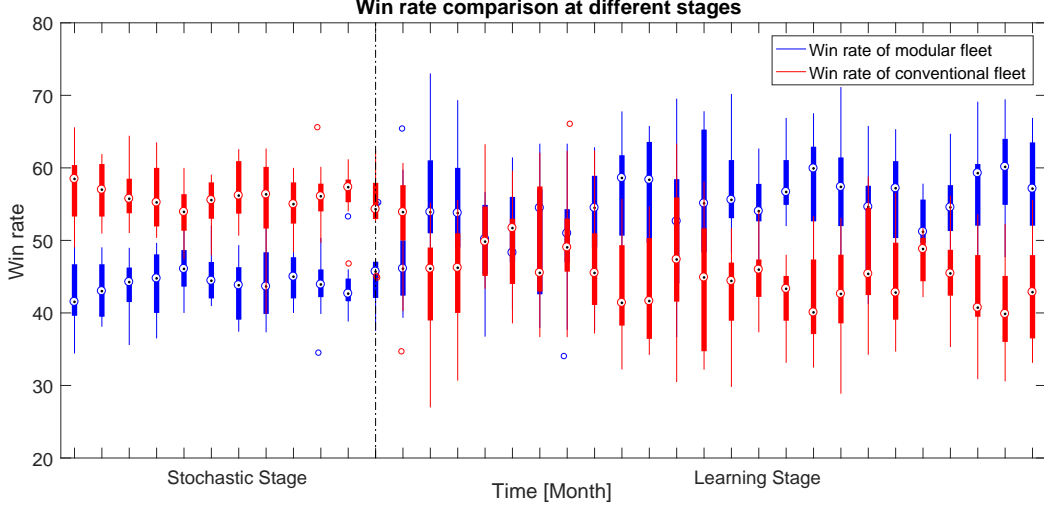


Figure 7: Comparison of win rate between modular fleet and conventional fleet

During the stochastic stage, dispatch agents from both teams place dispatch orders based on randomly selected strategy. The way of strategy selection and order achievement are also equivalent for both fleets. In order to explain the better performance of conventional fleet, it is necessary to explore the accuracy of both fleets in satisfying dispatch order. Mismatched attributes dramatically change the fleet performance: convoy with insufficient attributes may significantly raise the failure rate and damage; convoy with redundant attributes may increase the win rate slightly, but it can also contribute to the insufficiency in the short future because of the limitation of resources. Thus we calculate the amount of overused and insufficient convoy attributes during every month and compare the dispatch accuracy in Fig. 8.

Compared to the modular fleet, conventional fleet suffers remarkable redundancy. The higher redundancy comes from the rigidity of conventional fleet operation. As we know, modular fleet can real timely reconfigure itself to fit the dispatch order, however, conventional fleet can only wait for vehicles returned from field or recovery. This limitation is proven to hamper the ability of conventional fleet in satisfying the dispatch order. Once proper vehicles are scarce, conventional fleet has to use improper vehicles with little desired attributes to avoid the insufficiency. This rigidity in fleet operation is beneficial in improving the success rate during the stochastic stage, because its opponent cannot be aware of the unexpected additional attributes. However, once the opponent start to study the behavior, this advantage no longer exists. From

