JOURNAL OF AEROSPACE COMPUTING, INFORMATION, AND COMMUNICATION
Vol. 6, January 2009

Novel Goal-Based Weapon Target Assignment Doctrine

Hosam I. Mekawey* , M. S. Abd EL-Wahab† , and M. Hashem‡
Ain Shams University, Cairo, Egypt

DOI: 10.2514/1.34374

The weapon target assignment problem is about finding the optimal allocation of weapons to threats in a way that minimizes the expected damage inflicted on the defender assets. In 1986 Lloyd and Witsenhousen demonstrated that weapon target assignment was NP-complete with no exact algorithm. Researchers in this area are trying to provide an exact solution to special cases of the problem or heuristics that attempt to supply an approximate solution using a variety of tools and techniques from nonlinear network flows to artificial neural networks and genetic algorithms. A new approach for tackling the weapon target assignment problem is proposed in this paper. Such an approach is a novel goal-based system. The proposed novel approach combines state of the art goal-based optimization approach and the Hungarian method to preserve good performance under different air defense mission configurations. An air defense mission design and analysis package is developed to provide realistic air defense missions data to the algorithm. The proposed algorithm has the best performance when compared with other weapon target assignment doctrines.

I. Introduction

THE weapon target assignment (WTA) problem is considered one of the most prominent problems in operation research and military-oriented optimization. It has large applicability in the military, especially in air defense (AD). The WTA problem is to find a proper assignment of weapons to targets with the objective of minimizing the expected damage on the defended assets. The problem has been demonstrated to be NP-complete with no exact methods for large-sized problems. Manne [1] was the first to address this problem formally. Later, Braford [2] and Day [3] investigated the modeling issues of the problem. A comprehensive literature review on the WTA problem was provided by Maltin [4], Eckler and Burr [5], and Murphey [6]. The NP-completeness of the problem has been established by Lloyd and Witsenhausen [7]. DenBroader et al. [8] and Katter [9] established an exact algorithm for solving the problem for the special case when all weapons are identical, while Chang et al. [10] and Orlin [11] found an exact solution for the special case when each threat can receive one weapon only. Several heuristics have been proposed by Castanon et al. [12] for solving the problem. Such heuristics were based on nonlinear network flow. Artificial neural network (ANN) has been used by Wacholder [13] to provide approximate solution to the problem. Metler and Preston [14] have studied a suite of algorithms for solving the WTA problem efficiently. While Grant et al. [15] was the first to try to use genetic algorithms to tackle the problem, Zne-Jung Lee [16] used genetic algorithms

---

Received 2 September 2007; revision received 27 March 2008; accepted for publication 2 April 2008.

* Teaching Associate, Scientific Computing Department, Faculty of Computer and Information Sciences, Ain Shams University, Cairo, Egypt, Hosam_mekawey@asunet.shams.edu.eg, HMekawey@Microsoft.com, Hosam_mekawey@yahoo.com

† Professor at Scientific Computing Department, Faculty of Computer and Information Sciences, Ain Shams University, Cairo, Egypt, mswahab@gmail.com

‡ Professor at Information System Department, Faculty of Computer and Information Sciences, Ain Shams University, Cairo, Egypt, mhashem100@yahoo.com

with greedy eugenics to solve the problem efficiently. Green [17] elaborated a goal programming-based approach to the WTA problem. Ahuja et al. [18] proposed a construction heuristic and VLSN search algorithm.

A new approach for tackling the WTA problem is proposed in this paper. Such an approach is a novel goal-based WTA doctrine. The proposed doctrine defines five merits for each weapon-threat (WT) pair that indicate the WT pair's ability to contribute in achieving overall optimal WTA if chosen. The merits of each WT pair are combined into a single quality value. Different approaches for combining WT pair merits are investigated. To reduce the possibility of weapons saturation, load balance policy is designed and attached to the process of merits weight specification. A quality matrix is formed from quality values of all WT pairs. The Hungarian method [19–21] is employed to arrive at optimal assignment based on the quality matrix. The results of our doctrine are compared with other WTA doctrines.

The paper is organized as follows. In Sec. II, a mathematical formulation of WTA problem is introduced. The proposed novel goal-based WTA doctrine is presented in Sec. III. In Sec. IV, the developed system for air defense mission analysis and design used in the proposed novel WTA doctrine is introduced, factors taken into consideration when designing AD missions to be used by the proposed doctrine are indicated, sample analysis data of AD mission analysis performed by the developed AD mission analysis and design software is shown, and the results of employing the proposed algorithm to solve general WTA problems are presented. Other goal-based WTA doctrines are employed for comparison. The results showed the superiority of our doctrine. Finally, Sec. V concludes the paper. The AD mission modeling and analysis methodology used in the developed package is presented in Appendix A. Merits calculation equations are explained in detail in Appendix B. An investigation of the ANN approach in determining merits weights equations parameters is given in Appendix C. Appendix D shows a thumbnail view of different AD missions used in the simulation and results.

II. Problem Formulation

The objective of WTA problem, as defined in AD context, is to find the optimal assignment of defense weapons to the attacking threats that will cause maximum inflicted damage on the attacking threats, and minimizing the damage inflicted by the threats on the assets protected by such defense weapons in a given air defense mission. The AD mission consists of N threats, M weapon systems and G assets. The size of this problem ($N \times M$) may be very large. An asset is defined as any entity of importance that must be protected from the threat's attack. Each asset has a value assigned to it that represents how much damage it can tolerate. Different threats have different characteristics (average speed, ordnance damage, and so on). A weapon is assumed to have an inventory of "shots" or free engagement slots. An engagement slot consists of one missile and one free guidance channel. An engagement slot can be reserved for a planned use at a predicted time. In the case of reserved engagement slot, its guidance channel is only considered free before and after the time interval of the planned engagement it is reserved to. A missile is considered available for use if there are enough missiles to cover for the planned future engagements. Defense weapons may have different characteristics.

Let

$$A(t) = \begin{bmatrix} a_{11}(t) & \cdots & a_{1N}(t) \\ \vdots & a_{ij}(t) & \cdots \\ a_{M1}(t) & \cdots & a_{MN}(t) \end{bmatrix} \tag{1}$$

represent the engagement plan as a function in time, where $a_{ij}(t)$ is the engagement state between weapon i and threat j at time t, and its value can be 1 for engaged or 0 for not engaged.

The engagement start time must be greater than the detection time instance. The engagement must be ended before the threat leaves the weapon effective range or reaches its ordnance delivery point. There is only one engagement per WT pair.

An engagement may consume more than one missile. Number of consumed missiles in each engagement is given by

$$S = \begin{bmatrix} s_{11} & \cdots & s_{1N} \\ \cdots & s_{ij} & \cdots \\ s_{M1} & \cdots & s_{MN} \end{bmatrix} \tag{2}$$

where s_{ij} is the number of missiles fired from weapon i targeting threat j in their engagement.

The WTA problem handled here is the static WTA. In static WTA number of threats N is known and constant over time, number of weapons M is known and constant over time, and all weapons engage targets in a single stage. That is, after the assignment has been made no assessment for the outcome for another assignment stage is done. However, our approach can be easily extended to handle dynamic WTA, as shown later in the next section.

Any weapon can kill any threat in range with kill probability dependent on the interception point. Let

$$pk_{ij} = \begin{bmatrix} \mathrm{SSPK}_1 \\ \mathrm{SSPK}_2 \\ \cdots \\ \mathrm{SSPK}_{s_{ij}} \end{bmatrix} \tag{3}$$

represent a vector consists of the single shot kill probability (SSPK) of each missile fired during the engagement between weapon system i and threat j.

It is assumed that the defense knows:

1) *How many offensive threats there are at each moment during the attack*: which can be easily obtained by the defense sensors.
2) *Each threat flight path and which asset is targeted by each threat*: such information can be estimated from the defense awareness of the possible important assets in the area under control by this command and control center, and the flight path taken by the threat so far.
3) *A rough estimation of the possible ordnance delivery point of each threat*: which can be obtained from the identified type of the aircraft, as each aircraft has a certain mission profile (Hi–lo–hi, Lo–lo–lo, and so on), and the estimated flight path, targeted asset, and average speed.
4) *The ordnance lethality value of each threat*: can be deduced from the identified type of the attacking threat as each type of aircraft can carry certain types of ammunition and the defense can assume the worst case by assuming that it carries the most deadly ordnance that it can carry.
5) *The kill probability as a function in the interception point for each weapon system*: such information can be obtained from real-life test data of the weapon system kill probability distribution in its effective range for different types of threats.

The above information is not to be deduced manually by human operator in the Command and Control center. A program is to be developed to deduce such information based on a database that carries all threat types and specifications, weapons specifications and test data, and important assets in the area for which this command and control center is responsible. Such a program is indeed under development at this time. After detection and identification, the threats are prioritized according to their danger to the assets. Also, the threat's attack scheme is not changed during the attack. No missile should be fired if the interception will happen after the threat reaches its ordnance delivery position. It will be a waste of resources with no contribution in preventing the threat from achieving its mission. Each threat that leaks through the defense is assumed to make a successful ordnance delivery. The probability that threat j will survive all missiles fired on it by weapon system i is given by

$$P_{\mathrm{miss}_{ij}} = \coprod_{l=1}^{s_{ij}} (1 - \mathrm{SSPK}_l) \tag{4}$$

The probability that threat j will survive all missiles fired upon it by all weapon systems is given by

$$P_{\mathrm{miss}_j} = \coprod_{i=1}^{M} P_{\mathrm{miss}_{ij}} \tag{5}$$

Each threat is assumed to target only one asset. On the other hand, each asset may be targeted by more than one threat. The information regarding the threats targeting each asset is represented by matrix **B**

$$\mathbf{B} = \begin{bmatrix} O_{11} & \cdots & O_{1N} \\ \cdots & O_{ij} & \cdots \\ O_{G1} & \cdots & O_{GN} \end{bmatrix} \tag{6}$$

where O_{ij} can be 0 if asset i is not targeted by threat j, or 1 if asset i is targeted by threat j.

D_{gj} represents the expected damage inflected on asset g by threat j, which can be calculated as follows

$$D_{gj} = O_{gj} D_j P_{\text{miss}_j} \tag{7}$$

The WTA problem is to minimize the following cost function which represents the total expected damage inflected on all the assets from all the threats, by finding an optimal engagement plane $A*$ and optimal firing schedule $S*$

$$C(A, S) = \sum_{g=1}^{G} \text{Min} \left\{ \sum_{j=1}^{N} O_{gj} D_j P_{\text{miss}_j}, h_g \right\} \tag{8}$$

where h_g is the health of asset g representing an enumeration of how much damage the asset can tolerate before destruction. The expected damage D_{gj} is expressed in the same measure units as the asset health h_g. If the inflicted damage on the asset is greater than the asset health, then the inflected damage is considered equal to the asset health. The excess damage is considered wasted and should not have any effect on the solution. As an illustrating example, consider a threat j with expected damage $D_{gj} = 60$ (that is, it can destroy an asset with health = 60 health units). This threat targets an asset g that has health $h_g = 25$ health unit, then the asset will be totally destroyed, and there is 35 excess damage. Such excess damage is considered wasted, as the asset can be destroyed with only $D_{gj} = 25$. Thus this excess damage should not be considered as part of the total inflicted damage in the mission. A and S are the chosen engagement plan and firing schedule respectively. The minimization of the cost function in Eq. (8) is subject to the following constraints:

1) Each threat j can target only one asset.

$$\sum_{g=1}^{G} O_{gj} = 1 \tag{9}$$

2) Number of missiles fired by weapon system i cannot exceed the number of missiles in this weapon inventory.

$$\sum_{j=1}^{N} s_{ij} \leqslant \text{INV}_i \tag{10}$$

where INV_i is the number of missiles in weapon system i inventory before the mission starts.

3) Number of concurrent engagement by weapon system i cannot exceed number of guidance channels on this weapon system

$$\sum_{j=1}^{N} a_{ij}(t_c) \leqslant GC_i \quad \forall t_c \leqslant T_{\text{mission}} \tag{11}$$

where t_c is any given time instance before the mission end time T_{mission}, and GC_i is the total number of guidance channels on weapon system i.

