Article

Efficient Decision Approaches for Asset-Based Dynamic Weapon Target Assignment by a Receding Horizon and Marginal Return Heuristic

Kai Zhang , Deyun Zhou , Zhen Yang , Yiyang Zhao and Weiren Kong

School of Electronics and Information, Northwestern Polytechnical University, Xi'an 710072, China; dyzhou@nwpu.edu.cn (D.Z.); nwpuyz@mail.nwpu.edu.cn (Z.Y.); zhaoyiyang@mail.nwpu.edu.cn (Y.Z.); k@mail.nwpu.edu.cn (W.K.)

* Correspondence: zhangkainwpu@mail.nwpu.edu.cn

Received: 21 August 2020; Accepted: 11 September 2020; Published: 15 September 2020

Abstract: The weapon-target assignment problem is a crucial decision support in a Command and Control system. As a typical operational scenario, the major asset-based dynamic weapon target assignment (A-DWTA) models and solving algorithms are challenging to reflect the actual requirement of decision maker. Deriving from the “shoot–look–shoot” principle, an “observe–orient–decide–act” loop model for A-DWTA (OODA/A-DWTA) is established. Focus on the decide phase of the OODA/A-DWTA loop, a novel A-DWTA model, which is based on the receding horizon decomposition strategy (A-DWTA/RH), is established. To solve the A-DWTA/RH efficiently, a heuristic algorithm based on statistical marginal return (HA-SMR) is designed, which proposes a reverse hierarchical idea of “asset value–target selected–weapon decision.” Experimental results show that HA-SMR solving A-DWTA/RH has advantages of real-time and robustness. The obtained decision plan can fulfill the operational mission in the fewer stages and the “radical–conservative” degree can be adjusted adaptively by parameters.

Keywords: weapon target assignment; OODA; heuristic algorithm; combinatorial optimization; decision support system

1. Introduction

Weapon target assignment (WTA), which is also known as Weapon Allocation or Weapon Assignment (WA), refers to the reactive assignment of defensive weapons to counter identified threats [1]. With the development of advanced weapons and combat theory, it is difficult for human decision-makers to counter the fire allocation problem effectively in the complex operational environment [2]. WTA is studied as a critical problem in an intelligent decision support system in order to reduce the decision pressure of human decision-makers or replace them.

The WTA problem was first introduced by Manne in 1959 [3]. In the following decades, various types of WTA problems have evolved. From the decision process, the WTA problem can be divided into static WTA (SWTA) [4] and dynamic WTA (DWTA) [5,6]. The difference between SWTA and DWTA is whether the time is considered as a dimension. SWTA launches weapons in a salvo to maximize operational effectiveness. DWTA assigns the sequence of weapons for the equilibrium plan during multi-stage [7]. According to the operational mission, there are mainly target-based WTA (T-WTA) [8], asset-based WTA (A-WTA) [9] and sensor-WTA (S-WTA) [10,11] model. The T-WTA model adopts the kill effectiveness of weapons against targets as the optimization objective. The purpose of the A-WTA problem is to maximize the survival values of own assets. S-WTA considers the collaborations between sensors and weapons. This paper focuses on the A-DWTA problem.

A-DWTA problem is a more complex WTA problem, combining the characteristics of DWTA and A-WTA problems, and most WTA models can be seen as a special case of A-DWTA model. For example, the DWTA model can be considered as the zero-asset value of the A-DWTA model, and the A-WTA model can be considered as the one-stage model of the A-DWTA. Compared with the other WTA problems, the A-DWTA problem is less studied and the research frameworks are different. In the early research, a two-stage A-DWTA model is presented, which considers the dynamic as the repetition of the current stage and next stage. The defense is allowed to observe the outcomes of all engagement of the previous time stage before assigning and committing weapons for the present stage [12,13]. The decision variable of the two-stage model is the number of weapons assigned to the defense of assets, therefore this model has the assumption that the kill probability of a weapon-target pair depends solely on the asset to which the target is aimed [14,15]. Although the two-stage model considered the case of two-stage DWTA problem, the referenced conclusion can be extended to the multi-stage model [16–18]. With the advancement of optimization theory, the multi-stage model gradually becomes the major of WTA research problem [19]. The objective of the multi-stage model is to maximize the operational effectiveness from the current stage to the predetermined end stage, and the number and types of incoming targets of the remaining stages are known as a probability distribution.

The WTA problem is proved to be NP-complete, and the objective function is convex [20]. The exact algorithms solving the WTA model have a “dimension explosion” dilemma, which does not satisfy the real-time requirements in practical applications [21,22]. Hence, the alternate approaches, such as model transformation, heuristics algorithm, are studied to approximate the optimal solution [23–25]. The swarm intelligence algorithm, which is widely used to solve WTA models, has fewer limitations on constraints, and the computational complexity is less sensitive to the problem scale [26,27]. For example, Chen et al. [28] investigated the performance of three evolutionary algorithms, solving the proposed A-DWTA model. Chang et al. [29] put forward an improved artificial bee colony algorithm based on the ranking selection and elite guidance. Wang [30] designed a discrete particle swarm algorithm based on the intuitionistic fuzzy entropy. Lai and Wu [31] proposed an improved simplified swarm optimization with deterministic initialization and target exchange scheme. However, the population-based random search mechanism lacks rigorous theoretical proofs in terms of convergence. In recent years, researches have introduced heuristic information based on domain knowledge into the weapon-target assignment [24,32–34]. Domain knowledge generally comes from expert knowledge or empirical rules, and heuristic information based on domain knowledge construction can guide algorithms to perform the deterministic search, which can quickly obtain feasible near-optimal solutions. For example, Ahuja R K et al. [35] calculated the transfer return of weapon nodes and used it as heuristic information to optimize the initial solution obtained by the deterministic algorithm. Lee et al. [36] proposed the weapon-priority factor to obtain the good gene of a genetic algorithm. Xin et al. proposed two tabu search algorithms integrating the virtual permutation and local search to solve the A-DWTA problem, and [37] took the “weapon-target-stage” benefits in the A-DWTA model as heuristic information and distributed the more benefits in turn according to the principle of maximizing weapon killing efficiency. Leboucher et al. presented a dynamic weapon target assignment for area air defense [38], and proposed a multi-objective optimization method based on the evolutionary game theory to solve the A-DWTA problem, which has the advantage of being real-time [39]. Li et al. [40–43] investigated three multi-objective optimizers of DMOEA- ϵ C, NSGA-II, and MOEA/D-AWA on solving the multi-stage DWTA problem. Mei et al. [44] proposed a heuristic algorithm, which is derived from receding horizon control, to solve the DWTA problem. Chang [29] presented the target-priority factor with a random sequence and Cannikin Law to improve the initial population.

The objectives of the major A-DWTA models are derived from the classic static WTA model and are formulated as the asset value expectation on whole stages. After the multi-stage process of A-DWTA is over, the whole-stage return can indeed evaluate the decision-making performance. However, in the actual operational process, the C2 system makes it challenging to use the global

optimization goal as the decision metric of the current stage. There are three main reasons: (1) when calculating the global return, the variables of the subsequent stages can only be described by the probability distribution. Hence the farther prediction from the current stage, the lower the accuracy and real-time performance. (2) Due to the uncertainty of multi-stage evolution, the number of offensive and defensive stages is usually set artificially. The operational experience shows that both sides pursue killing targets in as few stages as possible and ending the combat quickly, so the number of stages should also be one optimization goal. (3) Under the premise that the number of stages can not be determined, the offensive and defensive parties' primary operational goal is to kill the target efficiently and high cost-effectively in the near stage. Therefore, the decision-maker usually only considers the assignment return of limited stages, rather than the overall stages. In the DWTa problem, the consensus is that the tactical operation should follow the "shoot-look-shoot" strategy [5,45]; that is, the current stage k decides the shooting plan by observing the situation after the shoot of stage $k - 1$, which is a typical Markov Decision Process (MDP) [46]. The "look" and "shoot" phases correspond to the "observe" and "act" phases of the known "observe-orient-decide-act" (OODA) loop. Hence, the A-DWTa model should be adequately described as an OODA loop system.

Optimizing the objective (21) also faces the following challenges: (1) the decision space is difficult to obtain the optimal global solution of the entire confrontation process while satisfying real-time performance. Assuming that the initial number of weapons is m , the number of targets is n , and there are s_m decision-making stages, then the search space scale of the dynamic weapon-target allocation problem is at least m^{s_m+n} . It is difficult to get the theoretical optimal solution for the medium and large-scale weapon-target allocation problem. (2) The uncertainty of multi-stage prediction. Decision-makers can only observe the system status of the current stage, such as target survival, asset survival, and weapon-target attack conditions. The certainty of the prediction information in the subsequent stages decreases with the extrapolation of the prediction horizon. Therefore, it is hard to achieve multi-stage global optimization by predicting system evolution.

Motivated by the above analysis, we described the OODA/A-DWTa loop system, and decomposed the A-DWTa model by receding horizon strategy (A-DWTa/RH). Then, a heuristic algorithm based on the statistical marginal return (HA-SMR) is proposed to obtain the optimal plan for a decision-maker. The contributions of this paper are described as follows:

1. An OODA/A-DWTa loop model is established for supporting the following A-DWTa decision model and solving algorithm.
2. To reflect the actual operational requirement, an A-DWTa decision model based on the receding horizon strategy is presented. The "radical-conservative" degree of the obtained plan, which relates to the number of decision stages, can be adaptively adjusted by the model parameter.
3. A heuristic algorithm based on statistical marginal return is proposed to solve the A-DWTa model, which has the advantage of robust and real-time. The extensive experiments demonstrate the effectiveness of the proposed approaches.

The rest of this paper is organized as follows. Section 2 gives our motivations and formulates the OODA/A-DWTa loop model. Section 3 formulates the objective and constraints of the A-DWTa decision model. Section 4 presents the proposed HA-SMR for the A-DWTa model. Section 5 verifies the proposed HA-SMR solving the A-DWTa problem by experimental studies. The conclusion is finally summarized in Section 6.

2. Ooda/A-Dwta Loop Model

The OODA loop theory was put forward by John Boyd in the 1970s, and his theoretical viewpoint is that the execution process of a fire strike can be divided into the observe-orient-decide-act cycle. In the conflict, the party who completes the OODA loop faster has the advantage. Therefore, both sides' antagonistic goal is to accelerate their own OODA cycle, cut off, or delay the enemy OODA cycle

process. After half a century of testing, the OODA loop theory has been widely used in military, economic warfare, competitive sports, and other fields.

The operational scenario of A-DWTA is: enemy units launch an attack to destroy the assets with military value but no defensive ability, such as military installations, personnel gathering placing, command and control nodes, weapons warehouses and ports. At this time, the defense unit with combat ability is the decision-maker and is to maximize the assets' survival value in multiple attack and defense stages by allocating weapon resources to kill the enemy targets, as shown in Figure 1.

Figure 1. Schematic diagram of the asset-based weapon target assignment (A-DWTA) scenario.

The A-DWTA problem based on the “shoot–look–shoot” mechanism conforms to the typical OODA loop theory. In the observe phase, the decision-maker updates the external information (target state, asset state) of the current stage. In the orient phase, the state information collected in the observation phase is decomposed and reconstructed, such as the threat assessment based on the target track information, the identification of the attack intention of the targets against assets, the feasibility of a weapon attacking a target, and the determination of attack request. If the termination conditions are met, the weapon system is guided to exit the operational environment, otherwise continue to execute the OODA loop. In the decide phase, the specific combat requirements are converted into objective functions, and the solution is solved according to the designed weapon–target allocation algorithm. In the act phase, execute the weapon–target assignment plan given in the decide phase, update weapon state information, and evaluate the damaging effect of targets and the value of assets. A diagram of the OODA/A-DWTA loop model is shown in Figure 2. The notation employed in the context is listed in Table 1.