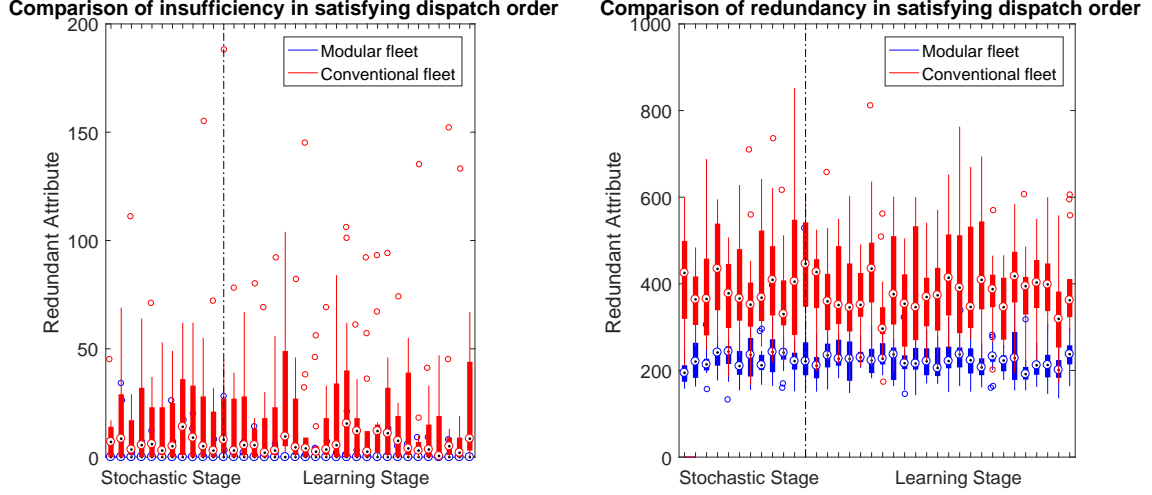


Figure 8: Comparison of mismatched convoy attributes

the failures at stochastic stage, modular fleet realize that conventional fleet intents to dispatch convoy with superfluous attributes. As a solution, they increase the attributes of dispatch order correspondingly. These redundant attributes are powerless in reacting to a well-armed enemy's convoy to make the conventional fleet stay the lead.

Besides the better understanding of enemy's behavior, intelligent agents also improve their understandings of the game along with time. To address the changes, we first denote the ability of convoy as maximum attributes can be achieved at each month, and compare the ability of both fleets in Fig.9. After entering the learning stage, both fleets raise the ability of convoy in all types of attributes, especially in firepower. Once modular fleet realize the importance of fire power to win the event by model f_s^b , the combat vehicles are rapidly formed from reconfiguration to boost the available fire power.

The swift reconfiguration of modular fleet lead to a dramatic increase of damage to the enemy in the first few months of learning stage, as shown in Fig.10. Although conventional fleet intend to increase the fire power to fight back, the limitation in vehicle structure results in a lower upper limit in the convoy ability. Thus, the difference in ability makes conventional fleet suffer higher damage from more dispatch, which forces the conventional fleet to operate in sub-healthy conditions for a long time.

The strategies used in learning stage are also distinct between two fleets. Fig.11 compares the proportion of strategies adopted by different fleets. After learning of game, both fleets prefer to select the defense strategy with large amount of firepower and fair amount of capacity, i.e., strategy 8,9. Because of the flexibility of fleet structure, modular fleet can be easily adapted to the vehicle damage and enemy's behavior, which leads to a better balance between different types of vehicles to perform a stronger strategy. The