4) A missile will not be fired unless interception is guaranteed to be inside the weapon effective range resulted in a nonzero kill probability.

$$\text{SSPK}_l > 0 \quad l = 1, 2, \ldots, s_{ij} \tag{12}$$

5) An asset g is either targeted or not targeted by threat j.

$$O_{gj} \in \{0, 1\}. \tag{13}$$

6) A weapon i and threat j are either engaged or not engaged at time t

$$a_{ij}(t) \in \{0, 1\} \tag{14}$$

7) All engagements with threat j must end before the mission end time T_{mission} and before the threat reaches its ordnance delivery position T_{ordnance_j}

$$a_{ij}(t) = 0, \forall t > \text{Min}(T_{\text{mission}}, T_{\text{ordnance}_j}), i < M, j < N \tag{15}$$

Where T_{ordnance_j} is the time instance at which threat j reaches its ordnance delivery position.

The above formulation contains a large number of parameters and constraints. However, to represent a realistic WTA problem all such constraints are needed, moreover such formulation is not completely realistic as it targets the static WTA problem. The proposed doctrine can be extended to handle the dynamic WTA, as will be discussed in the following section. That is why the WTA problem has been shown to be a NP-complete problem with no exact solution. Until now, researchers have tried to provide an exact solution to special cases of the problem or heuristics that try to provide good approximate solutions. However, none of the heuristics proposed can have their solution accuracy validated for large problem size, as no exact algorithm is known to solve the WTA problem.

Our proposed heuristic provides a novel approach that employs a goal-based system with the Hungarian Assignment method [19–21] to provide a state of the art doctrine that achieves superior results compared with other WTA doctrines, as will be shown next through the rest of the paper.

III. Proposed Goal-Based Novel WTA Doctrine

WTA problem in AD missions can be described as a system contains number of servers (defensive weapon systems) that provide service to customers (threats). Each server has an area of influence and can only provide service to customers in this area. Customers are on the move and servers can provide service to them only when they pass through the servers' area of influence. Servers have limited resources to be used in providing service. A customer may receive service from more than one server. The quality of service (QoS) given to a customer depends on the customer attributes and the attributes of the server providing the service. Customers may arrive concurrently to a server's area of influence. If the server is unable to serve all customers in its area (saturated server), then some of them will get zero QoS from that server. A control manager (command and control center) exists in the system. Its objective is to maximize the overall quality of service (overall inflicted damage on threats) provided by all the servers by trying to make optimal assignment of customers to servers. The problem is to predict the quality of service that would be given to certain customer if assigned to certain server. The predicted QoS is needed by the control manager to be able to make the optimal assignment. This paper tries to solve this problem by trying to identify the most prominent factors that affect quality of service given by a server (defensive weapon system) to a customer (threat). These factors are enumerated and used in computing the QoS for a server–customer (WT) pair. The resulted QoS matrix representing all possible server–customer assignments are used along with the Hungarian method [19–21] by the proposed doctrine to find a proper server–customer (WT) assignment through an iterative approach.

A. QoS Factors

From a careful study of the nature of WTA problem and air defense missions, a set of factors are proposed to be the prominent factors affecting quality of service given by weapon system i to threat j. As both weapon system and threat attributes affect quality of service, some of the factors are related to threat attributes while the others are related to weapon system attributes.

Threat-Related Factors and Threat Prioritization

The QoS factors related to customer j (threat j) attributes are proposed to be:

- Time remaining until threat j reaches its ordnance drop point and accomplishes its mission.
- Lethality of the ordnance the threat carries.
- Importance of the asset targeted by this threat.
- The amount of damage the asset can tolerate before destruction.

A customer (threat) prioritization rule shown in Eq. (16) has been tailored to enumerate the factors related to customer attributes and affecting the quality of service:

$$\text{Threat\_Priority} = \frac{\text{Min}\{1,\ (\text{InflictedDamage}/\text{AssetHealth})\} * \text{Asset\_Importance}}{\text{Min}\left\{(T_{\text{ordnance}}, \text{Max}\left\{\bigcup_{i=1}^{N} T_{W\text{Max}_{\text{Leave}_i}}\right\}\right\} / T_{\text{Mission}}} \tag{16}$$

$$\text{InflictedDamage} = \text{ExpectedOrdnanceDamage} * (1 - PK_{\text{cumulative}})$$

'InflictedDamage' represents the amount of damage inflicted by the threat ordnance multiplied by probability of reaching the drop point successfully. 'AssetHealth' is an enumeration of the amount of damage the asset can tolerate. The minimum between 1 and the ratio between the Inflicted damage and asset health is taken in the equation as any excess damage greater than the asset health inflicted by the threat is considered wasted and has no further effect on the overall result. 'AssetImportance' is an enumeration of the importance of the targeted asset. T_{ordnance} is the time of ordnance drop by the threat. $T_{W\text{Max}_{\text{Leave}_i}}$ is a set consists of a single element represents the time when the threat leaves. Maximum effective range of weapon i. $\text{Max}\{\bigcup_{i=1}^{N} T_{W\text{Max}_{\text{Leave}_i}}\}$ represents time of leaving the last weapon maximum effective range after which the threat cannot be engaged (customer cannot receive any service). $\text{Min}\{(T_{\text{ordnance}}, \text{Max}\{\bigcup_{i=1}^{N} T_{W\text{Max}_{\text{Leave}_i}}\}\}$ represents the time after which there is no engagement going to happen with this threat (whether the reason is because the threat leaves the last area of influence of a weapon system or because it successfully accomplished its mission by dropping its ordnance after that has no significance as the objective is to prevent the protected assets from receiving damage). T_{Mission} is the air defense mission end time. Dividing $\text{Min}\{(T_{\text{ordnance}}, \text{Max}\{\bigcup_{i=1}^{N} T_{W\text{Max}_{\text{Leave}_i}}\}\}$ by T_{Mission} makes measurements relative to the current air defense mission thus causing the calculated priority not to be affected by the different air defense missions configurations.

Threat priority is directly proportional to asset importance and the expected amount of damage the threat would inflict on the asset, while it is inversely proportional to the amount of time remaining until engagement with the threat becomes insignificant.

After calculating threat priorities, a normalization step takes place to make the priority range from zero to one.

Weapon System-Related Factors

The factors related to a weapon system (server) attributes and affecting QoS given by weapon System i to threat j are proposed to be:

- The overall cumulative kill probability that can be achieved by weapon system i regarding threat j.
- Available missiles in weapon system i inventory.
- Available free guidance channels in a weapon system i when threat j is in the weapon system area of influence.
- The saturation level of weapon system i if threat j is assigned to it. In other words, the effect of assigning threat j to weapon system i on the weapon system's ability to engage other threats concurrently or disjointedly in time.

The fourth factor actually does not affect the QoS given to threat j but it has an effect on the QoS given to the other threats by this weapon system, thus affecting the overall QoS provided by the system. These factors are enumerated by the computation of the following five merits values for the WT pair (server–customer pair):

1. Cumulative kill probability that can be achieved with the available engagement slots in the weapon system regarding the threat in concern, before the threat reaches its ordnance delivery point and achieves its mission.
2. The ratio between number of engagement slots (engagement slot consists of a single missile and a single free guidance channel) that will be free for use during weapon system engagement with this threat and the number of other threats that can be engaged *(engagable)* by this weapon system during this WT pair engagement. This merit measures weapon system ability to perform concurrent engagement if this threat is allocated to it. The smaller the value of this merit, the more this weapon system is susceptible to saturation.

3. The ratio between the number of engagement slots that will be available after this engagement ended and the number of other threats *engagable* by this system nonconcurrently with this engagement. This merit measures weapon system ability to perform nonconcurrent engagements if this threat is allocated to it. The smaller the value of this merit, the more this weapon system is susceptible to saturation.
4. The ratio between the remaining number of missiles in the weapon system inventory (after subtracting the missiles that will be needed to achieve the cumulative kill probability of this threat engagement) and the number of the other threats *engagable* by this system. This merit measures the effect of this engagement on the weapon fire power (missiles) that can be used in other possible engagements.
5. The complement of the ratio between the amount of time the engagement will take in the worst case and the total duration of the mission. The smaller this amount of time the quicker the used guidance channels in this engagement will be free to be used in another engagement.

All merit calculations are tailored to be relative to the air defense mission attributes. The second merit divides number of free engagement slots by number of other possible concurrent engagement to make it relative to the actual need for these engagement slots. The same has been done with merit 3 and merit 4. For the same reason, merit 5 has the engagement duration divided by the total mission duration. Figure 1 summarizes the QoS factors and their enumerations.

Weapon-Threat Pair QoS

To arrive at efficient WT assignment (and hopefully an optimal one), the control manager needs to know the QoS of all possible assignments before making any assignment actually. Such QoS values are represented by the quality

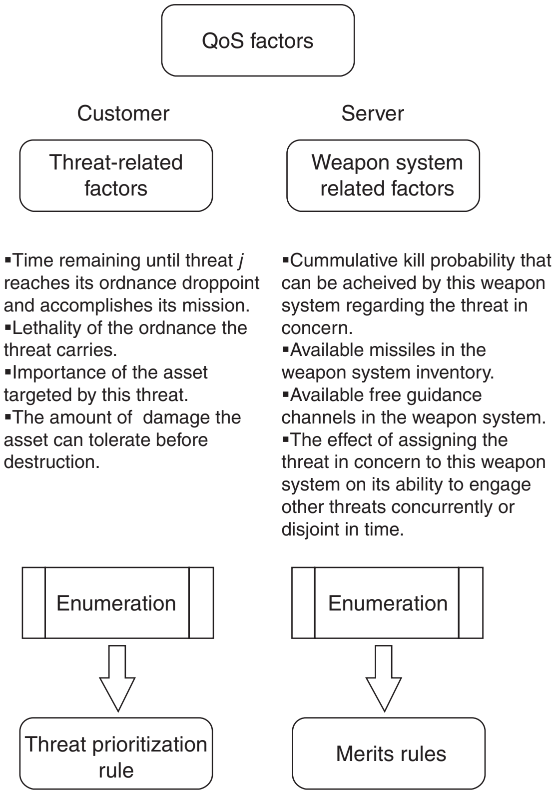


Fig. 1 Factors affecting QoS given by weapon system *i* to threat *j* in AD missions, and the results of their enumeration.

matrix **Q**

$$\mathbf{Q} = \begin{bmatrix} q_{11} & \cdots & q_{1N} \\ \vdots & q_{ij} & \cdots \\ q_{M1} & \cdots & q_{MN} \end{bmatrix} \tag{17}$$

where q_{ij} represents the QoS given to threat j if assigned to weapon system i. Such QoS value is determined by combining the calculated merits values for that WT pair using weighted summation. Each WT pair merit is assigned a weight to represent its importance. The sum of merits weights is a unit value, and they are all positive. The importance (weight) of a given merit type is not the same over all WT pairs. Merits weights are tailored to be function in threat priority to include the effect of threat-related factors. For example, the higher the threat priority is, the greater the importance of the weapon ability to kill it. This is resulted in a greater importance for the first merit. Furthermore, if the cumulative kill probability $= 0$, that is the weapon is incapable of inflicting any damage on the threat, then the QoS value should be set to 0 explicitly.

In addition, the weights values are affected by a load-balancing policy designed to try to prevent weapon systems from being saturated as much as possible as such saturation can have a significant effect on the overall outcome of the air defense mission. For example, if a weapon system is already saturated (threats cannot be engaged as the weapon system is busy or has no sufficient resources) the cumulative kill probability for those nonengaged threats become 0 for that system resulting in assignment to other systems although it is possible that the saturated system if free may provide better cumulative kill probability. Thus it is better to have some policy for load balancing to keep weapon systems far from saturation. Such a policy can have a bad effect on the overall kill probability by reducing the weight of the first merit as will be described next. For that, it is to be activated only when a system approaches near saturation to prevent greater losses. Near saturation state is a state when a system has the percentage of its free engagement slots from its overall engagement slots given at the start of the mission, drops below a specific threshold value $S_{\text{Threshold}}$. The load balance policy should increase the importance, by increasing the weights, of merits 2, 3, 4, and 5, especially merits 2 and 3, as they have great impact on saturation prevention as they represent the effect of the current engagement on the weapon system's ability to engage other threats. Such an increase in the weights of merits 2 through 5 is accompanied by a decrease in the weight of the first merit by the same amount, so that the sum of the weights is always 1 and all weights are nonnegative. The amount of weights shifts should be a function of the severity of the near saturation state; that is, how close the weapon is to saturation.