Figure 2. Diagram of the observe–orient–decide–act (OODA)/asset-based dynamic weapon target assignment (A-DWTA) loop model.

Table 1. Notation declaration.

Notation	Description
m	the number of available weapons at the initial stage;
n	the number of hostile targets at the initial stage;
l	the number of defense asset at the initial stage;
s	the index of decision stage, $s = 1, 2, \dots, s_m$;
$\mathbf{W}(s) = [w_i(s)]_{1 \times m}$	the weapon state of stage s ; $w_i(s) = 1$ denotes weapon i is available at the decide phase of stage s , otherwise $w_i(s) = 0$;
$\mathbf{T}(s) = [t_j(s)]_{1 \times n}$	the target state of stage s ; $t_j(s) = 1$ denotes target j is threatening at the observe phase of stage s , otherwise $t_j(s) = 0$;
$\mathbf{A}(s) = [a_k(s)]_{1 \times l}$	the asset state of stage s ; $a_k(s) = 1$ denotes asset k survives at the observe phase of stage s , otherwise $a_k(s) = 0$;
$\mathbf{P}(s) = [p_{ij}(s)]_{m \times n}$	the weapon–target kill probability of stage s , p_{ij} denotes the kill probability of weapon i against target j at the orient phase of stage s ;
$\mathbf{C}(s) = [c_{ij}(s)]_{m \times n}$	the weapon= target attack condition of stage s ; $c_{ij} = 1$ indicates that weapon i satisfies the attack condition of intercepting target j at stage s , otherwise $c_{ij} = 0$;
$\mathbf{Q}(s) = [q_{jk}(s)]_{n \times l}$	the target–asset intention matrix of stage s ; $q_{jk} > 0$ denotes that the attack intention of target j against asset k is assessed at the orient phase of stage s , and the destroy probability is q_{jk} , otherwise $q_{jk} = 0$;
$\mathbf{D}(s) = [d_{ij}(s)]_{m \times n}$	the weapon–target decision variable of stage s ; $d_{ij} = 1$ represents that weapon i is assigned to intercept target j at the decide phase of stage s , otherwise $d_{ij} = 0$;
$\mathbf{H}(s) = [h_j(s)]_{1 \times n}$	the target survival probability after the act phase of stage s ;
$\mathbf{U}^o(s) = [u_k^o(s)]_{1 \times l}$	the asset survival probability at the observe phase of stage s ;
$\mathbf{U}^a(s) = [u_k^a(s)]_{1 \times l}$	the asset survival probability after the act phase of stage s .

2.1. Observe Phase

In the OODA/A-DWTA loop, the observation model describes the survival state of assets and the damage state of targets at the current stage under the uncertain battlefield environment.

- (1) Target state observation. Let the current stage be s , and the target observed state $\mathbf{T}(s) = [t_j(s)]_{1 \times n}$ is the actual survival state of targets after the act phase of stage $s - 1$. Therefore, the observed target state at stage s should follow the Bernoulli distribution about the target survival probability at stage $s - 1$, namely $t_j(s) \sim B(1, h_j(s - 1))$:

$$\Pr(t_j(s) = r) = r \cdot h_j(s - 1) + (1 - r)(1 - h_j(s - 1)), \text{ for } r \in \{0, 1\}; j = 1, 2, \dots, n, \quad (1)$$

where $t_j(s) = 0$ indicates that the observed state of the target j in observation phase is damaged; conversely, $t_j(s) = 1$ represents target j as a threat target in observation phase; $h_j(s - 1)$ is the target survival probability expectation after the end of stage $s - 1$, which is given in the damage assessment of the act phase.

- (2) Asset state observation. Similarly, the asset observed state $\mathbf{A}(s)$ of stage s follows the Bernoulli distribution of the asset survival probability expectation of stage $s - 1$. The asset survival probability expectation is related to the target survival probability expectation in the act phase of stage $s - 1$, and the target observed state can be obtained in the stage s . Therefore, in the observation model of OODA/A-DWTA, the asset survival probability expectation $\mathbf{U}^o(s - 1) = [u_k^o(s - 1)]_{1 \times l}$ of stage $s - 1$ is modeled as the posterior conditional probability of the target observed state $\mathbf{T}(s)$ of stage s :

$$u_k^o(s - 1) = \Pr(a_k(s) = 1 | \mathbf{T}(s)) = a_k(s - 1) \prod_{j=1}^n (1 - q_{jk}(s - 1))^{t_j(s)}, \text{ for } k = 1, 2, \dots, l \quad (2)$$

The asset observed state $\mathbf{A}(s) = [a_k(s)]_{1 \times l}$ follows $a_k(s) \sim B(1, u_k^o(s-1))$:

$$\begin{aligned} \Pr(a_k(s) = r) &= r \cdot u_k^o(s-1) + (1-r)(1-u_k^o(s-1)) \\ &= r \cdot a_k(s-1) \prod_{j=1}^n (1-q_{jk}(s-1))^{t_j(s)} + (1-r) \left(1 - a_k(s-1) \prod_{j=1}^n (1-q_{jk}(s-1))^{t_j(s)} \right), \quad (3) \\ &\text{for } r = 0, 1; k = 1, 2, \dots, l \end{aligned}$$

where $a_k(s) = 0$ represents that the observed state of asset k at stage s is survival; conversely, $a_k(s) = 1$ represents that the observed state of asset k is destroyed.

2.2. Orient Phase

In the OODA/A-DWTA loop, the orient model is used to determine the attack request, and update the target-asset attack intention, weapon-target attack condition by the observation information.

- (1) Attack request determination. When any of the following conditions are met, the OODA/A-DWTA loop terminates, otherwise initiates an attack request and enters the decide phase: (a) all assets are destroyed, $\|\mathbf{A}(s)\|_1 = 0$, where $\|\cdot\|_1$ is the l1-norm; (b) all targets are damaged, $\|\mathbf{T}(s)\|_1 = 0$; (c) no available weapon, $\|\mathbf{W}(s)\|_1 = 0$. So the termination condition of OODA/A-DWTA is $\|\mathbf{T}\|_1 \cdot \|\mathbf{W}\|_1 \cdot \|\mathbf{A}\|_1 = 0$.
- (2) Target-asset attack intention. In actual combat, the identification of enemy target's attack intention requires the analysis of target altitude, distance, speed, acceleration, heading angle, azimuth, fire control radar state, maneuver type, and is predicted by domain knowledge. WTA research's focus lies in the decision model and algorithm of the command and control layer, not the methods of intention recognition. Hence we simplify the target-asset intention matrix $\mathbf{Q} = [q_{jk}]_{n \times l}$ to describe the enemy targets' attack state of each stage. $q(s)_{jk} = 0$ represents that the attack intention of target j against asset k is not recognized at stage s . Otherwise, it represents that target j against asset k , and the destroy probability is $q(s)_{jk}$.
- (3) Weapon-target attack condition. The condition of weapon-target attack is mainly determined by whether the time window of the fire control launch is satisfied. In DWTA, as the offensive and defensive stages recurse, the weapons that meet the attack conditions usually decrease. We introduce the lethality matrix $\mathbf{P} = [p_{ij}]_{m \times n}$ and feasibility matrix $\mathbf{C} = [c_{ij}]_{m \times n}$ to describe the weapon-target attack condition of each stage, where $p_{ij}(s) \in [0, 1)$ represents the destroy probability of weapon i against target j ; $c_{ij}(s) = 1$ indicates that weapon i meets the attack condition of intercepting target j , otherwise $c_{ij}(s) = 0$.

2.3. Decide Phase

As the critical issue of the OODA/A-DWTA loop, the decide model aims to obtain the optimal weapon-target assignment under the current situation and maximize the survival value expectation of assets.

$$\mathbf{D}(s)^* = \arg \max_{\mathbf{D}} F(\mathbf{V}, \mathbf{Q}, \mathbf{P}, \mathbf{A}, \mathbf{T}, \mathbf{W}, \mathbf{C}, \mathbf{D}, s), \quad (4)$$

where $F(\cdot)$ is the objective function of the A-DWTA decision model, which directly reflects the operational requirement of C2 system. The solving algorithm of Formula (4) reflects the intelligent degree of weapon-target assignment. As the focus of this paper, the A-DWTA decision model will be studied in detail in the next section.

2.4. Act Phase

According to the obtained weapons-target assignment $\mathbf{D}(s)$ at the current stage, the act model needs to carry out weapon states update, target damage assessment, and asset survival value assessment.

- (1) Weapon state update. After acting the weapon allocation plan, the current stage's weapon state $\mathbf{W}(s) = [w_i(s)]_{1 \times m}$ are determined by the previous stage's weapon states $\mathbf{W}(s-1)$ and the current stage's decision plan $\mathbf{D}(s)$.

$$w_i(s) = w_i(s-1) - \sum_{j=1}^n d_{ij}(s-1), \text{ for } i = 1, 2, \dots, m \quad (5)$$

where $w_i(s) = 0$ denotes weapon i is not available at stage s , otherwise $w_i(s) = 1$.

- (2) Target damage assessment. The target survival probability is mainly determined by the target state, the weapon-target kill probability and the decision plan at current stage.

$$h_j(s) = t_j(s) \prod_{i=1}^m (1 - p_{ij}(s))^{d_{ij}(s)}, \text{ for } j = 1, 2, \dots, n, \quad (6)$$

where (6), $h_j(s) \in [0, 1]$. If and only if $h_j(s) \in [0, 1]$, that is, the target j is confirmed as damaged by the observe phase of the stage s , $h_j(s) = 0$; if and only if $t_j(s) = 1$ and $\sum_{i=1}^m d_{ij} = 0$, that is, no weapon is not assigned to the target j with threat capability, $h_j(s) = 1$.

- (3) Asset damage assessment. The asset survival value expectation $\mathbf{F}(s)$ of stage s is related to the asset value $\mathbf{V}$ and the asset survival probability expectation in the act phase of stage s . In the act model of OODA/A-DWTA, the asset survival probability expectation $\mathbf{U}^a(s) = [u_k^a(s)]_{1 \times l}$ of stage s is modelled as the conditional probability under the target survival probability $\mathbf{H}(s)$

$$u_k^a(s) = u_k(s) | \mathbf{H}(s) = a_k(s) \prod_{j=1}^n (1 - h_j(s) q_{jk}(s)), \text{ for } k = 1, 2, \dots, l, \quad (7)$$

where the term $h_j(s) q_{jk}(s)$ represents the damage probability of target j attacking asset k in the act phase of stage s . The asset survival value expectation $\mathbf{F}(s) = [f_k(s)]_{1 \times l}$ is evaluated as

$$f_k(s) = v_k \cdot u_k^a(s) = v_k a_k(s) \prod_{j=1}^n \left(1 - t_j(s) q_{jk}(s) \prod_{i=1}^m (1 - p_{ij}(s))^{d_{ij}(s)} \right), \text{ for } k = 1, 2, \dots, l \quad (8)$$

where $f_k(s) \in [0, v_k]$. If and only if $a_k(s) = 0$, that is, the asset k is confirmed as destroyed by the observe phase of the stage s , $e_k(s) = 0$; if and only if $a_k(s) = 1$ and $\sum_{j=1}^n q_{jk}(s) = 0$, that is, no target has attack intention to the target j , $e_k(s) = v_k$.