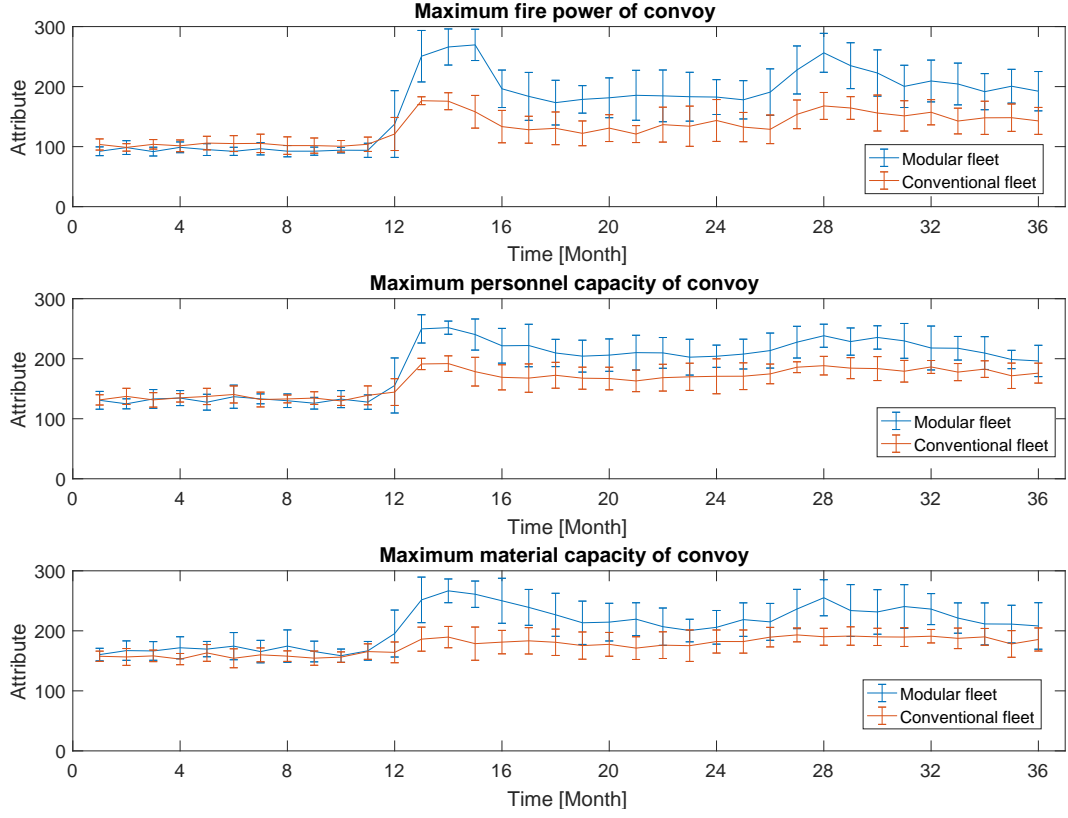


Figure 9: Maximum value of attribute carried by the convoy in different months

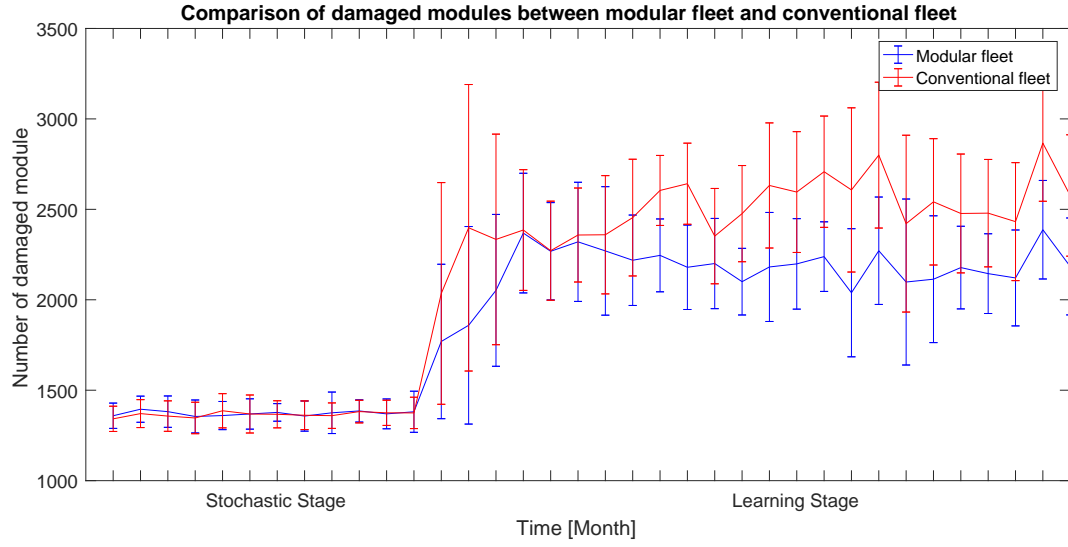


Figure 10: Modules damage occurs at each month

defense strategy selection also impacts the attack strategy. Compared to modular fleet, conventional fleet is much more likely to give up mission because of resources insufficiency. This weakness makes the modular fleet confident in dispatching little or even no combat vehicles to win the game. As an evidence, the proportion of strategy 1 used by modular fleet is much higher than that by conventional fleet. Meanwhile, modular fleet is more capable of performing aggressive strategies, i.e., strategy 8,9,10, more often than conventional fleet once a strong enemy is sensed.