The load balancing policy adjusts only the weights of the WT pairs involving the near saturated system with only the threats that have not yet been assigned to this weapon.

There are three general guidelines that govern merits weights estimation and shifting:

- First merit, that represents the cumulative kill probability, is of major importance as it has a substantial effect on the threat survivability and its importance increases when the threat priority increases.
- If the percentage of free engagement slots is less than the $S_{\text{Threshold}}$ value, then the load balance policy is activated. The less the percentage of engagement slots, the more important merits 2 through 5 (especially merits 2 and 3) and the less important merit 1 becomes (owing to load balance policy).
- When the load balance policy is active for a given WT pair, the greater the threat priority the greater the desire to destroy it. Thus, the shift in merit importance is lessened owing to load balancing policy for this pair, as an importance shift will lead to a decrease in the first merit importance. For the WT pairs containing the highest priority threat, there is no shift in merit importance at all owing to the load balance policy. For the WT pairs containing the lowest priority threats the shift in importance is at maximum. Between the highest and the lowest priority threats, the shift in merit importance is assumed to be a linear function in the threat priority. By applying this policy, the danger of saturation is resolved, sacrificing the possibility to inflict more damage on low priority threats than on high priority threats.

Such rules can be presented by the following proposed weights equations

$$\mathrm{Pr}_{\text{complement}} = (1 - \text{Priority}) \tag{18}$$

1st Merit weight

$$W_1 = (W_{\mathrm{hp}_1} * \text{Priority} + W_{\mathrm{lp}_1} * (1 - \text{Priority})) - \mathrm{Pr}_{\mathrm{complement}} W_{\mathrm{shift}} \Rightarrow \text{Load balance policy On} \tag{19}$$

$$W_1 = W_{\mathrm{hp}_1} * \text{Priority} + W_{\mathrm{lp}_1} * (1 - \text{Priority}) \Rightarrow \text{ Load Balance Policy Off} \tag{20}$$

ith Merit weight $i = 2, 3, 4, 5$

$$W_i = (W_{\mathrm{hp}_i} * \text{Priority} + W_{l\mathrm{p}_i} * (1 - \text{Priority})) + W_{\mathrm{hp}_i} \mathrm{Pr}_{\mathrm{complement}} W_{\mathrm{shift}} \Rightarrow \text{Load balance policy On} \tag{21}$$

$$W_i = W_{\mathrm{hp}_i} * \text{Priority} + W_{\mathrm{lp}_i} * (1 - \text{Priority}) \Rightarrow \text{ Load Balance Policy Off} \tag{22}$$

where ‘Priority’ is the priority of the threat in the WT pair for which we calculate the merits weights. W_{shift} is the weight shift value owing to load balance policy expressed as a factor between 0 and 1, where 1 means 100% shift in weights.

W_{hp_i} represents the weight of merit i of a given WT pair when the threat has the highest priority.

W_{lp_i} represents the weight of merit i of a given WT pair when the threat has the lowest priority and load balance policy is off.

When load balance policy is off, the ith merit weight W_i of a given WT pair is calculated by linear interpolation from W_{hp_i} and W_{lp_i}. Such interpolation is based on the priority of the threat in the WT pair of interest. The higher the threat priority, the closer W_i to W_{hp_i}. The amount of weight change is due to load balancing $= \mathrm{Pr}_{\mathrm{complement}} * W_{\mathrm{shift}}$. When load balance policy is active, the weight of the first merit is reduced by $\mathrm{Pr}_{\mathrm{complement}} * W_{\mathrm{shift}}$, so that the pair containing the highest priority threat will not suffer weight change in its first merit weight owing to load balance policy, and the pair containing the lowest priority threat will suffer maximum weight change equal to W_{shift} in its first merit weight owing to load balance policy. The weight amount reduced from the first merit weight is distributed to the other four merits weights based on W_{hp_i} of each of these merits, so that the summation of all five merits weights remains a unit value. The following constraints must hold to guarantee the validity of the above equations:

- $W_{\mathrm{shift}} < W_{\mathrm{lp}_1}$. If not, equation (19) may result in negative value.
- $W_{\mathrm{lp}_1} < W_{\mathrm{hp}_1}$. (the importance of the first merit for lowest priority threat is less than that of the highest priority threat)
- $\sum_{i=1}^{5} W_{\mathrm{hp}_i} = 1, \sum_{i=1}^{5} W_{\mathrm{lp}_i} = 1, \sum_{i=1}^{5} W_i = 1$. (summation of all 5 merits weights of a given WT pair is a unit value)

A method is needed for specifying the proper weight values of the five merits for the pairs involving the highest and lowest priority threats (represented by W_{hp} and W_{lp} vectors, respectively), the amount of importance shifting (W_{shift}) done by the load balance policy, and at what percentage of free slots ($S_{\mathrm{Threshold}}$) the load balance policy will be activated.

Merits Weight Parameters Estimation

To calculate the merits weights for each pair using Eqs. (19) to (22), optimal values for the weights equations parameters (W_{hp}, W_{lp}, W_{shift}, and $S_{\mathrm{Threshold}}$) must be specified first. Such parameters will be referred to as a *Solution Vector* from now on in this paper. At first, the possibility of using modular multi-layer-perceptron (MLP) type ANN has been investigated to find such optimal values for the solution vector. The result of such investigation is the conclusion that the investigated (MLP) type ANN cannot find the optimal solution vector, as the inflicted damage as a function in the weights is found to be nondifferentiable in this problem. As a result the gradient descent optimization method used by the neural network will fail to approach the minimum, as it takes steps proportional to the *negative* value of the gradient, and in this case the gradient is zero. The inflicted damage function is found to be a step function, which seems logical as the changes in weights cause a change in the assignment. The inflicted damage, resulting from the new assignment, presents a sudden change from the previous inflicted damage because there is no relation between them as they resulted from different assignments. A detailed description and results of such investigation is given in Appendix C.

After discovering that the damage function is a step function, several alternative approaches have been investigated. Sequential search approach was found more suitable, because if a sequential search is conducted with suitable step size the probability of finding the global minimum on step function is higher than on any other continuous function, because the global minimum is more likely to be an interval rather than a single point. The drawback of using the sequential search is that it can take an impossibly large amount of time to span all possible values of W_{hp}, W_{lp}., weight shift (W_{shift}), and load balance policy threshold $S_{Threshold}$. But if the logical constraints of merits equations are imposed, it will substantially reduce the size of the solution space without sacrificing the solutions that can be the global minimum. Such a reduction in the solution space size makes a sequential search usable for this problem. However, this search does take up a large amount of time, but this should be tolerated since it is performed only once, obtaining the optimal solution vector for global minimum gross damage for the AD missions used in the search. Once such a solution vector has been found, it will be used in the algorithm with any other possible AD mission. AD missions used in the sequential search algorithm are carefully designed to represent most possible situations.

Figure 2 is given below for the algorithm that employs sequential search to achieve the optimal value for the solution vector (W_{hp}, W_{lp},, W_{shift}, and $S_{Threshold}$) that will achieve minimum gross damage for all the missions. Gross damage is the summation of all missions damages for a given solution vector (W_{hp}, W_{lp},, W_{shift}, $S_{Threshold}$). The algorithm spans the solution space using the sequential search with appropriate step. For each valid solution vector, the proposed WTA doctrine, which will be illustrated shortly afterward, is used on a number of AD missions. For each mission, after allocating all threats, the overall mission expected damage is calculated and the information about the solution vector that caused the minimum expected damage for this mission is updated when needed. The same is done to the solution vector that caused the minimum gross damage for all the missions.

The process of finding the optimal solution vector (W_{hp}, W_{lp},, W_{shift}, $S_{Threshold}$) using sequential search is *not* to be executed every time the proposed doctrine is used. It is executed only once to obtain the optimal values, so that Eqs. (19)–(22) can be used in the proposed doctrine when it is applied.

B. Proposed WTA Doctrine

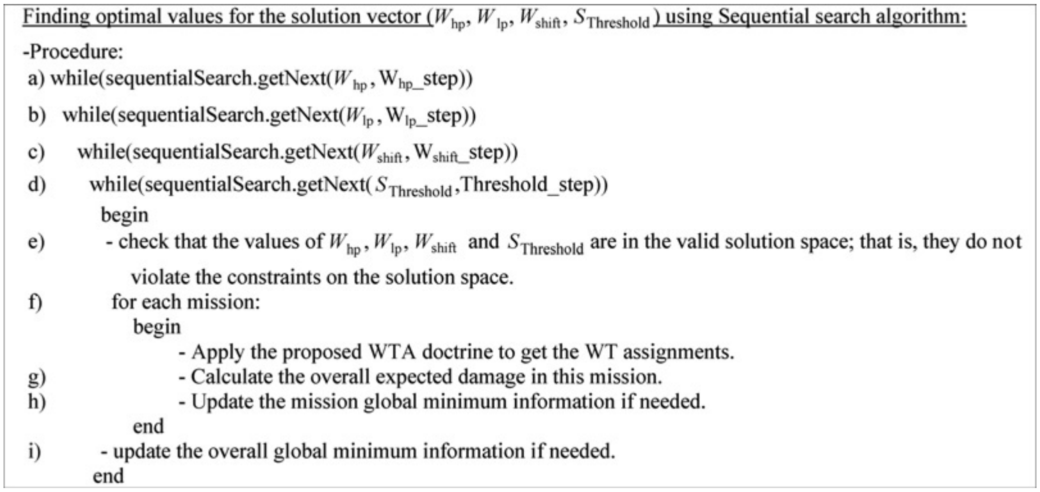
Figure 3 is given below showing the proposed WTA doctrine, which is explained as follows:

Initially:

Fictitious weapon systems are introduced in the quality matrix **Q** to make the number of weapon systems equal the number of threats (square quality matrix). The qualities of the pairs that contain fictitious weapons are set to zero in the quality matrix. Threats are prioritized normally according to the prioritization rule proposed in Eq. (16).

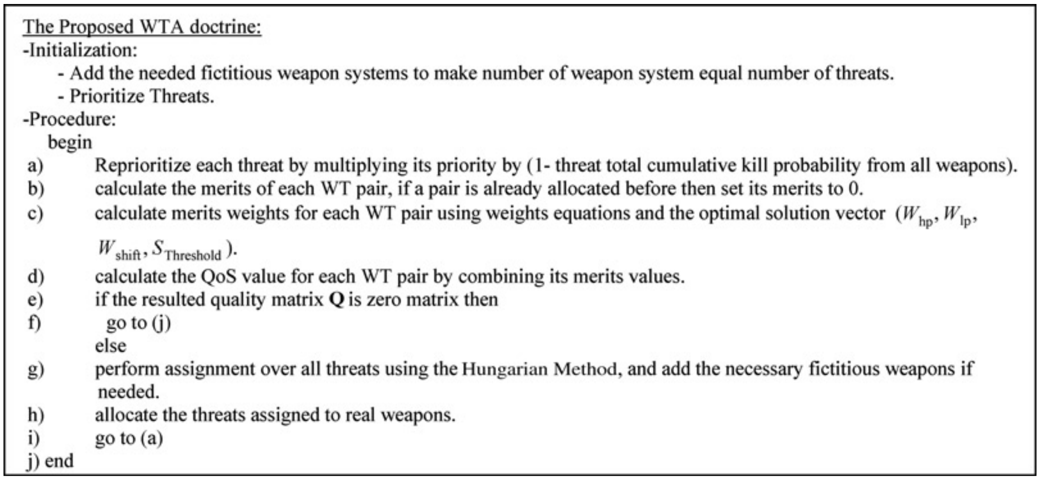
In each iteration:

Reprioritization of the threats is done by multiplying the threat priority by (1- threat total cumulative kill probability from all weapons allocated to it so far). Such a step is required to reduce the priority of the already allocated threats



```
Finding optimal values for the solution vector ($W_{hp}$, $W_{lp}$, $W_{shift}$, $S_{Threshold}$) using Sequential search algorithm:
-Procedure:
 a) while(sequentialSearch.getNext($W_{hp}$, $W_{hp}$_step))
 b)   while(sequentialSearch.getNext($W_{lp}$, $W_{lp}$_step))
 c)     while(sequentialSearch.getNext($W_{shift}$, $W_{shift}$_step))
 d)       while(sequentialSearch.getNext($S_{Threshold}$,Threshold_step))
            begin
 e)           - check that the values of $W_{hp}$, $W_{lp}$, $W_{shift}$ and $S_{Threshold}$ are in the valid solution space; that is, they do not
                violate the constraints on the solution space.
 f)           for each mission:
                begin
                    - Apply the proposed WTA doctrine to get the WT assignments.
 g)                 - Calculate the overall expected damage in this mission.
 h)                 - Update the mission global minimum information if needed.
                end
 i)           - update the overall global minimum information if needed.
            end
```

Fig. 2 Pseudo code of the algorithm for finding optimal values for (W_{hp}, W_{lp}, W_{lp}, W_{shift}, $S_{Threshold}$) using Sequential search algorithm.



```
The Proposed WTA doctrine:
-Initialization:
    - Add the needed fictitious weapon systems to make number of weapon system equal number of threats.
    - Prioritize Threats.
-Procedure:
   begin
 a)     Reprioritize each threat by multiplying its priority by (1- threat total cumulative kill probability from all weapons).
 b)     calculate the merits of each WT pair, if a pair is already allocated before then set its merits to 0.
 c)     calculate merits weights for each WT pair using weights equations and the optimal solution vector ($W_{hp}$, $W_{lp}$,
        $W_{shift}$, $S_{Threshold}$ ).
 d)     calculate the QoS value for each WT pair by combining its merits values.
 e)     if the resulted quality matrix Q is zero matrix then
 f)       go to (j)
        else
 g)     perform assignment over all threats using the Hungarian Method, and add the necessary fictitious weapons if
        needed.
 h)     allocate the threats assigned to real weapons.
 i)     go to (a)
j) end
```

Fig. 3 Pseudo code of the proposed WTA doctrine

so that the threats assigned before to fictitious weapons, or the threats that receive low cumulative kill probability previously, gain higher priority in this iteration. Merits of each WT pair are then calculated. After that merits weights for each WT pair are calculated using weights equations and the optimal solution vector (W_{hp}, W_{lp},, W_{shift}, $S_{\mathrm{Threshold}}$) found before using the sequential search. The QoS value of each pair is then calculated by merits weighted summation. If a threat is already allocated to the same weapon system in a previous iteration then this WT pair quality is explicitly set to zero so that it is not chosen again by the assignment algorithm. Then the assignment is done using the Hungarian assignment algorithm shown in Eqs. (19) to (21). Only threats assigned to real weapon systems are allocated to them.

When the quality matrix **Q** is found to be zero matrix, then all possible threats allocation took place for this mission and the doctrine is terminated. Figure 4 shows an illustration of the proposed WTA doctrine.

C. Optimal Assignment Using Hungarian Method

The objective is to arrive at the optimal assignment based on the QoS matrix. To achieve such assignment the Hungarian method shown in Eqs. (19) to (21) is used in step (g) from Fig. 3 that shows the proposed WTA doctrine. The Hungarian algorithm is a combinatorial optimization algorithm which solves assignment problems in polynomial time ($O(n^3)$). The algorithm models an assignment problem as an $m \times n$ cost matrix, where each element usually represents the cost of assigning the ith worker to the jth job. Here the quality matrix **Q** serves as the assignment problem model, as each element in QoS matrix represents the quality of assigning the ith weapon to the jth threat. In this case the Hungarian method arrives at an assignment that has the maximum overall quality value.

D. Possibility of Handling Dynamic WTA Problem

The Dynamic WTA problem is a more realistic WTA problem. In the dynamic WTA problem, new threats may emerge during the mission, weapon systems and assets may be destroyed. As a result, number of threats N, number of weapons M, and number of assets G is not constant over the mission time. The modification to the proposed doctrine to be able to handle the dynamic WTA problem is simple. In static WTA the steps of the proposed doctrine is performed once at the beginning of the mission, as soon as the required information for calculating the merits is deduced, to achieve the assignment once and for all. For the Dynamic WTA, instead of performing such doctrine once to obtain the assignment, the proposed doctrine is to be performed every time the number of threats, number of weapons, or number of assets changes. Thus, when a new threat emerges, or a weapon system or asset is destroyed, there should be no worry about the delay resulting from executing the doctrine multiple times as the execution time is governed by the execution time of the Hungarian method, which is $\mathrm{O}(\mathrm{n}^3)$ where n is max{N,M}.

This could happen in the middle of ongoing weapon-threat engagements. The guidance channels that are busy in the current engagements of the old assignment can be freed from the old assignment after completing the guidance of the current launched missiles.

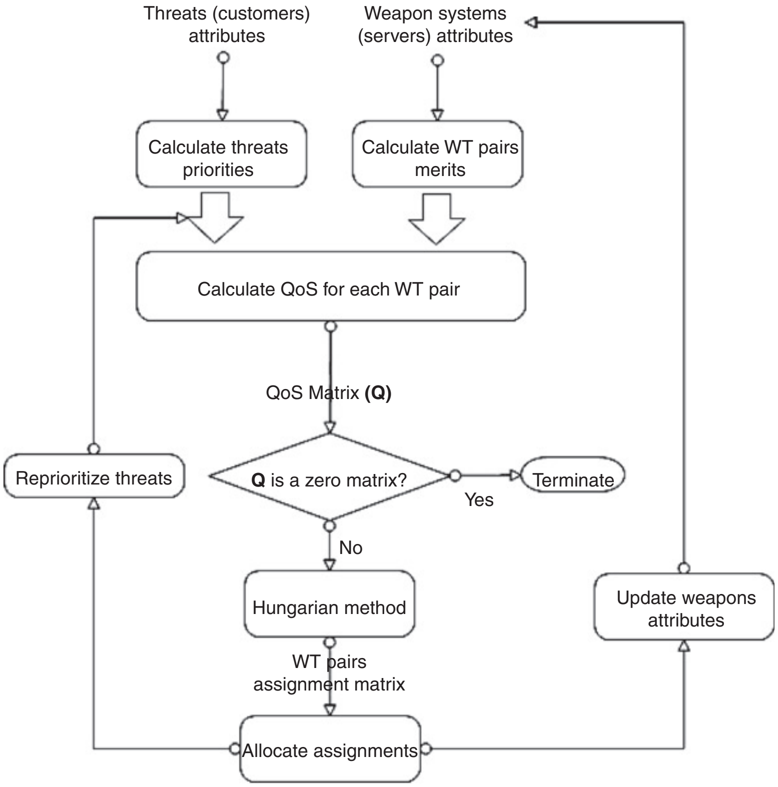


Fig. 4 An illustration of the proposed WTA doctrine.

IV. Simulation and Results

A. Air Defense Mission Analysis and Design (ADMAD) System

To be able to calculate the merits of a WT pair, an AD mission design and analysis package has been developed. In such package, each AD mission contains three types of entities: assets, weapon systems, and threats.

Asset

The asset is defined by position, health (represents the asset ability to sustain damage), and importance.

Weapon

A weapon entity is defined by specifying its position, number of guidance channels, number of missiles in inventory, nonprecision tracker range (detection range), average time needed for detection by the nonprecision tracker, average time needed to establish a firm track by the nonprecision tracker, precision tracker range, average time needed for precision tracking and weapon aiming, average time needed to perform kill assessment, maximum weapon range and minimum weapon range, average missile speed, and kill probability.

Threat

A threat entity is defined by initial position, average speed, flight path, ordnance delivery point, expected ordnance damage (expressed in the same units as assets health), the targeted asset, and threat priority which is calculated by the ADMAD package.

After defining the entities in the AD mission, ADMAD analyzes the mission to conclude the following information for each possible WT pair:

- Engagement Start time and end time.
- The time and threat position at which the threat is detected, and the threat position at this time.

- The time and threat position at which nonprecision Firm track is established.
- The time and threat position at which precision Firm track is established.
- The time and threat position at which the threat enters the weapon maximum range.
- The time and threat position at which the threat leaves the weapon maximum range.
- The time and threat position at which the threat enters the weapon minimum range.
- The time and threat position at which the threat leaves the weapon minimum range.
- The missiles launch times and the threat interception times.

The analysis that generates above information, is conformant with the constraints in the problem description and formulation. The information above is generated for each WT pair before any allocation of threats to weapons.

The AD mission modeling and analysis methodology employed in ADMAD package is introduced in Appendix A.

After generating the engagement information for WT pairs, the package calculates the threats priority and the merits of each WT pair. Appendix B provides detailed description on how the WT pair merits are calculated in the ADMAD package.

B. Air Defense Missions

Missions Design Considerations

Air defense missions are designed to be used in the proposed algorithm. Such missions were carefully designed to take into consideration most possible tactical situations encountered in any Air defense mission. These considerations include the following factors:

1. Different Attack Patterns

Threats attacks can be:

- Wave attack vs stream attack: the wave attack consists of a group of threats that approach the defended site from different directions with near simultaneous arrival times. On the other hand, stream attack consists of a group of threats that approach on the same bearing but with some time between each. Stream attack could result in high saturation possibility for the weapon systems in the sector that faces the incoming threats.
- Threats arrivals are closely spaced in time vs. distantly spaced in time: closely spaced arrivals results in small overall mission time, and an increase in concurrent engagement by the weapon systems, which leads to increase in the importance of guidance channels as a resource.

2. Weapon Systems Placement

Weapon system placement layout can be disjoint or layered. Layered defenses are defenses with overlapping weapon ranges. Such types of defense are usually employed in AD missions especially in protection of highly important, and/or closely spaced assets. Although layered defenses provide better protection, in such missions WTA decisions become more difficult and play a greater role in deciding the overall mission outcome.

Mission Description

Nine AD missions have been designed and exploited by the algorithm (see Figure 2) used in the search for an optimal solution vector. A thumbnail view of the tactical picture of these missions is given in Appendix D.

A detailed description of a sample mission is given next.

1. Tactical Picture

Figure 5 shows the sample mission tactical picture and layout. Two assets protected by layered defense which consists of four weapon systems. The attack force consists of two groups of threats. The first group targets asset1. It consists of four threats which arrive almost simultaneously to asset1. The other attack group targets asset2. Group2 consists of 5 threats which are spaced in time of arrivals. The two assets have the same importance and health. Each threat is *engagable* by any of the four weapon systems.

2. Mission Configuration

Mission detailed configurations are given below:

a- Assets:

There are two assets in the mission with identical health and importance. Table 1 shows the assets configurations.

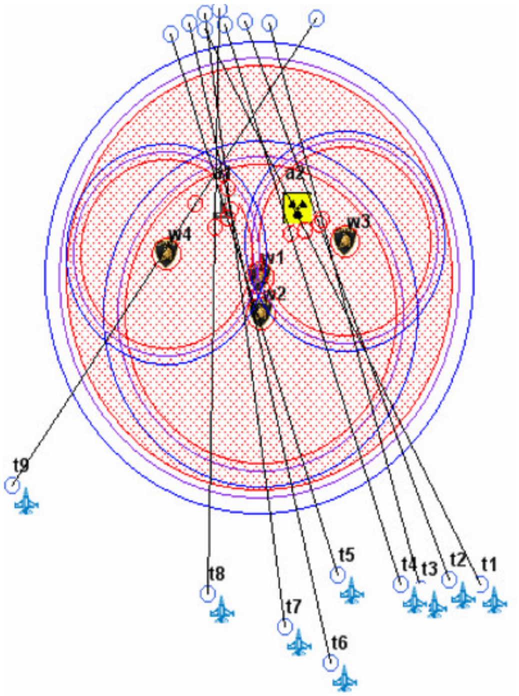


Fig. 5 Mission tactical picture and layout.

Table 1 Assets configurations

Asset Name	Health	Importance
a1	150	1
a2	150	1

b- Weapon Systems:

Each weapon system in Fig. 5 has different ranges indicated by circles in different colors:

- Detection range: Blue circle.
- Precision tracker range: Purple circle.
- Weapon effective range: Red dotted region.

Configurations for mission weapon systems are indicated in Table 2.

c- Threats:

The mission contains nine threats. The path traveled by each threat is drawn. Each threat has an ordnance delivery position indicated by small red circle on the threat path. Table 3 shows mission threats configurations.

3. Possible Engagements Analysis

There are 36 WT pairs. The results of all pairs engagements analysis are visualized in Fig. 6.