In summary, the model of OODA/A-DWTA loop model is

$$\begin{cases}
\Pr(t_j(s) = r) = r \cdot h_j(s-1) + (1-r)(1-h_j(s-1)), \text{ for } r = 0, 1; j = 1, 2, \dots, n \\
h_j(s-1) = t_j(s-1) \prod_{i=1}^m (1-p_{ij}(s-1))^{d_{ij}(s-1)}, \text{ for } j = 1, 2, \dots, n \\
\Pr(a_k(s) = r) = r \cdot u_k^o(s-1) + (1-r)(1-u_k^o(s-1)), \text{ for } r = 0, 1; k = 1, 2, \dots, l \\
u_k^o(s-1) = a_k(s-1) \prod_{j=1}^n (1-q_{jk}(s-1))^{t_j(s)}, \text{ for } k = 1, 2, \dots, l \\
\mathbf{D}^*(s) = \arg \max_{\mathbf{D}} F(\mathbf{V}, \mathbf{Q}, \mathbf{P}, \mathbf{A}, \mathbf{T}, \mathbf{W}, \mathbf{C}, \mathbf{D}, s) \\
f_k(s) = v_k \cdot u_k^a(s), \text{ for } k = 1, 2, \dots, l \\
u_k^a(s) = a_k(s) \prod_{j=1}^n (1-h_j(s)q_{jk}(s)), \text{ for } k = 1, 2, \dots, l \\
w_i(s) = w_i(s-1) - \sum_{j=1}^n d_{ij}(s-1), \text{ for } i = 1, 2, \dots, m \\
s_m = \{s \mid \|\mathbf{A}(s)\|_1 \cdot \|\mathbf{T}(s)\|_1 \cdot \|\mathbf{W}(s)\|_1 = 0\}
\end{cases} \quad (9)$$

3. A-Dwta/Rh Formulation

In the A-DWTA decision model, it is assumed that there are s_m stages of attack and defense decision-making before the enemy targets are annihilated, or assets are destroyed. In stage s , the problem variables are the destroy effectiveness $\mathbf{Q}$ of enemy targets against assets obtained by threat assessment; the survival value $\mathbf{V}$ of assets; the defensive effectiveness $\mathbf{P}$ of weapons against targets and the attack constraints $\mathbf{C}$ of weapon–target. The output variable is the defense decision matrix $\mathbf{D}$. The research goal of the A-DWTA problem is to solve the decision model through algorithms under the current situation $\{\mathbf{V}, \mathbf{Q}, \mathbf{P}, \mathbf{A}, \mathbf{T}, \mathbf{W}, \mathbf{C}, \mathbf{D}, s\}$, obtain an optimal firepower decision plan $\mathbf{D}^*$, and maximize the effectiveness of reducing enemy targets' threat to assets.

$$\mathbf{D} = \arg \max_{\mathbf{D}} F(\mathbf{V}, \mathbf{Q}, \mathbf{P}, \mathbf{A}, \mathbf{T}, \mathbf{W}, \mathbf{C}, \mathbf{D}, s) \quad (10)$$

where the design of the objective function F directly reflects the operational purpose of the command and control system, and the constraints C reflect the battlefield environment. The solving algorithm directly reflects the reasonable degree of firepower decision.

To solve the problem that the standard A-DWTA model pursues the theoretical optimal solution in mathematics and is not tightly integrated with the operational requirements, this section establishes a two-stage A-DWTA model based on the Receding Horizon Strategy (A-DWTA/RH).

3.1. Objective Building

The major objective function of A-DWTA model is derived from the classic static WTA model, and is formulated as the sum of the asset value expectation at each stage, such as

$$\max J_t(X^t) = \sum_{k=1}^l v_k \prod_{j=1}^n \left[1 - q_{jk} \prod_{s=t}^{s_m} \prod_{i=1}^m (1 - p_{ij}(s))^{x_{ij}(s)} \right] \text{ with } t \in \{1, 2, \dots, s_m\} \quad (11)$$

where s_m is the maximum number of decision stages; decision set $\mathbf{X}^t = \{X_t, X_{t+1}, \dots, X_{s_m}\}$ is the global decision sequence starting from the current stage t , where $X_s = [x_{ij}(s)]_{m \times n}$ is the local decision of stage s .

For solving the problem (11), the related algorithms are to set the maximum stages s_m , predict the targets and assets status of all subsequent stages at the current stage, search the current global optimal solution for stage one to s_m . In response to this problem, we believe that the maximum number of stages is one of the most crucial algorithm metrics in the A-DWTA problem. The primary task of decision-makers is still to kill the targets efficiently and protect assets from damage cost-effectively in

the fewest stages. Secondly, considering the uncertainty of weapons killing targets, it is still necessary to ensure the remaining weapons' ability to counter survival targets in subsequent stages. Following the above idea, we propose a two-stage A-DWTA model based on the receding horizon strategy. The design of the objective function follows two points: (1) from the perspective of the decision-maker, it should first ensure the efficient and cost-effective killing of targets at the current stage to achieve the purpose of protecting our assets. (2) Assuming the targets have not been killed after the act of the current stage, the remaining weapons should have sufficient countermeasures against the enemy target. In summary, the objective function of the A-DWTA/RH model is composed of two parts. The first part is the direct income expectation of weapons killing targets in the current stage, and the second part is the prediction return of the remaining weapons against the undamaged target in the next stage, as shown

$$\max F(\mathbf{V}, \mathbf{Q}, \mathbf{P}, \mathbf{A}, \mathbf{T}, \mathbf{W}, \mathbf{C}, \mathbf{D}, s) = \lambda f_1(\mathbf{V}, \mathbf{Q}, \mathbf{P}, \mathbf{A}, \mathbf{D}, s) + (1 - \lambda) f_2(\mathbf{V}, \mathbf{Q}, \mathbf{P}, \mathbf{U}^a, \mathbf{H}, \mathbf{W}, s), \quad (12)$$

where f_1 represents the absolute return expectation of the decision variable $\mathbf{b}$ at the current stage; f_2 represents the threat return of the remaining weapons $\overline{\mathbf{W}}(s)$ to the target under the prediction situation of the stage $s + 1$, and $\overline{\mathbf{W}}(s) = \mathbf{W}(s) \setminus \mathbf{D}(s)$ denotes the remaining available weapon set after the action process of stage s . $\mathbf{H}(s)$ and $\mathbf{U}^a(s)$ respectively represent the target survival probability and asset survival probability under the situation $\{\mathbf{V}, \mathbf{Q}, \mathbf{P}, \mathbf{A}, \mathbf{T}, \mathbf{W}, \mathbf{C}, s\}$ and the decision variable $\mathbf{D}(s)$, calculated according to the system state transfer Equation (9). λ is a weight parameter reflecting the decision maker's preference for f_1 and f_2 in the objective function. f_1 and f_2 are normalized.

3.1.1. Absolute Return Expectation at the Current Stage

The absolute return expectation at the current stage can be expressed by the combat effectiveness of the used weapons at the current stage: the survival value expectation of assets under the weapon-target assignment scheme. Let the current stage be the stage s and the decision variable is $\mathbf{D}(s)$, then the normalized combat effectiveness of the current stage is

$$f_1 = \frac{1}{\|\mathbf{V} \cdot \mathbf{A}(s)\|_1} \sum_{k=1}^l v_k a_k(s) \prod_{j=1}^n \left(1 - q_{jk}(s) h_j(s)\right). \quad (13)$$

Substitute the Formula (6) into the above formula to get

$$f_1 = \frac{1}{\|\mathbf{V} \cdot \mathbf{A}(s)\|_1} \sum_{k=1}^l v_k a_k(s) \prod_{j=1}^n \left(1 - q_{jk}(s) \prod_{i=1}^n (1 - p_{ij}(s))^{d_{ij}(s)}\right). \quad (14)$$

3.1.2. Return Evaluation of Remaining Weapons on Prediction Situation

Since the current stage cannot obtain the situation information such as the weapon/target/asset status and target attack intention of the next stage, it is necessary to design the return evaluation method of combat mission under the prediction situation. The designed evaluation method is: due to the event "weapons attack the target," the threat of the target to the asset will be attenuated, and the value expectation of the asset will rise. Therefore, according to the state distribution of the situation information in the next stage, the value expectation of assets under the situation of "threatening target-surviving asset" is calculated first. Then the value expectation under the situation of "surplus weapon-threatening target-surviving asset" is calculated. The difference between the two value expectation can be used as the return expectation of remaining weapons in the prediction situation.

Based on the above ideas, after the weapon-target assignment plan $\mathbf{D}(s)$ is implemented at the stage s , the OODA/A-DWTA loop established in Section 2 gives: the weapon state is $\mathbf{W}(s+1) = \mathbf{W}(s) \setminus \mathbf{D}(s)$, the target state is subject to $t_j(s+1) \sim B(1, h_j(s))$, and the asset state is subject to (6). The return expectation of each "remaining weapon-threat target-survival asset" subject to the above distribution can be used as the return of the remaining weapon in the prediction situation.

First, regardless of weapons $\mathbf{W}(s+1)$, the prediction threat of targets $\mathbf{T}(s+1)$ to assets $\mathbf{A}(s+1)$ can be evaluated as:

$$\sum_{k=1}^l \frac{1}{\|\mathbf{T}(s)\|_1} \sum_{j=1}^n v_k(s) u_k^a(s) \left(1 - h_j(s) q_{jk}(s)\right). \quad (15)$$

From the above formula, considering the damage of weapons $\mathbf{W}(s+1)$ on targets $\mathbf{T}(s+1)$, the prediction threat to assets $\mathbf{A}(s+1)$ can be evaluated as

$$\sum_{k=1}^l \frac{1}{\|\mathbf{T}(s)\|_1} \sum_{j=1}^n v_k(s) u_k^a(s) \left(1 - h_j(s) q_{jk}(s) \left(1 - \frac{\sum_{i=1}^n p_{ij}(s) w_i(s+1)}{\|\mathbf{W}(s+1)\|_1}\right)^{\frac{\|\mathbf{W}(s+1)\|_1}{\|\mathbf{T}(s)\|_1}}\right). \quad (16)$$

Normalize the difference value between Equations (15) and (16), and the prediction return of remaining weapons in the next stage is presented as

$$f_2 = \frac{1}{\|\mathbf{V}(s)\|_1 \|\mathbf{T}(s)\|_1} \sum_{k=1}^l \sum_{j=1}^n u_k^a(s) v_k(s) h_j(s) q_{jk}(s) \cdot \left(1 - \left(1 - \frac{1}{\|\mathbf{W}(s+1)\|_1} \sum_{i=1}^n p_{ij}(s) w_i(s+1)\right)^{\frac{\|\mathbf{W}(s+1)\|_1}{\|\mathbf{T}(s)\|_1}}\right). \quad (17)$$

3.2. Constraints

In the A-DWTA model, the necessary mathematical constraints include 0–1 integer constraints of decision variables, and total constraints of weapons, as shown below:

$$s.t. \begin{cases} d_{ij}(s) \in \{0, 1\}, \text{ for } i = 1, 2, \dots, m; j = 1, 2, \dots, n; s = 1, 2, \dots, s_m \\ \sum_{s=1}^{s_m} \sum_{j=1}^m d_{ij} \leq m, \text{ for } i = 1, 2, \dots, m \end{cases} \quad (18)$$

Besides, model constraints can be designed according to specific DWTA models. Commonly used model constraints include multi-target attack constraints and weapon consumption constraints on single-target.

$$s.t. \begin{cases} \sum_{j=1}^m d_{ij}(s) \leq m_i, \text{ for } i = 1, 2, \dots, m; s = 1, 2, \dots, s_m \\ \sum_{i=1}^m d_{ij}(s) \leq n_j, \text{ for } j = 1, 2, \dots, n; s = 1, 2, \dots, s_m \end{cases}, \quad (19)$$

where the first term is the multi-target attack constraint, limiting the number of targets that one weapon can attack at one stage. In the coding of WTA models, the weapon platform is usually disassembled into single-weapon, that is, $m_i = 1$. The second item is the weapon consumption constraint on single-target, which is the upper limit of the number of weapons allocated to one target at one stage. This paper considers that the motivation of constraint (19) is to control the consumption of weapons at each stage, prevent excessive consumption of weapons and the DWTA model degenerating into the SWTA model. In the auxiliary decision-making system, this constraint is not the operational index that the decision-maker concentrates on. Furthermore, the artificial setting of n_i itself adds an extra burden of decision-making and cannot directly reflect the improvement of combat effectiveness.