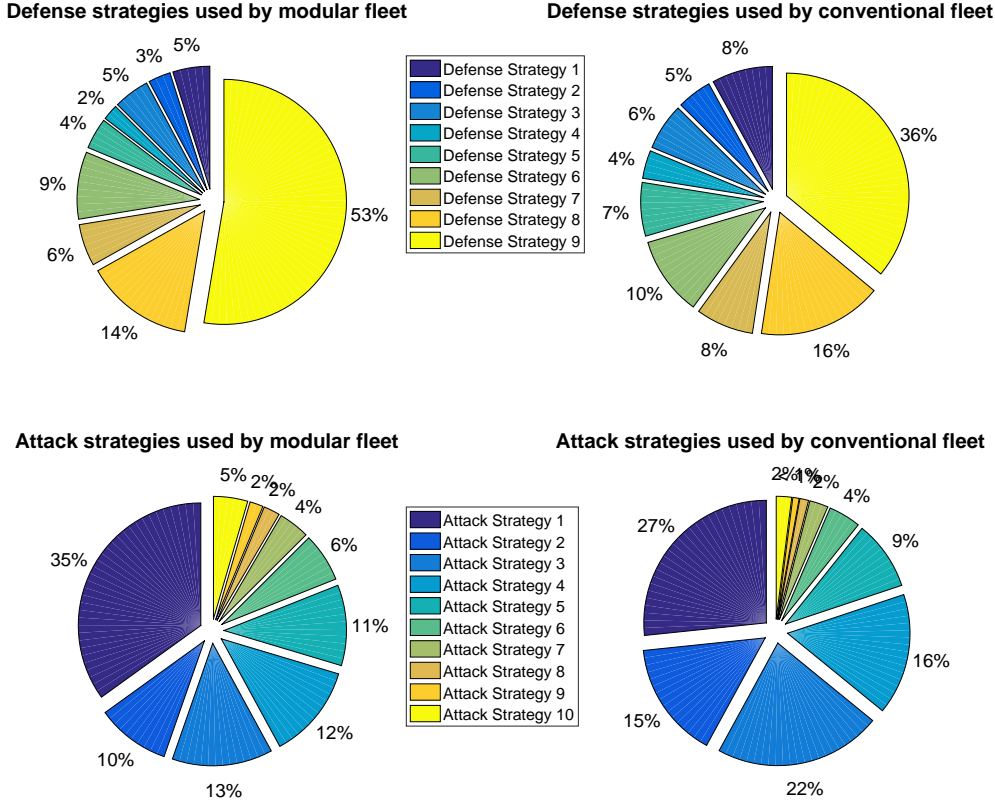


Figure 11: Proportion of defense strategy at different stages

To further investigate the improved performance of modularized fleet, we also compared the inference accuracy between different fleets. We denote the mean square error (MSE) between forecasted and actual convoy attributes as the metric to quantify the inference accuracy. As can be seen from the comparison in Fig.12. Inference errors are significantly high at the beginning of the learning stage, because agents are trained by the data from stochastic dispatch, which contributes little to forecasting the behavior of a rational competent. Along with the learning process, more combat and operation records, which are based on trained ABM, are generated, in which enemy's behaviors are more explainable. As a result, inference errors are significantly reduced in the following four months. However, the inference error keep fluctuating

during the rest of the learning stage, because both fleets keep checking and countering the other's behavior.