Events of each engagement are indicated by colored dots on the threat path. Several dots of the same color may appear on a threat path owing to visualization of all engagements involving this threat on the same path:

- *Detection event*: is indicated by aquamarine dot on the threat path (may not be obvious as the black dot representing nonprecision tracking may coincide with it as the two events time difference is very small).
- *Nonprecision firm tracking event*: is indicated by a black dot on the threat path.
- *Precision firm tracking event*: is indicated by a beige dot on the threat path.
- *Missile launch event*: indicated by an orange dot on the threat path.
- *Threat interception event*: indicated by a red dot on the threat path.

Table 2 Weapons configurations

Weapon name		W1	W2	W3	W4
Guidance channels		1	1	1	1
Missiles inventory		4	4	4	4
Detection radius[‡]		150	110	70	70
Avg. detection time (t0)		1	1	1	1
Avg. firm-track time* (t1)		0.2	0.2	0.2	0.2
Precision-tracker radius[‡]		140	100	65	65
Avg. precision- tracking and aiming time (t2)		2	2	2	2
Avg. kill assessment time (t_{ka})		5	4	3	3
Avg. missile speed (map points/s)[†]		7	7	7	7
Weapon effective range	Min[‡]	10	10	10	10
	Max[‡]	135	95	60	60
Kill probability (PK)		0.7	0.7	0.7	0.7

*Firm tracking done by nonprecision tracker. [†]The map size used is 512 × 512 points. [‡]All ranges are expressed in map points.

Table 3 Threats configurations

Threat name	Ordnance damage	Priority	Velocity (map points/s)	Targeted asset	Delivery point	
					X	**Y**
t1	50	0.4392324	5	a2	282	147
t2	50	0.6372458	5	a2	292	144
t3	50	0.6478215	5	a2	294	143
t4	50	0.774303	5	a2	272	151
t5	50	0.7134833	5	a1	228	141
t6	50	0	5	a1	230	137
t7	50	0.1941773	5	a1	228	126
t8	50	0.7748292	5	a1	219	144
t9	50	1	5	a1	205	129

Analysis of (w1, t1) Engagement

Figure 7 shows a detailed description of (w1, t1) engagement. Note that the dots representing detection time and nonprecision tracking time are nearly coincident as the time needed to establish nonprecision firm track is just 0.2 s, which is why only the black dot appears. In addition, once precision firm tracking is established, the 1st launch is initiated, which is why the beige dot representing the Precision firm track event is overwritten by an orange dot representing the first launch as, in this particular example, when precision firm tracking has been established, the threat is inside the weapon's effective range, permitting the first launch to take place immediately. Table 4 represents detailed analysis for (w1, t1) pair engagement. Table 5 shows the calculated merits for the (w1, t1) pair.

(w1, t1) merits are calculated using (B.1)-(B.5) in Appendix B.

C. Proposed WTA Doctrine

The algorithm, shown in Fig. 2, has been executed with the configuration shown in Table 6.

The global minimum of the gross damage over all missions found at solution vector given in Table 7.

After finding the global minimum shown in Table 7, sequential search was executed again with greater precision in the neighborhood of the resulted $W_{\text{hp}}$and $W_{\text{lp}}$of the global minimum solution vector. No better results were found after this higher precision neighborhood search. Expected damage and WTA of each mission at the global minimum configuration is given in Table 8.

Optimal minimum damage of each mission is given in Table 9.

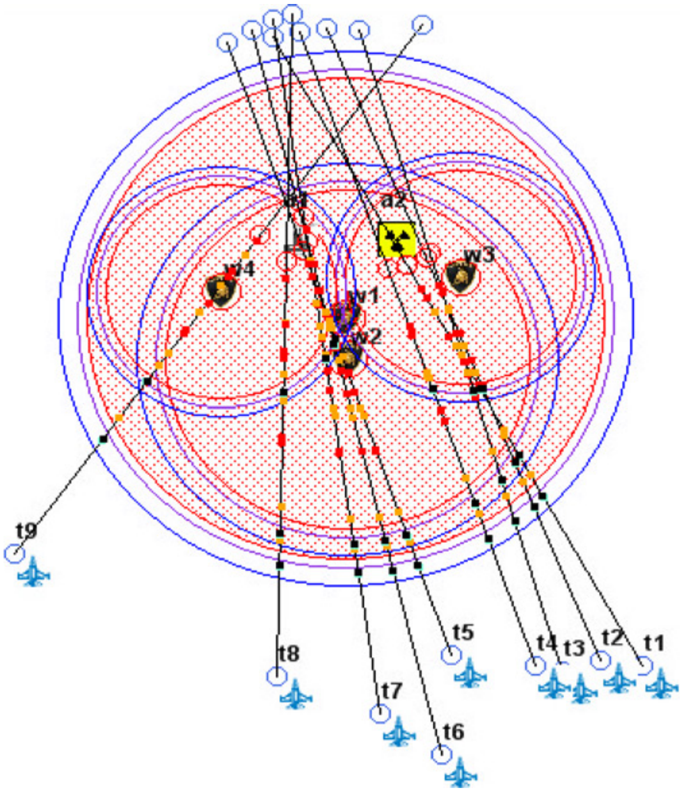


Fig. 6 Engagement analysis of possible WT pairs.

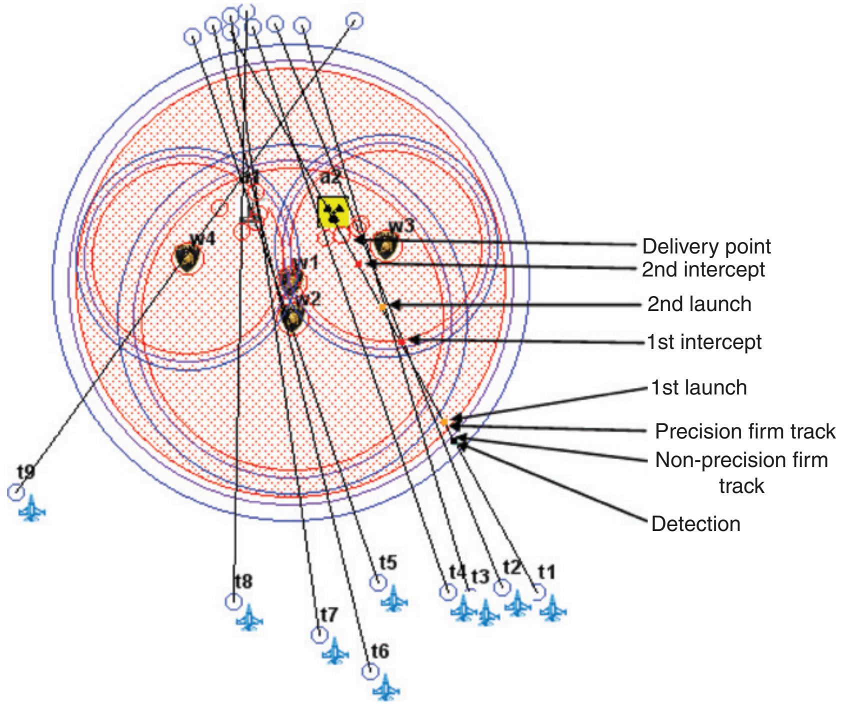


Fig. 7 Detailed description of (w1, t1) pair engagement.

Table 4 (w1, t1) pair engagement analysis

Engagement start time	Engagement end time	Delivery time	Detection time	NPF track time		PF track time	
22.553936	51.1096878	51.1096878	21.4948769	21.6948776		24.553936	
				Launch times		Intercept times	
Weapon maximum range entrance time	Weapon maximum range leave time	Weapon minimum range entrance time	Weapon minimum range leave time	1st launch time	2nd Launch time	1st intercept time	2nd intercept time
23.5440025	74.55736	not exist	not exist	24.553936	40.8449631	35.8449631	47.0873642

Table 5 (w1, t1) pair merits value

Merit1	Merit2	Merit3	Merit4	Merit5
Cumulative kill probability	Free engagement slots ratio during pair engagement	Free engagement slots ratio after pair engagement	Avg. missile count per engagement after this engagement	Complement of engagement time ratio
0.91	0	1	2	0.6278937

Table 6 Configuration parameters of the algorithm in Listing.1 used in finding optimal values for solution vector (W_{hp}, W_{lp}, W_{shift}, $S_{\mathrm{Threshold,S}}$) using sequential search

		Range	
	Step size	Min	Max
W_{hp}	0.1	(0,0,0,0,0)	(0.9,0.9,0.9,0.9,0.9)
W_{lp}	0.1	(0,0,0,0,0)	(0.9,0.9,0.9,0.9,0.9)
Weight shift (W_{Shift})	0.05	0	1
LoadBalance threshold ($S_{\mathrm{Threshold}}$)	0.1	0	1

Table 7 Optimal values for (W_{hp}, W_{lp}, W_{shift}, $S_{\mathrm{Threshold}}$)

Load balance threshold ($S_{\mathrm{Threshold}}$)	Weight shift (W_{shift})	W_{lp}	W_{hp}	Global minimum gross damage
[0,1]	[0,0.4]	(0.5, 0.2, 0.2, 0.1, 0)	(0.9, 0, 0, 0.1, 0)	1272.3

The relative error between optimal mission damage given in Table 9 and mission damage at the global minimum gross damage given in Table 8 is shown in Fig. 8, where relative error is computed by Eq. (23)

$$E_{\mathrm{relative}_i} = \frac{\mathrm{Damage}_i - \mathrm{Damage}_{\mathrm{optimal}_i}}{\sum_{k=1}^{9} \mathrm{Damage}_{\mathrm{optimal}_k}} \quad i = 1, 2, 3, 4, 5, 6, 7, 8, 9 \tag{23}$$

where i is the mission id.

The proposed algorithm has the best performance when compared with other goal-based WTA assignment doctrines. A goal-based assignment doctrine is an assignment doctrine that aims at achieving a minor simple goal which hopefully will lead to the achievement of the major goal of minimizing expected damage over the assets. Three goal-based doctrines for WTA assignment are suggested by Macfadzean [22] to establish the fact that there is no single WTA assignment doctrine that can achieve good results in all possible situations encountered in AD missions. A brief description of each doctrine is given below:

Table 8 WTA and expected damage of the nine AD missions, given in Appendix D, at the optimal values for the solution vector (W_{hp}, W_{lp}, W_{shift}, $S_{\text{Threshold}}$) given in Table 7

Mission	Expected damage	WT assignment (WTA)
Mission1	143.5	(w1,t2),(w1,t1),(w2,t4), (w2, t5), (w3, t6),(w3,t7).
Mission2	226.35	(w2, t2),(w1,t1),(w3,t5).
Mission3	194.85	(w1,t2),(w1,t1),(w2,t6), (w2,t2),(w3,t3),(w3,t4), (w4,t8),(w4,t3).
Mission4	199.5	(w1,t3),(w1,t5),(w2,t1), (w3,t4),(w4,t7),(w4,t3).
Mission5	129.85	(w1,t4),(w1,t7),(w2,t1), (w3,t3),(w3,t2),(w4,t5).
Mission6	38.4	(w1,t9),(w1,t8),(w2,t6), (w2,t4),(w3,t2),(w3,t1), (w3,t5),(w4,t7),(w4,t8), (w5,t3),(w5,t5),(w6,t1).
Mission7	114.35	(w1,t8),(w1,t2),(w1,t1), (w2,t4),(w2,t6),(w2,t2), (w3,t3),(w3,t1),(w4,t5), (w4,t9),(w4,t6).
Mission8	108.5	(w1,t2),(w1,t7),(w2,t4), (w2,t3),(w3,t6),(w3,t5).
Mission9	117	(w1,t2),(w1,t3),(w1,t1), (w2,t7),(w2,t1),(w2,t2), (w3,t6),(w3,t4),(w3,t7).