To quantify the cost-effectiveness ratio of the decision plan more accurately and reduce the burden on decision-makers, we convert the constraints (19) into the gradient limit of target survival

value based on the characteristics of the weapon consumption characteristic curve. The attenuation of firepower return is employed as constraint metric

$$\arg \min_{\mathbf{D}(s)} u_k^a(s) > \rho, \text{ for } k = 1, 2, \dots, l; s = 1, 2, \dots, s_m, \quad (20)$$

where ρ is the threshold parameter of weapon consumption. In summary, the proposed A-DWTA/RH model and constraints are

$$\left\{ \begin{array}{l} \max F(\mathbf{V}, \mathbf{Q}, \mathbf{P}, \mathbf{A}, \mathbf{T}, \mathbf{W}, \mathbf{C}, \mathbf{D}, s) = \lambda f_1(\mathbf{V}, \mathbf{Q}, \mathbf{P}, \mathbf{A}, \mathbf{D}, s) + (1 - \lambda) f_2(\mathbf{V}, \mathbf{Q}, \mathbf{P}, \mathbf{U}^a, \mathbf{H}, \mathbf{W}, s) \\ f_1 = \frac{1}{\|\mathbf{V} \cdot \mathbf{A}(s)\|_1} \sum_{k=1}^l v_k a_k(s) \prod_{j=1}^n \left(1 - q_{jk}(s) \prod_{i=1}^n (1 - p_{ij}(s))^{d_{ij}(s)} \right) \\ f_2 = \frac{1}{\|\mathbf{V}(s)\|_1 \|\mathbf{T}(s)\|_1} \sum_{k=1}^l \sum_{j=1}^n u_k^a(s) v_k(s) h_j(s) q_{jk}(s) \cdot \\ \left(1 - \left(1 - \frac{1}{\|\mathbf{W}(s+1)\|_1} \sum_{i=1}^n p_{ij}(s) w_i(s+1) \right)^{\frac{\|\mathbf{W}(s+1)\|_1}{\|\mathbf{T}(s)\|_1}} \right) \end{array} \right. \quad (21)$$

$$s.t. \left\{ \begin{array}{l} d_{ij}(s) \in \{0, 1\}, \text{ for } i = 1, 2, \dots, m; j = 1, 2, \dots, n; s = 1, 2, \dots, s_m \\ \sum_{s=1}^{s_m} \sum_{j=1}^n d_{ij}(s) \leq m, \text{ for } i = 1, 2, \dots, m \\ \sum_{j=1}^n d_{ij}(s) \leq 1, \text{ for } i = 1, 2, \dots, m; s = 1, 2, \dots, s_m \\ \arg \min_{\mathbf{D}(s)} u_k^a(s) > \rho, \text{ for } k = 1, 2, \dots, l; s = 1, 2, \dots, s_m \end{array} \right. \quad (22)$$

4. Ha-Smr Algorithm

This section analyzes the limitations of the heuristic information commonly used in the current WTA algorithms, proposes heuristic information based on statistical marginal return, and designs an A-DWTA solving algorithm based on this heuristic information.

4.1. Algorithm Framework

Disassembling the offensive and defensive relationship of the A-DWTA/RH decision model shows that there are two types of decision plans: (1) target–asset attack plan; (2) weapon–target assignment plan. It can be seen that the return flow of the objective function of the A-DWTA/RH model is “weapon–target–asset.” Therefore, the algorithm proposed in this paper is to perform the reverse hierarchical way of “asset value–target selected–weapon decision.” First, based on the perspective of asset defense, calculate the value loss expectation of each asset according to the current weapon–target plan, target–asset intention, and target–asset destroy probability, then select the resource with the maximal value loss expectation as the priority defense asset. Secondly, calculate the value loss expectation caused by each target attacking the priority defense asset, and select the target corresponding to the maximum value as the priority attack target. Finally, the priority weapon is solved for by heuristic information. The decision plan is updated, and the weapon is removed from the set of available weapons. Return to update the state of the decision-making system and repeat the above operations until the algorithm meets the termination condition. In the above iterative process, the diversity of the decision set and the algorithm’s real-time performance are maintained by the objective function.

The main loop of MOEA-CMWTa is outlined in Algorithm 1. The flow diagram of HA-SMR is shown in Figure 3.

Algorithm 1 Main loop of heuristic algorithm based on statistical marginal return (HA-SMR).**Input:** $\{V, P, Q, C, A, T, W, s\}$: the current situation.**Output:** D_b : the output solution of SMRA.

- 1: Initial the weapon-target decision matrix $D(s) = [d_{ij}(s)]_{m \times n}$, $\forall x_{ij} = 0$, the number of plans $u = 1$;
- 2: Calculate the target value expectation $V' = [v'_k]_{1 \times l}$, target-asset value loss expectation matrix $L(s) = [l_{jk}(s)]_{n \times l}$ by decision matrix $D(s)$, weapon-target kill probability $P(s)$, target-asset destroy probability $Q(s)$, and asset value $V(s)$.
- 3: Gain the priority defense asset a_p and priority attack target t_p by target-asset value loss expectation matrix $L(s)$;
- 4: Calculate the weapon-target marginal return expectation matrix $E = [e_{ij}]_{m \times n}$ according to the weapon-target situation $\{D(s), P(s), C(s)\}$;
- 5: Calculate the heuristic information $g_i, i = 1, 2, \dots, m$ of each weapon-target pair. Select d weapons as available weapons with higher priority;
- 6: Generate d individuals in the solution set $D = \{D^1(s), D^2(s), \dots, D^u(s)\}$ according to the priority weapons in Step 5 and move the used weapons out of the available weapon set in the respective decision space, then the number of solutions of feasible schemes is $u = u \cdot d$;
- 7: Update the weapon-target feasibility set $C = \{C^1(s), C^2(s), \dots, C^u(s)\}$ corresponding the set D . If the termination condition $C^h(s) = O$ is meet, remove the solution $D^h(s)$ out of the D , and update the historical optimal solution D_b ;
- 8: If the number of solutions exceeds the set upper limit of solution set, calculate the objective function of the solutions, and keep the optimal u_m solution individuals;
- 9: If solution set $D = \emptyset$, the algorithm terminates, otherwise return to Step 2.

Figure 3. Flow diagram of heuristic algorithm based on statistical marginal return.

4.1.1. Priority of Target and Asset

To facilitate the understanding of the algorithm proposed in this chapter, first define the key variable between the “weapon–target” and “target–asset” situations in the A-DWTA model: the equivalent conditional kill probability.

Definition 1 (Equivalent Conditional Kill Probability). *In the A-DWTA model, considering the determined weapon–target assignment plan, the expected kill probability of targets against assets under the “weapon–target–asset” situation is defined as equivalent conditional kill probability.*

First, let the current weapon–target decision matrix be $\mathbf{D}$, the target–asset destroy matrix be $\mathbf{Q}$, and the equivalent kill condition probability matrix of weapon–target–asset be $\mathbf{Q}' = [q'_{jk}]_{n \times l}$

$$q'_{jk} = q_{jk} | \{\mathbf{D}, \mathbf{P}\} = q_{jk} \prod_{i=1}^m (1 - p_{ij})^{d_{ij}}, \text{ for } j = 1, 2, \dots, n; k = 1, 2, \dots, l. \quad (23)$$

Second, calculate the target–asset loss value expectation matrix $\mathbf{L} = [g_{jk}]_{n \times l}$ based on the target–asset equivalent condition damage probability $\mathbf{Q}'(t)$ and the asset survival value $\mathbf{V}$

$$\mathbf{L} = \mathbf{Q}' \mathbf{I}_l \mathbf{V}, \quad (24)$$

where $\mathbf{I}_l = [1]_{l \times 1}$. At this time, the value of the element g_{jk} represents the loss value expectation of the two-tuple “target j –asset k ” under the current decision matrix $\mathbf{D}$. According to the maximum marginal return strategy, the maximum element in the loss value expectation matrix $\mathbf{D}$ is extracted to determine the priority defense asset A , and then the priority attack target B is selected, namely

$$\begin{cases} t_p = \arg \max_j l_{jk} \\ a_p = \arg \max_k l_{t_p k} \end{cases} \quad (25)$$

4.1.2. Weapon–Target Marginal Return Expectation

Theorem 1. *Under the current decision D and situation $\{\mathbf{V}, \mathbf{P}, \mathbf{Q}, \mathbf{C}, \mathbf{A}, \mathbf{T}, \mathbf{W}\}$, the weapon–target marginal return expectation matrix $\mathbf{E}$ of the A-DWTA model is*

$$\begin{cases} e_{ij} = v_{k_a} \prod_{h=1}^n (1 - q'_{hk_a})^{\frac{p_{ij} q_{jk_a}}{1 - q_{jk_a}}} \\ k_a = \{k | q_{jk} > 0\} \\ q'_{hk_a} = q_{hk_a} \prod_{i=1}^m (1 - p_{ij})^{d_{ij}} \end{cases}, \text{ for } i = 1, 2, \dots, m; j = 1, 2, \dots, n, \quad (26)$$

where e_{ij} represents the marginal return expectation of the value of asset k_a generated by weapon i attacking target j , and asset k_a is the attack intention of target j ; q'_{hk} denotes the equivalent kill condition probability of target h against asset k .

Proof of Theorem 1. Under the current decision $\mathbf{D}$ and situation $\{\mathbf{V}, \mathbf{P}, \mathbf{Q}, \mathbf{C}, \mathbf{A}, \mathbf{T}, \mathbf{W}\}$, the target–asset equivalent conditional killing probability is $\mathbf{Q}'$, and the value expectation of each asset is $\mathbf{V}' = [v'_k]_{1 \times l}$:

$$v'_k = v_k \prod_{j=1}^n (1 - q'_{jk}), \text{ for } k = 1, 2, \dots, l. \quad (27)$$

Take target j_t with attack intention to asset k_a , i.e., $q_{j_t k_a} > 0$, for example. If weapon i_w is available and assign $i_w - j_t$, then the equivalent conditional killing probability of target j_t to asset k_a under $i_w - j_t - k_a$ triad is

$$q''_{j_t k_a} = q'_{j_t k_a} (1 - p_{i_w j_t}). \quad (28)$$

Then the value expectation of asset k_a is

$$\begin{aligned} v''_{k_a} &= v_k (1 - q''_{j_t k_a}) \prod_{j \neq j_t}^n (1 - q'_{j k_a}) \\ &= v'_{k_a} \frac{1 - q'_{j_w k_a} (1 - p_{i_w j_t})}{1 - q'_{j_w k_a}}. \end{aligned} \quad (29)$$

Difference Formulas (28) and (29) shows that the transfer function from weapon i_w against target j_t to the marginal return expectation of asset k_a is

$$\begin{aligned} e_{i_w j_t} &= v''_{k_a} - v'_{k_a} \\ &= v'_{k_a} \frac{1 - q'_{j_w k_a} (1 - p_{i_w j_t})}{1 - q'_{j_w k_a}} - v'_{k_a} \\ &= v_{k_a} \prod_{j=1}^n (1 - q'_{j k_a}) \frac{p_{i_w j_t} q_{j_w k_a}}{1 - q'_{j_w k_a}}. \end{aligned} \quad (30)$$

Substitute the equivalent conditional kill probability d of Formula (23) into the above formula. Without loss of generality, the weapon–target marginal return expectation matrix $\mathbf{E} = [e_{ij}]_{m \times n}$ is

$$\begin{cases} e_{ij} = v_{k_a} \prod_{h=1}^n (1 - q'_{h k_a}) \frac{p_{ij} q_{j k_a}}{1 - q'_{j k_a}} \\ k_a = \{k | q_{jk} > 0\} \\ q'_{h k_a} = q_{h k_a} \prod_{i=1}^m (1 - p_{ij})^{d_{ij}} \end{cases}, \text{ for } i = 1, 2, \dots, m; j = 1, 2, \dots, n \quad (31)$$

□

According to the above proof, it can be seen that the matrix $\mathbf{E}$ has been constrained, and $e_{ij} = 0$ when the following three situations occur: (1) weapon i is unavailable; (2) target j has been destroyed; (3) weapon i is available, and target j is not damaged, but the $i - j$ pair does not meet the attackable condition, $c_{ij} = 0$.