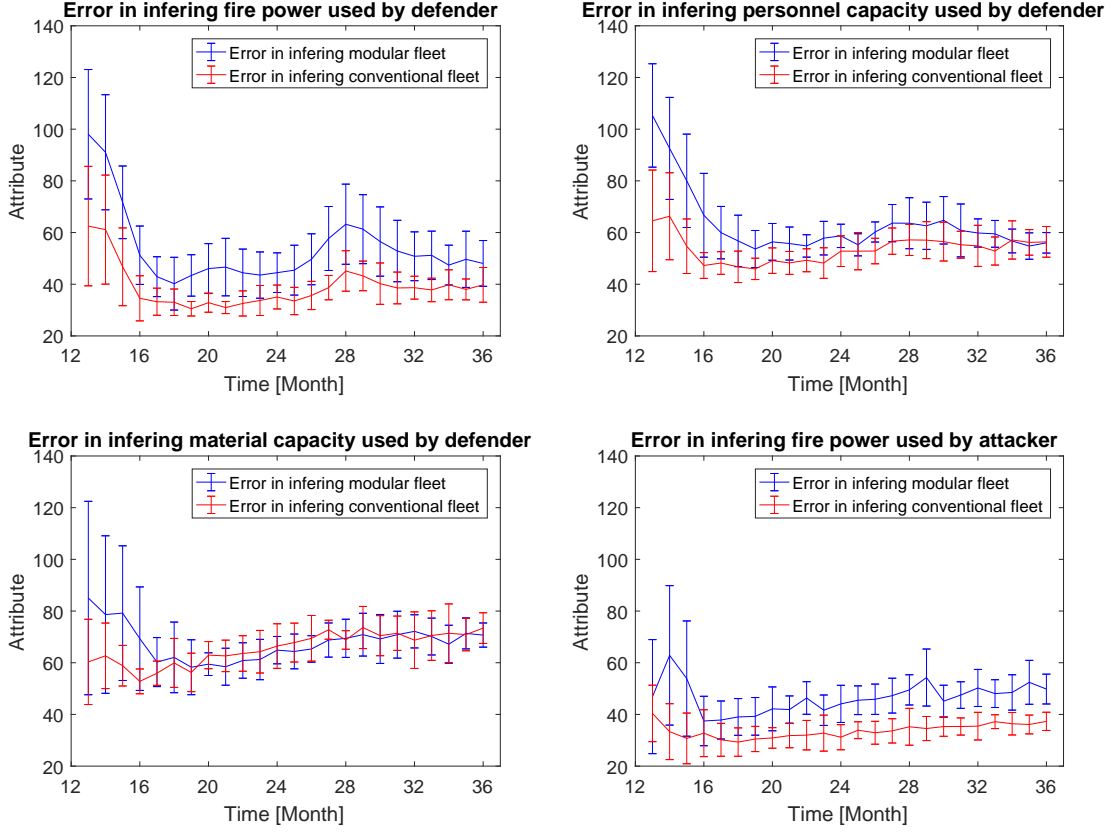


Figure 12: Comparison of inference accuracy between modular fleet and conventional fleet

The results also show that it is easier to infer the strategy of conventional fleet than modular fleet, especially in the attribute of firepower. It originates from the higher freedom in decision making after fleet modularity. As a defender, fleet usually needs to prepare a convoy with all types of attributes to satisfy the demands. With limited vehicle stocks, decision maker of conventional fleet has constrained choice of strategy. However, for the modular fleet, they could vary the dispatch strategy by real-time vehicle reconfiguration, i.e., reconfigure cargos to combat vehicles to achieve the switch from Strategy 4 to Strategy 8.

However, the burden of modularity is also significant, which is the high acquisition of capacity. According to Fig.13, modular fleet always requires more machines than conventional fleet because of additional ADR actions. It can also be observed that machine requirements are increased significantly once entering the learning stage, which comes from damage from smarter strikes by enemy. The higher losses in conventional fleet also shrink the difference in machine usage at learning stage. In this study, we only test the fleet performance at a certain capacity, studies investigating on the influence of capacity can be found in

the following literature (Li and Epureanu, 2017b, 2018).

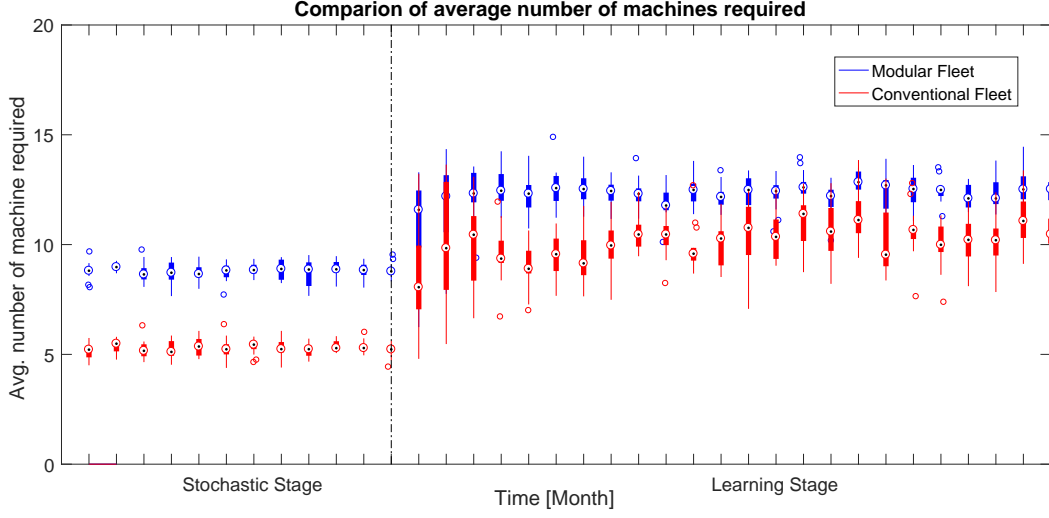


Figure 13: Comparison of machine requirements between modular fleet and conventional fleet