Table 9 Optimal assignment for each mission

	Expected damage	WT optimal assignment	Sample solution vectors			
			W_{hp}	W_{lp}	W_{shift}	$S_{\text{Threshold}}$
Mission1	143.5	(w1,t2), (w1,t1), (w2,t4), (w2,t5), (w3,t6), (w3,t7)	(0.9, 0, 0, 0.1, 0)	(0.5, 0.2, 0.2, 0.1, 0)	[0, 0.4]	[0,1]
Mission2	203.1	(w2,t5), (w1,t1), (w3, t7)	(0.9, 0.01, 0.01, 0.07, 0.01)	(0.7, 0.05, 0.14, 0.08, 0.03)	[0,1]	[0,1]
Mission3	146.2	(w1,t4), (w1,t9), (w1,t1), (w2,t2), (w2,t6), (w3,t3), (w3,t4), (w4,t8)	(0.9, 0.01, 0.01, 0.07, 0.01)	(0.7, 0.03, 0.05, 0.10, 0.12)	[0.15, 0.7]	[0,1]
Mission4	185.85	(w1,t3), (w1,t1), (w3,t4), (w3,t3), (w4,t7), (w4,t4)	(0.9, 0.01, 0.01, 0.07, 0.01)	(0.7, 0.03, 0.05, 0.10, 0.12)	[0, 0.7]	[0,1]
Mission5	94.85	(w1,t7), (w1,t6), (w1,t3), (w2,t1), (w3,t4), (w3,t3), (w3,t2), (w4,t5)	(0.9, 0.1,0,0,0)	(0.5, 0.2, 0.1, 0.1, 0.1)	[0.25, 0.5]	[0,1]
Mission6	38.4	(w1,t9), (w1,t8), (w2,t6), (w2,t4), (w3,t2), (w3,t1), (w3,t5), (w4,t7), (w4,t8), (w5,t3), (w5,t5), (w6,t1)	(0.9, 0, 0, 0.1, 0)	(0.5, 0.2, 0.2, 0.1, 0)	[0, 0.4]	[0,1]
Mission7	103.85	(w1,t8), (w1,t7), (w1,t1), (w2,t4), (w2,t6), (w2,t2), (w3,t3), (w3,t1), (w4,t5), (w4,t9), (w4,t6)	(0.9,0,0, 0.1,0)	(0.7, 0.1, 0.1, 0.1,0)	0.05	[0,1]
Mission8	102.2	(w1,t5), (w1,t2), (w2,t6), (w2,t3), (w3,t4), (w3,t7)	(0.9,0,0,0, 0.1)	(0.5, 0.3,0, 0.1, 0.1)	[0, 0.5]	[0,1]
Mission9	79.5	(w1,t2), (w1,t5), (w1,t1), (w2,t7), (w2,t1), (w2,t2), (w3,t4), (w3,t3), (w3,t6)	(0.9, 0.01, 0.01, 0.07, 0.01)	(0.7, 0.05, 0.01, 0.12, 0.12)	0	[0,1]

Closest Point of Approach (CPA) WTA Doctrine: Assuming the target flight path is predictable; the range at the closest point of approach (CPA) to each engagement system can be estimated. The target might be assigned to the system against which it will pass the closest. If the assignment is early enough, the minimum CPA criterion allows the maximum number of shots to be taken.

Maximum Number of Shots WTA Doctrine: The maximum number of shots assignment is based on the system that can achieve the most intercepts against the target within its maximum effective range.

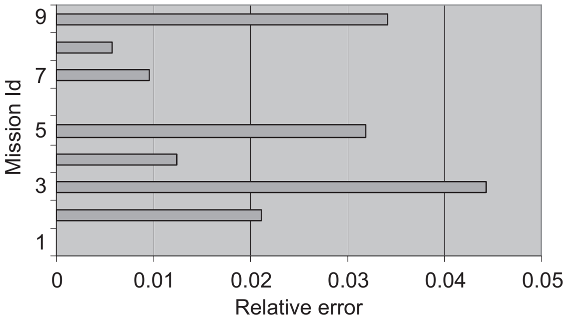


Fig. 8 Relative error for each AD mission.

Least-Engaged WTA Doctrine: If an individual weapon system is free, it can be paired with the threat that has been fired upon the least. Of course, that threat must be *engagable* by the free system. The least engaged assignment can minimize the potential for saturation because it automatically allocates systems and shots as equally as possible over the threat ensemble. This is particularly true when the individual weapon systems are equipped with only a single guidance channel.

The comparison results of the expected damage among the proposed doctrine and the above three doctrines are shown in Table 10 and Fig. 9, indicating superior results and better assignment achieved by the proposed doctrine.

The CPA and Max Shots doctrines have similar performances. The least engaged doctrine generally has bad performance except in layered defense type missions (missions 6, 7, 8, and 9). In such missions, the least engaged

Table 10 Comparison table among the proposed doctrine and CPA, MaxShots, and least engaged doctrines showing the expected damage of each mission

Mission:	Mission1	Mission2	Mission3	Mission4	Mission5	Mission6	Mission7	Mission8	Mission9
CPA	143.5*	203.1*	260.755	234.5	133.105	137.44	234.915	230†	157.85
MaxShots	143.5*	203.1*	260.755	231.35	167.16	165.82**	251.22†	216.386	162.11†
Least engaged	154†	240†	284.91**	300†	185.85**	128.455	205.35	213.5	146.85
Proposed algorithm	143.5*	226.35	194.85*	199.5*	129.85*	38.4*	114.35*	108.5*	117*

Best result for each mission is indicated by*; worst result is indicated by†

Fig. 9 Comparison chart of different WTA doctrines with the proposed doctrine through the nine AD missions in Appendix D.

doctrine achieves better assignment than the CPA and Max Shots doctrines. The proposed algorithm dramatically enhances the results as shown in Fig. 9 by providing better assignments.

V. Conclusion

In this paper a novel goal-based doctrine is proposed for the WTA problem. Load balance policy is a proposed for preventing weapon system saturation. Weapon-threat QoS factors are suggested and enumerated by a proposed threat prioritization rule and WT pair merits rules. The merit weight representing importance differs from one WT pair to another owing to differences in threat priority and the percentage of free engagement slots (which affects load balance policy activation). The neural network approach is investigated for determining optimal values for merit weights equations parameters. Such an approach fails to find a solution, revealing that the cost function is a nondifferentiable step function with respect to W_{hp} or W_{lp}, causing the failure of the gradient descent optimization method employed in the neural network learning algorithm. The sequential search algorithm is found to be more suitable in searching for the optimal configuration for merit weight equations parameters as the cost function, being a step function, most likely has its global minimum as an interval instead of single point. After finding the optimum merit equation parameters using the sequential search, the proposed doctrine is applied to nine AD missions, along with other goal-based WTA doctrines. The relative errors between the resulting damage for each mission at the global minimum weight parameters and each mission optimal damage value are calculated, showing that near optimal assignment is achieved in different air defense mission types by the proposed doctrine. The proposed algorithm has the best results when compared with the CPA, Maximum Shots, and least engaged goal-based WTA assignment doctrines.

Appendices

Appendix A: Air Defense Mission Modeling And Analysis Methodology

Mission analysis is a many-on-many proposition. However, its basic analytical building block is a method of handling one-on-one encounters. The method should remain tractable when applied to more complex situations. The methodology used has been cited in [22]. Its description is briefly given below.

Figure A1 shows some key parameters of a single-system encounter against a threat. Five range quantities are shown and are defined as follows:

R_{DS} = detection range of the search sensor.
R_{DT} = detection range of the precision track sensor.
$R_{W\,\mathrm{max}}$ = maximum effective weapon range.
$R_{W\,\mathrm{min}}$ = minimum weapon range.
R_{CPA} = range at CPA.

The search sensor will normally detect the target first in t_{detect} time interval, followed by designation to the tracking sensor. In a perfect case, a precision track will be established in time to allow the first-round intercept to occur at the maximum effective range of the weapon. This is not always the case. In Fig. A1, R_{DT} is drawn less than $R_{W\,\mathrm{max}}$ to show that the tracker range can be less than the maximum effective range of the weapon. We assume the weapon cannot be fired until a precision track has been established.

Figure A1 shows three intercepts. A sample probability of kill (PK) is indicated for the first two. The third intercept falls outside the maximum effective range of the weapon, resulting in a $\mathrm{PK} = 0$. In addition, the target range exceeds the precision-tracker range before the third intercept. Thus, for an intercept to have a nonzero kill probability, it must occur within:

- The maximum effective range of the weapon.
- The maximum-precision tracking range.

After threat detection, a firm track must be established. The firm track pertains to a nonprecision track, performed by the search sensor in this case. It will take t_{NPFTrack} time on average to do it. After the firm track has been established a weapon is assigned to the threat. Assignment results in the transmission of designation data to the precision tracking element. After a firm track is established by the precision tracker, assuming it will take t_{PFTrack} time on average, the weapon is aimed and the first launch starts.

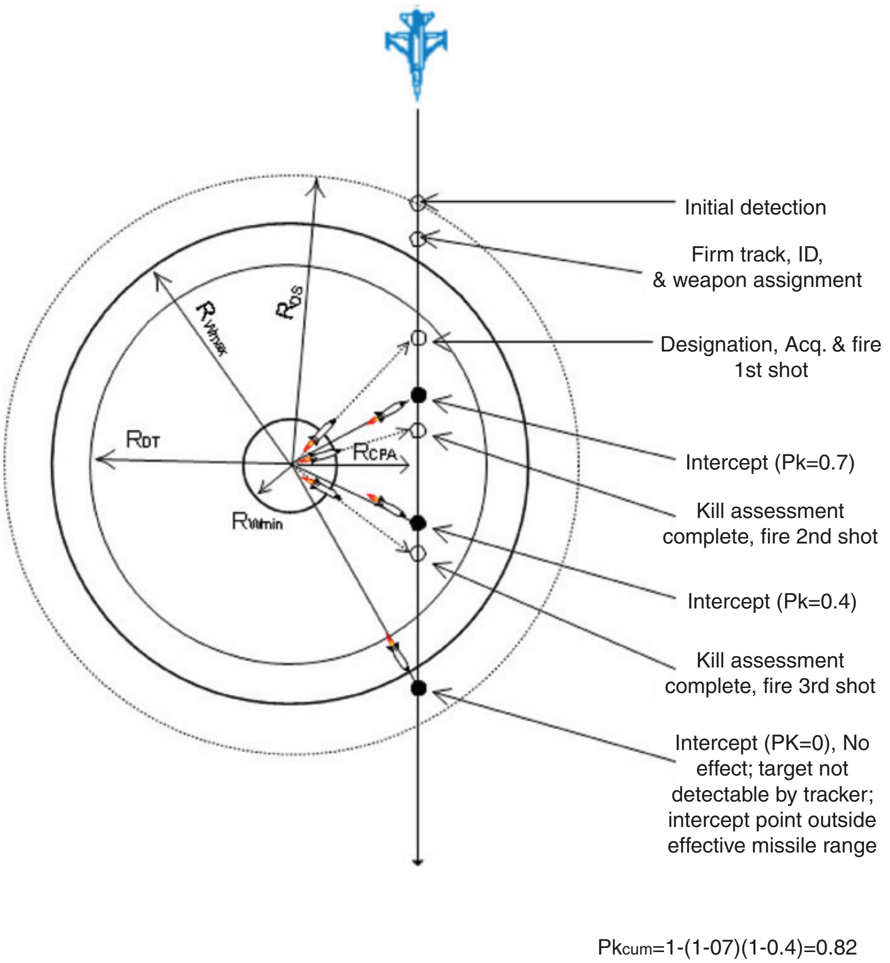


Fig. A1 A reproduction of an example one-on-one engagement figure given in [22].

Interception Time Calculation

Mission models are generally based on major simplifications when they account for the dynamics of weapon flight [22]. The missile is assumed to travel in a straight line, corresponding to a perfect collision course. The distance traveled can be represented as a linear function of time. Figure A2 shows the geometry of the interception time calculation problem.

L: distance between threat position and closest point of approach at launch time.

V_t : Threat speed.

V_m : Missile speed.

R_m : Distance traveled by Missile in time t.

R_t : Distance between the weapon system and the Threat.