The information flow diagram of the calculation of the weapon–target marginal return expectation is shown in Figure 4.

Figure 4. Information flow of the calculation of weapon–target marginal return expectation.

4.2. Heuristic Information Design

Under the reverse hierarchical solution strategy proposed above, the A-DWTA model has been transformed into a weapon selection problem similar to the SWTA model. In the WTA problem, the heuristic information of weapon selection is usually based on the construction of combat efficiency maximized from the current perspective [10,29,36]. First, calculate the comprehensive kill efficiency matrix $\mathbf{G} = [g_{ij}]_{m \times n}$ based on the target value $\mathbf{V}$ and the kill probability matrix $\mathbf{G} = [g_{ij}]_{m \times n}$.

$$g_{ij} = v_j p_{ij}, \text{ for } i = 1, 2, \dots, m; j = 1, 2, \dots, n. \quad (32)$$

According to different decision-making perspectives, the heuristic information based on killing effectiveness can be roughly divided into weapon-priority heuristic information and target-priority heuristic information. Weapon/target-priority heuristic information is widely used in large-scale WTA algorithms, which can improve the real-time performance and quickly obtain feasible solutions. However, the following analysis shows that the heuristic information based on the local perspective can easily make the search algorithm fall into the local optimal solution. Taking the 2*2 scale as an example, in Figure 5a, according to the decision process based on weapon-priority heuristic information, the return expectation of weapon 1 against target 1 is $0.8 \times 0.6 = 0.48$, and the return expectation of weapon 1 against target 2 is $0.6 \times 0.4 = 0.24$. The principle of maximizing return will assign weapon 1 to target 1, and the value expectation of target 1 is updated to $0.6 \times (1 - 0.8) = 0.12$. Continue to execute the allocation of weapon 2. Return the expectation of weapon 2 attacking target 1 is $0.2 \times 0.12 = 0.024$, and the return expectation of weapon 2 attacking target 2 is $0.4 \times 0.4 = 0.16$. Therefore, the allocation scheme of $(w_1 - t_1, w_2 - t_2)$ will be generated. Similarly, the same local optimal scheme can be obtained by using the target-priority heuristic information. However, the optimal scheme can be known $(w_1 - t_2, w_2 - t_1)$ through enumeration. In Figure 5b, the global optimal allocation scheme can be obtained by using the weapon/target-priority heuristic information.

Figure 5. Defects based on weapon/target priority heuristic information.

In Figure 5a, the weapon/target priority heuristic information makes the decision plan fall into local optimal, while in Figure 5b, the weapons/target priority heuristic information leads to the global optimal plan. By analyzing the internal reasons and mechanisms, the following inference is presented: heuristic information based on statistical marginal return can better overcome local optimal than heuristic information that maximizes marginal return from a single perspective.

The main reason why the maximum marginal return heuristic information easily falls into a locally optimal solution is that, in the case of certain coupling of weapon-target killing efficiency, the weapon-target assignment from the maximum marginal return is terrible under the global

perspective. For example, if the killing effect of weapon w_i against other targets is much smaller than that of the target t_j , the maximum marginal benefit heuristic information will execute the decision $w_i - t_j$, so that the threat weight of target j will decay rapidly. Then, the marginal benefit of the remaining weapon $W \setminus \{w_i\}$ attacking target t_j is minimal and is usually allocated to the other targets $T \setminus \{t_j\}$. If weapon w_i has a better killing effect on the other targets $T \setminus \{t_j\}$, and the remaining weapons $W \setminus \{w_i\}$ have a better killing effect on target t_j , the decision has fallen into a terrible local optimal solution.

In order to overcome the above coupling mechanism, this paper constructs heuristic information based on statistical marginal return: if a weapon has high killing efficiency against a few targets and low killing efficiency against the other targets, it is considered to have high execution priority and should give priority to the target. Otherwise, the target that the weapon can kill effectively may enter the attack position of other weapons, so that the marginal return of the target being pursued will be greatly reduced, and the weapon will be forced to be allocated to other targets with a lower return. When a weapon attacks each target with a small difference in its killing efficiency, it can be considered more applicable, that is, it has the characteristics of “jack of all trades” and low execution priority, and is assigned after the higher priority weapon is assigned. Analyzing Figure 5 shows that if the above ideas are used, the local optimal can be effectively avoided, and the theoretical optimal solution can be obtained.

The operation process of weapon selection based on statistical marginal return is as follows: assuming that the priority defense asset t_p and the priority attack target t_p are generated by Equation (25), and the marginal return matrix $E = [e_{ij}]_{m \times n}$ of “weapon–target–priority asset” is generated by Equation (31). The heuristic information $G = \{g_1, g_2, \dots, g_m\}$ of “weapon set–priority attack target–priority defense asset” is evaluated by E . A weapon is selected to target t_p by G . The statistical marginal return heuristic information should be constructed according to the following principles: (1) It can reflect the absolute marginal return of priority assets under the allocation of “weapon–priority attack target”; (2) It can reflect the relative return of the asset set when the weapon abandons the priority target and attacks another targets.

Taking $i = 1, 2, \dots, m$ as an example, heuristic information g_i can be disassembled into information pair (g_i^1, g_i^2) according to the above principles. g_i^1 represents the absolute marginal return of the priority asset a_k when the weapon w_i is allocated to the priority target t_j , namely e_{it_p} ; g_i^2 represents the relative return of asset set A when weapon w_i attacks the priority target t_p compared with another target. The relative return should first reflect the absolute priority of “ $w_i - t_p$ ” in the “ $w_i -$ target set T ” sequence, and secondly, reflect the dispersion degree of the distance between “ $w_i - t_p$ ” and “ $w_i -$ low-return target.” Therefore, from the perspective of weapon w_i , we define the target set whose killing efficiency is higher than $w_i - t_j$ as the high-return target set $T_p = \{j | e_{ij} > e_{it_p}\}$, and the target set lower than $w_i - t_j$ as the low-return target set $T_s = \{j | 0 < e_{ij} < e_{it_p}\}$, where the available weapon condition is not satisfied, that is, the weapon of $e_{ij} = 0$ do not enter this set division. So the relative return is designed as

$$e_2(i, t_p) = |T_p| + \frac{1}{|T_s|} \sqrt{\sum_{j \in T_s} (e_{it_p} - e_{ij})^2}, i = 1, 2, \dots, m \quad (33)$$

where the first term represents the absolute priority of $i - t_p$ in the $i - T$ sequence; the second term represents the dispersion degree of $i - t_p$ from the $i - T_s$ sequence, reflecting the relative priority of $i - t_p$. Since $|T_p| \in \mathbb{Z}$ and $\frac{1}{|T_s|} \sqrt{\sum_{j \in T_s} (w_{it_p} - p_{ij})^2} \in [0, 1)$, the heuristic information e_2 firstly depends

on the absolute priority, and then depends on the relative priority if the absolute priority is the same. The heuristic information pair constructed in this paper is

$$\begin{cases} (g_1, g_2) = \left(e_{it_p} |T_p| + \frac{1}{|T_s|} \sqrt{\sum_{j \in T_s} (e_{it_p} - e_{ij})^2} \right) \\ T_p = \{j | e_{ij} > e_{it_p}\} \\ T_s = \{j | 0 < e_{ij} < e_{it_p}\} \end{cases} \quad (34)$$

According to Equation (34), the information set $\mathbf{G} = \{g_1, g_2, \dots, g_m\}$ composed of each weapon heuristic information pair $g_i(\mathbf{E}, t_p) = (g_i^1, g_i^2)$ can be obtained, which is used as an indicator of each weapon's execution priority.

In summary, the A-DWTA decision method proposed in this paper is as follows: The current solution set of solution individuals is $D^r = \{\mathbf{D}_1, \mathbf{D}_2, \dots, \mathbf{D}_{u_m}\}$, where u_m is the upper limit of the solution set, and r is the current weapon consumption. Extract a single solution $\mathbf{D}_i$, solve the priority defense asset a_p and priority attack target t_p , calculate the statistical marginal return information set $G = \{(g_{11}, g_{12}), (g_{21}, g_{22}), \dots, (g_{d1}, g_{d2})\}$ of each weapon under " $a_p - t_p$," select the non-dominated information (g_{i1}, g_{i2}) from G to form a new information set G' . Set the maximum number of differentiation for a single solution as d . If the number of non-dominated information pairs is greater than the differentiation number, that is, $|G'| > d$, d information pairs corresponding to d usable weapons are selected in descending order of g_1 , and $\mathbf{D}_i$ is differentiated into K feasible solutions; otherwise, $\mathbf{D}_i$ is directly differentiated into $|G'|$ feasible solutions. Move the differentiated solutions to the differentiation set D' . Repeat the above process until all individuals in D complete the differentiation operation. Calculate each solution's fitness in D , and select u_m individuals with optimal fitness as the solution set D^{t+1} for the next iteration. Repeat the above process until the algorithm termination condition is met. The flowchart of HA-SMR is shown in Figure 6.

Figure 6. Information flow of the weapon selection mechanism by the proposed heuristic information.

In HA-SMR, the designed differentiation strategy also reflects the diversity of the solution set, which can effectively avoid local optimal solutions. The reason for the generation of differentiated individuals in descending order g_1 is that considering the combat objectives of asset defense, the absolute return of killing targets at the current stage should take priority over the relative return.

4.3. Constraint Handling

In WTA problems, model constraints are usually used as algorithm termination conditions. In the model constraints (22), the constraint of marginal return expectation of a weapon is proposed. This constraint not only optimizes the decision efficiency-cost ratio but also controls the algorithm flow adaptively. As in step 7 of Algorithm 1, the weapon–target feasibility $C(s)$ is updated as

$$c_{ij}(s) := \begin{cases} 1, & \text{if } u_{k_a}^a(\mathbf{D}(s)) \geq \rho \\ c_{ij}(s), & \text{otherwise} \end{cases}, \quad (35)$$

where asset k_a is the attack intention of target j as in Formula (31); ρ is the weapon consumption threshold of Formula (20). Formula (35) shows that the assignment $w_i - t_j$ will reduce the cost–effectiveness ratio, and w_i is allocated to other unmet threshold targets or retained for the next stage.

In the constraint conditions (22) of the A-DWTA model, the first three mathematical constraints have been satisfied by the coding method, and the fourth model constraint is also satisfied by the above termination condition. The dynamic constraints that need to be dealt with in the solving process are weapon–target feasibility constraints and target–resource feasibility constraints. According to the information flow of the algorithm, the observable variables connecting weapon, target, and resource are the weapon–target kill probability matrix $\mathbf{P}$, the weapon–target attack feasibility matrix $\mathbf{C}$, and the target–resource damage matrix $\mathbf{Q}$. The algorithm employs the $\{\mathbf{W}, \mathbf{T}, \mathbf{A}, \mathbf{C}\}$ to update $\mathbf{P}$ and $\mathbf{Q}$, that is, set the component corresponding to the infeasible variable to 0, and there is no assignment return. For example, if $p_{ij} = 0$ indicates the following constraints: (1) weapon i is not available; (2) target j has been killed; (3) weapon i is available, and target j survives, but the attack condition of $w_i - t_j$ is not satisfied, namely $c_{ij} = 0$. Similarly, if $q_{jk} = 0$ indicates: (1) target j has been killed; (2) asset k has been destroyed; (3) both target j and asset k are in a state of survival, but the attack intention of target j is not asset k . Through the above processing, the algorithm has been de-constrained, without additional constraint processing.