7. Conclusions

In this paper, we investigate the benefits and burdens from fleet modularization by simulating an attacker-defender game between modular fleet and conventional fleet. We simulate the fleet competition for three years which are divided into stochastic stage and learning stage. By contrasting the simulation results from two fleets, we find that conventional fleet stay the leads when both fleets are selecting strategies stochastically; modular fleet outperforms conventional fleet once intelligence of the decision maker is considered. With additional operational flexibility from on-base ADR actions, modular fleet exhibits a better adaptability in reacting to the enemy’s actions, higher upper limit in convoy formation and a more significant unpredictability from the additional flexibility in operation.

Acknowledgment

This research has been supported by the Automotive Research Center, a US Army Center of Excellence in Modeling and Simulation of Ground Vehicle Systems, headquartered at the University of Michigan. This support is gratefully acknowledged. The authors are solely responsible for opinions contained herein.

References

- Adhitya, A., Srinivasan, R., and Karimi, I. A. (2007). A model-based rescheduling framework for managing abnormal supply chain events. *Computers & Chemical Engineering*, 31(5):496–518.
- Atsalakis, G. S., Atsalaki, I. G., and Zopounidis, C. (2018). Forecasting the success of a new tourism service by a neuro-fuzzy technique. *European Journal of Operational Research*, 268(2):716–727.
- Azaiez, M. N. and Bier, V. M. (2007). Optimal resource allocation for security in reliability systems. *European Journal of Operational Research*, 181(2):773–786.
- Bayrak, A., Egilmez, M., Kaung, H., Xingyu, L., Park, J., Umpfenbach, E., Anderson, E., Gorsich, D., Hu, J., Papalambros, P., et al. (2016). A system of systems approach to the strategic feasibility of modular vehicle fleets. *IEEE Transactions on Systems Man and Cybernetics*.
- Chen, J. Y. and Barnes, M. J. (2013). Human-agent teaming for multi-robot control: A literature review. Technical report, Army Research Lab Aberdeen Proving Ground Md.
- Chen, K., Zhou, Y., and Dai, F. (2015). A lstm-based method for stock returns prediction: A case study of china stock market. In *Big Data (Big Data), 2015 IEEE International Conference on*, pages 2823–2824. IEEE.
- Dasch, J. M. and Gorsich, D. J. (2016). Survey of modular military vehicles: benefits and burdens. Technical report, Army Tank Automotive Research, Development and Engineering Center (TARDEC) Warren United States.
- Di Persio, L. and Honchar, O. (2016). Artificial neural networks approach to the forecast of stock market price movements. *International Journal of Economics and Management Systems*, 1:158–162.
- Dresher, M. (1961). Games of strategy: theory and applications. Technical report, Rand Corp Santa Monica CA.
- Evers, L., Barros, A. I., Monsuur, H., and Wagelmans, A. (2014). Online stochastic uav mission planning with time windows and time-sensitive targets. *European Journal of Operational Research*, 238(1):348–362.
- Fischer, T. and Krauss, C. (2017). Deep learning with long short-term memory networks for financial market predictions. *European Journal of Operational Research*.