Note that $R_m = R_t$ at interception time.

$$R_t^2 = CPA^2 + (L - V_t t)^2$$

$$R_m = V_m t$$

$$R_m = R_t$$

$$V_m t = \sqrt{CPA^2 + L^2 - 2V_t t + V_t^2 t^2}$$

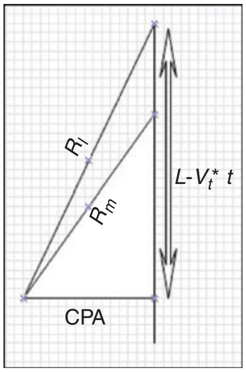


Fig. A2 Interception time calculation geometry.

$$V_m^2 t^2 = CPA^2 + L^2 - 2V_t t + V_t^2 t^2$$

$$(V_m^2 - V_t^2) t_{\text{Intercept}}^2 + 2V_t t_{\text{Intercept}} - CPA^2 - L^2 = 0$$

$$t_{\text{Intercept}} = \frac{-2V_t + \sqrt{4V_t^2 + 4(V_m^2 - V_t^2)(CPA^2 - L^2)}}{2(V_m^2 - V_t^2)} \tag{A1}$$

Note that $t_{\text{Intercept}}$ is real positive only when $V_m > V_t$.

Kill Assessment

It takes time to determine if the target is destroyed, and to determine if another shot will achieve an intercept within the effective range of the weapon. This time is very short if the target can be seen; it can be several seconds if post-intercept tracking data must be interpreted.

Appendix B: Merits Calculation

Five merits are calculated for each WT pair after prioritization of AD mission threats.

Merit1: Cumulative Kill Probability

Merit1 represents the cumulative kill probability given by

$$\text{Merit}_1 = P_{k_{\text{cummulative}}} = 1 - \prod_{k=1}^{S_{ij}} (1 - \text{SSPK}_k) \tag{B1}$$

where $\prod_{k=1}^{S_{ij}} (1 - \text{SSPK}_k)$ represents the probability that all intercepts will miss and the threat will survive. SSPK_k is the single shot kill probability for missile k, and S_{ij} is the number of missiles launched on threat j by weapon i. This merit captures information about the weapon's ability to destroy the threat.

Merit2: Free Engagement Slots Ratio During Pair Engagement

Merit2 represents the ratio between the number of free engagement slots and the number of *engagable* threats during this pair engagement.

$$\text{Merit}_2 = \frac{N_{\text{FreeSlots}}}{N_{\text{EngagableThreats}_{\text{During}}}} \tag{B2}$$

where $N_{\text{FreeSlots}}$ represents number of free engagement slots after allocating this WT pair, and $N_{\text{EngagableThreats}_{\text{During}}}$ represents number of *engagable* threats during this engagement. Note that if any engagement intersects in time with this pair engagement then its threat is considered one of the *engagable* threats during this engagement.

This merit captures information about the weapon's ability to engage other threats concurrently with this engagement, if the threat of this engagement is assigned to it. If this merit's value is less than 1, this means that this weapon system may be saturated.

Merit3: Free Engagement Slots Ratio after Pair Engagement

Merit3 represents the ratio between the free engagement slots if this pair is allocated and the number of *engagable* threats that have their engagement time interval not intersected with the current pair engagement interval. Such merit is calculated by

$$\mathrm{Merit}_3 = \frac{N_{\mathrm{FreeSlots}}}{N_{\mathrm{EngagableThreats}_{\mathrm{Before}}} + N_{\mathrm{EngagableThreats}_{\mathrm{After}}}} \tag{B3}$$

where $N_{\mathrm{FreeSlots}}$ represents number of free engagement slots after allocating this WT pair, $N_{\mathrm{EngagableThreats}_{\mathrm{Before}}}$ represents number of *engagable* threats that have their engagement interval before this engagement, and $N_{\mathrm{EngagableThreats}_{\mathrm{After}}}$ represents number of *engagable* threats that have their engagement interval after this engagement.

This merit captures information regarding the weapon's ability to engage other threats before or after this engagement, if the threat of this pair is assigned to it. If this merit value is less than 1, then this weapon system may be saturated if this threat is allocated to it.

Merit4: Average Missile Count Available for Each Engagable Threat

Merit4 represents the ratio between the remaining number of shots in the weapon system inventory, after allocating this weapon threat pair, and the number of remaining threats *engagable* by this weapon. Such ratio is calculated by

$$\mathrm{Merit}_4 = \frac{N_{\mathrm{Missiles}}}{N_{\mathrm{EngagableThreats}}} \tag{B4}$$

where N_{Missiles} is the remaining number of shots in the weapon system inventory if this pair is allocated, and $N_{\mathrm{EngagableThreats}}$ is the number of other threats *engagable* by this system. This merit captures the effect of assigning the threat in this pair to the weapon to the fire power (missiles) in other possible engagements that can be assigned to this weapon.

Merit5: Time Ratio Complement Consumed by Pair Engagement

Merit5 represents 1 minus the ratio between this pair engagement time and the overall AD mission time.

$$\mathrm{Merit}_5 = 1 - \frac{T_{\mathrm{Engagement}}}{T_{\mathrm{Mission}}} \tag{B5}$$

where $T_{\mathrm{Engagement}}$ is the amount of time the engagement will take. $T_{\mathrm{Engagement}}$ represents the duration in which a guidance channel is busy for this engagement. The engagement starts when the precision tracker starts tracking the threat and ends when the threat reaches the ordnance delivery point, or leaves weapon maximum range. T_{Mission} is the total duration of the mission.

This merit captures the effect of allocating this pair on the number of free guidance channels available. The greater the value of this merit the less time taken by this engagement and the sooner the occupied guidance channels will be free.

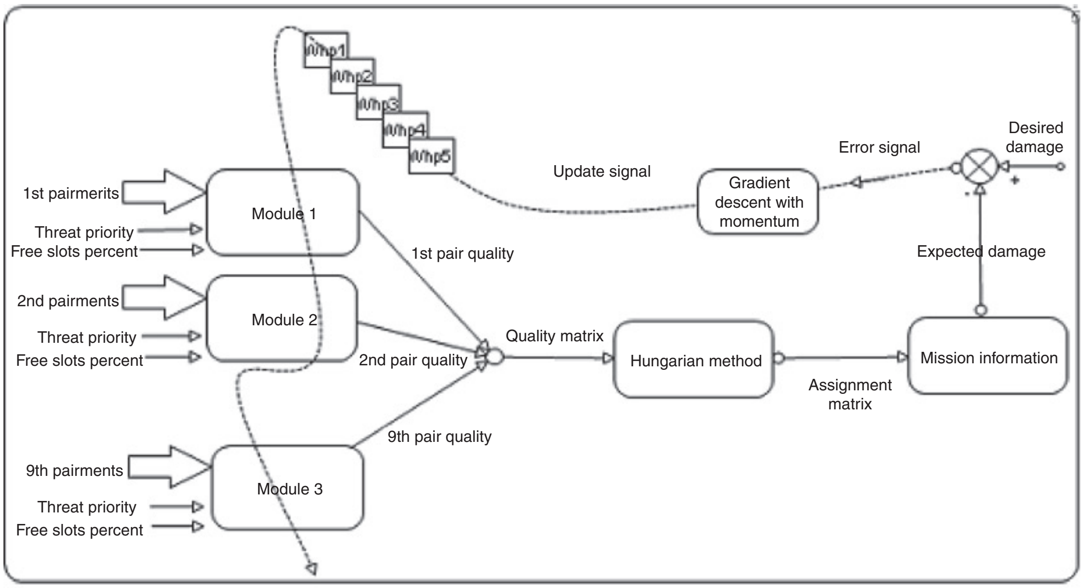
All merits equations have been tailored in such a way that the greater their values, the better and more suitable the WT pair they represent.

Appendix C: Failure of MLP type ANN in Estimating Merits Weights Equation Parameters

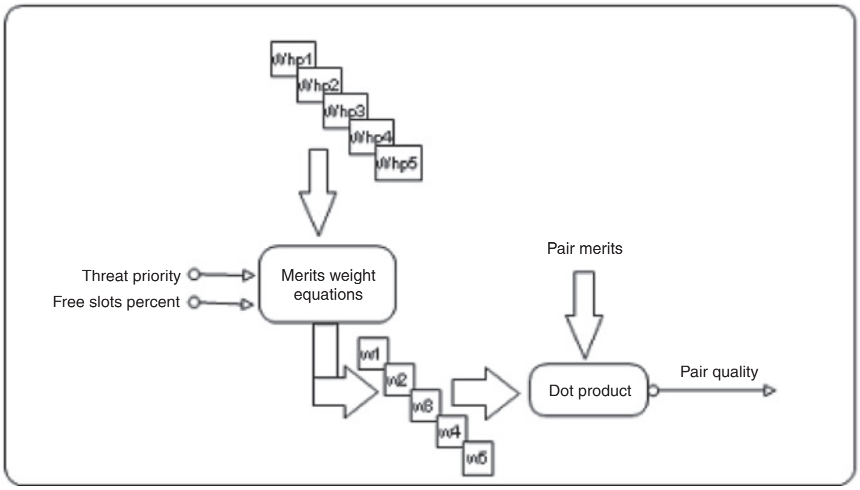
Investigation of ANN as a possible tool for specifying merits importance, and the amount of importance shifting done by the load balance policy, has been conducted. A modular multi-layer-perceptron (MLP) type ANN is designed.

Design

Figure C1 shows ANN design for a simple 3 × 3 WTA problem. Each module represents a WT pair. The input layer of each module consists of seven neurons representing the five merits value, threat priority, and free engagement



a)



b)

Fig. C1 ANN design for a simple 3 × 3 WTA problem is shown in (a). Design of single module from the nine modules of the example ANN is shown in (b).

slots percentage. No hidden layer is used. The output layer consists of a single neuron representing the overall quality value for this weapon threat pair. In each module no activation function is employed in the output neuron and instead of an adder for summing the weighted input signal, a custom processing unit is implemented to encapsulate the merits weight equations and perform the following tasks:

1. Use the synaptic weights after normalization as W_{hp}(or W_{lp}, based on the training phase) and substitute in Eqs. (19) to (22) to calculate merits weights.
2. Multiply the input signal by the resulted merits weights and summing the resulted weighted input to get the output signal which represents the WT pair QoS value.

The output of each module represents an element in the quality matrix **Q**. The quality matrix is then used by the Hungarian method to calculate the optimal assignment for such quality matrix. After that, the overall expected damage inflected on the assets is calculated based on mission data and the assignment generated by the Hungarian method. The cost is calculated based on the difference between the calculated expected damage and the desired expected damage that would result if the optimal assignment for this mission is chosen. The ANN uses the calculated cost to update the weights W_{hp} (or W_{lp}) in a way that reduces the cost. The algorithm used to update the weights is based on the gradient descent optimization method with momentum to reduce the possibility of being trapped in a local minimum instead of the global minimum. Note that all modules in the network share the same weight vector. The training should be done in two phases and for different $S_{\mathrm{Threshold}}$ and W_{shift} values.

In the first phase, we fixed the W_{lp} vector and tried to find W_{hp} using ANN. In the second phase, W_{hp} is fixed at the values we got from the first phase, and we tried to obtain W_{lp} . These two phases are repeated for different $S_{\mathrm{Threshold}}$ and W_{shift} values.

Steps of training the ANN is summarized as the following:

1. prioritize threats.
2. For each training phase:
    a. compute the merits values and the percentage of free engagement slots for each WT pair.
    b. use the neural network to acquire the overall quality value for each pairs by performing weighted sum on the merits.
    c. the Hungarian method is used to obtain the optimal assignment based on the resulted quality matrix.
    d. calculate the expected damage inflicted on the assets based on the assignment and mission data.
    e. calculate the error and update the weights using gradient descent to minimize that error.