5. Experimental Studies

This section verifies the superiority of the A-DWTA/RH model and the HA-SMR proposed in this paper by setting experimental cases and comparison algorithms.

5.1. Operational Scenario and Performance Metrics

In the experiment section, the target–asset destroy intention, weapon–target attack condition is initialized and evolved by the following methods.

- (1) Target–asset destroy intention. Assuming that the defender is unknown to the targets' attack strategy, and the threat assessment system identifies the target intention. The initialization method of target–asset destroy matrix is: (i) when targets are no fewer than assets ($n \leq l$), the target intention is generated randomly on the premise that each asset is assigned at least one target. (ii) When there are more targets than assets ($n > l$), the target intention is generated randomly under the premise that each target aims to different assets. The target–asset destroy probability is initialized as

$$q_{jk}(s) = q_l + (q_h - q_l) \cdot \text{rand}, \text{ for } j = 1, 2, \dots, m; k = 1, 2, \dots, n; s = 1, 2, \dots, s_m, \quad (36)$$

where q_h and q_l reflect the upper and lower bounds of target–asset vulnerability.

(2) Weapon–target attack condition. The weapon–target kill probability is initialized by

$$p_{ij}(1) = p_l + (p_h - p_l) \cdot \text{rand}, \text{ for } i = 1, 2, \dots, m; j = 1, 2, \dots, n, \quad (37)$$

where p_h and p_l denotes the upper and lower bounds of weapon–target kill probability. In the evolution of subsequent stages, due to the operational state's continuity, the weapon–target kill matrix of the current stage should be related to the weapon–target kill matrix, weapon state and target state of the previous stage. Let the transfer state of weapon–target kill probability follow

$$\Pr(p_{ij}(s) = x) = \begin{cases} 0, & \text{if } w_i(s-1)t_j(s-1) = 0 \\ \frac{1}{\sqrt{2\pi}\delta_p} e^{-\frac{(x-p_{ij}(s))^2}{2\delta_p^2}}, & \text{otherwise} \end{cases}, x \in \{0, 1\}; i = 1, 2, \dots, m; j = 1, 2, \dots, n; s = 2, 3, \dots, s_m. \quad (38)$$

Weapon–target feasibility matrix **C** of each stage refer to [37]

$$c_{jk}(s) = \frac{1}{2} \left[\text{sign} \left(\text{rand} - \left(c_l + \frac{(c_u - c_l)(s-1)}{s' - 1} \right) + 1 \right) \right], \quad (39)$$

for $s = 1, 2, \dots, s_m; j = 1, 2, \dots, n; k = 1, 2, \dots, l$

where $c_{ij}(s) = 1$ means that the launch condition of $i - j$ pair is met at stage s , otherwise $c_{ij}(s) = 0$; $\text{ratio}(s)$ is the probability threshold that weapon i can not launch for target j at stage s . The threshold decreases with the increase of the stage s , and the generated probability of zero elements in **c** increases, which meet the evolution trend of attack uncertainty in the DWTA problem. c_h and c_l are the upper and lower bounds of the function $\text{ratio}(s)$, and s' is set a constant not less than the maximum stage number s_m . $\text{sign}(\cdot)$ is equal to 1 if its argument is positive and -1 otherwise.

From the application background of A-DWTA, the normalized asset value (NAV) is proposed as the dynamic metric. The weapon consumption and the number of stages are taken as the mission completion metric. Algorithm complexity and compute time are used as the real-time metric. The above metrics cover the operational requirements of decision-makers.

5.2. Experiments on Comparison Algorithms

The research on the A-DWTA problem is less than the research on the other WTA problems, and the models and solving algorithms proposed by each study are different, so it is not easy to make comparison under the same experimental framework. We select research with a similar model framework as the comparison group and adopt the verified algorithms as comparison algorithms: Hybrid Genetic Algorithm (HGA); Hybrid Ant-Colony Optimization (HACO); Memetic Algorithm based on Greedy Local Search (MA-GLS); Rule-based Heuristic Algorithm (RHA). Except, RHA has no parameter setting, the parameters of the model and comparison algorithms are as follows:

- A-DWTA/RH model: the weight of the absolute return expectation at the current stage $\lambda_1 = 0.6$; the weight of the return evaluation of remaining weapons on prediction situation $\lambda_2 = 1 - \lambda_1 = 0.4$;
- HA-SMR: The capacity of solution set $u_m = 50$; the differentiation number $d = 4$; the threshold of weapon consumption $\rho = 0.85$;
- HGA: the population size $\mu = 100$; the number of elite individuals $\lambda = 100$; the crossover probability $p_c = 0.8$; the mutation probability $p_m = 0.2$;
- HACO: the population size $\mu = 200$; the train importance factor $\alpha = 1$; the visibility importance factor $\beta = 2$; the pheromone evaporation rate $\rho = 0.5$;
- MA-GLS: the population size $p_s = 200$; the crossover/mutation number of each generation is set to 10.

To complete the defense mission, the number of weapons should be no less than the number of targets. In order to compare the algorithms in different problem scales, the following method is used to generate experimental cases: randomly generate the number of targets n in different scale

intervals $[n_l, n_u]$; randomly generate the number of weapons m in the interval $[1.5n, 2n]$; and randomly generate the number of assets l in the interval $[0.5n, 1.5n]$, as shown in the formula. In the interval $[1, 100]$, the target number n is randomly generated with an interval of 10, and nine experimental cases generated are shown in Table 2.

Table 2. The case scales of the algorithm comparison experiment.

Scale	No.1	No.2	No.3	No.4	No.5	No.6	No.7	No.8	No.9
m	28	35	69	81	105	114	137	136	159
n	18	21	36	47	58	65	72	82	97
l	13	24	48	62	80	55	91	77	58

Let each comparison algorithm independently solve the above cases 30 times. The statistics of the experimental results are shown in Table 3, and the distribution of NAV and computational time is illustrated as box plots in Figure 7. A box plot is used to illustrate the distribution of metrics. The notches represent a robust estimate of the uncertainty about the medians for box-to-box comparison, and the symbol + denotes outliers.

Analyzing Table 3 and Figure 7, the mean value of NAV of HA-SMR is only 0.002~0.005 lower than HGA and HACO in case 1 and 2, slightly lower than HACO in case 4–7, and better than other algorithms in case 8 and 9. The reason is that the SMR algorithm has no apparent advantages in convergence at a relatively small scale. However, owing to the problem scale and population size, the convergences of HGA, HACO, and MA-GLS are worse than HA-SMR. RHA can be considered as a particular case of HA-SMR in a single search direction, so the NAV of RHA is lower than HA-SMR. HA-SMR and RHA are deterministic algorithms, and the mean square error of NSVA is 0 theoretically. In the other swarming intelligent algorithms, the mean square errors of NAV are between 0.001 and 0.005 and increase with the increase of problem size. In terms of real-time performance, HA-SMR outperforms than the other algorithms, followed by RHA. The computational time of HA-SMR and RHA is in the same order of magnitude, while the other algorithms' compute time is significantly higher than one order of magnitude. Similarly, the theoretical mean square error of two deterministic algorithms' compute time is 0, and the mean square error of the other algorithms' computational time increases with the problem scale, and is between 0 and 5 s. In conclusion, within the scale of W150T100A100, HA-SMR has better robustness in convergence, and balances the optimal and real-time effectively.

Besides NAV and computational time, minimizing the number of stages is also an essential operational requirement and a motivation of this paper. To verify the design of the SMR algorithm for "interception as soon as possible" combat requirements, Figure 8 shows the distribution of the actual number of stages obtained by each algorithm solving the same case 30 times independently.

According to Figure 8, the number of stages of the decision plan obtained by HA-SMR solving the A-DWTA/RH model is less than other algorithms. When the problem scale is small (case 1 to 4), the number of stages obtained by HA-SMR and RHA is two or three. Let the number of decision stages of the HA-SMR algorithm be s_m , then the number of stages obtained by HGA, HACO, and MA-GLS is distributed in the interval $[s_{\min}, 4]$. The proportion of four stages solved by HACO is significantly less than that of HGA and MA-GLS. In conclusion, the scheme obtained by the HA-SMR algorithm can complete the target killing task in as few stages as possible. To further illustrate the advantages of HA-SMR through the weapon–target–asset states at each stage, Figure 9 shows the mean values of NMVA, remaining weapons, and surviving targets from the observe phase of each stage.

(a) Boxplots of NAV values obtained by the comparison algorithms solving the nine cases

(b) Boxplots of compute time obtained by the comparison algorithms solving the nine cases

Figure 7. Boxplots of normalized asset value (NAV) and compute time obtained by the comparison algorithms solving the nine instances over 30 independent runs.

Table 3. Statistics of the results obtained by the comparison algorithms solving nine case over 30 independent runs.

Case	Mean Metric	Comparison Algorithms				
		HA-SMR	RHA	HGA	HACO	MA-GLS
No.1	NAV	0.9251 ± 0	0.9013 ± 0	0.9301 ± 0.0014	0.9295 ± 0.0017	0.9147 ± 0.0019
	CT	0.07 ± 0	0.03 ± 0	11.84 ± 0.37	15.83 ± 0.52	14.93 ± 0.22
No.2	NAV	0.9374 ± 0	0.8963 ± 0	0.9398 ± 0.0017	0.9416 ± 0.0013	0.9262 ± 0.0016
	CT	0.15 ± 0	0.09 ± 0	19.39 ± 0.54	23.85 ± 0.78	20.39 ± 0.46
No.3	NAV	0.8370 ± 0	0.8359 ± 0	0.8352 ± 0.0020	0.8364 ± 0.0019	0.8284 ± 0.0015
	CT	0.50 ± 0	0.28 ± 0	37.98 ± 0.55	49.56 ± 0.61	43.98 ± 0.50
No.4	NAV	0.9408 ± 0	0.9268 ± 0	0.9385 ± 0.0018	0.9417 ± 0.0021	0.9201 ± 0.0027
	CT	0.96 ± 0	0.52 ± 0	59.46 ± 0.54	61.93 ± 0.72	62.87 ± 0.79
No.5	NAV	0.9010 ± 0	0.8815 ± 0	0.8905 ± 0.0028	0.9118 ± 0.0035	0.9007 ± 0.0033
	CT	1.87 ± 0	1.16 ± 0	113.06 ± 1.49	126.65 ± 1.81	115.48 ± 2.36
No.6	NAV	0.9090 ± 0	0.8915 ± 0	0.8972 ± 0.0038	0.9107 ± 0.0030	0.8906 ± 0.0031
	CT	4.27 ± 0	3.22 ± 0	117.79 ± 1.96	124.66 ± 2.04	121.90 ± 1.72
No.7	NAV	0.8708 ± 0	0.8611 ± 0	0.8635 ± 0.0036	0.8815 ± 0.0034	0.8783 ± 0.0037
	CT	8.91 ± 0	5.52 ± 0	158.97 ± 2.96	165.49 ± 3.14	159.42 ± 2.92
No.8	NAV	0.7866 ± 0	0.7619 ± 0	0.7652 ± 0.0039	0.7710 ± 0.0041	0.7553 ± 0.0045
	CT	11.29 ± 0	8.11 ± 0	176.10 ± 2.58	189.55 ± 3.96	178.96 ± 3.16
No.8	NAV	8718 ± 0	0.8340 ± 0	0.8379 ± 0.0043	0.8651 ± 0.0047	0.8492 ± 0.0048
	CT	14.23 ± 0	11.57 ± 0	205.96 ± 3.81	217.91 ± 4.13	209.91 ± 3.63

Figure 8. Distribution of the number of stages obtained by the comparison algorithms solving case 1~9 of 30 runs independently.

Figure 9. Mean value of NAV, remaining weapons and surviving targets observed in each stage, obtained by the comparison algorithms solving case 1~9 over 30 independent runs.