- Hagan, M. T., Demuth, H. B., Beale, M. H., et al. (1996). *Neural network design*, volume 20. Pws Pub. Boston.
- Hausken, K. and Levitin, G. (2009). Minmax defense strategy for complex multi-state systems. *Reliability Engineering & System Safety*, 94(2):577–587.
- Hausken, K. and Zhuang, J. (2011). Defending against a terrorist who accumulates resources. *Military Operations Research*, pages 21–39.
- Hecht-Nielsen, R. (1992). Theory of the backpropagation neural network. In *Neural networks for perception*, pages 65–93. Elsevier.
- Jose, V. R. R. and Zhuang, J. (2013). Technology adoption, accumulation, and competition in multi-period attacker-defender games. *Military Operations Research*, 18(2):33–47.
- Kardes, E. and Hall, R. (2005). Survey of literature on strategic decision making in the presence of adversaries. Technical report, CREATE Homeland Security Center.
- Lam, K.-m. and Leung, H.-f. (2006). Formalizing risk strategies and risk strategy equilibrium in agent interactions modeled as infinitely repeated games. In *Pacific Rim International Workshop on Multi-Agents*, pages 138–149. Springer.
- Landa, M. d. (1991). *War in the age of intelligent machines*. Zone Books.
- Li, X. and Epureanu, B. I. (2017a). Intelligent agent-based dynamic scheduling for military modular vehicle fleets. In *IIE Annual Conference. Proceedings*, pages 404–409. Institute of Industrial and Systems Engineers (IISE).
- Li, X. and Epureanu, B. I. (2017b). Robustness and adaptability analysis of future military modular fleet operation system. In *ASME 2017 Dynamic Systems and Control Conference*, pages V002T05A003–V002T05A003. American Society of Mechanical Engineers.
- Li, X. and Epureanu, B. I. (2018). An agent-based approach for optimizing modular vehicle fleet operation. *Review of International Journal of Production Economics*.
- Lynch, J. (2015). On Strategic Unpredictability. <https://mwi.usma.edu/2015223on-strategic-unpredictability>. [Online; accessed 2 March 2015].

- Ma, X., Tao, Z., Wang, Y., Yu, H., and Wang, Y. (2015). Long short-term memory neural network for traffic speed prediction using remote microwave sensor data. *Transportation Research Part C: Emerging Technologies*, 54:187–197.
- Mikolov, T., Karafiát, M., Burget, L., Černocký, J., and Khudanpur, S. (2010). Recurrent neural network based language model. In *Eleventh Annual Conference of the International Speech Communication Association*.
- Onggo, B. S. and Karatas, M. (2016). Test-driven simulation modelling: A case study using agent-based maritime search-operation simulation. *European Journal of Operational Research*, 254(2):517–531.
- Paulson, E. C., Linkov, I., and Keisler, J. M. (2016). A game theoretic model for resource allocation among countermeasures with multiple attributes. *European Journal of Operational Research*, 252(2):610–622.
- Powell, R. (2007). Defending against terrorist attacks with limited resources. *American Political Science Review*, 101(3):527–541.
- Rezaee, M. J., Jozmaleki, M., and Valipour, M. (2018). Integrating dynamic fuzzy c-means, data envelopment analysis and artificial neural network to online prediction performance of companies in stock exchange. *Physica A: Statistical Mechanics and its Applications*, 489:78–93.
- Roos, P. and Nau, D. (2010). Risk preference and sequential choice in evolutionary games. *Advances in Complex Systems*, 13(04):559–578.
- Shinkman, P. D. (2014). Trashed: US Gear in Afghanistan to be Sold. <https://www.usnews.com/news/articles/2014/06/04/us-military-equipment-in-afghanistan-to-be-sold-scrapped>. [Online; accessed 4 April 2014].
- Wang, C. and Bier, V. M. (2011). Target-hardening decisions based on uncertain multiattribute terrorist utility. *Decision Analysis*, 8(4):286–302.
- Xu, J., Zhuang, J., and Liu, Z. (2016). Modeling and mitigating the effects of supply chain disruption in a defender–attacker game. *Annals of Operations Research*, 236(1):255–270.
- Yu, L., Wang, S., and Lai, K. K. (2009). An intelligent-agent-based fuzzy group decision making model for financial multicriteria decision support: The case of credit scoring. *European journal of operational research*, 195(3):942–959.

- Zhang, J., Zhuang, J., and Jose, V. R. R. (2018). The role of risk preferences in a multi-target defender-attacker resource allocation game. *Reliability Engineering & System Safety*, 169:95–104.
- Zhao, Z., Chen, W., Wu, X., Chen, P. C., and Liu, J. (2017). Lstm network: a deep learning approach for short-term traffic forecast. *IET Intelligent Transport Systems*, 11(2):68–75.
- Zhuang, J., Bier, V. M., and Alagoz, O. (2010). Modeling secrecy and deception in a multiple-period attacker–defender signaling game. *European Journal of Operational Research*, 203(2):409–418.