Table C1 Sample input for the example ANN shown in Fig. C1

	Input							Desired Signal
WT pair	Priority	freeSlots Percent	Pkcumm	AvgFreeSlots RatioDuring	AvgFreeSlots RatioAfter	AvgMissile CountRatioAfter	1-Engagement TimeRatio	Optimal assignment damage
(w1,t1)	0.2315	1	0.91	0	1	6	0.677066	24
…	…	…	…	…	…	…	…	…
(w1,t2)	0.1878	1	0.91	0	1	6	0.689689	24
…	…	…	…	…	…	…	…	…
(w1,t3)	0.1662	0	0	0	0	0	0	15
…	…	…	…	…	…	…	…	…
(w1,t4)	0.2102	0	0	0	0	0	0	59

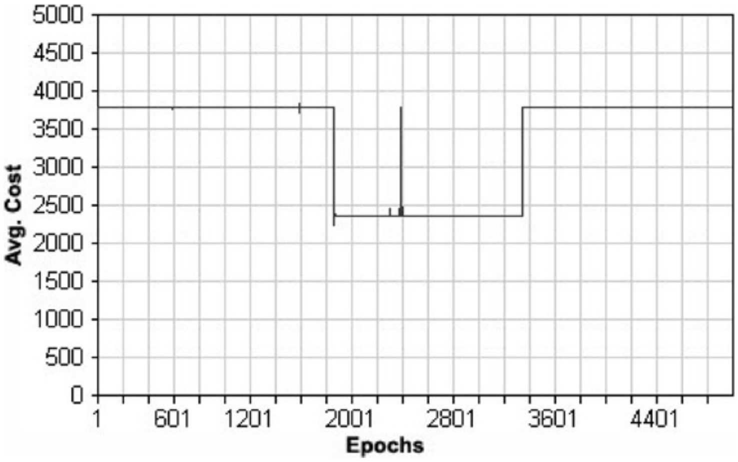
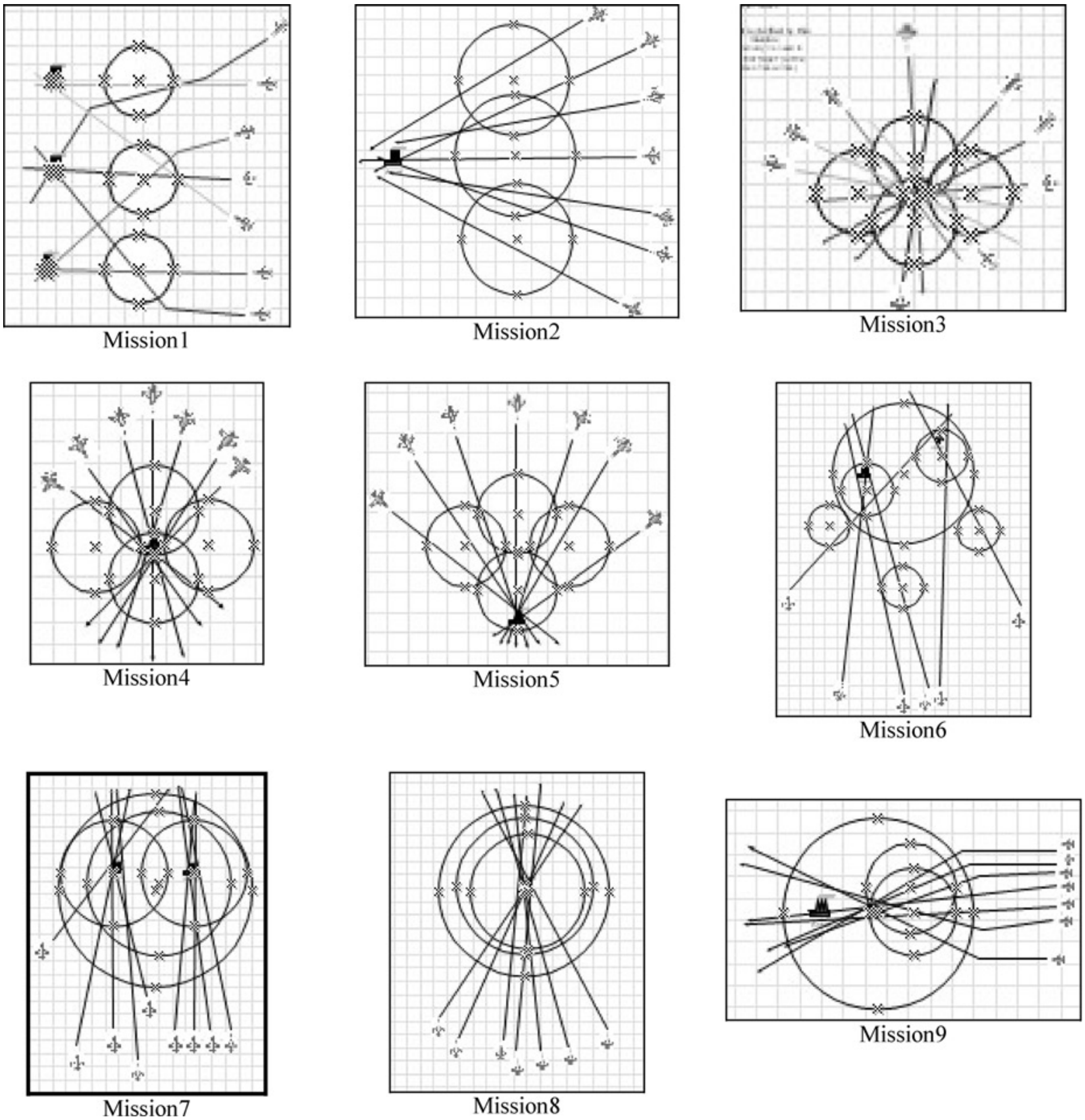


Fig. C2 Training results for the ANN represented by average cost vs epochs, showing that the average cost function is found to be nondifferentiable step function.

Table C2 Artificial neural network training configuration parameters

η	μ	Epochs
0.1	0.7	5000



Mission1

Mission2

Mission3

Mission4

Mission5

Mission6

Mission7

Mission8

Mission9

Fig. D1 Thumbnail view for the nine AD missions used in the simulation and results of the proposed WTA doctrine.

Steps for solving the WTA problem using the trained ANN network:

1) Prioritize threats.
2) Compute the merits values and the percentage of free engagement slots for each WT pair.
3) Use the neural network to get the overall quality value for each pairs by performing weighted sum on the merits.
4) The Hungarian method is used to get the optimal assignment based on the resulted quality matrix.

Sample Input

Each input to the whole network represents an assignment problem that consists of three weapons and three threats (see the example ANN design shown in Fig. C1), resulting in the nine pairs represented by the three modules of the ANN.

One mission is used in this prototype example ANN. Such a mission provides 42 different 3×3 assignment problems. A sample from the input data for the 1st module along with the desired value of such input is shown in Table C1. The desired signal is the expected damage of the optimal assignment.

Training Results

Training is conducted for 5000 iterations (epochs). Figure C2 shows the average cost of the 42 assignments versus training epochs in the first training phase to get W_{hp}while fixing W_{lp} at (0.5,0.15,0.15,0.05,0.15), weight shift (W_{shift}) $= 0.05$, and threshold ($S_{\mathrm{Threshold}}$) $= 0.1$. The average cost function is found to be a nondifferentiable function. No tendency to convergence was found during the training (even for different values of W_{lp}, W_{shift}, and $S_{\mathrm{Threshold}}$), as the gradient descent optimization method require the cost to be differentiable with respect to weights. Gradient descent method with momentum employs the following error correction rule:

$$\Delta W(n+1) = -\eta \nabla \xi(W) - \mu \Delta W(n) \tag{C1}$$

where η is gradient descent learning rate (step size), and μ is the momentum rate and $\nabla \xi(W)$ is the gradient vector of the cost function. Configuration parameters of the ANN are given in Table C2.

After investigating the results, it is found that the cost as a function in the weights E(W) is not differentiable in this problem. As a result the gradient descent method used by the neural network will fail to approach the minimum as it takes steps proportional to the *negative* value of the gradient, and in this case the gradient is zero. As shown in Fig.C2, the cost function is a step function, which seems logical as the changes in weights causes a change in the assignment. The expected damage, resulting from the new assignment, presents a sudden change from the previous expected damage because there is no relation between them as they resulted from different assignments.

Appendix D: Air Defense Missions

Nine missions were designed and used in the results and simulation. A thumbnail view for the topology of these missions is given in Fig. D1. Different patterns of attacks and different defensive topologies have been taken into account in the design of these missions.

References

[1] Manne, A. S., “A Target-Assignment Problem,” *Operations Research*, Vol. 6, No. 1, 1958, pp. 346–351.

[2] Braford, J. C., “Determination of Optimal Assignment of a Weapon System to Several Targets,” Technical Report AER-EITM-9, Vought Aeronautics, Dallas, TX, 1961.

[3] Day, R. H., “Allocating Weapons to Target Complexes by Means of Nonlinear Programming,” *Operations Research*, Vol. 14, No. 6, 1966, pp. 992–1013.

[4] Maltin, S. M., “A Review of the Literature on the Missile-Allocation Problem,” *Operations Research*, Vol. 18, No. 2, 1970, pp. 334–373.

[5] Eckler, A. R., and Burr, S. A., “*Mathematical Models of Target Coverage and Missile Allocation*,” Military Operations Research Society, Alexandria, VA, 1972.

[6] Murphey, R. A., “Target-Based Weapon Target Assignment Problems,” *Nonlinear Assignment Problems: Algorithms and Applications*, edited by P. M. Pardalos and L. S. Pitsoulis, Kluwer Academic, The Netherlands, 1999, pp. 39–53.

[7] Lloyd, S. P., and Witsenhausen, H. S., “Weapons Allocation is NP-complete,” *Proceedings of the 1986 Summer Conference on Simulation*, The Society for Modeling & Simulation International (SCS), Reno, NV, 1986, pp. 1054–1058.

[8] DenBroader G. G., Ellison, Jr. R. E., and Emerling, L., "On Optimum Target Assignments," *Operations Research*, Vol. 7, 1958, pp. 322–326.

[9] Katter, J. D., "A Solution of the Multi-Weapon, Multi-Target Assignment Problem," Working paper 26957, MITRE, 1986.

[10] Chang S. C., James, R. M., and Shaw, J. J., "Assignment Algorithm for Kinetic Energy Weapons in Boost Defense," *Proceedings of the IEEE 26th Conference on Decision and Control*, Inst. of Electrical and Electronics Engineers, New York, NY, 1987, pp. 1678–1683.

[11] Orlin, D., "Optimal Weapons Allocation Against Layered Defenses," *Naval Research Logistics,* Vol. 34, No. 5, 1987, pp. 605–616.
doi: 10.1002/1520-6750(198710)34:5<605::AID-NAV3220340502>3.0.CO;2-L

[12] Castanon, D. A., "Advanced Weapon-Target Assignment Algorithm Quarterly Report," TR-337, ALPHA TECH Inc., Burlington, MA, 1987.

[13] Wacholder, E., "A Neural Network-Based Optimization Algorithm for the Static Weapon-Target Assignment Problem," *ORSA Journal on Computing,* Vol. 1, No. 4, 1989, pp. 232–246.

[14] Metler, W. A., and Preston, F. L., "A Suite of Weapon Assignment Algorithms for a SDI Mid-couRse Battle Manager," NRL Memorandum Report 671, Naval Research Laboratory, Washington, DC, 1990.

[15] Grant, K. E., "Optimal Resource Allocation Using Genetic Algorithms," *Naval Review*, Naval Research Laboratory, Washington, DC, 1993, pp. 174–175.

[16] Zne-Jung Lee, "Efficiently Solving General Weapon-Target Assignment Problem by Genetic Algorithms with Greedy Eugenics," *IEEE Transactions on Systems, Man, and Cybernetics—Part B: Cybernetics*, Vol. 33, No. 1, 2003, pp. 113–121.

[17] Green, D. J., "An Integer Solution Heuristic for the Arsenal Exchange Model (AEM)," *Military Operations Research Society*, Vol. 3, No. 2, 1997, pp. 5–16.

[18] Ahuja, R. K., Kumar, A., Krishna, J., and Orlin, J. B., "Exact and Heuristic Methods for the Weapon Target Assignment Problem," MIT Sloan School of Management, Working Paper No. 4464-03, 2004.

[19] Frank, A., and Magyar, A., "The Hungarian Method and its Extensions" (in Hungarian), Szigma, Vol. 33, 1-2, 2002, pp. 13–44.

[20] Kuhn, H. W., "The Hungarian Method for the Assignment Problem," *Naval Research Logistic Quarterly*, Vol. 2, No. 1–2, 1955, pp. 83–97.
doi: 10.1002/nav.3800020109

[21] Kuhn, H. W., "On the Origin of the Hungarian Method," *History of Mathematical Programming: A Collection of Personal Reminescences*, edited by J. K. Lenstra, A. H. G. Rinnooy Kan, and A. Schrijver, CWI Amsterdam and North-Holland, Amsterdam, 1991.

[22] Macfadzean, R. H. M., "*Surface-Based Air Defense System Analysis*," Artech House, 2000, pp. 336–339.

Christopher Rouff
Associate Editor