In Figure 9, HA-SMR and RHA consume more weapons at the 1st stage than the other algorithms, making the observed number of threat targets and remaining weapons drop rapidly at the 2nd stage. However, due to RHA's greedy strategy, if the targets survive after the 1st stage, the remaining weapons in the 2nd or 3rd are less effective in killing the target. The statistical return strategy of HA-SMR maintains the remaining weapons still kill the surviving targets effectively. Due to the full-stage prediction mechanism, the other algorithms' decision strategy is more conservative than HA-SMR, and the remaining weapons are less than HA-SMR when the operational mission is completed. It can be seen that the HA-SMR solving A-DWTA/RH has a better efficiency–cost ratio.

5.3. Parameter Sensitivity

In this section, the effects of specific parameters in the A-DWTA/RH model and HA-SMR on operational missions are analyzed experimentally.

5.3.1. Model Weight λ

In the A-DWTA/RH model, parameters are used to weigh the weapon consumption's killing efficiency at the current stage against the counter-effectiveness of the remaining weapon. To analyze the influence on operational mission, $\lambda = [0.1 : 0.1 : 0.9]$ is set up for simulation experiments. The problem scale of the example was W100T50A50, and the other variables are generated as the previous experiment. Under the different values of λ , the following three operational mission termination conditions are set:

1. All assets are destroyed ($l(s_m) = 0$), defense mission fails, let $\Gamma = 0$.
2. Assets are not all damaged and targets are all killed ($l(s_m) > 0$ and $n(s_m) = 0$), the defense task is completed, let $\Gamma = 1$.
3. Assets have not been destroyed and targets have not been killed, but weapons have been consumed ($l(s_m) > 0$, $n(s_m) > 0$ and $m(s_m) = 0$). It is predicted that the defense mission will fail, and let $\Gamma = -1$.

The experimental results of each case under the above parameter settings are shown in Table 4.

Table 4. Experiment results of W100T50A50 under different values of λ .

Parameter λ	Stage s_m	Compute Time	NAV	Plan Fitness	Termination State
0.1	4	3.3302	0.3461	0.5391	Γ 1:W54T0A24
0.2	3	1.9247	0.5417	0.6369	Γ 1:W47T0A34
0.3	2	1.7992	0.8096	0.6713	Γ 1:W47T0A43
0.4	3	1.8254	0.7599	0.7235	Γ 1:W36T0A41
0.5	2	1.4043	0.8882	0.7640	Γ 1:W31T0A46
0.6	2	1.6805	0.9328	0.8012	Γ 1:W21T0A47
0.7	2	1.2787	0.9832	0.8604	Γ 1:W15T0A49
0.8	2	1.3525	0.9721	0.9057	Γ 1:W7T0A49
0.9	2	1.2706	0.9615	0.9448	Γ 1:W4T0A49

To further analyze the dynamic performance of HA-SMR solving the A-DWTA/RH under different λ value, Figure 10 gives the plot of the number of surviving assets, the number of surviving assets targets, the number of remaining weapons, NAV, and plan fitness observed at each stage.

Figure 10. The dynamic performance of the number of assets/targets/weapons, NAV and plan fitness under λ from 0.1 to 0.9.

In Table 4 and Figure 10, with the parameter increases from 0.1 to 0.9, the number of decision stages decreases from 4 to 2, and the compute time decreases from 3.3302s to 1.2706s. The reason is that A-DWTA/RH's objective F is composed of the current decision's kill effect f_1 and the remaining weapons' counter effect f_2 after execution. λ represents the weight of the former. When λ is small, the decision plan will be guided to limit the weapon consumption of the current stage and retain the countermeasures in subsequent stages, which leads to an increased number of stages and computational time. The maximum value of NAV is 0.9832 at $\lambda = 0.7$, and the fitness metric increases with the increase of λ . It can be seen that the objective F of A-DWTA/RH is more sensitive to f_1 than f_2 . Even when $\lambda = 0.9$, there is no situation of excessive weapon consumption at the current stage and no counter weapons at the later stage ($\Gamma = -1$). In the mission termination state, the defender can complete the mission ($\Gamma = 1$) under the different λ value. However, as the λ value increases, the proportion of the current stage's decision return increases, and the weapon consumption of the current stage also

increase. So the number of surviving assets increases when the operational mission is completed. It is worth noting that when λ is 0.7, 0.8, and 0.9, the number of surplus weapons keeps decreasing, 15, 7 and 4 respectively. However, the number of decision stages and surviving targets remains unchanged, 2 and 49, respectively. It can be seen that if λ is higher than 0.7, weapon resource is easily wasted. In conclusion, the model parameter λ has the effect of balancing the “radical–conservative” degree of the decision plan, and the reasonable value is between 0.6 and 0.8.

5.3.2. Problem Scale $\{M, N, L\}$

In case 9 of Table 4, the surplus number of weapons is only 4. Hence the relationship between the scale of weapon–target–asset directly affects the completion of the combat mission. In order to provide reference and basis for the configuration of completing the A-DWTA/RH mission, the dynamic relationship among the scale of weapon–target–asset is experimentally analyzed. In this experiment, the target number is fixed to 50. To ensure the possibility of completing the mission, the weapons should be more than targets. The weapon number is set to $m = [60 : 10 : 100]$ in turn, and the asset number is set to $l = [20 : 10 : 80]$. The model parameter is set to 0.8. The remaining variables are consistent with the previous experiment. As the scales and parameters of each case in this simulation are different, they can only be analyzed based on the win–loss indicator’s statistical data. The distribution of defense mission completion indicator is shown in Figure 11.

Illustrated by Figure 11, all assets are destroyed ($\Gamma = 0$) only when the asset number is 5, and the weapon number is 55, 60, and 70. The reason is that the asset number is too small, and the weapon number is only 10% to 20% higher than the incoming target number, which cannot guarantee the survival probability of assets in a few stages. When the weapon number is less than 70, and the assets number rises to the target number (50), the situation of $\Gamma = -1$ is basically obtained, that is, the weapons are fully used. The reason is that the weapon number is close to the target number, making it impossible to complete the defense mission. With the increase of asset number ($l > 50$), the target cannot destroy all assets in a few stages under the attack of weapons, so the weapons are consumed first. Finally, when the number of weapons is no less than 75, the probability of successful defense ($\Gamma = 1$) increases significantly. When the number of weapons is no less than 95, the proportion of $\Gamma = 1$ is 100%. In summary, regardless of the state of assets, the weapon number of 95% higher than the target number can ensure the completion of the defense mission.

5.3.3. Set Capacity U_m and Differentiation Number D

In HA-SMR, the solution set’s capacity u_m and the differentiation number d of a single solution are the essential parameters to control the search space. The following cases are set to verify the influence of u_m and d on HA-SMR: In the initial stage, weapon number m is set to 80; the target number n is set to 50; the asset number l is set to 50; the weapon–target kill probability $\mathbf{P}$ and the target–asset damage matrix $\mathbf{Q}$ are generated by a formula, and the weapon consumption threshold is set to 0.8. Let the capacity of solution set be $u_m = [10:10:100]$, and the differentiation number be $d = [1:1:10]$. Figure 12 gives the fitting surface of four metrics distribution under different solution set capacity u_m and differentiation number d .

Figure 12a shows that the capacity u_m and differentiation number d have no impact on the completion of the defense mission. The deterministic convergence mechanism of HA-SMR can give an effective solution under different values of u_m and d . In Figure 12b, the number of decision stages s_m decreases with the increase of the solution set’s capacity u_m , and increases with the increase of the differentiation number d . According to Figure 12c, increasing the values of u_m and d can improve NAV metric. In Figure 12d, when the value of u_m rises from 10 to 100 and the value of d rises from 1 to 10, the compute time increases from 0.0064 s to 20.1675 s. In conclusion, on the premise of satisfying the real-time, the capacity u_m of the solution set should be increased, and the number differentiation number d can be set to 5. Continue to increase the value of d do not significantly improve NAV, and burden the number of decision stages and compute time.

Figure 11. The distribution of mission completion indicator under the scale of $n = 50$, $55 \leq m \leq 100$, $5 \leq l \leq 100$.

Figure 12. Fitting surface of four performance metric distribution of different set capacity u_m and differentiation number d .

5.3.4. Threshold Parameter ρ

To verify the threshold ρ of weapon consumption in HA-SMR, the following cases are set: in the initial stage, the weapon number m is set to 100; the target number is set to 20; the asset number is set to 20; the weapon-target kill probability $\mathbf{P}$ and the target-asset destroy matrix $\mathbf{Q}$ are generated by a formula, and the solution set's capacity u_m and differentiation number d are set to 100 and 5. Let the threshold be $\rho = [0.1:0.1:0.1]$, and adopt HA-SMR to solve A-DWTA/RH model. Record the number of weapons consumed in the first stage under different ρ values, and the plot of NAV and NAV return with the increase of weapon consumption, as shown in Figure 13.

Figure 13. Plots of weapon consumption, NAV and NAV return under different values of ρ .

To further analyze HA-SMR dynamic performance under different ρ values, Figure 14 shows the mission completion state, the number of decision stages, the normalized asset value, the number of surplus weapons, and the number of surviving targets when the threshold ρ is 0.1 to 0.9. In Figure 14, the defense mission cannot be completed ($\Gamma = 0$) when the value of ρ is 0.1 and 0.2. The assets are all destroyed by targets. The reason is that the ρ value is too small to limit the weapon consumption at each stage, making the number of stages ($u_m = 5, 19$) and the number of surplus weapons large. When the value of ρ rises from 0.3 to 0.9, the state of defense mission is $\Gamma = 1$; the number of decision stages drops from 13 to 1; the number of surviving targets is 0, and the number of surplus weapons decreases from 73 to 2. The above metrics are relatively stable at ρ of 0.3 to 0.7. Like the above experimental conclusion, ρ of 0.7~0.9 has better cost-effectiveness, which can minimize the number of decision stages, maximize the asset value and avoid waste of weapon resources on the premise of ensuring the completion of defense tasks.

Figure 14. The mission state, the number of decision stages, NAV, the number of surviving targets, and the number of surplus weapons under different ρ values (0.1~0.9).

6. Conclusions and Future Work

The major A-DWTA researches adopt the global return model based on multi-stage prediction, which can not fully reflect the operational requirement at the current stage and burden computing resource. From the above motivation, the A-DWTA problem is first analyzed by OODA theory. Then the A-DWTA/RH decision model is established, and the solving algorithm HA-SMR is designed. Experimental results show that the obtained decision plan can complete the defense mission in fewer decision stages than several comparison algorithms. The “radical-conservative” degree of the obtained plan can be adjusted by the model parameter λ and the algorithm parameter ρ . The model weight λ is suggested as 0.6~0.8, and the damage threshold ρ is suggested as 0.7~0.9.

The proposed heuristic information based on statistical marginal return is not only limited to the A-DWTA model, but also has a certain universality for weapon selection of various WTA problems. In this paper, the heuristic information is designed by the qualitative analysis of domain knowledge, and lacks rigorous theoretical verification. Aiming at the WTA of specific scenarios, design the corresponding weapon selection information by quantitative derivation, which is the future challenge.

Author Contributions: Conceptualization, K.Z.; methodology, K.Z. and Z.Y.; data curation, Y.Z.; Software, K.Z.; formal analysis, W.K. and Z.Y.; project administration, D.Z.; writing—original draft, K.Z.; writing—review and editing, Y.Z.; funding acquisition, D.Z. All authors have read and agreed to the published version of the manuscript.

Funding: This work was supported in part by the National Natural Science Foundation of China under Grant 61603299 and Grant 61612385, and in part by the Fundamental Research Funds for the Central Universities under Grant 3102019ZX016.

Acknowledgments: The authors greatly appreciate the reviews valuable suggestions and the editors encouragement.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Roux, J.N.; Van Vuuren, J.H. Threat evaluation and weapon assignment decision support: A review of the state of the art. *ORiON* **2007**, *23*, 151–187. [\[CrossRef\]](#)
2. Yang, Z.; Zhou, D.; Kong, W.; Piao, H.; Zhang, K.; Zhao, Y. Nondominated Maneuver Strategy Set with Tactical Requirements for a Fighter Against Missiles in a Dogfight. *IEEE Access* **2020**, *8*, 117298–117312. [\[CrossRef\]](#)
3. Manne, A.S. A target-assignment problem. *Oper. Res.* **1958**, *6*, 346–351. [\[CrossRef\]](#)
4. Hosein, P.A.; Athans, M. Preferential Defense Strategies. Part I: The Static Case; this Issue. 1990. Available online: <https://core.ac.uk/download/pdf/4378993.pdf> (accessed on 14 September 2020).
5. Hosein, P.; Athans, M. *Preferential Defense Strategies. Part 2: The Dynamic Case*; LIDS-P-2003; MIT: Cambridge, MA, USA, 1990.
6. Cai, H.; Liu, J.; Chen, Y.; Wang, H. Survey of the research on dynamic weapon-target assignment problem. *J. Syst. Eng. Electron.* **2006**, *17*, 559–565.
7. Zhang, K.; Zhou, D.; Yang, Z.; Pan, Q.; Kong, W. Constrained Multi-Objective Weapon Target Assignment for Area Targets by Efficient Evolutionary Algorithm. *IEEE Access* **2019**, *7*, 176339–176360. [\[CrossRef\]](#)
8. Murphey, R.A. Target-based weapon target assignment problems. In *Nonlinear Assignment Problems*; Springer: Berlin, Germany, 2000; pp. 39–53.
9. Cho, D.H.; Choi, H.L. Greedy Maximization for Asset-Based Weapon-Target Assignment with Time-Dependent Rewards. *Coop. Control Multi-Agent Syst. Theory Appl.* **2017**, 115–139. Available online: <https://onlinelibrary.wiley.com/doi/10.1002/9781119266235.ch5> (accessed on 14 September 2020).
10. Xin, B.; Wang, Y.; Chen, J. An efficient marginal-return-based constructive heuristic to solve the sensor-weapon-target assignment problem. *IEEE Trans. Syst. Man Cybern. Syst.* **2018**, *49*, 2536–2547. [\[CrossRef\]](#)
11. Li, X.; Zhou, D.; Yang, Z.; Pan, Q.; Huang, J. A novel genetic algorithm for the synthetical Sensor-Weapon-Target assignment problem. *Appl. Sci.* **2019**, *9*, 3803. [\[CrossRef\]](#)
12. Kline, A.; Ahner, D.; Hill, R. The weapon-target assignment problem. *Comput. Oper. Res.* **2019**, *105*, 226–236. [\[CrossRef\]](#)
13. Hosein, P.A.; Athans, M. *Some Analytical Results for the Dynamic Weapon-Target Allocation Problem*; Technical Report; Massachusetts Inst of Tech Cambridge Lab for Information and Decision Systems: Cambridge, MA, USA, 1990.
14. Eckler, A.R.; Burr, S.A. *Mathematical Models of Target Coverage and Missile Allocation*; Technical Report; Military Operations Research Society: Alexandria, VA, USA, 1972.
15. Soland, R.M. Optimal terminal defense tactics when several sequential engagements are possible. *Oper. Res.* **1987**, *35*, 537–542. [\[CrossRef\]](#)
16. Hosein, P.; Athans, M. *The Dynamic Weapon-Target Assignment Problem*; Technical Report; 1989. Available online: <https://core.ac.uk/download/pdf/4380801.pdf> (accessed on 14 September 2020).
17. Leboucher, C.; Shin, H.S.; Siarry, P.; Chelouah, R.; Le Méneç, S.; Tsourdos, A. A two-step optimisation method for dynamic weapon target assignment problem. In *Recent Advances on Meta-Heuristics and Their Application to Real Scenarios*; InTech: Rijeka, Croatia, 2013; pp. 109–129.
18. Zhengrong, J.; Faxing, L.; Hangyu, W. Multi-stage attack weapon target allocation method based on defense area analysis. *J. Syst. Eng. Electron.* **2020**, *31*, 539–550. [\[CrossRef\]](#)
19. Jang, J.; Yoon, H.G.; Kim, J.C.; Kim, C.O. Adaptive Weapon-to-Target Assignment Model Based on the Real-Time Prediction of Hit Probability. *IEEE Access* **2019**, *7*, 72210–72220. [\[CrossRef\]](#)
20. Lloyd, S.P.; Witsenhausen, H.S. Weapons allocation is NP-complete. In Proceedings of the 1986 Summer Computer Simulation Conference, Reno, NV, USA, 28–30 July 1986; pp. 1054–1058.
21. Hocaoglu, M.F. Weapon target assignment optimization for land based multi-air defense systems: A goal programming approach. *Comput. Ind. Eng.* **2019**, *128*, 681–689. [\[CrossRef\]](#)
22. Cao, M.; Fang, W. Swarm Intelligence Algorithms for Weapon-Target Assignment in a Multilayer Defense Scenario: A Comparative Study. *Symmetry* **2020**, *12*, 824. [\[CrossRef\]](#)
23. Kline, A.G.; Ahner, D.K.; Lunday, B.J. Real-time heuristic algorithms for the static weapon target assignment problem. *J. Heuristics* **2019**, *25*, 377–397. [\[CrossRef\]](#)
24. Zhao, P.; Wang, J.; Kong, L. Decentralized Algorithms for Weapon-Target Assignment in Swarming Combat System. *Math. Probl. Eng.* **2019**, 2019, 8425403. [\[CrossRef\]](#)

25. Summers, D.S.; Robbins, M.J.; Lunday, B.J. An approximate dynamic programming approach for comparing firing policies in a networked air defense environment. *Comput. Oper. Res.* **2020**, *117*, 104890. [\[CrossRef\]](#)
26. Fu, G.; Wang, C.; Zhang, D.; Zhao, J.; Wang, H. A Multiobjective Particle Swarm Optimization Algorithm Based on Multipopulation Coevolution for Weapon-Target Assignment. *Math. Probl. Eng.* **2019**, *2019*, 1424590. [\[CrossRef\]](#)
27. Guo, D.; Liang, Z.; Jiang, P.; Dong, X.; Li, Q.; Ren, Z. Weapon-target assignment for multi-to-multi interception with grouping constraint. *IEEE Access* **2019**, *7*, 34838–34849. [\[CrossRef\]](#)
28. Chen, J.; Xin, B.; Peng, Z.; Dou, L.; Zhang, J. Evolutionary decision-makings for the dynamic weapon-target assignment problem. *Sci. China Ser. F Inf. Sci.* **2009**, *52*, 2006. [\[CrossRef\]](#)
29. Chang, T.; Kong, D.; Hao, N.; Xu, K.; Yang, G. Solving the dynamic weapon target assignment problem by an improved artificial bee colony algorithm with heuristic factor initialization. *Appl. Soft Comput.* **2018**, *70*, 845–863. [\[CrossRef\]](#)
30. Wang, Y.; Li, J.; Huang, W.; Wen, T. Dynamic weapon target assignment based on intuitionistic fuzzy entropy of discrete particle swarm. *China Commun.* **2017**, *14*, 169–179. [\[CrossRef\]](#)
31. Lai, C.M.; Wu, T.H. Simplified swarm optimization with initialization scheme for dynamic weapon–target assignment problem. *Appl. Soft Comput.* **2019**, *82*, 105542. [\[CrossRef\]](#)
32. Shin, M.K.; Lee, D.; Choi, H.L. Weapon-Target Assignment Problem with Interference Constraints using Mixed-Integer Linear Programming. *arXiv* **2019**, arXiv:1911.12567.
33. Lee, D.; Shin, M.K.; Choi, H.L. Weapon Target Assignment Problem with Interference Constraints. In Proceedings of the AIAA Scitech 2020 Forum, Orlando, FL, USA, 6–10 January 2020; p. 0388.
34. Xu, W.; Chen, C.; Ding, S.; Pardalos, P.M. A bi-objective dynamic collaborative task assignment under uncertainty using modified MOEA/D with heuristic initialization. *Exp. Syst. Appl.* **2020**, *140*, 112844. [\[CrossRef\]](#)
35. Ahuja, R.K.; Kumar, A.; Jha, K.; Orlin, J.B. *Exact and Heuristic Methods for the Weapon Target Assignment Problem*; INFORMS: Catonsville, MD, USA, 2003.
36. Lee, Z.J.; Su, S.F.; Lee, C.Y. Efficiently solving general weapon-target assignment problem by genetic algorithms with greedy eugenics. *IEEE Trans. Syst. Man Cybern. Part B Cybern.* **2003**, *33*, 113–121.
37. Xin, B.; Chen, J.; Peng, Z.; Dou, L.; Zhang, J. An efficient rule-based constructive heuristic to solve dynamic weapon-target assignment problem. *IEEE Trans. Syst. Man Cybern.-Part A Syst. Hum.* **2010**, *41*, 598–606. [\[CrossRef\]](#)
38. Leboucher, C.; Shin, H.; Le Méné, S.; Tsourdos, A.; Kotenkoff, A. Optimal weapon target assignment based on an geometric approach. *IFAC Proc. Vol.* **2013**, *46*, 341–346. [\[CrossRef\]](#)
39. Leboucher, C.; Shin, H.S.; Le Méné, S.; Tsourdos, A.; Kotenkoff, A.; Siarry, P.; Chelouah, R. Novel evolutionary game based multi-objective optimisation for dynamic weapon target assignment. *IFAC Proc. Vol.* **2014**, *47*, 3936–3941. [\[CrossRef\]](#)
40. Li, J.; Chen, J.; Xin, B.; Dou, L. Solving multi-objective multi-stage weapon target assignment problem via adaptive NSGA-II and adaptive MOEA/D: A comparison study. In Proceedings of the 2015 IEEE Congress on Evolutionary Computation (CEC), Sendai, Japan, 25–28 May 2015; pp. 3132–3139.
41. Li, J.; Chen, J.; Xin, B.; Dou, L.; Peng, Z. Solving the uncertain multi-objective multi-stage weapon target assignment problem via MOEA/D-AWA. In Proceedings of the 2016 IEEE Congress on Evolutionary Computation (CEC), Vancouver, BC, Canada, 24–29 July 2016; pp. 4934–4941.
42. Li, J.; Chen, J.; Xin, B.; Chen, L. Efficient multi-objective evolutionary algorithms for solving the multi-stage weapon target assignment problem: A comparison study. In Proceedings of the 2017 IEEE Congress on Evolutionary Computation (CEC), San Sebastian, Spain, 5–8 June 2017; pp. 435–442.
43. Li, J.; Chen, J.; Xin, B. Optimizing multi-objective uncertain multi-stage weapon target assignment problems with the risk measure CVaR. In Proceedings of the 2019 IEEE 15th International Conference on Control and Automation (ICCA), Edinburgh, UK, 16–19 July 2019; pp. 61–66.
44. Mei, Z.; Peng, Z.; Zhang, X. Optimal dynamic weapon-target assignment based on receding horizon control heuristic. In Proceedings of the 2017 13th IEEE International Conference on Control & Automation (ICCA), Ohrid, Macedonia, 3–6 July 2017; pp. 876–881.

45. Kalyanam, K.; Casbeer, D.; Pachter, M. Monotone optimal threshold feedback policy for sequential weapon target assignment. *J. Aerosp. Inf. Syst.* **2017**, *14*, 68–72. [[CrossRef](#)]
46. Duan, W.; Yuan, W.; Pan, L. Research on Algorithm for Dynamic Weapon Target Assignment Based on the Improved Markov Decision Model. In Proceedings of the 2018 37th Chinese Control Conference (CCC), Wuhan, China, 25–27 July 2018; pp. 8417–8424.

© 2020 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).